Master's Thesis

Takens' Theorem: Proof and Applications

Jordi Penalva Vadell

Supervised by

María Jesús Álvarez
Antonio E. Teruel

Advanced Physics and Applied Mathematics
2018

RESUME

In this work, we prove the Takens' Embedding Theorem and we give a more general version of the Theorem.

Firstly, we write a plotline to understand the theorem and the proofs. We follow mainly some articles, such as [1]. In some cases, we explain more the proofs and in others, we give some alternative proofs. We prove the Takens' Embedding Theorem, as it appears in his article [2] and through it, we arrive at a more general result. This result appears in some references, but as far as we know, there is no explicit proof of this generalization. Finally, we make some applications to understand the use of the theorem. Some applications are theoretical, and others are practical. The theoretical experiments are given by dynamical systems and the practical are mainly from harmonic signals.

Index

Chapter 1

Introduction

Takens' Embedding Theorem is a very widespread result into time-series analyses and it is commonly used into a lot of branches, such as psychology [3], physics [4], biology [5] or economy [6], where it is natural the usage of time-series.

In this chapter, we give some motivation of the Takens' Embedding Theorem and an informal explanation about how it works.

In the theorem, we talk about topological manifolds. The first attempt to formalize the concept of 'manifold' appears formally with Riemann, the years 1851, with his doctoral thesis, and 1854 in *Habilitationsvortrag*. Some attempts to understand the concept of manifold arose through the following years. In particular, Weyl introduces a manifold as a set of points which are locally homeomorphic to some real space. There is another approach, given by Dini, which states that manifolds are subsets of a real space defined near each point by expressing some of the coordinates in terms of the others, using differential functions.

Hassler Whitney proved that both concepts are equivalent. He did not only prove that these concepts are the same, but that manifolds can be thought as subsets of some Euclidean space, and thus in a real space. However, this result does not give a reconstruction of the manifold in the real space: it is true that we can 'embed' the manifold in some $\mathbb{R}^n$, for n sufficiently large, but we do not have its parametrization.

In 1981, Floris Takens published its article [2]. In this work, he presented a map that allows us to embed the manifold in some real space. However, his main objective was not related to manifolds, but related to dynamical systems. In this context, Takens' Theorem allows us to reconstruct the attractor set of a dynamical system. If the attractor set is a manifold, giving only partial information from the system the theorem assures to recover the attractor set. Since Takens' article, there have appeared a lot of attempts to generalize their results. One of the most notable ones was from Sauer, Yorke and Casdagli [7], that generalizes the theorem to attractor sets which are not more manifolds, but are fractal sets.

To introduce briefly the theorem, let us say that we have some signal, and we assume that this signal comes from some dynamical system having an attractor set M.

For example, we simulate a Morris-Lecar's model with some mathematical software. We take it from [8]. In that book, the model is two dimensional. However, we add a fast variable to generate a buster. Moreover, the differential equation is a piecewise linear continuous map. Hence, the solution of the system is a $\mathcal{C}^1$ map. The equations are

$$\begin{cases} C\dot{V} = f(V) - u - w + I \\ \\ \dot{w} = \dfrac{w_\infty(V) - w}{\tau_w(V)} \\ \\ \dot{u} = \dfrac{u_\infty(V) - u}{\tau_u(V)} \end{cases}$$

where

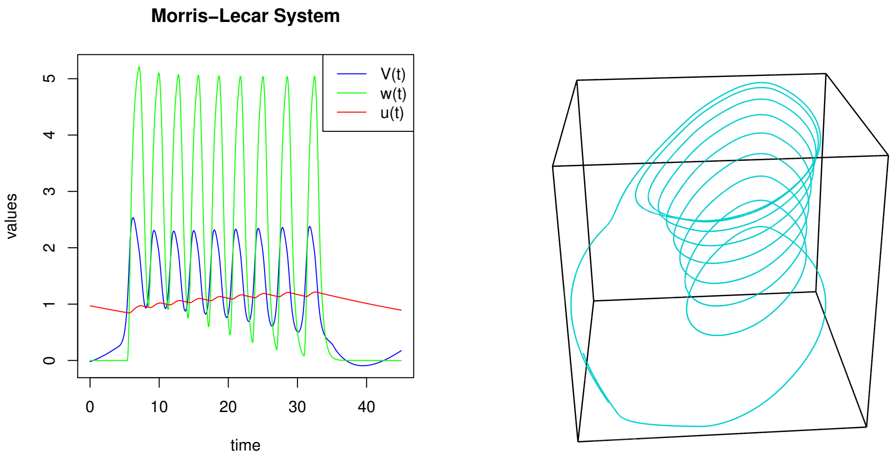
$$f(V) = \begin{cases} -V - d & V < -d, \\ m(V+d) & |V| \leq d, \\ k(V-d) + 2md & d < V < 2, \\ 2 - V + k(2-d) + 2md & V > 2 \end{cases}$$

$$w_\infty(V) = \begin{cases} 0 & V < 1, \\ s(V-1) & 1 < V < 2, \\ s & V > 2 \end{cases}$$

$$u_\infty(V) = \begin{cases} 0 & V < 1.5, \\ s(V-1.5) & 1.5 < V < 2.5, \\ s & V > 2.5 \end{cases}$$

$C = 1, d = 1/4$, $s = 5.5$, $m = -0.2$, $k = 2.5$, $I = 1.0571$, $\tau_w(V) = 0.5$ and $\tau_u(V) = 40$. This system has different attractor sets. In particular, a periodic orbit formed by small amplitude oscillations followed by one big amplitude oscillation. This orbit passes close to the point $(-0.02, 0, 0.97)$. Hence, starting at $(-0.02, 0, 0.97)$ we get three time-series, one for each variable. We can see them in Figure 1.1, jointly with its phase portrait.

We now focus on only one of these signals: for example V. From the integration method, we have some time-series with a temporal fixed step h. Let us take some



(a) (b)

Figure 1.1: Morris-Lecar system. 1.1(a) Solutions. 1.1(b) Phase portrait.

multiple $\tau = h \cdot p$. We apply the following function

$$\begin{array}{rccl} \Phi_\tau : & \mathbb{R} & \to & \mathbb{R}^3 \\ & V(t) & \mapsto & (V(t), V(t+\tau), V(t+2\tau)). \end{array}$$

We apply it for every $t = h \cdot \hat{p}$ from our signal, as long as is well-defined for $t + 2\tau$. We plot these 3-dimensional points and get some 3-dimensional figure. We can see the plot in Figure 1.2.

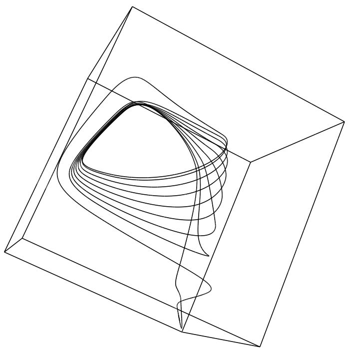


Figure 1.2: Φ_τ applied to V.

It's quite similar to the phase portrait. At first sight, it seems strange. From a simple map, and with a partial information of the system, we make a copy of the original one. In the manuscript, we talk about these kind of maps. Takens' Theorem tells us that, if the signal V has sufficiently information about its own system, we may 'reconstruct' it from the same V.

The contents of the manuscript are distributed as follows: in Chapter 2, we introduce the basic concepts concerning the theorem. In Chapter 3, we state the original version and we prove it, and other alternative versions. In Chapter 4, we give some examples of the usage of the theorem.

Chapter 2

Preliminary

In this chapter, we recall some basic notions which will be needed to introduce Takens' Theorem. First of all, we introduce some concepts about Topology. Secondly, we continue with Differential Topology. This branch is divided into two sections: Differential Topology and Function spaces. We finish this chapter with some basic concepts about Dynamical Systems.

2.1 General Topology

In this section, we recall some basic concepts about Topology. Most of them are classic in a Mathematical degree, but it is interesting to follow the plot that leads to Takens' Theorem. Moreover, for readers who have not follow a course on Topology they will find in this section the necessary concepts to understand and follow the main result. Nevertheless, the objective of these section is not to give a course on General Topology, but to give some definitions and the basic properties that we will use along the rest of the manuscript. If someone is interested in more details, I recommend the lectures of [9], [10] and [11]. In this section, we do not give any proof, since they appear in any book of Topology or they are intuitive.

First of all, we need to define the basic structure we work with:

Definition 2.1. Let $\mathcal{X}$ be a set. Let $\mathcal{T} \subseteq \mathcal{P}(\mathcal{X})$ be a subset of the powerset of $\mathcal{X}$ which satisfies:

(i) $\emptyset$, $\mathcal{X} \in \mathcal{T}$,

(ii) If U, $V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$,

(iii) If $\{U_i\}_{i\in I} \subset \mathcal{T}$, then $\cup_{i\in I} U_i \in \mathcal{T}$.

In this case, we say that $\mathcal{T}$ is a *topology* on $\mathcal{X}$. The elements of $\mathcal{T}$ are the *open sets* and the pair $(\mathcal{X}, \mathcal{T})$ is a *topological space*.

From the open sets, we can define the closed sets.

Definition 2.2. The *closed sets* of a topology are the complements of open sets. In other words, if $U \in \mathcal{T}$, where U is an open set from the topology $\mathcal{T}$, then its complementary U^c is a closed set.

We introduce some concepts that lead to define consistently a topology by a set.

Definition 2.3. A topology $\mathcal{T}$ on $\mathcal{X}$ is said to be *finer* than $\mathcal{T}'$ if $\mathcal{T}' \subset \mathcal{T}$. In this case, we also say that $\mathcal{T}'$ is *coarser*–or less fine– than $\mathcal{T}$. In the case that $\mathcal{T} \not\subset \mathcal{T}'$ and $\mathcal{T}' \not\subset \mathcal{T}$, we say that they are *incomparables*.

We know that intersection of topologies is also a topology.

Lemma 2.1. Consider two topologies $\mathcal{T}$ and $\mathcal{T}'$ on $\mathcal{X}$. Then, $\mathcal{W} = \mathcal{T} \cap \mathcal{T}'$ is a topology.

Lemma 2.1 allows us to prove the next result.

Lemma 2.2. Given a collection of subsets $\mathcal{S}$ of $\mathcal{X}$, there exists a topology $\mathcal{T}_{\mathcal{S}}$ such that for any other topology $\mathcal{T}$ such that $\mathcal{S} \subseteq \mathcal{T}$, then $\mathcal{T}_{\mathcal{S}} \subseteq \mathcal{T}$.

This result allows us to make a consistent definition.

Definition 2.4. Let $\mathcal{X}$ be a set and $\mathcal{S}$ a subset of the powerset. We define the *topology generated by* $\mathcal{S}$ as the coarser topology containing $\mathcal{S}$.

Sometimes, it is difficult to define explicitly the open sets of a topology. The most usual technique is using a set of generators that allows obtaining the open sets. The principal generators we are going to work with are bases and subbases.

Definition 2.5. A collection $\mathcal{S}$ of subsets of $\mathcal{X}$ is a *subbase* of a topological space $(\mathcal{T}, \mathcal{X})$ if $\mathcal{T}$ is a topology generated by $\mathcal{S}$.

Example 2.1. Consider the real space $\mathbb{R}$, and consider the topology generated by the subbase of all the open intervals $]a_i, b_i[$, with $a_i < b_i$. This is the usual topology that we will use working with the topological space $\mathbb{R}$.

It is possible to generalize the previous topology to $\mathbb{R}^n$ in the following way. Consider in $\mathbb{R}^n$ the topology generated by the subbase of all the open balls

$$B_\epsilon(x) = \{y \in \mathbb{R}^n : |x - y| < \epsilon\},$$

where $\epsilon > 0$, $x \in \mathbb{R}^n$ and $|\cdot|$ is the Euclidean norm.

Example 2.2. The *topological product* of a family of topological spaces $\{(\mathcal{X}_i, \mathcal{T}_i)\}_{i\in I}$ is the topological space

$$\prod_{i\in I}(\mathcal{X}_i, \mathcal{T}_i) = (\prod_{i\in I}\mathcal{X}_i, \prod_{i\in I}\mathcal{T}_i),$$

where $\prod_{i\in I}\mathcal{X}_i$ is the Cartesian product given by $(\mathcal{X}_i)_{i\in I}$ and $\prod_{i\in I}\mathcal{T}_i$ is the topology on $\prod_{i\in I}\mathcal{X}_i$ such that, if

$$\begin{array}{rccc} \pi_j : & \prod_{i\in I}\mathcal{X}_i & \to & \mathcal{X}_j \\ & (x_i)_{i\in I} & \mapsto & x_j \end{array}$$

is the j-projection, $j \in I$, then

$$\bigcup_{i\in I}\pi_i^{-1}(\mathcal{T}_i)$$

is a subbase of the topological product.

Definition 2.6. A *base* of a topological space $(\mathcal{X}, \mathcal{T})$ is a collection of open sets $\{B_i\}_{i\in I}$ of $\mathcal{X}$ such that for every open set U of $\mathcal{X}$, it can be written as an union of open sets of this base:

$$U = \bigcup_{j\in J} B_j.$$

The difference between bases and subbases is that the intersection of members of the base belongs to the base. However, it is not necessarily true in subbases.

Proposition 2.1. Let B be a collection of subsets of a set $\mathcal{X}$. B is a base of some topology on $\mathcal{X}$ if, and only if, it satisfies the following two properties:

(i) $\mathcal{X} = \bigcup \mathrm{B}$, that is, $\mathcal{X}$ is the union of all the open sets of B.

(ii) If $B_1, B_2 \in \mathrm{B}$, then $B_1 \cap B_2 = \cup_{i\in I} B_i$, where $B_i \in \mathrm{B}$, for some subset I.

Example 2.3. In $\mathbb{R}$, the set of intervals $B = \{]x, y[: \ x, y \in \mathbb{R}\}$, is a base of the first topology of the Example 2.1. It is because every intersection of open (not disjoint) intervals is also an open interval, and the union is the real space, so it is base of some topology.

It is the same for $\mathbb{R}^n$, because the set

$$B = \{\prod]x_i, y_i[: \ x_i, y_i \in \mathbb{R}\}$$

is also a topological base.

Definition 2.7. Let $(\mathcal{T}, \mathcal{X})$ be a topological space. Then $(\mathcal{X}, \mathcal{T})$ is second countable if it has some countable base.

It is usual to define second countable in terms of subbases, but it is equivalent to the previous one and along all the manuscript we only use this equivalence.

Example 2.4. Observe that the base given in Example 2.3 is not a countable base. However, if we take the set $\mathrm{B} = \{]p, q[: \ p, q \in \mathbb{Q}\}$, it is a countable base of the same topology. As in Example 2.3, we may say that it is a base. Moreover, if we have intervals $]a, b[$, with $a, b \in \mathbb{R}$, then we can choose two successions $\{a_i\}_{i\in\mathbb{N}} \in \mathbb{Q}$, $\{b_i\}_{i\in\mathbb{N}} \in \mathbb{Q}$, where $a_i \to a$, $b_i \to b$, $a < a_i < b_i < b$, for all i, and the union of opens $\cup_{i\in\mathbb{N}}]a_i, b_i[=]a, b[$. Hence, it generates the same intervals as in Example 2.1 and is a countable base, since it is generated by pairs $\{(p, q) : \ p, q \in \mathbb{Q}\}$ and $\mathbb{Q}$ is countable, so $\mathbb{Q}^2$ is also countable. The same argument is valid for $\mathbb{R}^n$.

Now, we continue with compact sets. It is necessary to introduce the covers.

Definition 2.8. A *cover* of a subset $\mathcal{S}$ of $\mathcal{X}$ is a collection of subsets $\mathcal{U} = \{U_j\}_{j\in J}$ in $\mathcal{X}$ such that $\mathcal{S} \subseteq \cup_{j\in J} U_j$. If J is finite, we say that $\mathcal{U}$ is a *finite cover*. If every U_j, $j \in J$ is an open set, then we say that $\mathcal{U}$ is an *open cover*.

Given two covers $\mathcal{U} = \{U_j\}_{j\in J}$ and $\mathcal{V} = \{V_i\}_{i\in I}$ of $\mathcal{S}$, if for every $j \in J$ there exists $i \in I$ such that $U_j = V_i$, then we say that $\mathcal{U}$ is a *subcover* of $\mathcal{V}$.

A covering $\{V_j\}_{j\in J}$ is a *refinement* of the covering $\{U_i\}_{i\in I}$ when each V_j is contained in some U_i.

Definition 2.9. We say that the subset $\mathcal{S}$ of $\mathcal{X}$ is *compact* if for every open cover of $\mathcal{S}$ there exists a finite subcover of $\mathcal{S}$.

Every closed subset of a compact set is also a compact set. We state this result as it is necessary, especially when we use partitions of unity.

Lemma 2.3. Let $\mathcal{X}$ be a topological space and A a compact set of $\mathcal{X}$. Every closed subset $B \subseteq A$ is also compact.

The next result is known as *Heine-Borel* Theorem. It characterizes the compact sets of a real topological space.

Theorem 2.1 (Heine-Borel)**.** A subset of $\mathbb{R}^n$ is compact if, and only if, it is a closed and bounded set.

Since in some proofs we use the sequentially compactness, we continue by introducing this concept.

Definition 2.10. A topological space $(\mathcal{X}, \mathcal{T})$ is said to be *sequentially compact* if every sequence in it has a convergent subsequence.

In general, compactenss and squentially compactness are different concepts. However, they are the same in metric spaces.

Proposition 2.2. In a metric space, compactness and sequentially compactness are equivalent.

This allows to prove the Lebesgue's Lemma. We give the proof in Appendix C. First we need to define the diameter.

Definition 2.11. Let $A \subset \mathcal{X}$ be a subset of a topological space. Its *diameter* $\operatorname{diam}(A)$ is the number $\sup\{d(x, y) : x, y \in A\}$.

Lemma 2.4 (Lebesgue's Lemma)**.** Let $\mathcal{X}$ be a compact metric space and let $\mathcal{U}$ be an open cover of $\mathcal{X}$. Then there exists a real number $\delta > 0$ such that any subset of $\mathcal{X}$ of diameter less than δ is contained in some member of $\mathcal{U}$. δ is called the *Lebesgue number* of $\mathcal{U}$.

Another tool that we will frequently use is the concept of Hausdorff space.

Definition 2.12. A topological space $(\mathcal{X}, \mathcal{T})$ is said to be *Hausdorff* if for every pair of points $x, y \in \mathcal{X}$ there exist open sets U_x and U_y such that $x \in U_x$, $y \in U_y$ and $U_x \cap U_y = \emptyset$.

Hausdorff spaces, jointly with second countable spaces, allow us to define the concept of manifold. Manifolds are the basic concept in all the Section 2.2.

We now define closure, denseness and interior.

Definition 2.13. Let A be a subset of a topological space $(\mathcal{X}, \mathcal{T})$ and $x \in \mathcal{X}$. Then, x is a *closure point* of A if for every open $x \in U_x$, we have $U_x \cap A \neq \emptyset$.

The *closure* of A is the set of all the closure points. We denote the closure by $\overline{A}$.

Proposition 2.3. Let A be a subset of a topological space $(\mathcal{X}, \mathcal{T})$.

(i) $\overline{A}$ is a closed subset of $\mathcal{X}$. Moreover, $\overline{A}$ is the smallest closed subset of the space $\mathcal{X}$ containing A.

(ii) A is closed in $\mathcal{X}$ if and only if $\overline{A} = A$.

(iii) $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ and $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Definition 2.14. A set A is *dense* in a topological space $(\mathcal{X}, \mathcal{T})$ if its closure is all the space; i.e. $\overline{A} = \mathcal{X}$.

We note that A is dense if, and only if, for every $x \in \mathcal{X}$ and every open $x \in U_x$, we have $U_x \cap A \neq \emptyset$, by the definition of closure. We usually use this definition.

Along the manuscript, we use some characteristic sets from the topological space.

Definition 2.15. A set A is *generic* in a topological space $(\mathcal{X}, \mathcal{T})$ when it is open and dense.

A generic set satisfies a lot of properties. In particular these two:

Proposition 2.4. Let A, B be two generic sets in a topological space $(\mathcal{X}, \mathcal{T})$. Then

- $A \cup B$ is generic.
- $A \cap B$ is generic.

Definition 2.16. Let $(\mathcal{X}, \mathcal{T})$ be a topological space and $p \in \mathcal{X}$. A *neighborhood* of p is a subset U of $\mathcal{X}$ that includes an open set V containing p,

$$p \in V \subseteq U.$$

Definition 2.17. Let $(\mathcal{X}, \mathcal{T})$ be a topological space, $x \in \mathcal{X}$ is an *interior point* of a subset A when A is a neighborhood of x.

The set of all elements that are interior of A make the *interior set* of A and we denote it by $\overset{\circ}{A}$.

We state in Proposition 2.5 the equivalent conditions of Proposition 2.3, but for interior sets.

Proposition 2.5. Let $(\mathcal{X}, \mathcal{T})$ be a topological space and $A \subseteq \mathcal{X}$.

(i) $\overset{\circ}{A}$ is the biggest open contained in A.

(ii) $A = \overset{\circ}{A}$ if and only if A is open.

(iii) $\overset{\circ}{A} \cup \overset{\circ}{B} \subset \overset{\circ}{\overbrace{A \cup B}}$ and $\overset{\circ}{A} \cap \overset{\circ}{B} = \overset{\circ}{\overbrace{A \cap B}}$.

Now, we talk about continuous applications between topological spaces.

Definition 2.18. Let $(\mathcal{X}, \mathcal{T})$ and $(\mathcal{Y}, \mathcal{T})$ be two topological spaces. A *continuous application* between $\mathcal{X}$ and $\mathcal{Y}$ is an application $f : \mathcal{X} \to \mathcal{Y}$ such that, for every element $x \in \mathcal{X}$, every neighborhood V of $f(x)$ in $\mathcal{Y}$ contains the image of some neighborhood U of x in $\mathcal{X}$.

From the previous definition we conclude that, if f is a continuous map, for every neighborhood V of $f(x)$, there exists a neighborhood $U \subset \mathcal{X}$ such that $f(U) \subset V$. This definition is useful. However, it is necessary to give some other equivalent definitions.

Proposition 2.6. Let $(\mathcal{X}, \mathcal{T})$ and $(\mathcal{Y}, \mathcal{T})$ be two topological spaces. An application $f : \mathcal{X} \to \mathcal{Y}$ is continuous if, and only if, the inverse image of an open set is also an open set. The same is true for closed set: that is, f is continuous if, and only if, the inverse image of a closed set is also closed.

Example 2.5. We can see the space of matrices with n^2, $n \in \mathbb{N}$ square real entries as a real space $\mathbb{R}^k$, with $k = n^2$. Consider the function given by the determinant

$$\begin{array}{rccc} \det : & \mathbb{R}^k & \to & \mathbb{R} \\ & x & \mapsto & \det x. \end{array}$$

Since det is a polynomial on the k entries, it is a continuous function. The point $0 \in \mathbb{R}$ is a closed set. Therefore, $\mathbb{R} \setminus \{0\}$ is an open set and thus $\det^{-1}(\mathbb{R} \setminus \{0\})$ is an open set of $\mathbb{R}^k$, by Proposition 2.6. Hence, the full rank matrices form an open set from the space of matrices.

The following equivalence is for metric spaces.[1]

Proposition 2.7. Let $(\mathcal{X}, \mathcal{T})$ and $(\mathcal{Y}, \mathcal{T})$ be two metric spaces. An application $f : \mathcal{X} \to \mathcal{Y}$ is continuous if, and only if, for every succession $\{x_n\} \to x$ in $\mathcal{X}$, then $\{f(x_n)\} \to f(x)$ in $\mathcal{Y}$.

Proposition 2.8. The composition of continuous applications is also a continuous application.

Example 2.6. Let $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be topological spaces. For every pair of applications $f : \mathcal{Z} \to \mathcal{X}$ and $g : \mathcal{Z} \to \mathcal{Y}$, the application

$$\begin{array}{rccl} h: & \mathcal{Z} & \to & \mathcal{X} \times \mathcal{Y} \\ & x & \mapsto & (f(x), g(x)) \end{array}$$

is continuous if, and only if, the functions f and g are continuous.

