Dimensionality Reduction for Dynamical Systems with Parameters

Welshman, Christopher

2014

MIMS EPrint: **2014.22**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

DIMENSIONALITY REDUCTION FOR DYNAMICAL SYSTEMS WITH PARAMETERS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2014

Christopher Welshman
School of Mathematics

Contents

Word count 33793

List of Tables

List of Figures

The University of Manchester

Christopher Welshman
Doctor of Philosophy
Dimensionality Reduction for Dynamical Systems with Parameters
April 30, 2014

Dimensionality reduction methods allow for the study of high-dimensional systems by producing low-dimensional descriptions that preserve the relevant structure and features of interest. For dynamical systems, attractors are particularly important examples of such features, as they govern the long-term dynamics of the system, and are typically low-dimensional even if the state space is high- or infinite-dimensional. Methods for reduction need to be able to determine a suitable reduced state space in which to describe the attractor, and to produce a reduced description of the corresponding dynamics. In the presence of a parameter space, a system can possess a family of attractors. Parameters are important quantities that represent aspects of the physical system not directly modelled in the dynamics, and may take different values in different instances of the system. Therefore, including the parameter dependence in the reduced system is desirable, in order to capture the model's full range of behaviour.

Existing methods typically involve algebraically manipulating the original differential equation, either by applying a projection, or by making local approximations around a fixed-point. In this work, we take more of a geometric approach, both for the reduction process and for determining the dynamics in the reduced space. For the reduction, we make use of an existing secant-based projection method, which has properties that make it well-suited to the reduction of attractors. We also regard the system to be a manifold and vector field, consider the attractor's normal and tangent spaces, and the derivatives of the vector field, in order to determine the desired properties of the reduced system.

We introduce a secant culling procedure that allows for the number of secants to be greatly reduced in the case that the generating set explores a low-dimensional space. This reduces the computational cost of the secant-based method without sacrificing the detail captured in the data set. This makes it feasible to use secant-based methods with larger examples.

We investigate a geometric formulation of the problem of dimensionality reduction of attractors, and identify and resolve the complications that arise. The benefit of this approach is that it is compatible with a wider range of examples than conventional approaches, particularly those with angular state variables. In turn this allows for application to non-autonomous systems with periodic time-dependence. We also adapt secant-based projection for use in this more general setting, which provides a concrete method of reduction.

We then extend the geometric approach to include a parameter space, resulting in a family of vector fields and a corresponding family of attractors. Both the secant-based projection and the reproduction of dynamics are extended to produce a reduced model that correctly responds to the parameter dependence. The method is compatible with multiple parameters within a given region of parameter space. This is illustrated by a variety of examples.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to thank John Brooke for his guidance and support, David Broomhead for his infectious enthusiasm, and Paul Glendinning for his encouragement. I would also like to thank all those involved in the CICADA project. I have benefited from the MAGIC programme of graduate mathematics lectures, which have had an impact on this work and beyond. This work was supported by EPSRC grant EP/C536452/1.

Chapter 1

Introduction

One of the lessons learned from the study of dynamical systems is that large, complex systems can exhibit simple behaviour, while small, simple systems can exhibit complex behaviour. The canonical example of the latter is the Lorenz system, in which a simple 3-dimensional system gives rise to a chaotic attractor. On the other hand, large, dissipative systems often possess simple long-term behaviour, such as fixed points and periodic orbits. These observations suggest that it may be possible to describe some large systems, at least partially, using small systems.

By 'large' and 'small', we are referring to the dimension of the state space of the system, which reflects the number of variables needed to quantify the system's configuration at any moment in time. In this work we investigate methods of obtaining low-dimensional models that describe the long-term dynamics of given high-dimensional systems. This is important as high-dimensional systems have a large computational cost associated with their simulation and analysis. We are therefore interested in *dimensionality reduction* methods for dynamical systems, and so we begin by briefly discussing the main types of dimensionality reduction methods and their uses.

1.1 Dimensionality Reduction

There are a number of situations in modern science in which one encounters high-dimensional spaces. These are particularly common when dealing with physical systems whose configuration is described by a large number of variables. The study of

such systems is made complicated by its dimension; for example, the observation, identification and classification of structure is more challenging due to the greater degrees of freedom and difficulties in visualizing the system. A further consequence of high-dimensionality is the large computational cost associated with a numerical treatment of the system, such as processing data or simulating dynamics. Thanks to modern technology we also often have an abundance of data, particularly from experimental measurements, and in some cases multiple mathematical models of a system, of varying shapes and sizes.

Despite their high-dimensionality, in many cases one finds that the structure of interest is of intrinsically lower dimension. Dimensionality reduction refers to mathematical methods that are designed to produce a low-dimensional description of a particular structure in a given high-dimensional system. We can classify existing methods into three classes: *data reduction*, *invariant manifolds* and *model reduction*. Each of these classes deals with rather different contexts and attempts to preserve structures relevant to their particular context. In this section we describe the main manifestations of dimensionality reduction and discuss their similarities and differences in both their objective and approach.

1.1.1 Data Reduction

In *data reduction*, one deals with a set of data points in a high-dimensional space, typically $\mathbb{R}^n$. Such data may be the result of experimental measurements, for example, a set of images of brain activity of a patient under varying stimuli. When analysing such data sets, one wants to discover particular structures in the data in order to identify corresponding relationships in the physical system. Simple examples of structure include statistical properties, such as correlation and clustering, while more advanced examples may involve reverse engineering the topological structure of the data, by regarding the data as living on a manifold embedded into the ambient space in a complicated way. Data reduction techniques involve obtaining a description of the data set in an ambient space of lower dimension, while preserving the structure of interest.

Principal Component Analysis

A simple method of reducing a high-dimensional data set is via Principal Component Analysis (PCA) [31, 24]. PCA produces an orthogonal basis for the space, with a corresponding set of positive weights that quantify the variance in each direction, called the principal components. The components are produced in order of decreasing variance. This allows one to keep the most significant principal components and discard the least significant components. By describing the data set using only the most significant components, one has obtained a description of the data of lower dimension that retains most of the statistical variance, i.e. the data remains 'spread out' in the subspace.

The PCA can be performed by use of a singular value decomposition (§2.1.18) of the data matrix (where each column describes a data point). The singular values are related to the principal components, and the left-singular vectors provide the corresponding orthonormal basis.

Multidimensional Scaling

Another classic method is Multidimensional Scaling (MDS) [9]. This method attempts to preserve the pairwise distances between data points in the low-dimensional space. The problem is phrased as an optimization problem for the low-dimensional data set, with a cost function given in terms of the difference in pairwise distances between the original and reduced data sets.

Nonlinear Methods

Nonlinear methods of data reduction regard the data as living on a low-dimensional manifold embedded into the ambient space, $\mathbb{R}^n$. Rather than seeking an explicit nonlinear map from the high-dimensional space to a low-dimensional space, most nonlinear methods attempt to produce a corresponding data set in a low-dimensional ambient space that is regarded as a coordinate system parameterizing the manifold. The low-dimensional parameterization is determined so that it preserves some of the properties of the original data, which are chosen to reflect the underlying manifold structure.

This often requires first determining the topological structure of the data set by constructing a neighbourhood graph, which quantifies which data points are 'near' to each other. This is then used to compute quantities to be preserved in the reduced data set.

Neighbourhood Graph

A neighbourhood graph is a graph with a vertex for each data point and edges defining the local neighbourhood of each point. There are two common ways to connect the vertices with edges in terms of a metric on the ambient space. The k-nearest neighbour approach creates edges to the k closest vertices to each point. This means a graph containing N vertices has at least $Nk/2$ edges in total. However, some inappropriately long edges may be generated by this method in regions of low density.

The fixed-radius neighbourhood approach creates edges for all pairs of points that are closer than a given distance, $\varepsilon > 0$. One must be careful to choose a suitable ε for use with the given data set, as the number of edges generated is sensitive to the density of sampled data points relative to the fixed radius.

Isomap

The Isomap method [45] takes the approach of MDS (preserving pairwise distances), but rather than using the Euclidean distance through the ambient space, uses an approximation of the geodesic distance along the underlying manifold. The method begins by constructing a neighbourhood graph with each edge weighted by its corresponding Euclidean distance. The graph is processed to compute the shortest distance between each pair of vertices through the graph. For a sufficiently dense data set, this shortest path approximates the geodesic distance along the underlying manifold. The pairwise distances are then reproduced by a low-dimensional data set.

Local Linear Embedding

The Local Linear Embedding (LLE) [38] is based on the fact that a local piece of a manifold can be approximated linearly. For each data point, the method constructs an approximation of the position of the point as a linear combination of its k nearest neighbours, with appropriate constraints placed on the weights. As a result, these

weights are invariant under rotations, scalings and translations of the data set in the ambient space. The reduction then proceeds by by seeking a low-dimensional data set that produces the same set of weights at each point.

Hessian Local Linear Embedding

The Hessian LLE method [11] is similar to the LLE, in that it computes local linear descriptions of the data set. The central object is the quadratic form $\mathcal{H}(f) = \int_{\mathcal{M}} \|H_f(m)\|_{\mathrm{F}}^2 \, \mathrm{d}m$, defined on smooth functions $f : \mathcal{M} \to \mathbb{R}$, where $H_f(m)$ is the Hessian of f at $m \in \mathcal{M}$. The key property of this quadratic form is that it has a $(d+1)$-dimensional kernel. Excluding the zero eigenvalue that corresponds to the 1-dimensional space of constant functions, this leaves a d-dimensional space containing functions that parameterize the manifold.

The method uses a neighbourhood graph of the data set to find a d-dimensional basis at each point that approximates each tangent space. These local bases are used to construct a matrix that describes the discretized version of the quadratic form (where the manifold becomes a finite data set, the smooth function becomes a finite set of values, and $\mathcal{H}$ becomes a symmetric matrix). An orthogonal basis for its kernel is chosen in which the data points are then expressed.

Another method that considers the local geometric structure is Local Tangent Space Alignment [51]. A quantitative comparison of these nonlinear methods can be found in [47].

Secant-Based Projection

Secant-based projection [2] seeks an orthogonal projection that preserves the secants generated by the data. Motivated by the Whitney embedding theorem, this results in an embedding of the manifold into the subspace. This approach will be covered in detail in §3.

1.1.2 Invariant Manifolds

When applying the notion of dimensionality reduction to the state space of a dynamical system, one can take advantage of invariant sets, whose orbits are confined

to a subset of state space. In the linear case, one may have an invariant subspace: a linear subspace invariant under the time evolution operator. In the more general case, one may have an invariant manifold: a smooth object invariant under the flow [23]. Attractors are particularly important examples of invariant manifolds, as their attracting properties mean they dictate the long-term dynamics of the system. Attractors are also often compact and of low dimension, which motivates the search for low-dimensional descriptions of the asymptotic dynamics.

As a result of these invariant structures, the reduction of dynamical systems possesses some similarities to nonlinear data reduction. In both cases one has a low-dimensional compact manifold embedded in an ambient space that one wishes to describe in a low-dimensional ambient space, and one also typically does not have an explicit description of this object. However, for dynamical reduction one not only has to produce a description of the geometric object, but also its dynamics in the low-dimensional space. One therefore typically requires a map relating the original state space and the reduced space in order to determine the corresponding reduced dynamics. Although there are several methods of finding such a relation, projections onto linear subspaces are a common theme, with the state space taken to be a vector space with some additional structure (such as a Hilbert space). Nonlinear reductions have also been developed [27, 42].

Galerkin Projection

The Galerkin projection is a method for obtaining a low-dimensional dynamical system to describe the dynamics in a linear subspace of the state space. Let $x \in X$ be the state variable with dynamics given by $\dot{x} = f(x)$. We can describe a subspace as the image of an orthonormal map $W : \hat{X} \to X$, with $W^\dagger W = \mathrm{id}_{\hat{X}}$, which means $P = WW^\dagger$ projects orthogonally onto the subspace $W(\hat{X})$. Assuming the attractor is preserved in $\hat{X}$ under $W^\dagger$, we can write the inverse as $x = W\hat{x} + r(\hat{x})$. Applying this to the original dynamics gives the reduced dynamics $\dot{\hat{x}} = W^\dagger f(W\hat{x} + r(\hat{x}))$. The Galerkin approach is to choose a projection such that $r(\hat{x}) = 0$ on the attractor, i.e. to project onto the subspace that the orbit explores. For a general subspace, r may be non-zero and the attractor's inverse projection $\hat{x} \mapsto x$ is a nonlinear map.

One of the limitations of the Galerkin procedure is that it demands a linear inverse.

Since the quality of the reduced dynamics is determined by how well this linear inverse approximates the true inverse, the quality of the reduced dynamics is intimately related to the dimension of the subspace. As a result, improving the reduced dynamics requires adding more modes/basis vectors (and hence dimensions) to the reduced space, even though there may be no geometric reason to do so. Methods that attempt to describe the nonlinear inverse are called nonlinear Galerkin methods, or approximate inertial manifolds [32, 43].

Another limitation is that the differential equation for the low-dimensional state, $\dot{\hat{x}} = W^\dagger f(W\hat{x} + r(\hat{x}))$, in general requires the evaluation of the high-dimensional nonlinear vector field, f, even if $r(\hat{x}) = 0$. Only for certain forms of f can the RHS be used to derive a function $\hat{f}(\hat{x})$ that can be evaluated without going through the high-dimensional space. To address this, methods have been developed that construct an approximation of $\hat{f}$, without algebraically deriving it from f. This bypasses the need for restrictions on the form of f and the inverse projection. Methods of this type were first developed by Broomhead and Kirby [4], who use radial basis functions to reproduce $\hat{f}$ and its derivatives on the attractor. A more recent method is the 'discrete empirical interpolation method' [6], which is used to approximate the (nonlinear part of) $\hat{f}$, and is an example of a mixed approach between the conventional top-down algebraic manipulation of f (used for the linear part) and the bottom-up approximation of $\hat{f}$ (used for the nonlinear part).

Proper Orthogonal Decomposition

Also known as the Karhunen-Loève Decomposition, the Proper Orthogonal Decomposition (POD) [20, 30] is a method for finding the best subspace for describing a given data set. It is equivalent to the PCA. The POD gives the optimal Galerkin projection, in the sense that it minimizes the mean square length of r over the data set,

$$\sum_{x\in\mathcal{M}} \left\| r(W^\dagger x) \right\|^2 = \sum_{x\in\mathcal{M}} \left\| (\mathrm{id}_X - P)x \right\|^2 ,$$

for a given dimension of the subspace. For this reason, the use of the POD subspace with the Galerkin procedure is referred to as the POD-Galerkin method.

Centre Manifold Reduction

Another use of invariant manifolds is in the study of the dynamics around a fixed point. Centre manifold reduction [26, 5, 18] involves producing a reduced description of the dynamics around a fixed point. Using the linearization of the system at the fixed point, p, the tangent space at p can be decomposed into three subspaces: stable, unstable and centre, $\mathrm{T}_p X = E^s \oplus E^u \oplus E^c$. The stable subspace corresponds to eigenvalues with negative real part, the unstable subspace to eigenvalues with positive real part, and the centre subspace to eigenvalues with zero real part. As they are eigenspaces, these three subspaces are therefore invariant subspaces of the linearized dynamics. Each of these subspaces has a corresponding invariant manifold to which it is the tangent space at the fixed point: the stable manifold, unstable manifold, and centre manifold. Centre manifold reduction involves obtaining a local description of the dynamics on the centre manifold around the fixed point, which can be of much lower dimension than that of the full state space.

Centre manifolds are important in the study of bifurcations, since local bifurcations involve an eigenvalue crossing the imaginary axis as a parameter is varied. This means the fixed point is non-hyperbolic and therefore has a nontrivial centre subspace.

Inertial Manifolds

Inertial manifolds are invariant manifolds that have additional properties. In particular, they are finite-dimensional objects that contain the global attractor and are exponentially stable. They are used in the study of infinite-dimensional dissipative systems to reduce the system to a finite-dimensional system that provides a simplified description of the asymptotic dynamics, without having to produce an explicit description of the attractor itself. The theory of inertial manifolds takes place within the setting of functional analysis and topological vector spaces. [35, 8, 34, 14, 39]

1.1.3 Model Reduction

In control engineering one deals with control systems, which are dynamical systems under the influence of external inputs, and with outputs that allow for a limited observation of the state of the system. A controller's objective is to influence the system

to produce desirable outputs by dynamically manipulating the available inputs. The inputs represent the physical components directly under the influence of the controller (such as a thermostat), the outputs correspond to sensors (such as a thermometer), and the dynamical system is the relevant laws of physics that connect the two (such as the heat equation). In order for a controller be able to perform its function it

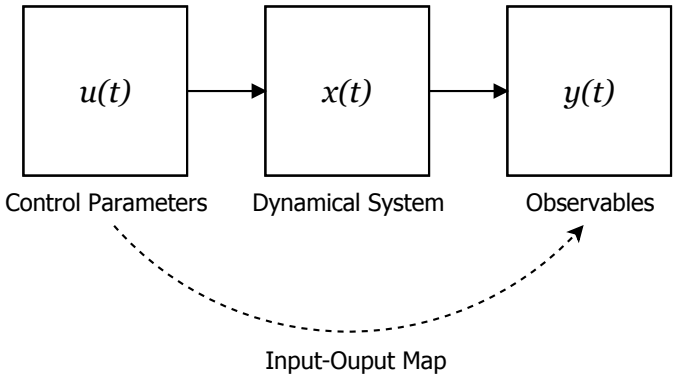


Figure 1.1: Control system.

needs to have an understanding of the relationship between inputs and outputs. Since a controller also needs to act in real-time, this model needs to be simple enough to be evaluated in a timely fashion. As the physics often produces high-dimensional state spaces, dimensionality reduction techniques are not only beneficial, but in many cases necessary, to reduce the dimension of the state space in order to produce a *reduced model* that can be evaluated with lower computational cost, while attempting to preserve the input-output behaviour found in the original model. This type of reduction is referred to as *model reduction*, or *model order reduction* (where 'order' refers to the dimension of the state space).

Linear Time-Invariant Systems

The simplest and most well-studied type of control system are linear time-invariant (LTI) systems [1], which are defined on the state space X, control space U, and observation space Y, all of which are vector spaces, with

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned} \tag{1.1}$$

where $A \in L(X)$, $B \in L(U, X)$, and $C \in L(X, Y)$ are linear maps. This gives a linear vector field under the influence of linear control, with the observation also given by a

linear map.

A reduced model is an LTI system with a lower-dimensional state space $\hat{X}$, related to the original model by the linear maps $W : \hat{X} \to X$ and $V : \hat{X} \to X$, satisfying $V^\dagger W = \mathrm{id}_{\hat{X}}$. These maps describe a projection, $WV^\dagger$, onto the subspace $W(\hat{X})$, which gives the reduced model

$$\begin{aligned} \dot{\hat{x}} &= V^\dagger AW\hat{x} + V^\dagger Bu \\ y &= CW\hat{x}, \end{aligned} \tag{1.2}$$

where we can define the reduced linear maps $\hat{A} = V^\dagger AW$, $\hat{B} = V^\dagger B$, and $\hat{C} = CW$. In the case $V = W$, the reduction is an orthogonal projection. The key part of model reduction methods therefore involves determining a suitable subspace for the reduced state space.

Proper Orthogonal Decomposition

As in §1.1.2, one can use the POD to determine a subspace. For a particular initial condition and inputs $u(t)$, one can generate a sequence of 'snapshots' from the orbit $x(t)$, and use this to compute the POD.

1.2 Parameters

Parameters are quantities that are used in the specification of a dynamical system. In models of physical systems, parameters can be used to represent aspects of the system not dynamically modelled in detail. These may be complicated microscopic processes, such as friction, whose detailed description is beyond the scope of the model, but whose influence can be included via a simple abstraction, which is configured by parameters. Such parameters are therefore of great importance to the scientists and engineers who produce the models.

The existing methods of dimensionality reduction have not explicitly addressed the problem of parameter dependence. If a dynamical system has a parameter space then, rather than a single attractor, we can have a family of attractors in the state space. Although a Galerkin procedure will retain any parameters, and this can correctly preserve some of the parameter dependence in some cases [36], the method does not

explicitly account for parameters and relies on choosing a subspace large enough to span all of the attractors simultaneously.

Note that in model order reduction, the 'parameters' are dynamically-varying control inputs, and the objective of the reduction is to preserve input-output behaviour [1], rather than particular geometric objects in the state space. In contrast to this,

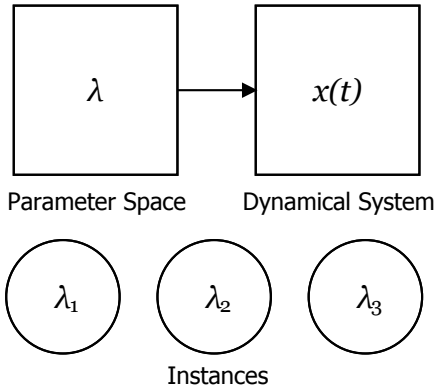


Figure 1.2: Dynamical system with static parameters.

we consider static parameters; an autonomous view is taken without any inputs or outputs. Studying the behaviour of a model by its response to variations in the values of these parameters can be done via numerical methods such as *computational steering* [7].

1.3 Sources of Relevant Examples

There are three main areas of application that are relevant to the dimensionality reduction of attractors.

1.3.1 Complex Systems

A complex system is a system comprised of many interacting components. The individual components and their interactions may be very simple, or may themselves be complex. In studying complex systems, one is particularly interested in *emergent* features, where the system as a whole exhibits behaviour or dynamically acquires properties that cannot be understood nor predicted from the study of the individual components and interactions. These kinds of system are commonly found in the biological sciences and in social science. Attractors may be regarded as one type of emergent

feature, or as an underlying mechanism that is responsible for emergent behaviour.

1.3.2 Biological Systems

As life forms need to be dynamic in order to survive and propagate, the systems that comprise them are necessarily dynamic. As a result, biology is full of rhythms, oscillations and cycles. The period, amplitude and shape of these oscillations is something that is of interest to biologists. For example, oscillations can contain fast and slow parts, such as spikes and waves, as in neurons or the heartbeat. The shape of oscillation corresponds to the geometry of the orbit in the state space.

As biological systems are also the product of evolutionary processes and in contact with an external environment, one would expect these systems to feature an element of robustness. When the external environment perturbs the system, the system needs to be able to recover from that perturbation in order to continue functioning and surviving. We would therefore expect to find stable structures in the state space of such systems, such as attractors.

1.3.3 PDEs

There are many physical systems, especially in physics and applied mathematics, whose dynamics are described by PDEs. The state of these systems is usually a scalar field or vector field over physical space, and the resulting state space is an infinite-dimensional function space. Even though the space is of infinite dimension, low-dimensional attractors can still be featured in the dynamics, particularly in systems featuring dissipation. In studying these systems, it is common to discretize the physical space in order to produce a model of finite dimension that is compatible with numerical simulation. However, the dimension of these discretizations is often still very high compared to the attractor, and reducing the number of samples in the discretization as a means to reduce the dimension further is undesirable due to the presence of spatial structure in the field.

1.4 Hypothesis

In this work, we focus on methods of dimensionality reduction for attractors and their dynamics. The objective of such a reduction is to produce a reduced model that describes an attractor and its dynamics in a low-dimensional state space. In order to do so we require a map between the original and reduced state spaces that preserves the attractor. An inverse of this dimension-reducing map may also be useful for converting reduced states back into original states (which carry a physical interpretation), and for determining the dynamics in the reduced space. These requirements make secant-based projection particularly appealing as a method of dimensionality reduction for attractors, as not only does it provide an explicit map (an orthogonal projection), but it also preserves the attractor by embedding it into the subspace.

Conventional methods of reduction, such as the POD-Galerkin, are applicable to state spaces with a vector space structure and use linear algebra to perform the reduction. In contrast to this, we can take a more general approach by regarding the state space to be a smooth manifold with dynamics provided by a vector field. Such an approach would allow for the reduction of systems with angular variables, which would not be compatible with existing methods, especially in the case of orbits winding around these circular dimensions. This would also allow for the problem to be formulated in terms independent of the algebraic structure/form used to describe the dynamics, which makes it applicable to a wide range of examples.

We are also interested in dynamical systems that have parameter spaces. In this case, one may have a family of attractors, each associated with a different parameter value. We would like the reduced model to respond to the parameter, producing an accurate low-dimensional description of each attractor over a given parameter region.

1.5 Research Questions

Although it has useful properties, one of the downsides to secant-based projection is that the number of secants is of the order N^2, where N is the number of data points sampled from the manifold. This is a potential barrier to scaling the method to manifolds of larger dimension, or of more complex geometry, which would require a larger number of samples to accurately capture. This raises the question of whether the

computational cost resulting from a large number of secants can be reduced, without sacrificing any relevant detail.

In formulating the problem of dimensionality reduction of attractors in terms of smooth manifolds, vector fields and smooth maps, what problems arise in this more general setting and can they be resolved? Does this more general formulation reduce to the conventional description in the Euclidean case?

Can we extend this formulation for a family of attractors corresponding to a parameter space? Can this be used to produce a reduced model parameterized by the original parameter space that produces the correct family of attractors in the low-dimensional space? What about bifurcations?

1.6 Approach and Evaluation Criteria

The approach taken builds on work by Broomhead and Kirby [2, 4]. Their dimension reduction method is based on the preservation of secants, which will be detailed in §3. This has several useful properties that make it particularly suited to the reduction of attractors, which other more common methods (such as the PCA/POD) lack. In order to obtain the dynamics for the reduced space, we will consider both the well-known Galerkin-style approach, and the optimization approach taken in [4].

Motivated by the geometric considerations of a secant-based approach, we will investigate an approach to the problem in terms of smooth manifolds and vector fields, rather than vector spaces and differential equations. This will allow us to focus on the vector field and its orbits, without being constrained to working within the particular algebraic structure used to specify the dynamics. This will be used to incorporate parameter dependence into the reduced model.

In order to evaluate the effectiveness of the method, the family of attractors, and their orbits, of the original and reduced systems will be compared in the reduced space for a variety of examples.

1.7 Contributions

The original contributions made in this work include the following.

- We develop a useful approximation to reduce the computational cost of secant-based methods. In particular, we introduce approximate secants, motivated by the practical constraint of finite data sets, which are used to reduce the total number of secants via a culling procedure. This also gives a method to estimate the dimension of the affine subspace explored by the attractor.

- We develop a geometric formulation of the problem of dimensionality reduction of attractors for general state space manifolds, and extend both secant-based projection and the reproduction of dynamics to this more general setting. We identify and resolve the resulting issues with stability in a manner that allows for practical implementation.

- We extend the formulation of the problem to allow for a parameter space with a corresponding family of attractors. The secant-based projection and reproduction of dynamics are adapted to produce a reduced model parameterized by the original parameter space that can accurately describe the family of attractors in the reduced space.

- Source code of the C++ implementation has been made publicly available at `https://github.com/cwzx/DRDSP`. Funding was obtained to further develop the implementation of this work from the myGrid project as part of EPSRC grant EP/G026238/1.

At the time of writing we have two publications [49, 50] and a third is planned.

1.8 Outline

This thesis is organized as follows.

- In §2 we revise the relevant mathematical background in order to arm the reader with the required vocabulary, notation and reference that is used in the later chapters.

- In §3 we discuss the existing secant-based projection in more detail, and in particular the use of optimization over Grassmann manifolds to obtain a projection. We then introduce a new secant culling procedure to reduce computational cost,

and investigate its properties, which we demonstrate are useful for other purposes, such as estimating the submanifold's dimension of exploration.

- In §4 we formulate an approach to the dimensionality reduction of attractors produced by vector fields on general state space manifolds. We consider the properties required of both the dimension-reducing map, and the reduced vector field, including stability and the derivatives of the vector field. This results in interesting complications in contrast to the linear case, which we resolve. We then adapt the existing secant-Grassmann projection and radial basis function methods for use in this more general setting.

- In §5 we consider the introduction of a parameter space to the problem, where a family of vector fields produces a family of attractors. We extend the geometric formulation to this problem and use it to develop a new method for producing a reduced family of vector fields, parameterized by the original parameter space, that reproduces the family of attractors in a low-dimensional space.

- In §6 we present a number of example systems, including the reproduction of period-doubling bifurcations in the Rössler system, the limit cycles of a non-autonomous double pendulum, the limit cycles of the Brusselator PDE, and a torus in a dynamo PDE.

- In §7 we conclude with a summary, a discussion of the method's limitations, and ideas for future work.

Chapter 2

Background Mathematics

In this chapter, we revise the background mathematics relevant to the work found in the following chapters. In particular we cover multilinear algebra, differential geometry, dynamical systems, and optimization over Riemannian manifolds. The aim is to arm the reader with a concise, modern summary of the concepts and notation that will be used later, while presenting detail when it is relevant.

2.1 Linear and Multilinear Algebra

Linear algebra is the study of vector spaces and their morphisms, while multilinear algebra deals with products of vector spaces, particularly with themselves and their dual, and multilinear maps. Linear algebra is traditionally presented in terms of matrices, especially at an undergraduate level, which allows for a more numerical approach to the subject, while multilinear algebra is often considered separately, as a more abstract extension of linear algebra. However, multilinear algebra can provide a more elegant view of certain aspects of linear algebra, especially when taking a geometric perspective; this is the approach we shall take. [16, 17, 10]

2.1.1 Linear Maps

Given vector spaces V and W over a field of scalars, $\mathbb{F}$, a linear map (also called a linear operator or transformation) $A : V \to W$ preserves the linear structure, i.e. $A(\lambda x + y) = \lambda A(x) + A(y)\ \forall x, y \in V,\ \lambda \in \mathbb{F}$. Linear maps preserve the null vector, $A(0) = 0$, and the composition of linear maps is linear. Both the kernel and image

of a linear map are vector subspaces (of V and W respectively); the dimension of the kernel is known as the *nullity*, while the dimension of the image is called the *rank*. The Rank-Nullity Theorem relates these dimensions: $\dim(\ker A)+\dim(\operatorname{im} A) = \dim V$. We denote the set of linear maps from V to W as $L(V,W)$. This set naturally inherits a linear structure from the vector spaces, with null map $0(x) = 0$, making it an $\mathbb{F}$-vector space of dimension $\dim V \cdot \dim W$. For the case $V = W$, linear operators are endomorphisms of vector spaces, denoted $L(V) = L(V,V)$. This vector space has the additional structure of multiplication in terms of composition, $(AB)(x) = A(B(x))$. This multiplication distributes over addition and is bilinear, with the identity operator $1(x) = \mathrm{id}_V(x) = x$. This gives $L(V)$ the structure of a unital associative algebra. The set of invertible linear operators forms a group under composition, called the general linear group, $\mathrm{GL}(V)$.

2.1.2 Matrix Description

Each linear map $A : V \to W$ can be described by a matrix. A particular matrix is determined by choosing frames for V and W. Let $(e_1, \cdots, e_n)$ and $(f_1, \cdots, f_m)$ be frames for V and W respectively, we can then write $w = A(v)$ as

$$\sum_j w^j f_j = \sum_i v^i A(e_i). \tag{2.1}$$

As $A(e_i)$ is a vector in W, we can express it in the frame of W, $A(e_i) = \sum_j A^j{}_i f_j$,

$$\sum_j w^j f_j = \sum_j \left(\sum_i A^j{}_i v^i \right) f_j. \tag{2.2}$$

This gives us the relation between the components of w and v,

$$w^j = \sum_i A^j{}_i v^i. \tag{2.3}$$

Thus a linear map can be written as the matrix equation,

$$\begin{bmatrix} w^1 \\ \vdots \\ w^m \end{bmatrix} = \begin{bmatrix} A^1{}_1 & \cdots & A^1{}_n \\ \vdots & \ddots & \vdots \\ A^m{}_1 & \cdots & A^m{}_n \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}, \tag{2.4}$$

where the ith column contains the components of $A(e_i)$ in the frame of W. The matrix algebra of $m \times n$ matrices corresponds to the algebra of $L(V,W)$. Therefore invertible

linear operators correspond to invertible matrices. Under a change in bases, described by invertible linear operators $S \in \mathrm{GL}(V)$ and $T \in \mathrm{GL}(W)$, the matrix description transforms according to $A \mapsto TAS^{-1}$.

We make use of the Einstein summation convention, where a repeated index in a term (one superscript, one subscript) implies a summation over that index, $A^i{}_j x^j \Leftrightarrow \sum_j A^i{}_j x^j$, and the sum is over the full range unless otherwise specified.

2.1.3 Norms

A *norm* on a real or complex vector space V is a map $\|\cdot\| : V \to \mathbb{R}$ that satisfies the following properties.

- **Null:** $\|x\| = 0 \;\Leftrightarrow\; x = 0$,
- **Homogeneity**: $\|\lambda x\| = |\lambda| \, \|x\| \quad \forall \lambda \in \mathbb{F}, x \in V$,
- **Triangle inequality**: $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$.

A vector space with a norm is called a *normed vector space*. Norms are non-negative, which provides a notion of 'length' for vectors. A vector with unit length, $\|v\| = 1$, is called a unit vector. Any non-zero vector can be 'normalized' by dividing it by its norm to produce a unit vector, $v \mapsto v/\|v\|$. A norm induces a metric, $d(x, y) = \|x - y\|$, which in turn gives the vector space a topology. For finite-dimensional vector spaces, all norms are topologically equivalent, and under this topology the norm and all linear maps are continuous. We can therefore take continuity for granted in finite-dimensional linear algebra, but this is not the case for infinite-dimensional vector spaces.

2.1.4 Inner Product

An inner product, $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$, maps each pair of vectors in V to a scalar in $\mathbb{F}$ and satisfies the following axioms. Here $\mathbb{F}$ is either $\mathbb{C}$ or $\mathbb{R}$.

- **Positive definiteness**: $\langle x, x \rangle \geq 0 \quad \forall x \in V$
- $\langle x, x \rangle = 0 \Leftrightarrow x = 0$
- **Linearity**: $\langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle \quad \forall x, y, z \in V, \; \lambda \in \mathbb{F}$

- **Conjugate symmetry**: $\overline{\langle y, x\rangle} = \langle x, y\rangle \quad \forall x, y \in V$

This means the inner product is conjugate-linear in the second argument; in the real case, the inner product is linear in both. A vector space with an inner product is called an *inner product space*. In components, the inner product of two arbitrary vectors is $\langle x, y\rangle = g_{ij} x^i \, \overline{y^j}$, where $g_{ij} := \langle e_i, e_j\rangle$. A pair of vectors, $x, y \in V$, are said to be *orthogonal* if $\langle x, y\rangle = 0$.

An inner product induces a norm, via $\|x\|^2 = \langle x, x\rangle$. This is a well-defined norm as a result of the Cauchy-Schwarz inequality, $|\langle x, y\rangle| \leq \|x\| \, \|y\|$, which ensures the induced norm satisfies the triangle inequality. This inequality also motivates the use of the inner product as a definition of angle between two (non-zero) vectors in a real inner product space, via $\cos\theta_{xy} = \langle x, y\rangle \, / \, \|x\| \, \|y\|$, since the right-hand side is bounded in $[-1, 1]$. The angle satisfies $\theta_{xx} = 0$, $\theta_{x,-x} = \pi = 180°$, and, for a pair of orthogonal vectors, $\theta_{xy} = \pi/2 = 90°$.

In the real case, we have the useful addition rule

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,\langle x, y\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\,\|x\| \, \|y\| \cos\theta_{xy}, \end{aligned} \tag{2.5}$$

which is the abstract generalization of Pythagoras' theorem.

2.1.5 Dual Vector Space

Given a vector space, V over $\mathbb{F}$, we can define the *dual space*, V^*, as the set of all linear functionals, $V \to \mathbb{F}$. The dual space inherits the linear structure of a vector space from V, and the elements of V^* are called *covectors*. A finite-dimensional vector space and its dual are isomorphic. Given a basis for V, $(e_1, \cdots, e_n)$, we write the corresponding basis for V^* using superscript notation, $(e^1, \cdots, e^n)$, defined by $e^i(e_j) = \delta^i{}_j$, which implies that for an arbitrary vector $x \in V$, $e^i(x) = x^i$. A particular isomorphism, denoted $* : V \to V^* : v \mapsto v^*$, is induced by a non-degenerate bilinear form B, via $v^*(x) = B(x, v)$. The non-degeneracy property is $B(x, v) = 0 \ \forall x \Rightarrow v = 0$, which ensures that $*$ is an isomorphism.

At this point it is useful to borrow some notation from differential geometry. If (V, g) is a real inner product space, the inner product is a non-degenerate bilinear form, and the pair of isomorphisms induced by the inner product are called the *musical*

isomorphisms, denoted $\flat : V \to V^*$ and $\sharp : V^* \to V$. In components these are given by $(x^\flat)_i = g_{ij}x^j =: x_i$ and $(\omega^\sharp)^i = g^{ij}\omega_j =: \omega^i$, where (g^{ij}) is the matrix inverse of (g_{ij}). Notationally these isomorphisms raise and lower the indices of the components, hence the analogy with musical notation. The induced inner product on V^* is given by $\langle \alpha, \beta \rangle = g(\alpha^\sharp, \beta^\sharp) = g_{ij}\alpha^i\beta^j = g^{ij}\alpha_i\beta_j$.

2.1.6 Direct Sum

The direct sum of two $\mathbb{F}$-vector spaces, V and W, is an $\mathbb{F}$-vector space denoted $V \oplus W$. The set is given by the Cartesian product of the underlying sets, and the vector space structure is given by

$$(v, w) + (x, y) = (v + x, w + y)$$
$$\lambda(v, w) = (\lambda v, \lambda w)$$

The null vector is therefore the pairing of the individual nulls, $(0, 0)$. As the name suggests, the dimension of the sum is given by $\dim(V \oplus W) = \dim V + \dim W$. For multiple products, we use the notation $V^{\oplus k} := \underbrace{V \oplus \cdots \oplus V}_{k \text{ times}}$. The direct sum can be given an inner product in terms of the inner products of V and W,

$$\langle (v, w), (x, y) \rangle_{V \oplus W} = \langle v, x \rangle_V + \langle w, y \rangle_W .$$

Under this inner product, V and W are orthogonal in $V \oplus W$, by identifying V with $(V, 0)$, and W with $(0, W)$.

If W is a subspace of V, and g is an inner product on V, we can define the *orthogonal complement* of W in V with respect to g as

$$W^\perp = \{x \in V \,:\, g(x, w) = 0 \ \forall w \in W\} .$$

This is a subspace of V, which contains all vectors orthogonal to W. The space V can then be written as the orthogonal decomposition $V = W \oplus W^\perp$, i.e. an arbitrary vector is the sum of two orthogonal components.

2.1.7 Tensor Product

The tensor product, denoted $\otimes$, is a bilinear product of of two $\mathbb{F}$-vector spaces, V and W, that produces a new $\mathbb{F}$-vector space. The tensor product space $V \otimes W$ is the

vector space spanned by elements of the form $v \otimes w$. Not all elements of $V \otimes W$ can be expressed as a product of a pair of vectors from V and W – those that can are called *pure* or *decomposable* – but rather a general element is a linear combination of decomposable elements.

Extending the definition to multiple tensor products gives the kth *tensor power* of V,

$$T^k(V) := V^{\otimes k} := \underbrace{V \otimes \cdots \otimes V}_{k \text{ times}}. \tag{2.6}$$

This can be extended to include the dual space, V^*,

$$T^r_s(V) := \underbrace{V \otimes \cdots \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s \text{ times}}. \tag{2.7}$$

Elements of this space are called tensors of type (r, s). Given a basis for V, (e_i) and the corresponding dual basis (e^j), a tensor T is a linear combination

$$T = T^{i_1 \cdots i_r}{}_{j_1 \cdots j_s}\ e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}, \tag{2.8}$$

where $T^{i_1 \cdots i_r}{}_{j_1 \cdots j_s} \in \mathbb{F}$ are the components of the tensor, T, in the basis $(e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s})$. As each index runs over the n dimensions of V, the dimension of the tensor space is $\dim T^r_s(V) = n^{r+s}$.

The tensor product can be applied to a pair of tensors of types (r, s) and (p, q) to produce a tensor of type $(r+p, s+q)$. By convention, the $(0, 0)$ tensor space is defined to be the underlying field, $\mathbb{F}$ (which is a vector space over itself), therefore the tensor product of a scalar and a vector is just scalar multiplication, $\lambda \otimes v = \lambda v$.

Tensors as Multilinear Maps

A tensor of type (r, s) can be interpreted as a multilinear form $(V^*)^r \times V^s \to \mathbb{F}$. As a map, we can evaluate the tensor using r covectors, $(\phi_1, \cdots, \phi_r)$, and s vectors, $(v_1, \cdots, v_s)$, to give a scalar value, via

$$\begin{aligned} &T(\phi_1, \cdots, \phi_r, v_1, \cdots, v_s) \\ &= T^{i_1 \cdots i_r}{}_{j_1 \cdots j_s}\ \phi_1(e_{i_1}) \otimes \cdots \otimes \phi_r(e_{i_r}) \otimes e^{j_1}(v_1) \otimes \cdots \otimes e^{j_s}(v_s) \\ &= T^{i_1 \cdots i_r}{}_{j_1 \cdots j_s}\ (\phi_1)_{i_1} \cdots (\phi_r)_{i_r} (v_1)^{j_1} \cdots (v_s)^{j_s}. \end{aligned} \tag{2.9}$$

Each covector produces a scalar value when evaluated by a vector, and the tensor product of scalars is just scalar multiplication.

Similarly a $(1,1)$ tensor can be regarded as a linear operator $A : V \to V$, by evaluating only the covector part,

$$\begin{aligned} A(v) &= A^i{}_j \; e_i \otimes e^j(v) \\ &= A^i{}_j v^j \; e_i. \end{aligned} \tag{2.10}$$

This is a natural isomorphism between tensors and operators, $V \otimes V^* \cong L(V)$, which generalizes to linear maps between different vector spaces, $W \otimes V^* \cong L(V, W)$.

Trace of a Linear Map

The *trace* is a function $\mathrm{tr} : L(V) \to \mathbb{F}$ that assigns a scalar value to each linear map. It can be defined as a linear functional that acts on pure tensors of type $(1,1)$ by

$$\mathrm{tr}(v \otimes \alpha) = \alpha(v). \tag{2.11}$$

The trace of an arbitrary $(1,1)$ tensor is then given by linearity,

$$\begin{aligned} \mathrm{tr}(A) &= A^i{}_j \; \mathrm{tr}(e_i \otimes e^j) \\ &= A^i{}_j \delta^j{}_i \\ &= A^i{}_i, \end{aligned} \tag{2.12}$$

i.e. it is the sum of the diagonal elements of the matrix description in any basis. The trace has the product rule

$$\mathrm{tr}(AB) = \mathrm{tr}(BA), \tag{2.13}$$

the trace of the null map is 0, and the trace of the identity is the dimension of V.

2.1.8 Exterior Product

The *wedge product* or *exterior product* is a product of a vector space V with itself. Like the tensor product, the wedge product is bilinear, but is also alternating, $x \wedge y = -y \wedge x$ $\forall x, y \in V$. The vector space spanned by elements of the form $x \wedge y$ is denoted $V \wedge V$. Elements that can be expressed as a pair $x \wedge y$ are called *decomposable*.

Extending the definition to multiple wedge products allows us to define the kth *exterior power* of V,

$$\Lambda^k(V) = \underbrace{V \wedge \cdots \wedge V}_{k \text{ times}}. \tag{2.14}$$

Elements of $\Lambda^k(V)$ are called *k-vectors* and elements of $\Lambda^k(V^*)$ are called *k-forms*. Given a basis for V, $(e_1, \cdots, e_n)$, we can write $T \in \Lambda^k(V)$ as a linear combination of the wedge products of the basis vectors, $e_{i_1} \wedge \cdots \wedge e_{i_k}$. However, the alternating property of the wedge product means an arbitrary linear combination is non-unique. By removing terms that are null or linearly dependent, the basis for $\Lambda^k(V)$ can be written as

$$\{e_{i_1} \wedge \cdots \wedge e_{i_k} \,:\, 1 \leq i_1 < \cdots < i_k \leq n\}\,, \tag{2.15}$$

and the component description can be made unique by requiring $T^{i_1 \cdots i_k}$ to be totally anti-symmetric,

$$\begin{aligned} T &= \sum_{i_1 < \cdots < i_k} T^{i_1 \cdots i_k}\; e_{i_1} \wedge \cdots \wedge e_{i_k} \\ &= \frac{1}{k!}\, T^{i_1 \cdots i_k}\; e_{i_1} \wedge \cdots \wedge e_{i_k}. \end{aligned} \tag{2.16}$$

This also allows us to identify k-vectors with anti-symmetric k-tensors,

$$\frac{1}{k!}\, T^{i_1 \cdots i_k}\; e_{i_1} \wedge \cdots \wedge e_{i_k} \;\longleftrightarrow\; T^{i_1 \cdots i_k}\; e_{i_1} \otimes \cdots \otimes e_{i_k}. \tag{2.17}$$

The dimension is given by the binomial coefficient

$$\dim \Lambda^k(V) = \left(\begin{array}{c} n \\ k \end{array} \right) := \frac{n!}{k!(n-k)!}, \tag{2.18}$$

where $n = \dim V$. Note that $\dim \Lambda^k(V) = \dim \Lambda^{n-k}(V)$; for this reason n-vectors

k	$\dim \Lambda^k(V)$
0	1
1	n
2	$\frac{1}{2}n(n-1)$
$\vdots$	$\vdots$
$n-2$	$\frac{1}{2}n(n-1)$
$n-1$	n
n	1

Table 2.1: Dimension of the exterior powers.

are referred to as pseudo-scalars, $(n-1)$-vectors are referred to as pseudo-vectors, and $\Lambda^n(V)$ is referred to as the 'top' exterior power. The terminology 'bivector' is also used for 2-vectors.