We use the fact that continuous functions preserve compactness in a lot of proofs.

Lemma 2.5. Let $\mathcal{X}$ and $\mathcal{Y}$ be two topological spaces (with their respective topologies) and $f : \mathcal{X} \to \mathcal{Y}$ a continuous application between these two topological spaces. If $\mathcal{S}$ is a compact subset of $\mathcal{X}$, then $f(\mathcal{S})$ is also a compact subset of $\mathcal{Y}$.

The interest of the general topology relies on finding properties which are preserved by applications between topological spaces. The most used are homeomorphisms.

Definition 2.19. An application f between two topological spaces $(\mathcal{X}, \mathcal{T})$ and $(\mathcal{Y}, \mathcal{T}')$ is an *homeomorphism* if f is bijective, continuous and the inverse f^{-1} is also continuous.

Properties which are preserved by homeomorphisms are called *topological properties*. The compactness, for example, is a topological property, since it is preserved by continuous applications.

2.2 Differential Topology

In this section, we recall some basic notions on Differential Topology. We start defining manifolds, a topological object indispensable to work with Takens' Theorem, and next we give some properties related with them.

[1] Actually, we can relax the condition, but we only use this on metric spaces.

In the widest sense, a manifold is defined as a topological space, which is locally homeomorphic to some $\mathbb{R}^n$. This classical definition is too weak. Therefore, it is habitual to define a manifold with other topological restrictions. In our case, we will consider the following definition, that is the most common one in the literature.

Definition 2.20. A *manifold* M of *dimension* m is a topological Hausdorff space, second countable, such that for every point on M, there exists an open neighborhood of the point that is homeomorphic to an open set of $\mathbb{R}^m$.

Sometimes, we say that M is a m-dimensional manifold. Observe that despite a manifold is locally homeomorphic to $\mathbb{R}^m$, it is not necessary to be globally homeomorphic to some $\mathbb{R}^m$.

Example 2.7.

– Every $\mathbb{R}^m$ space is itself a manifold. We only have to take $\mathbb{R}^m$ as an open cover and the identity $\mathrm{id} : \mathbb{R}^m \to \mathbb{R}^m$ as the homeomorphism. This is a straightforward case of a manifold that is globally homeomorphic to $\mathbb{R}^m$.

– The circumference

$$S^1 = \{f(t) = (cos(t), sin(t)) \in \mathbb{R}^2 : t \in \mathbb{R}\}$$

is a 1−dimensional manifold. the image of the open sets $U_1 =]-\pi, \pi[$ and $U_2 =]0, 2\pi[$ by f are charts that cover all the graph. However, it is not homeomorphic to $\mathbb{R}$. We have not introduced connected spaces, but it is a property that connected spaces are preserved by homeomorphisms. In this case, if we remove one point from the circumference, the graph is connected, but if we take off a point of the real line, we obtain two connected components. For more details about connected spaces, see any book on general topology.

(a) (b)

Figure 2.1: 2.1(a) One component. 2.1(b) Two components.

– The eight '8' curve is not a manifold. Since it is a curve, it should be a 1-dimensional manifold. However, if we choose the intersection point, any neighborhood of this point is homeomorphic to $\mathbb{R}$. In fact, if we take a neighborhood of the point, and after that we remove this point, we obtain four components, except in the case when we take all the graph, that we obtain two components. In this case, if we remove another point, we obtain only one component, thus it is neither homeomorphic to $\mathbb{R}$ (see Figure 2.2).

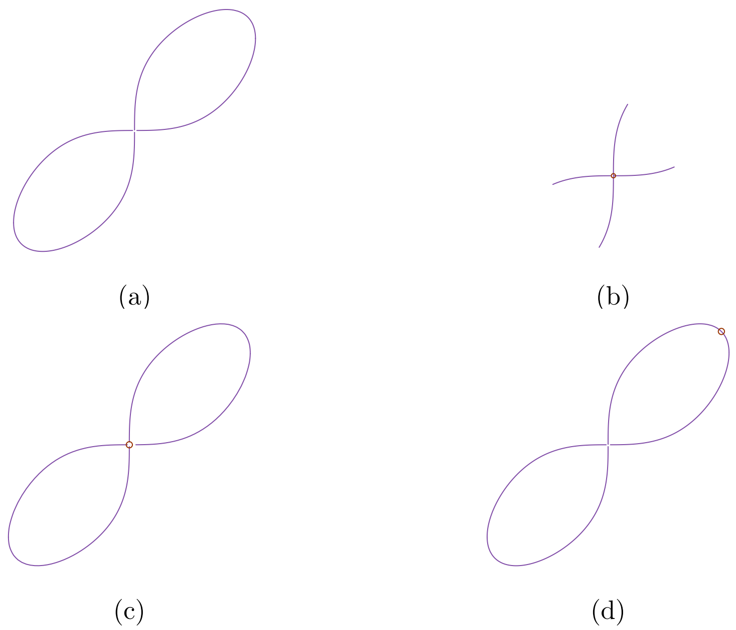


Figure 2.2: In 2.2(a), we have the '8' curve. In 2.2(b), we have some neighborhood of the center and if we remove the center point, we have 4 components and hence it is not homeomoprhic to $\mathbb{R}$. In 2.2(c), we remove again the center point, but since the figure is complete, we obtain two components. If we want to see that it is not homeomorphic to $\mathbb{R}$, we use another point, such as in figure 2.2(d), where we only have 1 component.

Definition 2.21. A *local chart* (or a *chart*) is a pair $(\mathcal{U}, h)$, such that $\mathcal{U} \subset M$ is an open set and $h : \mathcal{U} \to \mathbb{R}^n$ is an homeomorphism into its image. The set $\mathcal{U}$ is called *domain*.

An *atlas* is a collection of local charts such that their domains cover M.

Atlas allow us to cover all the manifold, even though M is not homeomorphic to $\mathbb{R}^m$. For example, to cover a circumference we need two or more local charts. The function h is composed by n functions, say $\mu_i : \mathcal{U} \to \mathbb{R}$, $i = 1, \ldots, n$ such that $h(x) = (\mu_1(x), \ldots, \mu_n(x))$. These functions are called *coordinate charts*.

Definition 2.22. Let N be a n-dimensional manifold and $M \subseteq N$ a m-dimensional manifold, with $m \leq n$. It is said that M is a *submanifold* of N if every point in M has a local chart which can be obtained from a chart (V, g) of N, by restricting g to $V \cap M$ and dropping the last $n - m$ coordinates.

Example 2.8. By using the previous definition, one can check that the sphere of radius 1, $\mathcal{S}^1$, is a 2-dimensional manifold. We take two partial parametrizations of the sphere, given by

$$\begin{cases} x(\theta, \phi) = \cos\theta \cos\phi, \\ y(\theta, \phi) = \cos\theta \sin\phi, \\ z(\theta, \phi) = \sin\theta. \end{cases}$$

The first one for $\theta \in]-\pi/2, \pi/2[$, $\phi \in]0, 2\pi[$ and the second one θ in the same interval and $\phi \in]-\pi, \pi[$. If we fix $\theta = 0$, we have $x = \cos\phi$, $y = \sin\phi$ and $z = 0$ and we get the circumference of Example 2.7. Thus, the circumference is a submanifold of the sphere.

If two local charts $(\mathcal{U}, h)$ and $(\mathcal{V}, g)$ share domains (in other words, $\mathcal{U} \cap \mathcal{V} \neq \emptyset$), the transformations

$$hg^{-1} : g(\mathcal{U} \cap \mathcal{V}) \to \mathbb{R}^m$$
$$gh^{-1} : h(\mathcal{U} \cap \mathcal{V}) \to \mathbb{R}^m$$

are functions from open sets of $\mathbb{R}^m$ to open sets of $\mathbb{R}^m$, because $\mathcal{U}$ and $\mathcal{V}$ are open sets and and h, g homeomorphisms.

Definition 2.23. Consider M a manifold and the local charts described previously. If hg^{-1} and gh^{-1} are r times differentiable, we say that the charts are $\mathcal{C}^r-$*related*. The set of charts that are $\mathcal{C}^r-$related gives a *differential structure*. An atlas with all the charts $\mathcal{C}^r-$related is an atlas $\mathcal{C}^r-$*differentiable*. In this case, the differential structure is in the whole atlas and it gives a $\mathcal{C}^r-$*manifold*.

With the previous definitions, we have classified manifolds with respect to their differentiability. In the following definition, we classify applications between two $\mathcal{C}^r$-manifolds.

Definition 2.24. Let M, N be two $\mathcal{C}^r$ manifolds m and n dimensional, respectively. Let $f : M \to N$ be a function between these two $\mathcal{C}^r$ manifolds. The function f is $\mathcal{C}^s-$*differentiable* (with $s \leq r$) if, for every point $p \in M$, there are local charts $(\mathcal{U}, h)$ of M with $p \in \mathcal{U}$, and $(\mathcal{V}, g)$ of N with $f(p) \in \mathcal{V}$, such that $gfh^{-1} : h(\mathcal{U} \cap f^{-1}V) \to \mathbb{R}^n$ is s times continuously differentiable at $h(p)$.

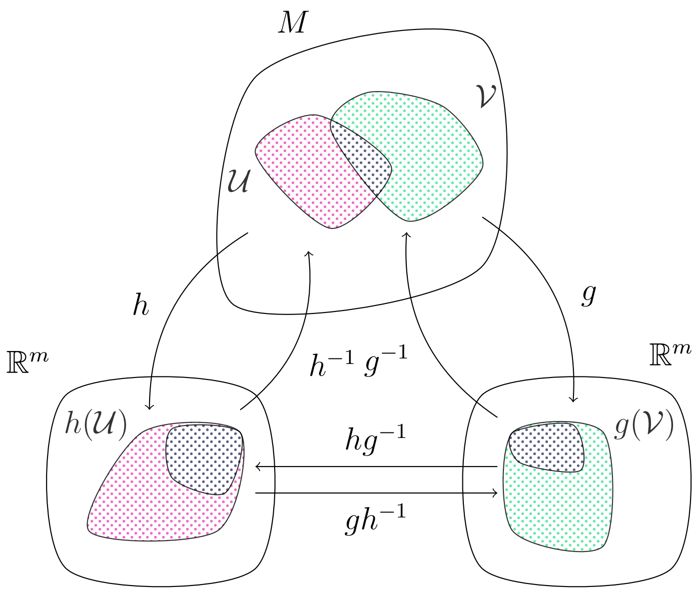


Figure 2.3: Manifold diagram.

In general, manifolds are at first sight topological spaces, but we do not know if there is an environment where the manifold lives in. Recall that a manifold is not necessarily homeomorphic to $\mathbb{R}^n$, hence we cannot assume that it is embedded in any $\mathbb{R}^n$ (this problem will be solved by Whitney's Theorem, see Theorem E.1 in Appendix E). We classify f from gfh^{-1}, since gfh^{-1} is an application between real spaces. Now, we set a property that preserves the topological and geometric properties between manifolds.

Definition 2.25. Let $f : M \to N$ be a function between two $\mathcal{C}^r$-manifolds m and n dimensional, respectively, and we assume that f is $\mathcal{C}^s$-differentiable, $1 \leq s \leq r$. Let $(\mathcal{U}, h)$ be a local chart of M and $p \in \mathcal{U}$. We say that f is an *immersion* at p if $Dgfh^{-1}(h(p))$ is injective. If the derivative is surjective at p then f is said to be *submersive* at p. A function that is immersive everywhere is an *immersion*, and a function that is submersive everywhere is a *submersion*.

A function f is an *embedding* if it is immersive and furthermore, f is homeomorphic to its image, i.e. $f : M \to f(M)$ is an homeomorphism.

Embeddings preserve all topological and geometrical properties. At first sight, it seems difficult to distinguish between embedding and immersion. The following three results explain their differences.

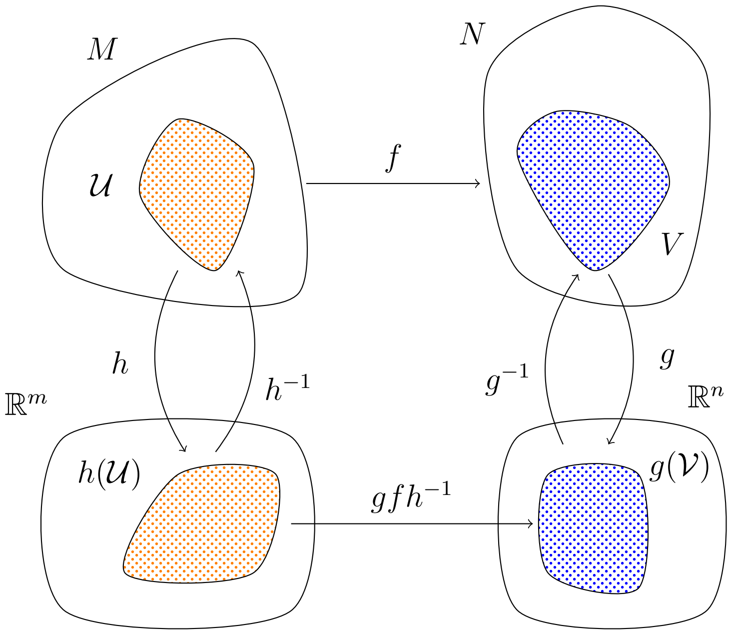


Figure 2.4: Manifold diagram.

Proposition 2.9. Let M, N be two compact manifolds (with dimensions m and n, respectively) and $f : M \to N$ an immersion. If f is injective, then f is also an embedding.

Proof. Let f be a continuous application between two manifolds M and N, where f is an injective immersion. Since M is compact, every closed subset $A \subseteq M$ is compact, and the image of compact sets through a continuous application is also compact. In this case, this compact image is carried homeomorphically through some finite charts to $\mathbb{R}^n$, because N is compact and it allows a finite atlas. Their image is a closed subset, because compacts are closed sets in $\mathbb{R}^n$, by Theorem 2.1. Since the charts are homeomorphisms, every subset of the compact set is closed and the finite union of closed sets is also closed. Therefore, $f(A)$ is a closed set and thus f is a closed map. Hence, f^{-1} is continuous. □

Proposition 2.9 tells us that the only difference between an immersion and an embedding in compact manifolds is the injectivity. In $\mathbb{R}$ the two concepts are equivalent.

Proposition 2.10. Let $f : \mathbb{R} \to \mathbb{R}$ be immersive. Then f is an embedding.

Proof. We only have to choose the local chart $(\mathbb{R}, \mathrm{id})$. Then the derivative is always different from zero. In this case, f is continuous and monotone. Now, if it is monotone, then it is injective and, in applications between real spaces, injectivity assures an inverse continuous everywhere, hence f is an homeomorphism into its image. □

Finally, we see that immersions are locally embeddings. This argument is very common in differential topology.

Proposition 2.11. Let M, N be two manifolds (with dimensions m and n, respectively), $p \in M$ and $f : M \to N$ immersive at p. Then there exists a neighborhood of p such that f is an embedding.

Proof. Since f is immersive, we can choose local charts (U, h), (V, g) such that $Dgfh^{-1}(h(p))$ is injective: that is, the derivative has full rank. By the Inverse Function Theorem, we have that gfh^{-1} is a local diffeomorphism at some neighborhood of $h(p)$, thus f is an homeomorphism at some neighborhood and then an embedding. □

We include two results related to functions between manifolds. We are going to use these two results at various points. They will be proved in the Appendices A and B respectively, as their proof is long.

Lemma 2.6. Let M and N be manifolds with dimensions m and n respectively, $m < n$. If $f : M \to N$ is a $\mathcal{C}^1$ function, then $N \setminus f(M)$ is dense in N.

Lemma 2.7. Let M and N be manifolds with dimensions m and n respectively, $m > n$, and $f : M \to N$ be a $\mathcal{C}^1$ function. Consider $q \in N$. If f is submersive at every p such that $f(p) = q$, then the set $f^{-1}(q)$ is a submanifold of M, of dimension $m - n$.

We want now to introduce the bump functions. These functions are a basic tool in differential topology. Consider the function

$$\begin{array}{rccl} \lambda : & \mathbb{R} & \to & \mathbb{R} \\ & t & \mapsto & \begin{cases} 0 & \text{if } t \leq 0, \\ e^{-t^{-2}} & \text{if } t > 0. \end{cases} \end{array}$$

This function is $\mathcal{C}^\infty$, since it is a composition of $\mathcal{C}^\infty$ functions everywhere except at 0. In this case, for $t > 0$, the derivatives of λ are $q(t)e^{-t^{-2}}$, where $q(t)$ is a rational function. Therefore $q(t)e^{-t^{-2}} \to 0$ as $t \to 0$. Consequently

$$\lim_{t \to 0^+} \frac{d^n}{dt^n}\lambda(t) = \lim_{t \to 0^-} \frac{d^n}{dt^n}\lambda(t) = 0 = \frac{d^n}{dt^n}\lambda(0)$$

and it is well defined.

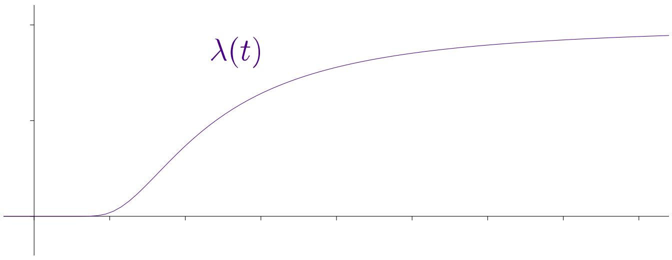


Figure 2.5: Representation of the λ function described as above.

Note that $0 \leq \lambda(t) \leq 1$, because $e^{-t^{-2}}$ is monotone for $t \geq 0$, $\lim_{t\to 0} e^{-t^{-2}} = 0$ and $\lim_{t\to\infty} e^{-t^{-2}} = 1$. As a result, $\lambda(t) = 0$ if and only if $t \leq 0$.

Consider $\epsilon > 0$ and the function

$$\phi_\epsilon(t) = \lambda(t) \cdot (\lambda(t) + \lambda(\epsilon - t))^{-1}.$$

Observe that

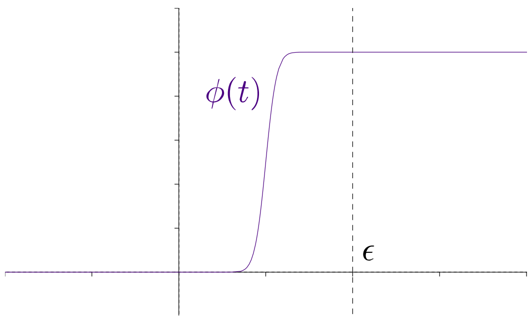


Figure 2.6: Representation of the ϕ function described as above.

$$\lambda(t) + \lambda(\epsilon - t) = \begin{cases} e^{-(\epsilon-t)^{-2}} & t < 0, \\ e^{-(\epsilon-t)^{-2}} + e^{-t^{-2}} & 0 \leq t < \epsilon, \\ e^{-t^{-2}} & t \geq \epsilon. \end{cases}$$

As a result, $\lambda(t) + \lambda(\epsilon - t) \neq 0$ and therefore $(\lambda(t) + \lambda(\epsilon - t))^{-1}$ is a $\mathcal{C}^\infty$ map. Consequently, $\phi_\epsilon(t)$ is also $\mathcal{C}^\infty$, as it is the product of $\mathcal{C}^\infty$ functions. It is clear that $\phi_\epsilon(t) > 0$, because every function in its domain of definition is positive and we make products and sums. Moreover,

$$0 < \lambda(\epsilon - t) \iff \lambda(t) < \lambda(\epsilon - t) + \lambda(t) \iff \lambda(t) \cdot (\lambda(t) + \lambda(\epsilon - t))^{-1} < 1.$$

Hence, $0 \leq \phi_\epsilon \leq 1$ and $\phi_\epsilon(t) = 0$ when $\lambda(t) = 0$ and that is when $t \leq 0$. Likewise, when $t \geq \epsilon$, we have

$$\phi_\epsilon(t) = \frac{\lambda(t)}{\lambda(t)} = 1.$$

Thus, $\phi_\epsilon(t) = 1$ if and only if $t \geq \epsilon$.

Definition 2.26. Let $\epsilon, r > 0$. We define a *bump function* $\psi_{\epsilon,r} = \psi$ as

$$\begin{array}{rccl} \psi: & \mathbb{R}^n & \to & \mathbb{R} \\ & x & \mapsto & \psi(x) = 1 - \phi_\epsilon(|x| - r), \end{array}$$

where $|\cdot|$ is a norm, typically the Euclidean norm.

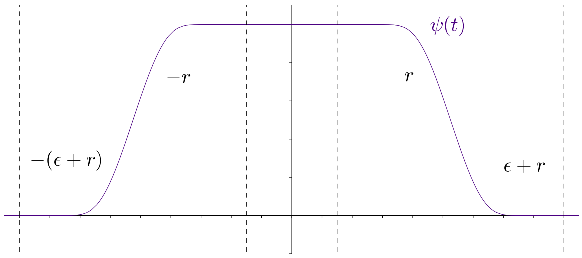


Figure 2.7: Representation of the $\psi : \mathbb{R} \to \mathbb{R}$, described as in Definition 2.26.

Proposition 2.12. Let $\psi_{\epsilon,r}(x) = \psi(x)$ be a bump function. Then:

(i) $0 \leq \psi(x) \leq 1$, for all $x \in \mathbb{R}^n$,

(ii) $\psi(x) = 1$ if and only if $x \in \mathrm{B}(r)$, where

$$\mathrm{B}(r) = B(0, r) = \{x \in \mathbb{R}^n : |x| < r\},$$

(iii) $\psi(x) = 0$ if and only if $|x| \geq r + \epsilon$,

(iv) $\psi \in \mathcal{C}^\infty$.

Proof. (i) It is clear, because $0 \leq \phi_\epsilon \leq 1$ and then $\psi(x) = 1 - \phi_\epsilon(|x| - r) \in [0, 1]$.

(ii) In this case,

$$\psi(x) = 1 - \phi_\epsilon(|x| - r) = 1 \Leftrightarrow \phi_\epsilon(|x| - r) = 0 \Leftrightarrow |x| - r \leq 0 \Leftrightarrow |x| \leq r \Leftrightarrow x \in \mathrm{B}(r).$$

(iii) Similarly

$$\psi(x) = 1 - \phi_\epsilon(|x| - r) = 0 \Leftrightarrow \phi_\epsilon(|x| - r) = 1 \Leftrightarrow |x| - r \geq \epsilon \Leftrightarrow |x| \geq r + \epsilon.$$

(iv) Since ϕ_ϵ is differentiable, we shall concern with differentiability when $|x| = 0$ (that is, x is the zero on $\mathbb{R}^n$). However, in $-r$ the function ϕ_ϵ is locally constant. Therefore, at this point the function will be always differentiable.

□

Bump functions have the property of compact support. This property allows us to shift functions locally.

Definition 2.27. The *support* of a function $f : \mathcal{X} \to \mathbb{R}^n$ is the closure of the set of values where the function does not vanish. Sometimes we will write as $\operatorname{supp} f$,

$$\operatorname{supp} f := \overline{\{x \in \mathcal{X} : f(x) \neq 0\}}.$$

Corollary 2.1. Every bump function has compact support.

Proof. It follows from Proposition 2.12, since $\psi(x) \neq 0$ if and only if $|x| < r + \epsilon$ and its closure is compact, by Theorem 2.1. □

Another tool is the partition of unity. Partitions of unity allow us to go from local to global properties. For more details on partitions of unity, see [12].

Definition 2.28. A collection of subsets $\{U_i\}_{i \in I}$ of a topological space $\mathcal{X}$ is called *locally finite* if for each point $x \in \mathcal{X}$ there exists a neighborhood V intersecting only finitely many U_i.

Lemma 2.8. Any open covering $\{A_j\}_{j \in J}$ of a m-dimensional manifold M has a countable, locally finite refinement $\{(U_i, h_i)\}_{i \in I}$ by local charts such that

(i) $h_i(U_i) = B(0,3)$ and

(ii) $\{V_i = h_i^{-1}(B(0,1))\}$ is still covering of M.