A consequence of the alternating property is that the wedge product of set of vectors is only non-zero if the set is linearly independent. This explains why $\Lambda^k(V)$

is the trivial vector space for all $k > n$, since, for an n-dimensional space, a linearly-independent set of vectors contains at most n vectors.

The wedge product of a p-vector α with a q-vector β is a $(p+q)$-vector. When reversing the order of the product, $\alpha \wedge \beta = (-1)^{pq}\ \beta \wedge \alpha$, since there are pq individual wedges to be reversed, each contributing a factor of -1,

$$\alpha \wedge \beta = \frac{1}{p!q!}\alpha^{i_1 \cdots i_p}\beta^{j_1 \cdots j_q}\ e_{i_1} \wedge \cdots \wedge e_{i_p} \wedge e_{j_1} \wedge \cdots \wedge e_{j_q}. \tag{2.19}$$

The component description of the product must be anti-symmetrized, by summing over permutations, which gives

$$(\alpha \wedge \beta)^{i_1 \cdots i_{p+q}} = \frac{1}{p!q!} \sum_{\sigma \in S_{p+q}} \mathrm{sgn}(\sigma)\ \alpha^{i_{\sigma(1)} \cdots i_{\sigma(p)}}\beta^{i_{\sigma(p+1)} \cdots i_{\sigma(p+q)}}, \tag{2.20}$$

where S_k is the Symmetric group of permutations of $\{1, \cdots, k\}$, and sgn() is the signature of the permutation: $+1$ for even and -1 for odd.

Since k-forms can be identified with anti-symmetric $(0, k)$ tensors, they can be used as alternating multilinear forms, $\alpha : V^k \to \mathbb{F}$, given by

$$\alpha(v_1, \cdots, v_k) = \alpha_{i_1 \cdots i_k}\ (v_1)^{i_1} \cdots (v_k)^{i_k}. \tag{2.21}$$

Notation

The Levi-Civita symbol is a useful notation,

$$\varepsilon_{i_1 \cdots i_n} = \varepsilon^{i_1 \cdots i_n} = \begin{cases} +1, & (i_1, \cdots, i_n) \text{ is an even permutation of } (1, \cdots, n), \\ -1, & (i_1, \cdots, i_n) \text{ is an odd permutation of } (1, \cdots, n), \\ 0, & \text{otherwise,} \end{cases}$$

which satisfies

$$\varepsilon_{i_1 \cdots i_n}\varepsilon^{i_1 \cdots i_n} = n!. \tag{2.22}$$

The following anti-symmetrization notation is also convenient,

$$\begin{aligned} T^{[i_1 \cdots i_p]} &:= \frac{1}{p!} \sum_{\sigma \in S_p} \mathrm{sgn}(\sigma)\ T^{i_{\sigma(1)} \cdots i_{\sigma(p)}} \\ &= \frac{1}{p!(n-p)!}\varepsilon^{i_1 \cdots i_p j_1 \cdots j_{n-p}}\varepsilon_{k_1 \cdots k_p j_1 \cdots j_{n-p}}T^{k_1 \cdots k_p}. \end{aligned} \tag{2.23}$$

Therefore the components of the wedge product are given by

$$(\alpha \wedge \beta)^{i_1 \cdots i_{p+q}} = \frac{(p+q)!}{p!q!}\alpha^{[i_1 \cdots i_p}\beta^{i_{p+1} \cdots i_{p+q}]}. \tag{2.24}$$

2.1.9 Determinant

Given a linear operator $A : V \to V$ on an n-dimensional vector space, we can define a corresponding linear operator on the exterior power, $\Lambda^k A : \Lambda^k V \to \Lambda^k V$ via

$$(\Lambda^k A)(v_1 \wedge \cdots \wedge v_k) := A(v_1) \wedge \cdots \wedge A(v_k). \tag{2.25}$$

For the case $k = n$, we have $\dim \Lambda^n V = 1$, and so the action of this linear operator corresponds to scalar multiplication. The determinant of A is then defined as this scalar,

$$A(e_1) \wedge \cdots \wedge A(e_n) = (\det A)\ e_1 \wedge \cdots \wedge e_n. \tag{2.26}$$

Clearly, the inverse of $\Lambda^n A$ is given by $\Lambda^n A^{-1}$, and so the determinant of the inverse must be the reciprocal. Hence invertible linear maps have non-zero determinant. Other useful properties of the determinant are readily obtainable from this definition with very little effort:

- $\det A^{-1} = 1/\det A$,
- $\det 0 = 0$,
- $\det \mathrm{id}_V = 1$,
- $\det(\lambda A) = \lambda^n \det A$,
- $\det \prod_i A_i = \prod_i \det A_i$.

Due to the product rule and the inverse rule, the determinant is independent of the choice of basis. An explicit formula for the determinant can also be obtained from the definition by expanding the LHS by linearity,

$$\det A = \varepsilon_{i_1 \cdots i_n} A^{i_1}{}_1 \cdots A^{i_n}{}_n. \tag{2.27}$$

The first few dimensions are as follows,

$$\begin{aligned} n = 1: &\quad \det A = A^1{}_1, \\ n = 2: &\quad \det A = A^1{}_1 A^2{}_2 - A^2{}_1 A^1{}_2, \\ n = 3: &\quad \det A = A^1{}_1 A^2{}_2 A^3{}_3 - A^1{}_1 A^3{}_2 A^2{}_3 - A^2{}_1 A^1{}_2 A^3{}_3 \\ &\qquad + A^2{}_1 A^3{}_2 A^1{}_3 + A^3{}_1 A^1{}_2 A^2{}_3 - A^3{}_1 A^2{}_2 A^1{}_3. \end{aligned} \tag{2.28}$$

This component formula can be used to define the determinant for an arbitrary matrix, and therefore for the inner product in a particular basis, g_{ij}.

2.1.10 Orientation

Two frames for V that are related by the invertible linear operator $J \in \mathrm{GL}(V)$ are said to have the same *orientation* if $\det J > 0$, else have opposite orientations. This is related to even and odd permutations of the wedge product respectively. This classifies frames on V into two equivalence classes. An *orientation form* is a way of specifying a choice of standard orientation for V. It is a normalized n-form, $\omega \in \Lambda^n(V^*)$, which is unique up to the choice of orientation. Given a particular frame that has the standard orientation, the orientation form can be expressed in terms of the inner product, g, as

$$\omega = \sqrt{\det g}\; e^1 \wedge \cdots \wedge e^n, \tag{2.29}$$

$$\omega_{i_1 \cdots i_n} = \sqrt{\det g}\; \varepsilon_{i_1 \cdots i_n}, \tag{2.30}$$

where the role of the inner product is to provide the normalization factor. A frame $(f_i)_{i=1}^n$ has the standard orientation if $\omega(f_1, \cdots, f_n) > 0$. An orientation form can be thought of as specifying a choice of the 'positive' direction on the 1-dimensional space $\Lambda^n(V^*)$.

2.1.11 Hodge Dual

As noted above, since the dimension of the kth exterior power is equal to that of the $(n-k)$th exterior power, we can seek an isomorphism between the two. The Hodge star $* : \Lambda^k(V^*) \to \Lambda^{n-k}(V^*)$ is a particular isomorphism between k-forms and $(n-k)$-forms that is generated by an inner product and choice of orientation specified by an orientation form, ω. It is defined by

$$*\left(e^{i_1} \wedge \cdots \wedge e^{i_k}\right) = \frac{1}{(n-k)!}\; g^{i_1 j_1} \cdots g^{i_k j_k}\; \omega_{j_1 \cdots j_n}\; e^{j_{k+1}} \wedge \cdots \wedge e^{j_n} \tag{2.31}$$

For a k-form β, this gives

$$(*\beta)_{i_1 \cdots i_{n-k}} = \frac{1}{k!}\; \omega_{i_1 \cdots i_{n-k} j_1 \cdots j_k}\; \beta^{j_1 \cdots j_k} \tag{2.32}$$

Note that $*\omega = 1$, and $*1 = \omega$. The Hodge star also satisfies

$$** \beta = (-1)^{k(n-k)} \beta, \tag{2.33}$$

which means $*^2 = 1$ whenever n is odd, or k is even.

2.1.12 Cross Product

Consider the product of two 1-forms, α and β, defined by $*(\alpha \wedge \beta)$. The result of this product is an $(n-2)$-form, given by

$$(*(\alpha \wedge \beta))_{i_1 \cdots i_{n-2}} = \sqrt{\det g}\ \varepsilon_{i_1 \cdots i_{n-2} j_1 j_2}\ \alpha^{j_1} \beta^{j_2} \tag{2.34}$$

In the case $n = 3$, the result is also a 1-form. Since we can identify 1-forms with vectors, we can use this to define a bilinear anti-symmetric vector product $V \times V \to V$ called the *cross product*,

$$v \times w := (*(v^\flat \wedge w^\flat))^\sharp, \tag{2.35}$$

giving

$$(v \times w)^k = \sqrt{\det g}\ g^{ka} \varepsilon_{aij}\, v^i w^j. \tag{2.36}$$

If the basis is orthogonal, then g is diagonal $(g_{ij}) = \mathrm{diag}(g_1, g_2, g_3)$, giving

$$\begin{aligned}
(v \times w)^1 &= \sqrt{\frac{g_2 g_3}{g_1}}(v^2 w^3 - v^3 w^2) \\
(v \times w)^2 &= \sqrt{\frac{g_3 g_1}{g_2}}(v^3 w^1 - v^1 w^3) \\
(v \times w)^3 &= \sqrt{\frac{g_1 g_2}{g_3}}(v^1 w^2 - v^2 w^1).
\end{aligned} \tag{2.37}$$

In an orthonormal basis this reduces to the well-known *ad hoc* construction from vector calculus,

$$\begin{aligned}
(v \times w)^1 &= v^2 w^3 - v^3 w^2 \\
(v \times w)^2 &= v^3 w^1 - v^1 w^3 \\
(v \times w)^3 &= v^1 w^2 - v^2 w^1.
\end{aligned} \tag{2.38}$$

2.1.13 Eigenvalues

Consider a linear operator $A : V \to V$. The eigenvalue equation is

$$\begin{aligned}
A(v) &= \lambda v \\
(\lambda 1 - A)(v) &= 0
\end{aligned} \tag{2.39}$$

Clearly this has the trivial solution, $v = 0$. For $(\lambda 1 - A)$ invertible, this is the only solution. Hence to find nontrivial solutions we consider the non-invertible case. A linear operator is non-invertible when its determinant is zero,

$$\det(\lambda 1 - A) = 0 \tag{2.40}$$

The values of λ that satisfy this equation are the *eigenvalues* of A. Let $n = \dim V$. Using the definition of the determinant,

$$\det(\lambda 1 - A)\ e_1 \wedge \cdots \wedge e_n = (\lambda 1 - A)(e_1) \wedge \cdots \wedge (\lambda 1 - A)(e_n), \tag{2.41}$$

and expanding the brackets via the multilinearity of the wedge gives a degree n polynomial in λ, called the *characteristic polynomial*,

$$p_A(\lambda) := \det(\lambda 1 - A) = \sum_{i_1 \cdots i_n} \varepsilon_{i_1 \cdots i_n} \prod_{j=1}^{n} (\lambda \delta^{i_j}_{\ j} - A^{i_j}_{\ j}). \tag{2.42}$$

For real or complex vector spaces, the polynomial has n roots in the complex plane. For example, the case $n = 2$ gives

$$\lambda^2 - (A^1_{\ 1} + A^2_{\ 2})\lambda + (A^1_{\ 1} A^2_{\ 2} - A^2_{\ 1} A^1_{\ 2}) = 0. \tag{2.43}$$

For each eigenvalue λ_i, the vectors that satisfy the corresponding eigenvalue equation are the *eigenvectors*,

$$A(v_i) = \lambda_i v_i. \tag{2.44}$$

Each eigenvector is determined up to scalar multiplication, i.e. if v_i is an eigenvector for λ_i, then μv_i is also an eigenvector for λ_i, for all $\mu \in \mathbb{F}$. For this reason we can choose to normalize the eigenvectors (if we have a norm). The set of eigenvectors corresponding to a given eigenvalue form a subspace, called an eigenspace. If a set of eigenvectors form a basis for V we call this an 'eigenbasis'. If $(a_1, \cdots, a_n)$ is an eigenbasis of A, with eigenvalues $(\lambda_1, \cdots, \lambda_n)$, then in this basis A becomes diagonal,

$$A = \sum_i \lambda_i\ a_i \otimes a^i. \tag{2.45}$$

The determinant of a linear operator is the product of its eigenvalues,

$$\begin{aligned} A(a_1) \wedge \cdots \wedge A(a_n) &= \lambda_1 \cdots \lambda_n\ a_1 \wedge \cdots \wedge a_n \\ &= (\det A)\ a_1 \wedge \cdots \wedge a_n, \end{aligned} \tag{2.46}$$

$$\det A = \prod_{i=1}^{n} \lambda_i. \tag{2.47}$$

The trace can also be expressed in terms of the eigenvalues,

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i. \tag{2.48}$$

2.1.14 Adjoint

Given two inner product spaces, (V, g) and (W, h), and a linear map $A : V \to W$, we can define the adjoint $A^\dagger : W \to V$, via

$$g(v, A^\dagger(w)) = h(A(v), w) \quad \forall v \in V,\ w \in W. \tag{2.49}$$

The adjoint is also a linear map. In components this is

$$(A^\dagger)^i{}_a = h_{ab}\ \overline{A^b{}_j}\ g^{ji}. \tag{2.50}$$

For orthonormal frames, $g_{ij} = \delta_{ij}$ and $h_{ab} = \delta_{ab}$, the adjoint corresponds to the matrix conjugate-transpose, but this is not the case for non-orthonormal frames. The map $\dagger : L(V, W) \to L(W, V)$ is conjugate-linear and an involution, $(A^\dagger)^\dagger = A$ and $(AB)^\dagger = B^\dagger A^\dagger$. In the real case, the complex conjugate disappears, so the conjugate-linearity of $\dagger$ becomes full linearity. The determinant of the adjoint is given by $\det A^\dagger = \det g \cdot \overline{\det A} \cdot \det g^{-1} = \overline{\det A}$. For an invertible map, the adjoint is also invertible, with $(A^\dagger)^{-1} = (A^{-1})^\dagger$.

2.1.15 Orthonormal Maps

An *orthonormal* map $A : V \to W$ is a map between inner product spaces that preserves the inner product,

$$g(x, y) = h(A(x), A(y)) \quad \forall x, y \in V, \tag{2.51}$$

i.e. it is a linear isometry. It follows that the adjoint is the left-inverse of an orthonormal map, $A^\dagger A = \mathrm{id}_V$. This means that A is injective, and therefore W must be large enough to accommodate V, i.e. $\dim W \geq \dim V$. Orthonormal maps preserve the norm and angles (which are induced by the inner product).

2.1.16 Orthogonal Operators

Given a symmetric bilinear form, g, an *orthogonal* operator is an invertible operator $A : V \to V$ that preserves g, i.e. $g(A(x), A(y)) = g(x, y)$. These operators collectively form a group under composition, called the orthogonal group, denoted $\mathrm{O}(V, g)$.

On real inner product spaces, the inner product is a symmetric bilinear form. For the corresponding orthogonal operators, the adjoint is the inverse, $A^{-1} = A^\dagger$. It follows

that the determinant has unit modulus, $|\det A| = 1$. Orthogonal operators describe rotations of the vector space about the null vector. Rotations can be classified as *proper* (those with $\det A = 1$) or *improper* (those with $\det A = -1$). Proper rotations preserve the orientation of the space, while improper rotations reverse it. The proper rotations form a subgroup, called the special orthogonal group, $\mathrm{SO}(V, g)$, and coincide with the conventional usage of the word 'rotation'.

By choosing an orthonormal frame for V, the operators are identified with invertible matrices, the adjoint with the matrix transpose, and the inner product with the dot product. The corresponding orthogonal groups of matrices are denoted $\mathrm{O}(n, \mathbb{R})$ and $\mathrm{SO}(n, \mathbb{R})$, or just $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, where $n = \dim V$.

2.1.17 Projections

A linear operator $P : V \to V$ that satisfies $P^2 = P$ is called a *projection*, or a *projector*. This means a projection leaves vectors in its image unchanged, $P(x) = x \ \forall x \in P(V)$. Therefore, two projections with the same image agree on the image, but may have different kernels, i.e. the way in which they 'collapse' V onto $P(V)$ differs. A projection is determined by specifying its image and kernel. This is because we can write $V = \operatorname{im} P \oplus \ker P$ (since these subspaces are linearly independent and their dimensions sum to $\dim V$), and the behaviour of P on both of these subspaces is fixed (it preserves the image and nullifies the kernel). It is common to refer to the dimension of the subspace $P(V)$ as the dimension of the projection. Let $A : V \to W$ and $B : W \to V$. The composition BA is a projection if $AB = \mathrm{id}_W$.

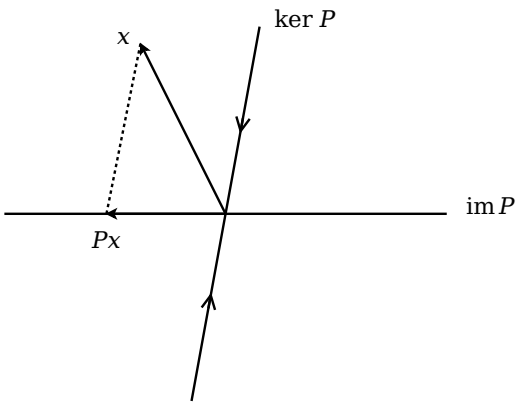


Figure 2.1: Projection P.

An *orthogonal projection* is a projection whose image is orthogonal to its kernel, i.e. $g(Px, y - Py) = 0$ for all $x, y \in V$. It follows that $P = P^\dagger P$. Since $P = P^2$, this corresponds to $P = P^\dagger$, i.e. P is self-adjoint. The kernel of an orthogonal projection is therefore the orthogonal complement of the image, $\ker P = (\operatorname{im} P)^\perp$. For a given subspace, the orthogonal projection onto that subspace is unique, since its kernel is determined by its image (for a fixed inner product).

If A is an orthonormal map then $AA^\dagger$ is an orthogonal projection onto the image of A, since $A^\dagger A = \mathrm{id}_W$. We can therefore write an arbitrary $x \in V$ as the orthogonal decomposition

$$x = AA^\dagger x + (\mathrm{id}_V - AA^\dagger)x, \tag{2.52}$$

where $AA^\dagger x$ is the component of x in the subspace $A(W)$, and $(\mathrm{id}_V - AA^\dagger)x$ is the component in the orthogonal complement, $A(W)^\perp$.

The norm of a projection satisfies,

$$\|Px\| \leq \|x\|, \tag{2.53}$$

with equality for $x \in \operatorname{im} P$, and $\|Px\| = 0$ for $x \in \ker P$. The inequality follows by decomposing an arbitrary x into these two components, and using the triangle inequality of the norm. For an orthogonal projection described by an orthonormal map A, it follows that

$$\left\|A^\dagger x\right\| \leq \|x\|. \tag{2.54}$$

2.1.18 Singular Value Decomposition

In the context of linear maps, the *singular value decomposition* (SVD) can be described as a decomposition of a linear map $V \to W$, thought of as an element of $W \otimes V^*$, into a weighted sum of pure components. Let $A \in W \otimes V^*$, $n = \dim V$ and $m = \dim W$. The SVD of A is a decomposition of the form

$$A = \sum_{k=1}^{\min(m,n)} \sigma_k \, w_k \otimes (v_k)^\flat, \tag{2.55}$$

where $(\sigma_k)_{k=1}^{\min(m,n)}$ are non-negative real numbers called the *singular values*, and both $(w_k)_{k=1}^m$ and $(v_k)_{k=1}^n$ are orthonormal frames (for W and V respectively) called the left- and right-*singular vectors* respectively. The terms in the decomposition are ordered

such that $\sigma_1 \geq \cdots \geq \sigma_{\min(m,n)}$. The pure components are mutually orthogonal and linearly independent, which means that if r is the number of non-zero singular values, then $r = \text{rank}(A)$. Furthermore, the first r left-singular vectors provide an orthonormal frame for the image of A. By evaluating the map with an arbitrary $x \in V$, we get

$$A(x) = \sum_{k=1}^{\min(m,n)} g(x, \sigma_k v_k)\, w_k. \tag{2.56}$$

Under the adjoint, the left- and right-singular vectors are exchanged,

$$A^\dagger = \sum_{k=1}^{\min(m,n)} \sigma_k\, v_k \otimes (w_k)^\flat. \tag{2.57}$$

Also note that $A^\dagger(w_i) = \sigma_i\, v_i$ and $A(v_i) = \sigma_i\, w_i$ for $1 \leq i \leq \min(m,n)$, and null otherwise. Combining these gives $AA^\dagger(w_i) = \sigma_i^2 w_i$ and $A^\dagger A(v_i) = \sigma_i^2 v_i$, i.e. the singular values are the square roots of the (largest) eigenvalues of $AA^\dagger$ and $A^\dagger A$, and the singular vectors are their respective eigenvectors. In components,

$$A^i{}_j = \sum_{k=1}^{\min(m,n)} \sigma_k\, (w_k)^i (v_k)_j = g_{jc} \sum_{k=1}^{\min(m,n)} \sigma_k\, (w_k)^i \overline{(v_k)^c}. \tag{2.58}$$

Given orthonormal frames for V and W, the matrix description can be written as the matrix factorization,

$$A = \mathcal{W}\Sigma\mathcal{V}^\dagger, \tag{2.59}$$

where $\mathcal{W} \in \mathbb{F}^{m\times m}$ and $\mathcal{V} \in \mathbb{F}^{n\times n}$ are unitary matrices whose columns describe the left- and right-singular vectors respectively, $\Sigma \in \mathbb{F}^{m\times n}$ is a diagonal matrix containing the singular values, and $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

Low Rank Approximations

The SVD can be used to obtain an approximation of A with a given rank of $d < r$. This is given by truncating the SVD expansion after the first d terms,

$$A_d = \sum_{k=1}^{d} \sigma_k\, w_k \otimes (v_k)^\flat, \tag{2.60}$$

i.e. by keeping the largest d singular values, and discarding the rest.

Using the norm induced by the inner product on $W \otimes V^*$, the error in this approximation is bounded by the sum of the discarded singular values,

$$\begin{aligned} \|A - A_d\| &= \left\| \sum_{k=d+1}^{r} \sigma_k \, w_k \otimes (v_k)^\flat \right\| \\ &\leq \sum_{k=d+1}^{r} \sigma_k, \end{aligned} \tag{2.61}$$

where we have used the triangle inequality of the norm, and

$$\left\| w_k \otimes (v_k)^\flat \right\| = \|w_k\| \, \|v_k\| = 1. \tag{2.62}$$

It can be shown that A_d is the best rank d approximation of A in this norm, by the Eckart-Young theorem.

2.2 Differential Geometry

Differential geometry involves extending many of the concepts from multilinear algebra to vector bundles over differentiable manifolds. Doing so allows for the study of curved spaces and quantities defined on them, such as vector fields. [25, 28, 48, 22]

2.2.1 Vector Bundles

A vector bundle can be thought of as a parameterized family of vector spaces. Let F be a finite-dimensional vector space, E and B be spaces, called the *total* and *base* spaces respectively, and $\pi : E \to B$ be a continuous surjection called the *bundle projection*. The total space is *locally trivial*, i.e. for each open set $U \subset B$, $\pi^{-1}(U)$ is homeomorphic to $U \times F$. This means that locally E looks like the product space $B \times F$. A *local trivialization* for U is a particular homeomorphism, $\varphi : \pi^{-1}(U) \to U \times F$, that maps $\pi^{-1}(\{x\})$ to $\{x\} \times F$, which allows us to identify $E_x := \pi^{-1}(\{x\})$, the *fibre* at x, with F. We say E is a *vector bundle* over B with fibre F.

One may take a slightly more general definition, in which the dimension of the fibre is not required to be constant over the base, however, the continuity of the projection means that the dimension must at least be constant over connected components.

For each pair of local trivializations, $(\varphi_\alpha, U_\alpha)$ and (φ_β, U_β), with non-empty intersection, we can define the transition map $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{GL}(F)$ via $(x, g_{\alpha\beta}(x)(v)) =$

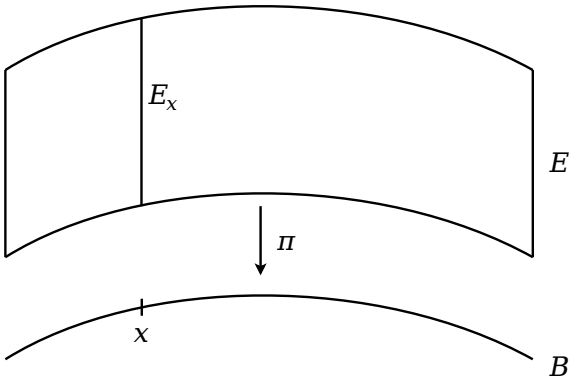


Figure 2.2: Vector bundle.

$\varphi_\alpha \varphi_\beta^{-1}(x, v)$. These are invertible linear operators that convert between different local descriptions of the bundle; they satisfy $g_{\alpha\alpha}(x) = \mathrm{id}_F$, $g_{\alpha\beta}(x)^{-1} = g_{\beta\alpha}(x)$, and the cocycle condition on triple intersections, $g_{\alpha\beta}(x) g_{\beta\gamma}(x) = g_{\alpha\gamma}(x)$.

Bundle Maps

A morphism between vector bundles $\pi_1 : E_1 \to B_1$ and $\pi_2 : E_2 \to B_2$ is given by a pair of continuous maps, $\psi : E_1 \to E_2$ and $\phi : B_1 \to B_2$, such that $\pi_2 \circ \psi = \phi \circ \pi_1$ and ψ acts linearly on fibres, i.e. $\pi_1^{-1}(\{x\}) \to \pi_2^{-1}(\{\phi(x)\})$ is a linear map for all $x \in B_1$.

$$\begin{array}{ccc} & \overset{\psi}{} & \\ E_1 & \longrightarrow & E_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ B_1 & \longrightarrow & B_2 \\ & \underset{\phi}{} & \end{array}$$

Sections

A *section* of a vector bundle is a continuous map $s : B \to E$ such that $\pi \circ s = \mathrm{id}_B$, i.e. it is a continuous choice of an element in F for each point in B. The space of sections is denoted $\Gamma(E)$. This space can be given an algebraic structure by pointwise vector addition and pointwise scalar multiplication by continuous functions $C(B, \mathbb{F})$. This is not quite a vector space, since $C(B, \mathbb{F})$ is only a ring, rather than a field, and so $\Gamma(E)$ is a module over $C(B, \mathbb{F})$. Alternatively, $\Gamma(E)$ can be regarded as a (usually infinite-dimensional) vector space over $\mathbb{F}$.

Due to their local triviality, vector bundles admit a local basis of sections, i.e. for each open set, one can choose a set of sections $(e_1, \cdots, e_n)$ defined on the open set which, at each point, provide a basis for the corresponding fibre. However, only trivial bundles admit a global basis of sections.

Fibrewise Operations

Most of the operations and products of vector spaces discussed in §2.1 can be extended to vector bundles over the same base space by *fibrewise* application. At each point on the base each bundle has a fibre that is a vector space, and these vector spaces can be producted in the usual way. For example, the direct sum bundle $E_1 \oplus E_2$, the tensor product bundle $E_1 \otimes E_2$, the exterior power bundle $\Lambda^k E$ and the dual vector bundle E^*.

2.2.2 Manifolds

In differential topology and differential geometry, we deal with a base space that has additional structure, in that it locally looks like $\mathbb{R}^n$. This allows us to use our existing knowledge of $\mathbb{R}^n$ to describe objects that are globally very different. An n-dimensional topological manifold is a second-countable Hausdorff space, $\mathcal{M}$, that is locally homeomorphic to $\mathbb{R}^n$. An open set U with a corresponding map, $\varphi : U \to \mathbb{R}^n$, that maps U homeomorphically to an open subset of $\mathbb{R}^n$, is called a *chart*. A collection of charts that covers $\mathcal{M}$ is called an *atlas*. Therefore each chart provides a local description of the manifold via a set of n numbers called *coordinates*, $(x^1, \cdots, x^n) \in \mathbb{R}^n$.

As with vector bundles, we can define transition maps for manifolds. Transition maps give a 'change of coordinates', from one copy of $\mathbb{R}^n$ to another. Consider two charts $(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) with non-empty intersection, $U_{\alpha\beta} = U_\alpha \cap U_\beta$. The transition map $g_{\alpha\beta} : \varphi_\beta(U_{\alpha\beta}) \to \varphi_\alpha(U_{\alpha\beta})$ is given by $g_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}|_{\varphi_\beta(U_{\alpha\beta})}$. The transition maps satisfy $g_{\alpha\alpha} = \mathrm{id}$, $g_{\beta\alpha} = (g_{\alpha\beta})^{-1}$, and $g_{\alpha\gamma} = g_{\alpha\beta} \circ g_{\beta\gamma}$ on triple intersections.

Additional structure can be given to the manifold by requiring the transition maps to satisfy additional properties. In differential topology, we require the transition maps to be differentiable. An atlas is said to be of differentiability class C^k if all of its transition maps are k-times differentiable (as partial functions from $\mathbb{R}^n$ to $\mathbb{R}^n$). Since the union of two C^k atlases is a C^k atlas, there is a *maximal* C^k atlas that contains

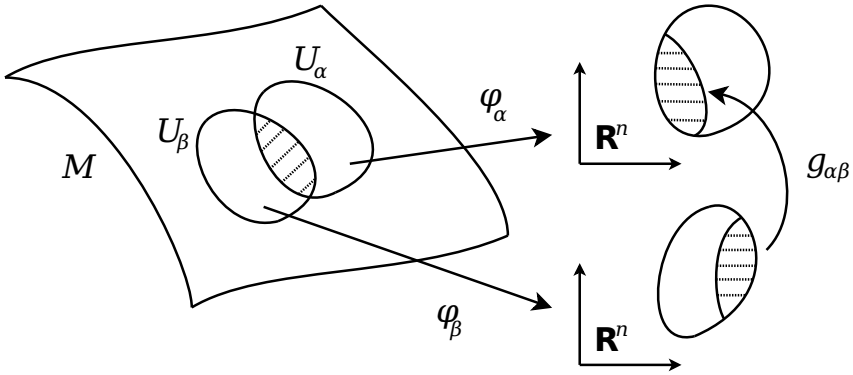


Figure 2.3: Transition map between charts.

all others. This maximal C^k atlas is called the *differential structure* of $\mathcal{M}$, which can be thought of as an object containing all coordinate systems on $\mathcal{M}$. A manifold with such a differential structure is described as a C^k manifold. A C^∞ manifold is called a smooth manifold.

Maps

A map between two manifolds, $\phi : \mathcal{M} \to \mathcal{N}$ is C^k if it is k-times differentiable in any pair of charts, e.g. for charts (φ, U) on $\mathcal{M}$ and (ψ, V) on $\mathcal{N}$, the map $\psi \circ \phi \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ is k-times differentiable. A smooth map is a C^∞ map and a C^k-diffeomorphism is a C^k bijection with C^k inverse.

For a function $f : \mathcal{N} \to \mathbb{R}$, the *pullback* of f by ϕ is a function $\phi^* f : \mathcal{M} \to \mathbb{R}$ given by $\phi^* f = f \circ \phi$.

The Tangent Bundle

Manifolds and vector bundles are similar topological constructions in that both have a notion of simple local descriptions that can be 'glued' together to describe the whole object. When given a differential structure, a manifold possesses a canonical vector bundle called the *tangent bundle*, whose fibres are n-dimensional real vector spaces called *tangent spaces*. As the name suggests, the tangent space at a particular point contains the velocities by which one can move along the manifold. Consider a point $p \in \mathcal{M}$, with a path $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ such that $\gamma(0) = p$. We can define the notion of a tangent vector to the path at p by taking a derivative along the path – however we first need to assign some values in $\mathbb{R}$ to the path in order to have the algebraic

structure required by the ordinary derivative. Consider a smooth function $f : \mathcal{M} \to \mathbb{R}$ on the manifold. This can be composed with the path γ to give $f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}$, which can be differentiated at p, giving a linear map $f \mapsto (f \circ \gamma)'(0)$. This allows us to define an equivalence relation on paths through p, as those that produce the same derivative for each function f. Each equivalence class is a *derivation* on the vector space of smooth functions, i.e. a linear map $X_p : C^\infty(\mathcal{M}, \mathbb{R}) \to \mathbb{R}$ that satisfies the product rule,

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g), \tag{2.63}$$

where function addition and multiplication are defined pointwise. The set of derivations is a real n-dimensional vector space called the tangent space, denoted $\mathrm{T}_p\mathcal{M}$. Applying this construction to each point on the manifold produces a real n-dimensional vector bundle over $\mathcal{M}$, denoted $\mathrm{T}\mathcal{M}$, called the tangent bundle, with fibres $\mathrm{T}_p\mathcal{M}$. For a C^k manifold, the tangent bundle is itself a manifold, possessing a C^{k-1} differential structure.

The coordinate description of the base manifold provided by the charts gives a natural basis for the tangent space at each point, called the coordinate basis. Let $(x^1, \cdots, x^n) \in \mathbb{R}^n$ be local coordinates in a neighbourhood of $p \in \mathcal{M}$. The corresponding partial derivatives $(\partial_1, \cdots, \partial_n)$ at p are a set of n linearly-independent derivations that provide a basis for $\mathrm{T}_p\mathcal{M}$, i.e. an arbitrary tangent vector is $X_p = X_p^i\, \partial_i$. The partial derivative ∂_i is the tangent vector in the direction of increasing x^i. The corresponding dual basis is denoted $(\mathrm{d}x^1, \cdots, \mathrm{d}x^n)$, which describes the *cotangent* space $\mathrm{T}_p^*\mathcal{M}$, and the cotangent bundle, $\mathrm{T}^*\mathcal{M}$. The meaning of this notation will be discussed in §2.2.5.

Scalar and Vector Fields

A *scalar field* is a section of the trivial bundle $\mathcal{M} \times \mathbb{R}$, i.e. a continuous $\mathbb{R}$-valued function on $\mathcal{M}$. We shall deal with smooth scalar fields unless otherwise specified. A *vector field* on $\mathcal{M}$ is a section of the tangent bundle, $X \in \Gamma(\mathrm{T}\mathcal{M})$. We usually deal with C^k vector fields, denoted $\mathfrak{X}^k(\mathcal{M})$ or smooth vector fields, $\mathfrak{X}(\mathcal{M}) = \mathfrak{X}^\infty(\mathcal{M})$. As tangent vectors are defined as derivations on functions, the action of a vector field on a scalar field, f, gives another scalar field, denoted Xf, which is given pointwise by $(Xf)(p) = X_p(f)$. We can also define a Lie bracket on smooth vector fields as the

commutator

$$[X, Y]f := X(Yf) - Y(Xf). \tag{2.64}$$

In a general basis $(E_i)_{i=1}^n$ of vector fields this is

$$[X, Y]^i = X(Y^i) - Y(X^i) + C^i{}_{jk} X^j Y^k. \tag{2.65}$$

where $[E_j, E_k] = C^i{}_{jk} E_i$. This indexed set of scalar fields, $C^i{}_{jk}$, is anti-symmetric in the lower two indices, $C^i{}_{jk} = -C^i{}_{kj}$. If the scalar fields $C^i{}_{jk}$ vanish everywhere (i.e. the basis commutes), we call the basis *holonomic*. For a coordinate basis, the basis vectors are the partial derivatives, $E_i = \partial_i$, which commute and are therefore holonomic.

A tensor field of type (r, s) is a section of the tensor power bundle, $T^r_s(\mathrm{T}\mathcal{M})$.

Maps

Given a C^1 map, $\phi : \mathcal{M} \to \mathcal{N}$, there is a corresponding map between the tangent bundles, called the *differential*, $\phi_* : \mathrm{T}\mathcal{M} \to \mathrm{T}\mathcal{N}$, which is defined by $\phi_*(p, X_p) = (\phi(p), (\phi_*)_p X_p)$ where $(\phi_*)_p$ is a linear map between fibres $(\phi_*)_p : \mathrm{T}_p\mathcal{M} \to \mathrm{T}_{\phi(p)}\mathcal{N}$ defined by

$$((\phi_*)_p X_p) f = X_p(\phi^* f). \tag{2.66}$$

In coordinates this is

$$((\phi_*)_p)^i{}_j = \partial_j \phi^i|_p, \tag{2.67}$$

The differential has alternate notations $D\phi$, $\mathrm{d}\phi$ (not to be confused with the exterior derivative), and $\mathrm{T}\phi$. If $(\phi_*)_x$ is injective for each $x \in \mathcal{M}$ then ϕ is called an *immersion*. If $(\phi_*)_x$ is surjective for each $x \in \mathcal{M}$, then ϕ is called a *submersion*. An *embedding* is an injective immersion that is also a homeomorphism onto its image. An injective immersion of a compact manifold is an embedding.

If ϕ is an injective immersion, the differential can be applied to a vector field X on $\mathcal{M}$ to give a vector field $\phi_* X$ on $\phi(\mathcal{M})$ called the *pushforward* of X by ϕ, giving $(\phi_* X) f = X(\phi^* f)$ and $(\phi_* X)^i = \partial_j \phi^i X^j$. The pushforward of the Lie bracket is given by the Lie bracket of the pushforward,

$$\phi_*[X, Y] = [\phi_* X, \phi_* Y]. \tag{2.68}$$

Paths and Loops

A *path* on $\mathcal{M}$ is a continuous map $\gamma : I \to \mathcal{M}$, where I is an interval of the real line, and a *loop* is a continuous map $\gamma : S^1 \to \mathcal{M}$. By interpreting γ as a map between manifolds, we can define the tangent to a path/loop in terms of the differential γ_* by choosing the value $1 \in \mathbb{R}$ to be pushed-forward, corresponding to $\partial_t \in \mathrm{T}_t I$. Note that both $\mathrm{T}_t I$ and $\mathrm{T}_t S^1$ coincide with $\mathbb{R}$, as they are both 1-dimensional. This allows us to define the 'time derivative' of the path/loop,

$$\dot{\gamma}(t) = (\gamma_*)_t(\partial_t) \in \mathrm{T}_{\gamma(t)}\mathcal{M}. \tag{2.69}$$

One can regard the tangent vector as the velocity along the path, and the choice of 1 as the corresponding rate of increase in the parameterization.

Flow of a Vector Field

Given a smooth vector field, X, at each $p \in \mathcal{M}$ the vector field defines an ODE in local coordinates for a smooth path $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$,

$$\dot{\gamma}(t) = X_{\gamma(t)}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\gamma^i(t) = X^i_{\gamma(t)}, \tag{2.70}$$

with $\gamma(0) = p$. As a result of the existence and uniqueness of solutions for ODEs, there is exactly one such path that satisfies this equation, called an *integral curve* of X. One can use these integral curves to define diffeomorphisms, $\Phi^t_X : \mathcal{M} \to \mathcal{M}$, for small t, by moving each point along its local integral curve,

$$\Phi^t_X(p) = \gamma(t). \tag{2.71}$$

Note that γ depends on both p and X, as does its domain, $(-\varepsilon, \varepsilon)$.

If the integral curves can be defined for all $t \in \mathbb{R}$, the vector field is called *complete*, and Φ^t_X is a diffeomorphism for any t. The map $t \mapsto \Phi^t_X$ is therefore a one-parameter group of diffeomorphisms, $\mathbb{R} \to \mathrm{Diff}(\mathcal{M})$, with multiplication $\Phi^t_X \circ \Phi^s_X = \Phi^{t+s}_X$, identity $\Phi^0_X = \mathrm{id}_{\mathcal{M}}$, and inverse $(\Phi^t_X)^{-1} = \Phi^{-t}_X$.

Lie Derivative

The Lie derivative is a way of defining the derivative of a vector field over a manifold. Its definition is motivated by analogy with the ordinary derivative of elementary

calculus,

$$\begin{aligned}(\mathcal{L}_X Y)_p &= \lim_{t\to 0}\frac{1}{t}\left((\Phi_X^{-t})_* Y_{\Phi_X^t(p)} - Y_p\right), \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_X^{-t})_* Y_{\Phi_X^t(p)},\end{aligned} \tag{2.72}$$

where the vector field X provides the 'direction' in which the derivative of Y is being taken. In other words, this is the derivative of Y along the integral curves of X. The result is also a vector field.

It can be shown that this coincides with the Lie bracket,

$$\mathcal{L}_X Y = [X, Y]. \tag{2.73}$$

As a result, some authors introduce the Lie derivative as the Lie bracket. Although this is simpler, it is void of motivation and lacks direct interpretation. However, it does demonstrate that the Lie derivative is antisymmetric, $\mathcal{L}_X Y = -\mathcal{L}_Y X$, and that the result depends on the first partial derivatives of both X and Y. This means that in order to compute the Lie derivative at a point, one needs both vector fields X and Y to be defined in a neighbourhood of the point.

The Lie derivative is extended to covector fields by the product rule

$$X(\alpha(Y)) = (\mathcal{L}_X \alpha)(Y) + \alpha(\mathcal{L}_X Y), \tag{2.74}$$

and then to arbitrary tensor fields by

$$\mathcal{L}_X(T \otimes S) = \mathcal{L}_X T \otimes S + T \otimes \mathcal{L}_X S, \tag{2.75}$$

where the Lie derivative of a scalar field is just the usual derivative, $\mathcal{L}_X f = X(f)$.

Metric Tensors

A metric tensor for a vector bundle E is a smooth section of $E^* \otimes E^*$ that provides a symmetric non-degenerate bilinear form on each fibre. If g is a metric then at each point we have a $g_p : E_p \times E_p \to \mathbb{F}$ that can be applied to sections X, Y to produce a scalar function $g(X, Y)$,

$$g(X, Y)(p) = g_p(X_p, Y_p). \tag{2.76}$$

It therefore provides musical isomorphisms between the bundle and its dual, $\flat : E \to E^*$, $\sharp : E^* \to E$. For the tangent bundle it is a type $(0, 2)$ tensor field, which in

coordinates is

$$g = g_{ij}\ \mathrm{d}x^i \otimes \mathrm{d}x^j, \tag{2.77}$$

and is applied to a pair of vector fields to produce a scalar field,

$$g(X, Y) = g_{ij} X^i Y^j. \tag{2.78}$$

We can extend the metric to arbitrary tensor fields, S, T of type (p, q),

$$g(S, T) = g_{i_1 j_1} \cdots g_{i_p j_p} g^{a_1 b_1} \cdots g^{a_q b_q} S^{i_1 \cdots i_p}_{a_1 \cdots a_q} T^{j_1 \cdots j_p}_{b_1 \cdots b_q}.$$

i.e. corresponding indices are contracted with the usual metric. Some authors include a pre-factor of $1/p!q!$ for later convenience, we shall not.

A Riemannian metric is a metric tensor for the tangent bundle that is positive definite. It therefore gives a fibrewise inner product, which can be used to induce a fibrewise norm on the tangent bundle. A smooth manifold with a Riemannian metric is called a Riemannian manifold.

Metrics that are not positive definite are often described as pseudo-Riemannian, or semi-Riemannian. These are commonly encountered in modern physics, where both time and physical space are combined into a single spacetime manifold, with the temporal dimension contributing a negative eigenvalue to the spacetime metric.

Pullback Metric

If $\phi : \mathcal{M} \to \mathcal{N}$ is an injective immersion, and h is a metric on $\mathcal{N}$, we can use ϕ to induce a metric g on $\mathcal{M}$, via

$$g(X, Y) = h(\phi_* X, \phi_* Y). \tag{2.79}$$

In coordinates this is

$$g_{ij} = h_{ab}\ \phi^a{}_{,i}\ \phi^b{}_{,j}, \tag{2.80}$$

where we have made use of the comma notation for partial derivatives, $\phi^a{}_{,i} := \partial_i \phi^a$. This is also called the pullback metric of h by ϕ, denoted $\phi^* h$.

Adjoint

Let $\phi : \mathcal{M} \to \mathcal{N}$ be an injective C^1 map between Riemannian manifolds. Then we can define the adjoint of the differential, $\phi^\dagger_* : \mathrm{T}_{\phi(\mathcal{M})}\mathcal{N} \to \mathrm{T}\mathcal{M}$, as the map that sends

$(\phi(x), w)$ to $(x, ((\phi_*)_x)^\dagger w)$, i.e. we are applying the linear algebraic adjoint (§2.1.14) to the differential at each point, and using the injectivity of ϕ to select which tangent space. Note that although we require ϕ to be injective, ϕ_* may not be injective.

If ϕ is an immersion and the metric on $\mathcal{M}$ is the pullback by ϕ, then the differential is an orthonormal linear map at each point and we have $\phi_*^\dagger \phi_* = \mathrm{id}_{\mathrm{T}\mathcal{M}}$. Also $\phi_* \phi_*^\dagger : \mathrm{T}_{\phi(\mathcal{M})}\mathcal{N} \to \mathrm{T}_{\phi(\mathcal{M})}\mathcal{N}$ is the orthogonal projection onto $\phi_*(\mathrm{T}\mathcal{M})$ at each point.

Normalizing the Coordinate Basis

A coordinate chart $(x^1, \cdots, x^n)$ gives a natural basis for the tangent spaces via the partial derivatives $(\partial_1, \cdots, \partial_n)$. Since the coordinate basis is also holonomic, it provides a convenient basis in which to work.

In many cases, the coordinate basis is orthogonal with respect to a given metric, g. This means the metric is diagonal when expressed in this basis, $g = \mathrm{diag}(g_1, \cdots, g_n)$. If one wishes to normalize the basis to produce an orthonormal basis, we must divide by the scale factors, $\sqrt{g_i}$,

$$(\partial_1, \cdots, \partial_n) \mapsto (\hat{e}_1, \cdots, \hat{e}_n) = \left(\frac{1}{\sqrt{g_1}} \partial_1, \cdots, \frac{1}{\sqrt{g_n}} \partial_n \right). \tag{2.81}$$

However, after normalization the basis may no longer be holonomic. The Lie bracket is

$$[\hat{e}_i, \hat{e}_j] = \frac{1}{2} \frac{1}{\sqrt{g_i g_j}} \left(\frac{\partial_j g_i}{g_i} \partial_i - \frac{\partial_i g_j}{g_j} \partial_j \right), \tag{2.82}$$

and so the normalized basis is only holonomic if $\partial_j g_i = 0$ for $i \neq j$, i.e. g_i is only a function of the corresponding coordinate x^i.

Product Manifold

If $\mathcal{M}$ and $\mathcal{N}$ are topological manifolds, the Cartesian product $\mathcal{M} \times \mathcal{N}$ is a topological manifold under the product topology. The dimension of the product is the sum of the dimensions of the individual manifolds, $\dim \mathcal{M} \times \mathcal{N} = \dim \mathcal{M} + \dim \mathcal{N}$. Given a chart (ϕ, U) for $\mathcal{M}$ and (ψ, V) for $\mathcal{N}$, the map $U \times V \to \mathbb{R}^{m+n} : (x, y) \mapsto (\phi(x), \psi(y))$ is a chart for the product manifold. If the manifolds are C^r, the product is also C^r.

2.2.3 Connection

A linear connection on a real vector bundle E is a map $\nabla : \Gamma(E) \to \Gamma(E \otimes \mathrm{T}^*\mathcal{M})$ that satisfies $\nabla(fs) = s \otimes \mathrm{d}f + f\nabla s$, where s is a section and f is a scalar field. Given a vector field $X \in \Gamma(\mathrm{T}\mathcal{M})$, the connection defines a *covariant derivative*, $\nabla_X : \Gamma(E) \to \Gamma(E)$, via $\nabla_X s = (\nabla s)(X)$, giving

$$\nabla_X(fs) = X(f)s + f\nabla_X s. \tag{2.83}$$

The covariant derivative is linear in X and additive in s,

$$\begin{aligned} \nabla_X(s_1 + s_2) &= \nabla_X s_1 + \nabla_X s_2 \\ \nabla_{X_1+X_2} s &= \nabla_{X_1} s + \nabla_{X_2} s \\ \nabla_{fX} s &= f\nabla_X s, \end{aligned} \tag{2.84}$$

and provides a notion of a directional derivative over the manifold.

Given local bases (e_i) for $\mathrm{T}_x\mathcal{M}$ and (f_j) for E_x, we can consider

$$\nabla_{e_j} f_k = \nabla_j f_k = \omega_j{}^i{}_k f_i, \tag{2.85}$$

where we have defined the *connection coefficients* of ∇ in the given bases as the indexed set of scalar fields, $\omega_j{}^i{}_k$. Applying this to an arbitrary section, s, gives

$$\begin{aligned} \nabla_X s &= X^j(\nabla_j s)^i f_i \\ &= X^j(e_j(s^i) + \omega_j{}^i{}_k s^k) f_i. \end{aligned} \tag{2.86}$$

The connection on E can be used to define a corresponding connection on E^* by requiring the product rule

$$X(\alpha(s)) = (\nabla_X \alpha)(s) + \alpha(\nabla_X s), \tag{2.87}$$

where $X \in \Gamma(\mathrm{T}\mathcal{M})$, $s \in \Gamma(E)$, and $\alpha \in \Gamma(E^*)$, which gives

$$\nabla_j f^i = -\omega_j{}^i{}_k f^k \tag{2.88}$$

and

$$\begin{aligned} \nabla_X \alpha &= X^j(\nabla_j \alpha)_k f^k \\ &= X^j(e_j(\alpha_k) - \omega_j{}^i{}_k \alpha_i) f^k. \end{aligned} \tag{2.89}$$

For the simplest case of the trivial bundle, $\mathcal{M} \times \mathbb{R}$, the covariant derivative is just the usual directional derivative of scalar fields,

$$\nabla_X f = X(f). \tag{2.90}$$

The above can be used to extend the connection to arbitrary tensor product bundles by the product rule

$$\nabla_X(T \otimes S) = \nabla_X T \otimes S + T \otimes \nabla_X S. \tag{2.91}$$

This means for a type (p, q) tensor field,

$$\begin{aligned}(\nabla_i A)^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q} &= e_i(A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}) \\ &\quad + \omega_i{}^{\mu_1}{}_j A^{j\mu_2\cdots\mu_p}_{\nu_1\cdots\nu_q} + \cdots + \omega_i{}^{\mu_p}{}_j A^{\mu_1\cdots\mu_{p-1}j}_{\nu_1\cdots\nu_q} \\ &\quad - \omega_i{}^{j}{}_{\nu_1} A^{\mu_1\cdots\mu_p}_{j\nu_2\cdots\nu_q} - \cdots - \omega_i{}^{j}{}_{\nu_q} A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_{q-1}j}.\end{aligned} \tag{2.92}$$

For connections on the tangent bundle, the connection coefficients are often written using the notation $\Gamma^i{}_{jk} = \omega_j{}^i{}_k$.

Torsion

The *torsion* of a connection on the tangent bundle is a map $T : \Gamma(\mathrm{T}\mathcal{M}) \times \Gamma(\mathrm{T}\mathcal{M}) \to \Gamma(\mathrm{T}\mathcal{M})$ given by

$$T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y]. \tag{2.93}$$

This is bilinear and antisymmetric, $T(X, Y) = -T(Y, X)$, which means T is an antisymmetric type-$(1, 2)$ tensor field, with components

$$T^i{}_{jk} = 2\Gamma^i{}_{[jk]} - C^i{}_{jk}. \tag{2.94}$$

This means that for a torsionless connection, the Lie derivative can be expressed simply in terms of the covariant derivative,

$$\mathcal{L}_X Y = \nabla_X Y - \nabla_Y X. \tag{2.95}$$

Levi-Civita Connection

The Levi-Civita connection is a connection on the tangent bundle that is

- compatible with the metric,

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \tag{2.96}$$

- torsion-free, $T(X, Y) = 0$.

For a given Riemannian metric, the Levi-Civita connection is unique and its connection coefficients are expressed in terms of the metric as

$$\Gamma^i{}_{jk} = \frac{1}{2} g^{ia}(e_j(g_{ak}) + e_k(g_{ja}) - e_a(g_{jk}) + C_{ajk} + C_{kja} - C_{jak}) \tag{2.97}$$

where $C_{ajk} = g_{ai} C^i{}_{jk}$. In a coordinate basis, these are called the *Christoffel symbols*,

$$\Gamma^i{}_{jk} = \frac{1}{2} g^{ia}(\partial_j g_{ak} + \partial_k g_{ja} - \partial_a g_{jk}). \tag{2.98}$$

The Christoffel symbols generated by a pullback metric can be expressed in terms of the inducing map, ϕ, as

$$\Gamma^i{}_{jk} = (\phi_*^\dagger)^i{}_a \phi^a{}_{,jk}. \tag{2.99}$$

Geodesics

A geodesic is a C^2 path $\gamma : I \to \mathcal{M} : t \mapsto \gamma(t)$ that is locally the shortest path between two points. In order to use the word 'shortest', we first require a distance metric on $\mathcal{M}$. The length of a path on $\mathcal{M}$ is given by

$$L(\gamma) = \int_a^b \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}\, \mathrm{d}t. \tag{2.100}$$

This functional can be minimized via variational calculus. In local coordinates, geodesics are described by the 2nd-order ODE

$$\ddot\gamma^i + \Gamma^i{}_{jk} \dot\gamma^j \dot\gamma^k = 0, \tag{2.101}$$

called the *geodesic equation*, where $\Gamma^i{}_{jk}$ are the Christoffel symbols. Therefore an initial position $\gamma(0)$ and velocity $\dot\gamma(0)$ is sufficient to determine the path via an initial value problem. When the Christoffel symbols vanish, the geodesic equation describes paths with constant velocity, i.e. straight lines, in local coordinates.

If the geodesics extend for all time $t \in \mathbb{R}$, the manifold is called *geodesically complete*. Compact manifolds are geodesically complete, by the Hopf-Rinow theorem.

Parallel Translation

Let $\gamma : I \to \mathcal{M}$ be a path and E be a vector bundle. The notion of a parallel translation is the method of moving a vector in the fibre at $\gamma(a)$ along the path to the fibre at

$\gamma(b)$ in a consistent manner. In other words a parallel translation is an isomorphism between fibres generated by a path, $P(\gamma)_{ab} : E_{\gamma(a)} \to E_{\gamma(b)}$. These isomorphisms are consistent in the sense that the parallel translation along a concatenation of two paths coincides with the composition of the individual parallel translations, and that the parallel translation along the trivial path $\gamma(t) = p$ is the identity, $P_{ab}(\gamma) = \mathrm{id}_{E_p}$.

A section V of the bundle defined on the path, $\gamma(I)$, is called parallel if $\nabla_{\dot{\gamma}} V = 0$, where ∇ is the connection on E. This is given locally by

$$\dot{V}^i + \omega_j{}^i{}_k \, \dot{\gamma}^j \, V^k = 0. \tag{2.102}$$

Note that a geodesic can be described as a parallel translation of the velocity vector along itself.

Exponential Map

The exponential map is a (partial) map from each tangent space to the base manifold, $\exp_p : \mathrm{T}_p\mathcal{M} \to \mathcal{M}$, that depends on the choice of connection. The map is defined in terms of geodesics as $\exp_p(X) = \gamma(1)$, where γ is the geodesic with $\gamma(0) = p$ and $\dot{\gamma}(0) = X$. If the manifold is geodesically complete, the exponential map's domain is the whole tangent space.

2.2.4 Curvature

There are two notions of curvature for a manifold: intrinsic and extrinsic. A manifold's intrinsic curvature is independent of its embedding into an ambient space, whereas the extrinsic curvature is dependent on the embedding.

Intrinsic Curvature

The existence of intrinsic curvature is the topic of Gauss' *Theorema Egregium* (remarkable theorem), which shows that the Gaussian curvature (a particular quantification of curvature for surfaces in $\mathbb{R}^3$) is independent of the embedding, even though its construction explicitly depends on it. Therefore one can talk about the (intrinsic) curvature of a manifold without needing an ambient space to refer to.

In modern differential geometry, the intrinsic curvature of a (semi-)Riemannian manifold is a function of the metric, via the Levi-Civita connection. The Riemann

tensor $R \in \Gamma(T^1_3(\mathrm{T}\mathcal{M}))$ is given by

$$R(X,Y)Z = \left(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\right) Z. \tag{2.103}$$

It measures the extent to which the covariant derivative is non-commutative. Geometrically, this quantifies how Z changes when it undergoes parallel translation around an infinitesimal closed loop defined by X and Y. In a coordinate basis, the partial derivatives of the three vector fields cancel out, and so the Riemann tensor can be expressed in terms of the Christoffel symbols and their derivatives,

$$R^i{}_{jkl} = \partial_k \Gamma^i{}_{jl} - \partial_l \Gamma^i{}_{jk} + \Gamma^i{}_{ka}\Gamma^a{}_{jl} - \Gamma^i{}_{la}\Gamma^a{}_{jk}, \tag{2.104}$$

where the indices are given in the order $R(\partial_k, \partial_l)\partial_j = R^i{}_{jkl}\partial_i$.

From the Riemann tensor one can define the Ricci tensor, Ric, as a $(0,2)$ tensor field defined as

$$\mathrm{Ric}(Z,Y) = \mathrm{tr}(X \mapsto R(X,Y)Z). \tag{2.105}$$

In coordinates this is just the contraction of the first and third indices of the Riemann tensor,

$$\mathrm{Ric}(\partial_i, \partial_j) = R_{ij} = R^k{}_{ikj}. \tag{2.106}$$

The Ricci scalar is given by the trace of the $(1,1)$ Ricci tensor,

$$R = R^j{}_j = g^{ji} R_{ij}. \tag{2.107}$$

Extrinsic Curvature

Given a Riemannian manifold $(\mathcal{M}, g)$ and an embedding $\varphi : \mathcal{N} \to \mathcal{M}$, the extrinsic curvature of $\mathcal{N}$ in $\mathcal{M}$ can be thought of as a measure of the difference between the intrinsic curvature of $\mathcal{N}$ (as measured by the pullback metric), and the curvature of $\mathcal{N}$ as measured using g in $\mathcal{M}$. In other words it is the 'additional' curvature that has been introduced to $\mathcal{N}$ by its embedding into $\mathcal{M}$.

Let ∇^g be the Levi-Civita connection for g, and let h be the pullback metric $h = \varphi^* g$, with connection ∇^h. At each point on $\varphi(\mathcal{N})$, let P be the orthogonal projection onto the tangent space $\mathrm{T}\varphi(\mathcal{N})$. Then for vector fields defined on $\varphi(\mathcal{N})$, we can perform the orthogonal decomposition of the covariant derivative,

$$\nabla^g_X Y = P(\nabla^g_X Y) + (\mathrm{id} - P)(\nabla^g_X Y). \tag{2.108}$$

The normal component is defined as the *second fundamental form*,

$$\mathrm{I\!I}(X,Y) = (\mathrm{id} - P)(\nabla^g_X Y), \tag{2.109}$$

which is a $(0,2)$ tensor field on $\varphi(\mathcal{N})$ whose values lie in the normal bundle of $\varphi(\mathcal{N})$ in $\mathcal{M}$. For tangential vector fields X and Y, $\mathrm{I\!I}$ is symmetric,

$$\begin{aligned}\mathrm{I\!I}(Y,X) &= (\mathrm{id} - P)(\nabla^g_Y X)\\ &= (\mathrm{id} - P)(\nabla^g_X Y) - (\mathrm{id} - P)(\mathcal{L}_X Y)\\ &= \mathrm{I\!I}(X,Y),\end{aligned} \tag{2.110}$$

since $\mathcal{L}_X Y$ is tangential.

If X and Y are vector fields on $\mathcal{N}$, then $\varphi_* X$ and $\varphi_* Y$ are tangential vector fields on $\varphi(\mathcal{N})$. In this case, the tangential component of (2.108) coincides with $\nabla^h_X Y$, i.e. $P(\nabla^g_{\varphi_* X}(\varphi_* Y)) = \varphi_* \nabla^h_X Y$, giving

$$\nabla^g_{\varphi_* X}(\varphi_* Y) = \varphi_* \nabla^h_X Y + \mathrm{I\!I}(\varphi_* X, \varphi_* Y). \tag{2.111}$$

This leads to the Gauss equation, which relates the curvatures,

$$\begin{aligned}g(R^g(\varphi_* X, \varphi_* Y)\varphi_* Z, W) &= g(\varphi_*(R^h(X,Y)Z), W)\\ &\quad + g(\mathrm{I\!I}(\varphi_* X, \varphi_* Z), \mathrm{I\!I}(\varphi_* Y, W))\\ &\quad - g(\mathrm{I\!I}(\varphi_* Y, \varphi_* Z), \mathrm{I\!I}(\varphi_* X, W)).\end{aligned} \tag{2.112}$$

This is rather messy, so we can drop some notation by identifying $\mathcal{N}$ with $\varphi(\mathcal{N})$, $\mathrm{T}\mathcal{N}$ with $\mathrm{T}\varphi(\mathcal{N})$, and X with $\varphi_* X$, etc.

$$\begin{aligned}g(R^g(X,Y)Z, W) &= g(R^h(X,Y)Z, W)\\ &\quad + g(\mathrm{I\!I}(X,Z), \mathrm{I\!I}(Y,W))\\ &\quad - g(\mathrm{I\!I}(Y,Z), \mathrm{I\!I}(X,W)).\end{aligned} \tag{2.113}$$

Note that the difference between R^g and R^h is given in terms of $\mathrm{I\!I}$,

$$\begin{aligned}g(R^g(X,Y)Z - R^h(X,Y)Z, W) &= g(\mathrm{I\!I}(X,Z), \mathrm{I\!I}(Y,W))\\ &\quad - g(\mathrm{I\!I}(Y,Z), \mathrm{I\!I}(X,W)).\end{aligned} \tag{2.114}$$

Therefore it is the normal components of the covariant derivatives that are relevant to expressing the extrinsic curvature.

2.2.5 Differential Forms

Differential p-forms are sections of the pth exterior power of the cotangent bundle, $\Gamma(\Lambda^p(\mathrm{T}^*\mathcal{M}))$, which we denote as simply $\Omega^p(\mathcal{M})$. The operations described in §2.1.8, such as the exterior product, can be extended to differential forms by fibrewise application. We shall refer to a differential p-forms as just p-forms from here on. A p-form provides an alternating multilinear map, which takes p vector fields, and gives a scalar field as a value.

Exterior Derivative

The exterior derivative is a map from p-forms to $(p+1)$-forms, $\mathrm{d} : \Omega^p(\mathcal{M}) \to \Omega^{p+1}(\mathcal{M})$. For a 0-form, f (a scalar field), the exterior derivative is defined as the 1-form that gives the directional derivative, $\mathrm{d}f(X) = X(f)$. Since the coordinates on the manifold are a set of (locally defined) scalar fields, $(x^1, \cdots, x^n)$, we can consider the exterior derivatives $(\mathrm{d}x^1, \cdots, \mathrm{d}x^n)$, which turns out to be the dual basis for the cotangent space corresponding to the partial derivatives, i.e. $\mathrm{d}x^i(\partial_j) = \partial_j(x^i) = \delta^i_j$.

The exterior derivative of a p-form, α, is a $(p+1)$-form given by

$$\mathrm{d}\alpha = \frac{1}{p!}\, \partial_i \alpha_{j_1 \cdots j_p}\, \mathrm{d}x^i \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_p}. \tag{2.115}$$

Anti-symmetrizing this description gives

$$\mathrm{d}\alpha_{i j_1 \cdots j_p} = (p+1)\, \partial_{[i} \alpha_{j_1 \cdots j_p]}. \tag{2.116}$$

This has the following properties, where α is a p-form and β is a q-form,

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta, \tag{2.117}$$

$$\mathrm{d}(\alpha + \beta) = \mathrm{d}\alpha + \mathrm{d}\beta, \tag{2.118}$$

$$\mathrm{d}^2\alpha = 0. \tag{2.119}$$

A p-form, α, is *exact* if there exists a $(p-1)$-form, β, such that $\alpha = \mathrm{d}\beta$. A p-form, α, is *closed* if $\mathrm{d}\alpha = 0$. Note that exact implies closed. Given $\phi : \mathcal{M} \to \mathcal{N}$ with a p-form α on $\mathcal{N}$, the pullback form, denoted $\phi^*\alpha$, is a p-form on $\mathcal{M}$ defined by

$$(\phi^*\alpha)(X_1, \cdots, X_p) = \alpha(\phi_* X_1, \cdots, \phi_* X_p), \tag{2.120}$$

which gives

$$(\phi^*\alpha)_{j_1\cdots j_p} = \alpha_{i_1\cdots i_p}\partial_{j_1}\phi^{i_1}\cdots\partial_{j_p}\phi^{i_p}. \tag{2.121}$$

The pullback satisfies

$$\phi^*(\alpha\wedge\beta) = \phi^*\alpha\wedge\phi^*\beta, \tag{2.122}$$

$$\phi^*(\alpha+\beta) = \phi^*\alpha+\phi^*\beta, \tag{2.123}$$

$$\phi^*\mathrm{d}\alpha = \mathrm{d}(\phi^*\alpha). \tag{2.124}$$

Lie Derivative

As with tensor fields, we can apply the Lie derivative to differential forms. For a given vector field, X, the Lie derivative maps p-forms to a p-forms, $\mathcal{L}_X : \Omega^p(\mathcal{M}) \to \Omega^p(\mathcal{M})$. It satisfies the product rule under the wedge product,

$$\mathcal{L}_X(\alpha\wedge\beta) = \mathcal{L}_X\alpha\wedge\beta + \alpha\wedge\mathcal{L}_X\beta, \tag{2.125}$$

and commutes with the exterior derivative,

$$\mathrm{d}(\mathcal{L}_X\alpha) = \mathcal{L}_X(\mathrm{d}\alpha). \tag{2.126}$$

Orientation and the Volume Form

A vector bundle is called *orientable* if its transition maps have positive determinant. This allows for a choice of orientation on each fibre that is consistent over the bundle. A differentiable manifold is said to be orientable if its tangent bundle is an orientable vector bundle.

On an orientable manifold we can specify a choice of orientation using a *volume form*, which is an n-form that provides a fibrewise orientation form for the tangent bundle. For the coordinate basis, this is

$$\nu = \sqrt{\det g}\,\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^n \tag{2.127}$$

$$\nu_{i_1\cdots i_n} = \sqrt{\det g}\,\varepsilon_{i_1\cdots i_n}. \tag{2.128}$$

Gradient

The *gradient* of a scalar field is a vector field defined by $\operatorname{grad} f = (\mathrm{d}f)^\sharp$ and given in coordinates by

$$(\operatorname{grad} f)^i = g^{ij}\partial_j f = g^{ij}f_{,j}. \tag{2.129}$$

It follows that $g(\operatorname{grad} f, X) = X(f) = \mathrm{d}f(X)$.

Curl

The *curl* of a p-form, α, is a $(n-p-1)$-form defined by

$$\operatorname{curl} \alpha = *(\mathrm{d}\alpha). \tag{2.130}$$

Curl in 3D On an oriented 3-manifold, the curl of a 1-form is also a 1-form. Hence we can define the curl of a vector field using the musical isomorphisms,

$$\operatorname{curl} X = (\operatorname{curl}(X^\flat))^\sharp. \tag{2.131}$$

For a coordinate basis and a diagonal metric, $(g_{ij}) = \operatorname{diag}(g_1, g_2, g_3)$, this gives

$$\begin{aligned}
(\operatorname{curl} X)^1 &= \frac{1}{\sqrt{\det g}} \left(\partial_2(g_3 X^3) - \partial_3(g_2 X^2)\right) \\
(\operatorname{curl} X)^2 &= \frac{1}{\sqrt{\det g}} \left(\partial_3(g_1 X^1) - \partial_1(g_3 X^3)\right) \\
(\operatorname{curl} X)^3 &= \frac{1}{\sqrt{\det g}} \left(\partial_1(g_2 X^2) - \partial_2(g_1 X^1)\right)
\end{aligned} \tag{2.132}$$

If we wish to normalize the basis, $(\partial_1, \partial_2, \partial_3) \mapsto (\hat{e}_1, \hat{e}_2, \hat{e}_3) = \left(\frac{\partial_1}{\sqrt{g_1}}, \frac{\partial_2}{\sqrt{g_2}}, \frac{\partial_3}{\sqrt{g_3}}\right)$, this becomes

$$\begin{aligned}
(\operatorname{curl} X)^1 &= \frac{1}{\sqrt{g_2 g_3}} \left(\partial_2(\sqrt{g_3} X^3) - \partial_3(\sqrt{g_2} X^2)\right) \\
(\operatorname{curl} X)^2 &= \frac{1}{\sqrt{g_1 g_3}} \left(\partial_3(\sqrt{g_1} X^1) - \partial_1(\sqrt{g_3} X^3)\right) \\
(\operatorname{curl} X)^3 &= \frac{1}{\sqrt{g_1 g_2}} \left(\partial_1(\sqrt{g_2} X^2) - \partial_2(\sqrt{g_1} X^1)\right).
\end{aligned} \tag{2.133}$$

For example, in cylindrical coordinates (ρ, ϕ, z), with metric $g = \operatorname{diag}(1, \rho^2, 1)$, the normalization of the basis is $(\partial_\rho, \partial_\phi, \partial_z) \mapsto (\partial_\rho, \frac{1}{\rho}\partial_\phi, \partial_z)$, and in this basis the curl is

$$\begin{aligned}
(\operatorname{curl} X)^\rho &= \frac{1}{\rho}\partial_\phi X^z - \partial_z X^\phi \\
(\operatorname{curl} X)^\phi &= \partial_z X^\rho - \partial_\rho X^z \\
(\operatorname{curl} X)^z &= \frac{1}{\rho}\left(\partial_\rho(\rho X^\phi) - \partial_\phi X^\rho\right).
\end{aligned} \tag{2.134}$$

2.3 Dynamical Systems

Dynamical systems concerns the abstract study of quantities that change over time according to a fixed rule of time evolution. There are two main notions of time, referred to as continuous time and discrete time, quantified by the reals and integers respectively, which correspond to the two main classes of time evolution: flows and iterated maps. Stochastic (containing random/probabilistic contributions) and hybrid (including both continuous and discrete parts) dynamics are also studied. As a result, dynamical systems theory involves the use of many different areas of mathematics, including group theory, topology, analysis, differential geometry, linear algebra, and measure theory. Numerical and computational methods are also important tools. [23, 21, 44, 40, 46]

A *dynamical system* is a collection $(X, \mathcal{T}, \psi)$ of three objects: a *state space*, X, a *time* monoid, $(\mathcal{T}, +)$, and an *evolution* function, $\psi : \mathcal{T} \times X \to X$. The state space, X, contains all possible configurations the system can take and is sometimes called a 'phase space' or a 'configuration space'. The monoid $\mathcal{T}$ is usually $\mathbb{R}$ or $\mathbb{Z}$ with addition, and is used to quantify the notion of time and serves as a parameterization of the evolution of the system. The evolution function ψ evolves the system forwards in time from its current state, by a specified time interval, causing a change in state. The evolution function is also written as $\psi_t(x) := \psi(t, x)$; it satisfies $\psi_{t+s} = \psi_t \circ \psi_s = \psi_s \circ \psi_t$ and $\psi_0 = \mathrm{id}_X$.

2.3.1 Discrete Time

In discrete-time dynamics, the time evolution is produced by repeated application of a map $f : X \to X$, and thus the evolution function is simply

$$\psi_n(x) = f^n(x) = \underbrace{f \circ \cdots \circ f}_{n \text{ times}}(x). \tag{2.135}$$

Time is therefore quantified by either non-negative integers, $\mathcal{T} = \mathbb{N}_0$, or, if f is invertible, the integers, $\mathcal{T} = \mathbb{Z}$. The orbit of an iterated map therefore produces a sequence of states, $\{x_0, x_1, x_2, \cdots\}$, corresponding to jumps in the state space.

2.3.2 Continuous Time

In continuous-time dynamics, time is quantified by the reals, $\mathcal{T} = \mathbb{R}$, and the time evolutions are given by a one-parameter group $\mathbb{R} \to \Psi : t \mapsto \psi_t$. The group Ψ therefore has identity ψ_0, inverse $\psi_t^{-1} = \psi_{-t}$, and a right-action on the state space, X, known as a *flow*. The orbit of a flow is a path through the state space.

2.3.3 Systems of ODEs

The flow is usually described in terms of a system of ODEs. When written in first-order form, these describe a vector field V on X in local coordinates, $V_{x(t)} = \dot{x}(t)$, where X is a smooth manifold. If this vector field is complete, then the integral curves of the vector field exist for all time and generate a one-parameter group of diffeomorphisms, $\psi_t = \Phi_V^t \in \mathrm{Diff}(X)$.

2.3.4 PDEs and Infinite-Dimensional State Spaces

Partial differential equations (PDEs) typically describe the evolution of scalar or vector fields on some physical space, $\phi(x, t)$,

$$F(x, \phi, \partial_x\phi, \partial_t\phi, \cdots) = 0 \tag{2.136}$$

The state of such a system is therefore an infinite-dimensional function space, usually quantified by a real or complex Hilbert space, $\mathcal{H}$,

$$\dot{\phi} = f(\phi), \quad \phi \in \mathcal{H}. \tag{2.137}$$

If the physical space is discretized, for example by use of a finite element scheme, the state space is approximated by a finite-dimensional space, with dynamics described by a system of ODEs,

$$\dot{\phi}_i = f_i(\phi_1, \cdots, \phi_N), \tag{2.138}$$

where i indexes the points used in the discretization of the physical space, and each ϕ_i is the field value at the corresponding point.

2.3.5 Invariant Structures in State Space

Given an *initial condition* $x_0 \in X$, the *orbit* of the system is the set of states, $\{\psi_t(x_0) : t \in \mathcal{T}\}$, that are reached from x_0 via ψ.

An *invariant set*, $\mathcal{S}$, is a subset of X such that all states in the subset remain in the subset under evolution, i.e. $\psi_t(\mathcal{S}) \subseteq \mathcal{S}\ \forall t$. The simplest example of an invariant set is a *fixed point*, or *equilibrium*, which is a single point $x^* \in X$, such that $\psi_t(x^*) = x^*$, i.e. if the system is in a fixed point configuration, it remains there indefinitely.

A *periodic orbit* is a subset $L \subset X$, such that $\psi_{t+\tau}(x) = \psi_t(x)\ \forall x \in L\ \forall t$, where τ is called the *period* of the orbit. Note that $\psi_{t+n\tau}(x) = \psi_t(x)\ \forall n \in \mathbb{N}$, so we must define the period τ to be the smallest such value in order to make it unique. In the case of an iterated map, a periodic orbit is a finite sequence of points that repeats indefinitely, while in the case of a flow, it is a closed loop in the state space.

More generally, one has invariant manifolds. A submanifold $\mathcal{M} \subset X$ is invariant under a diffeomorphism, $f : X \to X$, if $f(\mathcal{M}) = \mathcal{M}$. If the diffeomorphism is the flow of a vector field V on X, i.e. $f(x) = \Phi_V^t(x)$, then $\mathcal{M}$ is invariant under all time-t maps if V is tangent to $\mathcal{M}$. This means the orbits of the flow on $\mathcal{M}$ remain on $\mathcal{M}$ forever.

An *attractor*, A, is an invariant set that attracts nearby states from its *basin of attraction*, $B_A \subset X$, and that contains no proper subset with these properties.

A *limit cycle* is a periodic orbit of a flow that is isolated. This means that there are no neighbouring loops.

2.3.6 Stability of Invariant Sets

Stability refers to the properties of the time evolution near to an invariant set. This depends on the behaviour of the time evolution in a neighbourhood of the set.

An invariant set $\mathcal{S}$ is *Liapunov stable* if points nearby $\mathcal{S}$ stay nearby under (forwards) time evolution. For any neighbourhood U of $\mathcal{S}$, there is a neighbourhood V of $\mathcal{S}$ such that $\psi_t(V) \subseteq U$ for all $t > 0$.

An invariant set $\mathcal{S}$ is *asymptotically stable* if points nearby $\mathcal{S}$ tend towards $\mathcal{S}$ eventually. There is a neighbourhood U of $\mathcal{S}$ such that $\lim_{t\to\infty} \psi_t(x) \in \mathcal{S}\ \forall x \in U$. If $\mathcal{S}$ is both Liapunov stable and asymptotically stable it is called *stable*.

Linear Systems

A linear dynamical system has state space $\mathbb{R}^n$ and time evolution given by a linear differential equation

$$\dot{x} = Ax, \tag{2.139}$$

where $A \in \mathbb{R}^{n\times n}$. Although most systems of interest are nonlinear, linear systems are useful tools for studying the local behaviour of nonlinear systems around a fixed point.

The solution to a linear ODE is given by the exponential map, $\exp : \mathbb{R}^{n\times n} \to \mathrm{GL}(n,\mathbb{R})$,

$$x(t) = \exp(At)\, x(0), \tag{2.140}$$

where $\mathbb{R}^{n\times n}$ can be regarded as the tangent space to the Lie group $\mathrm{GL}(n,\mathbb{R})$ at the identity. When expressed in an eigenbasis of A, the complexified problem becomes diagonal,

$$\dot{x}^j = \lambda_j x^j, \tag{2.141}$$

and so does the solution,

$$x^j(t) = e^{\lambda_j t}\; x^j(0). \tag{2.142}$$

By writing $\lambda = \alpha + i\beta$,

$$e^{\lambda t} = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)) \tag{2.143}$$

shows the real part of λ describes an exponential growth for $\mathrm{Re}(\lambda) > 0$ and decay for $\mathrm{Re}(\lambda) < 0$, while the imaginary part describes an oscillation.

Due to its linearity, every point in the kernel of A is a fixed point. Therefore the origin is always a fixed point in a linear system. Given a nonlinear system on $\mathbb{R}^n$, $\dot{x} = f(x)$, with fixed point $x^* \in X$, one can construct the linearization around x^*,

$$\dot{\xi} = J_{x^*}\xi, \tag{2.144}$$

where J is the Jacobian of f (the matrix of partial derivatives) evaluated at x^*, and $\xi = x - x^*$ is the relative position.

Monodromy Matrix

Given a trajectory $x(t)$, the *monodromy matrix* M is a time-dependent matrix that describes how perturbations to the orbit change over time,

$$\xi(t) = M(t)\xi_0. \tag{2.145}$$

It is given in terms of the Jacobian by the *variational equation*,

$$\dot{M}(t) = J_{x(t)}M(t). \tag{2.146}$$

For a periodic orbit of period τ, $M(\tau)$ describes how perturbations change due to one period's worth of time evolution. The eigenvalues of $M(\tau)$ therefore describe the stability of the periodic orbit and are called the *Floquet multipliers*. Every periodic orbit has a unit Floquet multiplier, which describes motion along the orbit (the tangential component). The remaining $n-1$ multipliers affect stability (the normal components) – if they all have modulus less than one then all perturbations contract, corresponding to a stable periodic orbit.

Normally Hyperbolic Invariant Manifolds

Normally hyperbolic invariant manifolds (NHIMs) [13, 23] are a class of invariant manifolds that have particular stability properties. A NHIM, $\mathcal{M} \subset X$, is a compact manifold, invariant under a diffeomorphism $f : X \to X$ such that along $\mathcal{M}$ the tangent bundle admits a splitting into tangential, stable and unstable subbundles,

$$\mathrm{T}_{\mathcal{M}}X = \mathrm{T}\mathcal{M} \oplus E^s \oplus E^u, \tag{2.147}$$

i.e. there are no normal directions that are neither stable nor unstable. Both the stable bundle E^s and the unstable bundle E^u must be f_*-invariant, along with $\mathrm{T}\mathcal{M}$, which is automatically f_*-invariant due to $\mathcal{M}$ being f-invariant.

Compact NHIMs are particularly important examples of invariant manifolds as they persist under small differentiable perturbations of the vector field. The simplest example of a NHIM is a hyperbolic fixed-point, $\mathcal{M} = \{x^*\}$, $\mathrm{T}_{x^*}\mathcal{M} = \{0\}$. If x^* is a hyperbolic fixed point, there is a neighbourhood of x^* in which the full dynamics is topologically equivalent to the linearization around x^*, by the Hartman-Grobman theorem.

2.3.7 Parameters and Bifurcations

Typically, an evolution function, ψ, is partly specified by a number of parameters. In the continuous-time case, this gives a family of ODEs, $\dot{x} = f(x; \lambda)$, describing a family of vector fields. The parameters are held constant during the evolution of the system, however the time evolution may take on different parameter values in different instances of the system. There may be a region of parameter space of relevance, for example in which physically plausible values lie. Therefore it is necessary to study how

the features of the dynamics, such as the invariant sets and their stability, respond to changes in the parameter to understand the full range of behaviours possessed by the system.

Structural stability

Structural stability is a notion of stability that applies to perturbations of the system itself (e.g. via a parameter). Although two time evolutions may be different, possessing different orbits, we are often primarily interested in qualitative features such as the number, and type of, invariant sets, rather than the particular geometric shape of the orbits.

Two time evolutions on a state space X, ψ, ϕ, are said to be *topologically equivalent* if there is a homeomorphism $h : X \to X$ that maps the orbits of one onto the orbits of the other. Furthermore, the two systems are said to be *topologically conjugate* if $h\phi = \psi h$. This is stronger than the topological equivalence above, as not only does it map orbits to orbits (as sets), but it also preserves the way time increases along each orbit.

A system with parameter $\lambda \in U$ is said to be *structurally stable* at λ if there is a neighbourhood of λ in which each system is topologically conjugate to the system at λ.

Bifurcations

When varying the parameters of a model, the structures present in the state space can change in response. Even if the vector field depends smoothly on the parameters, the structures resulting from its flow can change in a wide variety of complicated ways. In the simplest case, structures may move around in state space, or may vary in geometry. The stability properties of structures may also vary – for example, a stable fixed point may become unstable. Structures can also be created and destroyed, or change in dimension, such as a limit cycle becoming a torus.

A parameter value at which the system is structurally unstable is called a *bifurcation* point. These are therefore points at which a (topologically) significant change occurs in the orbits of the system. Bifurcations are classified as being *local* or *global*, depending on the scale and impact of the changes. Local bifurcations have an impact

that can be confined to a small piece of the state space, usually a neighbourhood of a non-hyperbolic fixed point. Global bifurcations, on the other hand, involve the collision of invariant structures that produces topological changes in the orbits of the system throughout the state space.

The *normal form* of a bifurcation is the simplest low-dimensional system, with the fewest parameters necessary, that exhibits the bifurcation. They are given by low-degree polynomial vector fields. This allows for the bifurcation to be studied in a simple example, and allows one to detect such bifurcations in more complicated systems by transforming the system into the normal form.

2.3.8 Symmetry

Let G be a group with actions on sets X and Y. A function $f : X \to Y$ is said to be G-*equivariant* if $f(gx) = gf(x)$ for all $g \in G$. In the smooth case, we deal with a Lie group G with smooth action on a smooth manifold, X. We can use the action on a manifold X to define an action on its tangent bundle $\mathrm{T}X$ given by the differentials g_* (i.e. we are thinking of each g as a smooth map $X \to X$ and g_* as its differential). This is a well-defined action since $(\mathrm{id}_X)_* = \mathrm{id}_{\mathrm{T}X}$ and $(gh)_* = g_*h_*$. If $\phi : X \to Y$ is a smooth G-equivariant map then the differential ϕ_* is a smooth G-equivariant map under the tangent bundle actions.

A dynamical system is said to have a symmetry group G if its time evolutions are G-equivariant, i.e. they commute with the group G. For a flow this is

$$g \circ \Phi_V^t = \Phi_V^t \circ g \quad \forall t \in \mathbb{R},\ \forall g \in G. \tag{2.148}$$

i.e. $\Phi_V^t = g^{-1} \circ \Phi_V^t \circ g$. By taking the time derivative of this we get the corresponding equation for the vector field,

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g\Phi_V^t(x) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_V^t(gx) \\ (g_*)_x V_x &= V_{gx} \\ (g_*V)_{gx} &= V_{gx} \\ (g_*V)_y &= V_y \quad \forall y \in X,\ \forall g \in G, \end{aligned} \tag{2.149}$$

i.e. the vector field is equal to its pushforward, $g_*V = V$. This means the underlying vector field V is G-equivariant as a map $X \to \mathrm{T}X$.

The existence of a symmetry places constraints on the dynamics of the system. For example, if x is fixed by g (i.e. $gx = x$) then g also fixes every point on the orbit of x, $g\Phi_V^t(x) = \Phi_V^t(x)\ \forall t$. Therefore if we denote by $\mathrm{fix}(H)$ the set of points fixed by $H \subseteq G$, then $\mathrm{fix}(H)$ is a subset of X invariant under the flow, and $H \subset J \Rightarrow \mathrm{fix}(H) \supseteq \mathrm{fix}(J)$.

For each point $x \in X$ we can define an *isotropy* subgroup, Σ_x, as the set of elements in G that fix x. This is a subgroup of G that is constant along each orbit of the flow.

As a result, if $\mathcal{M}$ is an attractor in X and $g(\mathcal{M}) \cap \mathcal{M} = \varnothing$ then we can regard $g(\mathcal{M})$ to be another 'copy' of the same attractor elsewhere in the state space – the symmetry maps the orbits of $\mathcal{M}$ onto the orbits of $g(\mathcal{M})$.

2.4 Optimization on Riemannian Manifolds

Optimization deals with finding a point in a space that is 'optimal'. The definition of optimality is usually given in terms of a cost function, which is a real-valued function on the space, where the optimal point minimizes the cost. There are various types of spaces on which one can define optimization problems, but the particular type of optimization relevant to this work is optimization over Riemannian manifolds, with Euclidean space as a special case. [19, 33]

Optimization over a manifold can be viewed from an intrinsic perspective as unconstrained optimization over a curved space. Alternatively, given an embedding of the manifold into an ambient Euclidean space, one can take an extrinsic view, where the manifold is a constraint with respect to optimization over the ambient space.

2.4.1 Smooth Cost Function

Let $(\mathcal{M}, g)$ be an n-dimensional Riemannian manifold and $S : \mathcal{M} \to \mathbb{R}$ be a smooth cost function on the manifold that we wish to minimize. Given an initial point $x_0 \in \mathcal{M}$, we seek a sequence $(x_1, \cdots, x_N)$ such that $S(x_{i+1}) < S(x_i) \quad \forall i$. An optimization algorithm is an iterative method that computes x_{k+1} from $(x_1, \cdots, x_k)$. For optimization over a manifold, we will consider algorithms that follow these three steps.

1. Determine a suitable tangent vector as a search direction, $\Lambda_k \in \mathrm{T}_{x_k}\mathcal{M}$. This tangent determines a geodesic path, $\gamma_k(t)$, with $\gamma_k(0) = x_k$ and $\dot{\gamma}(0) = \Lambda_k$.

2. Perform a *line search* of the function $S \circ \gamma_k : \mathbb{R} \to \mathbb{R}$ to determine how 'far' to move in the search direction, i.e. a real number $\alpha_k > 0$ that approximately minimizes $S \circ \gamma_k$.

3. Move along the geodesic in the search direction by the amount determined in the line search, i.e. $x_{k+1} = \gamma_k(\alpha_k)$.

Such algorithms aim to find a local minimum of S near to x_0.

We will use the notation τV_k to denote the parallel translation of a vector V_k from x_k to x_{k+1} along the geodesic γ_k,

$$\tau V_k := P(\gamma_k)_{0\alpha_k}(V_k). \tag{2.150}$$

2.4.2 Line Search

Given a search direction, Λ, we need to find a suitable $\alpha > 0$ such that $S(\gamma(\alpha)) < S(\gamma(0))$. This is determined by minimizing the smooth real function $\phi = S \circ \gamma$ in a process called a *line search*. This is itself an optimization problem, albeit a simpler one than the main problem as it is 1-dimensional. However, as ϕ is nonlinear, the problem may be difficult to solve exactly. Despite this, an approximate solution can be sufficient so long as it results in a cost decrease, since this constitutes only a single step in the optimization of S. This is called an *inexact line search*.

Let $\phi(t) = S(\gamma(t))$ and $G = \operatorname{grad} S$. The derivative ϕ' is the directional derivative of S along the path γ,

$$\phi' = \mathrm{d}S(\dot{\gamma}) = g(G, \dot{\gamma}). \tag{2.151}$$

2.4.3 Wolfe Conditions

The *Wolfe conditions* are designed to assist in the inexact line search by providing criteria to evaluate whether a given α is good enough to terminate the line search. These conditions consider the change in the value of the cost function, and the curvature of the cost function.

1. The first condition considers the decrease in the cost function relative to the gradient,

$$\frac{\psi(\alpha) - \psi(0)}{\psi'(0)} \leq \mu. \tag{2.152}$$

2. The second condition considers the ratio of gradients

$$\frac{|\psi'(\alpha)|}{|\psi'(0)|} \leq \eta. \tag{2.153}$$

Here μ and η are constants which are typically chosen to be approximately $\mu = 0.00001$ and $\eta = 0.9$. If α satisfies both of these conditions, the line search is terminated.

In order to perform the inexact line search a number of methods can be employed. We shall outline the two popular ones.

Bounding The Minimum

Methods for minimizing $\phi(t)$ commonly require the turning point first be bounded in an interval $[a, b]$, i.e. $\phi'(a) < 0$ and $\phi'(b) > 0$. Since the search direction, Λ, is chosen to be a descent direction, we can choose $a = 0$.

In order to find an upper bound, we use the value of α from the previous optimization step, denoted α_{old}. If $\psi'(\alpha_{\text{old}}) > 0$ we can use $b = \alpha_{\text{old}}$. If not, α_{old} is repeatedly multiplied by a factor of $\xi > 1$ until $\psi'(\xi^n \alpha_{\text{old}}) > 0$, we then use $b = \xi^n \alpha_{\text{old}}$. We use $\xi = 2$ by default.