Proof. From the second countability property of M, we assume that there exists some open base of the topology of M that is countable. Since it is an open base, every open set of the manifold can be written as the union of elements of the base. As a result, we can make some chart basis $\{U_i\}_{i \in I}$, intersecting them, if necessary, where every chart will be the union of open sets $\{P_i\}_{i \in I}$, with $\overline{P_i}$ compact. Since U_j is a chart, it could be identify as some $\mathbb{R}^m$, and every $\mathbb{R}^m$ admits an open basis of balls $\{B_l\}_{l \in L}$, where every $\overline{B_l}$ is compact. If we take the balls around the rationals with rational radius, this base is countable and we can consider that $\{P_i\}_{i \in I}$ is countable.

We define an increasing sequence of compact sets $\{K_i\}_{i \in I}$ as follows

– $K_0 = \emptyset$,

– $K_1 = \overline{P_1}$, and

– $K_{i+1} = \overline{P_1 \cup \cdots \cup P_r}$, where $i \leq 1$ and $r > 1$ is the first integer such that $K_i \subset P_1 \cup \cdots \cup P_r$.

We note that

- $M = \bigcup_{i \in I} K_i \setminus \overset{\circ}{K}_{i-1} = K_1 \cup (K_2 \setminus \overset{\circ}{K}_1) \cup \ldots$. It is clear, because we can say that $K_i = K_{i-1} \cup (K_i \setminus \overset{\circ}{K}_{i-1})$; the left-right inclusion is evident, because $K_{i-1} \subset K_i$ and $K_i \setminus \overset{\circ}{K}_{i-1} \subset K_i$ and then their union is also a subset of K_i. For the other inclusion, we have the chain $\overset{\circ}{K}_{i-1} \subset K_{i-1} \subset K_i$. If $p \in K_i$, there are two options: $p \in K_{i-1}$ or $p \notin K_{i-1}$. In the first case there is no problem. In the second case, $p \notin \overset{\circ}{K}_{i-1}$ and then $p \in K_i \setminus \overset{\circ}{K}_{i-1}$.
- Every $K_i \setminus \overset{\circ}{K}_{i-1}$ is compact, because K_i is a closed set, $(\overset{\circ}{K}_{i-1})^c$ is also a closed set that can be written as

$$K_i \setminus \overset{\circ}{K}_{i-1} = K_i \cap (\overset{\circ}{K}_{i-1})^c.$$

  Then it is a closed subset of a compact set.
- With a similar argument, we could write

$$M = \bigcup_{i \in I} K_{i+2} \setminus \overset{\circ}{K}_{i-1}$$

  where the sets $K_{i+2} \setminus \overset{\circ}{K}_{i-1}$ are compacts.

If $p \in A_j$, then $p \in K_{i+2} \setminus \overset{\circ}{K}_{i-1}$ for some i, since their union covers M. Consider a local chart $\{(U_{p,j}, h_{p,j})\}$, with $U_{p,j} \subset (K_{i+2} \setminus \overset{\circ}{K}_{i-1}) \cap A_j$ and $h_{p,j}(U_{p,j}) = B(0,3)$. If we vary p and j, we can cover with open sets the compact set $K_{i+1} \setminus \overset{\circ}{K}_i$. Consider, for every p and j, the subset $V_{p,j} = h^{-1}(B(0,1))$. The $\{V_{p,j}\}_{p \in A_j}$ are subsets of $U_{p,j}$ that also cover $K_{i+1} \setminus \overset{\circ}{K}_i$.

Varying p and j, for every i we have cover sets $\{V_{i,p,j}\}_{p,j}$ such that admit a finite base of covers $\{V_{i_{(n)},p,j}\}_{p,j}$ for $K_{i+1} \setminus \overset{\circ}{K}_i$. The same argument applies with the covers $\{U_{i,p,j}\}_{p,j}$. These sets are a locally finite refinement, because there are finite covers on every band $K_i \setminus \overset{\circ}{K}_{i-1}$, and they only intersect with other finite covers $K_{i+1} \setminus \overset{\circ}{K}_i$ and $K_{i-1} \setminus \overset{\circ}{K}_{i-2}$; hence, on every point p we only have finitely many open sets. □

We denote the locally finite refinement described in Lemma 2.8 as a *regular covering*.

Definition 2.29. A *partition of unity* of a manifold M is a collection of functions $f_j : M \to [0,1]$, $j \in \Lambda$, such that

(i) $\{\operatorname{supp} f_j = \overline{f_j^{-1}(\mathbb{R} \setminus \{0\})}\}$ is locally finite,

(ii) $\sum_{j\in\Lambda} f_j(p) = 1$, $\forall p \in M$.[2]

A partition of unity is *subordinate* to an open cover $\{U_i\}_{i\in I}$ when $\forall j \in \Lambda$, $\operatorname{supp} f_j \subseteq U_i$ for some i.

Theorem 2.2. Given a regular covering $\{(U_i, h_i)\}$ of a manifold, there exists a partition of unity $\{f_i\}_{i\in I}$ subordinate to it with $f_i > 0$ on $V_i = h_i^{-1}(B(0,1))$ and $\operatorname{supp} f_i \subseteq h_i^{-1}(\overline{\mathrm{B}(0,2)})$.

Proof. Consider a bump function $\psi_{1,1} = \psi$. That is, $\psi(p) = 1$ for $|p| \leq 1$ and $\psi(p) = 0$ for $|p| \geq 2$ (Proposition 2.12). We can define the bump function on the manifold using the local charts (U_i, h_i) in the following way:

$$\begin{array}{rcccccl} \psi \circ h_i = \psi_i : & M & \to & \mathbb{R}^m & \to & \mathbb{R} \\ & p & \mapsto & h_i(p) & \mapsto & \psi(h_i(p)) = \psi_i(p). \end{array}$$

We have

- $\operatorname{supp} \psi_i \subseteq h_i^{-1}(\overline{B(0,2)})$:

$$\begin{aligned} \operatorname{supp} \psi_i &\overset{\text{def}}{=} \operatorname{supp} \psi \circ h_i \overset{\text{def supp}}{=} \overline{(h_i^{-1} \circ \psi^{-1})(\mathbb{R} \setminus \{0\})} \\ &\overset{(i)}{=} \overline{h_i^{-1}(B(0,2))} \overset{(ii)}{\subseteq} h_i^{-1}(\overline{B(0,2)}). \end{aligned}$$

  In (i), we know that $\psi^{-1}(\mathbb{R} \setminus \{0\}) = B(0,2)$. In (ii), $\overline{h_i^{-1}(B(0,2))}$ is the smallest closed set that contains $h_i^{-1}(B(0,2))$ and $h_i^{-1}(\overline{B(0,2)})$ is a closed set (since it is the inverse image of a closed set by a continuous function) that contains $h_i^{-1}(B(0,2))$.

- $\forall p \in \overline{V_i}$, $\psi_i(p) = 1$. We know that $V_i = h_i^{-1}(B(0,1))$. If $p \in \overline{h_i^{-1}(B(0,1))}$, then $p \in h_i^{-1}(\overline{B(0,1)})$, like in (ii). Hence, for all $p \in \overline{V}_i$, $h_i(p) \in \overline{B(0,1)}$ and therefore $|p| \leq 1$;

$$\psi_i(p) = \psi(h_i(p)) = 1.$$

[2]Since the support is locally finite, for every x, there are finitely $f_j(x)$ non-zero and hence the sum is finite.

Finally, we can define the partition of unity as

$$f_i = \frac{\psi_i}{\sum_{j\in I}\psi_j},\ i \in I.$$

and this is well defined:

- $f_i \geq 0$, since every $\psi_i \geq 0$.
- $\operatorname{supp} f_i = \operatorname{supp}\psi_i$, because f_i and ψ are non-zero in the same domain. Furthermore, $\operatorname{supp}\psi_i$ is locally finite, since $\operatorname{supp}\psi_i \subseteq \psi_i^{-1}(\overline{B(0,2)}) \subset h_i^{-1}(B(0,3))$, and therefore at every point $\operatorname{supp}\psi_i$ only intersects at most the finite ψ_j that h_i intersects with h_j, as $\{h_i\}_{i\in I}$ is locally finite.
- $\displaystyle\sum_{i\in I} f_i = \sum_{i\in I}\frac{\psi_i}{\sum_{j\in I}\psi_j} = \frac{\sum_{i\in I}\psi_i}{\sum_{j\in I}\psi_j} = 1.$
- $f_i > 0$ on V_i, since in $\overline{V}_i$, $f_i = 1$.

□

2.3 Function Spaces

In this section, we define some topologies for special function spaces. In the literature, there are two main topologies defined for function spaces; the weak topology and the strong topology. In our case, we use the weak topology. However, in some spaces these two topologies are equivalent, so that the obtained results could be also applied there. Part of this section is strongly inspired in [13].

First of all, we introduce the set of linear maps, since it is often referred along the manuscript. For example, we use the differential map that is a linear map.

Definition 2.30. We write the set of all linear applications between two vecctor spaces $\mathcal{X}$ and $\mathcal{Y}$ as $\mathcal{L}(\mathcal{X},\mathcal{Y})$.

Example 2.9. Let $\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$ be the set of linear applications between these two real spaces. If $f \in \mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$, then

$$\begin{array}{rccl} f: & \mathbb{R}^m & \to & \mathbb{R}^n \\ & x & \mapsto & f(x) = Ax. \end{array}$$

where A is the matrix associated to the linear map f in some basis.

From now on, it is considered that $\mathcal{X}$ and $\mathcal{Y}$ have finite dimension.

We will need to define a norm for this space.

Definition 2.31. We define the operator norm of a linear map $f \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ as

$$\|f\| := \sup\{|f(x)| : x \in \mathcal{X}, \|x\| \leq 1\}.$$

As we only work with norms between real spaces, we can alternatively define the norm of the linear map in terms of the associated matrix, that is

$$\|A\| := \|f\| = \sup\{|f(x)| : x \in \mathcal{X}, \|x\| \leq 1\} = \sup\{\|Ax\| : x \in \mathcal{X}, \|x\| \leq 1\}.$$

It is easy to prove that the definition of $\|A\|$ does not depend on the bases of the vector space. We use some equivalences.

Proposition 2.13. Let $f \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. The following sets are the same:

(i) $\sup\{\|f(x)\| : x \in \mathcal{X},\ \|x\| \leq 1\}$,

(ii) $\sup\{\|f(x)\| : x \in \mathcal{X},\ \|x\| = 1\}$,

(iii) $\sup\{\frac{\|f(x)\|}{\|x\|} : x \in \mathcal{X},\ \|x\| \neq 0\}$.

Injective linear maps are of especial interest, because they form an open set inside the set of linear maps. Before proving this, we prove the next result.

Proposition 2.14. A linear map $f \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is injective if, and only if, there exists $\alpha > 0$ such that $\|f(x)\| \geq \alpha\|x\|$, for all $x \in \mathcal{X}$.

Proof. $\Rightarrow$) Let f be an injective linear map. Then, f is bijective into its image. Hence we can consider the inverse f^{-1}, that will be also an injective linear map. In this case, since $f \not\equiv 0$, $\|f\|$ and $\|f^{-1}\|$ are non-zero. Observe that, by Proposition 2.13,

$$\|f^{-1}\| = \sup\{\frac{\|f^{-1}(f(x))\|}{\|f(x)\|} : f(x) \in f(\mathcal{X}),\ \|f(x)\| \neq 0\}.$$

Hence, for all $x \in \mathcal{X}$

$$\|f^{-1}\| \geq \frac{\|f^{-1}(f(x))\|}{\|f(x)\|} \Longrightarrow \|f(x)\| \geq \frac{1}{\|f^{-1}\|}\|x\|.$$

$\Leftarrow$), Let $\|f(x)\| \geq \alpha\|x\|$, for all $x \in \mathcal{X}$. We only have to see that the kernel is the trivial one. Therefore,

$$f(x) = 0 \Rightarrow \alpha\|x\| \leq \|f(x)\| = 0 \Rightarrow \alpha\|x\| \leq 0.$$

Then, as $\alpha > 0$ and $\|x\| \geq 0$, $\alpha\|x\| \geq 0$ and therefore $\alpha\|x\| = 0$. This is possible only if $\|x\| = 0$; that is $x = 0$. □

Proposition 2.14 leads to the next theorem:

Theorem 2.3. The set of all injective linear transformations $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is open.

Proof. Let $f \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be injective. Then, by Proposition 2.14, there exists $\alpha > 0$ such that $\|f(x)\| \geq \alpha\|x\|$. Choose in $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ an open ball $\mathrm{B} = B(f, \alpha/2)$. If $g \in \mathrm{B}$, then $\|f - g\| < \alpha/2$. We want to find $\delta > 0$ such that $\|g(x)\| > \delta\|x\|$.

$$\alpha\|x\| \leq \|f(x)\| = \|(f - g + g)(x)\| = \|(f - g)(x) + g(x)\| \leq \|(f - g)(x)\| + \|g(x)\|.$$

Isolating the operator norm of $g(x)$, we have $\|g(x)\| \geq (\alpha - \|f - g\|)\|x\|$. Since $\alpha/2 > \|f - g\|$, we can take $\delta = \alpha - \|f - g\| > 0$. □

Now, we define the weak and strong topology. Let M and N be $\mathcal{C}^r$-manifolds, with $r > 0$ and finite. We write

$$\mathcal{C}^r(M, N) = \{f : M \to N : f \in \mathcal{C}^r\text{-differentiable}\}.$$

Let $f \in \mathcal{C}^r(M, N)$ and take two local charts (U, h) and (V, g) of M and N, respectively. Choose some compact set $K \subset U$ such that $f(K) \subset V$. Let $0 < \epsilon \leq \infty$. We define a *weak subbasic neighborhood*

$$\mathcal{N}^r(f; (U, h), (V, g), K, \epsilon) \tag{2.1}$$

as the set of functions $\tilde{f} \in \mathcal{C}^r(M, N)$ such that $\tilde{f}(K) \subset V$ and

$$\|D^k(gfh^{-1})(x) - D^k(g\tilde{f}h^{-1})(x)\| < \epsilon$$

for all $x \in h(K)$, $k = 0, \ldots, r$.

Definition 2.32. The *weak topology* on $\mathcal{C}^r(M, N)$ is the topology generated by the subbase described in (2.1). We write $\mathcal{C}^r_w(M, N)$.

In compact manifolds, the weak topology works fine with the behavior of a map; this means that if two functions belong to a same open set, then they are close in compact domains. However, when the manifold is not compact, the weak topology does not control well the behavior of a map at the infinity. In this case, two functions could be close in certain local charts, but as the functions change charts they may separate. Therefore, we may find some open set that contains two functions, but these two functions are in fact very far. This does not happen in the strong topology. Therefore, in comapct maps it is preferable to work with the strong topology, that we define in the following.

Let $H = \{(U_i, h_i)\}_{i\in I}$ be a locally finite set of charts on a manifold M. Let $K = \{K_i\}_{i\in I}$ be a family of compact subsets of M, with $K_i \subset U_i$. Let $G = \{(V, g)\}_{i\in I}$ be a family of charts on N. Given a family of positive numbers $\epsilon = \{\epsilon_i\}_{i\in I}$, if $f \in \mathcal{C}^r(M, N)$ maps each K_i into V_i, we define a *strong basic neighborhood*

$$\mathcal{N}^r(f; H, G, K, \epsilon) \tag{2.2}$$

to be the set of $\tilde{f} \in \mathcal{C}^r(M, N)$ such tat for all $i \in I$, $\tilde{f}(K_i) \subset V_i$ and

$$\|D^k(g_i f h_i^{-1})(x) - D^k(g_i \tilde{f} h_i^{-1})(x)\| < \epsilon_i,$$

for all $x \in h_i(K_i)$, $k = 0, \dots, r$.

Definition 2.33. The *strong topology* on $\mathcal{C}^r(M, N)$ is the topology generated by the base described in (2.2). We write $\mathcal{C}^r_s(M, N)$.

Recall that a set is always a subbase of some topology. However, it is not true for a base. We should prove this property for the strong topology.

Proposition 2.15. The sets given by (2.2) form a base of the $\mathcal{C}^r_s(M, N)$ topology.

Proof. We prove the two properties on Proposition 2.1.

– Let f be a function between the manifolds. We can write

$$\mathcal{C}^r(M, N) = \bigcup_{f\in\mathcal{C}^r(M,N)} \mathcal{N}^r(f; H, G, K, \epsilon),$$

where $\epsilon = \{\epsilon_i\}_{i\in I}$, $\epsilon_i > 0$.

– Consider two elements of the base $\mathcal{N}_1$ and $\mathcal{N}_2$. We are going to prove that $\mathcal{N}_1 \cap \mathcal{N}_2$ is the union of elements of the base. We have,

$$\mathcal{N}_\lambda = \mathcal{N}^r(f_\lambda; H, G, K, \epsilon_\lambda)$$

for some f_λ and $\epsilon_\lambda = \{\epsilon_{\lambda,i}\}_{i\in I}$ $(\lambda = 1, 2)$. Fix $(U_i, h_i) \in H$, $(V_i, g_i) \in G$ and K_i. Let $f_j \in \mathcal{N}_1 \cap \mathcal{N}_2$. We have

$$\| D^k g_i f_\lambda h_i^{-1}(h_i p) - D^k g_i f_j h^{-1}(h_i p) \| < \epsilon_{\lambda,i},$$

for all $h_i(p) \in h_i(K_i)$, $k = 0, \ldots, r$. Since K_i is compact, $h_i(K_i)$ is compact and the function $L(h_i p) = \| g_i f_\lambda h_i^{-1}(h_i p) - g_i f_j h^{-1}(h_i p) \|$ is a continuous function defined on a compact domain, then it has a maximum, say $\epsilon_{\lambda,i,j} < \epsilon_{\lambda,i}$. Thus, we define

$$\hat{\epsilon}_{\lambda,i,j} = \frac{\epsilon_{\lambda,i} - \epsilon_{\lambda,i,j}}{2} > 0.$$

We consider $\epsilon_{i,j} = \min_{\lambda=1,2}\{\hat{\epsilon}_{\lambda,i,j}\}$. Then, the open set $\mathcal{N}_{f_j} = \mathcal{N}^r(f_j; H, G, K, \epsilon_j)$, $\epsilon = \{\epsilon_{i,j}\}_{i\in I}$ is contained in $\mathcal{N}_1 \cap \mathcal{N}_2$. Let $f \in \mathcal{N}_{f_j}$, then

$$\begin{aligned}
\| D^k g_i f_\lambda h_i^{-1}(h_i p) - D^k g_i f h_i^{-1}(h_i p) \| &= \| D^k g_i f_\lambda h_i^{-1}(h_i p) - D^k g_i f_j h_i^{-1}(h_i p) + \\
&\quad D^k g_i f_j h_i^{-1}(h_i p) - D^k g_i f_j h_i^{-1}(h_i p) \| \\
&\le \| D^k g_i f_\lambda h_i^{-1}(h_i p) - D^k g_i f_j h_i^{-1}(h_i p) \| + \\
&\quad \| D^k g_i f_j h_i^{-1}(h_i p) - D^k g_i f h_i^{-1}(h_i p) \| \\
&< \epsilon_{\lambda,i,j} + \frac{\epsilon_{\lambda,i} - \epsilon_{\lambda,i,j}}{2} = \frac{\epsilon_{\lambda,i} + \epsilon_{\lambda,i,j}}{2} \\
&< \frac{\epsilon_{\lambda,i} + \epsilon_{\lambda,i}}{2} = \epsilon_{\lambda,i}.
\end{aligned}$$

Hence, $f \in \mathcal{N}_\lambda$, for $\lambda = 1, 2$. Therefore, $\mathcal{N}_{f_j} \subset \mathcal{N}_1 \cap \mathcal{N}_2$. Thus, we have

$$\bigcup_{f_j \in \mathcal{N}_1 \cap \mathcal{N} 2} \mathcal{N}_{f_j} = \mathcal{N}_1 \cap \mathcal{N}_2.$$

□

We give some necessary results.

Proposition 2.16. Let M, N be compact manifolds. Then $\mathcal{C}_s^r(M, N) = \mathcal{C}_w^r(M, N)$.

Proof. Let $\mathcal{N}_s \in \mathcal{C}^r_s(M,N)$. Then we can write $\mathcal{N}_s$ as the union of members of the basis (2.2):

$$\begin{aligned}\mathcal{N}_s &= \bigcup_{j\in J} \mathcal{N}^r_j(f_i; H_j, G_j, K_j, \epsilon_j)\\ &= \bigcup_{j\in J}\bigcap_{i\in I} \mathcal{N}^r_{ij}(f_{ij}, (U_{ij}, h_{ij}), (V_{ij}, g_{ij}), K_{ij}, \epsilon_{ij}).\end{aligned}$$

Moreover, let $\mathcal{N}_w \in \mathcal{C}^r_w(M,N)$. Then this open set can be written as the union of a finite set of intersections as follows:

$$\mathcal{N}_w = \bigcup_{j\in J} \bigcap_{\substack{i\in I\\ \text{finite}}} \mathcal{N}^r_{ij}(f_{ij}, (U_i, h_i), (V_i, g_i), K_i, \epsilon_{ij}).$$

There are only two differences between $\mathcal{N}_w$ and $\mathcal{N}_s$. Firstly, the subindex i in $\mathcal{N}_w$ is finite. Secondly, ϵ_{ij} in $\mathcal{N}_w$ may be infinite. However, since M and N are compact sets, we can choose a finite locally finite atlas, and hence I is finite. Moreover, in compact sets, we achieve the maximum of the set for every function f_j, say A_j. Since there are finite functions, we have a maximum $\max_{j\in I}\{A_j\} = A$. Hence, it is the same to choose $\epsilon_{ij} = A$ or choose $\epsilon_{ij} = \infty$. □

Proposition 2.16 tells us that every property that is satisfied in one of these topologies is also satisfied in the other topology, as long as we work on compact sets. In this case, we write the open sets of the basis as $\mathcal{N}^r$.

Now, we define two basic sets. These sets extend the result of Lemma 2.3.

Definition 2.34. We denote by $\mathrm{Imm}^r(M,N)$ the set of $\mathcal{C}^r$ immersions between the manifolds M and N. Moreover, we denote by $\mathrm{Emb}^r(M,N)$ the set of $\mathcal{C}^r(M,N)$ embeddings.

Theorem 2.4. The set $\mathrm{Imm}^r(M,N)$ is open in $\mathcal{C}^r_s(M,N)$, for $r \geq 1$.

Proof. First of all, we have the equality

$$\mathrm{Imm}^r(M,N) = \mathrm{Imm}^1(M,N) \cap \mathcal{C}^r(M,N),$$

because $f \in \mathrm{Imm}^r(M,N)$ if, and only if, $Dgfh^{-1}$ is injective (for convenient charts) and f is r times differentiable and continuous. These conditions happen if, and only if, $f \in \mathrm{Imm}^1(M,N)$ and $f \in \mathcal{C}^r(M,N)$.