Once the bounds have been established, an iterative procedure is used to refine the interval, via one of the following methods.

Bisection Method

The bisection method begins with a bounding interval $[a, b]$. The interval is refined by testing the mid-point, $c = (a + b)/2$. If $\phi'(c)$ is negative, c becomes the new lower bound; if it is positive, c becomes the new upper bound.

False Position

The false position method begins with a bounding interval $[a, b]$. The interval is refined by testing the point

$$c = \frac{a\phi'(b) - b\phi'(a)}{\phi'(b) - \phi'(a)}, \tag{2.154}$$

which is the root of secant from $(a, \phi'(a))$ to $(b, \phi'(b))$. If $\phi'(c)$ is negative, c becomes the new lower bound; if it is positive, c becomes the new upper bound.

2.4.4 Methods for Choosing a Search Direction

Gradient Descent

The gradient descent method uses the search direction $\Lambda_k = -(\operatorname{grad} S)_{x_k}$,

$$(\Lambda_k)^i = -g^{ij}\partial_j S|_{x_k}. \tag{2.155}$$

This is the simplest method and is easy to implement, but it converges slowly and therefore takes a large number of iterations to obtain an accurate result.

Conjugate Gradient

In the conjugate gradient method, the search direction for the kth step is determined by a combination of the gradient at the current position and the $(k-1)$th search direction. Let G_k be the gradient of S at x_k, and Λ_k be the kth search direction. The $(k-1)$th search direction, Λ_{k-1} is first parallely translated from x_{k-1} to x_k along the previous geodesic γ_{k-1}, giving $\tau\Lambda_{k-1}$. This is then combined with the gradient to give the search direction,

$$\Lambda_k = -G_k + \beta_k\,\tau\Lambda_{k-1}, \tag{2.156}$$

using a factor $\beta_k \in \mathbb{R}$, which is determined by one of three popular methods.

- Fletcher-Reeves
$$\beta_k = \frac{g(G_k, G_k)}{g(G_{k-1}, G_{k-1})} \tag{2.157}$$
- Polak-Ribière
$$\beta_k = \frac{g(G_k, G_k - \tau G_{k-1})}{g(G_{k-1}, G_{k-1})} \tag{2.158}$$
- Hestenes-Stiefel
$$\beta_k = \frac{g(G_k, G_k - \tau G_{k-1})}{g(\tau G_{k-1}, G_k - \tau G_{k-1})} \tag{2.159}$$

Newton

The Newton method makes use of the second derivatives of the cost function. The Hessian of S can be defined as a $(1,1)$ tensor field,

$$(\operatorname{Hess} S)(X) = \nabla_X(\operatorname{grad} S). \tag{2.160}$$

In coordinates this is

$$\begin{aligned}(\text{Hess}\, S)^i{}_j &= g^{ik}\nabla_j\partial_k S \\ &= g^{ik}(\partial_j\partial_k S - \Gamma^a{}_{jk}\partial_a S).\end{aligned} \tag{2.161}$$

Note that some authors define the Hessian as a type $(0,2)$ tensor field, but for our purposes it's more convenient to use the corresponding $(1,1)$ tensor field. The search direction Λ is then determined by solving the linear equation

$$(\text{Hess}\, S)(\Lambda) = -\text{grad}\, S. \tag{2.162}$$

2.4.5 Least-Squares Problems

Let (V, g) and (W, h) be real inner product spaces and $f_k : V \to W$ be a number of smooth (possibly nonlinear) functions and $y_k \in W$ be data points. A least-squares problem is a cost function $S : V \to \mathbb{R}$ given by

$$S(x) = \sum_k \|y_k - f_k(x)\|^2, \tag{2.163}$$

where the norm is induced by h. We can express the optimization in terms of the derivatives of f_k, rather than S. This gives the gradient

$$(\text{grad}\, S)^i = -2g^{ic}h_{ab}\sum_k \partial_c(f_k)^a\ (y_k - f_k(x))^b \tag{2.164}$$

$$\text{grad}\, S = -2\sum_k (\nabla f_k)^\dagger (y_k - f_k(x)) \tag{2.165}$$

and the Hessian

$$(\text{Hess}\, S)^i{}_j = -2g^{ic}h_{ab}\sum_k \left[\partial_j\partial_c(f_k)^a\ (y_k - f_k)^b - \partial_c(f_k)^a\ \partial_j(f_k)^b\right] \tag{2.166}$$

$$\begin{aligned}(\text{Hess}\, S)^i{}_j &= -2\sum_k g^{ic}h_{ab}\partial_j\partial_c(f_k)^a\ (y_k - f_k)^b \\ &\quad + 2\sum_k ((\nabla f_k)^\dagger \nabla f_k)^i{}_j\end{aligned} \tag{2.167}$$

2.4.6 Linear Least-Squares Fitting Problems

Let V and W be real inner product spaces, and $(x_k \in V)_{k=1}^N$ and $(y_k \in W)_{k=1}^N$ be data sets. Let $A \in L(V, W)$ be a linear map. By setting $f_k(A) = Ax_k$, the optimization

problem for A is a least-squares

$$S(A) = \sum_{k=1}^{N} \|y_k - Ax_k\|^2 . \tag{2.168}$$

By setting grad $S = 0$, the optimal A is then given by the linear problem,

$$A^i{}_j \left(\frac{1}{N} \sum_{k=1}^{N} (x_k)^j (x_k)^a \right) = \frac{1}{N} \sum_{k=1}^{N} (y_k)^i (x_k)^a. \tag{2.169}$$

By using the block data matrices, this is

$$A(\mathcal{X}\mathcal{X}^{\mathrm{T}}) = \mathcal{Y}\mathcal{X}^{\mathrm{T}}, \tag{2.170}$$

where $\mathcal{X} = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix}$ and $\mathcal{Y} = \begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}$.

2.4.7 Affine Least-Squares Fitting Problems

Let V and W be real inner product spaces, and $(x_k \in V)_{k=1}^N$ and $(y_k \in W)_{k=1}^N$ be data sets. We seek the best affine map $V \to W$ that relates the two data sets. By writing the affine map in the form

$$A(x - \bar{x}) + b, \tag{2.171}$$

where $\bar{x}$ is the mean of the data set,

$$\bar{x} = \frac{1}{N} \sum_{i=k}^{N} x_k, \tag{2.172}$$

the solution to the least-squares problem

$$S = \sum_{k=1}^{N} \|y_k - (A(x_k - \bar{x}) + b)\|^2 \tag{2.173}$$

can be decoupled into two separate parts. The solution for b is given by

$$b = \frac{1}{N} \sum_{k=1}^{N} y_k = \bar{y}, \tag{2.174}$$

i.e. the mean of the values, independent of A. The solution for the linear term is given by

$$A^i{}_j \left(\frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x})^j (x_k - \bar{x})^a \right) = \frac{1}{N} \sum_{k=1}^{N} (y_k)^i (x_k - \bar{x})^a, \tag{2.175}$$

which can be written in block matrices as

$$A(\mathcal{X}\mathcal{X}^{\mathrm{T}}) = \mathcal{Y}\mathcal{X}^{\mathrm{T}}, \tag{2.176}$$

where $\mathcal{X} = \begin{bmatrix} x_1 - \bar{x} & \cdots & x_N - \bar{x} \end{bmatrix}$ and $\mathcal{Y} = \begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}$.

Chapter 3

Secant-Based Projection

In this chapter, we begin by explaining an existing secant-based projection method, the Whitney embedding theorem that motivates it, and its use of optimization over the Grassmann manifold to find a good projection (§3.1–3.7). In §3.8 we detail an affine inverse of the orthogonal projection, which will be important for later chapters. Then from §3.9 onwards, we develop the theory of approximate secants – geometric objects that are produced by allowing a secant's direction to be perturbed. Whereas secants correspond to lines, approximate secants correspond to infinite double cones. The properties of these objects are investigated and used to develop a practical method of secant culling, which can reduce the computational cost of secant-based methods, and can also be used to estimate the dimension of the hyperplane explored by the attractor.

3.1 Introduction

In 2000, Broomhead and Kirby [2] introduced a new approach to the dimensionality reduction of an embedded manifold based on the preservation of its secants. This approach was motivated by a proof of the (easy) Whitney embedding theorem, which makes use of secants in order to show that an m-dimensional manifold can be embedded into $\mathbb{R}^{2m+1}$. The method finds an orthogonal projection onto a subspace of the ambient space that preserves the secants and therefore embeds the manifold into the subspace. This makes it particularly suited for application to attractors of dynamical systems, as not only does it preserve the attractor, but it also provides an explicit

smooth map between the high- and low-dimensional ambient spaces, allowing for the vector field to be pushed forward, which is essential for determining the dynamics in the low-dimensional space. An additional benefit of method is a guarantee that these projections are, in a technical sense, easy to find, provided the dimension of the projection is at least $2m + 1$, which makes practical implementations feasible. Other methods of dimensionality reduction that attempt to preserve a manifold – especially those focusing on data reduction – often do not produce an explicit map between the two spaces, and are therefore not well-suited to the problem of dynamics. The application of this method to attractors was investigated in [4].

3.2 Secants and Projections

Given an inner product space X and a pair of distinct points $x, y \in X$, the *secant* is the unique line in X that intersects x and y. This line can be parameterized as $\gamma_{xy} : \mathbb{R} \to X : t \mapsto x + (y - x)t$. For a subset $Y \subset X$, we can generate a set of secants by each distinct pair of points,

$$\Gamma(Y) = \{\gamma_{xy} \,:\, x \neq y,\ x, y \in Y\}\,. \tag{3.1}$$

A secant-preserving projection is a projection $P : X \to X$ that preserves the secants generated by Y, i.e. $t \mapsto P(\gamma(t))$ is injective for all $\gamma \in \Gamma(Y)$. This is true as long as $P(y - x) \neq 0$ for all $x \neq y \in Y$. Therefore for each pair of points, we only need to check the vector $(y - x)$. As P is linear, we in fact only need to check the direction parallel to $(y - x)$, and hence we can consider the set of directions,

$$\mathcal{K}(Y) = \left\{ \frac{x - y}{\|x - y\|} \,:\, x \neq y,\ x, y \in Y \right\}, \tag{3.2}$$

which we call *unit secants*. If $P(k) \neq 0\ \forall k \in \mathcal{K}(Y)$, then P preserves the secants generated by Y. Note that if Y is a finite set of N points, then $|\mathcal{K}(Y)| = N(N - 1)$.

3.3 Whitney Embedding Theorem

The (easy) *Whitney embedding theorem* is a key result in differential topology [22]. It states that for a compact differentiable manifold, $\mathcal{M}$, of dimension m, there exists an

embedding $\mathcal{M} \to \mathbb{R}^{2m+1}$. This is important for the dimensionality reduction of attractors of dynamical systems, as it means that it is feasible to describe the attractor in an ambient space whose dimension is the same order of magnitude as the attractor itself. Moreover, the proof of the Whitney embedding theorem provides a direct motivation to using a secant-based approach to dimension reduction. As a result, we give a qualitative description of the role of secants in the proof.

To prove the Whitney embedding theorem, one first establishes the existence of an embedding of the manifold $\mathcal{M} \to \mathbb{R}^n$ for some $n > 0$. Then, given such an embedding, one shows that the dimension n can be systematically reduced until one arrives at $n = 2m+1$. It is this second part that makes use of secants and projections. Since the manifold is assumed to be compact, it is sufficient to produce an injective immersion, i.e. to consider the injectivity of both the map and its differential.

The first part takes advantage of the compactness of $\mathcal{M}$ by using a finite subcover of an atlas. By extending each of the chart maps to the whole of $\mathcal{M}$ in a smooth way, one constructs a smooth map of $\mathcal{M}$ into $\mathbb{R}^n$ where $n = mk + k$ and k is the (finite) number of open sets in the subcover. This map is shown to be an injective immersion and hence an embedding.

In order to prove the existence for dimension $2m+1$, one starts with an embedding $j : \mathcal{M} \to \mathbb{R}^n$ with $n > 2m + 1$ and proves the existence of a projection onto a codimension-1 subspace such that the composition with j remains an embedding. Every codimension-1 projection P has a kernel spanned by a unit vector $v \in \mathbb{R}^n$. As discussed in the previous section, as long as this vector is not parallel to any secant, the projection preserves the secants generated by $\mathcal{M}$ and is therefore injective on $\mathcal{M}$. Therefore we are interested in viewing $\mathcal{K}(\mathcal{M})$ as a subset of the unit sphere, S^{n-1}. The proof demonstrates that under the assumed $n > 2m + 1$, the unit secants are nowhere dense in S^{n-1}, and so codimension-1 secant-preserving projections are, in this topological sense, easy to find.

In order for the composition $P \circ j$ to be an embedding, P must also be an immersion on $\mathrm{T}\mathcal{M}$, i.e. $(P_*)_x$ must be injective on each $\mathrm{T}_x\mathcal{M}$. Note that since P is linear, $(P_*)_x = P$. We can identify each tangent vector $v \in \mathrm{T}_x\mathcal{M}$ with a vector in $\mathbb{R}^n$, by $\mathrm{T}\mathcal{M} \to \mathbb{R}^n : (x, v) \mapsto (j_*)_x(v)$. Since j is an immersion, each non-zero tangent vector is non-zero in $\mathbb{R}^n$ and therefore has a corresponding unit vector, which is a point in

S^{n-1}. This gives the map $\tau : \mathrm{T}\mathcal{M} - \mathcal{M} \to S^{n-1} : (x, v) \mapsto (j_*)_x(v)/\left\|(j_*)_x(v)\right\|$, where $\mathcal{M}$ is identified with the subset of $\mathrm{T}\mathcal{M}$ with zero vectors, $(x, 0)$. Since all tangential directions are included in the set of secants, we need only consider the set of secants generated by $\mathcal{M}$. Therefore, by preserving the secants, $P \circ j$ is an injective immersion and hence an embedding.

This argument can then be applied iteratively, producing a sequence of codimension-1 projections that reduce the dimension until $n = 2m+1$. By using orthogonal projections, the projection is uniquely determined by its kernel. Since the kernels in the sequence are mutually orthogonal, the composition of the orthogonal projections is an orthogonal projection, by Proposition A.2.8.

3.4 Orthogonal Projections and Subspaces

Given a real inner product space, (X, g), an orthogonal projection $P : X \to X$ is uniquely determined by the subspace onto which it projects. A subspace can be identified as the image of a linear injection, $W : \hat{X} \to X$. This can be used to pull back the inner product from X to $\hat{X}$, giving $W^\dagger W = \mathrm{id}_{\hat{X}}$, i.e. W is an orthonormal map. The orthogonal projection onto $W(\hat{X})$ is then given by $P = WW^\dagger$. Note that under orthogonal transformations $\mathcal{Q} \in \mathrm{O}(\hat{X}, \hat{g})$,

- $W\mathcal{Q}(\hat{X}) = W(\hat{X})$,
- $(W\mathcal{Q})^\dagger(W\mathcal{Q}) = \mathrm{id}_{\hat{X}} = W^\dagger W$,
- $(W\mathcal{Q})(W\mathcal{Q})^\dagger = WW^\dagger = P$,

i.e. the subspace, orthonormality, and projection are invariant. This sets up an equivalence relation on orthonormal maps,

$$W_1 \sim W_2 \Leftrightarrow \exists \mathcal{Q} \in \mathrm{O}(\hat{X}, \hat{g}) \,:\, W_1 = W_2\mathcal{Q}. \tag{3.3}$$

We can therefore identify each subspace of X with an orthogonal projection P, and with the equivalence class $[W]$ containing orthonormal maps related by orthogonal transformations. These two descriptions are related by $P = WW^\dagger$, which is well-defined.

3.5 The Grassmann Manifold

Given a vector space X of dimension n, the set of d-dimensional subspaces of X forms the *Grassmann manifold*, $\mathrm{Gr}_d(X)$, also known as the *Grassmannian*. The Grassmannian is a compact smooth manifold that provides a topological structure that parameterizes linear subspaces of a given dimension. Since, under a given inner product, each subspace corresponds to an orthogonal projection, we can regard $\mathrm{Gr}_d(X)$ as a configuration space for orthogonal projections of a given dimension.

In the real case, the Grassmannian has dimension $d(n-d)$. The Grassmannian can be given a geometric structure, which permits optimization over the Grassmannian to be performed. Although in differential geometry one usually works with intrinsic descriptions of a manifold, by describing the manifold and its geometry extrinsically, in terms of orthonormal maps, the geodesics and parallel translation can be expressed in closed form, rather than as nonlinear ODEs that must be numerically integrated. This makes optimization over the Grassmannian a feasible method for finding a good orthogonal projection.

3.5.1 Projective Space and the Plücker Embedding

The Grassmannian $\mathrm{Gr}_1(X)$, the set of 1-dimensional subspaces, is called the *projective space*, denoted $\mathbb{P}(X)$. Let $W : \hat{X} \to X$ be an orthonormal map that identifies a d-dimensional subspace, $W(\hat{X})$. Each basis for the subspace is given by a basis for $\hat{X}$, by $w_i = W(e_i)$. We can show that the wedge product of the d-frame $(w_1, \cdots, w_d)$ only depends on the subspace up to scalar multiplication. To see this, consider a pair of frames for $\hat{X}$, which are related by an invertible linear operator $J \in \mathrm{GL}(\hat{X})$, i.e. $\tilde{w}_i = WJ(e_i)$, then

$$\begin{aligned} \tilde{w}_1 \wedge \cdots \wedge \tilde{w}_d &= WJ(e_1) \wedge \cdots \wedge WJ(e_d) \\ &= (\det J)\ W(e_1) \wedge \cdots \wedge W(e_d) \\ &= (\det J)\ w_1 \wedge \cdots \wedge w_d. \end{aligned} \tag{3.4}$$

This means we can identify each subspace, described by the d-frame $(w_1, \cdots, w_d)$, with a line in $\Lambda^d X$, spanned by $w_1 \wedge \cdots \wedge w_d$. This map $\mathrm{Gr}_d(X) \to \mathbb{P}(\Lambda^d X)$ is in fact an embedding of the Grassmannian into the projective space of the exterior power,

called the *Plücker embedding*. Note that $w_1 \wedge \cdots \wedge w_d$ is non-zero since the vectors are linearly independent, and $\det J \neq 0$.

3.6 Cost Function on the Grassmannian

In order to systematically find a good projection, we can define a cost function on the Grassmannian, $\mathcal{F} : \mathrm{Gr}_d(X) \to \mathbb{R}^+$, and use optimization to find a local minimum. In order to be a well-defined function on the Grassmannian, the cost function may need to satisfy certain symmetries depending on how we are parameterizing the manifold.

- Using orthogonal projections $P : X \to X$ of rank d.
- Using orthonormal maps $W : \hat{X} \to X$, with $\dim \hat{X} = d$, and the orthogonal symmetry
$$\mathcal{F}(W) = \mathcal{F}(W\mathcal{Q}) \quad \forall \mathcal{Q} \in \mathrm{O}(\hat{X}, \hat{g}). \tag{3.5}$$
- Using decomposable d-vectors, $w = w_1 \wedge \cdots \wedge w_d \in \Lambda^d(X)$, with scale invariance,
$$\mathcal{F}(w) = \mathcal{F}(\lambda w) \quad \forall \lambda \neq 0 \in \mathbb{R}. \tag{3.6}$$

Since the objective is to preserve the secants, the cost function is constructed to reflect this objective in terms of the lengths of the projected unit secants,

$$\mathcal{F}(W) = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \left\| W^\dagger k \right\|^{-1}. \tag{3.7}$$

Note that in practice we will be using a finite sample of points from the manifold with a corresponding finite set of secants. This cost function is sensitive to individual secants, so that if any secant becomes small under the projection, the contribution to the cost function can become arbitrarily large. This sensitivity ensures that every secant is preserved by the optimization (at least those in the data set). This is a well-defined function on the Grassmannian as it satisfies the orthogonal symmetry $\mathcal{F}(W\mathcal{Q}) = \mathcal{F}(W)$, since orthogonal maps preserve the inner product and norm. It is also invariant under translations of the data, so the data does not need to be centred on the origin.

In terms of the projection P, the cost function can also be written as

$$\mathfrak{F}(P) = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \|Pk\|^{-1}. \tag{3.8}$$

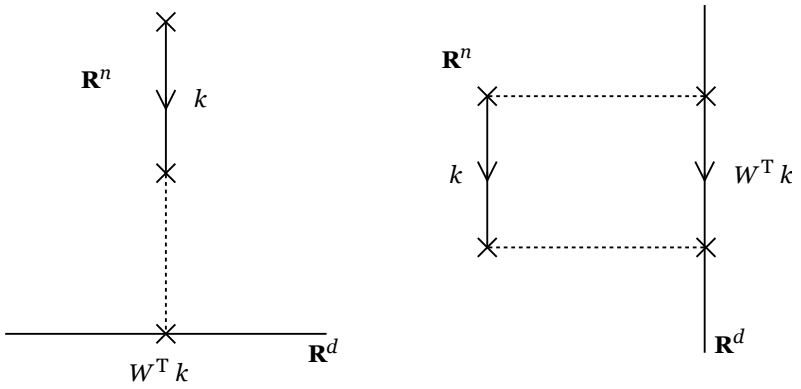


Figure 3.1: Illustration of a good (right) and bad (left) projection of a secant. The bad projection gives a projected length of zero, corresponding to a non-injective map.

Note that $\mathfrak{F}(WW^\dagger) = \mathcal{F}(W)$, since $\left\|WW^\dagger k\right\| = \left\|W^\dagger k\right\|$. The partial derivatives of the cost function are

$$\left(\frac{\partial \mathcal{F}}{\partial W}\right)_i^{\ j} = \frac{\partial \mathcal{F}}{\partial W^i_{\ j}} = -\frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} \frac{k_i (W^\dagger k)^j}{\left\|W^\dagger k\right\|^3}. \tag{3.9}$$

3.7 Geometry and Optimization on the Grassmannian

In order to perform the optimization of the cost function over the Grassmannian, we refer the reader to [12] for the details; a concise summary is given here. By parameterizing the Grassmannian using equivalence classes of orthonormal maps, $[W]$, we can make use of the ambient space of linear maps $L(\hat{X}, X)$ and its algebraic structure. In this representation, the tangent space at $[W]$ corresponds to the set of linear maps $\Delta \in L(\hat{X}, X)$ satisfying $W^\dagger \Delta = 0$. This is well-defined since $(W\mathcal{Q})^\dagger \Delta = \mathcal{Q}^\dagger W^\dagger \Delta = 0$. The dimension of these tangent spaces is therefore given by the dimension of $\hat{X}$ (the maximum rank of Δ) multiplied by the nullity of $W^\dagger$ (the dimension of the kernel), which is $d(n-d)$. Note that $(\mathrm{id} - WW^\dagger)A$ is a tangent vector for any $A \in L(\hat{X}, X)$.

3.7.1 Riemannian Metric

The Grassmannian can be given a Riemannian metric in terms of the ambient space by

$$g_{[W]}(\Delta_1, \Delta_2) = \langle \Delta_1, \Delta_2 \rangle = \mathrm{tr}(\Delta_1 \Delta_2^\dagger), \tag{3.10}$$

which is independent of W and is just the extension of the inner products on the underlying vector spaces to the tensor space $X \otimes \hat{X}^*$, described in §2.1.7 and §2.2.2. Note that under this metric W is orthogonal to the tangent space at $[W]$, i.e. it is in the normal space. In fact, $WW^\dagger A$ is orthogonal to the tangent space for any A. Therefore $\mathrm{id} - WW^\dagger$ is the orthogonal projection onto the tangent space, giving the tangential component $(\mathrm{id} - WW^\dagger)A$, and the normal component $WW^\dagger A$, of an arbitrary map $A \in L(\hat{X}, X)$.

It is convenient to use the matrix description of the linear maps by choosing orthonormal frames for X and $\hat{X}$, under which W becomes an $n \times d$ matrix, the adjoint becomes the transpose, and the inner product is Frobenius.

3.7.2 Gradient and Geodesics

The gradient of the cost function $\mathcal{F}$ (with respect to the Grassmannian) is given by the tangential component of the gradient with respect to the ambient space,

$$\operatorname{grad} \mathcal{F} = (I_n - WW^{\mathrm{T}}) \frac{\partial \mathcal{F}}{\partial W}. \tag{3.11}$$

The geodesic along the Grassmannian from initial point $[W]$ and initial velocity $\Lambda = A\Sigma B^{\mathrm{T}}$ (compact singular value decomposition) is given by

$$W(t) := \begin{bmatrix} WB & A \end{bmatrix} \begin{bmatrix} \cos \Sigma t \\ \sin \Sigma t \end{bmatrix} B^{\mathrm{T}}. \tag{3.12}$$

Note that $\sin(\operatorname{diag}(x_1, \cdots, x_n)) = \operatorname{diag}(\sin x_1, \cdots, \sin x_n)$ and similar for cos. The simplest choice of velocity is the gradient descent, given by $\Lambda = -\operatorname{grad} \mathcal{F}$, although other methods, such as conjugate gradient and Newton, are also possible (2.4.4).

3.7.3 Line Search and Parallel Translation

The step amount, t, is determined by a line search of the function $\psi(t) = \mathcal{F}(W(t))$. Typical line search methods require the derivative $\psi'(t)$, which is given by

$$\psi'(t) = \langle G(t), \Lambda(t) \rangle, \tag{3.13}$$

where $G(t)$ is the gradient of $\mathcal{F}$ at $W(t)$, and $\Lambda(t)$ is the tangent to the geodesic at $W(t)$. Since geodesics can be described as parallel translations of the velocity vector

along itself, $\Lambda(t)$ corresponds to the parallel translation of the initial velocity Λ at W along the geodesic to $W(t)$. The parallel translation of a tangent vector Δ at W along the geodesic with initial velocity $\Lambda = A\Sigma B^{\mathrm{T}}$ is given by

$$\Delta(t) := \left(\begin{bmatrix} WB & A \end{bmatrix} \begin{bmatrix} -\sin \Sigma t \\ \cos \Sigma t \end{bmatrix} A^{\mathrm{T}} + (I_n - AA^{\mathrm{T}}) \right) \Delta. \tag{3.14}$$

3.8 An Affine Inverse

By preserving the secants, the projection embeds the submanifold generating the secants into the subspace. Although the projection is a linear map, in general the inverse of the projection restricted to the submanifold is nonlinear. Here we consider a simple type of nonlinear inverse, an affine map, and consider the cases in which an affine inverse is an accurate description of the true inverse.

A point x_i can be written as $x_i = W\hat{x}_i + z_i$, where $\hat{x}_i := W^{\dagger}x_i$ and $z_i := (\mathrm{id} - WW^{\dagger})x_i$. We shall refer to $W\hat{x}_i$ as the horizontal component and z_i as the vertical component. The projected length of a unit secant, $k_{ij} := (x_i - x_j)/\left\|x_i - x_j\right\|$, corresponds to

$$\left\|W^{\dagger}k_{ij}\right\|^2 = 1 - \frac{\left\|z_i - z_j\right\|^2}{\left\|x_i - x_j\right\|^2}, \tag{3.15}$$

so minimizing $\left\|W^{\dagger}k_{ij}\right\|^{-1}$ corresponds to minimizing $\left\|z_i - z_j\right\|$. This means that, in a sense, the optimization tries to make z uniform over the data set – the minimum possible value the cost function can take is when $z_i = z_j \; \forall i, j$. If the optimal projection is good enough, i.e. the variance of z is low enough, we can approximate the inverse of the projection with the affine map

$$\hat{x} \mapsto W\hat{x} + \bar{z}, \tag{3.16}$$

where $\bar{z}$ is the mean of z_i over the data set,

$$\bar{z} := (\mathrm{id} - WW^{\dagger})\frac{1}{N}\sum_i x_i = (\mathrm{id} - WW^{\dagger})\bar{x}. \tag{3.17}$$

The mean square error in the affine inverse approximation over the data corresponds to the variance of z, which can be bounded from above,

$$\begin{aligned} \frac{1}{N}\sum_i \left\|x_i - (W\hat{x}_i + \bar{z})\right\|^2 &= \frac{1}{N}\sum_i \left\|z_i - \bar{z}\right\|^2 \\ &\leq (1 - \kappa_{\min}^2)\frac{1}{N^2}\sum_{ij} \left\|x_i - x_j\right\|^2, \end{aligned} \tag{3.18}$$

where $\kappa_{\min} := \min_{k \in \mathcal{K}} \left\| W^\dagger k \right\|$ is the minimum projected length. Therefore we can place a requirement on $\kappa_{\min}$ to be close to 1 in order for the inverse to be well approximated by an affine map. If the optimal W does not satisfy this requirement, the dimension d is too small (see [3] for a method of finding a suitable d).

3.8.1 Inverse of Convex Polytopes

We can also consider how this error extends to points not included in the data set by looking at the linear interpolation of data points, or, more generally, convex polytopes formed from the data. Let $r(z) := \|z - \bar{z}\|$ be the residual. Given a set of k vertices, $(v_1, \cdots, v_k) \subset \mathbb{R}^n$, every point in the convex hull $\text{Conv}(v_1, \cdots, v_k)$ has a residual that is bounded from above by the maximum vertex residual,

$$r(p) \leq \max_i r(v_i) \quad \forall p \in \text{Conv}(v_1, \cdots, v_k). \tag{3.19}$$

This can be proven using the norm's triangle inequality together with the fact that convex combinations have non-negative coefficients that sum to unity. Since linear maps preserve convexity, the vertical component of a convex combination of data points is a convex combination of the vertical components of the data points. This means that if the affine map is a good approximation on the data, and points on the attractor not included in the data are well-approximated by convex combinations of the data, then the affine map is a good approximation for the whole attractor.

3.9 Approximate Secants

In order to use the secant-based method in practice, we must use a finite sample of data points from the manifold and the resulting finite set of secants in order to evaluate the cost function (3.7) and perform the optimization over the Grassmannian. Secants that are parallel (or anti-parallel) produce the same contribution to this cost function, since the contribution depends only on the norm of the projected unit secant. Furthermore, because these contributions are continuous with respect to the unit secant, secants that are approximately parallel produce approximately the same contribution to the cost function. We therefore investigate the consequences of *approximate secants*, where we allow each secant to be perturbed by a small amount. Whereas unit secants describe

lines through the origin, approximate secants correspond to infinite double cones. In this section we study the properties of objects and devise two useful applications for dimensionality reduction that we will discuss in §3.10.

Consider $\mathbb{R}^n$ with the Euclidean inner product $\langle\,,\rangle$. We can allow for approximate parallelism by defining a tolerance in terms of the inner product.

Definition 3.9.1. Two unit secants (k_a, k_b) are said to be approximately parallel if

$$|\langle k_a, k_b\rangle|^2 \geq 1 - \varepsilon, \tag{3.20}$$

where $0 \leq \varepsilon \leq 1$ is a given tolerance.

Since $\langle k_a, k_b\rangle = \cos\theta_{ab}$, we can express the tolerance as an angle,

$$\theta := \arccos(\sqrt{1-\varepsilon}). \tag{3.21}$$

The definition above then becomes $\cos^2\theta_{ab} \geq \cos^2\theta$, with angular tolerance $0 \leq \theta \leq \pi/2$. As a binary relation on the set of unit secants, this is reflexive and symmetric, but not transitive.

Given a secant with direction $k \in \mathbb{R}^n$, and an angular tolerance, θ, the set of lines through the origin that are approximately parallel to k defines an infinite double cone with apex fixed to the origin,

$$C_\theta^n(k) = \left\{x \in \mathbb{R}^n \,:\, |\langle x, k\rangle|^2 \geq \|x\|^2\,\|k\|^2\cos^2\theta\right\}. \tag{3.22}$$

Clearly this description is invariant under non-zero scaling, $C_\theta^n(\lambda k) = C_\theta^n(k)$, so we can choose k to be of unit length without loss of generality. From here on we shall use the word 'cone' to refer to an infinite cone with apex fixed to the origin. We shall describe $\operatorname{span}\{k\}$ as the axis of the cone. Each double n-cone contains all lines through the origin that are approximately parallel to its axis.

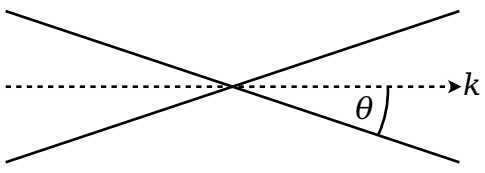


Figure 3.2: Double cone of size θ and direction k in 2-dimensions, $C_\theta^2(k)$.

3.9.1 Symmetry of the Double Cone

Symmetries of the double cone are maps $f : \mathbb{R}^n \to \mathbb{R}^n$ that leave the cone invariant, $f(C_\theta^n(k)) = C_\theta^n(k)$.

Proposition 3.9.2. Let $Q \in \mathrm{O}(n, \mathbb{R})$ be an orthogonal map with an eigenvector k. Then Q is a symmetry of $C_\theta^n(k)$.

Proof. An orthogonal transformation satisfies $Q^\dagger = Q^{-1}$ and has eigenvalues of unit modulus. Therefore,

$$\begin{aligned} |\langle Qx, k\rangle|^2 &= \left|\left\langle x, Q^\dagger k\right\rangle\right|^2 \\ &= \left|\left\langle x, Q^{-1} k\right\rangle\right|^2 \\ &= \left|\left\langle x, \frac{1}{\lambda} k\right\rangle\right|^2 \\ &= |\langle x, k\rangle|^2 \\ &\geq \|x\|^2 \|k\|^2 \cos^2\theta \\ &= \|Qx\|^2 \|k\|^2 \cos^2\theta. \end{aligned} \tag{3.23}$$

Hence $x \in C_\theta^n(k) \Rightarrow Qx \in C_\theta^n(k)$ and similar for the inverse. □

For example, a reflection in the codimension-1 subspace orthogonal to k,

$$x \mapsto x - 2\frac{\langle x, k\rangle}{\|k\|^2} k \tag{3.24}$$

is an orthogonal map featuring an eigenvector k with eigenvalue -1,

$$k \mapsto -k. \tag{3.25}$$

Any rotation around the axis k is an orthogonal map with eigenvector k and eigenvalue 1, although the existence of rotational axes depends on the dimension of the space.

Scalar multiplication is also a symmetry, $x \in C_\theta^n(k) \Leftrightarrow \lambda x \in C_\theta^n(k)\ \forall \lambda \neq 0 \in \mathbb{R}$.

3.9.2 Intersection with the Unit Sphere

Although the cones are infinite in extent, their intersection with the unit $(n-1)$-sphere is of finite area. We can make use of the area of intersection to determine a lower bound for the number of cones required to cover the space, and an upper bound for the packing problem, as a function of the cones' angular size, θ.

The intersection of a single cone with the unit $(n-1)$-sphere coincides with a *spherical cap*. Spherical caps are defined as the pieces of a sphere produced by slicing the sphere with a codimension-1 hyperplane.

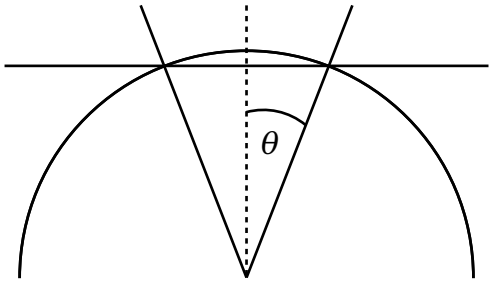


Figure 3.3: Schematic of the cone-sphere intersection and the corresponding plane-sphere intersection, resulting in a spherical cap.

Lemma 3.9.3. Let an $(n-1)$-dimensional hyperplane be described by the injective affine map $A : \mathbb{R}^{n-1} \to \mathbb{R}^n$ such that the linear part $L(y) := A(y) - A(0)$ is orthogonal to the translational part $A(0)$, i.e. $\langle A(0), L(y)\rangle = 0$ for all $y \in \mathbb{R}^{n-1}$. The intersection of the hyperplane with the unit sphere is an $(n-2)$-dimensional sphere with centre $A(0)$ and radius $\sqrt{1 - \|A(0)\|^2}$, provided $\|A(0)\| < 1$.

Proof. The intersection with the unit sphere is

$$A(\mathbb{R}^{n-1}) \cap S^{n-1} = \left\{A(y) \,:\, \|A(y)\| = 1,\ y \in \mathbb{R}^{n-1}\right\} \tag{3.26}$$

Since $\|A(y)\|^2 = \|A(0)\|^2 + \|L(y)\|^2 + 2\langle A(0), L(y)\rangle = \|A(0)\|^2 + \|L(y)\|^2$, this means $\|L(y)\|^2 = 1 - \|A(0)\|^2$ on the sphere. As $\|A(0)\|$ is constant, this equation describes a $(n-2)$-sphere in $A(\mathbb{R}^{n-1}) \subset \mathbb{R}^n$ of radius $\sqrt{1 - \|A(0)\|^2}$, with centre $A(0)$. Clearly if $\|A(0)\| > 1$ the intersection is empty. □

Lemma 3.9.4. The intersection of the boundary of a single n-cone of angle θ and direction $\hat{k}$ with the unit $(n-1)$-sphere is equal to the intersection of an $(n-1)$-plane with the unit $(n-1)$-sphere.

Proof. The intersection of the boundary of the single n-cone with the unit sphere is

$$\left\{x \in \mathbb{R}^n \,:\, \left\langle x, \hat{k}\right\rangle = \cos\theta,\ \|x\| = 1\right\}. \tag{3.27}$$

We can write an arbitrary $x \in \mathbb{R}^n$ as the orthogonal decomposition,

$$x = \mathrm{proj}_{\hat{k}}\, x + (\mathrm{id}_{\mathbb{R}^n} - \mathrm{proj}_{\hat{k}})x = x_{\parallel} + x_{\perp}, \tag{3.28}$$

which means on the intersection $x_\parallel = \cos\theta\,\hat{k}$ and $\|x_\perp\|^2 = 1 - \cos^2\theta$. Since $x_\parallel$ is constant on the intersection, we can consider it fixed, which leaves $x_\perp$ describing an $(n-1)$-dimensional subspace. Hence we can identify $x_\parallel$ with the translational part, and $x_\perp$ with the linear part, of an injective affine map

$$A(y) = \cos\theta\ \hat{k} + L(y) \tag{3.29}$$

describing an $(n-1)$-plane in $\mathbb{R}^n$ with intersection corresponding to $\|A(y)\| = 1$. □

For $0 < \theta < \pi/2$, the area of the intersection of the double n-cone with the unit sphere is therefore given by the area of a disjoint pair of antipodal spherical caps. This area can be expressed relative to the total area of the sphere by [29]

$$A_\theta^{(n)} = I_{\sin^2\theta}\left(\frac{n-1}{2}, \frac{1}{2}\right), \tag{3.30}$$

where I is the regularized incomplete beta function,

$$I_x(a,b) = \frac{B(x;a,b)}{B(a,b)}, \tag{3.31}$$

$B(x;a,b)$ is the incomplete beta function,

$$B(x;a,b) = \int_0^x t^{a-1}(1-t)^{b-1}\,\mathrm{d}t, \tag{3.32}$$

and $B(a,b) = B(1;a,b)$ is the beta function. It is straightforward to compute the derivative of $A_\theta^{(n)}$,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}A_\theta^{(n)} = \frac{2\sin^{n-2}\theta}{B(\frac{n-1}{2}, \frac{1}{2})}. \tag{3.33}$$

We can use this to determine the lowest-order non-zero term in the Taylor series in θ around 0. When taking the higher derivatives of $A_\theta^{(n)}$, each derivative will consist of terms containing a constant coefficient with powers of $\sin\theta$ and $\cos\theta$. Since any term containing a positive power of $\sin\theta$ will vanish when evaluated at 0, the first non-zero term in the series will be one whose corresponding derivative contains a term without $\sin\theta$. The first derivative to satisfy this is the $(n-1)$th derivative, which will contain a single such term,

$$\frac{\mathrm{d}^{n-1}}{\mathrm{d}\theta^{n-1}}A_\theta^{(n)} = \frac{2}{B(\frac{n-1}{2}, \frac{1}{2})}\left((n-2)!\cos^{n-2}\theta + \cdots\right). \tag{3.34}$$

There other terms in this derivative must contain positive powers of $\sin\theta$ and are therefore irrelevant. The resulting Taylor series is therefore

$$A_\theta^{(n)} = \frac{2\theta^{n-1}}{(n-1)B(\frac{n-1}{2}, \frac{1}{2})} + \mathcal{O}(\theta^n), \tag{3.35}$$

which can be used as an approximation for small θ. Note that this is an $(n-1)$-dimensional area, and so it makes sense that the first non-zero term contains the $(n-1)$th power of θ.

The Covering Problem

The covering problem asks how many double n-cones of size θ are required to cover the whole of $\mathbb{R}^n$? We can compute a lower bound for this problem by using the area of intersection with the unit sphere. In order for a set of m cones generated by vectors $(k_1, \cdots, k_m)$ to cover the space, $\bigcup_{i=1}^{m} C_\theta^n(k_i) = \mathbb{R}^n$, their areas of intersection with the sphere must sum to at least the area of the sphere. This gives a lower bound for the number of cones needed,

$$m \geq \left\lceil 1/A_\theta^{(n)} \right\rceil . \tag{3.36}$$

Alternatively, if m is given and n is unknown, assuming the cones cover the space gives an upper bound for n.

The Packing Problem

Dual to the covering problem is the packing problem, which asks how many double n-cones of angle θ can be packed into $\mathbb{R}^n$ with no nontrivial intersections? In this case we define a trivial intersection to be the singleton $\{0\}$, since by definition every cone contains the origin. In order for a set of m cones generated by vectors $(k_1, \cdots, k_m)$ to be packed into $\mathbb{R}^n$, their interiors must have pairwise disjoint intersections with the unit sphere. The total area of intersections can therefore be at most the area of the sphere, which gives an upper bound on m in terms of n,

$$m \leq \left\lfloor 1/A_\theta^{(n)} \right\rfloor . \tag{3.37}$$

Alternatively if m is given and n is unknown, this gives a lower bound for n. In other words, if n is smaller than this lower bound then $\mathbb{R}^n$ isn't large enough to pack m double n-cones of size θ.

3.9.3 Intersection with a Subspace

Proposition 3.9.5. Let $W : \mathbb{R}^d \to \mathbb{R}^n$ be an orthonormal map that defines a d-dimensional subspace of $\mathbb{R}^n$. The intersection of the subspace $W(\mathbb{R}^d)$ with a double

n	$\left\lfloor 1/A^{(n)}_{\pi/36} \right\rfloor$	$\left\lfloor 1/A^{(n)}_{\pi/18} \right\rfloor$
2	18	9
3	262	65
4	3550	445
5	46097	2903
6	584062	18451

Table 3.1: Table for $\theta = \pi/36 = 5^\circ$ and $\theta = \pi/18 = 10^\circ$

cone $C^n_\theta(k)$ corresponds to a double cone $C^d_\alpha(z)$ in $\mathbb{R}^d$, with $z = W^\dagger k$ and

$$\alpha = \arccos\left(\frac{\|k\|\cos\theta}{\|W^\dagger k\|}\right) \tag{3.38}$$

provided $\left\|W^\dagger k\right\| \geq \|k\|\cos\theta$, i.e. $W(\mathbb{R}^d) \cap C^n_\theta(k) = W(C^d_\alpha(z))$.

Proof. The intersection is

$$W(\mathbb{R}^d) \cap C^n_\theta(k) = \left\{Wy : |\langle Wy, k\rangle|^2 \geq \|Wy\|^2 \|k\|^2 \cos^2\theta, \quad y \in \mathbb{R}^d\right\}. \tag{3.39}$$

Let $z \in \mathbb{R}^d$ define a double cone in $\mathbb{R}^d$, $C^d_\alpha(z)$. Under W this is

$$W(C^d_\alpha(z)) = \left\{Wy : |\langle y, z\rangle|^2 \geq \|y\|^2 \|z\|^2 \cos^2\alpha, \quad y \in \mathbb{R}^d\right\}. \tag{3.40}$$

By definition of the adjoint, $\langle Wy, k\rangle = \left\langle y, W^\dagger k\right\rangle$, and since W is orthonormal, $\|Wy\| = \|y\|$. Therefore (3.39) and (3.40) coincide when $z = W^\dagger k$ and

$$\left\|W^\dagger k\right\|^2 \cos^2\alpha = \|k\|^2 \cos^2\theta. \tag{3.41}$$

This only has a solution for α when $\left\|W^\dagger k\right\| \geq \|k\|\cos\theta$. □

Corollary 3.9.6. If $k \in W(\mathbb{R}^d)$ then $W(\mathbb{R}^d) \cap C^n_\theta(k) = W(C^d_\theta(W^\dagger k))$.

Proof. Since $k \in W(\mathbb{R}^d)$, there exists $z \in \mathbb{R}^d$ such that $k = Wz$. Hence, since W is orthonormal, $z = W^\dagger k$ and $\left\|W^\dagger k\right\| = \|z\| = \|Wz\| = \|k\|$. It follows that $\cos^2\alpha = \cos^2\theta$ and since $0 \leq \theta \leq \pi/2$, $\alpha = \theta$. □

3.9.4 Orthogonal Projection of Double Cones

Proposition 3.9.7. Let $W : \mathbb{R}^d \to \mathbb{R}^n$ be an orthonormal map. If $k \in W(\mathbb{R}^d)$ then $W^\dagger(C^n_\theta(k)) = C^d_\theta(W^\dagger k)$.