Since $\mathcal{C}^r(M,N)$ is the total space, we only have to see that $\mathrm{Imm}^1(M,N)$ is open. Let $f : M \to N$ be a $\mathcal{C}^1$ immersion. We are going to choose a convenient neighborhood of f, which is contained in $\mathrm{Imm}^r(M,N)$. Let $G = \{(V_j, g_j)\}_{j\in J}$ be an atlas of N and $H = \{(U_i, h_i)\}_{i\in I}$ an atlas of M. We choose U_i such that $f(U_i) \subseteq V_{j(i)}$, for some open chart of N (for example, if we have some cover $\{\hat{U}_i\}_{i\in\Lambda}$, we can intersect every open set with all the open sets in the chart V, and we will have a refinement of the set). We can send every U_i into the open ball $B(0,3)$ (through another refinement if it is necessary), where the closure of every U_i is compact. Let $K = \{K_i\}_{i\in I}$ be a compact cover of M with $K_i \subseteq U_i$. The set

$$A_i = \{D(g_i f h_i^{-1})(x) : x \in h(K_i)\}$$

is compact, since we can write A_i as

$$A_i = D(g_i f h_i^{-1})(h(K_i)),$$

and $h(K_i)$ is the image of a compact set by a continuous map, and $D(g_i f h_i^{-1})$ is also continuous. Moreover, A_i is a compact set of linear maps from $\mathbb{R}^m$ to $\mathbb{R}^n$. We know by Theorem 2.3 that the set of all linear maps is open in the vector space $\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$, and because of that there exists $\epsilon_l > 0$ such that $B \in \mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$ is injective if $\|B - D\| < \epsilon_i$, and $D \in A_i$. If we take $\epsilon = \{\epsilon_i\}_{i\in I}$, then every element of $\mathcal{N}^1(f; H, G, K, \epsilon)$ is an immersion. $\square$

In the previous result we have proved that the set of immersions is an open set. We are going to prove that the set of embeddings is also an open set. We state a previous lemma that will be proved in Appendix D.

Lemma 2.9. Let $U \subset \mathbb{R}^m$ be an open set and $W \subset U$ an open set with compact closure $\overline{W} \subset U$. Let $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^1$ embedding. There exists $\epsilon > 0$ such that if $g : U \to \mathbb{R}^n$ is $\mathcal{C}^1$ and

$$\|Dg(p) - Df(p)\| < \epsilon \text{ and } |g(p) - f(p)| < \epsilon$$

for all $p \in W$, then $g|_W$ is an embedding.

Theorem 2.5. The set $\mathrm{Emb}^r(M,N)$ of $\mathcal{C}^r(M,N)$ is open in $\mathcal{C}^r_S(M,N)$, $r \geq 1$.

Proof. Again, with the same argument as in Theorem 2.4, we can say that

$$\mathrm{Emb}^r(M,N) = \mathrm{Emb}^1(M,N) \cap \mathcal{C}^r(M,N).$$

Thus we only have to prove the theorem for $r = 1$. Let $f \in \mathrm{Emb}^1(M,N)$. We have to show that there exists a neighborhood of f where every function is an embedding. We can take:

(i) A locally finite atlas $H = \{(U_i, h_i)\}_{i \in I}$ of M, by Lemma 2.8.

(ii) A set of local charts $G = \{(V_i, g_i)\}_{i \in I}$ of N, where $f(U_i) \subset V_i$, as in Theorem 2.4.

(iii) An open cover $\{K_i\}_{i \in I}$ of M, where $\overline{K_i} \subset U_i$ is compact.

(iv) $\epsilon_i > 0$ such that if

$$g \in \mathcal{N}_0 = \mathcal{N}^r(f; H, G, K, \epsilon),$$

where $\epsilon = \{\epsilon_i\}_{i \in I}$, then $g(\overline{W}_i) \subset V_i$ and $g|_{K_i}$ is a $\mathcal{C}^r$ embedding, by Lemma 2.9.

(v) A neighbourhood $\mathcal{N}_2$ such that every open set in $\mathcal{N}_2$ is an immersion.

Since f is an embedding, for every $i \in I$, there exist disjoint open sets A_i, B_i in N such that $f(\overline{K_i}) \subset A_i$ and $f(M \setminus U_i) \subset B_i$, because $\overline{K_i}$ and $M \setminus U_i$ are disjoint sets (recall that $\overline{K_i} \subset U_i$, hence $M \setminus U_i \subset M \setminus \overline{K_i}$). Therefore, we can find a neighbourhood $\mathcal{N}_1$ of f in the $\mathcal{C}^r_s(M,N)$ topology such that if $\hat{f} \in \mathcal{N}_1$, then

$$\hat{f}(\overline{K_i}) \subset A_i, \quad \hat{f}(M \setminus K_i) \subset B_i.$$

We want to see that every $\hat{f} \in \mathcal{N}_0 \cap \mathcal{N}_1 \cap \mathcal{N}_2$ is an embedding: that is, an immersion that carries homeomorphically into its image. Since $\hat{f} \in \mathcal{N}_2$, we only have to see the second one.

- $\hat{f}$ is injective. Suppose that x, y are disjoint points of M. Since $\cup K_i = M$, in particular $\cup \overline{K}_i = M$, then $x \in \overline{K}_i$, for some $i \in I$. If $y \in U_i$, then $\hat{f}(x) \neq \hat{f}(y)$, since $\hat{f}|_{U_i}$ is injective. Else, $y \in M \setminus K_i$ ($(M \setminus K_i) \cup U_i = M$, because $K_i \subset \overline{K}_i \subset U_i$), where $\hat{f}(x) \in A_i$, $\hat{f}(y) \in B_i$ and these two are disjoint open sets, hence $\hat{f}(x) \neq \hat{f}(y)$.

- $\hat{f}: M \to \hat{f}(M)$ is an homeomorphism. Since $\hat{f}$ is continuous and injective, we must show that $\hat{f}^{-1}$ is continuous. We use the argument of continuity by sequence. Consider a sequence $\{y_n\}_{n\in\mathcal{N}} \in M$ such that $\hat{f}(y_n) \to \hat{f}(x)$. We have to show that $y_n \to x$.

  Let $x \in K_i$, then $\hat{f}(x) \in A_i$ (we consider $\hat{f}(K_i) \subset A_i$). Therefore we have only a finite number of $\hat{f}(y_n)$ in B_i, since there exists n_0 such that, if $\hat{f}(y_n) \to \hat{f}(x)$, then $\hat{f}(y_n) \in A_i$, for all $n \geq n_0$. Hence, there is only a finite number of $y_n \in K_i$. Finally, since $\hat{f}|_{K_i} : K_i \to \hat{f}(K_i)$ is an homeomorphism, then $y_n \to x$.

□

2.4 Dynamical Systems

In the literature, there are a lot of relations between differential topology and dynamical systems. Floris Takens makes another connection between these two branches of Mathematics. The goal of this section is to introduce the concepts about dynamical systems that allow us to understand the implications of the Takens' Theorem.

We introduce the formal definition of dynamical system, as described in [14].

Definition 2.35. A *dynamical system* is a semi-group $(G, \odot)$ acting on a space M, that is, there is a family of transformations on M, $\{T_g\}_{g\in G}$, and a map

$$\begin{array}{rccl} T: & G \times M & \to & M \\ & (g, x) & \mapsto & T_g(x) \end{array}$$

such that $T_g \circ T_h = T_{g\odot h}$.

There are a lot of phenomena that give rise to a dynamical system.

Definition 2.36. Consider $G = \mathbb{Z}$ or $G = \mathbb{N}$. Fix a function f defined over M. A *discrete dynamical system* is given by the action

$$\begin{array}{rccl} T: & G \times M & \to & M \\ & (n, x) & \mapsto & T_n(x) = f^n(x), \end{array}$$

that is, the n-times composition of f.

$$(T_n \circ T_m)(x) = f^n(f^m(x)) = f^{n+m}(x) = T_{n+m}(x).$$

By agreement, $f^0 = \text{id}$ and $f^{-n} = (f^n)^{-1}$, if $n > 0$.

Example 2.10. Let (a, b) be a point on the circumference of radius r, centered at the origin. We may write $a = r\sin(x_0)$ and $b = r\cos(x_0)$. We construct a discrete dynamical system such that $x_n = r\sin(x_0 + nk)$ and $y_n = r\cos(x_0 + nk)$. Therefore, by expanding the sinus and the cosinus of a sum of angles,

$$\begin{aligned} x_{n+1} &= r\sin((x_0 + nk) + k) = r\sin(x_0 + nk)\cos(k) + r\cos(x_0 + nk)\sin(k) \\ &= x_n\cos(k) + y_n\sin(k), \\ y_{n+1} &= r\cos((x_0 + nk) + k) = r\cos(x_0 + nk)\cos(k) - r\sin(x_0 + nk)\sin(k) \\ &= y_n\cos(k) - x_n\sin(k). \end{aligned}$$

This is a dynamical system that walks on the circumference. We can write it as a rotation matrix

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} \cos(k) & \sin(k) \\ -\sin(k) & \cos(k) \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \cos(k) & \sin(k) \\ -\sin(k) & \cos(k) \end{pmatrix}^{n+1}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = f^{n+1}((x_0, y_0)),$$

where f is a linear application. If $\frac{2\pi}{k}$ is irrational, then the orbit is dense on the circumference. However, if it is rational, the system is periodic: that is, there is some $N > 0$ such that $x_N = x_0$ and $y_N = y_0$.

Definition 2.37. Consider $G = \mathbb{R}$. Let ϕ be the flow of an autonomous ordinary differential equation (ODE). As usually, we denote $\phi(t; (0, x_0)) = \phi(t; x_0)$, where $\phi(t; x_0)$ is the solution that in time 0 passes through x_0. A *continuous dynamical system* is given by the action

$$\begin{aligned} T: \ \mathbb{R} \times M &\to M \\ (t, x_0) &\mapsto T_t(x_0) = \phi(t; x_0), \end{aligned}$$

Therefore,

$$(T_t \circ T_s)(x_0) = T_t(\phi(s; x_0)) = \phi(t; \phi(s; x_0)) = \phi(t + s; x_0) = T_{t+s}(\phi(t_0)).$$

The equality $\phi(t; \phi(s; x_0)) = \phi(t + s; x_0)$ is a consequence of the uniqueness of solutions of ODE, since $\phi(t + s; x_0)$ at time $t = 0$ passes through $\phi(s; x_0)$.

Example 2.11. Consider the differential equation

$$\begin{cases} x' = -y + x(1 - \sqrt{x^2 + y^2}), \\ y' = x + y(1 - \sqrt{x^2 + y^2}). \end{cases}$$

Trivially, one can check that $(x, y) = (0, 0)$, and $(x, y) = (\cos t, \sin t)$ are solutions of the previous system. Furthermore, there is a branch of solutions

$$(x(t; (x_0, y_0)), y(t; (x_0, y_0))) = \Big(\frac{k_0 e^{t+\theta_0}}{1 + k_0 e^{t+\theta_0}}\cos(t + \theta_0), \frac{k_0 e^{t+\theta_0}}{1 + k_0 e^{t+\theta_0}}\sin(t + \theta_0)\Big), \quad (2.3)$$

where $r_0 = x_0^2 + y_0^2$, $k_0 = r_0/(1 - r_0)$ and $\theta_0 = \arctan \frac{y_0}{x_0}$. Then, if $\bar{x}_0 = (x_0, y_0)$, the dynamical system is

$$\begin{array}{rlcl} T: & \mathbb{R} \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (t, \bar{x}_0) & \mapsto & T_t(\bar{x}_0) = (x(t; (x_0, y_0)), y(t; (x_0, y_0))). \end{array}$$

Even when we restrict ourselves to discrete and continuous dynamical systems, the notion of dynamical system can be defined over much more abstract situations, as in the next exemple, where we consider a permutation group acting on a vector space.

Example 2.12. Let $M = \mathrm{Pol}(\mathbb{K})$ be the space of power series in a field $\mathbb{K}$ and $\mathfrak{S}$ the set of permutations of infinite elements. We define an action

$$\begin{array}{rlcl} T: & \mathfrak{S} \times M & \to & M \\ & (\sigma, p) & \mapsto & T_\sigma(p) = p^\sigma, \end{array}$$

where

$$p(x) = \sum_{i \in I} a_i x^i, \quad p^\sigma(x) = \sum_{i \in I} a_i x^{\sigma(i)}.$$

In this case, we have built a dynamical system from a permutation set and a given polynomial;

$$\begin{aligned} T_\sigma \circ T_{\tilde{\sigma}}(p(x)) &= T_\sigma \left(\sum_{i \in I} a_i x^{\tilde{\sigma}(i)} \right) = \sum_{i \in I} a_i x^{\sigma \tilde{\sigma}(i)} \\ &= T_{\sigma \circ \tilde{\sigma}}(p(x)). \end{aligned}$$

We focus especially on continuous dynamical systems. We might also work on the other explained dynamical systems, but we shall center first on continuous ones.

We recall that, given an autonomous Ordinary Differential Equation

$$\begin{cases} \dot{x} = F(x), \\ x(0) = x_0, \end{cases} \tag{2.4}$$

where $x = x(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^n$, a solution is a $\mathcal{C}^1$-function $\phi(t; x_0)$ such that satisfies the ODE. We denote F as the *vector field*. If F satisfies a Lipschitz condition in a domain, there exists uniqueness of solutions in this domain. Therefore, we suppose until the end of the manuscript that the vector field satisfies a Lipschitz condition in all the domain. It is well known that given an ODE with uniqueness of solutions in the whole domain,there exists a differential conjugate ODE with uniqueness of solutions and they are defined in the whole $\mathbb{R}$. It is not restrictive, to assume the ODEs we are working with have solutions defined in the whole real line.

Definition 2.38. Consider an ODE (2.4). A point p such that $F(p) = 0$ is called a *singular point* of F. The other points are called *regular points* of F.

We note that a singular point corresponds to a constant solution of the differential equation; that is, if p is a singular point, then $\phi(t;p) = p$ is a solution of the equation.

Example 2.13. Consider the ODE described in Example 2.11. In this case, the only singular point is $(x, y) = (0, 0)$. The other points are regular points.

Definition 2.39. A *periodic solution* for (2.4) is a solution $\phi(t; x_0)$ such that there exists a time $\tau > 0$ such that $\phi(t + \tau; x_0) = \phi(t; x_0)$.

Definition 2.40. A point $p \in \mathbb{R}^n$ is an *ω-limit* point for the solution through x_0, that is $\phi(t; x_0)$, if there is a sequence $t_n \to \infty$ such that $\lim_{n\to\infty} \phi(t_n; x_0) = p$. The set of all ω-limit points of the solution through x_0 is the ω-limit set of $\phi(t; x_0)$ and it is denoted by $\omega(\phi(t; x_0))$.

The α-limit points and the α-limit set $\alpha(\phi(t; x_0))$ are defined similarly, but instead of $t_n \to \infty$ considering $t_n \to -\infty$ in the definition of ω-limit.

A *limit set* is the ω-limit set or the α-limit set for a differential equation.

Example 2.14. Consider the ODE described in Example 2.11. As we can see,

$$\lim_{t\to\infty} \frac{k_0 e^{t+\theta_0}}{1 + k_0 e^{t+\theta_0}} = 1.$$

Hence, all the solutions in (2.3) except the singular point $(0, 0)$ ten to the solution $(x, y) = (\cos t, \sin t)$, as t tends to $+\infty$. Therefore, the circumference is the ω-limit set of all the solutions, except for $(0, 0)$. Moreover, $(0, 0)$ is the α-limit of all the solutions starting in the open unit circle.

Definition 2.41. A *positively invariant set* A from a flow $\phi(t, x)$ is a set such that if $\phi(t_0, x) \in A$ for some t_0, then $\phi(t, x) \in A$ for all $t \geq t_0$.

Definition 2.42. A *stable set* S from a continuous dynamical system of flow $\phi(t, x)$ is a set such that there exists a neighborhood B of S satisfying that if $\phi(t_0, x) \in B$, then $\gamma_0(t, x) \in B$ for all $t \geq t_0$.

Furthermore, if there exists a neighbourhood B such that, for every negihbourhood $B' \subseteq B$, if $\phi(t_0, x) \in B$, then there exists $t_1 \geq t_0$ such that $\phi(t_0, x) \in B'$ for every $t \geq t_1$, then S is also *assimptotically stable set.*

Definition 2.43. An *attracting set* of an ODE is a closed, positively invariant and asymptotically stable set. An *attractor* of an ODE is an attracting set which contains a dense orbit.

Chapter 3

Takens' Embedding Theorem

In this chapter, we enunciate and prove the Takens' Embedding Theorem. The proof appears firstly in [2]. We follow the proof described in [1] and complement the demonstration with [15]. In our case, we give a little more details and finish the proof with a relaxed condition. In [1], the author states Takens' Theorem assuming that the functions are $\mathcal{C}^2$ functions. At the end of the article, he writes about the reduction of this condition to $\mathcal{C}^1$ functions. By following these ideas, in Section 3.9 we completely prove the theorem for $\mathcal{C}^1$ flows.

In all this chapter, we consider M as a compact manifold of dimension m. For this proof, we have to make some stages. We divide the chapter in sections where we build every stage. Until the Section 3.7 and unless stated otherwise, we consider y as a variable function and a fixed $\phi \in \mathrm{Dif}^2(M)$ with the following properties:

(i) The periodic points of ϕ with period less than or equal to $2m$ are finite in number.

(ii) If x is any periodic point with period $k \leq 2m$, then the eigenvalues of the derivative of ϕ^k at x are all distinct.

This pair of functions (ϕ, y) leads to a family of functions

$$\begin{array}{rccl} \Phi_{(\phi,y;k)}: & M & \to & \mathbb{R}^k \\ & x & \mapsto & \Phi_{(\phi,y;k)}(x) = (y(x), y(\phi(x)), \ldots, y(\phi^{k-1}(x))). \end{array} \tag{3.1}$$

We refer by $\Phi_{(\phi,y)} = \Phi_{(\phi,y;2m+1)}$, since it is a special case. It is called *delay map*. We state the mainly theorem:

Theorem 3.1 (Takens' Embedding Theorem)**.** Let M be a compact manifold of dimension m. For pairs (ϕ, y), with $\phi \in \mathrm{Dif}^2(M)$, $y \in \mathcal{C}^2(M, \mathbb{R})$, it is a generic property that the map $\Phi(\phi, y)$ is an embedding.

We note that in some sections we will not use all this conditions. For example, in section 3.1, we only consider $\phi \in \mathrm{Dif}^1(M)$. To prove Takens' Theorem, we must prove the genericity of the theorem. Therefore, we shall prove the openess and denseness part. In section 3.1 we prove the openess part for a fixed $\phi \in \mathrm{Dif}^2(M)$ and from Section 3.3 to 3.6 we prove the denseness part for the same ϕ. In Section 3.8 we prove the openess and denseness part for the general theorem.

3.1 Openness of the set of embeddings

We start with the following result.

Lemma 3.1. Fix $\phi \in \mathrm{Dif}^1(M)$. The function

$$\begin{array}{rrcl} F_1: & \mathcal{C}^1(M,\mathbb{R}) & \to & \mathcal{C}^1(M,\mathbb{R}) \\ & y & \mapsto & y \circ \phi \end{array}$$

is continuous.

Proof. Let $\{(U_i, h_i) : i \in I\}$ be a finite regular covering for M. We ensured in Chapter 2 that we can always take a finite regular covering for a manifold M. Hence, we can take sets $W = \{W_i\}_{i\in I}$, where $W_i = h_i^{-1}B(0,1)$ and the set W covers M. Let $(\mathbb{R}, \mathrm{id})$ be the local chart for $\mathbb{R}$. Given any neighborhood $\mathcal{N} \in \mathcal{C}^1(M,\mathbb{R})$ of $y \circ \phi$, because or the definition of neighborhood, we can choose an open set contained in $\mathcal{N}$, and furthermore, tjos $\mathcal{N}$ will be intersections of elements of the base, hence

$$\mathcal{N} = \cap_{i\in I}\mathcal{N}^1(y \circ \phi, (U_i, h_i), (\mathbb{R}, \mathrm{id}), \overline{W}_i, \epsilon'),$$

for some $\epsilon' > 0$ sufficiently small. We have to show that there exists a neighborhood $\mathcal{N}(\epsilon) = \cap_{i\in I}\mathcal{N}^1(y\circ\phi, (U_i, h_i), (\mathbb{R}, \mathrm{id}), \overline{W}_i, \epsilon)$, of y such that if $\hat{y} \in \mathcal{N}(\epsilon)$, then $F_1(\hat{y}) \in \mathcal{N}$; that is, $F_1(\mathcal{N}(\epsilon)) \subseteq \mathcal{N}$, so that F_1 is continuous. We have to prove it for some $\epsilon > 0$.

The sets W_i, $i \in I$, cover M, and since ϕ is a diffeomorphism on M, so do the sets $\phi^{-1}W_i$, because

$$\bigcup_{i\in I}\phi^{-1}W_i = \phi^{-1}\bigcup_{i\in I}W_i = M.$$

Also, the sets $\phi^{-1}W_i \cap W_j$, $i, j \in I$, also cover M:

$$\bigcup_{i,j\in I}(\phi^{-1}W_i \cap W_j) = \bigcup_{j\in I}(\bigcup_{i\in I}(\phi^{-1}W_i) \cap W_j) = \bigcup_{j\in I}W_j = M.$$

Then, we can consider one of the sets $\phi^{-1}W_i \cap W_j$ to be a non-empty set. The closure $\overline{\phi^{-1}W_i \cap W_j} \subseteq \overline{\phi^{-1}W_i} \cap \overline{W_j} = \phi^{-1}\overline{W_i} \cap \overline{W_j}$, hence for some $i, j \in I$, $\phi^{-1}\overline{W_i} \cap \overline{W_j}$ is not empty and it is compact, since every $\overline{W}_k$, $k \in I$ is compact, because W_i is the inverse image of a bounded real set and intersection of closed and compact sets is a compact set. We note that the map

$$h_i\phi h_j^{-1} : h_j(\phi^{-1}\overline{W}_i \cap \overline{W}_j) \subset \mathbb{R}^m \to \mathbb{R}^m.$$

Hence, the derivative is a map $Dh_i\phi h_j^{-1} : h_j(\phi^{-1}\overline{W}_i \cap \overline{W}_j) \to \mathbb{R}^{m\times m}$. Since ϕ is a $\mathcal{C}^1$-diffeomorphism, it is also continuous. Hence, the matrix norms of these derivatives are bounded, because the image remains into a compact set on $\mathbb{R}^{m\times m}$ and every compact set of a real space has a maximum value. Hence $\|Dh_i\phi h_j^{-1}(u)\| < A_{i,j}$, for all $u \in h_j(\phi^{-1}\overline{W}_i \cap \overline{W}_j)$. Since M is compact, we have only finitely W_j that cover M, so we have finitely many of $\phi^{-1}W_i \cap W_j$ and we can find a single A which is an upper bound for $\{A_{i,j} : i, j \in I\}$. We are going to find ϵ. Let $\hat{y} \in \mathcal{N}(\epsilon)$ and let $x \in \overline{W}_j$. Thus, there is some $i \in I$ such that $x \in \phi^{-1}\overline{W}_i \cap \overline{W}_j$. Consider the image through ϕ, $x' = \phi(x)$. Since $x \in \phi^{-1}\overline{W}_i$, then $\phi(x) \in \overline{W}_i$. Therefore,

$$\begin{aligned}
\|\hat{y} \circ \phi h_j^{-1}(h_j x) - y \circ \phi h_j^{-1}(h_j x)\| &= \|\hat{y}(\phi(x)) - y(\phi(x))\| \\
&= \|\hat{y}(x') - y(x')\| \\
&= \|\hat{y}h_i^{-1}(h_i x') - y h_i^{-1}(h_i x')\| \\
&< \epsilon.
\end{aligned}$$

The inequality holds because $\hat{y}, y \in \mathcal{N}(\epsilon)$. We need then $\epsilon < \epsilon'$. In addition,

$$\begin{aligned}
\|D\hat{y}\phi h_j^{-1}(h_j x) - Dy\phi h_j^{-1}(h_j x)\| &= \|D\hat{y}h_i^{-1}h_i\phi h_j^{-1}(h_j x) - Dyh_i^{-1}h_i\phi h_j^{-1}(h_j x)\| \\
&= \|D\hat{y}h_i^{-1}(h_i\phi h_j^{-1}h_j x)Dh_i\phi h_j^{-1}(h_j x) \\
&\quad - Dyh_i^{-1}(h_i\phi h_j^{-1}h_j x)Dh_i\phi h_j^{-1}(h_j x)\| \\
&= \|D\hat{y}h_i^{-1}(h_i\phi x)Dh_i\phi h_j^{-1}(h_j x) \\
&\quad - Dyh_i^{-1}(h_i\phi x)Dh_i\phi h_j^{-1}(h_j x)\| \\
&= \|D\hat{y}h_i^{-1}(h_i x')Dh_i\phi h_j^{-1}(h_j x) \\
&\quad - Dyh_i^{-1}(h_i x')Dh_i\phi h_j^{-1}(h_j x)\| \\
&= \|(D\hat{y}h_i^{-1}(h_i x') - Dyh_i^{-1}(h_i x'))Dh_i\phi h_j^{-1}(h_j x)\| \\
&\leq \|D\hat{y}h_i^{-1}(h_i x') - Dyh_i^{-1}(h_i x')\|\|Dh_i\phi h_j^{-1}(h_j x)\| \\
&< \epsilon A.
\end{aligned}$$

Hence we want $\epsilon A < \epsilon'$. We must take $\epsilon < \min\{\epsilon', \epsilon'/A\}$. In this case, if $\hat{y} \in \mathcal{N}(\epsilon)$, $F_1(\hat{y}) = \hat{y} \circ \phi \in \mathcal{N}(\epsilon)$ and $\hat{y} \circ \phi \in \mathcal{N}$. Hence, F_1 is continuous. □

Corollary 3.1. The function $F_n : \mathcal{C}^1(M, \mathbb{R}) \to \mathcal{C}^1(M, \mathbb{R})$, for $n \in \mathbb{Z}^+$, defined by $y \mapsto y \circ \phi^n$ is continuous.