Proof. By definition of the adjoint, $\left\langle W^\dagger x, W^\dagger k\right\rangle = \left\langle x, WW^\dagger k\right\rangle$. Since k is in the image of W, $WW^\dagger k = k$, $\left\langle W^\dagger x, W^\dagger k\right\rangle = \langle x, k\rangle$, and $\|k\| = \left\|WW^\dagger k\right\| = \left\|W^\dagger k\right\|$. We also have $\|x\| \geq \left\|W^\dagger x\right\|$. Hence

$$\begin{aligned} W^\dagger C_\theta^n(k) &= \left\{W^\dagger x \,:\, x \in C_\theta^n(k)\right\} \\ &= \left\{W^\dagger x \,:\, |\langle x, k\rangle|^2 \geq \|x\|^2 \|k\|^2 \cos^2\theta,\ x \in \mathbb{R}^n\right\} \\ &= \left\{W^\dagger x \,:\, \left|\left\langle W^\dagger x, W^\dagger k\right\rangle\right|^2 \geq \left\|W^\dagger x\right\|^2 \left\|W^\dagger k\right\|^2 \cos^2\theta,\ x \in \mathbb{R}^n\right\}. \end{aligned} \tag{3.42}$$

As $W^\dagger$ is surjective,

$$\begin{aligned} &\left\{W^\dagger x \,:\, \left|\left\langle W^\dagger x, W^\dagger k\right\rangle\right|^2 \geq \left\|W^\dagger x\right\|^2 \left\|W^\dagger k\right\|^2 \cos^2\theta,\ x \in \mathbb{R}^n\right\} \\ &= \left\{y \in \mathbb{R}^d \,:\, \left|\left\langle y, W^\dagger k\right\rangle\right|^2 \geq \|y\|^2 \left\|W^\dagger k\right\|^2 \cos^2\theta\right\} \\ &= C_\theta^d(W^\dagger k). \end{aligned} \tag{3.43}$$

□

Theorem 3.9.8. Let $W : \mathbb{R}^d \to \mathbb{R}^n$ be an orthonormal map, $k \in W(\mathbb{R}^d)$, and $P = WW^\dagger$ be the orthogonal projection onto $W(\mathbb{R}^d)$. Then $P(C_\theta^n(k)) = W(\mathbb{R}^d) \cap C_\theta^n(k)$.

Proof. By Proposition 3.9.7 and Corollary 3.9.6. □

3.9.5 Secants and Convex Combinations

Let X_1 and X_2 be finite sets of points in $\mathbb{R}^n$, and let

$$\mathcal{S}(X_1, X_2) = \{y - x \,:\, x \in X_1,\ y \in X_2\} \tag{3.44}$$

be the set of secant vectors from X_1 to X_2. Consider the convex hulls of X_1 and X_2. Pick an arbitrary point in both, which are given by the convex combinations

$$\begin{aligned} x &= \sum_i \alpha_i x_i \\ y &= \sum_j \beta_j y_j, \end{aligned} \tag{3.45}$$

where $x_i \in X_1$, $y_j \in X_2$, and both sets of coefficients are non-negative and sum to unity. Is the secant $y - x$ in the convex hull of $\mathcal{S}(X_1, X_2)$? Let $k_{ij} = y_j - x_i$ then the two coincide when

$$\begin{aligned} \sum_j \beta_j y_j - \sum_i \alpha_i x_i &= \sum_{ij} \gamma_{ij} k_{ij} \\ &= \sum_j \left(\sum_i \gamma_{ij}\right) y_j - \sum_i \left(\sum_j \gamma_{ij}\right) x_i, \end{aligned} \tag{3.46}$$

hence

$$\begin{aligned} \alpha_i &= \sum_j \gamma_{ij} \\ \beta_j &= \sum_i \gamma_{ij}. \end{aligned} \tag{3.47}$$

If the convex coefficients (γ_{ij}) are given, this determines (α_i) and (β_j) as valid convex coefficients, since

$$\begin{aligned} &\sum_{ij} \gamma_{ij} = 1 \Leftrightarrow \sum_i \alpha_i = 1 \\ &\sum_{ij} \gamma_{ij} = 1 \Leftrightarrow \sum_j \beta_j = 1 \\ &\gamma_{ij} \geq 0 \ \forall ij \Rightarrow \alpha_i \geq 0 \ \forall i \\ &\gamma_{ij} \geq 0 \ \forall ij \Rightarrow \beta_j \geq 0 \ \forall j. \end{aligned} \tag{3.48}$$

This gives a map $\operatorname{Conv} \mathcal{S}(X_1, X_2) \to \operatorname{Conv} X_1 \times \operatorname{Conv} X_2$, and we also have $\operatorname{Conv} X_1 \times \operatorname{Conv} X_2 \to \mathcal{S}(\operatorname{Conv} X_1, \operatorname{Conv} X_2) : (x, y) \mapsto (y - x)$, which when composed gives the map $\operatorname{Conv} \mathcal{S}(X_1, X_2) \to \mathcal{S}(\operatorname{Conv} X_1, \operatorname{Conv} X_2)$. This map sends secants to themselves, and therefore $\operatorname{Conv} \mathcal{S}(X_1, X_2) \subseteq \mathcal{S}(\operatorname{Conv} X_1, \operatorname{Conv} X_2)$.

If (α_i) and (β_j) are given convex coefficients, then any set of scalars (γ_{ij}) that satisfy the two equations (3.45) automatically satisfies $\sum_{ij} \gamma_{ij} = 1$, however they must be required to be non-negative in order to be valid convex coefficients.

3.9.6 Relation to the Wedge Product

Qualitatively, the double cone contains all lines through the origin that are approximate to the central axis of the cone. We can also quantify this using the norm of the wedge product – the norm is greatest when the vectors are perpendicular and zero when (anti-)parallel. We can therefore formulate the double cone in terms of the norm of the wedge product, rather than the inner product.

For decomposable bivectors, the inner product on $\Lambda^2(V)$ can be written in terms of the inner product on V as

$$\frac{1}{2} g(x \wedge y, v \wedge w) = g(x, v) g(y, w) - g(x, w) g(y, v), \tag{3.49}$$

and therefore the norm,

$$\begin{aligned}\frac{1}{2}\|v \wedge w\|^2 &= \|v\|^2 \|w\|^2 - |g(v,w)|^2 \\ &= \|v\|^2 \|w\|^2 \sin^2\theta_{vw}.\end{aligned} \tag{3.50}$$

Hence the definition of the double n-cone $C^n_\theta(k)$ can be written as

$$C^n_\theta(k) = \left\{ x \in \mathbb{R}^n \,:\, \|x \wedge k\| \leq \|x\| \|k\| \sqrt{2} \sin\theta \right\}. \tag{3.51}$$

3.10 Secant Culling Optimization

One of the limitations of secant-based methods is that the number of secants is of the order N^2, which can make the method expensive when N is large. Placing limitations on the size of N is not ideal, since a complicated submanifold may naturally require a large number of data points to accurately capture its geometry. However, we can seek to lower the computational cost by pre-processing the set of unit secants $\mathcal{K}$ to produce a new set of unit secants $\tilde{\mathcal{K}}$ whose cardinality is lower. We shall use the term *culling* to describe the process of removing secants from $\mathcal{K}$ that are either redundant, or approximately redundant, for our problem. The justification for this approach is based on two observations that we will now discuss.

Firstly, secants that are parallel or anti-parallel produce the same contribution to the cost function (3.7), since the contribution to the cost function depends only on the norm of the projected unit secant. This means we can replace a set of secants that are mutually parallel or anti-parallel by a single representative secant, weighted by the number of secants it is representing. Therefore the cost function can be rewritten as

$$\mathcal{F}(W) = \frac{1}{|\mathcal{K}|} \sum_{k \in \tilde{\mathcal{K}}} \rho_k \left\| W^\dagger k \right\|^{-1}, \tag{3.52}$$

where ρ_k is a positive integer weight for k that satisfies $\sum_{k\in\tilde{\mathcal{K}}} \rho_k = |\mathcal{K}|$, and $\tilde{\mathcal{K}}$ is the set of representative secants.

The second observation is that if the submanifold is exploring a low-dimensional space (i.e. it is confined to a d-dimensional hyperplane in $\mathbb{R}^n$), then regardless of how large N is, the secants span a space of dimension at most d. If we allow for approximate parallelism in the culling process, up to a given tolerance, then there will be a finite number of representative secants needed to account for every possibility

in the d-dimensional subspace. As we will show momentarily, the maximum number of representative secants only depends on the dimension d and the tolerance – it is independent of both n and N.

3.10.1 The Representative Secants

The culling process produces a set of unit secants $(k_1, \cdots, k_m)$ that satisfy $k_i \notin C_\theta^n(k_j)$, $i \neq j$, i.e. k_i is not approximately parallel to k_j. Note that the cones may still have nontrivial intersections. However, we can show that the corresponding set of cones with half-angle, $C_{\theta/2}^n(k_i)$, have no nontrivial intersections, $C_{\theta/2}^n(k_i) \cap C_{\theta/2}^n(k_j) = \{0\}$, $i \neq j$, and therefore their intersections with the unit sphere are pairwise disjoint.

Proposition 3.10.1. Let $C_\theta^n(k_1)$ and $C_\theta^n(k_2)$ be double cones. If $k_1 \notin C_\theta^n(k_2)$ and $k_2 \notin C_\theta^n(k_1)$ then $C_\beta^n(k_1) \cap C_\beta^n(k_2) = \{0\}$ when $\beta \leq \theta/2$.

Proof. Since $k_1 \notin C_\theta^n(k_2)$ and $k_2 \notin C_\theta^n(k_1)$,

$$|\langle k_1, k_2 \rangle|^2 < \|k_1\|^2 \|k_2\|^2 \cos^2 \theta. \tag{3.53}$$

Hence $\cos^2 \alpha < \cos^2 \theta$, where α is the angle between k_1 and k_2. By Lemma B.0.10, two double cones of angle β intersect nontrivially when $\cos^2 \alpha \geq \cos^2 2\beta$, therefore two cones of angle $\beta \leq \theta/2$ have no nontrivial intersections. □

As these half-angle cones have no nontrivial intersections, we can apply the results of the double cone packing problem, §3.9.2. Specifically, we have the upper bound on the number of representative secants,

$$m \leq \left\lfloor 1/A_{\theta/2}^{(n)} \right\rfloor. \tag{3.54}$$

However, this upper bound can be greatly improved in the case that the set generating the secants is confined to a d-dimensional hyperplane within $\mathbb{R}^n$, where $d < n$.

3.10.2 Confinement to a Hyperplane

Lemma 3.10.2. Let $Y \subset \mathbb{R}^n$. If Y is confined to a d-dimensional hyperplane, given by the injective affine map $A : \mathbb{R}^d \to \mathbb{R}^n$, then the secant vectors generated by Y are confined to a d-dimensional subspace of $\mathbb{R}^n$.

Proof. For each pair of points (y_1, y_2) in Y, since $Y \subset A(\mathbb{R}^d)$, there is a unique pair of points (x_1, x_2) in $\mathbb{R}^d$ such that $A(x_1) = y_1$ and $A(x_2) = y_2$. Since A is affine, $y_1 - y_2 = A(x_1) - A(x_2) = L(x_1 - x_2)$ as the translational terms cancel, leaving only the linear part, L. Hence since L is a rank d linear map, each secant vector lives in the d-dimensional subspace given by the image of L. □

This means the unit secants lie in a d-dimensional slice of the double cones, which are d-dimensional double cones with the same angles, by Corollary 3.9.6. Therefore, when the attractor explores a d-dimensional hyperplane of the state space, we have the improved upper bound for the number of representative secants,

$$m \leq \left\lfloor 1/A_{\theta/2}^{(d)} \right\rfloor . \tag{3.55}$$

When d is small, this upper bound can be much smaller than the total number of secants (Table 3.1), in which case the secant culling procedure is guaranteed to remove a large number of secants from the data set.

3.10.3 Implementation of Secant Culling

Let $(k_1, \cdots, k_M)$ be the total set of unit secants and $(\rho_i \in \mathbb{N}_0)_{i=1}^M$ be counts, one for each secant. The algorithm compares each pair of secants for approximate parallelism. If the two are approximately parallel, the count of one is incremented, and the count of the other is zeroed. After the algorithm has terminated, the secants with a zero count are the culled secants, and those with non-zero counts are the representative secants.

An unsigned integer data type can be used for the counts. Since a large number of secants may be culled, the value of the non-zero counts may be quite large. Therefore the data type needs to have enough bits of precision to avoid overflow during incrementation. We recommend adding an explicit test to detect overflow, and to increase the size of the data type if overflow occurs. Note that the total number of secants, M, may be very large, and each has a count, so using large integers may not be feasible due to limited amounts of memory (which is also being used to store the high-dimensional data set among other things). We found that 16-bit unsigned integers were sufficient for our larger examples. In fact, we can use the upper bound on the number of representative secants (3.55) to give a lower bound for the maximum count, by dividing

Algorithm 1 Secant culling implementation.

```
Initialize ρ_i ← 1 ∀i.
for i = 1, ··· , M do
    if ρ_i = 0 then
        continue
    end if
    for j = i + 1, ··· , M do
        if ρ_j = 0 then
            continue
        end if
        if |⟨k_i, k_j⟩|^2 ≥ cos^2 θ then
            ρ_i ← ρ_i + 1
            ρ_j ← 0
        end if
    end for
end for
```

Note that 'continue' means to jump to the next iteration of the current loop, as in the C/C++ languages.

the culled secants evenly between the representative secants,

$$\left\lceil \frac{M}{\left\lfloor 1/A_{\theta/2}^{(d)} \right\rfloor} \right\rceil \leq \max_i \rho_i \leq M. \tag{3.56}$$

For example, with 10 million secants exploring a 3-dimensional subspace and a tolerance of 10°, the max count is at least 38168, which would require a data type of at least 16-bits and at most 24-bits.

3.10.4 Minimum Projected Length

When working with the smaller set of representative secants, rather than the full set of secants, we are no longer able to determine the minimum projected length, $\kappa_{\min}$, or may not want to iterate over the full set of secants to compute it. However, we can compute a lower bound for this value in terms of the representative secants and the tolerance.

Let $\tilde{k}$ be a representative unit secant and $C_\theta^n(\tilde{k})$ be its double n-cone. Also, let

$\left\|W^\dagger \tilde{k}\right\| = \cos\alpha$. Then for all unit secants k in $C_\theta^n(\tilde{k})$,

$$\begin{aligned} \left\|W^\dagger k\right\| &\geq \cos(\alpha+\theta) \\ &= \cos\alpha\cos\theta - \sin\alpha\sin\theta \\ &= \left\|W^\dagger \tilde{k}\right\|\cos\theta - \sqrt{1-\left\|W^\dagger \tilde{k}\right\|^2}\sin\theta. \end{aligned} \tag{3.57}$$

For a given angle $0 \leq \theta < \pi/2$, this lower bound is strictly increasing w.r.t. $\left\|W^\dagger \tilde{k}\right\|$ in the domain $[0,1]$. For $\sin\theta = 0$ this is trivial. For $\sin\theta \neq 0$, it has a single turning point at $-\cos\theta$ with value -1 and second derivative $1/\sin^2\theta$ (a minimum). When $0 < \theta < \pi/2$, this turning point is below 0. Therefore the minimum projected length of all secants can be given a lower bound in terms of the minimum projected length of the representative secants,

$$\begin{aligned} \kappa_{\min} &= \min_{k\in\mathcal{K}} \left\|W^\dagger k\right\| \\ &\geq \tilde{\kappa}_{\min}\cos\theta - \sqrt{1-\tilde{\kappa}_{\min}^2}\sin\theta, \end{aligned} \tag{3.58}$$

where

$$\tilde{\kappa}_{\min} = \min_{\tilde{k}\in\tilde{\mathcal{K}}} \left\|W^\dagger \tilde{k}\right\|. \tag{3.59}$$

Chapter 4

Reproducing Dynamics From General State Space Manifolds

Most conventional methods of dimensionality reduction for dynamical systems are designed for state spaces with a vector space structure, and make use of linear projections onto subspaces. In this chapter, we formulate the problem for general state space manifolds and investigate both the reproduction of the dynamics on the attractor, and the reproduction of the stability of the attractor in the reduced space. This introduces interesting complications in contrast to the case of linear projections, which we resolve in a way that allows for a practical implementation. We then adapt the secant-Grassmann projection method discussed in §3 for use with state space manifolds beyond $\mathbb{R}^n$. We present two practical implementations, a Galerkin-style method applicable in a special case, and an optimization approach used in [4], which is applicable in generality.

4.1 Introduction

The goal of a method of dimension reduction is to produce a low-dimensional system that captures a particular structure found in a given high-dimensional system. In the case of attractors of dynamical systems, the relevant structure consists of a low-dimensional manifold together with the dynamics on the manifold. Also of relevance is the stability of the attractor, which describes the attractor's relationship with its ambient state space.

In many examples of dynamical systems, the state space of the system has an algebraic structure that is used to describe its dynamics. Particularly common are vector spaces with some additional structure, such as Euclidean spaces in the finite-dimensional case, and topological vector spaces (e.g. Hilbert/Banach) in the infinite-dimensional case. This algebraic/analytic structure is used to express the time evolution, for example, by a differential equation. Because of this, many existing methods centre around using the algebraic structure to perform the reduction, both to map the high-dimensional space to a low-dimensional space, and to determine the reduced dynamics. For example, orthogonal projections are a common choice of dimension-reducing map on an inner product space, such as those produced by the POD / KL / SVD, and the Galerkin projection is a method for finding reduced dynamics by algebraically manipulating the original differential equation.

In contrast to this, we can take a more geometric approach by regarding the system as a state space manifold with dynamics provided by a vector field, independent of any algebraic structure that may be used to describe the vector field. This allows us to focus on the relevant structure (the orbits in the state space), while not being constrained to working within a particular algebraic structure, such as the linear algebra commonly used in the literature. The limitations of a direct linear algebraic approach are particularly felt when dealing with nonlinear dynamics, where one must resort to local approximations (such as a Taylor polynomial around a fixed point) or by constraining oneself to particular classes of nonlinearity, in order to be able to fully apply linear methods.

An additional benefit to taking a geometric approach is that it allows for examples that contain angular variables. A simple example of this is the pendulum, whose state consists of an angular position and angular velocity, i.e. the cylinder $S^1 \times \mathbb{R}$. The angular nature of this position variable must be taken into account in order to properly identify periodic behaviour in orbits that wind around the cylinder. However, an algebraic treatment of this system usually regards the state space to be $\mathbb{R}^2$ (i.e. using the universal cover of the circle), in which the periodicity of the angle must be manually accounted for.

For a system that contains an explicit time dependence, one may describe its

dynamics as a time-dependent vector field, corresponding to a non-autonomous differential equation. The downside to this interpretation is that the manifold no longer enjoys the properties of a conventional state space. In particular, the future orbit is no longer determined solely by the current state – the current time is also required. It therefore explicitly couples the parameterization of the orbit (i.e. time) to the dynamics of the system. However, an alternate perspective can be taken in which the non-autonomous system is made autonomous by encapsulating the time dependence in additional state variables. In the case of periodic time dependence, the time dependence is determined by an angular variable that forms the circle, S^1, and the new state space is the product manifold of S^1 with the original state space. The orbits of this new system wind around the circle, and hence the angular nature of the extra variable must be acknowledged in order to correctly identify periodic behaviour in the whole system. By allowing for general state space manifolds, we can use our approach with some non-autonomous examples by encapsulating the explicit time dependence as additional state.

4.2 Use of Data From Numerical Simulations

One of the main obstacles in developing a method of dimensionality reduction for attractors is that one does not have an explicit description of the attractor. The attractor is an object that results from the long-term behaviour of the orbits of the flow of the vector field, and in general one does not have a closed form expression of the flow, which would require solving a large system of nonlinear differential equations. One therefore doesn't have any information about the properties of an attractor, such as its dimension, shape, or particular location in the state space. Instead, data points from numerical simulations can be used to approximately describe the attractor.

When obtaining data one must take care to obtain data that explores the full extent of the attractor, with sufficient density to capture high-frequency details. If there are multiple attractors each with a basin of attraction, one must obtain data from each.

4.3 Geometric Formulation

We begin by formulating the problem of the reduction of attractors and reproduction of their dynamics in a general geometric setting. Autonomous systems can be thought of as a state space manifold with a vector field. The flow of the vector field provides the dynamical evolution of the system, which produces orbits that are integral curves of the underlying vector field. We are interested in attractors that are invariant submanifolds under the flow of the vector field. Geometrically, this invariance corresponds to the vector field being tangent to the submanifold.

Let X be the state space as a smooth manifold and $V \in \mathfrak{X}(X)$ be a smooth vector field. Let $\mathcal{M} \subset X$ be a compact submanifold without boundary invariant under the flow of V. We wish to map the original system (X, V) to a reduced system $(\hat{X}, \hat{V})$ by obtaining the map $\varphi : X \to \hat{X}$. Note that in this general formulation, both X and $\hat{X}$ are smooth manifolds, and while (X, V) is given, we have the freedom to choose $\hat{X}$ such that φ and $\hat{V}$ are easy to find.

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & \hat{X} \\ V \downarrow & & \downarrow \hat{V} \\ \mathrm{T}X & \xrightarrow{\varphi_*} & \mathrm{T}\hat{X} \end{array}$$

We will describe φ as the state-space map/dimension-reducing map. As we wish to reproduce the attractor and its dynamics, we require the map to satisfy certain properties. In particular, we require φ to be a smooth map that embeds $\mathcal{M}$ into $\hat{X}$. This makes it possible to reproduce the orbits on the submanifold $\varphi(\mathcal{M})$ by the flow of a vector field on $\hat{X}$ and ensures no discontinuous jumps are introduced into the reduced dynamics.

4.3.1 Preserving the Flow

Given a state-space mapping, φ, that satisfies the above properties, we want to reproduce the dynamics on $\mathcal{M}$ in the reduced space. This means the corresponding vector field $\hat{V}$ on $\hat{X}$ is chosen such that φ preserves the flow on $\mathcal{M}$,

$$\Phi^t_{\hat{V}}(\varphi(x)) = \varphi\left(\Phi^t_V(x)\right) \qquad \forall x \in \mathcal{M},\ \forall t, \tag{4.1}$$

where Φ_V^t is the flow of V for a time t. As one would expect, taking the derivative of this shows that we require,

$$\hat{V}_{\varphi(x)} = \varphi_* V_x \qquad \forall x \in \mathcal{M}, \tag{4.2}$$

where $\varphi_* : \mathrm{T}X \to \mathrm{T}\hat{X}$ is the differential of φ.

4.4 Reproducing a Stable Attractor

Although matching the tangent vectors is sufficient to reproduce the flow on the submanifold itself, we would also like to reproduce the stability of the submanifold in order to produce a stable attractor in the reduced space. This is especially important since the reduced vector field may be subject to approximation error in a practical implementation. In order to achieve this, the behaviour of the vector field in the neighbourhood of $\varphi(\mathcal{M})$ must be specified. We therefore must consider the derivatives of the vector field on $\hat{X}$ in directions normal to the submanifold.

4.4.1 Tangent and Normal Spaces

In order to produce a stable attractor in the reduced system, we can make use of knowledge of the original vector field and its derivatives along the attractor, however we need to be careful as we are dealing with vector fields on two different manifolds, each with its own metric and connection. This raises the question of how to make use of the original vector field to specify the derivatives of the reduced vector field in a meaningful way that has the desired interpretation.

Let (X, g) and $(\hat{X}, \hat{g})$ be Riemannian manifolds, V be a vector field on X, $\mathcal{M} \subset X$ be a V-invariant submanifold of X, and $\varphi : X \to \hat{X}$ be a smooth map that embeds $\mathcal{M}$ into $\hat{X}$. We can make a few observations regarding the normal and tangent bundles of $\mathcal{M}$ in X and $\varphi(\mathcal{M})$ in $\hat{X}$.

- Since φ embeds $\mathcal{M}$ into $\hat{X}$, the tangent spaces of $\mathcal{M}$ are preserved under φ_*, i.e. $\varphi_*(\mathrm{T}\mathcal{M}) = \mathrm{T}\varphi(\mathcal{M})$.

- The normal bundle of $\mathcal{M}$ in X, $\mathrm{N}\mathcal{M}$, may not be preserved. In fact, since we are dealing with dimension-reducing maps, it is likely that the normal bundle will be collapsed significantly by φ_*.

- Since $\varphi_*|_{\mathrm{T}_{\mathcal{M}}X}$ is not necessarily surjective, the tangent spaces of $\hat{X}$ along $\varphi(\mathcal{M})$ may contain directions not present in $\varphi_*(\mathrm{T}_{\mathcal{M}}X)$. Also note that $\varphi_*(\mathrm{T}_{\mathcal{M}}X)$ may not be a vector bundle since the rank of φ_* may not be constant over connected components of $\mathcal{M}$. This means each tangent space can be written as

$$\mathrm{T}_{\varphi(x)}\hat{X} = \operatorname{im}(\varphi_*)_x \oplus (\operatorname{im}(\varphi_*)_x)^\perp, \qquad x \in \mathcal{M}, \tag{4.3}$$

  where $(\operatorname{im}(\varphi_*)_x)^\perp$ is only empty if $(\varphi_*)_x$ is surjective.

As a result of these observations, the reduced tangent space at each point can be written as

$$\begin{aligned} \mathrm{T}_{\varphi(x)}\hat{X} &= \operatorname{im}(\varphi_*)_x \oplus (\operatorname{im}(\varphi_*)_x)^\perp \\ &= (\varphi_*)_x(\mathrm{T}_x\mathcal{M}) \oplus (\varphi_*)_x(\mathrm{N}_x\mathcal{M}) \oplus (\operatorname{im}(\varphi_*)_x)^\perp \\ &= \mathrm{T}_{\varphi(x)}\varphi(\mathcal{M}) \oplus (\varphi_*)_x(\mathrm{N}_x\mathcal{M}) \oplus (\operatorname{im}(\varphi_*)_x)^\perp \end{aligned} \tag{4.4}$$

i.e. we have directions that are tangential to the attractor, all of which are preserved in the reduced space, the image of directions normal to the original attractor, and directions normal to the attractor in the reduced space that are not mapped onto by φ_*. Note that $(\varphi_*)_x(\mathrm{N}_x\mathcal{M})$ may not be orthogonal to $\mathrm{T}_{\varphi(x)}\varphi(\mathcal{M})$ in the reduced space.

This is important for the stability of the attractor, as we need to specify the directional derivative of the vector field in all of the normal directions, and a lack of surjectivity of φ_* means that there are directions for which the original vector field cannot provide any information about stability. This is reflected in the kernel of the adjoint of φ_*, since at each point $\ker\varphi_*^\dagger = (\operatorname{im}\varphi_*)^\perp$, by Theorem A.1.1.

4.4.2 Pushforward of the Covariant Derivative

One method of using the original vector field to specify the derivatives of the reduced vector field is to compute the derivatives of V on X and then apply φ_* to push the result forward to $\hat{X}$,

$$\hat{\nabla}_{\hat{Z}}\hat{V} = \varphi_*\nabla_Z V. \tag{4.5}$$

In order to relate each $\hat{Z} \in \mathrm{T}\hat{X}$ to a $Z \in \mathrm{T}X$, we can require $\hat{g}(\hat{Z}, \hat{\nabla}_{\hat{Z}}\hat{V}) = g(Z, \nabla_Z V)$,

$$\begin{aligned} \hat{g}(\hat{Z}, \hat{\nabla}_{\hat{Z}}\hat{V}) &= \hat{g}(\hat{Z}, \varphi_*\nabla_Z V) \\ &= g(\varphi_*^\dagger\hat{Z}, \nabla_Z V). \end{aligned} \tag{4.6}$$

Therefore $\hat{g}(\hat{Z}, \hat{\nabla}_{\hat{Z}}\hat{V}) = g(Z, \nabla_Z V)$ when $Z = \varphi_*^\dagger \hat{Z}$, giving

$$\hat{\nabla}_{\hat{Z}}\hat{V} = \varphi_* \nabla_{\varphi_*^\dagger \hat{Z}} V. \tag{4.7}$$

We can now observe two issues with this method. Firstly, by Theorem A.1.1, $\varphi_*^\dagger \hat{Z} = 0$ for all $\hat{Z}$ in the orthogonal complement of $\operatorname{im} \varphi_*$. This means that, as expected, this approach can only be used to specify the derivatives for directions in the image of φ_*. Therefore, if φ_* is non-surjective, there are directions whose derivative cannot be specified in terms of the original vector field.

Secondly, since we are computing the derivative on X with ∇ and then pushing the result forward to $\hat{X}$, the values of these directional derivatives always lie within the image of φ_*. This means that this approach cannot capture the extrinsic curvature introduced by φ, which is quantified by the normal components of the derivatives (§2.2.4).

Lemma 4.4.1. Let (X, g) be a Riemannian manifold with Levi-Civita connection ∇, and V be a smooth vector field on X. If $g(\nabla_{e_i} V, e_j) = -c_i \delta_{ij}$, where $c_i > 0$, for all $i, j = 1, \cdots, k$, then $g(\nabla_Z V, Z) < 0$ for all $Z \neq 0 \in \operatorname{span}\{e_1, \cdots, e_k\}$.

Proof. Let $Z = Z^i e_i$, then

$$\begin{aligned} g(\nabla_Z V, Z) &= \sum_{ij} Z^i Z^j g(\nabla_{e_i} V, e_j) \\ &= -\sum_{ij} Z^i Z^j c_i \delta_{ij} \\ &= -\sum_i \left|Z^i\right|^2 c_i \\ &< 0. \end{aligned} \tag{4.8}$$

□

4.4.3 Constructing a Vector Field in the Neighbourhood

Rather than computing the derivative on X with ∇, we can instead construct a vector field in a neighbourhood of $\varphi(\mathcal{M})$ in $\hat{X}$ and then use $\hat{\nabla}$ to compute the derivative. This has the advantage of being able to capture the extrinsic curvature introduced by φ. Let $\mathcal{N}$ be a neighbourhood of $\varphi(\mathcal{M})$ in $\hat{X}$, and let $\psi : \mathcal{N} \to X$ be a smooth

map that provides a non-unique choice of inverse. Then define a vector field on $\mathcal{N}$ by $Y_p = \varphi_* V_{\psi(p)}$. The covariant derivative of this is given in coordinates by

$$\hat{\nabla}_j Y^i = \left(\partial_a \varphi^i \; \partial_c V^a + \partial_c \partial_a \varphi^i \; V^a\right) \hat{\partial}_j \psi^c + \hat{\Gamma}^i{}_{jb} \, \partial_a \varphi^b \, V^a \tag{4.9}$$

Note the three terms on the RHS: the first is the derivative in X, the second accounts for the extrinsic curvature introduced by φ, and the third accounts for the intrinsic curvature of $\hat{X}$.

This approach is better than the pushforward of the covariant derivative as it is able to capture extrinsic curvature. However, it is still not appropriate for directions in $(\mathrm{im}\,(\varphi_*)_x)^\perp$, due to a lack of information in the original.

4.4.4 Enforcing Stability

The original vector field cannot provide any information about the derivatives for directions in $(\mathrm{im}\,(\varphi_*)_x)^\perp$. In order to specify the derivatives in these directions, we can manually enforce stability. However to do so, we need to ensure that this does not disrupt the dynamics on the attractor itself, i.e. the derivatives in the tangential directions are unaffected.

To solve this problem, we can look for an orthogonal decomposition of each tangent space along $\varphi(\mathcal{M})$, $\mathrm{T}_{\hat{x}}\hat{X} = E_{\hat{x}} \oplus F_{\hat{x}}$. If $\mathrm{T}_{\hat{x}}\varphi(\mathcal{M}) \subseteq E_{\hat{x}} \subseteq \mathrm{im}\,(\varphi_*)_x$ then we can use the original vector field to specify derivatives for directions in $E_{\hat{x}}$, and manually specify the derivatives for directions in $F_{\hat{x}}$. This condition is required as we do not wish to disturb the dynamics on the attractor, so $E_{\hat{x}}$ must contain the tangents to the attractor.

Let A be a $(1,1)$ tensor at $\hat{x}$ that provides an orthogonal projection onto $E_{\hat{x}}$. This allows us to decompose an arbitrary vector $\hat{Z} \in \mathrm{T}_{\hat{x}}\hat{X}$ as

$$\hat{Z} = \hat{Z}_\parallel + \hat{Z}_\perp := A\hat{Z} + (\mathrm{id} - A)\hat{Z}. \tag{4.10}$$

As $\hat{Z}_\parallel$ is in the image of $(\varphi_*)_x$, we can use information from the original model to specify the derivative in this direction, using §4.4.3.

For the component $\hat{Z}_\perp$, the original vector field cannot provide any information, and so we can enforce local stability in this direction. To achieve this we require $\hat{g}(\hat{\nabla}_{\hat{Z}_\perp}\hat{V}, \hat{Z}_\perp) < 0$, and we can choose to specify the derivative as

$$\hat{\nabla}_{\hat{Z}_\perp}\hat{V} = -k\hat{Z}_\perp, \tag{4.11}$$

where $k > 0$ is a parameter that dictates the strength of attraction (we use $k = 1$ by default). The two directional derivatives can be combined via linearity to give

$$\begin{aligned}\hat{\nabla}_{\hat{Z}}\hat{V} &= \hat{\nabla}_{\hat{Z}_\parallel}\hat{V} + \hat{\nabla}_{\hat{Z}_\perp}\hat{V} \\ &= \hat{\nabla}_{A\hat{Z}}Y - k(\mathrm{id} - A)\hat{Z},\end{aligned} \tag{4.12}$$

which is a linear equation in $\hat{Z}$. This allows us to specify the derivatives of $\hat{V}$ on the submanifold in all directions on $\hat{X}$. In coordinates, this is

$$\hat{\nabla}_j \hat{V}^i = (\hat{\nabla}_a Y)^i A^a_j - k(\delta^i_j - A^i_j). \tag{4.13}$$

The only problem that remains is to determine the orthogonal projection A from the available information. We typically don't have an explicit description of the attractor or its tangent spaces, but we do know φ. Therefore we can choose $E_{\hat{x}} = \mathrm{im}\,(\varphi_*)_x$.

4.4.5 Decomposition via SVD

If we choose $E_{\hat{x}} = \mathrm{im}\,(\varphi_*)_x$, we can construct A at each point by taking an SVD of $(\varphi_*)_x$.

Proposition 4.4.2. Let $r = \mathrm{rank}(\varphi_*)_x$ and $(u_a)_{a=1}^r$ be the first r left-singular vectors of $(\varphi_*)_x$. Then

$$A = \sum_{a=1}^r u_a \otimes (u_a)^\flat, \qquad A^i{}_j = \hat{g}_{jk}\sum_{a=1}^r (u_a)^i (u_a)^k, \tag{4.14}$$

is an orthogonal projection onto $\mathrm{im}\,(\varphi_*)_x$.

Proof. A pure (1,1) tensor $v \otimes \alpha$ under the adjoint is $(v \otimes \alpha)^\dagger = \alpha^\sharp \otimes v^\flat$. Hence,

$$A^\dagger = \sum_{a=1}^r (u_a \otimes (u_a)^\flat)^\dagger = \sum_{a=1}^r u_a \otimes (u_a)^\flat = A \tag{4.15}$$

A is self-adjoint. Also,

$$\begin{aligned}A^2 &= \sum_{a=1}^r \sum_{b=1}^r u_a \otimes (u_a)^\flat (u_b) \otimes (u_b)^\flat \\ &= \sum_{a=1}^r \sum_{b=1}^r \delta_{ab}\ u_a \otimes (u_b)^\flat \\ &= \sum_{a=1}^r u_a \otimes (u_a)^\flat \\ &= A\end{aligned} \tag{4.16}$$

A is a projection. Hence A is an orthogonal projection. The image is given by $\operatorname{span}\{u_1, \cdots, u_r\} = \operatorname{im}(\varphi_*)_x$. □

When φ is a submersion, then $\operatorname{im}\varphi_* = \mathrm{T}_{\varphi(X)}\hat{X}$ and no decomposition is needed, so we can choose $A = \mathrm{id}$.

4.5 A Worked Example: The Cylinder

To illustrate the above, we consider a simple example. Let the state space be the cylinder, $X = S^1 \times \mathbb{R}$ with $(\theta, z) \in X$. Let V be a vector field on the cylinder of the form

$$V = \partial_\theta + f(\theta, z)\partial_z, \tag{4.17}$$

where $f : X \to \mathbb{R}$ is a smooth function. As a system of ODEs this is

$$\begin{bmatrix} \dot{\theta} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 1 \\ f(\theta, z) \end{bmatrix}. \tag{4.18}$$

This means all orbits wind around the cylinder with an angular velocity of 1. Let the attractor $\mathcal{M}$ be a periodic orbit winding around the cylinder.

We can embed the cylinder into $\mathbb{R}^3$ in the usual way, and then orthogonally project onto the xy-plane. This gives the state-space map $\varphi : X \to \mathbb{R}^2 : (\theta, z) \mapsto (x, y) = (\cos\theta, \sin\theta)$. Let the inverse approximation be $\psi : \mathbb{R}^2 \setminus \{0\} \to X : (x, y) \mapsto (\operatorname{atan2}(y, x), c)$, where $c \in \mathbb{R}$ is a constant. We can compute the following quantities, which we express using matrix notation,

$$\varphi_* = \begin{bmatrix} -\sin\theta & 0 \\ \cos\theta & 0 \end{bmatrix} = \begin{bmatrix} -y & 0 \\ x & 0 \end{bmatrix} \tag{4.19}$$

$$\psi_* = \begin{bmatrix} -y & x \\ 0 & 0 \end{bmatrix} = \varphi_*^\dagger \tag{4.20}$$

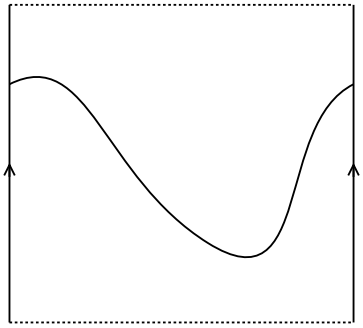
$$\varphi_* V = \begin{bmatrix} -y \\ x \end{bmatrix} \tag{4.21}$$

$$\partial V = \begin{bmatrix} 0 & 0 \\ \partial_\theta f & \partial_z f \end{bmatrix} \tag{4.22}$$

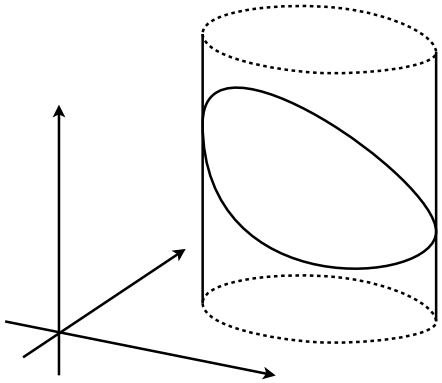
$$(\partial\partial\varphi)V = \begin{bmatrix} -\cos\theta & 0 \\ -\sin\theta & 0 \end{bmatrix} = \begin{bmatrix} -x & 0 \\ -y & 0 \end{bmatrix} \tag{4.23}$$

$$\varphi_* \partial V = \begin{bmatrix} -y & 0 \\ x & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \partial_\theta f & \partial_z f \end{bmatrix} = 0 \tag{4.24}$$

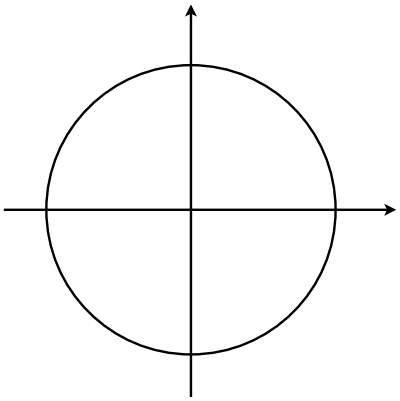
$$\hat{\nabla}(\varphi_* V_{\psi(x,y)}) = \begin{bmatrix} -x & 0 \\ -y & 0 \end{bmatrix} \begin{bmatrix} -y & x \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} xy & -x^2 \\ y^2 & -xy \end{bmatrix} \tag{4.25}$$



(a) $S^1 \times \mathbb{R}$



(b) $S^1 \times \mathbb{R} \to \mathbb{R}^3$



(c) $S^1 \times \mathbb{R} \to \mathbb{R}^3 \to \mathbb{R}^2$

Figure 4.1: Schematic of an orbit winding around the cylinder, its embedding into $\mathbb{R}^3$, and projection onto $\mathbb{R}^2$.

Since φ is non-submersive, we need to decompose the tangent space in $\mathbb{R}^2$. The image of φ_* is just the tangent space to the circle, and its orthogonal complement is the radial direction. At a point (x, y) the tangential component is spanned by $(-y, x)$, and the radial component is spanned by (x, y). The derivative along the orbit gives

$$\hat{\nabla}_{\varphi_* V}(\varphi_* V_{\psi(x,y)}) = \begin{bmatrix} xy & -x^2 \\ y^2 & -xy \end{bmatrix} \begin{bmatrix} -y \\ x \end{bmatrix} = - \begin{bmatrix} x \\ y \end{bmatrix}, \tag{4.26}$$

which is the expected radial acceleration associated with circular motion. The orthogonal projection onto the tangential component is given by

$$\begin{aligned} A &= (-y, x) \otimes (-y, x)^\flat \\ &= \begin{bmatrix} -y \\ x \end{bmatrix} \begin{bmatrix} -y & x \end{bmatrix} \\ &= \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}. \end{aligned} \tag{4.27}$$

Therefore the reduced vector field on $\mathbb{R}^2$ should have values on the unit circle given by

$$\hat{V}_{(x,y)} = \varphi_* V = \begin{bmatrix} -y \\ x \end{bmatrix}, \tag{4.28}$$

and derivatives on the circle of

$$\begin{aligned} \hat{\nabla}\hat{V}_{(x,y)} &= \nabla V_{\psi(x,y)} A - k(I - A) \\ &= \begin{bmatrix} xy & -x^2 \\ y^2 & -xy \end{bmatrix} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} - k \begin{bmatrix} 1 - y^2 & xy \\ xy & 1 - x^2 \end{bmatrix} \\ &= \begin{bmatrix} xy & -x^2 \\ y^2 & -xy \end{bmatrix} - k \begin{bmatrix} x^2 & xy \\ xy & y^2 \end{bmatrix} \\ &= \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} y & -x \end{bmatrix} - k \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix}. \end{aligned} \tag{4.29}$$

Note that these derivatives always take values in the radial direction.

We can also illustrate what would happen if one were to compute the derivatives on X and then push the result forward to X using the method of §4.4.2. If X is given the Riemannian metric induced by its embedding into $\mathbb{R}^3$, then we can compute the following quantities.

$$\alpha_* = \begin{bmatrix} -\sin\theta & 0 \\ \cos\theta & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -y & 0 \\ x & 0 \\ 0 & 1 \end{bmatrix} \tag{4.30}$$

$$(\partial\partial\alpha)V = \begin{bmatrix} -\cos\theta & 0 \\ -\sin\theta & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -x & 0 \\ -y & 0 \\ 0 & 0 \end{bmatrix} \tag{4.31}$$

$$g = \begin{bmatrix} x^2 + y^2 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{4.32}$$

$$\alpha_*^\dagger = \begin{bmatrix} -y & x & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.33}$$

$$\varphi_* \alpha_*^\dagger = \begin{bmatrix} -y & 0 \\ x & 0 \end{bmatrix} \begin{bmatrix} -y & x & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} y^2 & -xy & 0 \\ -xy & x^2 & 0 \end{bmatrix} \tag{4.34}$$

Therefore the pushforward of the covariant derivative is

$$\begin{aligned} \varphi_* \nabla V &= \varphi_* \partial V + \varphi_* \alpha_*^\dagger \, (\partial \partial \alpha) V \\ &= 0 + \begin{bmatrix} y^2 & -xy & 0 \\ -xy & x^2 & 0 \end{bmatrix} \begin{bmatrix} -x & 0 \\ -y & 0 \\ 0 & 0 \end{bmatrix} \\ &= 0. \end{aligned} \tag{4.35}$$

This demonstrates that the method of constructing a vector field in a neighbourhood is superior to using the pushforward of the covariant derivative, since the latter fails to produce derivatives that describe the circular motion in the plane, which is a result of the embedding of the state space.

4.6 Special Case: Multilinear Series

Here we define a special case where the family of reduced vector fields can be obtained directly using a top-down algebraic manipulation of the original. This case generalizes many of the examples that are used with a Galerkin procedure, with linear maps used for the projection and its inverse, and a particular form of nonlinear dynamics that allows for the computation of low-dimensional reduced dynamics. This includes local approximations of general nonlinearities, since a Taylor polynomial also has this form. We refer to this form of dynamics as a multilinear series.