Proof. The case F_1 ($n = 1$) is done in Lemma 3.1. Assume that F_n is continuous. By induction,

$$F_{n+1}(y) = y \circ \phi^{n+1} = (y \circ \phi^n) \circ \phi = F_1(y \circ \phi^n) = F_1(F_n(y)),$$

hence F_{n+1} is the composition of two continuous functions, F_1 and F_n. By Proposition 2.8, we have that F_{n+1} is continuous. □

Corollary 3.2. The function

$$\begin{array}{rccl} \mathcal{F}: & \mathcal{C}^1(M, \mathbb{R}) & \to & \mathcal{C}^1(M, \mathbb{R}^k) \\ & y & \mapsto & \Phi_{(\phi,y;k)} = (y, \ldots, y\phi^{k-1}). \end{array}$$

is continuous.

Proof. The identity is always a continuous function and we know (by Corollary 3.1) that the other components are continuous. Since the Cartesian product of continuous functions is a continuous function (Example 2.6), $\mathcal{F}$ is a continuous function. □

Hence, the set of functions described in (3.1) is a set of continuous functions. In particular, the delay map $\Phi_{(\phi,y)}$ is a continuous function.

Proposition 3.1. Let M be a compact manifold, $\phi : M \to M$ a diffeomorphism, and K a compact subset of M. Then the set of functions

$$\mathcal{Y} = \{y \in \mathcal{C}^1(M, \mathbb{R}) : \Phi_{(\phi,y;k)} \text{ immersive on } K\}$$

where $\Phi_{(\phi,y;k)} : M \to \mathbb{R}^k$ is the map (3.1), it is an open set in $\mathcal{C}^1(M, \mathbb{R})$.

Proof. Consider

$$S = \{f : M \to \mathbb{R}^k : f \text{ is immersive on } K\}. \tag{3.2}$$

We have proved in Theorem 2.4 that S is an open set, and we note that $\mathcal{F}^{-1}(S) = \mathcal{Y}$. Because of the continuity of $\mathcal{F}$, we have $\mathcal{Y}$ is open. □

We note that the set of embeddings of $\mathcal{C}^1(M, \mathbb{R}^k)$ forms an open set. Therefore, the same argument applies to injective immersions on K and hence we have the following proposition:

Proposition 3.2. Let M be a compact manifold, $\phi : M \to M$ a diffeomorphism, and K be a compact subset of M. Then the set of functions

$$\mathcal{Y}_e = \{y \in \mathcal{C}^1(M, \mathbb{R}) : \Phi_{(\phi,y;k)} \text{ embedding on } K\}$$

where $\Phi_{(\phi,y;k)} : M \to \mathbb{R}^k$ is the map (3.1), it is an open set in $\mathcal{C}^1(M, \mathbb{R})$.

Therefore, if $\Phi_{(\phi,y;k)}$ is an immersion or an embedding, for every $\hat{y}$ in a neighbourhood of y, $\Phi_{(\phi,\hat{y};k)}$ is also an immersion or an embedding.

3.2 Measurement Functions

We have proved that the set of measurement functions y that makes $\Phi_{(\phi,y)}$ an embedding is open in $\mathcal{C}^1(M, \mathbb{R})$. We will show that it is also dense. Let y be a measurement function such that $\Phi_{(\phi,y)}$ is not an embedding. We must find some function y' in every neighborhood of y with this property. We construct y' such that

$$y' = y + \sum_{i=1}^{N} a_i \psi_i, \tag{3.3}$$

where $N \in \mathbb{N}$ is finite, $a_i \in \mathbb{R}$ and $\psi_i : M \to \mathbb{R}$ is differentiable. To ensure that $y' \in \mathcal{C}^1(M, \mathbb{R})$, we need as a hypothesis $\psi_i \in \mathcal{C}^r$, $r \geq 1$.

Lemma 3.2. Let $y : M \to \mathbb{R}$ be $\mathcal{C}^1$ and let $\psi_i : M \to \mathbb{R}$, $i = 1, \ldots, N$ be $\mathcal{C}^r$, $r \geq 1$, for all i, where N is finite. Let $a = (a_1, \ldots, a_N) \in \mathbb{R}^N$. For each neighbourhood $\mathcal{N}$ of y, there is some $\delta > 0$ such that if $\|a\| < \delta$, the function y' defined as in Equation (3.3) belongs to $\mathcal{N}$.

Proof. Let $\{(U_i, h_i) : i \in I\}$ be a finite regular covering for M, with $\overline{W}_i \subset U_i$. Since $\mathcal{N}$ is a neighbourhood, there exists some open subset such that $\mathcal{N} = \cap_{i \in I} \mathcal{N}^1(y; (U_i, h_i), (\mathbb{R}, \text{id}), \overline{W}_i, \epsilon)$, for some $\epsilon > 0$.

For each $1 \leq j \leq N$, and each $i \in I$, the map $\psi_j h_i^{-1} : h_i(\overline{W}_i) \to \mathbb{R}$ is well defined, because $h_i^{-1} : h_i(\overline{W}_i) \to \overline{W}_i$, and $\psi_j : \overline{W}_i \subset U_i \subset M \to \mathbb{R}$. Furthermore, since both functions are continuous, $\psi_j h_i^{-1}$ is also continuous. We know that $\overline{W}_i$ is compact, so it is $\psi_j h_i^{-1}(h_i(\overline{W}_i))$. Then, the function is bounded by some constant $B_{i,j}$. Since we have a finite atlas, we may take $B_j = \max\{B_{i,j},\ i \in I\}$.

We use an induction argument. Suppose $N = 1$, and say $a_1 = a$, $\psi_1 = \psi$. We take $y' = y + a\psi$. If $x \in \overline{W}_i$, then

$$\begin{aligned}\|y'h_i^{-1}(h_i x) - yh_i^{-1}(h_i x)\| &= \|(y + a\psi)h_i^{-1}(h_i x) - yh_i^{-1}(h_i x)\| \\ &= \|a\psi h_i^{-1}(h_i x)\| = |a| \cdot \|\psi h_i^{-1}(h_i x)\| \\ &\leq |a|B_{i,1} \leq |a|B_1.\end{aligned}$$

So in this case we need $|a|B_1 < \epsilon$. Similarly, the derivatives $D\psi_j h_i^{-1}$ are continuous functions, so there is a bound $B'_{i,j}$. In case $N = 1$, we have

$$\begin{aligned}\|Dy'h_i^{-1}(h_i x) - Dyh_i^{-1}(h_i x)\| &= \|D(y + a\psi)h_i^{-1}(h_i x) - Dyh_i^{-1}(h_i x)\| \\ &= \|aD\psi h_i^{-1}(h_i x)\| = |a| \cdot \|\psi Dh_i^{-1}(h_i x)\| \\ &\leq |a|B'_{i,1} \leq |a|B'_1.\end{aligned}$$

So, we need also $|a|B'_1 < \epsilon$. If $\|a\| < \delta < \min\{\frac{\epsilon}{B_1}, \frac{\epsilon}{B'_1}\}$ (B_1 and B'_1 are non-zero), then $\|y'h_i^{-1}(h_i x) - yh_i^{-1}(h_i x)\| < \epsilon$ and $\|Dy'h_i^{-1}(h_i x) - Dyh_i^{-1}(h_i x)\| < \epsilon$. We have that $y' \in \mathcal{N}$.

Suppose that it is true for N and we want to prove for $N + 1$. Assuming that

$$y' = y + \sum_{j=1}^{N} a_j \psi_j$$

remains in the same neighbourhood $\mathcal{N}$, we want to show that

$$y'' = y + \sum_{j=1}^{N+1} a_j \psi_j.$$

Since y' is also $\mathcal{C}^1(M, \mathbb{R})$, we have that for the case $N = 1$ applied to y', there is some neighbourhood $\mathcal{N}' \subseteq \mathcal{N}$ where $y', y'' \in \mathcal{N}'$, for some $\delta > 0$, hence $y'' \in \mathcal{N}$. □

We can see in the previous proof that compactness allows y' to be near y. If M is not compact, this result is false.

We call at the process of finding some y' close to y an adjustment of y. We look for a chain of adjustments that gives rise to an embedding.

3.3 Immersion on Periodic Orbits

We define P_l as the set of periodic points of ϕ, with period less than or equal to l. For example, if $l = 1$, then $x \in P_1$ is a fixed point of ϕ. If $l = 2$, we also have $\phi^2(x) = x$.

In this section, we build some adjustments on y as in Section 3.2. We seek the map $\Phi_{(\phi,y;k)}$ to be injective restricted to the set P_l, for $l \in \mathbb{Z}^+$.

Proposition 3.3. The set

$$\mathcal{Y} = \{y \in \mathcal{C}^1(M, \mathbb{R}) : \Phi_{(\phi,y;k)} \text{ is injective restricted to } P_l\},$$

is dense.

Proof. Suppose that $y \notin \mathcal{Y}$ and we want to find some $y' \in \mathcal{Y}$ in a neighborhood of y. Since $y \notin \mathcal{Y}$, we have some pair of different points x_1 and x_2 in P_l such that $\Phi_{(\phi,y;k)}(x_1) = \Phi_{(\phi,y;k)}(x_2)$ and thus $y(x_1) = y(x_2)$. Since M is Hausdorff, there exists some local chart (U_1, h_1), where U_1 is carried homeomorphically to the ball $B(0, 3)$, centered in $h_1(x_1)$ and it does not contain x_2. We define a function

$$\begin{array}{rccl} \lambda : & M & \to & \mathbb{R} \\ & x & \to & \lambda(x) = \begin{cases} \psi(h_1(x)), & \text{for } x \in h_1^{-1}B(0,3), \\ 0 & \text{otherwise}, \end{cases} \end{array}$$

where $\psi : \mathbb{R}^m \to \mathbb{R}$ is a bump function having support in $B(0, 3)$ and equal to 1 on $B(0, 1)$. Hence

$$y' = y + a\lambda,$$

where $a \in \mathbb{R}$ will have sufficiently small norm. Hence, for every $a > 0$, we have

$$y'(x_1) = y(x_1) + a\lambda(x_1) = y(x_1) + a,$$
$$y'(x_2) = y(x_2) + a\lambda(x_2) = y(x_2).$$

Note that $\lambda(x_1) = 1$, because $x_1 \in h_1^{-1}(B(0, 1))$ (it is the center) and therefore $y'(x_1) = y(x_1) + a > y(x_2) = y'(x_2)$, hence $y'(x_1) \neq y'(x_2)$. Taking a arbitrarily small, we have that y' remains in a neighbourhood $\mathcal{N}$ of y, by Lemma 3.2. Since the set P_l is finite, we only need to make a finite number of these adjustments until we arrive at some y' such that, for every pair of points $x_i \neq x_j$ in P_l, then $y'(x_i) \neq y'(x_j)$. Therefore, y' is injective in P_l and this implies that $\Phi_{(\phi,y';k)}$ is injective in P_l. □

Corollary 3.3. The set $\mathcal{Y}$ described in Proposition 3.3 is also generic.

Proof. The denseness part is consequence of Proposition 3.3. The open part is consequence of Proposition 3.1, taking P_l as a compact set, since it is a finite set. □

We want $\Phi_{(\phi,y;k)}$ to be an injective immersion on P_l, for generic y. Therefore, we want that $Dg_i\Phi_{(\phi,y;k)}h_i^{-1}(h_ix_i) = D\Phi_{(\phi,y;k)}h_i^{-1}(h_ix_i)$[1] to be full rank at every $x_i \in P_l$.

Proposition 3.4. Let $\Phi_{(\phi,y;k)}$ be as in (3.1). Hence, $\Phi_{(\phi,y;k)}$ is an injective immersion on P_l, for generic y.

Proof. Suppose that we have $x_1, x_2, \ldots, x_p$ different points such that $\phi(x_j) = x_{j+1}$, for $1 \leq j < p$ and $\phi(x_p) = x_1$. Since M is Hausdorff, we can find open balls B_i, $h_1(x_1) \in B_1, \ldots, h_p(x_p) \in B_p$ disjoints.

Consider the question of immersivity at x_1: that is, the rank of $D\Phi_{(\phi,y;k)}h_1^{-1}(h_1x_1)$. Let $2 \leq j \leq p$. The matrix rows are

$$\begin{array}{lll} ip+1 \Rightarrow & Dy\phi^{ip}h_1^{-1}(h_1x_1) & = Dyh_1^{-1}h_1\phi^{ip}h_1^{-1}(h_1x_1) \\ & & = Dyh_1^{-1}(h_1x_1)Dh_1\phi^{ip}h_1^{-1}(h_1x_1) \\ ip+j \Rightarrow & Dy\phi^{(i+1)p-(p+1-j)}h_1^{-1}(h_1x_1) & = Dy\phi^{-(p+1-j)}h_1^{-1}h_1\phi^{(i+1)p}h_1^{-1}(h_1x_1) \\ & & = Dy\phi^{-(p+1-j)}h_1^{-1}(h_1x_1)Dh_1\phi^{(i+1)p}h_1^{-1}(h_1x_1) \end{array}$$

Let $v_1 = Dyh_1^{-1}(h_1x_1)$, $v_j = Dy\phi^{-(p+1-j)}$, for $j = 2, \ldots, p$ and $J = Dyh_1\phi^p h_1^{-1}(h_1x_1)$. We call $v_j = \sum_i \alpha_{j,i}e_j$, for $j = 1, \ldots, p$. We state that $J^s = Dh_1\phi^s h_1^{-1}(h_1x_1)$. The case $s = 1$ is trivially true. Suppose that it is satisfied for s and we want to prove this properly for $s+1$.

$$\begin{aligned} Dh_1\phi^{s+1}h_1^{-1}(h_1x_1) &= Dh_1\phi^s\phi h_1^{-1}(h_1x_1) = Dh_1\phi^s h_1^{-1}h_1\phi h_1^{-1}(h_1x_1) \\ &= Dh_1\phi^s h_1^{-1}(h_1x_1)Dh_1\phi h_1^{-1}(h_1x_1) = J^sJ = J^{s+1}. \end{aligned}$$

Hence, $v_jJ^{s-1} = \sum_i \alpha_{j,i}\lambda_j^{s-1}e_j$, where λ_i are the distinct eigenvalues of J and the matrix of the derivative takes the form

$$D\Phi_{(\phi,y;k)}h_1^{-1}(h_1x_1) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pm} \\ \alpha_{11}\lambda_1 & \alpha_{12}\lambda_2 & \cdots & \alpha_{1m}\lambda_m \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{11}\lambda_1^{\rho_1} & \alpha_{12}\lambda_2^{\rho_1} & \cdots & \alpha_{1m}\lambda_m^{\rho_1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{p1}\lambda_1^{\rho_p} & \alpha_{p2}\lambda_2^{\rho_p} & \cdots & \alpha_{pm}\lambda_m^{\rho_p} \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \end{pmatrix}.$$

[1] We recall that $g_i \equiv \mathrm{id}$, since the local chart for a real space is itself through the identity.

In this case, $\rho_j = \lfloor k/p \rfloor$ and if $k = ip+r$, with r the residual, then for every $j > r$ these columns do not appear on the matrix and then we consider $\rho_j = \lfloor k/p \rfloor - 1$. Since we want to know if the determinant is non-zero, we can rearrange the rows of the matrix:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{11}\lambda_1 & \alpha_{12}\lambda_2 & \cdots & \alpha_{1m}\lambda_m \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{11}\lambda_1^{\rho_1} & \alpha_{12}\lambda_2^{\rho_1} & \cdots & \alpha_{1m}\lambda_m^{\rho_1} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{p1}\lambda_1^{\rho_p} & \alpha_{p2}\lambda_2^{\rho_p} & \cdots & \alpha_{1m}\lambda_m^{\rho_p}. \end{pmatrix} \tag{3.4}$$

We want to find m linearly independent rows of matrix (3.4). Hence, we consider the first m rows and we have a matrix $m \times m$. They will be linearly independent if, and only if, the determinant is not zero. If the α_i's are real, then considered as a function the determinant is a polynomial. If the polynomial does not vanish identically, its zeros form a closed, nowhere dense set. In fact, if there is some neighborhood where the polynomial vanishes identically, and since the polynomials are analytic, its Taylor expansion will be identically zero and hence it is the polynomial zero. Thus, we look for some α_i's such that the determinant is non-zero. If we take the values

$$(\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1m}, \alpha_{21}, \ldots, \alpha_{pm}) = (1, 1, \ldots, 1, \lambda_1^{\rho_1+1}, \ldots, \lambda_p^{\sum_{j=1}^{p-1}\rho_0+1})$$

we call a Vandermonde matrix. Its determinant is non-zero for $\lambda_i \neq \lambda_j$, $i \neq j$. The set of values that makes the Jacobi full rank is open and dense in $\mathbb{R}^{mp}$. Therefore, we can find near $v = (v_1, \ldots, v_p)$ some v' such that $\|v - v'\|$ is arbitrarily small in norm and $\{v_1', v_1'J, \ldots, v_1'J^{\rho_1}, v_2', \ldots, v_p'J^{\rho_p}\}$ spans $\mathbb{R}^m$.

Let y be such that $\Phi_{(\phi,y;k)}$ is not an immersion in x_1. We define

$$\begin{array}{rccl} \lambda_{j,i}: & M & \to & \mathbb{R} \\ & x & \mapsto & \begin{cases} \pi_{j,i}(x)\psi(h_j x) & \text{for } x \in h_j^{-1}B_j, \\ 0 & \text{otherwise,} \end{cases} \end{array}$$

for $1 \leq i \leq m$ and $1 \leq j \leq p$, where $\pi_{j,i}$ is the i-th coordinate map of h_j and B_j are the open balls centered in $h_j(x_j)$. Thus, we define

$$y' = y + \sum_{j=1}^{p}\sum_{i=1}^{m} a_{j,i}\lambda_{j,i}.$$

We want $a = (a_{i,j})_{i,j}$ to be sufficiently small in norm and $\Phi_{(\phi,y;k)}$ be immersive in x_1. We shall take a such that $Dy'\phi^{-j}h_1^{-1}(h_1x_1) = v_j'$. Since M is Hausdorff, we can take the open balls B_j . Therefore, $\lambda_{j,i}(x)$ and $\lambda_{j',i'}(x)$ can not be both non-zero, for $j \neq j'$. Hence, the derivatives are

$$Dy'h_1^{-1}(h_1x_1) = Dyh_1^{-1}(h_1x_1) + a_{1,\bullet}.$$

Hence, $a_{1,\bullet} = v_1' - v_1$. In addition,

$$\begin{aligned} Dy'\phi^{-(c-1)}h_1^{-1}(h_1x_1) &= Dy'h^{-1}_{p-(c-1)+1}h_{p-(c-1)+1}\phi^{-(c-1)}h_1^{-1}(h_1x_1) \\ &= Dy'h^{-1}_{p-(c-2)}(h_{p-(c-2)}x_{p-(c-1)+1})Dh_{p-(c-2)}\phi^{-(c-1)}h_1^{-1}(h_1x_1) \\ &= D(y + \textstyle\sum_{j,i} a_{j,i}\lambda_{j,i})h^{-1}_{p-(c-2)}(h_{p-(c-2)}x_{p-(c-1)+1})A_c \\ &= (Dyh^{-1}_{p-(c-2)}(h_{p-(c-2)}x_{p-(c-1)+1}) + a_{c,i})A_c \\ &= Dy\phi^{-(c-1)}h_1^{-1}(h_1x_1) + a_{c,i}A_c \\ &= v_c + a_{c,i}A_c \end{aligned}$$

where $A_c = Dh_{p-(c-2)}\phi^{-(c-1)}h_1^{-1}(h_1x_1)$, $v_c = Dy\phi^{-(c-1)}h_1^{-1}(h_1x_1)$ and the fifth equality holds since

$$\begin{aligned} Dy\phi^{-(c-1)}h_1^{-1}(h_1x_1) &= Dyh^{-1}_{p-(c-2)}h_{p-(c-2)}\phi^{-(c-1)}h_1^{-1}(h_1x_1) \\ &= Dyh^{-1}_{p-(c-2)}(h_{p-(c-2)}x_{p-(c-2)})Dh_{p-(c-2)}\phi^{-(c-1)}h_1^{-1}(h_1x_1). \end{aligned}$$

Note that $Dh_\delta\phi^\rho h_1(h_1x_1) \in \mathbb{R}^{m\times m}$ and since ϕ is a diffeomorphism, $h_1, h_1^{-1} \in \mathcal{C}^1$, then A_c is invertible and $a_{c,i} = (v_c' - v_c)A_c^{-1}$.

Therefore, $\Phi_{(\phi,y';k)}$ is an immersion in x_1, and we can take a with arbitrarily small norm. In a neighborhood of y, there is some neighborhood $\mathcal{N}$ such that every $y' \in \Phi_{(\phi,y;k)}$ is an immersion on x_1. If we repeat this step for every periodic point of period p, we can make $\Phi_{(\phi,y;k)}$ immersive in all the points of period p, since we have a finite number of periodic points of period p. Finally, we repeat this step for every point of period $\leq l$ to embed the whole P_l.

We have discussed the case when J has real eigenvalues. It remains to prove the case of J having complex eigenvalues. In this case, if $\lambda_l \in \mathbb{C}$ is an eigenvalue, $\overline{\lambda}_l$ too. Let λ_l, $l = 1, \ldots, m$ be the eigenvalues:

$$\{\lambda_1, \ldots, \lambda_c, \lambda_{c+1}, \ldots, \lambda_{2c}, \lambda_{2c+1}, \ldots, \lambda_m\},$$

with $\overline{\lambda}_l = \lambda_{c+l}$, for $l = 1, \ldots, c$, and $\lambda_i \in \mathbb{R}$, for $l > 2c$. If $1 \leq l \leq c$, then $\lambda_l = \alpha_l + i\beta_l$

and $\lambda_{c+l} = \alpha_l - i\beta_j$. We use the basis $\{e_1, \overline{e}_1, \ldots, e_c, \overline{e}_c, e_{2c+1}, \ldots, e_m\}$. Therefore

$$\begin{aligned} v_j J^\rho &= \gamma_{j1} e_1 J^\rho + \overline{\gamma}_{j1} \overline{e}_1 J^\rho + \cdots + \gamma_{jc} e_c J^\rho + \overline{\gamma}_{jc} \overline{e}_c J^\rho + \cdots + \gamma_{jm} e_m J^\rho \\ &= \gamma_{j1} \lambda^\rho e_1 + \overline{\gamma}_{j1} \overline{\lambda}^\rho \overline{e}_1 + \cdots + \gamma_{jc} \lambda^\rho e_c + \overline{\gamma}_{jc} \overline{\lambda}^\rho \overline{e}_c + \cdots + \gamma_{jm} \lambda^\rho e_m. \end{aligned}$$

In this case, the Jacobian is

$$\begin{pmatrix} v_1 \\ \vdots \\ v_p J^{\rho_p} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \overline{\gamma}_{11} & \cdots & \gamma_{1,2c+1} & \cdots & \gamma_{1,m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma_{p1} & \overline{\gamma}_{p1} & \cdots & \gamma_{p,2c+1} & \cdots & \gamma_{p,m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma_{11}\lambda_1^{\rho_p} & \overline{\gamma}_{p1}\overline{\lambda_1^{\rho_p}} & \cdots & \gamma_{1,2c+1}\lambda_{2c+1}^{\rho_p} & \cdots & \gamma_{p,m}\lambda_p^{\rho_p} \end{pmatrix} \begin{pmatrix} e_1 \\ \overline{e}_1 \\ \vdots \\ e_{2c+1} \\ \vdots \\ e_m \end{pmatrix}$$

and if we permute the rows and we set all the γ's conveniently, we recall again a Vandermonde matrix. □

3.4 Immersion on M

Let $f \in \mathcal{C}^1(M, \mathbb{R}^k)$. If its derivative at some point $p \in M$ is injective, there exists some neighbourhood U of p such that $f : U \to f(U)$ is an embedding, by the Inverse Function Theorem. Therefore, since P_l is an injective immersion, for every $x_i \in P_l$ there exists a neighborhood that is embedded in $\mathbb{R}^k$, by $\Phi_{(\phi,y;k)}$. We recall from Whitney's Embedding Theorem E.1 that every compact manifold is metrizable. Thus, we can choose a neighbourhood $b_i(r_i, x_i)$, that is an open ball in the manifold centered in x_i with radius r_i. Hence, given $r > 0$, for all r_i P_l is an injective immersion in the neighbourhood. As a result, it does not map two points in the same image. However, we may find balls whose images intersect. Nevertheless, it is possible to find disjoint balls. We prove it in the following.