Let X be a vector space. Define a *multilinear series* as a map $M : X \to X$ of the form

$$M(x) = \sum_{k=0}^{m} M_k(\underbrace{x, \cdots, x}_{k \text{ times}}), \tag{4.36}$$

where $M_k : X^k \to X$ is a k-linear map in X, and $m \in \mathbb{N}_0$ is the *degree* of M. Since each multilinear term is being evaluated by a single repeated argument, we can require each M_k to be totally symmetric without loss of generality. Given a basis for X, M can be written in components as

$$(M(x))^i = \sum_{k=0}^{m} (M_k)^i{}_{j_1 \cdots j_k} x^{j_1} \cdots x^{j_k}, \tag{4.37}$$

Addition and scalar multiplication can be defined in the usual way, $(M + N)(x) = M(x) + N(x)$ and $(\lambda M)(x) = \lambda M(x)$, to give the set of degree m multilinear series a vector space structure. The dimension of this vector space is

$$n \sum_{k=0}^{m} \binom{n+k-1}{k} = n \binom{m+n}{m} = \frac{(m+n)!}{m!(n-1)!}, \tag{4.38}$$

where $n = \dim X > 0$. The space of degree m multilinear series is closed under composition with linear maps A,

$$\begin{aligned} (M \circ A(x))^i &= \sum_{k=0}^{m} (M_k)^i{}_{a_1 \cdots a_k} A^{a_1}{}_{j_1} \cdots A^{a_k}{}_{j_k} x^{j_1} \cdots x^{j_k}, \\ (A \circ M(x))^i &= \sum_{k=0}^{m} A^i{}_a (M_k)^a{}_{j_1 \cdots j_k} x^{j_1} \cdots x^{j_k}. \end{aligned} \tag{4.39}$$

It also satisfies

$$M(\lambda x) = \sum_{k=0}^{m} \lambda^k (M_k)^i{}_{j_1 \cdots j_k} x^{j_1} \cdots x^{j_k} \tag{4.40}$$

and $M(0) = M_0$.

Let X and $\hat{X}$ be vector spaces and $\varphi : X \to \hat{X}$ be a surjective linear map. Since we are dealing with vector spaces and linear maps, φ and φ_* coincide. Let ψ be a linear map defined on the whole of $\hat{X}$. Let the original vector field be given by a multilinear series for state dependence,

$$V_x = \sum_{k=0}^{m} M_k(\underbrace{x, \cdots, x}_{k \text{ times}}). \tag{4.41}$$

Then we can apply (4.2),

$$\begin{aligned} \hat{V}_{\hat{x}} &= \varphi_* \, V_{\psi(\hat{x})} \\ &= \varphi_* \sum_k M_k(\psi(\hat{x}), \cdots, \psi(\hat{x})) \\ &= \sum_k \hat{M}_k(\hat{x}, \cdots, \hat{x}) \end{aligned} \tag{4.42}$$

where $\hat{M}_k : \hat{X}^k \to \hat{X}$ is an k-linear map in $\hat{X}$ given by

$$(\hat{M}_k)^{\mu}{}_{\nu_1 \cdots \nu_k} = \varphi^{\mu}_{\ i} \ (M_k)^{i}{}_{j_1 \cdots j_k} \ \psi^{j_1}{}_{\nu_1} \cdots \psi^{j_k}{}_{\nu_k}. \tag{4.43}$$

One can pre-compute these quantities to produce an explicitly low-dimensional vector field. It is straightforward to verify that this reduced vector field has derivatives that satisfy (4.12),

$$\begin{aligned} \hat{\nabla}_{\hat{Z}}(\hat{V}_{\hat{x}}) &= \sum_k k\hat{M}_k(\hat{Z}, \hat{x}, \cdots, \hat{x}) \\ &= \sum_k k\varphi_* M_k(\psi(\hat{Z}), \psi(\hat{x}), \cdots, \psi(\hat{x})) \\ &= \hat{\nabla}_{\hat{Z}}(\varphi_* V_{\psi(\hat{x})}). \end{aligned} \tag{4.44}$$

This shows that the common top-down approach to obtaining reduced dynamics produces vector fields that satisfy the properties described in §4.3.1 and §4.4, assuming a linear inverse. In practice, a linear map may only be an approximation of the true (nonlinear) inverse.

4.7 Using Secant-Grassmann Projection with State Space Manifolds

To make use of the secant-based projection method discussed in §3, the attractor needs to be an embedded submanifold of some ambient vector space in order to make use of its secants. This makes attractors of a dynamical system with state space $\mathbb{R}^n$ a compatible example. However, when dealing with more general state space manifolds, we do not have this required structure.

In order to use secant-based projection in this case, we can embed the state space X into $\mathbb{R}^n$ for some n by the Whitney embedding theorem, and choose the reduced state space to be $\hat{X} = \mathbb{R}^d$. If $\alpha : X \to \mathbb{R}^n$ is the chosen embedding, and the orthonormal map $W : \mathbb{R}^d \to \mathbb{R}^n$ describes the orthogonal projection onto the subspace $W(\mathbb{R}^d)$, then the dimension-reducing map is $\varphi = W^\dagger \circ \alpha$, and its differential is $\varphi_* = W^\dagger \alpha_*$ with adjoint $\varphi_*^\dagger = \alpha_*^\dagger W$. By using the Euclidean inner product $\langle \, , \rangle$ on this intermediate space, $\mathbb{R}^n$, the embedding induces the pullback metric on X,

$$g(V_1, V_2) = \langle \alpha_* V_1, \alpha_* V_2 \rangle \, , \tag{4.45}$$

and similarly W induces the Euclidean inner product on $\mathbb{R}^d$ by definition. This makes α_* an orthonormal map, $\alpha_*^\dagger \alpha_* = \mathrm{id}_{TX}$. This is useful, as we typically aren't given a particular metric on the state space of a dynamical system.

The set of unit secants can then be generated by $\alpha(\mathcal{M})$ in $\mathbb{R}^n$,

$$\mathcal{K}(\alpha(\mathcal{M})) = \left\{ \frac{x-y}{\|x-y\|} \ : \ x \neq y, \ x, y \in \alpha(\mathcal{M}) \right\}. \tag{4.46}$$

4.7.1 Inverse

In §4.4 we require a (non-unique) choice of inverse of φ in a neighbourhood of the attractor, denoted ψ, in order to determine the reduced dynamics. In particular we only need the differential ψ_* along $\varphi(\mathcal{M})$ to specify the stability, (4.9). In §3.8 we detailed a simple affine inverse for the orthogonal projection generated by the secant-Grassmann method. Therefore if we choose $\psi = \alpha^L(W\hat{x}+\overline{z})$, where α^L is a left-inverse of α, we can compute the differential, $\psi_* = \alpha_*^L W$. Since α_* is orthonormal, the adjoint is its left-inverse, therefore $\psi_* = \alpha_*^\dagger W = \varphi_*^\dagger$. In practice the affine inverse of the orthogonal projection will be an approximate inverse, and so $\psi_* \approx \varphi_*^\dagger$.

4.7.2 Computing the Adjoint

In order to compute the adjoint of the differential of the embedding, $\alpha_*^\dagger$, we need the metric g on X and the Euclidean metric δ on $\mathbb{R}^n$. Since g is the pullback of δ by α, it is defined in terms of α_*.

Theorem 4.7.1. Let (X, g) and $(\hat{X}, h)$ be Riemannian manifolds and $\phi : X \to \hat{X}$ be a local isometry, i.e. g is the pullback of h by ϕ, $g = \phi^* h$. Then the adjoint of the differential $\phi_*^\dagger$ is given by

$$[\phi_*^\dagger] = \left([\phi_*]^{\mathrm{T}}[h][\phi_*]\right)^{-1} [\phi_*]^{\mathrm{T}}[h], \tag{4.47}$$

where $[\phi_*]$, $[g]$ and $[h]$ denote the matrix descriptions of the tensors at each point in arbitrary frames.

Proof. The pullback metric is

$$[g] = [\phi_*]^{\mathrm{T}}[h][\phi_*], \tag{4.48}$$

and the definition of the adjoint is

$$[\phi_*^\dagger] = [g]^{-1}[\phi_*]^{\mathrm{T}}[h]. \tag{4.49}$$

Therefore

$$[\phi_*^\dagger] = \left([\phi_*]^{\mathrm{T}}[h][\phi_*]\right)^{-1}[\phi_*]^{\mathrm{T}}[h]. \tag{4.50}$$

□

Corollary 4.7.2. Let (X, g) and $(\hat{X}, h)$ be Riemannian manifolds and $\phi : X \to \hat{X}$ be a local isometry, i.e. g is the pullback of h by ϕ, $g = \phi^* h$. Then the adjoint of the differential $\phi_*^\dagger$ in an orthonormal frame for $\mathrm{T}\hat{X}$ and an arbitrary frame for $\mathrm{T}X$ is given by

$$[\phi_*^\dagger] = \left([\phi_*]^{\mathrm{T}}[\phi_*]\right)^{-1}[\phi_*]^{\mathrm{T}}. \tag{4.51}$$

Corollary 4.7.3. Let (X, g) and $(\hat{X}, h)$ be Riemannian manifolds and $\phi : X \to \hat{X}$ be a local isometry, i.e. g is the pullback of h by ϕ, $g = \phi^* h$. Then the adjoint of the differential $\phi_*^\dagger$ in orthonormal frames is given by

$$[\phi_*^\dagger] = [\phi_*]^{\mathrm{T}}. \tag{4.52}$$

4.8 Reduced Vector Field by Optimization

Given a dimension-reducing map, φ, we need to obtain a reduced vector field, $\hat{V}$, that has the desired properties discussed in §4.3 and §4.4. In order to determine the reduced vector field for a general nonlinear original vector field, we can use the optimization approach taken in [4], and apply it to general state space manifolds. This approach uses a finite sample of snapshots from the attractor, $\mathcal{X} \subset \mathcal{M}$, and matches tangent vectors and derivatives at each point with those in the reduced space by minimizing a cost function. The cost function $S : \hat{U} \to \mathbb{R}$ is defined on a parameter space, $\hat{U}$, that parameterizes a family of reduced vector fields. As we are dealing with an attractor, a set of snapshots can be easily obtained by numerically flowing along the vector field.

The cost function is constructed using the least-squares residuals from the vector field and its derivatives. Let S be given by the weighted sum

$$S(\beta) := \omega_1 S^{(1)}(\beta) + \omega_2 S^{(2)}(\beta), \tag{4.53}$$

where

$$\begin{aligned} S^{(1)}(\beta) &:= \sum_{x\in\mathcal{X}} \left\| \varphi_* V - \hat{V}(\beta) \right\|^2 \\ S^{(2)}(\beta) &:= \sum_{x\in\mathcal{X}} \left\| \hat{\nabla}(\varphi_* V)A - k(I-A) - \hat{\nabla}\hat{V}(\beta) \right\|^2 . \end{aligned} \tag{4.54}$$

The norms are those induced by the metric. We can use the weights ω_1 and ω_2 to divide out any scale factors,

$$\begin{aligned} \omega_1^{-1} &:= \sum_{x\in\mathcal{X}} \|\varphi_* V\|^2 , \\ \omega_2^{-1} &:= \sum_{x\in\mathcal{X}} \left\| \hat{\nabla}(\varphi_* V)A - k(I-A) \right\|^2 . \end{aligned} \tag{4.55}$$

If $\hat{V}_{\hat{x}}$ is linearly dependent on $\beta \in \hat{U}$, then this optimization problem is linear, i.e. $\operatorname{grad} S = 0$ is a linear equation for β. In this case, let

$$\begin{aligned} \hat{V}^i &= B^i{}_c \beta^c \\ \hat{\nabla}_a \hat{V}^i &= C^i{}_{ac} \beta^c , \end{aligned} \tag{4.56}$$

then the optimal $\beta \in \hat{U}$ is given by the solution to the linear problem $\mathcal{A}\beta = \mathcal{B}$, where

$$\begin{aligned} \mathcal{A}_{kc} &:= \sum_{x\in\mathcal{X}} \left(\omega_1 \hat{g}_{ij} B^j{}_k B^i{}_c + \omega_2 \hat{g}_{ij} \hat{g}^{ab} C^j{}_{bk} C^i{}_{ac} \right) \\ \mathcal{B}_k &:= \sum_{x\in\mathcal{X}} \left(\omega_1 \hat{g}_{ij} (\varphi_* V)^i B^j{}_k + \omega_2 \hat{g}_{ij} \hat{g}^{ab} C^j{}_{bk} \left(\hat{\nabla}(\varphi_* V)A - k(I-A) \right)^i{}_a \right) . \end{aligned} \tag{4.57}$$

4.8.1 Radial Basis Functions

The particular family of reduced vector fields, $\hat{V} : \hat{U} \to \mathfrak{X}(\hat{X})$, used in [4] was constructed from radial basis functions, which can provide a good fit to a wide range of examples. This allows us describe the reduced dynamics without having prior knowledge of its form. Let the reduced parameter space be $\hat{U} = \mathbb{R}^{d(d+J)}$, i.e. a parameter $\beta \in \hat{U}$ consists of a $d \times d$ matrix, L, and J weights, $w_\gamma \in \mathbb{R}^d$, corresponding to centres $c_\gamma \in \mathbb{R}^d$, $1 \le \gamma \le J$. The centres are considered to be fixed *a priori*. The generic vector field on $\hat{X}$ is then expressed as a linear term with a sum of radial basis functions, each with a weight and a centre,

$$\hat{V}_{\hat{x}}(\beta) = L\hat{x} + \sum_{\gamma=1}^{J} w_\gamma \phi(\|\hat{x} - c_\gamma\|). \tag{4.58}$$

The function $\phi : \mathbb{R} \to \mathbb{R}$ can be any sensible radial basis function; we use $\phi(r) = r^2 \log r$ by default. As the centres are fixed, $\beta \mapsto \hat{V}_{\hat{x}}(\beta)$ is a linear map. For practical application it is convenient (and better numerically conditioned) to use block matrices, rather than the tensor notation of (4.57). The vector field can be written as

$$\hat{V}_{\hat{x}} = \mathcal{Z} \begin{bmatrix} \hat{x} \\ \Phi(\hat{x}) \end{bmatrix}, \tag{4.59}$$

where

$$\mathcal{Z} := \begin{bmatrix} L & w_1 & \cdots & w_J \end{bmatrix}, \qquad \Phi(\hat{x}) := \begin{bmatrix} \phi(\|\hat{x} - c_1\|) \\ \vdots \\ \phi(\|\hat{x} - c_J\|) \end{bmatrix}.$$

Note that $\mathcal{Z}$ is the block matrix description of $\beta \in \hat{U}$. By further forming the block matrices

$$\mathcal{Y}_1 := \varphi_* \begin{bmatrix} V_{x_1} & \cdots & V_{x_N} \end{bmatrix}, \qquad \mathcal{C}_1 := \begin{bmatrix} \hat{x}_1 & \cdots & \hat{x}_N \\ \Phi(\hat{x}_1) & \cdots & \Phi(\hat{x}_N) \end{bmatrix},$$

$$\mathcal{Y}_2 := \begin{bmatrix} \xi(x_1) & \cdots & \xi(x_N) \end{bmatrix}, \qquad \mathcal{C}_2 := \begin{bmatrix} & I_{d,Nd} & \\ \Psi(\hat{x}_1) & \cdots & \Psi(\hat{x}_N) \end{bmatrix},$$

where

$$\begin{aligned} \Psi_{jb}(\hat{x}) &:= \phi'(\|\hat{x} - c_j\|) \frac{(\hat{x} - c_j)^b}{\|\hat{x} - c_j\|}, \\ \xi_{ij}(x) &:= (\hat{\nabla}_k Y)^i A^k_{\ j} - k(\delta^i_{\ j} - A^i_{\ j}), \end{aligned} \tag{4.60}$$

we can write the linear problem as

$$\left(\sum_{h=1}^{2} \omega_h \mathcal{C}_h \mathcal{C}_h^{\mathrm{T}} \right) \mathcal{Z}^{\mathrm{T}} = \sum_{h=1}^{2} \omega_h \mathcal{C}_h \mathcal{Y}_h^{\mathrm{T}}, \tag{4.61}$$

for $\mathcal{Z}$. These matrices have sizes $\mathcal{Y}_1 \in \mathbb{R}^{d \times N}$, $\mathcal{C}_1 \in \mathbb{R}^{(d+J) \times N}$, $\mathcal{Y}_2 \in \mathbb{R}^{d \times Nd}$, $\mathcal{C}_2 \in \mathbb{R}^{(d+J) \times Nd}$, $\mathcal{Z} \in \mathbb{R}^{d \times (d+J)}$.

Placing Radial Basis Centres

Before the problem for the reduced dynamics can be solved, the centres $(c_1, \cdots, c_J)$ for the radial basis functions need to be placed in $\mathbb{R}^d$. We adopt a brute force strategy by generating random configurations. To do this we compute an axis-aligned bounding

box that contains all attractors in the parameter region. This can be obtained by iterating over all data points and recording the minimum and maximum coordinates on each axis. Given a bounding box, we select the position of each centre from a uniform distribution over the box. Since this takes place in the low-dimensional space it is typically cheaper to generate multiple configurations, solve for each, then choose the best configuration than it is to use nonlinear optimization on the combined problem.

On a practical note, we found that scaling up the size of the bounding box to around 1.5 times the default data box tended to result in lower cost solutions.

4.9 Reduction with Symmetry

In the presence of a symmetry group on the original system, we can ask whether the symmetry can be preserved in the reduced system. Let G be a symmetry group for the original system (X, V), i.e. the vector field is G-equivariant. Let $\mathcal{M}$ be an attractor that is also G-invariant, i.e. $g(\mathcal{M}) = \mathcal{M}$ for all $g \in G$. If we give G an action on the reduced state space $\hat{X}$ and require $\varphi : X \to \hat{X}$ to be G-equivariant then $\varphi(\mathcal{M})$ is G-invariant and the pushforward $\varphi_* V$ along $\varphi(\mathcal{M})$ is a G-equivariant vector field. Therefore in order to preserve the symmetry, the task is to determine an action of G on $\hat{X}$ such that φ is G-equivariant.

With a secant-based method of reduction, our dimension-reducing map consists of an embedding and an orthogonal projection, $\varphi = W^\dagger \circ \alpha$. Since we are free to choose the embedding α, and free to extend the symmetry to the embedding space, we focus on the orthogonal projection and finding actions that make $W^\dagger$ G-equivariant, $W^\dagger g = g W^\dagger \ \forall g \in G$.

Theorem 4.9.1. Let X and $\hat{X}$ be inner product spaces and $W : \hat{X} \to X$ be an orthonormal map. Given a (possibly nonlinear) action of G on $\hat{X}$, the corresponding action on X,

$$gx = Wg(W^\dagger x) + (\mathrm{id}_X - WW^\dagger)x \tag{4.62}$$

is a well-defined action under which W and $W^\dagger$ are G-equivariant.

Proof. $ex = x$ is straightforward and $h(gx)$ is given by

$$\begin{aligned} h(gx) &= Wh(g(W^\dagger x)) + (\mathrm{id}_X - WW^\dagger)(Wg(W^\dagger x) + (\mathrm{id}_X - WW^\dagger)x) \\ &= Wh(g(W^\dagger x)) + (\mathrm{id}_X - WW^\dagger)x \\ &= W(hg)(W^\dagger x) + (\mathrm{id}_X - WW^\dagger)x \\ &= (hg)x. \end{aligned} \tag{4.63}$$

Hence this is a well-defined action on X. Under these actions,

$$\begin{aligned} W^\dagger gx &= W^\dagger W g(W^\dagger x) + W^\dagger(\mathrm{id}_X - WW^\dagger)x \\ &= g(W^\dagger x) \end{aligned} \tag{4.64}$$

and

$$\begin{aligned} g(W\hat{x}) &= Wg(W^\dagger W\hat{x}) + (\mathrm{id}_X - WW^\dagger)W\hat{x} \\ &= Wg\hat{x}, \end{aligned} \tag{4.65}$$

so both W and $W^\dagger$ are G-equivariant. □

This gives a way of generating an action on X given an action on $\hat{X}$. Since W is G-equivariant, we can also relate these actions via $g\hat{x} = W^\dagger g(W\hat{x})$. For example, translation in $\hat{X}$, $\hat{x} \mapsto \hat{x}+a$, corresponds to translation in X by $x \mapsto x+Wa$. However, what we really want is a way of generating an action on $\hat{X}$ given one on X.

Theorem 4.9.2. Let X and $\hat{X}$ be inner product spaces and let G have a (possibly nonlinear) action on X. Let $W : \hat{X} \to X$ be an orthonormal map such that $W(\hat{X})$ is a G-invariant subspace of X. The corresponding action on $\hat{X}$,

$$g\hat{x} = W^\dagger g(W\hat{x}) \tag{4.66}$$

is a well-defined action under which W is G-equivariant.

Proof. Since $\operatorname{im} W$ is G-invariant, $Pg(W\hat{x}) = g(W\hat{x})$, where $P = WW^\dagger$ is the orthogonal projection onto $\operatorname{im} W$. $ex = x$ is straightforward and $h(gx)$ is given by

$$\begin{aligned} h(g\hat{x}) &= W^\dagger h(WW^\dagger g(W\hat{x})) \\ &= W^\dagger h(g(W\hat{x})) \\ &= W^\dagger(hg)(W\hat{x}) \\ &= (hg)\hat{x} \end{aligned} \tag{4.67}$$

Hence this is a well-defined action on $\hat{X}$. Under these actions,

$$\begin{aligned} Wg\hat{x} &= WW^\dagger g(W\hat{x}) \\ &= g(W\hat{x}) \end{aligned} \tag{4.68}$$

W is G-equivariant. □

Theorem 4.9.3. Let X and $\hat{X}$ be real inner product spaces and let G act orthogonally on X. Let $W : \hat{X} \to X$ be an orthonormal map such that $W(\hat{X})$ is a G-invariant subspace of X. The corresponding action on $\hat{X}$,

$$g\hat{x} = W^\dagger g(W\hat{x}) \tag{4.69}$$

is a well-defined action under which W, $W^\dagger$ and $P = WW^\dagger$ are G-equivariant.

Proof. By Theorem 4.9.2, the action is well-defined and W is G-equivariant. Since G acts orthogonally, $(\operatorname{im} W)^\perp$ is also an invariant subspace, which by Theorem A.1.1 coincides with $\ker W^\dagger = \ker P$. Hence

$$\begin{aligned} Pgx &= PgPx + Pg(\mathrm{id}_X - P)x \\ &= PgPx \\ &= gPx, \end{aligned} \tag{4.70}$$

P is G-equivariant. Also,

$$\begin{aligned} g(Px) &= Pg(x) \\ Wg(W^\dagger x) &= WW^\dagger g(x) \\ g(W^\dagger x) &= W^\dagger g(x), \end{aligned} \tag{4.71}$$

$W^\dagger$ is G-equivariant. □

Therefore, we have a method of preserving the symmetry under the conditions that the group acts orthogonally, and the reduced state space is chosen to be an invariant subspace of the group action.

The simplest nontrivial example of an orthogonal symmetry is $\mathbb{Z}_2$ acting by reflection. For this group action, every subspace is an invariant subspace, and the resulting action on the reduced space is also reflection.

Another example is a group of rotations with a common axis of rotation. Note that the 'axis' may have more than one dimension, depending on the dimension of the state

space. Since both the axis and its orthogonal complement are invariant subspaces, we can choose to project onto the orthogonal complement. In this case the resulting action on the reduced space is rotation in the subspace.

Chapter 5

Dimensionality Reduction with Parameter Dependence

In this chapter we develop a method of dimensionality reduction that is able to reproduce a family of attractors associated to a parameter space. We extend the formalism developed in the previous chapter to include a parameter space, and adapt secant-Grassmann projection for use with multiple attractors. We develop a practical implementation by extending the optimization approach from §4.8 by constructing a linear optimization problem over a space of affine maps.

5.1 Introduction

The parameters of a model are often a set of individual real numbers, $(\lambda_i \in \mathbb{R})_{i=1}^{l}$. For each of these, one may have an interval of $\mathbb{R}$ which is of interest, or of physical relevance, $[a_i, b_i]$. The total parameter space under investigation is then given by the product of these intervals, $U = [a_1, b_1] \times \cdots \times [a_l, b_l]$, which is a convex subset of $\mathbb{R}^l$. We can therefore think of U as a smooth manifold and also make use of linear algebra on $\mathbb{R}^l$. In the presence of a parameter space, we can regard the system as a family of vector fields.

5.2 Formulation

If we have a parameter space then, rather than a single vector field, we have a parameterized family of vector fields. As the attractor results from the flow of the vector field, we can have a corresponding family of attractors. Even though the resulting attractor can change in a wide variety of complicated ways, especially over bifurcations, the tangent vectors often depend smoothly on the parameter. This is something we can take advantage of, by reproducing the relevant parts of the underlying vector fields and allowing the attractors to emerge naturally, rather than attempting to describe them explicitly.

Let X be the state space as a smooth manifold and U be the parameter space. Let a smooth family of smooth vector fields be given by the map $V : U \to \mathfrak{X}(X)$. We will use the notation $V^\lambda = V(\lambda)$. We are interested in a given original system (X, U, V) such that for each parameter value $\lambda \in U$, the flow of V^λ has an attractor $\mathcal{M}_\lambda \subset X$. The objective is to construct a reduced system $(\hat{X}, \hat{U}, \hat{V})$. In general, the state-space map that relates X to $\hat{X}$ may depend on the parameter. Let $\varphi : X \times U \to \hat{X}$ be a smooth map and let $\varphi^\lambda : X \to \hat{X} : x \mapsto \varphi(x, \lambda)$. We also need to relate the original parameters to the parameters of the reduced system. Let the parameter map $Q : U \to \hat{U}$ be a smooth map. The objective is to obtain a φ and Q such that each attractor's vector field and derivatives are reproduced by $\hat{V}(Q(\lambda))$ in $\hat{X}$ under φ^λ, as in §4.

5.2.1 Product Manifold

Consider the product manifold $X \times U$. Let $i^\lambda : X \to X \times U : x \mapsto (x, \lambda)$ and $j_x : U \to X \times U : \lambda \mapsto (x, \lambda)$ be inclusions, and $\pi_X : X \times U \to X$ and $\pi_U : X \times U \to U$ be the canonical projections. The tangent space of a product manifold is the direct sum of the component spaces, $\mathrm{T}_{(x,\lambda)}(X \times U) = \mathrm{T}_x X \oplus \mathrm{T}_\lambda U$. We refer to the X component as horizontal, and the U component as vertical. Each vector field on the product manifold can be written as a sum of a horizontal and vertical vector field.

The smooth family of vector fields on X can be expressed as a single smooth vector field on the product manifold, $\mathcal{V} \in \mathfrak{X}(X \times U)$, via

$$\mathcal{V}_{(x,\lambda)} = (i^\lambda_*)_x V^\lambda_x. \tag{5.1}$$

This vector field is horizontal,

$$((\pi_X)_*)_{(x,\lambda)}\mathcal{V}_{(x,\lambda)} = V_x^\lambda \tag{5.2}$$

$$(\pi_U)_*\mathcal{V} = 0. \tag{5.3}$$

The reduced system can also be formulated as a product manifold, $\hat{X} \times \hat{U}$, with a horizontal vector field, $\hat{\mathcal{V}} \in \mathfrak{X}(\hat{X} \times \hat{U})$ satisfying $(\pi_{\hat{U}})_*\hat{\mathcal{V}} = 0$. In this formulation, the reduction consists of a smooth map $\xi : X \times U \to \hat{X} \times \hat{U}$ that preserves horizontality. If ξ is of the form $\xi(x, \lambda) = (\varphi(x, \lambda), Q(\lambda))$ then ξ preserves horizontality, since $\partial Q/\partial x = 0$,

$$\begin{aligned}
[(\xi_*)_{(x,\lambda)}\mathcal{V}_{(x,\lambda)}] &= \begin{bmatrix} \partial\varphi/\partial x & \partial\varphi/\partial\lambda \\ \partial Q/\partial x & \partial Q/\partial\lambda \end{bmatrix} \begin{bmatrix} V_x^\lambda \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} (\varphi_*^\lambda)_x & \partial\varphi/\partial\lambda \\ 0 & (Q_*)_\lambda \end{bmatrix} \begin{bmatrix} V_x^\lambda \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} (\varphi_*^\lambda)_x V_x^\lambda \\ 0 \end{bmatrix}.
\end{aligned} \tag{5.4}$$

5.2.2 Family of Attractors

For each $\lambda \in U$ let $\mathcal{M}_\lambda \subset X$ be a compact submanifold without boundary invariant under the flow of V^λ. Note that we are not assuming that the family $(\mathcal{M}_\lambda)_{\lambda\in U}$ is a submanifold of $X \times U$, since the dimension of the attractor may change with λ, producing singular points; rather each horizontal slice of the product manifold contains an attractor.

As in §4.3, we need to embed each individual attractor into the reduced state space, i.e. φ^λ is an embedding of $\mathcal{M}_\lambda$ into $\hat{X}$ for each $\lambda \in U$. The parameter map, Q, allows for a given parameter λ to determine a vector field on the reduced state space, $\hat{V}(Q(\lambda))$. Given a parameter-dependent dimension-reducing map, φ, that embeds the family of attractors into the reduced space, we seek a parameter map such that for each parameter, the reduced vector field has the properties discussed in §4, i.e. it reproduces the flow on the attractor, and has derivatives that are stable.

$$\hat{V}(Q(\lambda)) = \varphi_*^\lambda V(\lambda) \quad \forall x \in \mathcal{M}_\lambda \ \forall\lambda \in U \tag{5.5}$$

In §4.4 we need to choose an inverse within a neighbourhood of the attractor. Let $\mathcal{N}_\lambda$ be a neighbourhood of $\varphi^\lambda(\mathcal{M}_\lambda)$ and $\mathcal{N} = \cup_{\lambda\in U}\mathcal{N}_\lambda$. Let $\psi : \mathcal{N} \times U \to X$ be a smooth

map that coincides with the inverse of φ^λ on each attractor, i.e.

$$\psi^\lambda(\varphi^\lambda(x)) = x \quad \forall x \in \mathcal{M}_\lambda \; \forall \lambda \in U. \tag{5.6}$$

and $\psi^\lambda : \mathcal{N} \to X : \hat{x} \mapsto \psi(\hat{x}, \lambda)$. Then $Y_p^\lambda = \varphi_*^\lambda V_{\psi(p,\lambda)}^\lambda$ is a smooth family of vector fields defined on $\mathcal{N}$, which can be used to specify the derivatives of the reduced vector fields along each submanifold.

5.2.3 Parameter-Independent Reduction

In order to simplify the problem, we can seek a single state-space map that is suitable for the whole parameter region, i.e. φ is independent of λ. In this case we write $\varphi : X \to \hat{X}$ and $Q : U \to \hat{U}$.

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & \hat{X} \\ U & \xrightarrow{Q} & \hat{U} \\ V \big\downarrow & & \big\downarrow \hat{V} \\ \mathfrak{X}(X) & & \mathfrak{X}(\hat{X}) \end{array}$$

Doing so allows us to decouple the problem into two parts; first, finding a φ that satisfies the necessary topological requirements, and then using φ to find a Q that produces the corresponding dynamics in the reduced space. Even though φ does not depend on the parameter, the inverse of $\varphi|_{\mathcal{M}_\lambda}$ can indeed depend on the parameter, and therefore ψ retains a parameter dependence.

5.3 Extending Secant-Grassmann Projection

We can extend the secant-Grassmann projection method discussed in §3 and §4.7 to a family of submanifolds indexed by a parameter space, U, in a straightforward manner. Let $\alpha : X \to \mathbb{R}^n$ be an embedding of the state space into an ambient Euclidean space. Each submanifold, $\alpha(\mathcal{M}_\lambda)$, generates a set of unit secants, $\mathcal{K}(\alpha(\mathcal{M}_\lambda))$. For a finite sample of these secants, $\mathcal{K}_\lambda \subset \mathcal{K}(\alpha(\mathcal{M}_\lambda))$, we have the cost function (3.7),

$$\mathcal{F}_\lambda(W) = \frac{1}{|\mathcal{K}_\lambda|} \sum_{k \in \mathcal{K}_\lambda} \left\| W^\dagger k \right\|^{-1}, \tag{5.7}$$

where $W : \mathbb{R}^d \to \mathbb{R}^n$ is an orthonormal map. We can construct a single cost function using a finite collection of parameter values, $\mathcal{U} \subset U$, via

$$\mathcal{G}(W) = \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} \mathcal{F}_\lambda(W). \tag{5.8}$$

This is a well-defined cost function on the Grassmannian as it satisfies the orthogonal symmetry $\mathcal{G}(W\mathcal{Q}) = \mathcal{G}(W)$. Like the individual cost functions, this takes values in the range $[1, \infty]$. For finite values, all sampled secants on all sampled submanifolds are preserved. Therefore minimizing $\mathcal{G}$ obtains a good projection for the family of submanifolds in the region of parameter space described by $\mathcal{U}$. Note that we do not include secants generated by pairs of points with differing parameter values. This is because two attractors from different parameter values are allowed to intersect both in the original space and in the reduced space. The minimum projected length becomes

$$\kappa_{\min} = \min_{\lambda \in \mathcal{U}} \min_{k \in \mathcal{K}_\lambda} \left\| W^\dagger k \right\|. \tag{5.9}$$

5.3.1 Inverse of the State-Space Map

In order to determine the reduced dynamics, we need to make use of an inverse of the state-space map for each attractor. By using a parameter-independent state-space map, consisting of an embedding and an orthogonal projection, the inverse has the general form

$$\psi(\hat{x}, \lambda) = \alpha^L(W\hat{x} + \eta(\hat{x}, \lambda)), \tag{5.10}$$

where $\eta : \mathbb{R}^d \times U \to \mathbb{R}^n$ is a smooth map that gives the vertical component, $W^\dagger \eta(\hat{x}, \lambda) = 0$. Note that since the embedding α and the orthonormal map W are parameter-independent, the parameter dependence of ψ is confined to the vertical component given by η. We will also use the notation $\eta^\lambda(\hat{x}) = \eta(\hat{x}, \lambda)$.

As discussed in §4.4.3 and §4.8, we in fact only require the derivative of the inverse (in (4.9)), which is

$$\begin{aligned} \psi_*^\lambda &= \alpha_*^\dagger W + \alpha_*^\dagger \eta_*^\lambda \\ &= \varphi_*^\dagger + \alpha_*^\dagger \eta_*^\lambda. \end{aligned} \tag{5.11}$$

In practice we will use an approximation of the true inverse by choosing a function η to fit the available data. For a given α and W, the vertical component map η should

(approximately) satisfy

$$\eta(W^\dagger \alpha(x), \lambda) = (\mathrm{id} - WW^\dagger)\alpha(x) \quad \forall x \in \mathcal{M}_\lambda. \tag{5.12}$$

For each λ, this can be phrased as an optimization problem with least-squares cost

$$\sum_{y \in \alpha(\mathcal{X}_\lambda)} \left\| \eta^\lambda(W^\dagger y) - (\mathrm{id} - WW^\dagger)y \right\|^2 \tag{5.13}$$

for the map η^λ. Combining the optimization problems over the parameter space gives the single cost function

$$\frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} \frac{1}{|\mathcal{X}_\lambda|} \sum_{y \in \alpha(\mathcal{X}_\lambda)} \left\| \eta(W^\dagger y, \lambda) - (\mathrm{id} - WW^\dagger)y \right\|^2 \tag{5.14}$$

for the function η.

An Affine Vertical Component

In order to obtain an inverse whose derivative is parameter-independent, we can look for an η that is affine in the parameter λ and independent of $\hat{x}$,

$$\eta(\hat{x}, \lambda) = B(\lambda - \bar{\lambda}) + c, \tag{5.15}$$

where $c \in \mathbb{R}^n$, $B \in L(\mathbb{R}^l, \mathbb{R}^n)$, and $\bar{\lambda}$ is the mean parameter value over the parameter samples,

$$\bar{\lambda} := \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} \lambda. \tag{5.16}$$

The derivative of this is clearly null, $\eta^\lambda_* = 0$, and therefore the corresponding inverse does indeed have a parameter-independent derivative, $\psi^\lambda_* = \varphi^\dagger_*$. By writing the affine map in this form, we can determine the optimal B and c separately, as explained in §2.4.7. Assuming c is given, the optimal B is given by the linear problem

$$B^i{}_j \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (\lambda - \bar{\lambda})^j (\lambda - \bar{\lambda})^b = \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (\bar{z}_\lambda)^i (\lambda - \bar{\lambda})^b, \tag{5.17}$$

which is independent of the choice of c, where

$$\bar{z}_\lambda := \frac{1}{|\mathcal{X}_\lambda|} \sum_{y \in \alpha(\mathcal{X}_\lambda)} (\mathrm{id} - WW^\dagger)y \tag{5.18}$$

is the mean vertical component over the data set for parameter λ.

Similarly, assuming B is given and c is to be found, the solution to the least-squares problem for c is

$$c = \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} \bar{z}_\lambda, \tag{5.19}$$

which is independent of B. The inverse is then given in closed form for both $\hat{x}$ and λ as

$$\psi(\hat{x}, \lambda) = \alpha^L(W\hat{x} + B(\lambda - \bar{\lambda}) + c). \tag{5.20}$$

We can also show that $\eta(\hat{x}, \lambda)$ is orthogonal to the subspace $W(\mathbb{R}^d)$.

Proposition 5.3.1. $\eta(\hat{x}, \lambda)$ is orthogonal to the subspace $W(\mathbb{R}^d)$ for all λ.

Proof. By Theorem A.1.1, the orthogonal complement of the subspace $W(\mathbb{R}^d)$ is given by the kernel of $W^\dagger$. Therefore we need to verify $W^\dagger \eta(\hat{x}, \lambda) = 0$.

$$\begin{aligned} W^\dagger c &= \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} W^\dagger \bar{z}_\lambda \\ &= 0 \end{aligned} \tag{5.21}$$

Applying $W^\dagger$ to (5.17) gives

$$\begin{aligned} (W^\dagger)^a{}_i B^i{}_j \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (\lambda - \bar{\lambda})^j (\lambda - \bar{\lambda})^b &= (W^\dagger)^a{}_i \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (\bar{z}_\lambda)^i (\lambda - \bar{\lambda})^b, \\ &= \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (W^\dagger \bar{z}_\lambda)^a (\lambda - \bar{\lambda})^b, \\ &= 0. \end{aligned} \tag{5.22}$$

Since B is defined by the linear problem (5.17), $\frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} (\lambda - \bar{\lambda})^j (\lambda - \bar{\lambda})^b$ has a right-inverse, and so $W^\dagger B = 0$. Hence $W^\dagger \eta^\lambda = 0$. □

5.4 Special Case: Multilinear Series

In §4.6 we considered a special case where both state spaces are vector spaces, the reduction is linear, and the vector field is given by a multilinear series. In this case a Galerkin approach can be used to determine the reduced dynamics. We now extend this to include parameter dependence.

Let X and $\hat{X}$ be vector spaces and $\varphi : X \to \hat{X}$ be a surjective linear map. Let $\psi(\hat{x}, \lambda)$ be linear in both $\hat{x}$ and λ. Let the original vector field be given by a multilinear

series for state and parameter dependence,

$$V_x^\lambda = \sum_{pq} M_{pq}(\underbrace{x, \cdots, x}_{p \text{ times}}; \underbrace{\lambda, \cdots, \lambda}_{q \text{ times}}). \tag{5.23}$$

where $M_{pq} : X^p \times U^q \to \hat{X}$ is p-linear in X and q-linear in U. Then we can choose $\hat{U} = U$ and $Q = \mathrm{id}_U$, and apply (4.2),

$$\begin{aligned} \hat{V}_{\hat{x}}^\lambda &= \varphi_* \, V_{\psi^\lambda(\hat{x})}^\lambda \\ &= \varphi_* \sum_{pq} M_{pq}(\psi^\lambda(\hat{x}), \cdots, \psi^\lambda(\hat{x}); \lambda, \cdots, \lambda) \\ &= \sum_{pq} \hat{M}_{p,p+q}(\hat{x}, \cdots, \hat{x}; \lambda, \cdots, \lambda). \end{aligned} \tag{5.24}$$

Since $\psi(\hat{x}, \lambda)$ is multilinear,

$$\psi(\hat{x}, \lambda) = \psi^j{}_{\nu b} \hat{x}^\nu \lambda^b, \tag{5.25}$$

$\hat{M}$ is given by

$$(\hat{M}_{p,p+q})^\mu{}_{\nu_1 \cdots \nu_p \;\; b_1 \cdots b_p a_1 \cdots a_q} = \varphi^\mu_{\;\; i} \; (M_{pq})^i{}_{j_1 \cdots j_p \;\; a_1 \cdots a_q} \; \psi^{j_1}{}_{\nu_1 b_1} \cdots \psi^{j_p}{}_{\nu_p b_p}. \tag{5.26}$$

The error in the reduced vector field is given by

$$\begin{aligned} &\left\| \hat{V}^\lambda_{\varphi(x)} - \varphi_* V^\lambda_x \right\| \\ &= \left\| \varphi_* \sum_{pq} \left[M_{pq}(\psi^\lambda(\varphi(x)), \cdots, \psi^\lambda(\varphi(x)); \lambda, \cdots, \lambda) - M_{pq}(x, \cdots, x; \lambda, \cdots, \lambda) \right] \right\|. \end{aligned} \tag{5.27}$$

Note that the error vanishes when $\psi^\lambda(\varphi(x)) = x$, and therefore for a given φ, the error is determined by the quality of the inverse ψ.

5.5 Obtaining a Parameter Map by Optimization

In §4.8 we used optimization to obtain a single reduced vector field, which we can think of as corresponding to a fixed parameter in U. We now extend this approach to the parameter map, where we are interested in obtaining the map $Q : U \to \hat{U}$ from the original parameter that determines the original vector field to the reduced parameter that determines the reduced vector field.

In order to find a suitable mapping we can consider a space of such maps, and select the optimal map by extending the optimization approach from §4.8. Let $\mu \in \mathbb{R}^p$

parameterize a family of maps $Q_\mu : U \to \hat{U}$. We can construct an optimization problem for μ using a finite sample of original parameter values, $\mathcal{U} \subset U$, to extend the single-parameter case. For each $\lambda \in \mathcal{U}$ we want the original vector field to match the reduced vector field, $\hat{V}(Q_\mu(\lambda))$, in the sense described above. To this end we construct the cost function $T : \mathbb{R}^p \to \mathbb{R}$,

$$T(\mu) := \frac{1}{|\mathcal{U}|} \sum_{\lambda \in \mathcal{U}} S_\lambda(Q_\mu(\lambda)), \tag{5.28}$$

where S_λ is the single-parameter cost function for λ, (4.54). If $\mu \mapsto Q_\mu(\lambda)$ is a linear map from $\mathbb{R}^p$ to $\hat{U}$ for all $\lambda \in U$, then $\operatorname{grad} T = 0$ describes a linear problem for μ. Note that each map Q_μ does not need to be linear itself, just the parameterization. In this case, let $Q_\mu(\lambda) = D\mu$, where D is a linear map (w.r.t. μ) which may contain nonlinear dependence on λ. The optimal μ is then given by the solution to the linear problem

$$\left(\sum_{\lambda \in \mathcal{U}} D^{\mathrm{T}} \mathcal{A} D \right) \mu = \sum_{\lambda \in \mathcal{U}} D^{\mathrm{T}} \mathcal{B}, \tag{5.29}$$

where $\mathcal{A}$ and $\mathcal{B}$ are as defined in (4.57) (both depend on λ implicitly). We choose Q to be Q_μ with optimal μ.

5.5.1 An Affine Parameter Map

If the original tangent vectors have a simple dependence on the parameter space, for example an affine map, then the projection of these vectors, $\varphi_* V^\lambda$, are also affinely dependent on the parameter. For a more general parameter dependence, a first-order Taylor expansion around a particular parameter point will give an affine map that approximates the parameter dependence in a neighbourhood of the point. Therefore if the parameter region U is small enough such that an affine map is an accurate approximation, we can seek an affine parameter map $Q : U \to \hat{U}$, and a reduced family of vector fields $\hat{V}$ that is also affine in $\hat{U}$. The composition $\hat{V} \circ Q$ is then an affine map.

The space of affine maps is parameterized by $\mu \in \hat{U} \oplus L(U, \hat{U})$, i.e. $\mu = (q_0, q_1)$, where $q_0 \in \hat{U}$ and q_1 is a linear map from U to $\hat{U}$. This is clearly linear in μ,

$$Q_\mu(\lambda) = q_0 + q_1 \lambda. \tag{5.30}$$

With Radial Basis Functions

The concrete family of reduced vector fields used in §4.8.1 is linearly dependent on $\hat{U}$ and therefore compatible with an affine parameter dependence in the original system. Recall that J is the number of radial basis functions, d is the dimension of the reduced space, and l is the dimension of the parameter space. To make this compatible with the block matrix representation, we write $\mathcal{Z} = \Theta\Lambda$, where $\Theta \in \mathbb{R}^{d\times(d+J)(l+1)}$ contains the μ-dependence, and $\Lambda \in \mathbb{R}^{(d+J)(l+1)\times(d+J)}$ contains the λ-dependence,

$$\Theta = \begin{bmatrix} z_0 & z_1 & \cdots & z_l \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} I \\ \lambda^1 I \\ \vdots \\ \lambda^l I \end{bmatrix}, \tag{5.31}$$

where $l = \dim U$ and I is the $(d+J)\times(d+J)$ identity matrix. The linear problem to solve for Θ is then given by

$$\left(\sum_{\lambda\in\mathcal{U}} \Lambda \left(\sum_{h=1}^{2} \omega_h \mathcal{C}_h \mathcal{C}_h^{\mathrm{T}}\right) \Lambda^{\mathrm{T}}\right) \Theta^{\mathrm{T}} = \sum_{\lambda\in\mathcal{U}} \Lambda \left(\sum_{h=1}^{2} \omega_h \mathcal{C}_h \mathcal{Y}_h^{\mathrm{T}}\right). \tag{5.32}$$

Chapter 6

Examples

To demonstrate the method, we consider some examples. A variety of examples are used, both low-dimensional simple systems and discretizations of PDEs, which are high-dimensional. Although the method supports multiple parameters, we shall consider examples where a single parameter is varied. In each case, we consider a region of (1-dimensional) parameter space, and for a number of parameter values we generate a data set of points from the attractor numerically. Using these data sets, we apply the secant-Grassmann projection method to obtain a projection that is good for the parameter region (§5.3). This projection is then used to obtain a family of reduced vector fields, by finding an affine map between the original parameter space and reduced parameter space (§5.5.1). The family of attractors produced by the reduced model is compared with the projection of those of the original system in the low-dimensional space. We use the strategy from §4.8.1 for placing the radial basis centres.