Proposition 3.5. Let $\Phi_{(\phi,y;k)}$ be an injective immersion on P_l. For every pair $x_i, x_j \in P_l$, there exist two open balls B_i and B_j such that the images do not intersect.

Proof. Since $\Phi_{(\phi,y;k)}$ is a continuous application in a metric space, it satisfies the equivalence of continuity by successions. Suppose that every ball $b_i(r_i, x_i)$ and $b_j(r_j, x_j)$ always have some pair of points with the same image. Thus, we have two successions $\{z_\lambda\}_{\lambda \in \mathbb{Z}^+}$ and $\{\hat{z}_\lambda\}_{\lambda \in \mathbb{Z}^+}$ such that

$$\begin{aligned} z_\lambda &\in b_i(1/\lambda, x_i) \\ \hat{z}_\lambda &\in b_j(1/\lambda, x_j) \end{aligned}$$

where $z_\lambda \to x_i$, $\hat{z}_\lambda \to x_j$ and $\Phi_{(\phi,y;k)}(z_\lambda) = \Phi_{(\phi,y;k)}(\hat{z}_\lambda)$. Then, by continuity, we have

$$\lim_{\lambda\to\infty} \Phi_{(\phi,y;k)}(z_\lambda) = \lim_{\lambda\to\infty} \Phi_{(\phi,y;k)}(\hat{z}_\lambda) \Longrightarrow \Phi_{(\phi,y;k)}(x_i) = \Phi_{(\phi,y;k)}(x_j),$$

but $\Phi_{(\phi,y;k)}$ is injective in P_l: contradiction. Hence, there is some pair of disjoint balls centered in x_i and x_j, such that their images do not intersect. □

Corollary 3.4. Let P_l be an injective immersion on $\Phi_{(\phi,y;k)}$. Then for every $x_i \in P_l$, there exists an open set $b_i(r_i, x_i)$ such that if $x_j \in P_l$ $(x_j \neq x_i)$, with $b_j(r_j, x_j)$, then the images of the open sets are disjoint.

Proof. For every x_i and every x_j $(x_i \neq x_j)$, by Proposition 3.5, there exists a ball B_i with radius $r_{i(j)}$. Since P_l is finite, we choose $r_i = \min_j\{r_{i(j)}\}$. □

All in all, we have that $\Phi_{(\phi,y;k)}$ is an injective immersion on the union of the $b_j(r_j, x_j)$'s. Then, for every $b_j(r_j, x_j)$, we can choose another ball b_i centered in the points P_l and contained into these $b_j(r_j, x_j)$'s. The union of the closure of these b_j's a is closed and bounded set, thus it is a compact set by Theorem 2.1. Since the open sets depend on y, this set also depends on y. We denote the closure subsets as V_y. V_y is a compact neighborhood that contains P_l. We define $V_y = \cup_i \bar{b}_i$.

The goal is to make an immersion in the whole M. We start by covering M with compact sets and make a delay map which is an immersion of one of these sets. There, we can make another delay map which is an immersion in the other set, without modifying the first immersion, and we repeat the process until we cover the entire M.

Now, we announce the theorem. We give a condition of smoothness and for the first time a condition of dimensionality. As we have previously exposed, to make an immersion it is necessary to arrive at dimension $2m$: hence, it is the dimension condition. Moreover, because of Lemma 2.6, we must consider $y \in \mathcal{C}^2(M, \mathbb{R})$.

Theorem 3.2. Let M be a compact manifold of dimension m. Let $\phi : M \to M$ be a diffeomorphism, with the following properties: firstly, the periodic points of ϕ with periods $k \leq 2m$ are finite in number, and secondly, if x is a periodic point with period $k \leq 2m$, then the eigenvalues of the Jacobian matrix of ϕ^k at x are all distinct. Then for generic $y \in \mathcal{C}^2(M.\mathbb{R})$, the map $\Phi_{(\phi,y;k)}$ is an immersion, for $k \geq 2m$.

It resembles the first version of Takens' Theorem we are proving. In this case, we do not need that $k \geq 2m+1$; we only need $k \geq 2m$. The reason is that immersions allow auto-intersections. Usually, to break the auto-intersections, we must arrive to dimension $2m + 1$.

Proof. Consider the set V_y defined as in the previous discussion, $V_y = \cup_i \bar{b}_i$. Since V_y is compact, by Proposition 3.1, the delay embeddings of the set form an open set. Thus, there is a neighbourhood of y, say $\mathcal{U}_y \subseteq \mathcal{C}^2(M, \mathbb{R})$, such that for every $\hat{y} \in \mathcal{U}_y$, $\Phi_{(\phi,\hat{y};k)}$ is an embedding of V_y.

Consider an atlas of M. We construct a new atlas, by intersecting all the $\{b_i\}_{i\in I}$ with every local chart of the given atlas. With this atlas, by Lemma 2.8 we make a regular covering $\{(U_i, h_i)\}_{i\in I}$ such that:

- Every $U_i \subseteq b_i$,
- $U_i = h_i^{-1} = B(0, 3)$,
- $W_i = h_i^{-1} = B(0, 1)$ still cover M, and hence cover P_k. Furthermore, $\overline{W}_i$ is a compact subset of $\bar{b}_i$.

We can find an atlas for P_k^c. Since P_k^c is the complementary of a closed set, it is an open set and hence for each element x, there exists a neighbourhood of x, say $U_x \subset P_k^c$. Moreover, the set of points $\{x, \phi x, \ldots, \phi^k x\}$ are all distinct, because if x is periodic, it has at least period greater than k, by the hypothesis of the theorem. By the Hausdorff property, we can choose U_x sufficiently small such that the sets $U_x, \phi U_x, \ldots, \phi^k U_x$ are disjoint. For every U_x, we can suppose it is contained in some local chart (we only have to intersect with some local chart that contains x) and hence for every $x \in P_k^c$ we obtain a new local chart, (U_x, h_x), with $U_x = h_x^{-1} B(0, 3)$. From this, we may construct a new regular covering for P_k^c, with $W_x = h_x^{-1} B(0, 1)$ a cover of the set.

Thus, we have two families of sets: $\{W_x\}_{x\in P_k^c}$ and $\{W_i\}_{x_i\in P_k}$. The union of them is an open cover of M. Since M is compact, only finitely many of them cover M, and since every $x_i \in P_k$ is only contained in W_i, this subcover contains every one of these sets. Let β be the number of points in P_k. We relabel the sets such that W_i, $1 \leq i \leq \beta$ are the sets containing the periodic points, and W_i, $\beta < i \leq l$ are the sets contained in P_k^c. The corresponding charts (U_i, h_i) form the required atlas.

By construction, for every $\hat{y} \in \mathcal{U}_y$, $\Phi_{(\phi,\hat{y};k)}$ is an embedding of the compact set of the $\cup_{i=1}^{\beta} \overline{W}_i \subset \cup_{i=1}^{\beta} \bar{b}_i$. Hence, it is also an immersion. We shall adjust the measurement function to make an immersion of the remaining $\overline{W}_i$'s.

Suppose that $i > \beta$ is the smaller index for which $\Phi_{(\phi,y;k)}$ fails to be an immersion of $\overline{W}_i$. Let $x \in U_i$. The jacobi matrix is

$$D\Phi_{(\phi,y)}(h_i x) = \begin{pmatrix} \frac{\partial y h_i^{-1}}{\partial u_1}(u) & \frac{\partial y h_i^{-1}}{\partial u_2}(u) & \cdots & \frac{\partial y h_i^{-1}}{\partial u_m}(u) \\ \frac{\partial y \phi h_i^{-1}}{\partial u_1}(u) & \frac{\partial y \phi h_i^{-1}}{\partial u_2}(u) & \cdots & \frac{\partial y \phi h_i^{-1}}{\partial u_m}(u) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y \phi^k h_i^{-1}}{\partial u_1}(u) & \frac{\partial y \phi^k h_i^{-1}}{\partial u_2}(u) & \cdots & \frac{\partial y \phi^k h_i^{-1}}{\partial u_m}(u) \end{pmatrix}, \qquad (3.5)$$

where $u = h_i(x)$. For some $u \in h_i\overline{W}_i$, the matrix (3.5) has not full rank. We shall make it full ranked by some small perturbation of y. Suppose that the first s columns of the matrix (3.5) are linearly independent for all $u \in h_i\overline{W}_i$. Let $\lambda : \mathbb{R}^m \to \mathbb{R}$ be a bump function equal to 1 on $\overline{B}(0,1)$ and having support in $B(0,2)$. We recall that $W_i = h_i^{-1}B(0,1)$. Let $\mu_j : U_i \to \mathbb{R}$ be the coordinate functions of h_i, for $j = 1, \ldots, m$. We define

$$\begin{array}{rccl} \psi : & M & \to & \mathbb{R} \\ & x & \mapsto & \begin{cases} \mu_{s+1}(x)\lambda(h_i(x)) & \text{if } x \in U_i, \\ 0 & \text{otherwise.} \end{cases} \end{array}$$

If $x \in \overline{W}_i$, then $x \in \overline{h_i^{-1}B(0,1)} = h_i^{-1}\overline{B}(0,1)$. Thus, $h_i(x) \in \overline{B}(0,1)$. This means that $\lambda(h_i(x)) = 1$ and hence $\psi(x) = \mu_{s+1}(x)$. Furthermore, ψ has support in U_i and hence $\psi \circ \phi^{-j}$ has support in $\phi^j U_i$. Thus, we define $\psi_j = \psi \circ \phi^{-j}$ for $0 \le j \le k$. We have seen that the sets $U_x, \phi U_x, \ldots, \phi^k U_x$ are disjoint, thus the ψ_j's have disjoint support.

Now, we construct a measurement function

$$y' = y + \sum_{j=0}^{k} a_{j+1}\psi_j.$$

Note that $y' \in \mathcal{C}^2$, because μ_{s+1} is a coordinate function of h_i and thus it is at least $\mathcal{C}^2$, the bump function is $\mathcal{C}^\infty$ and products and compositions of $\mathcal{C}^2$ functions are also a $\mathcal{C}^2$ function. If $x \notin U_i, \phi U_i, \ldots, \phi^k U_i$, then $\psi_j = 0$ for every $j = 0, \ldots, k$ and hence $y'(x) = y(x)$.

We recall that if $u \in h_i\overline{W}_i$, then $h_i^{-1}(u) \in \overline{W}_i$. Since $\overline{W}_i \subset U_i$, we have $h_i^{-1}(u) \in U_i$

and therefore $\phi^\beta h_i^{-1}(u) \in \phi^\beta U_i$. Hence:

$$\begin{aligned} y'\phi^\beta h_i^{-1}(u) &= y\phi^\beta h_i^{-1}(u) + a_{\beta+1}\psi_\beta\phi^\beta h_i^{-1}(u)) \\ &= y\phi^\beta h_i^{-1}(u) + a_{\beta+1}\psi(\phi^{-\beta}\phi^\beta h_i^{-1}(u)) \\ &= y\phi^\beta h_i^{-1}(u) + a_{\beta+1}\psi(h_i^{-1}(u)) \\ &= y\phi^\beta h_i^{-1}(u) + a_{\beta+1}\mu_{s+1}(h_i^{-1}(u)) \\ &= y\phi^\beta h_i^{-1}(u) + a_{\beta+1}u_{s+1}. \end{aligned}$$

Therefore, if we derivate the last equality:

$$\frac{\partial y'\phi^\beta h_i^{-1}}{\partial u_{s+1}}(u) = \frac{\partial y\phi^\beta h_i^{-1}}{\partial u_{s+1}}(u) + a_{\beta+1}.$$

Hence, we see that we only modify the $s+1$-th column of (3.5), and the effect is to add the vector $(a_1, \dots, a_{k+1})^T$. That is true for all $u \in h_i\overline{W}_i$. If $u \notin h_i\overline{W}_i$, we have $\psi_j = 0$ and hence $y'(x) = y(x)$.

Let $x \in U_i$. We have that the first s columns of the matrix (3.5) are full ranked, for every $x \in \overline{W}_i$. Let $J_s(x)$ be the matrix formed with these columns. Since the differential is continuous, $J_s : U_i \to \mathbb{R}^{(k+1)\times s}$ is a continuous function. We know (Example 2.5) that the full rank matrices form an open subset of this space. Hence, there is a neighborhood $X \subset U_i$, with $\overline{W}_i \subset X$, such that for every point in X, the first s columns of (3.5) are linearly independent.

Now we define

$$\begin{aligned} \mathcal{S}: \quad & \mathbb{R}^s \times X \quad \to \quad \mathbb{R}^{k+1} \\ & (\alpha_1, \dots, \alpha_s, x) \mapsto \sum_{j=1}^{s} \alpha_j \begin{pmatrix} \frac{\partial y h_i^{-1}}{\partial u_j}(u) \\ \frac{\partial y\phi h_i^{-1}}{\partial u_j}(u) \\ \vdots \\ \frac{\partial y\phi^k h_i^{-1}}{\partial u_j}(u) \end{pmatrix} - \begin{pmatrix} \frac{\partial y h_i^{-1}}{\partial u_{s+1}}(u) \\ \frac{\partial y\phi h_i^{-1}}{\partial u_{s+1}}(u) \\ \vdots \\ \frac{\partial y\phi^k h_i^{-1}}{\partial u_{s+1}}(u) \end{pmatrix}, \end{aligned}$$

where $u = h_i(x)$. Since y and ϕ are $\mathcal{C}^2$ by hypothesis, the derivatives are $\mathcal{C}^1$ and hence $\mathcal{S} \in \mathcal{C}^1$. Moreover, the matrix (3.5) has m columns and this implies that $s \leq m-1$. The dimension of $\mathbb{R}^s \times X \subset \mathbb{R}^s \times \mathbb{R}^m$ is at most $2m-1$. By Lemma 2.6, the complement of $\mathcal{S}(\mathbb{R}^s \times X)$ is dense in $\mathbb{R}^\beta$, for $\beta \geq 2m$. Hence, we can find a vector $a = (a_1, \dots, a_\beta)^T \in \mathbb{R}^\beta$ such that $a \notin \mathcal{S}(\mathbb{R}^s \times X)$. For this choice, the first $s+1$ columns of the Jacobi matrix 3.5 must be linearly independent for all $x \in \overline{W}_i$.

We repeat the process until we make all the columns linearly independents. Moreover, since we can find immersions of $\overline{W}_i$ using arbitrary perturbations, by Lemma 3.2 we can find y which is an immersion of any $\overline{W}_j$, with $j < i$, and also in i. Therefore, we repeat the argument for all $i = s, \ldots, l$ and thus we get an immersion on M. $\square$

3.5 Orbit Segments

We have just seen that the map $\Phi_{(\phi,y;k)}$ is generically an immersion, for $k \geq 2m$. We have to see that it is also generically an embedding, for some k sufficiently big. We know that by compactness we only need to show that $\Phi_{(\phi,y;k)}$ is injective. We need an intermediate step. We recall the next definition:

Definition 3.1. Let M be a manifold, $x \in M$ and $\phi : M \to M$. We call an *orbit segment of x by ϕ* to the collection of points $\{x, \phi x, \ldots, \phi^l x\}$, for some $l \in \mathbb{N}$.

We only consider the case where $l = 2m$. We note that, if we check injectivity, we look for pairs (x, x') such that $\Phi_{(\phi,y;k)}(x) \neq \Phi_{(\phi,y;k)}(x')$. However, if x and x' are points that are on the same periodic orbit of period less than or equal to $4m$, the orbit segments of x and x' overlap: that is, $x = \phi^j x'$ or $x' = \phi^r x$, for some $0 \leq j, r \leq 2m$. Hence, we first create a delay map with the property that for every x in an orbit segment, $x \notin P_{2m}$, it does not share an image under $\Phi_{(\phi,y;k)}$. It follows from the next lemma.

Lemma 3.3. Let y' be such that $\Phi_{(\phi,y';k)}$ is an injective immersion on V_y. In every neighbourhood of y' in $\mathcal{C}^2(M, \mathbb{R})$ there is a function, say y'', such that for every $x \in M$, and j in the range $1 \leq j \leq 2m$, $\Phi_{(\phi,y'';k)}(x) \neq \Phi_{(\phi,y'';k)}(\phi^j x)$, unless $x = \phi^j x$.

Proof. Let j be the smallest value for which the lemma is not true. Let $S = \cap_{i=0}^{2m} \phi^{-i} V_y$. If $x \in S$, $x \in \phi^{-r} V_y$ and $\phi^r x \in V_y$ for every $0 \leq r \leq 2m$. Furthermore, since V_y is an embedding, we have that $\Phi_{(\phi,y'';k)}(x) \neq \Phi_{(\phi,y'';k)}(x')$ for every y'' in a neighbourhood of y'.

Let $T = \overline{S^c}$. We note that $T \cup S = M$. Hence, if we prove for T we finish the proof. We recall that S is a neighbourhood of P_{2m}. Therefore, if $x \in T$, then $x \notin P_{2m}$, since the points of the boundary are not interior points. Therefore, the orbit segment of x by ϕ, that is $\{x, \phi x, \ldots, \phi^{2m} x\}$, are all different points. By the Hausdorff property, we may take a neighbourhood of x, we call it U_x, such that $U_x, \phi U_x, \dots, \phi^{2m}U_x$ are all disjoint. We need to consider two cases:

- Case 1: If x is not a periodic point whose period is between $2m+1$ and $4m$. Then we can find a neighbourhood of x, U_x such that $U_x, \phi U_x, \dots, \phi^{4m}U_x$ are all disjoints. As usual, we can assume that U_x is the domain of a chart from a regular covering $\{(U_x, h_x)\}$ and $V_x = h_x^{-1}(B(0,1))$. Without loss of generality and since we use it in the following case, we call $V_x = X_x$ in case 1.
- Case 2: If x has period r, where $2m+1 \leq r \leq 4m$. We now find U_x such that $U_x, \dots, \phi^{r-1}U_x$ are all disjoint and again we take U_x as the domain from a regular covering, with V_x the same as in case 1. We define $X_x = V_x \cap \phi^{-r}V_x$. It is clear that $x \in X_x$, because $x \in V_x$ and $\phi^r x = x$. Moreover, X_x is an open set, since it is the intersection of two open sets.

We note that in the two cases if $2m+j < r$, none of the sets $\phi^{2m+1}\overline{X}_x, \dots, \phi^{2m+j}\overline{X}_x$ intersect $\cup_{l=0}^{2m}\phi^l U_x$, since $V_x \subset U_x$ strictly, and moreover by regular covering $\phi^p \overline{V}_x \subset \phi^p U_x$. Furthermore, in case 2, if $2m+j \geq r$, none of $\phi^{2m+1}\overline{X}_x, \dots, \phi^{r-1}\overline{X}_x$ intersect again at $\cup_{l=0}^{2m}\phi^l U_x$. In addition, since $\phi^r \overline{X} = \phi \overline{V}_x \cap \overline{V}_x$, we have the set of inclusions, $\phi^r \overline{X}_x \subset \overline{V}_x, \phi^{r+1}\overline{X}_x \subset \phi \overline{V}_x, \dots, \phi^{2m+j}\overline{X}_x \subset \phi^{2m+j-r}\overline{V}_x$.

Since T is a closure set, it is also a closed set contained in a compact set. Therefore, it is a compact set. Hence, from the open cover of T, $\{X_x : x \in T\}$, we can extract a finite cover $\{X_i\}_{i=1,\dots,N}$, with $\{(U_i, h_i)\}$ its corresponding charts. Suppose that for $1 \leq i' < i$, and every $x \in X_{i'}$, $\Phi_{(\phi,y')}(x) \neq \Phi_{(\phi,y')}(\phi^j x)$ for every $1 \leq j \leq 2m$, but there is some j such that $\Phi_{(\phi,y')}(x) = \Phi_{(\phi,y')}(\phi^j x)$, for $x \in X_i$. We define

$$\begin{aligned} \lambda : \; M &\to \mathbb{R} \\ x &\mapsto \lambda(x) = \begin{cases} \psi(h_i x) & \text{if } x \in U_i, \\ 0 & \text{if } x \notin U_i. \end{cases} \end{aligned}$$

Moreover, we define $\lambda_l = \lambda \circ \phi^{-l}$, for $l = 0, \dots, 2m$. Since $\operatorname{supp} \lambda \subset U_i$, then $\operatorname{supp} \lambda_l \subset \phi^l U_i$. We seek for y'' sufficiently close to y', and hence we define

$$y'' = y' + \sum_{l=0}^{2m} a_l \lambda_l. \tag{3.6}$$

If a_l is sufficiently small in norm then by Lemma 3.2, y'' is sufficiently close to y'. Hence we may show that it satisfies the conditions stated in the lemma.

For all $x \in \overline{X}_i$, we have $x \in \overline{V}_i$ and $\phi^l x \in \phi^l \overline{V}_i$. Thus, $\psi(h_i x) = 1$ and

$$\lambda_l(\phi^l x) = \lambda \circ \phi^{-l}(\phi^l x) = \lambda(x) = \psi(h_i x) = 1,$$

for $l = 0, \dots, 2m$. Hence

$$\begin{array}{rcl} y''(x) & = & y'(x) + a_0, \\ y''(\phi x) & = & y'(\phi x) + a_1, \\ \vdots & \vdots & \vdots \\ y''(\phi^{2m} x) & = & y'(\phi^{2m} x) + a_{2m}. \end{array}$$

We shall discuss the values of $y''(\phi^j x)$.

- Case 1, or Case 2 with $2m + j < k$. The points $\phi^j x, \phi^{j+1} x, \dots, \phi^{2m} x$ lay in $\phi^j \overline{V}_i, \phi^{j+1} \overline{V}_i, \dots, \phi^{2m} \overline{V}_i$, respectively. In this case, $\lambda_l(x) = 1$, for $l = j, \dots, 2m$. Moreover, recall that in Case 1 and 2 if $2m + j < k$, none of the sets $\phi^{2m+p} \overline{X}_x$, $p = 1, \dots, j$ intersect $\cup_{l=0}^{2m} \phi^l U_x$. Hence, $\phi^{2m+1} x, \dots, \phi^{2m+j} x$ lay outside $\cup_{l=0}^{2m} \phi^l U_x$. As a result, $\lambda_l(x) = 0$, for $l = 2m+1, \dots, 2m+j$. All in all, we have

$$\begin{array}{rcccc} y''(\phi^j x) & = & y'(\phi^j x) & + & aj, \\ \vdots & \vdots & \vdots & & \vdots \\ y''(\phi^{2m} x) & = & y'(\phi^{2m} x) & + & a_{2m}, \\ y''(\phi^{2m+1} x) & = & y'(\phi^{2m+1} x), & & \\ \vdots & \vdots & \vdots & & \\ y''(\phi^{2m+j} x) & = & y'(\phi^{2m+j} x). & & \end{array}$$

- Case 2 with $2m + j \geq r$. We have:

  (i) $\phi^j x \in \phi^j \overline{V}_i, \dots, \phi^{2m} x \in \phi^{2m} \overline{V}_i$.

  (ii) $\phi^{2m+1} x, \dots, \phi^{r-1} x$ are outside $\cup_{i=0}^{2m} \phi^l U_i$.

  (iii) $\phi^k x \in \overline{V}_i, \dots, \phi^{2m+j} x \in \phi^{2m+j-r} \overline{V}_i$. Hence

$$\begin{array}{rcccc} y''(\phi^j x) & = & y'(\phi^j x) & + & aj, \\ \vdots & \vdots & \vdots & & \vdots \\ y''(\phi^{2m} x) & = & y'(\phi^{2m} x) & + & a_{2m}, \\ y''(\phi^{2m+1} x) & = & y'(\phi^{2m+1} x), & & \\ \vdots & \vdots & \vdots & & \\ y''(\phi^{r-1} x) & = & y'(\phi^{r-1} x), & & \\ y''(\phi^r x) & = & y'(\phi^r x) & + & a_0, \\ \vdots & \vdots & \vdots & & \vdots \\ y''(\phi^{2m+j} x) & = & y'(\phi^{2m+j} x) & + & a_{2m+j-r}. \end{array}$$

We use the same argument with the two cases, altough we concentrate on Case 2. Note that for any $x \in \overline{X}_i$

$$\Phi_{(\phi,y'')}(x) - \Phi_{(\phi,y'')}(\phi^j x) = \Phi_{(\phi,y')}(x) - \Phi_{(\phi,y')}(\phi^j x) + Aa,$$

where $a = (a_0, \ldots, a_{2m})$ and A is a diagonal matrix such that it has 1's on the diagonal, a upper diagonal of 1 from the j's column and a lower diagonal of 1 from the $r-j+1$'s row. If $2m - j + 1 - (r - j) \geq m + 1$, A has at least rank $m + 1$; otherwise, it happens $r - j > m + 1$ or $2m - (2m - j + 1) > m + 1$. Either the first $r - j$ columns or the last $2m - (2m - j + 1)$ columns has at least rank $m + 1$, since they are triangular matrices. Hence, A has rank $r > m + 1$. Let L be the r-dimensional subspace of $\mathbb{R}^{2m+1}$ which is the image of $\mathbb{R}^{2m+1}$ under A. Let $P : \mathbb{R}^{2m+1} \to L$ be the orthogonal projection onto L. We define

$$\begin{array}{rccl} F: & U_i & \to & L \\ & x & \mapsto & P(\Phi_{(\phi,y')}(x) - \Phi_{(\phi,y')}(\phi^j x)). \end{array}$$

$F \in \mathcal{C}^2$ and then by Lemma 2.6, $L \setminus F(U_i)$ is dense. Let $b \in L - F(U_i)$ with arbitrarily small norm. Let $(\mathrm{Ker}A)^\perp$ be the orthogonal complement of the kernel of A. We state that there exists an unique $b' \in (\mathrm{Ker}A)^\perp$ such that $b = Ab'$. In fact, if there is another $d' \neq b'$, we have

$$Ab' - Ad' = A(b' - d') = 0 \Longrightarrow b' - d' \in \mathrm{Ker}A.$$

However, we know that $(\mathrm{Ker}A)^\perp$ is a subspace, thus $b' - d' \in (\mathrm{Ker}A)^\perp$ and in addition, $\mathrm{Ker}A \oplus (\mathrm{Ker}A)^\perp = \mathbb{R}^{2m+1}$. Hence, $b' - d' = 0$ and then $b' = d'$. Therefore, there is a unique $b' \in (\mathrm{Ker}A)^\perp$ such that $b = Ab'$. If $\|b\|$ is arbitrarily small, $\|b'\|$ is too. Hence, $\Phi_{(\phi,y'';k)}(x) \neq \Phi_{(\phi,y'';k)}(\phi^j x)$ for all $x \in \overline{X}_i$.

Let

$$\Lambda = \{y \in \mathcal{C}^2(M, \mathbb{R}) : \Phi_{(\phi,y;k)}(x) - \Phi_{(\phi,y;k)}(\phi^j x) \neq 0, \ \forall x \in \overline{X}_i\}.$$

Thus, we state that Λ is an open set. To prove it, we need to show that every $y \in \Lambda$ is an interior point. Consider $y \in \Lambda$. Then, for every $x \in \overline{X}_i$, $\Phi_{(\phi,y;k)}(x) - \Phi_{(\phi,y;k)}(\phi^j x) \neq 0$. Thus, there exists some coordinate s such that $y\phi^s x - y\phi^{s+j} x \neq 0$. By the sign conservation property, there exists an open set $\mathcal{U}_x$ such that for all $x' \in \overline{\mathcal{U}}_x$, $y\phi^s x' - y\phi^{s+j} x \neq 0$ and thus $\|y\phi^s x' - y\phi^{s+j} x'\| > 0$.

For every x, we have some s_x such that $y\phi^{s_x} x - y\phi^{s_x+j} x \neq 0$ and hence a family of neighborhoods $\{\mathcal{U}_x\}_{x \in \overline{X}_i}$ that covers $\overline{X}_i$. Since the set is compact, we can take a finite cover, say $\{\mathcal{U}_i\}_{i=1}^{\hat{n}}$. We take

$$\delta_i = \min_{x \in \overline{\mathcal{U}}_i} \{\|y\phi^{s_x}x - y\phi^{s_x+j}x\|\} > 0$$

and thus $\delta = \frac{1}{2}\min_{i=1,\ldots,\hat{n}}\{\delta_i\}$. Let $\mathcal{N} = \cap_l \mathcal{N}(y; (U_l, h_l), (\mathbb{R}, \mathrm{id}), C_l, \delta/2)$. We take U_l such that it has only one point of each orbit, as we have built previously. We want to prove that

$$\begin{aligned} \hat{y} \in \mathcal{N} \;&\Rightarrow\; \Phi_{(\phi,\hat{y};k)}(x) - \Phi_{(\phi,\hat{y};k)}(\phi^j x) \neq 0, \;\forall x \in \overline{X}_i \\ &\Leftrightarrow\; \text{There exists } s \text{ such that } \hat{y}\phi^s x - \hat{y}\phi^{s+j}x \neq 0, \;\forall x \in \overline{X}_i \\ &\Leftrightarrow\; \text{There exists } s \text{ such that } \|\hat{y}\phi^s x - \hat{y}\phi^{s+j}x\| \neq 0, \;\forall x \in \overline{X}_i \end{aligned}$$

Suppose that there exsits $x \in \overline{X}_i$ such that $\hat{y}\phi^s x - \hat{y}\phi^{s+j}x = 0$, for all s. Take $y\phi^s x - y\phi^{s+j}x \neq 0$. Then

$$\begin{aligned} \delta &< \|y\phi^s x - y\phi^{s+j}x\| = \|y\phi^s x - \hat{y}\phi^s x + \hat{y}\phi^s x - \hat{y}\phi^{s+j}x + \hat{y}\phi^{s+j}x - y\phi^{s+j}x\| \\ &\leq \|y\phi^s x - \hat{y}\phi^s x\| + \|\hat{y}\phi^s x - \hat{y}\phi^{s+j}x\| + \|\hat{y}\phi^{s+j}x - y\phi^{s+j}x\| \\ &< \delta/2 + \|\hat{y}\phi^s x - \hat{y}\phi^{s+j}x\| + \delta/2 = \delta + \|\hat{y}\phi^s x - \hat{y}\phi^{s+j}x\|. \end{aligned}$$

Therefore, $\|\hat{y}\phi^s x - \hat{y}\phi^{s+j}x\| > 0$ and this means $\hat{y}\phi^s x - \hat{y}\phi^{s+j}x \neq 0$, which is a contradiction. Hence, if $\hat{y} \in \mathcal{N}$, then $\hat{y} \in \Lambda$. Consequently, every $\hat{y} \in \Lambda$ is interior and finally Λ is an open set.

We can make a series of adjustments of the form (3.6) each of which establishes the property on each $\overline{X}_i$ and since we have a finite number of them, we make a finite number of adjustments to generate a function y'' such that $\Phi_{(\phi,y'';k)}(x) - \Phi_{(\phi,y'';k)}(\phi^j x) \neq 0$ for all $x \in T$. Moreover, we make again adjustments for every j until $j = 2m$, since it will be a finite number of adjustments. Therefore, for $x \in T$ and $1 \leq j \leq 2m$, $\Phi_{(\phi,y;k)}(x) \neq \Phi_{(\phi,y;k)}(\phi^j x)$. □

As well as we did with period points, we can extend the previous lemma on a neighbourhood of the orbit segments. First of all, we prove the following proposition.

Proposition 3.6. Let $\Phi_{(\phi,y')}$ be an immersion on M. There exists $\epsilon > 0$ and a neighborhood $\mathcal{U}'_y$ of y such that if $\hat{y} \in \mathcal{U}'_y$, then $\Phi_{(\phi,\hat{y})}$ is an immersion of M, an embedding of V_y and $\Phi_{(\phi,\hat{y})}(x) \neq \Phi_{(\phi,\hat{y})}(x')$, whenever $x \neq x'$ and $d(x,x') \leq \epsilon$, where d is the distance induced by the metric of the manifold.

Proof. If $\Phi_{(\phi,y')}$ is an immersion of M, then for each point $x \in M$ there is an open neighbourhood of x, we call N_x, such that $\Phi_{(\phi,y')}$ is an embedding of N_x, by Proposition

2.11. As a result of Whitney Embedding Theorem (see Corollary E.1), there is a metric over the manifold M. Hence, we can find a closed ball $\overline{\beta}_x$ centered at x and contained in N_x. Since we consider every $x \in M$, we have that $\cup_{x\in M}\beta_x = M$. Moreover, M is compact and therefore we only need a finite number of these open sets:

$$\bigcup_{i=1}^{n'} \beta_i = M \Longrightarrow \bigcup_{i=1}^{n'} \overline{\beta}_i = M.$$

Thus, $\overline{\beta}_i$ is a compact cover. Every $\overline{\beta}_i$ is embedded by $\Phi_{(\phi,y')}$, since it is contained in some N_x. We choose one of these $\overline{\beta}_i$. There is a neighborhood of y' such that, by Proposition 3.1, we have an open set of embeddings for every $\overline{\beta}_i$. We call $\mathcal{W}_i$. Since there is a finite number of $\overline{\beta}_i$, we have a finite number of open sets $\mathcal{W}_i$ where every $\overline{\beta}_i$ is an embedding: thus this intersection is an open set, $\hat{\mathcal{U}}_y = \cap_{i=1}^{n'}\mathcal{W}_i$. Since $y \in \mathcal{W}_i$, then $y \in \hat{\mathcal{U}}_y$. Moreover, since $y \in \mathcal{U}_y$, we may consider $\mathcal{U}'_y = \mathcal{U}_y \cap \mathcal{W}_y$, and then $\mathcal{U}'_y$ satisfies the same as $\mathcal{W}_y$ and furthermore $\mathcal{W}_y \subseteq \mathcal{U}_y$. By Lebesgue's Lemma C.1, there exists some number $\epsilon > 0$ (the Lebesgue's number) such that every closed ball of radius ϵ at any point of M is contained in the interior of $\overline{\beta}_i$ for at least one i. Therefore, for every $x \in M$, the ball $B(x,\epsilon)$ is embedded by $\Phi_{(\phi,\hat{y})}$, for every $\hat{y} \in \mathcal{U}'_y$.

All in all, if $\hat{y} \in \mathcal{U}'_y$, then $\Phi_{(\phi,y')}$ is an immersion of M, an embedding of V_y, since $y \in \mathcal{U}_y$, and $\Phi_{(\phi,\hat{y})}(x) \neq \Phi_{(\phi,\hat{y})}(x')$, whenever $x \neq x'$ and $d(x,x') \leq \epsilon$, since in the ball $B(x,\epsilon)$ it is an embedding and thus injective. □

Now, we can prove the the following lemma:

Lemma 3.4. Let y'' be a function as in Lemma 3.3. There is a number $\delta > 0$ such that: if $x, x' \in M$, $x \neq x'$, and $d(\phi^i x, \phi^j x') < \delta$ for some $0 \leq i, j \leq r$, then $\Phi_{(\phi,y'')}(x) \neq \Phi_{(\phi,y'')}(x')$.

Proof. We prove it by contradiction. Suppose that there is a sequence $\delta_n \to 0$ of positive numbers such that for each n, for every pair of points $x_n \neq x'_n$ and all the integers $0 \leq i_n, j_n \leq r$ such that $d(\phi^{i_n} x_n, \phi^{j_n} x'_n) < \delta_n$, we have $\Phi_{(\phi,y'')}(x) = \Phi_{(\phi,y'')}(x')$.

Since M is compact and a metric space, we have that it is also sequentially compact and therefore we can consider two sequences $\{x_n\}_{n\in\mathbb{N}} \to x$, and $\{x'_n\}_{n\in\mathbb{N}} \to x'$. Moreover, since the number of values of i_n and j_n goes from 0 to r, we can take only a finite number of them and thus we can also consider that the previous sequences have the same i_n and j_n values. For simplicity, say $i = i_n$ and $j = j_n$. Since ϕ

is a continuous function, its composition is also a continuous function and we have the sequential continuity. Hence, $\phi^i x_n \to \phi^i x$ and $\phi^j x'_n \to \phi^j x'$. Since $\delta_n \to 0$, $d(\phi^i x_n, \phi^j x'_n) \to 0$ and thus by continuity $\phi^i x = \phi^j x'$. Since $x \neq x'$, we have $i \neq j$. Let $i < j$ (the case $i > j$ is similar). Hence, $x = \phi^{j-i} x'$ and $r \geq j - i > 0$. Thus, x, x' belong to the same segment orbit. However, since $\Phi_{(\phi,y'')}$ is a continuous function and $x_n \to x$, $x'_n \to x'$, we have $\Phi_{(\phi,y'')}(x) = \Phi_{(\phi,y'')}(x')$.

Since x_n and x'_n tend to the same limit, for every $\epsilon > 0$, there exists n_0 such that $d(x_n, x'_n) < \epsilon$, for any $n \geq n_0$. If we take the Lebesgue's number of Proposition 3.6 we have that $\Phi_{(\phi,y'')}$ is an embedding of the ball $B(x_n, \epsilon)$. However, since $x_n \neq x'_n$, it can not happen that $\Phi_{(\phi,y'')}(x) = \Phi_{(\phi,y'')}(x')$ (since it is injective): this is a contradiction that comes from supposing that there is not a number $\delta > 0$ that satisfies the lemma. □

We have just proved that distinct points are not mapped to the same image by $\Phi_{(\phi,y'')}$ if their orbit segments are sufficiently close together. However, pairs of points that are not close among them require a different approach.

3.6 Injectivity on M

We shall extend the result in Lemma 3.4 to the other par of points of M. We recall that if $\Phi_{(\phi,y'')}$ is injective, then

$$\Phi_{(\phi,y'')}(x) \neq \Phi_{(\phi,y'')}(x') \Longrightarrow \Phi_{(\phi,y'')}(x) - \Phi_{(\phi,y'')}(x') \neq 0$$

for every pair of points $x \neq x'$. Hence, if $\Delta = \{(x,x) : x \in M\}$, $\Phi_{(\phi,y'')}$ is injective if the map

$$\begin{array}{rcl} (M \times M) \setminus \Delta & \to & \mathbb{R}^k \\ (x, x') & \mapsto & \Phi_{(\phi,y'')}(x) - \Phi_{(\phi,y'')}(x') \end{array}$$

does not contain the zero. We prove the injectivity part with this argument.

Let y'' be as in Lemma 3.4, with δ its Lebesgue's number. Consider V_y. The set $M \setminus \mathring{V}_y = M \cap (\mathring{V}_y)^c$ is a closed subset of a compact set, and hence a compact set. Let

$$\mathcal{Z} = \bigcup_{j=0}^{2m} \phi^j (M \setminus \mathring{V}_y).$$

Since $\phi^j(M \setminus \mathring{V}_y)$ is the image of a compact set by a continuous function, $\mathcal{Z}$ is a finite union of compact sets and hence compact. We claim the following proposition:

Proposition 3.7. Let $\mathcal{Z}$ as above. There exists a finite covering $\{U_l\}_{l=1}^N$, such that:

(i) For each $l = 1, \dots, N$ and $0 \leq i, j \leq 2m$, $\phi^{-i}U_l \cap \phi^{-j}U_l = \emptyset$ unless $i = j$.

(ii) For each $l = 1, \dots, N$, the diameter of U_l is less than δ.

Proof. (i) Since $P_{2m} \subset V_y$, then for every $x \in \mathcal{Z}$, the points $x, \phi^{-1}x, \dots, \phi^{-2m}x$ are all distinct. Therefore, since M is Hausdorff, we can find an open set U_x containing x such that $U_x, \phi^{-1}U_x, \dots, \phi^{-2m}U_x$ are all disjoint. Hence, we have that $\{U_x : x \in \mathcal{Z}\}$ satisfies the property 1.

(ii) For every $x \in \mathcal{Z}$, if we take $\hat{U}_x = U_x \cap B(x, \delta/2)$ the cover $\{\hat{U}_x\}_{x\in\mathcal{Z}}$ satisfies the property 1 and also the property 2.

Finally, we only have to reduce $\{U_x\}$ to a finite cover: but we recall that $\mathcal{Z}$ is a compact set and thus, we can choose a finite cover $\{U_l\}_{l=1}^N$ from the cover $\{\hat{U}_x\}_{x\in\mathcal{Z}}$. □

Let $\mathcal{Z}$ be as above. We build a partition of unity $\psi_l : M \to \mathbb{R}$, $l = 1, \dots, N$ on $\mathcal{Z}$, subordinate to a cover as in Proposition 3.7. Let

$$\begin{aligned} W = \; & \{(x, x') : d(\phi^i x, \phi^j x') \geq \delta, \text{ for all } 0 \leq 1, j \leq k, \\ & \text{and either } x \notin \overset{\circ}{V}_y \text{ or } x' \notin \overset{\circ}{V}_y\}. \end{aligned} \tag{3.7}$$

First of all, $W \subset M \times M$. Moreover, if $W_{i,j} = \{(x, x') : d(\phi^i x, \phi^j x) \geq \delta\}$, then

$$W = \bigcup_{0\leq i,j\leq k} W_{i,j} \cap \big(((\overset{\circ}{V}_y)^c \times M) \cup (M \times (\overset{\circ}{V}_y)^c)\big),$$

Every $W_{i,j}$ is a closed set, since d is a continuous map and the preimage by a continuous map of a compact set is also a closed set. Therefore, W is a finite union of closed sets, hence a closed set. Let

$$y_\epsilon = y'' + \sum_{i=1}^{N} \epsilon_l \psi_l,$$

for $\epsilon = (\epsilon_1, \dots, \epsilon_N) \in \mathbb{R}^N$. This will make $\Phi_{(\phi,y_\epsilon;k)}$ an injective map, for $k \geq 2m + 1$. To show that, we study the map

$$\begin{aligned} \Psi : \; & M \times M \times \mathbb{R}^N && \to && \mathbb{R}^k \\ & (x, x', \epsilon) && \mapsto && \Phi_{(\phi,y_\epsilon;k)}(x) - \Phi_{(\phi,y_\epsilon;k)}(x'). \end{aligned}$$

Firstly, we assign charts to $M \times M \times \mathbb{R}^N$. Let $\{(h_p, V_p)\}_{p\in\Lambda}$ be an atlas for M. Then, $\{(g_{p,q}, V_p \times V_q \times \mathbb{R}^N)\}_{p,q\in\Lambda}$, where $g_{p,q}(x, x', \epsilon) = (h_p(x), h_q(x'), \epsilon)$, is an atlas for $M \times M \times \mathbb{R}^N$. Let $h_p(x) = u$ and $h_q(x') = u'$. We study the Jacobian $D\psi g_{p,q}^{-1}(u, u', 0)$. Since

$$\psi g_{p,q}^{-1}(u, u', \epsilon) = \Phi_{(\phi,y_\epsilon;k)}(h_p^{-1}(u)) - \Phi_{(\phi,y_\epsilon;k)}(h_q^{-1}(u')),$$

then,

$$D\psi g_{p,q}^{-1}(u,u',\epsilon) = D\Phi_{(\phi,y_\epsilon;k)}(h_p^{-1}(u)) - D\Phi_{(\phi,y_\epsilon;k)}(h_q^{-1}(u')),$$

If $z=(z_1,\dots,z_s)$, we denote by f_z the derivative matrix of f with respect to $z_1,\dots,z_s$. Therefore, we have

$$\psi g_{p,q}^{-1}(u,u',0))_u = D\Phi_{(\phi,y_0;k)}(h_p^{-1}(u)) = D\Phi_{(\phi,y'';k)}(h_p^{-1}(u)),$$

$$\psi g_{p,q}^{-1}(u,u',0)_{u'} = -D\Phi_{(\phi,y_0;k)}(h_q^{-1}(u')) = -D\Phi_{(\phi,y'';k)}(h_q^{-1}(u')),$$

and for every ϵ_l,

$$\frac{d\Phi_{(\phi.y_\epsilon;k)}h_s^{-1}(\hat{u})}{\partial\epsilon_l} = \Big(\frac{\partial y_\epsilon h_s^{-1}(\hat{u})}{\partial\epsilon_l}, \frac{\partial y_\epsilon \phi h_s^{-1}(\hat{u})}{\partial\epsilon_l}, \dots, \frac{\partial y_\epsilon \phi^k h_s^{-1}(\hat{u})}{\partial\epsilon_l}\Big),$$

where $s=p,q$ and $\hat{u}=u,u'$. All in all, the Jacobian matrix is formed by the three following submatrices:

$$D\psi g_{p,q}^{-1}(u,u',0) = \Big(D\Phi_{(\phi,y'';k)}(h_p^{-1}(u)) \;\Big|\; -D\Phi_{(\phi,y'';k)}h_q^{-1}(u') \;\Big|\; A(x)-A(x')\Big), \quad (3.8)$$

where $A(x)$ is a $k\times N$ matrix whose elements are given by

$$A_{i,l}(x) = \frac{\partial y_\epsilon \phi^{i-1} h_s^{-1}}{\partial\epsilon_l}(u).$$

Furthermore,

$$\begin{aligned}\frac{\partial y_\epsilon \phi^{i-1} h_s^{-1}}{\partial\epsilon_l}(\hat{u}) &= \frac{\partial y'' + \sum_{i=0}^N \epsilon_i\psi_i}{\partial\epsilon_l} = \frac{\partial\epsilon_l\psi_l(\phi^{i-1}h_s^{-1})}{\partial\epsilon_l}(\hat{u}) \\ &= \psi^l\phi^{i-1}h_s^{-1}(\hat{u}).\end{aligned} \quad (3.9)$$

Our first assert is that the matrix (3.8) forms a basis for $\mathbb{R}^k$, when $(x,x')\in W$.

Proposition 3.8. Consider the Jacobian matrix (3.8). If $(x,x')\in W$, then the columns of (3.8) form a basis for $\mathbb{R}^k$, for $k\geq 2m+1$.