The first example illustrates the accurate reproduction of parameter dependence in period-doubling bifurcations without any dimension reduction; the second illustrates a non-autonomous system with a state consisting of multiple angular variables; the third illustrates a high-dimensional system produced by a discretized PDE.

In the future, we plan to conduct a more comprehensive survey of applications of the method to various examples, including more complex PDEs, and examples with large numbers of angular variables that the POD-Galerkin method is not compatible with. Examples with multiple parameters can also be investigated. Such an undertaking is beyond the scope of this work, in which we have focused on developing the approach.

6.1 Rössler

The Rössler system [37] is a well-known 3-dimensional dynamical system featuring a chaotic attractor. In particular it features the classic route to chaos via period-doubling bifurcations of a limit cycle. Since the system is already of low dimension, we do not attempt to reduce its dimension, but rather use it as a test of the method's ability to the reconstruct dynamics with parameter variation. This is a particularly good example as it features not only a family of attractors with bifurcations, but the individual attractors also feature slow and fast parts. This is particularly noticeable in the z-coordinate, which rapidly spikes to large values. This will test the method's ability to reproduce vector fields with large variations over the attractor.

The system has state $(x, y, z) \in \mathbb{R}^3 = X$ and parameters $(a, b, c) \in \mathbb{R}^3$ with a vector field given by

$$V_{(x,y,z)} = \begin{bmatrix} -y - z \\ x + ay \\ b + z(x - c) \end{bmatrix} \tag{6.1}$$

We fix $a = b = 0.1$ and vary c. At $c = 4$ the system possesses a period-1 limit cycle. We consider the parameter region $U = [4, 8.8]$, which contains a number of period-doubling bifurcations (from period 1 to 8), and choose 21 equally-spaced parameter values from this interval. Since we are not reducing the dimension, $\hat{X} = X$, $\varphi = \mathrm{id}_X$, and $\psi_\lambda = \mathrm{id}_X$.

The derivative of the vector field is

$$\nabla V_{(x,y,z)} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ z & 0 & x - c \end{bmatrix}. \tag{6.2}$$

Since the φ is a submersion, we do not need to decompose the tangent space.

6.1.1 Results

All simulations were performed with an RK4 integrator with a time-step of 0.001. For reference the period-1 limit cycle at $c = 4$ has a period of approximately 6 units of time. To reproduce the dynamics, 40 radial basis centres were used. This number was determined by trial-and-error to be large enough to give a good reproduction of the

attractors over the majority of the parameter region while keeping the computational cost reasonable. We found that the radial basis function $\phi(r) = r^3$ was a superior choice for this example over the thin-plate spline.

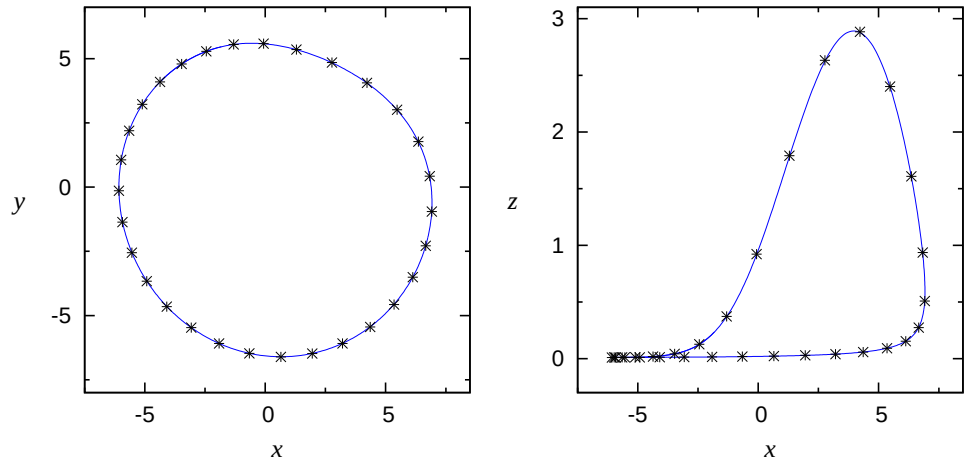
The reconstruction obtained demonstrates an accurate reproduction of both individual attractors and their period-doubling bifurcations. Fig. 6.1 compares a selection of individual attractors and Fig. 6.2 shows the two bifurcation diagrams for the parameter region.

To quantify the accuracy of the reproduction, we use several measures of error:

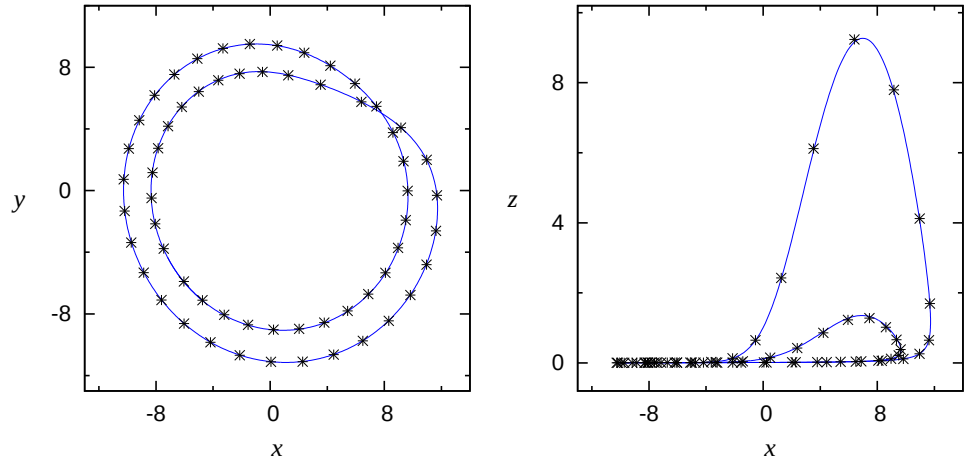
- The per-sample error,
$$E(t) = \|\hat{x}(t) - x(t)\| .$$
- The root-mean-square (RMS) error over N samples,
$$E_{\text{RMS}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|\hat{x}(t_i) - x(t_i)\|^2}.$$
- The maximum error over N samples,
$$E_{\max} = \max_i \|\hat{x}(t_i) - x(t_i)\| .$$
- The set distance – the maximum of the minima over N samples,
$$E_{\text{maxmin}} = \max_i \min_j \|\hat{x}(t_i) - x(t_j)\| .$$

As these are absolute errors it is important to take them in context with the scale of the attractors in the state space. At $c = 4$ the attractor is contained within a ball of diameter 14 units, increasing to 36 units at $c = 8.8$. Fig. 6.3 shows the time series errors for a number of parameter values over a time interval of 100 units. Fig. 6.4 shows the parameter dependence of the various error meansurements.

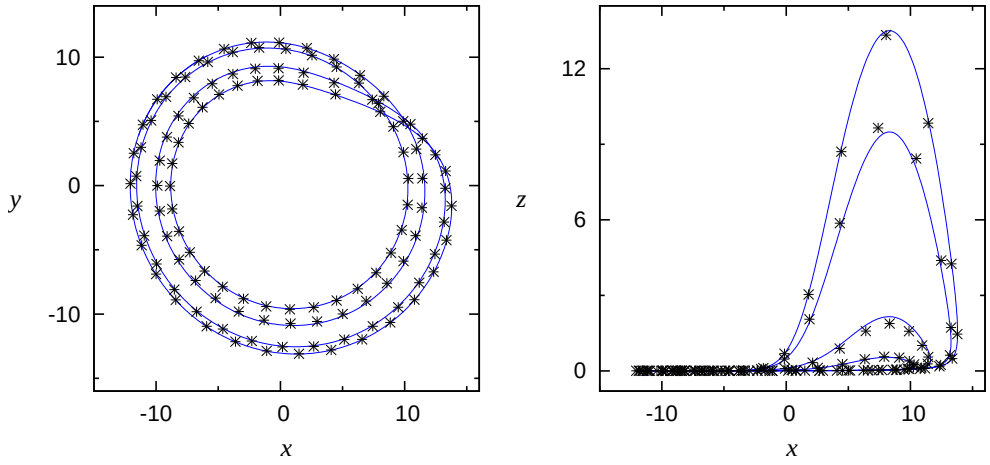
As we are not doing any dimensionality reduction in this particular case, we can also compare the Floquet multipliers of the original and reconstruction. One of the two non-unity multipliers is approximately zero, the other is real and shown in Fig. 6.5. Since the optimization explicitly tries to reproduce the derivatives of the vector field along the orbit, we expect the Floquet multipliers to be a close match to those of the original.



(a) $c = 4$

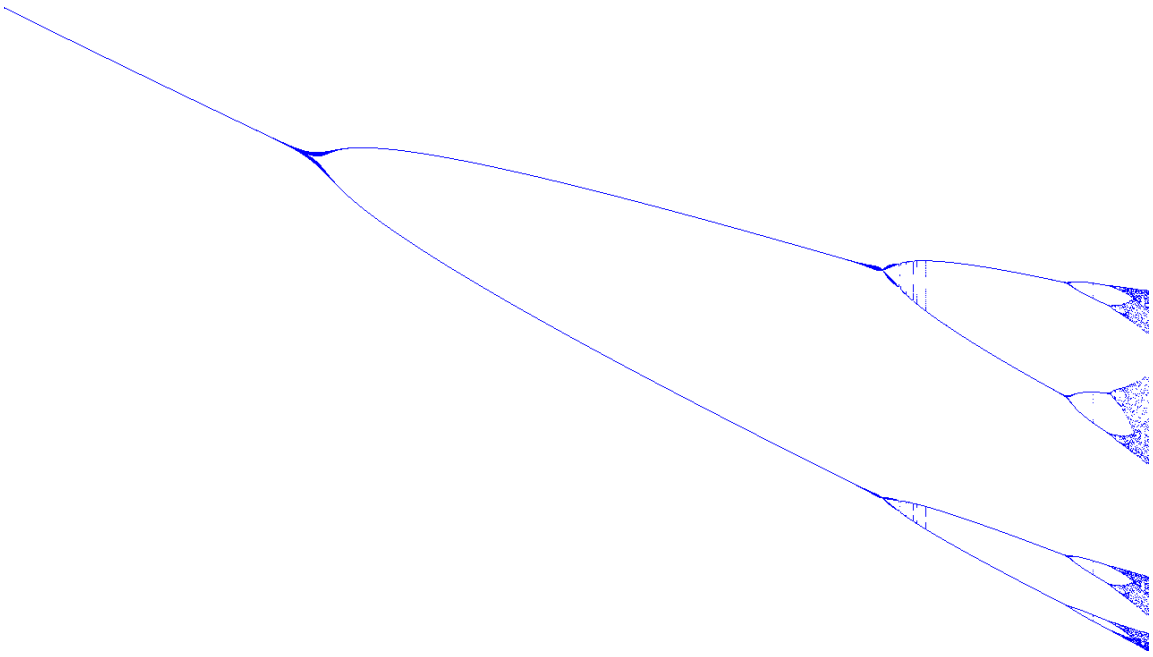


(b) $c = 7$

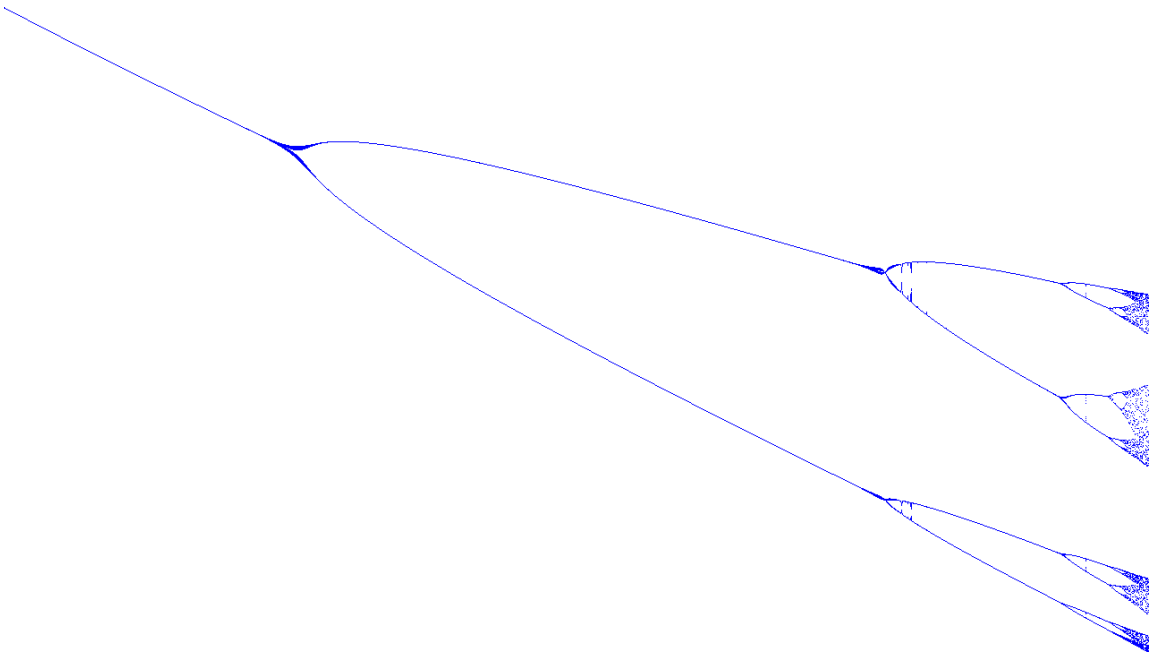


(c) $c = 8.32$

Figure 6.1: The Rössler system. Comparison of the attractors of the original model with those of the reconstruction in the full 3-dimensions. Points are from the original model and lines are from the reconstruction.

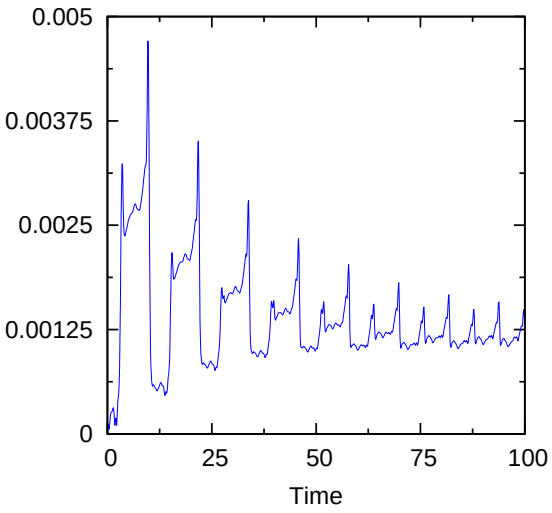


(a) Original

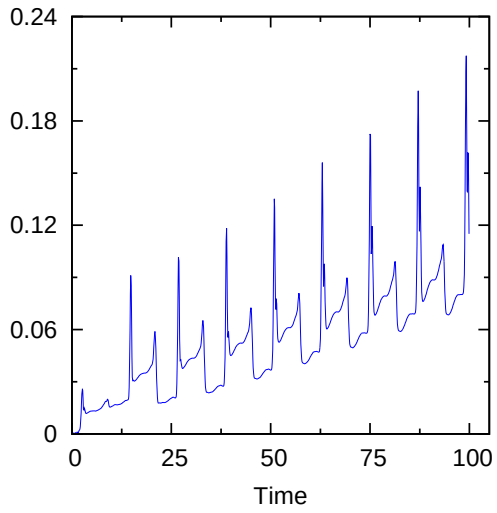


(b) Reconstruction

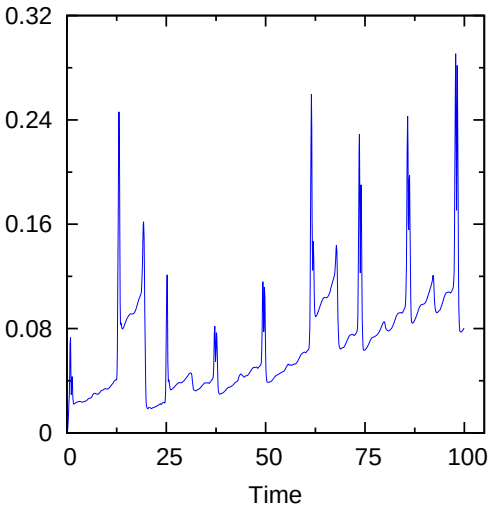
Figure 6.2: The Rössler system. Bifurcation diagrams for both the original and reconstruction over the investigated parameter region. Sampled values are the x-coordinates when the orbit crosses from negative-y to positive-y. Samples were captured over a time interval of 500 units.



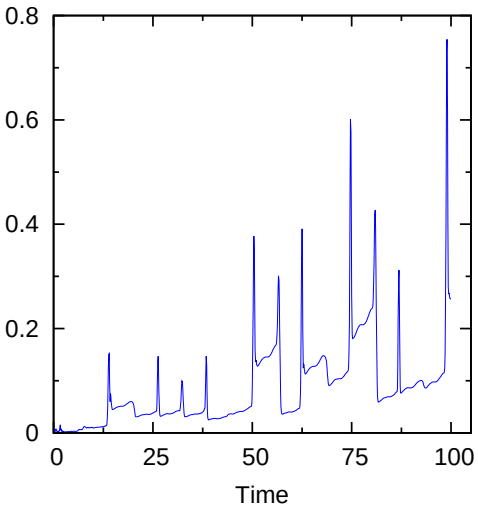
(a) $c = 4.48$



(b) $c = 6.4$



(c) $c = 8.32$



(d) $c = 8.8$

Figure 6.3: The Rössler system. Absolute error in the time series over a time interval of 100 units. The spikes coincide with the fast parts of the attractor.

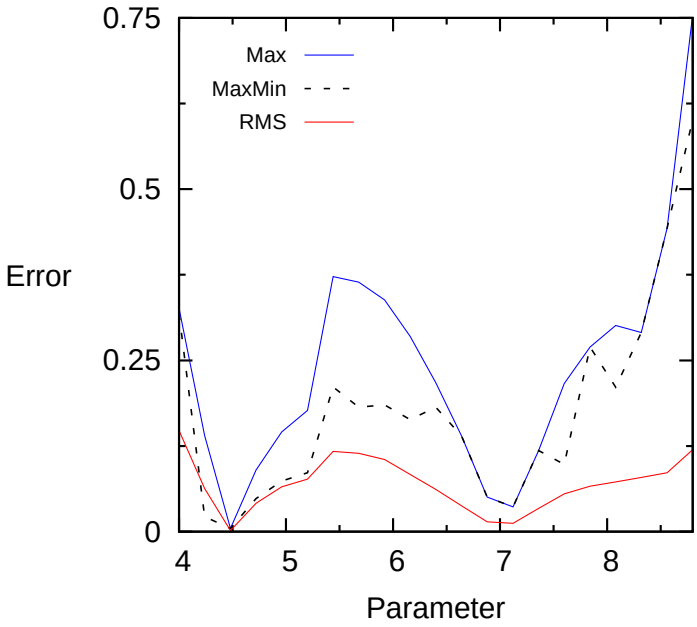


Figure 6.4: The Rössler system. Parameter dependence of the absolute errors in the time series over a time interval of 100 units. The error worsens towards the upper end of the parameter region as the system moves towards chaos.

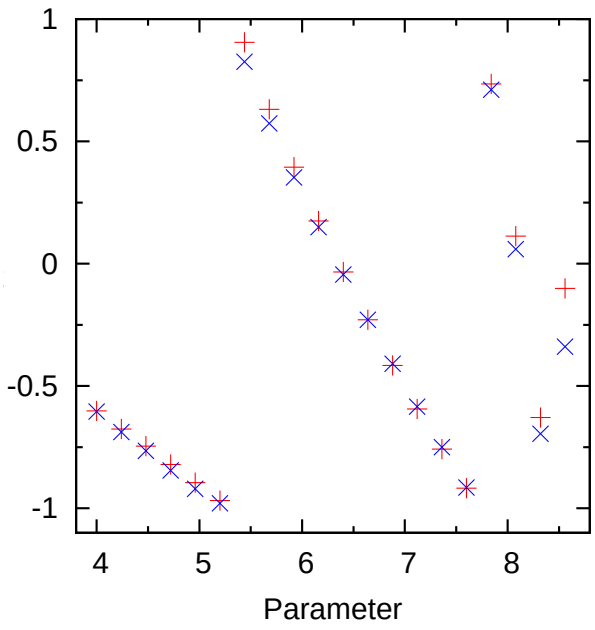


Figure 6.5: The Rössler system. Parameter dependence of the Floquet multipliers. Plusses are from the original, crosses are from the reconstruction. The period of each limit cycle was determined by manually inspecting the time series and the Floquet multipliers were computed by numerically integrating the variational equation over the period.

6.2 Pendulum

A forced damped double pendulum [41] has a state space given by $X = \mathbb{T}^3 \times \mathbb{R}^2$ with coordinates $(\phi, \theta, \psi, v_\phi, v_\theta)$. The angle ψ encapsulates an explicit periodic time dependence in order to make the system autonomous. The vector field is given by

$$\begin{aligned}
\dot{\phi} &= v_\phi \\
\dot{\theta} &= v_\theta \\
\dot{\psi} &= \Omega \\
\dot{v}_\phi &= -f_5(\phi) v_\theta^2 - \delta_2 v_\phi - f_3(\psi) \sin\phi \cos\theta \\
\dot{v}_\theta &= \left(f_1(\phi) v_\theta v_\phi - \delta_1 v_\theta - f_2(\phi) f_3(\psi) \sin\theta\right) / f_4(\phi),
\end{aligned} \tag{6.3}$$

where

$$\begin{aligned}
f_1(\phi) &= (3L + 2\cos\phi)\sin\phi \\
f_2(\phi) &= (M+2)L + \cos\phi \\
f_3(\psi) &= 1 + A\Omega^2 \cos\psi \\
f_4(\phi) &= (M+3)L^2 + 3L\cos\phi + \cos^2\phi \\
f_5(\phi) &= f_1(\phi)/2
\end{aligned}$$

We use the parameter values $M = 7$, $L = 1.1$, $\delta_1 = 0.245$, $\delta_2 = 0.245$, $A = 0.15$, and vary Ω in the region $U = [1.8, 1.82]$. A simple flat torus embedding is used to embed X into $\mathbb{R}^8$,

$$\alpha(\phi, \theta, \psi, v_\phi, v_\theta) = (\cos\phi, \sin\phi, \cos\theta\sin\theta, \cos\psi, \sin\psi, v_\phi, v_\theta). \tag{6.4}$$

We choose this particular embedding due to its simplicity and its ability to generalize to higher dimensions.

6.2.1 Derivatives

$$\begin{aligned}
f_1'(\phi) &= -2 + (3L + 4\cos\phi)\cos\phi \\
f_2'(\phi) &= -\sin\phi \\
f_3'(\psi) &= -A\Omega^2 \sin\psi \\
f_4'(\phi) &= -(3L + 2\cos\phi)\sin\phi \\
f_5'(\phi) &= f_1'(\phi)/2
\end{aligned}$$

$$\begin{aligned}\partial_\phi\dot{\phi} = \partial_\theta\dot{\phi} = \partial_\psi\dot{\phi} = \partial_{v_\theta}\dot{\phi} = \partial_\phi\dot{\theta} = \partial_\theta\dot{\theta} \\ = \partial_\psi\dot{\theta} = \partial_{v_\phi}\dot{\theta} = \partial_\phi\dot{\psi} = \partial_\theta\dot{\psi} = \partial_\psi\dot{\psi} = \partial_{v_\phi}\dot{\psi} = \partial_{v_\theta}\dot{\psi} = 0\end{aligned} \tag{6.5}$$

$$\partial_{v_\phi}\dot{\phi} = \partial_{v_\theta}\dot{\theta} = 1 \tag{6.6}$$

$$\begin{aligned}\partial_\phi\dot{v_\phi} &= -f_5'(\phi)v_\theta^2 - f_3(\psi)\cos\phi\cos\theta \\ \partial_\theta\dot{v_\phi} &= f_3(\psi)\sin\phi\sin\theta \\ \partial_\psi\dot{v_\phi} &= -f_3'(\psi)\sin\phi\cos\theta \\ \partial_{v_\phi}\dot{v_\phi} &= -\delta_2 \\ \partial_{v_\theta}\dot{v_\phi} &= -f_1(\phi)v_\theta\end{aligned} \tag{6.7}$$

$$\begin{aligned}\partial_\phi\dot{v_\theta} &= \frac{f_1'(\phi)v_\theta v_\phi - f_2'(\phi)f_3(\psi)\sin\theta}{f_4(\phi)} - \frac{\left[f_1(\phi)v_\theta v_\phi - \delta_1 v_\theta - f_2(\phi)f_3(\psi)\sin\theta\right]f_4'(\phi)}{[f_4(\phi)]^2} \\ \partial_\theta\dot{v_\theta} &= \frac{-f_2(\phi)f_3(\psi)\cos\theta}{f_4(\phi)} \\ \partial_\psi\dot{v_\theta} &= \frac{-f_2(\phi)f_3'(\psi)\sin\theta}{f_4(\phi)} \\ \partial_{v_\phi}\dot{v_\theta} &= \frac{f_1(\phi)v_\theta}{f_4(\phi)} \\ \partial_{v_\theta}\dot{v_\theta} &= \frac{f_1(\phi)v_\phi - \delta_1}{f_4(\phi)}\end{aligned} \tag{6.8}$$

6.2.2 Embedding Derivatives

Let $(x^1, \cdots, x^8) = \alpha(\phi, \theta, \psi, v_\phi, v_\theta)$. Then the non-zero first partial derivatives are as follows.

$$\begin{aligned}
\partial_\phi \alpha^1 &= -x^2 \\
\partial_\phi \alpha^2 &= x^1 \\
\partial_\theta \alpha^3 &= -x^4 \\
\partial_\theta \alpha^4 &= x^3 \\
\partial_\psi \alpha^5 &= -x^6 \\
\partial_\psi \alpha^6 &= x^5 \\
\partial_{v_\phi} \alpha^7 &= 1 \\
\partial_{v_\theta} \alpha^8 &= 1
\end{aligned} \tag{6.9}$$

Therefore we can compute the pullback metric in the coordinate basis,

$$\begin{aligned}
[g] &= [\alpha_*]^{\mathrm{T}} [\alpha_*] \\
&= \begin{bmatrix} -x^2 & x^1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -x^4 & x^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -x^6 & x^5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -x^2 & 0 & 0 & 0 & 0 \\ x^1 & 0 & 0 & 0 & 0 \\ 0 & -x^4 & 0 & 0 & 0 \\ 0 & x^3 & 0 & 0 & 0 \\ 0 & 0 & -x^6 & 0 & 0 \\ 0 & 0 & x^5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\
&= I
\end{aligned} \tag{6.10}$$

Therefore the coordinate basis on $\mathrm{T}X$ is orthonormal under the pullback metric.

The non-zero second partial derivatives are

$$\begin{aligned}
\partial_{\phi\phi}\alpha^1 &= -x^1 \\
\partial_{\phi\phi}\alpha^2 &= -x^2 \\
\partial_{\theta\theta}\alpha^3 &= -x^3 \\
\partial_{\theta\theta}\alpha^4 &= -x^4 \\
\partial_{\psi\psi}\alpha^5 &= -x^5 \\
\partial_{\psi\psi}\alpha^6 &= -x^6.
\end{aligned} \tag{6.11}$$

6.2.3 Results

Using 5 parameter values, $\mathcal{U} = \{1.8, 1.805, 1.81, 1.815, 1.82\}$, and 1000 state samples from each attractor, a projection was found from $\mathbb{R}^8$ to $\mathbb{R}^4$ using the secant-Grassmannian method. Since the composition of the embedding and orthogonal projection is non-submersive, the reduced tangent space needed to be decomposed using the SVD method detailed in §4.4.4. The dynamics in $\mathbb{R}^4$ was obtained using 100 radial basis functions, which produced stable attractors that correspond to those of the original (Fig. 6.6).

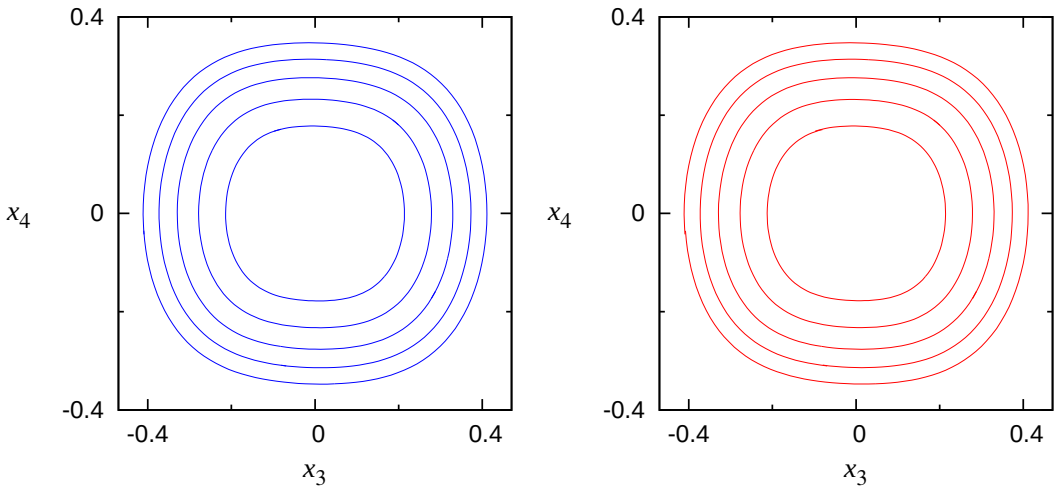


Figure 6.6: Double Pendulum. Attractors from the original model (left) and reduced model (right) for several parameter values. The x_3x_4-plane is shown. In the x_1x_2-plane, all attractors form the unit circle in both original and reduced systems.

The double pendulum demonstrates an example that is not compatible with the POD-Galerkin method, due to both the state space structure and the type of nonlinearities in the dynamics. Although the example is quite low-dimensional, and is only

reduced by a single dimension, it does demonstrate the method can successfully deal with the additional complications raised by such systems.

6.3 Brusselator

The Brusselator is a reaction diffusion equation that describes an autocatalytic chemical reaction, notable for its oscillatory dynamics [15]. Its dynamics are specified by the coupled PDEs for the pair of scalar fields (ϕ_1, ϕ_2) given by

$$\begin{aligned} \partial_t \phi_1 &= a - (b+1)\phi_1 + \phi_1^2 \phi_2 + d_1 \nabla^2 \phi_1 \\ \partial_t \phi_2 &= b\phi_1 - \phi_1^2 \phi_2 + d_2 \nabla^2 \phi_2. \end{aligned} \tag{6.12}$$

We choose a 2-dimensional physical space and apply a 32×32 square spatial discretization with periodic boundaries, giving a 2048-dimensional state space, $X = \mathbb{R}^{2048}$. We fix the parameters $a = d_1 = d_2 = 1$ and vary b in the range $U = [2.5, 2.6]$ with 11 parameter samples and 272 state samples per attractor.

6.3.1 Discretization

For the spatial discretization, we use a centred space scheme,

$$(\nabla^2 \phi)_{ij} = \frac{\phi_{i+1,j} - 2\phi_{ij} + \phi_{i-1,j}}{(\Delta x)^2} + \frac{\phi_{i,j+1} - 2\phi_{ij} + \phi_{i,j-1}}{(\Delta y)^2}, \tag{6.13}$$

giving the system of ODEs,

$$\begin{aligned} (\dot{\phi}_1)_{ij} &= a + (\phi_1)_{ij}^2 (\phi_2)_{ij} - (b+1)(\phi_1)_{ij} + d_1 (\nabla^2 \phi_1)_{ij} \\ (\dot{\phi}_2)_{ij} &= b(\phi_1)_{ij} - (\phi_1)_{ij}^2 (\phi_2)_{ij} + d_2 (\nabla^2 \phi_2)_{ij}. \end{aligned} \tag{6.14}$$

6.3.2 Derivatives

The partial derivatives of the finite-dimensional vector field are

$$\frac{\partial(\nabla^2\phi)_{ij}}{\partial\phi_{\mu\nu}} = \frac{\delta_{i+1,\mu}\delta_{j\nu} - 2\delta_{i\mu}\delta_{j\nu} + \delta_{i-1,\mu}\delta_{j\nu}}{(\Delta x)^2} + \frac{\delta_{i\mu}\delta_{j+1,\nu} - 2\delta_{i\mu}\delta_{j\nu} + \delta_{i\mu}\delta_{j-1,\nu}}{(\Delta y)^2} \tag{6.15}$$

$$\frac{\partial(\dot{\phi}_1)_{ij}}{\partial(\phi_1)_{\mu\nu}} = (2(\phi_1)_{ij}(\phi_2)_{ij} - b - 1)\delta_{i\mu}\delta_{j\nu} + d_1\frac{\partial(\nabla^2(\phi_1))_{ij}}{\partial(\phi_1)_{\mu\nu}} \tag{6.16}$$

$$\frac{\partial(\dot{\phi}_1)_{ij}}{\partial(\phi_2)_{\mu\nu}} = (\phi_1)_{ij}^2\delta_{i\mu}\delta_{j\nu} \tag{6.17}$$

$$\frac{\partial(\dot{\phi}_2)_{ij}}{\partial(\phi_1)_{\mu\nu}} = (b - 2(\phi_1)_{ij}(\phi_2)_{ij})\delta_{i\mu}\delta_{j\nu} \tag{6.18}$$

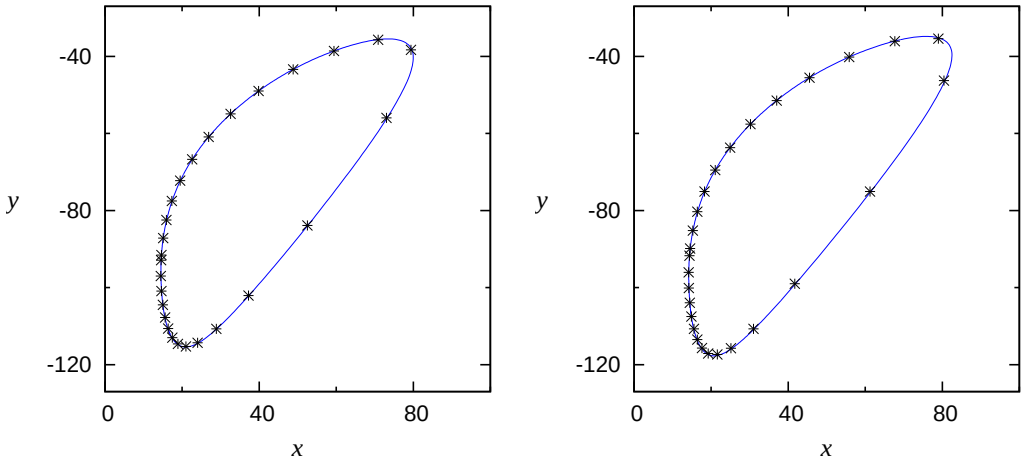
$$\frac{\partial(\dot{\phi}_2)_{ij}}{\partial(\phi_2)_{\mu\nu}} = -(\phi_1)_{ij}^2\delta_{i\mu}\delta_{j\nu} + d_2\frac{\partial(\nabla^2\phi_2)_{ij}}{\partial(\phi_2)_{\mu\nu}}. \tag{6.19}$$

6.3.3 Results

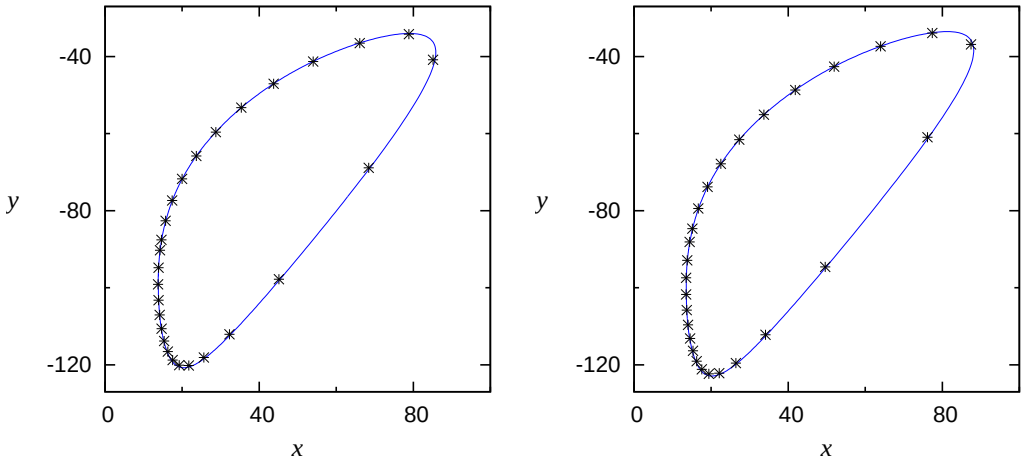
For each parameter value there were 36856 secants with 16 to 18 secants remaining after secant culling. The fact that at most 18 secants remain is a reflection of the limit cycles exploring a 2-dimensional subspace, as Table 3.1 indicates. This is confirmed by the singular values of the combined data set, the first few of which are 5046, 1471, 15.69, 11.73. The 2-dimensional subspace represents 99.998% of the signal energy. Applying the method obtains a projection onto $\hat{X} = \mathbb{R}^2$. The dynamics in $\hat{X}$ are reproduced with 30 radial basis functions, producing the corresponding family of limit cycles, as shown in Fig. 6.7, and Fig. 6.8 shows the family of attractors on a single plot.

6.3.4 POD-Galerkin

The POD-Galerkin method is very suitable for describing the limit cycles of the Brusselator, as the limit cycles explore a subspace that is approximately 2-dimensional. The 2-dimensional POD produces a projection onto this subspace, whose linear inverse is highly accurate as a result. As the dynamics are of the form in §4.6, the Galerkin dynamics can also be evaluated without going through the high-dimensional space. However, we note that the POD requires performing an SVD of a large data matrix (2048×2992 in our example), which took much longer than our secant-based method with the secant culling optimization (50 seconds for secant culling followed



(a) $b = 2.5$ (b) $b = 2.53$



(c) $b = 2.57$ (d) $b = 2.6$

Figure 6.7: The Brusselator. Comparison of the attractors of the original model with the reduced model. Points are from the original model and lines are from the reduced model.

by 2 seconds for optimization over the Grassmannian, compared with 38 minutes for the POD). The attractors of the 2-dimensional POD-Galerkin system are shown in Fig. 6.9 and Fig. 6.10.

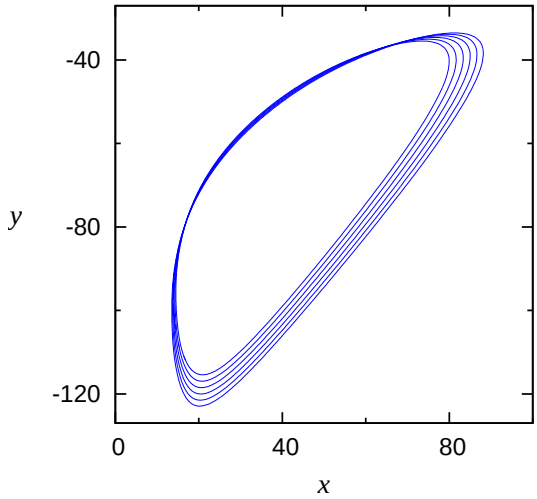
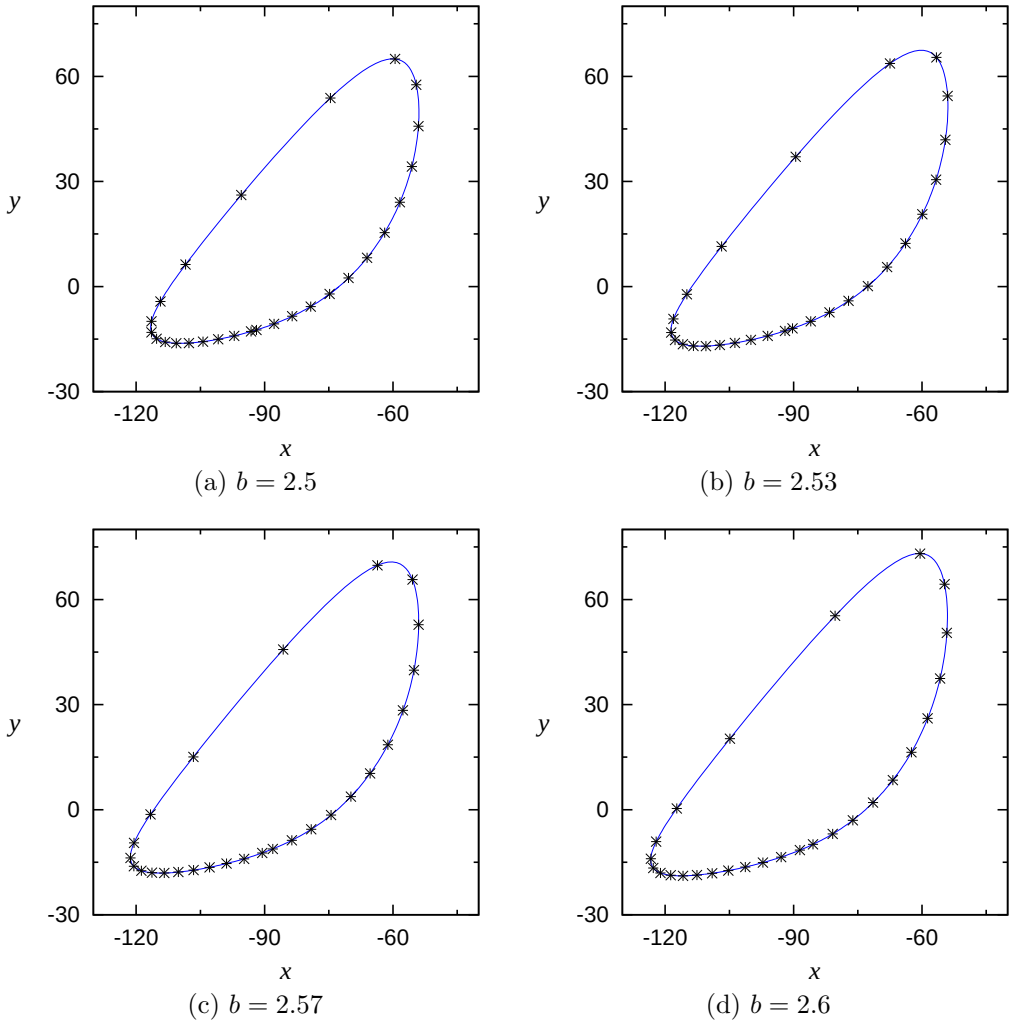


Figure 6.8: The Brusselator. Attractors for several parameter values from the reduced model.



(a) $b = 2.5$ (b) $b = 2.53$

(c) $b = 2.57$ (d) $b = 2.6$

Figure 6.9: The Brusselator. POD-Galerkin. Comparison of the attractors of the original model with the reduced model. Points are from the original model and lines are from the reduced model.

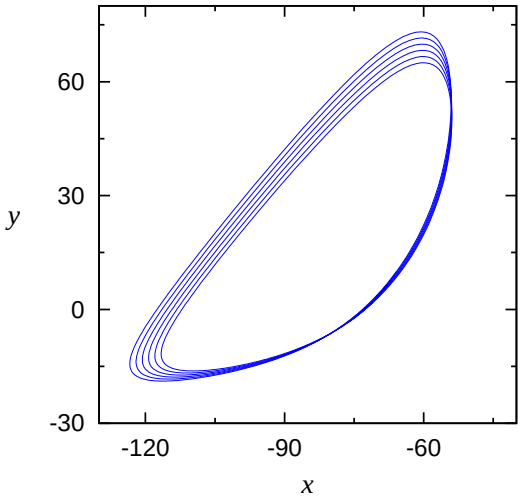


Figure 6.10: The Brusselator. POD-Galerkin. Attractors for several parameter values from the reduced model.

6.4 Dynamo

A 3-dimensional physical space is described by a Riemannian manifold with toroidal coordinates (η, θ, φ) and metric

$$[g] = \mathrm{diag}(g_\eta, g_\theta, g_\varphi) = \frac{1}{c^2}\mathrm{diag}(1, 1, \sinh^2\eta), \tag{6.20}$$

where $c := \cosh\eta - \cos\theta$, given in the coordinate basis. We also make use of the normalized non-holonomic basis

$$(\hat{\eta}, \hat{\theta}, \hat{\varphi}) = \left(\frac{1}{\sqrt{g_\eta}}\partial_\eta, \frac{1}{\sqrt{g_\theta}}\partial_\theta, \frac{1}{\sqrt{g_\varphi}}\partial_\varphi\right) = c\left(\partial_\eta, \partial_\theta, \frac{1}{\sinh\eta}\partial_\varphi\right). \tag{6.21}$$

and the coordinate $s := \exp(\eta_0 - \eta)$, with η_0 constant. A vector field, B, is defined by a pair of scalar fields, $a(\eta, \theta)$ and $b(\eta, \theta)$, via

$$B = \mathrm{curl}\,(a\hat{\varphi}) + b\hat{\varphi}, \tag{6.22}$$

which gives

$$\begin{aligned} B &= B^\eta\hat{\eta} + B^\theta\hat{\theta} + B^\varphi\hat{\varphi} \\ B^\eta &= c\partial_\theta a - a\sin\theta \\ B^\theta &= a\left(\frac{\cosh\eta\cos\theta - 1}{\sinh\eta}\right) + cs\partial_s a \\ B^\varphi &= b. \end{aligned} \tag{6.23}$$

The time evolution of B is therefore determined by the time evolution of a and b, which is prescribed by the coupled PDEs

$$\begin{aligned} \partial_t a &= C_\alpha \alpha b + D^2 a, \\ \partial_t b &= C_\omega F + C_\alpha G + D^2 b. \end{aligned} \tag{6.24}$$

C_α and C_ω are constants, and $\alpha(\eta, \theta, a, b)$, $F(\eta, \theta, a)$ and $G(\eta, \theta, a, b)$ are scalar fields. D^2 is the diffusion operator, given by

$$D^2 A := -g(\mathrm{curl}\,(\mathrm{curl}\,(A\hat{\varphi})), \hat{\varphi}) \tag{6.25}$$

$$D^2 = c^2\left[s^2\partial_s^2 + s\left(1 - \coth\eta + \frac{\sinh\eta}{c}\right)\partial_s + \partial_\theta^2 - \frac{\sin\theta}{c}\partial_\theta - \frac{1}{\sinh^2\eta}\right]. \tag{6.26}$$

$$\alpha := \frac{\sin\theta}{1 + \alpha_B \left\|B\right\|^2}, \tag{6.27}$$

α_B is a constant. The norm is given by

$$\|B\|^2 = (B^\eta)^2 + (B^\theta)^2 + (B^\varphi)^2. \tag{6.28}$$

$$F := -\frac{3}{2}\varpi^{-3/2}\left[(1-\cosh\eta\cos\theta)\partial_\theta a - s\sinh\eta\sin\theta\partial_s a\right], \tag{6.29}$$

where $\varpi := \sinh\eta/c$.

$$G := -\alpha D^2 a - cs\partial_s\alpha\left(\frac{1-\cosh\eta\cos\theta}{\sinh\eta}a + cs\partial_s a\right) - c\partial_\theta\alpha\left(c\partial_\theta a - a\sin\theta\right) \tag{6.30}$$

6.4.1 Discretization

We discretize the space in (s,θ) using a rectangular lattice with separations Δs and $\Delta\theta$. s has range $[0,1]$ and θ is periodic, $[0,2\pi)$. For a field ϕ we use the notation $\phi_{ij} := \phi(s_i,\theta_j)$,

$$\begin{aligned}\dot{a}_{ij} &= C_\alpha\alpha_{ij}b_{ij} + (D^2a)_{ij},\\ \dot{b}_{ij} &= C_\omega F_{ij} + C_\alpha G_{ij} + (D^2b)_{ij}.\end{aligned} \tag{6.31}$$

The spatial derivatives are discretized using a central difference scheme,

$$\partial_s\phi_{ij} = \frac{\phi_{i+1,j}-\phi_{i-1,j}}{2\Delta s}, \qquad \partial_\theta\phi_{ij} = \frac{\phi_{i,j+1}-\phi_{i,j-1}}{2\Delta\theta} \tag{6.32}$$

$$\partial_s^2\phi_{ij} = \frac{\phi_{i+1,j}-2\phi_{ij}+\phi_{i-1,j}}{(\Delta s)^2}, \qquad \partial_\theta^2\phi_{ij} = \frac{\phi_{i,j+1}-2\phi_{ij}+\phi_{i,j-1}}{(\Delta\theta)^2} \tag{6.33}$$

$$\partial_{s\theta}\phi_{ij} = \frac{\phi_{i+1,j+1}-\phi_{i+1,j-1}-\phi_{i-1,j+1}+\phi_{i-1,j-1}}{4\Delta s\Delta\theta} \tag{6.34}$$

6.4.2 Derivatives

The partial derivatives of the finite-dimensional vector field are

$$\begin{aligned}\frac{\partial\dot{a}_{ij}}{\partial a_{\mu\nu}} &= C_\alpha\frac{\partial\alpha_{ij}}{\partial a_{\mu\nu}}b_{ij} + \frac{\partial(D^2a)_{ij}}{\partial a_{\mu\nu}}\\ \frac{\partial\dot{a}_{ij}}{\partial b_{\mu\nu}} &= C_\alpha\frac{\partial\alpha_{ij}}{\partial b_{\mu\nu}}b_{ij} + C_\alpha\alpha_{ij}\delta_{i\mu}\delta_{j\nu}\\ \frac{\partial\dot{b}_{ij}}{\partial a_{\mu\nu}} &= C_\omega\frac{\partial F_{ij}}{\partial a_{\mu\nu}} + C_\alpha\frac{\partial G_{ij}}{\partial a_{\mu\nu}}\\ \frac{\partial\dot{b}_{ij}}{\partial b_{\mu\nu}} &= C_\alpha\frac{\partial G_{ij}}{\partial b_{\mu\nu}} + \frac{\partial(D^2b)_{ij}}{\partial b_{\mu\nu}}\end{aligned} \tag{6.35}$$

The derivatives of the discretized spatial derivatives are

$$\frac{\partial\partial_s\phi_{ij}}{\partial\phi_{\mu\nu}} = \frac{\delta_{j\nu}}{2\Delta s}(\delta_{\mu,i+1}-\delta_{\mu,i-1}), \qquad \frac{\partial\partial_\theta\phi_{ij}}{\partial\phi_{\mu\nu}} = \frac{\delta_{i\mu}}{2\Delta\theta}(\delta_{\nu,j+1}-\delta_{\nu,j-1})$$

$$\frac{\partial\partial_s^2\phi_{ij}}{\partial\phi_{\mu\nu}} = \frac{\delta_{j\nu}}{(\Delta s)^2}(\delta_{\mu,i+1} - 2\delta_{\mu i} + \delta_{\mu,i-1}), \qquad \frac{\partial\partial_\theta^2\phi_{ij}}{\partial\phi_{\mu\nu}} = \frac{\delta_{i\mu}}{(\Delta\theta)^2}(\delta_{\nu,j+1} - 2\delta_{\nu j} + \delta_{\nu,j-1})$$

$$\frac{\partial\partial_{s\theta}\phi_{ij}}{\partial\phi_{\mu\nu}} = \frac{1}{4\Delta s\Delta\theta}(\delta_{\mu,i+1}\delta_{\nu,j+1} - \delta_{\mu,i-1}\delta_{\nu,j+1} - \delta_{\mu,i+1}\delta_{\nu,j-1} + \delta_{\mu,i-1}\delta_{\nu,j-1})$$

The other derivatives are

$$\frac{\partial F_{ij}}{\partial a_{\mu\nu}} = -\frac{3}{2}\varpi_{ij}^{-3/2}\left[(1-\cosh\eta_i\cos\theta_j)\frac{\partial\partial_\theta a_{ij}}{\partial a_{\mu\nu}} - s_i\sinh\eta_i\sin\theta_j\frac{\partial\partial_s a_{ij}}{\partial a_{\mu\nu}}\right]$$

$$\begin{aligned}\frac{\partial(D^2\phi)_{ij}}{\partial\phi_{\mu\nu}} = c_{ij}^2\Bigg[&s_i^2\frac{\partial\partial_s^2\phi_{ij}}{\partial\phi_{\mu\nu}} + s_i\left(1-\coth\eta_i+\frac{\sinh\eta_i}{c_{ij}}\right)\frac{\partial\partial_s\phi_{ij}}{\partial\phi_{\mu\nu}}\\ &+\frac{\partial\partial_\theta^2\phi_{ij}}{\partial\phi_{\mu\nu}} - \frac{\sin\theta_j}{c_{ij}}\frac{\partial\partial_\theta\phi_{ij}}{\partial\phi_{\mu\nu}} - \frac{\delta_{i\mu}\delta_{j\nu}}{\sinh^2\eta_i}\Bigg]\end{aligned} \tag{6.36}$$

$$\begin{aligned}\frac{\partial G_{ij}}{\partial a_{\mu\nu}} =& -\frac{\partial\alpha_{ij}}{\partial a_{\mu\nu}}(D^2a)_{ij} - \alpha_{ij}\frac{\partial(D^2a)_{ij}}{\partial a_{\mu\nu}}\\ &- c_{ij}s_i\frac{\partial\partial_s\alpha_{ij}}{\partial a_{\mu\nu}}\left(\frac{1-\cosh\eta_i\cos\theta_j}{\sinh\eta_i}a_{ij} + c_{ij}s_i\partial_s a_{ij}\right)\\ &- c_{ij}s_i\partial_s\alpha_{ij}\left(\frac{1-\cosh\eta_i\cos\theta_j}{\sinh\eta_i}\delta_{i\mu}\delta_{j\nu} + c_{ij}s_i\frac{\partial\partial_s a_{ij}}{\partial a_{\mu\nu}}\right)\\ &- c_{ij}\frac{\partial\partial_\theta\alpha_{ij}}{\partial a_{\mu\nu}}(c_{ij}\partial_\theta a_{ij} - a_{ij}\sin\theta_j)\\ &- c_{ij}\partial_\theta\alpha_{ij}\left(c_{ij}\frac{\partial\partial_\theta a_{ij}}{\partial a_{\mu\nu}} - \delta_{i\mu}\delta_{j\nu}\sin\theta_j\right)\end{aligned} \tag{6.37}$$

$$\begin{aligned}\frac{\partial G_{ij}}{\partial b_{\mu\nu}} =& -(D^2a)_{ij}\frac{\partial\alpha_{ij}}{\partial b_{\mu\nu}}\\ &- c_{ij}s_i\frac{\partial\partial_s\alpha_{ij}}{\partial b_{\mu\nu}}\left(\frac{1-\cosh\eta_i\cos\theta_j}{\sinh\eta_i}a_{ij} + c_{ij}s_i\partial_s a_{ij}\right)\\ &- c_{ij}\frac{\partial\partial_\theta\alpha_{ij}}{\partial b_{\mu\nu}}(c_{ij}\partial_\theta a_{ij} - a_{ij}\sin\theta_j)\end{aligned} \tag{6.38}$$

$$\begin{aligned}\frac{\partial\alpha_{ij}}{\partial\phi_{\mu\nu}} &= \frac{-\alpha_{ij}\alpha_B}{1+\alpha_B\left\|B\right\|_{ij}^2}\frac{\partial}{\partial\phi_{\mu\nu}}\left\|B\right\|_{ij}^2\\ \frac{\partial\left\|B\right\|_{ij}^2}{\partial a_{\mu\nu}} &= 2\left(B_{ij}^\eta\frac{\partial B_{ij}^\eta}{\partial a_{\mu\nu}} + B_{ij}^\theta\frac{\partial B_{ij}^\theta}{\partial a_{\mu\nu}}\right)\\ \frac{\partial\left\|B\right\|_{ij}^2}{\partial b_{\mu\nu}} &= 2b_{ij}\delta_{i\mu}\delta_{j\nu}\end{aligned} \tag{6.39}$$

$$\begin{aligned}\frac{\partial B_{ij}^\eta}{\partial a_{\mu\nu}} &= c_{ij}\frac{\partial\partial_\theta a_{ij}}{\partial a_{\mu\nu}} - \delta_{i\mu}\delta_{j\nu}\sin\theta_j\\ \frac{\partial B_{ij}^\theta}{\partial a_{\mu\nu}} &= \delta_{i\mu}\delta_{j\nu}\left(\frac{\cosh\eta_i\cos\theta_j - 1}{\sinh\eta_i}\right) + c_{ij}s_i\frac{\partial(\partial_s a)_{ij}}{\partial a_{\mu\nu}}\end{aligned} \tag{6.40}$$

6.4.3 Boundary Conditions

At $s = 0$, the boundary rows are given by the average of the adjacent row,

$$\begin{aligned} \dot{a}_{1j} &= \frac{1}{N_\theta} \sum_{k=1}^{N_\theta} \dot{a}_{2k} \quad \forall j \\ \dot{b}_{1j} &= \frac{1}{N_\theta} \sum_{k=1}^{N_\theta} \dot{b}_{2k} \quad \forall j. \end{aligned} \tag{6.41}$$

At $s = 1$, the row of boundary points in a are given by a linear combination of the two adjacent rows.

$$\dot{a}_{N_s,j} = \sum_{k=1}^{N_\theta} A^{(1)}_{jk} \; \dot{a}_{N_s-1,k} + \sum_{k=1}^{N_\theta} A^{(2)}_{jk} \; \dot{a}_{N_s-2,k}, \tag{6.42}$$

where the matrices $A^{(1)}$ and $A^{(2)}$ are given. b is zero on the boundary,

$$\dot{b}_{N_s,j} = 0 \quad \forall j. \tag{6.43}$$

The derivatives are

$$\begin{aligned} \frac{\partial \dot{a}_{1j}}{\partial \phi_{\mu\nu}} &= \frac{1}{N_\theta} \sum_{k=1}^{N_\theta} \frac{\partial \dot{a}_{2j}}{\partial \phi_{\mu\nu}} \\ \frac{\partial \dot{b}_{1j}}{\partial \phi_{\mu\nu}} &= \frac{1}{N_\theta} \sum_{k=1}^{N_\theta} \frac{\partial \dot{b}_{2k}}{\partial \phi_{\mu\nu}}. \end{aligned} \tag{6.44}$$

$$\frac{\partial \dot{a}_{N_s,j}}{\partial \phi_{\mu\nu}} = \sum_{k=1}^{N_\theta} A^{(1)}_{jk} \; \frac{\partial \dot{a}_{N_s-1,k}}{\partial \phi_{\mu\nu}} + \sum_{k=1}^{N_\theta} A^{(2)}_{jk} \; \frac{\partial \dot{a}_{N_s-2,k}}{\partial \phi_{\mu\nu}}, \tag{6.45}$$

$$\frac{\partial \dot{b}_{N_s,j}}{\partial \phi_{\mu\nu}} = 0. \tag{6.46}$$

6.4.4 Results

The constants were fixed to $\eta_0 = 1$, $\alpha_B = 1$, and $C_\omega = -1000$. The spatial discretization consisted of 41×160 points, resulting in a state space of $X = \mathbb{R}^{13120}$. At the parameter value of $C_\alpha = 4$, the system exhibits quasi-periodic behaviour on a torus. Using 1326 data points to describe this attractor, there were 878475 secants and 5436 secants remaining after secant culling (more than 99% of the secants were culled). A projection was found onto $\mathbb{R}^4$, which is shown in Fig. 6.11. This represents a huge reduction in dimension, 13120 to 4. The secant culling procedure resulted in approximately half the time requirement versus no culling, a saving of around 24 hours in this example.

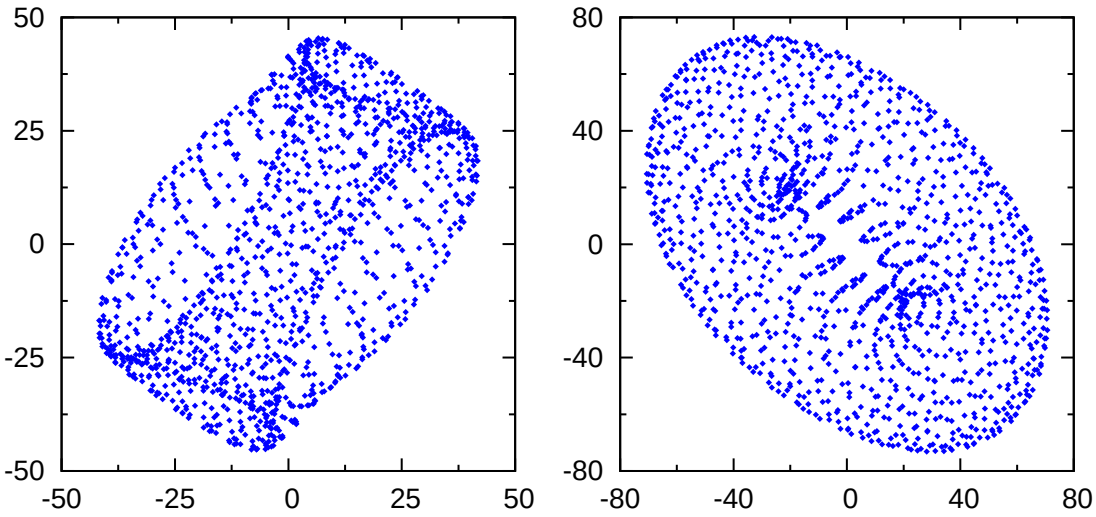


Figure 6.11: Torus in the Dynamo system. Optimal orthogonal projection onto a 4-dimensional subspace.

Chapter 7

Conclusion

7.1 Summary

In this work we have considered the problem of describing a low-dimensional attractor and its dynamics in a low-dimensional ambient space. We have made use of an existing secant-based projection method due to its desirable topological features, rather than other projection methods (such as the PCA) which do not possess these features.

There are three primary original contributions developed in this work. Firstly, we have proposed the idea of *approximate secants*, which correspond to double infinite cones, and studied their properties (§3.9). The two applications of this are a secant culling procedure that can greatly reduce the number of secants in a data set (§3.10), and a method of estimating the dimension of the plane explored by the attractor (§3.10.2).

Secondly, we have formulated the problem from a differential geometric perspective, which allows for general state space manifolds with dynamics given by a vector field (§4). We discussed and resolved issues with reproducing a stable attractor using the derivatives of the vector field that arise due to a lack of submersiveness in the dimension-reducing map (§4.4). We showed that our more general criteria is satisfied by the special case of a Galerkin-style linear projection (§4.6). For the more general case, we adapted an existing optimization method to produce a practical implementation using radial basis functions (§4.8).

Thirdly, we extended our geometric formulation to the problem of parameter dependence, where a family of attractors are to be reproduced in a low-dimensional space

(§5). We began by formulating the problem in full generality, where the dimension-reducing map is permitted to depend on the parameter (§5.2). We considered a product manifold formulation in which the family of vector fields becomes a single horizontal vector field on the product manifold (§5.2.1). We then simplified the problem by restricting ourselves to parameter-independent reduction, which permits a decoupling of the problem into two stages – obtaining a single dimension-reducing map, and then obtaining a corresponding parameter map (§5.2.3). This decoupling allows for the secant-based projection to be extended straightforwardly to include a parameter space (§5.3). A similar extension was used to develop an optimization method for obtaining the parameter map, by constructing a linear optimization problem over a space of affine maps (§5.5.1).

We demonstrated the dimensionality reduction with parameters in §6 using a variety of example systems, including a nonautonomous system, a high-dimensional discretized PDE, and a small system with period-doubling bifurcations.

7.2 Limitations of the Method

The method requires data sets containing samples of points from the attractors for multiple parameter values. The role of this data is to describe the attractors in order to identify their location, shape and size within the state space. Although the data can be produced by a numerical simulation of the original system, care must be taken to ensure the data fully explores each attractor, and that a sufficient number of parameter values are sampled. This data requirement makes the approach less feasible for investigating many parameters. For example, if the parameter space under investigation is of the form $[a_1, b_1] \times \cdots \times [a_l, b_l] \subset \mathbb{R}^l$ and each interval is sampled N times, the total number of parameter samples is N^l, which increases exponentially with the parameter dimension, l.

In order to use an orthogonal projection in the presence of a parameter space, we made use of an affine inverse approximation §5.3.1, as the derivative of the affine inverse is independent of the parameter (5.11). This allowed us to determine the reduced model without needing to construct the parameter dependence of the inverse.

However, this affine inverse is only accurate when the each attractor explores a low-dimensional hyperplane. In order to use a non-affine inverse, the parameter dependence of the inverse needs to be constructed.

The method requires the *a priori* choice of a family of reduced vector fields. The particular family of reduced vector fields used in the method consists of radial basis functions, which work well for a variety of examples. However, adding more radial basis functions to the summation (4.58) increases the dimension of the reduced parameter space $\hat{U}$, which in turn increases the size of the linear problem for the parameter map, potentially leading to rank deficiency. This can be a limiting factor on the size of the parameter region used in the reduction.

The particular parameter map used is affine, which may place limits on the size of the parameter space under consideration if the original parameter dependence is non-affine. However, the approach can support non-affine parameter maps.

7.3 Future Work

The work completed so far suggests a number of avenues for further research.

7.3.1 Secant-Based Method For Classification Problems

With a small modification, the secant-Grassmann projection method (§3) lends itself to classification problems. A classification problem is a type of data reduction problem in which one has a typed data set (each data point has a 'type', e.g. colour). The objective is to learn which regions of space correspond to each type so that a data point of unknown type can be correctly classified based on its position. Reducing the dimension of such a data set while preserving the separation of types would be useful before the analysis and learning stages.

By using only the secants generated by pairs of points with different types, one can use the secant-Grassmann method to obtain an orthogonal projection that best separates the different types from each other, without caring about the separation of same-type points. This approach would have an advantage over classic methods such as the PCA, which is unable to distinguish between the relevant variance (spreading out points of different type) from the irrelevant variance (spreading out points of the

same type).

7.3.2 Dimensionality Reduction with Noise

Models of physical systems are often deterministic, being derived from fundamental physical laws. However in large systems, there may be microscopic details or external influences that have not been included in the deterministic model. Rather than modelling these in detail as complicated deterministic processes, one can model them as a simple but stochastic contribution to the dynamics. This noise can then account for small fluctuations around the driving deterministic dynamics.

The presence of noise can produce behaviour not possible in a purely deterministic system. For example, multistability occurs where the noise causes an orbit to move from one basin of attraction to another nearby basin, allowing a single orbit to encounter multiple attractors. Noise can also destroy features of the deterministic model that are sensitive to perturbations. Such features would therefore be unlikely to be observed in the physical system.

Random Variables on the Tangent Bundle

When reducing the dimension of a noisy system, rather than treating the system as a completely stochastic system, we can separate the deterministic and stochastic parts. We can therefore take a geometric approach with the deterministic part of the system, and regard the noise to be produced by random variables on the tangent spaces of the state space. One could then produce a corresponding set of random variables on the reduced tangent space that have the appropriate distribution.

7.3.3 A Zoo of Models

In formulating the problem of dimensionality reduction with parameter dependence (§5), we saw that the problem is equivalent to taking two dynamical systems (with different state and parameter spaces), and obtaining a relationship between them such that their dynamics agree on a subset of the state space (the attractors) for a region of parameter space. In this work, we were free the choose the second system (the reduced model) to make it easy to obtain a solution.

A similar problem is produced by taking two different models of a physical system (which may have very different state and parameter spaces) and seeking a relationship between them. This situation is common in sciences such as biology, where the systems studied are large and complex, which results in a 'zoo' of models of varying shapes, sizes, and levels of complexity, yet all describing the same underlying physical system. For example, the most complex models are often those derived from fundamental physics applied to all components of the system. The parameters in these models are often physically meaningful, but the state space is high-dimensional and the dynamics are highly nonlinear. One is therefore constrained to numerical investigations with such models. At the other end of the spectrum, there are small simple models that are not derived from physics, but are constructed abstractly in order to mimic observations made in either experiment or complex models. The parameters in these systems do not usually have any direct physical interpretation, although the dynamics is simple enough to permit analysis.

The approaches taken in this work could be adapted for finding connections between different models of a physical system. This could allow the parameters of a simple abstract model to be converted into the physically-meaningful parameters of a complex model.

Appendix A

Linear Algebra

A.1 Adjoint

Theorem A.1.1. Let $A : X \to \hat{X}$ be a linear map between inner product spaces. Then $\ker A^\dagger = (\operatorname{im} A)^\perp$.

Proof. Since inner products are non-degenerate, we have

$$g(x, y) = 0 \ \forall y \Leftrightarrow x = 0. \tag{A.1}$$

Given $\hat{x} \in (\operatorname{im} A)^\perp$,

$$\hat{g}(\hat{x}, Ay) = 0 \ \forall y \in X. \tag{A.2}$$

By the definition of the adjoint,

$$g(A^\dagger \hat{x}, y) = 0 \ \forall y \in X. \tag{A.3}$$

By the non-degeneracy of the inner product, this implies $A^\dagger \hat{x} = 0$, i.e. $\hat{x} \in \ker A^\dagger$. Similarly, given $\hat{x} \in \ker A^\dagger$,

$$g(A^\dagger \hat{x}, y) = 0 \ \forall y \in X. \tag{A.4}$$

By the definition of the adjoint,

$$\hat{g}(\hat{x}, Ay) = 0 \ \forall y \in X, \tag{A.5}$$

hence $\hat{x} \in (\operatorname{im} A)^\perp$. Therefore $\ker A^\dagger = (\operatorname{im} A)^\perp$. □

Theorem A.1.2. Let $A : X \to \hat{X}$ be a linear map between inner product spaces. Then $\operatorname{im} A^\dagger = (\ker A)^\perp$.

Proof. By Theorem A.1.1,

$$\begin{aligned} \ker A &= \ker(A^\dagger)^\dagger \\ &= (\operatorname{im} A^\dagger)^\perp \end{aligned} \tag{A.6}$$

Hence $\operatorname{im} A^\dagger = (\ker A)^\perp$. □

Theorem A.1.3. Let $A : X \to \hat{X}$ be a linear map between inner product spaces. Then $X = \ker A \oplus \operatorname{im} A^\dagger$ and $\hat{X} = \operatorname{im} A \oplus \ker A^\dagger$ are orthogonal decompositions of the domain and codomain respectively.

Proof. By Theorem A.1.1 and Theorem A.1.2. □

Corollary A.1.4. Let $A : X \to X$ be a self-adjoint linear operator on an inner product space. Then X can be written as the orthogonal decomposition $X = \operatorname{im} A \oplus \ker A$.

Theorem A.1.5. Let $A : X \to \hat{X}$ be a linear map between inner product spaces. Then $\operatorname{rank} A = \operatorname{rank} A^\dagger$.

Proof. By the rank-nullity theorem we have

$$\begin{aligned} \dim \operatorname{im} A + \dim \ker A &= \dim X \\ \dim \operatorname{im} A^\dagger + \dim \ker A^\dagger &= \dim \hat{X}. \end{aligned} \tag{A.7}$$

By Theorem A.1.3, $X = \ker A \oplus \operatorname{im} A^\dagger$ and $\hat{X} = \operatorname{im} A \oplus \ker A^\dagger$, and so

$$\begin{aligned} \dim \ker A + \dim \operatorname{im} A^\dagger &= \dim X \\ \dim \operatorname{im} A + \dim \ker A^\dagger &= \dim \hat{X}. \end{aligned} \tag{A.8}$$

Combining these gives $\dim \operatorname{im} A = \dim \operatorname{im} A^\dagger$. □

A.2 Projections

Lemma A.2.1. Let P_1 and P_2 be projections. If $[P_1, P_2] = 0$ then $P_1 P_2$ is a projection. Furthermore, if P_1 and P_2 are orthogonal projections then so is $P_1 P_2$.

Proof.

$$\begin{aligned} (P_1 P_2)^2 &= P_1 P_2 P_1 P_2 \\ &= P_1 \left(P_1 P_2 - [P_1, P_2] \right) P_2 \\ &= P_1^2 P_2^2 - P_1 \left[P_1, P_2 \right] P_2 \\ &= P_1 P_2 - P_1 \left[P_1, P_2 \right] P_2. \end{aligned} \tag{A.9}$$

If $[P_1, P_2] = 0$ then $(P_1P_2)^2 = P_1P_2$ and so P_1P_2 is a projection. Also consider the adjoint,

$$\begin{aligned}(P_1P_2)^\dagger &= (P_2P_1 + [P_1, P_2])^\dagger \\ &= P_1^\dagger P_2^\dagger + [P_1, P_2]^\dagger\end{aligned} \tag{A.10}$$

If $[P_1, P_2] = 0$, $P_1^\dagger = P_1$ and $P_2^\dagger = P_2$, then P_1P_2 is self-adjoint. Therefore P_1P_2 is an orthogonal projection. □

Lemma A.2.2. If P is a projection then $\mathrm{id} - P$ is a projection with kernel $\mathrm{im}\, P$ and image $\ker P$. Furthermore, if P is an orthogonal projection, $\mathrm{id} - P$ is an orthogonal projection.

Proof. If $P^2 = P$ then

$$\begin{aligned}(\mathrm{id} - P)^2 &= P^2 - 2P + \mathrm{id}^2 \\ &= \mathrm{id} - P,\end{aligned} \tag{A.11}$$

so $\mathrm{id} - P$ is a projection. Note that $(\mathrm{id} - P)x = x - x = 0$ for all $x \in \mathrm{im}\, P$ and $(\mathrm{id} - P)x = x$ for all $x \in \ker P$. Therefore $\mathrm{im}\,(\mathrm{id} - P) = \ker P$ and $\ker(\mathrm{id} - P) = \mathrm{im}\, P$.

Furthermore, if P is self-adjoint then

$$\begin{aligned}(\mathrm{id} - P)^\dagger &= \mathrm{id}^\dagger - P^\dagger \\ &= \mathrm{id} - P\end{aligned} \tag{A.12}$$

and so $\mathrm{id} - P$ is an orthogonal projection. □

Lemma A.2.3. If P_1 and P_2 are projections that anti-commute, $\{P_1, P_2\} = 0$, then $P_1 + P_2$ is a projection. Furthermore, if P_1 and P_2 are orthogonal projections, $P_1 + P_2$ is an orthogonal projection.

Proof. If $P^2 = P$ and $\{P_1, P_2\} = 0$ then

$$\begin{aligned}(P_1 + P_2)^2 &= P_1^2 + P_1P_2 + P_2P_1 + P_2^2 \\ &= P_1 + P_2 + \{P_1, P_2\} \\ &= P_1 + P_2\end{aligned} \tag{A.13}$$

so $P_1 + P_2$ is a projection.

Furthermore, if P_1 and P_2 are self-adjoint then

$$(P_1 + P_2)^\dagger = P_1^\dagger + P_2^\dagger = P_1 + P_2 \tag{A.14}$$

and so $P_1 + P_2$ is an orthogonal projection. □

Lemma A.2.4. If P_1 and P_2 are orthogonal projections with orthogonal images, then $P_1P_2 = P_2P_1 = 0$ and so $[P_1, P_2] = 0$ and $\{P_1, P_2\} = 0$.

Proof. The kernel of an orthogonal projection is the orthogonal complement of its image. Therefore $\operatorname{im} P_1 \subseteq (\operatorname{im} P_2)^\perp = \ker P_2$ and $\operatorname{im} P_2 \subseteq (\operatorname{im} P_1)^\perp = \ker P_1$. Hence $P_1P_2 = P_2P_1 = 0$, $[P_1, P_2] = 0$ and $\{P_1, P_2\} = 0$. □

Lemma A.2.5. If P_1 and P_2 are orthogonal projections with orthogonal kernels, then $[P_1, P_2] = 0$.

Proof. By Lemma A.2.2, $\mathrm{id} - P_1$ and $\mathrm{id} - P_2$ are orthogonal projections onto $\ker P_1$ and $\ker P_2$ respectively. Since these have orthogonal images, applying Lemma A.2.4 gives

$$\begin{aligned} 0 &= [\mathrm{id} - P_1, \mathrm{id} - P_2] \\ &= [\mathrm{id}, \mathrm{id}] - [\mathrm{id}, P_2] - [P_1, \mathrm{id}] + [P_1, P_2] \\ &= [P_1, P_2]\,. \end{aligned} \tag{A.15}$$

□

Lemma A.2.6. If P_1 and P_2 are orthogonal projections with orthogonal kernels, then $\ker P_1P_2 = \ker P_1 \oplus \ker P_2$.

Proof. $\ker P_2 \subseteq (\ker P_1)^\perp = \operatorname{im} P_1$ and $\ker P_1 \subseteq (\ker P_2)^\perp = \operatorname{im} P_2$, therefore the kernel of one projection is preserved by the other. This means we can write

$$\begin{aligned} \ker P_1P_2 &= (\operatorname{im} P_2 \cap \ker P_1) \oplus \ker P_2 \\ &= \ker P_1 \oplus \ker P_2. \end{aligned} \tag{A.16}$$

□

Proposition A.2.7. If P_1 and P_2 are orthogonal projections with orthogonal images, then $P_1 + P_2$ is an orthogonal projection.

Proof. By Lemma A.2.4 and Lemma A.2.3. □

Proposition A.2.8. If P_1 and P_2 are orthogonal projections with orthogonal kernels, then P_1P_2 is an orthogonal projection with kernel $\ker P_1 \oplus \ker P_2$.

Proof. By Lemma A.2.5, Lemma A.2.1 and Lemma A.2.6. □

Proposition A.2.9. Let $\{v_1, \cdots, v_d\}$ be a set of mutually orthogonal vectors. The orthogonal projection with kernel span $\{v_1, \cdots, v_d\}$ can be decomposed into the sequence of codimension-1 orthogonal projections,

$$P = P_{v_1} \cdots P_{v_d}, \tag{A.17}$$

which is independent of the order of composition.

Proof. Since the vectors are mutually orthogonal, by Proposition A.2.8 the RHS is an orthogonal projection with $\ker P_{v_1} \cdots P_{v_d} = \ker P_{v_1} \oplus \cdots \oplus \ker P_{v_d} = \text{span}\{v_1, \cdots, v_d\}$. Therefore, as orthogonal projections are uniquely determined by their kernels, $P = P_{v_1} \cdots P_{v_d}$. By Lemma A.2.5 the codimension-1 projections commute and therefore the decomposition is independent of the order. □

Appendix B

Cones

Lemma B.0.10. Let $C_\theta^n(k_1)$ and $C_\theta^n(k_2)$ be double cones with $\theta < \pi/2$. The two cones have a nontrivial intersection iff $\cos^2\alpha \geq \cos^2 2\theta$, where $\cos\alpha = \langle k_1, k_2\rangle \,/\, \|k_1\|\,\|k_2\|$.

Proof. If k_1 and k_2 are linearly dependent then $\cos^2\alpha = 1 \geq \cos^2 2\theta$ for any θ, and the two cones are equal. Therefore we consider the case when k_1 and k_2 are linearly independent. We can reduce the n-dimensional problem to two dimensions by projecting onto the 2-dimensional subspace containing k_1 and k_2. Let $W : \mathbb{R}^2 \to \mathbb{R}^n$ be an orthonormal map such that $W(\mathbb{R}^2) = \operatorname{span}\{k_1, k_2\}$. Since $\theta < \pi/2$, the intersection of the kernel of $W^\dagger$ with the cones is $\{0\}$. By Proposition 3.9.7, $W^\dagger$ applied to both cones produces a pair of double 2-cones in $\mathbb{R}^2$ of angle θ. Therefore if the intersection of the cones in $\mathbb{R}^2$ is trivial, the intersection of the original cones is trivial. Similarly, if the 2-cones intersect nontrivially, the original cones intersect nontrivially. This allows us to reduce the n-dimensional case to the 2-dimensional case, which is straightforward. □

Lemma B.0.11. Let $C_\theta^n(k)$ be a double cone and $P : \mathbb{R}^n \to \mathbb{R}^n$ be an orthogonal projection with $k \in \operatorname{im} P$. Then $PC_\theta^n(k) \subseteq C_\theta^n(k)$.

Proof. By Proposition 3.9.7, $PC_\theta^n(k)$ is a double cone of dimension $\operatorname{im} P$ and angle θ with axis Pk. For each $y \in PC_\theta^n(k)$,

$$|\langle y, Pk\rangle|^2 \geq \|y\|^2 \,\|Pk\|^2 \cos^2\theta. \tag{B.1}$$

Since $k \in \operatorname{im} P$, $Pk = k$ and

$$|\langle y, k\rangle|^2 \geq \|y\|^2 \,\|k\|^2 \cos^2\theta. \tag{B.2}$$

Hence $y \in C_\theta^n(k)$. □

Lemma B.0.12. Let $C_\theta^n(k_1)$ and $C_\theta^n(k_2)$ be double cones and $P : \mathbb{R}^n \to \mathbb{R}^n$ be the orthogonal projection onto $\operatorname{span}\{k_1, k_2\}$. Then $PC_\theta^n(k_1) \cap PC_\theta^n(k_2) = \{0\}$ iff $C_\theta^n(k_1) \cap C_\theta^n(k_2) = \{0\}$.

Proof. By Lemma B.0.11. □

Bibliography

[1] A. C. Antoulas, D. C. Soresnsen, and S. Gugercin, *A survey of reduction methods for large-scale systems*, Contemp. Math. **280** (2001), 193–219.

[2] D. S. Broomhead and M. Kirby, *A New Approach to Dimensionality Reduction: Theory and Algorithms*, SIAM J. Appl. Math. **60** (2000), no. 6, 2114–2142.

[3] D. S. Broomhead and M. J. Kirby, *The Whitney Reduction Network: A Method for Computing Autoassociative Graphs*, Neural Comput. **13** (2001), no. 11, 2595–2616.

[4] ———, *Dimensionality Reduction Using Secant-Based Projection Methods: The Induced Dynamics in Projected Systems*, Nonlinear Dynam. **41** (2005), 47–67.

[5] J. Carr, *Applications of centre manifold theory*, Applied Mathematical Sciences Series, no. v. 35, Springer-Verlag, 1981.

[6] S. Chaturantabut and D. C. Sorensen, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comput. **32** (2010), no. 5, 2737–2764.

[7] J. Chin, J. Harting, S. Jha, P. V. Coveney, A. R. Porter, and S. M. Pickles, *Steering in computational science: mesoscale modelling and simulation*, Appl. Numer. Math. **43** (2003), 9–44.

[8] I. Chueshov, *Introduction to the theory of infinite-dimensional dissipative system*, University Lectures in Contemporary Mathematics, ACTA, 2002.

[9] T. F. Cox and M. A. A. Cox, *Multidimensional scaling*, second ed., Chapman & Hall, 2001.

[10] C. W. Curtis, *Linear algebra: An introductory approach*, Allyn and Bacon, 1968.

[11] David L. Donoho and Carrie Grimes, *Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data*, Proc. Natl. Acad. Sci. USA **100** (2003), no. 10, 5591–5596.

[12] A. Edelman, T. A. Arias, and S. T. Smith, *The Geometry of Algorithms with Orthogonality Constraints*, SIAM J. Matrix Anal. Appl. **20** (1998), no. 2, 303–353.

[13] Neil Fenichel, *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J. **21** (1972), 193–226.

[14] C. Foias, M.S. Jolly, I.G. Kevrekidis, G.R. Sell, and E.S. Titi, *On the computation of inertial manifolds*, Phys. Lett. A **131** (1988), no. 78, 433–436.

[15] P. Glansdorff and I. Prigogine, *Thermodynamic theory of structure, stability, and fluctuations*, Wiley, 1971.

[16] W. H. Greub, *Multilinear algebra*, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 1967.

[17] ———, *Linear algebra*, fourth ed., Graduate Texts in Mathematics, Springer, 1975.

[18] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Applied Mathematical Sciences, no. v. 42, Springer, 1983.

[19] Uwe Helmke and John B. Moore, *Optimization and dynamical systems*, Würzburg, Germany, 1994.

[20] Michael Hinze and Stefan Volkwein, *Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control*, Dimension Reduction of Large-Scale Systems (Peter Benner, Danny C. Sorensen, Volker Mehrmann, Timothy J. Barth, Michael Griebel, David E. Keyes, Risto M. Nieminen, Dirk Roose, and Tamar Schlick, eds.), Lecture Notes in Computational Science and Engineering, vol. 45, Springer Berlin Heidelberg, 2005, 10.1007/3-540-27909-1_10, pp. 261–306.

[21] M. Hirsch, S. Smale, and R. Devaney, *Differential Equations, Dynamical Systems, and an introduction to Chaos*, 2nd ed., Elsevier, 2004.

[22] M. W. Hirsch, *Differential Topology*, Springer-Verlag, 1976.

[23] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant Manifolds*, Springer-Verlag, 1977.

[24] I. T. Jolliffe, *Principal Component Analysis*, second ed., Springer, October 2002.

[25] Jürgen Jost, *Riemannian geometry and geometric analysis*, 5th ed., Springer, 2008.

[26] Al Kelley, *The stable, center-stable, center, center-unstable, unstable manifolds*, J. Differential Equations **3** (1967), no. 4, 546 – 570.

[27] Michael Kirby and Rick Miranda, *Nonlinear reduction of high-dimensional dynamical systems via neural networks*, Phys. Rev. Lett. **72** (1994), 1822–1825.

[28] S. Lang, *Introduction to differentiable manifolds*, Universitext (1979), Springer, 2002.

[29] S. Li, *Concise formulas for the area and volume of a hyperspherical cap*, Asian Journal of Mathematics and Statistics **4** (2011), no. 1, 66–70.

[30] Bernd R. Noack, Michael Schlegel, Marek Morzynski, and Gilead Tadmor, *Galerkin method for nonlinear dynamics*, Reduced-Order Modelling for Flow Control (Bernd R. Noack, Marek Morzyski, and Gilead Tadmor, eds.), CISM Courses and Lectures, vol. 528, Springer Vienna, 2011, pp. 111–149 (English).

[31] K. Pearson, *On lines and planes of closest fit to systems of points in space*, Philosophical Magazine **2** (1901), no. 6, 559–572.

[32] Giuseppe Rega and Hans Troger, *Dimension reduction of dynamical systems: Methods, models, applications*, Nonlinear Dynam. **41** (2005), 1–15, 10.1007/s11071-005-2790-3.

[33] W. Ring and B. Wirth, *Optimization methods on riemannian manifolds and their application to shape space*, SIAM J. Optim. **22** (2012), no. 2, 596–627.

[34] A. J. Roberts, *The utility of an invariant manifold description of the evolution of a dynamical system*, SIAM J. Math. Anal. **20** (1989), no. 6, 1447–1458.

[35] J. Robinson, *Some closure results for inertial manifolds*, J. Dynam. Differential Equations **9** (1997), 373–400, 10.1007/BF02227487.

[36] J. D. Rodriguez and L. Sirovich, *Low-dimensional dynamics for the complex Ginzburg-Landau equation*, Phys. D **43** (1990), no. 1, 77 – 86.

[37] O. E. Rössler, *An equation for continuous chaos*, Phys. Lett. A **57** (1976), 397–398.

[38] Sam T. Roweis and Lawrence K. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, Science **290** (2000), no. 5500, 2323–2326.

[39] Björn Schmalfuss and Klaus R. Schneider, *Invariant manifolds for random dynamical systems with slow and fast variables*, J. Dynam. Differential Equations **20** (2008), no. 1, 133–164.

[40] R. Seydel, *Practical bifurcation and stability analysis*, Interdisciplinary Applied Mathematics, Springer, 2010.

[41] A. C. Skeldon, *Dynamics of a parametrically excited double pendulum*, Phys. D **75** (1994), 541–558.

[42] N. Smaoui, *Linear versus nonlinear dimensionality reduction of high-dimensional dynamical systems*, SIAM J. Sci. Comput. **25** (2004), no. 6, 2107–2125.

[43] Alois Steindl and Hans Troger, *Methods for dimension reduction and their application in nonlinear dynamics*, Internat. J. Solids Structures **38** (2001), no. 1013, 2131 – 2147.

[44] S.H. Strogatz, *Nonlinear dynamics and chaos*, Studies in nonlinearity, Westview Press, 2008.

[45] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science **290** (2000), no. 5500, 2319–2323.

[46] G. Teschl, *Ordinary differential equations and dynamical systems*, Graduate studies in mathematics, American Mathematical Society, 2012.

[47] F. S. Tsai, *Comparative study of dimensionality reduction techniques for data visualization*, Journal of Artificial Intelligence **3** (2010), 119–134.

[48] F. W. Warner, *Foundations of differentiable manifolds and lie groups*, Graduate Texts in Mathematics, Springer, 1983.

[49] C. Welshman and J. Brooke, *Dimensionality reduction of dynamical systems with parameters*, AIP Conf. Proc. **1510** (2012), no. 1, 263–266.

[50] C. Welshman and J. Brooke, *Dimensionality reduction of dynamical systems with parameters: A geometric approach*, SIAM J. Appl. Dyn. Syst. **13** (2014), no. 1, 493–517.

[51] Zhenyue Zhang and Hongyuan Zha, *Principal manifolds and nonlinear dimensionality reduction via local tangent space alignment*, SIAM J. Sci. Comput. **26** (2004), 313–338.