Proof. Since $\Phi_{(\phi,y'';k)}$ is immersive, we know that the first $2m$ columns of (3.8) span at least an m-dimensional subspace of $\mathbb{R}^k$. However, we discard this columns and concentrate on the submatrix $A(x)-A(x')$. We show that this matrix has at least k independent columns. From (3.9), we have

$$A_{il}(x) - A_{il}(x') = \psi_l\phi^{i-1}(x) - \psi_l\phi^{i-1}(x').$$

Firstly, we can show that each column of the matrix $A(x) - A(x')$ has at most one non-zero element. Suppose that for some l there are two different i, j such that $A_{il}(x) - A_{il}(x') \neq 0$ and $A_{jl}(x) - A_{jl}(x') \neq 0$. Therefore, at least one of $A_{il}(x)$, $A_{il}(x')$ was non-zero, and the same for $A_{jl}(x)$, $A_{jl}(x')$:

- $A_{il}(x)$ and $A_{jl}(x)$ cannot be both non zero. On the contrary, suppose that we have $A_{il}(x) = \psi\phi^{i-1}(x) \neq 0$ and $A_{jl}(x) = \psi\phi^{j-1}(x) \neq 0$. Hence, $\phi^{i-1}(x) \in \operatorname{supp}\psi_l$ and $\phi^{j-1}(x) \in \operatorname{supp}\psi_l$. Therefore, $\phi^{i-1}(x), \phi^{j-1}(x) \in U_l$ and hence $x \in \phi^{-(i-1)}U_l \cap \phi^{-(j-1)}U_l \neq \emptyset$: but we have taken U_l as in Proposition 3.7 and thus $\phi^{-(i-1)}U_l \cap \phi^{-(j-1)}U_l = \emptyset$. The same is true for $A_{il}(x')$ and $A_{jl}(x')$.
- $A_{il}(x)$ and $A_{jl}(x')$ cannot both be non-zero. On the contrary, suppose that $A_{il}(x) = \psi_l\phi^{i-1}(x) \neq 0$ and $A_{jl}(x') = \psi_l\phi^{j-1}(x') \neq 0$. Hence, $\phi^{i-1}(x), \phi^{j-1}(x') \in \operatorname{supp}\psi_l$ and thus $\phi^{i-1}(x), \phi^{j-1}(x') \in U_l$. Finally, by Proposition 3.7, the diameter of U_l is less than δ and then $d(\phi^{i-1}(x), \phi^{j-1}(x')) < \delta$: therefore, $(x, x') \notin W$ (recall the definition of W in (3.7)): but we had taken (x, x') in W. The same is true for $A_{il}(x')$ and $A_{jl}(x)$.

Therefore, if $A_{il}(x) \neq 0$, then $A_{jl}(x) = A_{jl}(x') = 0$, which contradicts the assumption that $A_{jl}(x) - A_{jl}(x') \neq 0$. Hence $A_{il}(x) = 0$, and then $A_{il}(x') \neq 0$; but again this implies that $A_{jl}(x) = A_{jl}(x') = 0$. We conclude that each column has at most one non-zero element.

Now we are going to show with similar arguments as above that every row of $A(x) - A(x')$ has at least one non-zero element. Since $(x, x') \in W$, at least one of x or x' is in $M \setminus \mathring{V}_y$. Without loss of generality, we assume that $x \in M \setminus \mathring{V}_y$. Then,

$$\phi^{i-1}(x) \in \phi^{i-1}M \setminus \mathring{V}_y \subset \bigcup_{j=0}^{2m} \phi^j(M \setminus \mathring{V}_y) = \mathcal{Z},$$

for $1 \leq i \leq 2m+1$. Since ψ_l comes from a partition of unity and $\mathcal{Z}$ is a closed set, we have $\sum_{l=1}^{N} \psi_l\phi^{i-1}(x) = 1$. Hence, for every $1 \leq i \leq 2m+1$, there must be some l, $1 \leq l \leq N$ such that $\psi_l\phi^{i-1}(x) \neq 0$ (if not, the sum cannot be one): that is, for every i there is an l such that $A_{il}(x) \neq 0$. If we suppose that $A_{il}(x') \neq 0$, then $\psi_l\phi^{i-1}(x) \neq 0$ and $\psi_l\phi^{i-1}(x') \neq 0$, that is $\phi^{i-1}(x)$ and $\phi^{i-1}(x')$ lay in $\operatorname{supp}\psi_l$. Thus, the two points belong to U_l and $d(\phi^{i-1}(x), \phi^{i-1}(x')) < \delta$, which implies that $(x, x') \notin W$. Since we have considered $(x, x') \in W$, we conclude that $A_{il}(x') = 0$ and $A_{il}(x) - A_{il}(x') \neq 0$. Therefore $A(x) - A(x')$ has at least one non-zero element in every row.

Since $A(x) - A(x')$ has at least one non-zero element in every row and at most one non-zero element in every column, we have:

- The matrix have at least as many columns as rows. Otherwise, some column should have more than two ones.
- The matrix must be full rank. That is, the rank of the matrix is the number of rows, $k \geq 2m + 1$, since we have as many columns as rows. We can choose a submatrix of dimension $k \times k$ and then it would have exactly a non-zero element for each row and each column, and thus the determinant would be non-zero.

$\square$

All in all, by Proposition 3.8, the rank of $D\Psi g_{p,q}^{-1}(u, u', 0)$ must be k and hence Ψ is submersive at $(x, x', 0)$, for every $(x, x') \in W$. Thus, for every $(x, x', 0)$ the derivative is full ranked and by continuity there is an open subset containing this point throughout which the derivative is full rank. The union of these open sets is also an open set $X \subset M \times M \times \mathbb{R}^N$ that covers W and $\Psi|_X$ is a submersion. By Lebesgue's Lemma, there is an $\eta > 0$ such that every closed ball $\overline{B((x, x', 0), \hat{\epsilon})} \subset X$, with $\hat{\epsilon} \leq \eta$. Thus, if $\|\epsilon\| < \eta$, then $d((x, x', 0), (x, x', \epsilon)) \leq \eta$ and $W \times \{\epsilon\} \subset X$.

Since $\Psi|_X : X \to \mathbb{R}^k$ is a submersion, by Lemma 2.7 we have that $\Psi|_X^{-1}(0)$ is a submanifold of X, with dimension

$$\dim(\Psi|_X^{-1}(0)) = \dim(M \times M \times \mathbb{R}^N) - \dim(\mathbb{R}^k) = 2m + N - k \leq N - 1.$$

Consider the projection

$$\begin{array}{rccc} \pi : & X & \to & \mathbb{R}^N \\ & (x, x', \epsilon) & \mapsto & \epsilon, \end{array}$$

and its restriction $\hat{\pi} = \pi|_{\Psi|_X^{-1}(0)}$. We note that $\hat{\pi}$ is a map between a manifold of dimension $2m+N-k$ to $\mathbb{R}^N$, with a greater dimension N. By Lemma 2.6 we have that $\mathbb{R}^N \backslash \hat{\pi}(\Psi|_X^{-1}(0))$ is dense. Thus, we have some ϵ with arbitrarily small norm such that ϵ is not in the range of $\hat{\pi}$. Thus, for pairs $(x, x') \in W$ we have $\Psi_{(\phi, y_\epsilon; k)}(x) - \Psi_{(\phi, y_\epsilon; k)}(x') \neq 0$.

All in all we have:

- A neighborhood $\mathcal{N}_1$ such that the pairs (x, x') with $d(\phi^i x, \phi^j x') < \delta$ for some $0 \leq i, j \leq r$, satisfy $\Psi(x, x', \epsilon) \neq 0$, by Lemma 3.4.
- A neighborhood $\mathcal{N}_2$ such that if $x, x' \in V_y$, then $\Psi(x, x', \epsilon) \neq 0$.

- Hence, (x, x') is from the two previous sets or $(x, x') \in W$. Thus, we can take ϵ arbitrarily small in norm such that $y_\epsilon \in \mathcal{N}_1 \cap \mathcal{N}_2$ and $W \times \{0\}$ under Ψ does not contain 0.

Finally, this means $\Phi_{(\phi, y_\epsilon; k)}$ is injective on M. Since it is an injective immersion, it is an embedding. Now, we can state the Restrictive Takens' Theorem:

Theorem 3.3 (Restrictive Takens' Embedding Theorem)**.** Let M be a compact manifold of dimension m. For pairs (ϕ, y), with $\phi \in \mathrm{Dif}^2(M)$, $y \in \mathcal{C}^2(M, \mathbb{R})$, it is a generic property that the map $\Phi_{(\phi,y;k)}$ is an embedding, for $k \geq 2m + 1$.

3.7 Transversality

In this section we present an advanced result in Differential Topology concerning periodic points. We will prove that nondegenerate periodic points are also generic. We follow mainly the lecture of [16].

Definition 3.2. Let M and N be manifolds, and $p \in M$. Suppose $f, g : M \to N$ are smooth maps with $f(p) = g(p) = q$.

(i) f has first order contact with g at p if $(df)_p = (dg)_p$ as mapping of $T_pM \to T_qN$.

(ii) f has kth order contact with g at p if $(df) : TM \to TN$ has $(k-1)st$ order contact with (dg) at every point in T_pM. This is written as $f \sim_k g$ at p.

(iii) Let $J^k(M, N)_{p,q}$ be the set of equivalence classes under "$\sim_k$ at p" of mappings $f : M \to N$, where $f(p) = q$.

(iv) Let $J^k(M, N) = \bigcup_{(p,q)\in M\times N} J^k(M, N)_{(p,q)}$. It is a disjoint union. An element $\sigma \in J^k(M, N)$ is called a *k-jet* from M to N.

(v) The *k-jet of* f is the map

$$\begin{array}{rccl} j^k f : & M & \to & J^k(M, N)_{p,f(p)} \\ & p & \mapsto & j^k f(p) \end{array}$$

where $j^k f(p)$ is the equivalence class of f in $J^k(X, Y)_{p,f(p)}$.

(vi) Let $X^{(s)} = \{(x_1, \ldots, x_s) \in X^s : x_i \neq x_j, \text{ for } i \neq j\}$. We call the source map

$$\begin{array}{rll} \alpha: & J^k(M,N) & \to M \\ & \sigma \in J^k(M,N)_{p,q} & \mapsto p. \end{array}$$

α is well-defined, since $J^k(M,N)$ is a disjoint union. This map can be naturally extend to $\alpha^s : J^k(M,N)^s \to M^s$. It is direct to prove that $f \sim_k g$ is an equivalence relation. Furthermore, note that $J^0(M,N) = M \times N$, hence f has $\sim_0$ contact with g at p if, and only if, $f(p) = g(p)$, and $j^0 f(p) = (p, f(p))$ is the graph of f.

(vii) Let $J^k_s(M,N) = (\alpha^s)^{-1}(M^s)$. It is a subset of $J^k(M,N)^s$. We extend in the natural way a multidimensional jet as

$$\begin{array}{rll} j^k_s: & M^{(s)} & \to J^k_s(M,N) \\ & (x_1,\ldots,x_s) & \mapsto j^k_s f(x_1,\ldots,x_s) = (j^k f(x_1),\ldots,j^k f(x_s)). \end{array}$$

Definition 3.3. Let M and N be smooth manifolds and $f : M \to N$ a smooth mapping. Let W be a submanifold of N and $p \in M$. Then f intersects W transversally at p if either

(i) $f(p) \notin W$ or

(ii) $f(p) \in W$ and $T_{f(p)}N = T_{f(p)}W + (df)_p(T_pM)$.

Example 3.1. Let $f(x) = (x, x^2)$ and consider $W = \{(1, y) : y \in \mathbb{R}\}$ and $p_1 = (1, 1)$, and $p_2 = (-1, 1)$. Since $p_2 \notin W$, f intersects W transversally at p_2. Contrarily, $p_1 \in W$. However, we have

$$T_{(1,2)}\mathbb{R}^2 = T_{(1,2)}W + (df)_1(T_1\mathbb{R}) \Longrightarrow \mathbb{R}^2 = \langle(0,1)\rangle + \langle(1,2)\rangle.$$

If $A \subseteq M$ then f intersects W transversally on A if f intersects W transversally at every $p \in A$. Finally, f intersects W transversally if f intersects transversally on M.

Definition 3.4. A subset G of M is said to be *residual* if it is a countable intersection of open and dense sets.

Theorem 3.4 (Multijet Transversality Theorem)**.** Let M and N be smooth manifolds with W a submanifold of $J^k_s(M,N)$. Then, the set

$$T_W = \{f \in \mathcal{C}^\infty(M,N) : \ j^k_s f \text{ intersects transversally } W\}$$

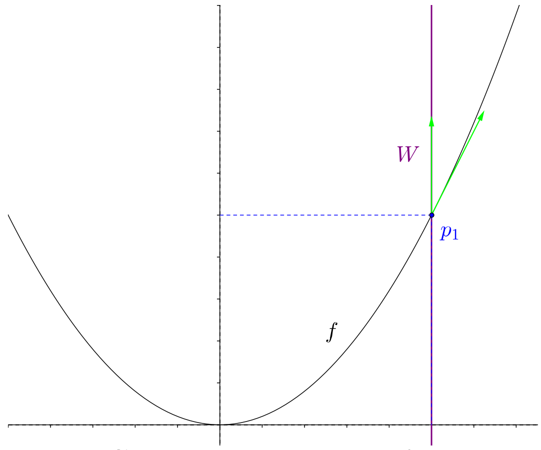


Figure 3.1: Graphic description of transversally.

is a residual subset of $\mathcal{C}^\infty(M, N)$. Moreover, if W is compact, then T_W is also open; hence, generic.

The proof is not part of our work, since it requires a lot of results that they go beyond the manuscript. However, Theorem 3.4 allows us to show a property about periodic points on a compact set.

Corollary 3.5. Let $f : M \to M$, $f \in \text{Dif}(M)$.

- The set $\{f \in \text{Dif}(M)$: fixed points of f which are nondegenerate$\}$ is generic in $\text{Dif}(M)$.
- Nondegenerate fixed points are isolated.

Lemma 3.5. Let M be a compact manifold and suppose $q \geq 1$. The set

$$\{f \in \mathcal{C}^k(M, M) : \ f \text{ is nondegenerate and has only finitely many periodic points of periods } p \leq q\}$$

is generic.

The next result is independent from the previous results, but it is very related with them and it allows to generalize the restrictive Takens' Embedding Theorem.

Proposition 3.9. Let $M_{n\times n}(\mathbb{R})$ be the set of matrices $n \times n$ with real coefficients. Thus, the set of matrices with distinct eigenvalues are open and dense.

Proof. Let A be a matrix. We can write $A = SJS^{-1}$, with J its Jordan's canonical form. We say $J = S^{-1}AS$. We can add a small perturbation $\lambda\text{diag}(1, \ldots, n)$, where

$\text{diag}(a_1 \ldots, a_n)$ is the diagonal matrix, and $\lambda \in \mathbb{R}$. In this case, $J + \lambda\text{diag}(1, \ldots, n)$ has only finite values of λ such that this matrix has at least two eigenvalues with the same value. If a_i are the diagonal values of J, then the new diagonal elements are $a_i + i\lambda$. Thus, if we define

$$\lambda_{i,j} = -\frac{a_j - a_i}{j - i}, \ i < j,$$

and substitute λ by $\lambda_{i,j}$, for some i, j such that $i < j$, then the resulting matrix has at least two eigenvalues with the same value and at most $n \cdot (n-1)/2$ distinct values $\lambda_{i,j}$. Therefore, taking any other λ the eigenvalues will be different. Hence,

$$J + \lambda\text{diag}(1, \ldots, n) = S^{-1}AS + \lambda\text{diag}(1, \ldots, n)$$
$$\Downarrow$$
$$S(J + \lambda\text{diag}(1, \ldots, n))S^{-1} = A + \lambda S\text{diag}(1, \ldots, n)S^{-1}.$$

Let $k = |S\text{diag}(1, \ldots, n)S^{-1}|$. If $|z| < \frac{\epsilon_2}{k}$, then $A + \lambda S\text{diag}(1, \ldots, n)S^{-1}$ is a matrix that belongs to a neighborhood of A of radius ϵ_2. If we take $\epsilon = \min\{\epsilon_1, \epsilon_2\}$, this matrix also has distinct eigenvalues. Thus, the set of matrices with distinct eigenvalues is dense.

Moreover, if A has distinct eigenvalues, the characteristic polynomial of A has distinct roots and hence there is a neighborhood of this polynomial such that all the polynomials (polynomials with degree equal to n) on this neighborhood has distinct roots. Therefore, this set is also an open set. □

Therefore, matrices with distinct eigenvalues are generic. Since the intersection of generic sets is also a generic set, we conclude that for generic f, we have fixed points and its fixed points have distinct eigenvalues. Since the periodic points are fixed points of f^k, the periodic points have distinct eigenvalues too.

3.8 Takens' Embedding Theorem

In the previous section we have proved a restrictive version of Takens' Embedding Theorem. Now, we prove the original one that appears in [2].

In this case, with the conclusions of Section 3.7, we may prove the denseness part quickly. However, the openess part needs again some extra work. We state the theorem:

Theorem 3.5 (Takens' Embedding Theorem)**.** Let M be a compact manifold of dimension m. For pairs (ϕ, y), with $\phi \in \text{Dif}^2(M)$, $y \in \mathcal{C}^2(M, \mathbb{R})$, it is a generic property that the map $\Phi(\phi, y; k)$ is an embedding, for $k \geq 2m + 1$.

We note that it is a little different from Theorem 3.1. This version is a bit more general: however, we always want to embed in the smaller space, hence we work generally with $k = 2m+1$ and thus with $\Phi_{(\phi,y)}$.

Theorem 3.6. Let M be a compact manifold of dimension m. The set of pairs (ϕ, y), with $\phi \in \mathrm{Dif}^2(M)$, $y \in \mathcal{C}^2(M,\mathbb{R})$, such that the map $\Phi(\phi, y; k)$ is an embedding, is dense, for $k \geq 2m+1$.

Proof. Let

$$A = \{\phi \in \mathrm{Dif}^2(M) : \sharp\{x \in P_{2m} : x \text{ has distinct eigenvalues}\} \text{ is finite}\}.$$

A is open and dense on $\mathrm{Dif}^2(M)$, as we have seen in Section 3.7. Now, let $V \subseteq \mathrm{Dif}^2(M) \times \mathcal{C}^2(M,\mathbb{R})$ be the set of pairs (ϕ, y) such that $\Phi_{(\phi,y;k)}$ is an embedding. For every $\phi \in A$ there is an open dense subset $O_x \in \mathcal{C}^2(M,\mathbb{R})$, such that $\{(x, y) : y \in O_x\} \subset V$. Hence, if we take $(\hat{x}, \hat{y}) \in \mathrm{Dif}^2(M) \times \mathcal{C}^2(M,\mathbb{R})$, there is some x near $\hat{x}$ such that $x \in A$, since A is dense. Furthermore, since O_x is a dense subset, we can find $y \in O_x$ near $\hat{y}$. Therefore, V is dense in $\mathrm{Dif}^2(M) \times \mathcal{C}^2(M,\mathbb{R})$. $\square$

If we want to check if these pairs also form an open set, we need to proceed in some steps. As in the restrictive version, we do not use the number $2m+1$, neither the $\mathcal{C}^2$ condition. Therefore, we fix some positive integer p and we consider the map

$$\begin{array}{rccl} \mathcal{F}^{(2)}: & \mathrm{Dif}^1(M) \times \mathcal{C}^1(M,\mathbb{R}) & \to & \mathcal{C}^1(M,\mathbb{R}) \\ & (\phi, y) & \mapsto & \Phi_{(\phi,y;p)}. \end{array} \tag{3.10}$$

Lemma 3.6. The function

$$\begin{array}{rccl} F_1: & \mathrm{Dif}^1(M) \times \mathcal{C}^1(M,\mathbb{R}) & \to & \mathcal{C}^1(M,\mathbb{R}) \\ & (\phi, y) & \mapsto & y \circ \phi \end{array}$$

is continuous.

Proof. Let $\{(U_i, h_i)\}_{i \in I}$ be a finite regular covering for M, with $U_i = h_i^{-1}B(0,3)$. Then $\{W_i = h_i^{-1}B(0,1)\}_{i \in I}$ still covers M. Since ϕ is a diffeomorphism, ϕ^{-1} is a diffeomorphism too. We note that

$$\bigcup_{i \in I} \phi^{-1} W_i = \phi^{-1} \bigcup_{i \in I} W_i = \phi^{-1} M = M.$$

Hence, the collection $\{\phi^{-1}W_i\}_{i \in I}$ is a cover of M. Let $\{(V_j, g_j)\}_{j \in J}$ be a locally finite refinement of $\{\phi^{-1}W_i\}_{i \in I}$. We may suppose that J is finite, since M is compact. Let

$\{X_j = g_j^{-1}B(0,1)\}_{j\in J}$. This set still covers M. Because the cover of V_j's is subordinate to the cover of $\phi^{-1}W_i$'s, for each V_j, there is some W_i, which we call $W_{i(j)}$, such that $\phi V_j \subset W_{i(j)}$.

Let $\epsilon > 0$. For any pair (ϕ, y), the functions $yh_i^{-1} : h_i\overline{W}_i \to \mathbb{R}$ are uniformly continuous, since it is a continuous function defined over a compact set. Therefore, there is a $\delta_i > 0$ such that if $\|u' - u\| < \delta_i$, then $|yh_i^{-1}(u') - yh_i^{-1}(u)| < \epsilon$. Furthermore, since we have taken a finite atlas, we have a finite number of these functions. Hence, we can take $\delta_1 = \min_{i\in I}\{\delta_i\} > 0$ and δ_1 works for all $i \in I$. In addition, we note that

$$\begin{aligned} yh_i^{-1} : h_i\overline{W}_i \subset \mathbb{R}^m \to \mathbb{R} &\quad\Rightarrow\quad Dyh_i^{-1} : h_i\overline{W}_i \to \mathbb{R}^{m\times 1} = \mathbb{R}^m \\ h_i\phi g_j^{-1} : g_j\overline{X}_j \subset \mathbb{R}^m \to \mathbb{R}^m &\quad\Rightarrow\quad Dh_i\phi g_j^{-1} : g_j\overline{X}_j \to \mathbb{R}^{m\times m}. \end{aligned}$$

The derivatives Dyh_i^{-1} and $Dh_i\phi g_j^{-1}$ are continuous, since all the functions are $\mathcal{C}^1$. Furthermore, they have compact domains. Hence, they are bounded. For every $i \in I$, $j \in J$, we can find A_i and $B_{i,j}$ such that $\|Dyh_i^{-1}(u)\| < A_i$ for all $u \in h_i\overline{W}_i$ and $\|Dh_i\phi g_j^{-1}(u)\| < B_{i,j}$ for all $u \in g_j\overline{X}_j$. Since we have a finite number of i's and j's, we may take $A = \min_{i\in I}\{A_i\}$ and $B = \min_{i\in I, j\in J}\{B_{i,j}\}$. Moreover, since we have a continuous function over a compact domain, the function Dyh_i^{-1} is uniformly continuous: given $\epsilon > 0$ there exists $\hat{\delta}_i > 0$ such that $\|Dyh_i^{-1}(u) - D_yh_i^{-1}(u)\| < \epsilon$, for all $\|u' - u\| < \hat{\delta}_i$, for every i, and since we have a finite number of this functions, we may take again $\delta_2 = \min_{i\in I}\{\hat{\delta}_i\}$.

Now, given any neighbourhood of $y \circ \phi$ in $\mathcal{C}^1(M, \mathbb{R})$, there is a neighbourhood of the form $\mathcal{N} = \cap_{j\in J}\mathcal{N}^1(y \circ \phi; (V_j, g_j), (\mathbb{R}, id)\overline{X}_j, \epsilon')$ contained within it. We choose δ sufficiently small that the following inequalities are satisfied:

- $|yh_i^{-1}(u')| < \epsilon'/2$, for all $\|u' - u\| < \delta$, $u, u' \in h_i\overline{W}_i$, and $i \in I$.
- $\|Dyh_i^{-1}(u') - Dyh_i^{-1}(u)\| < \epsilon'/(3B)$, for all $\|u' - u\| < \delta$, $u, u' \in h_i\overline{W}_i$, and $i \in I$.
- $\delta < B$ and $\delta < \epsilon'/(3A)$.

For example, $\delta = \min\{\delta_1, \delta_2, \epsilon'/4, \epsilon'/(6A), B/2\}$. Also, we choose $\epsilon < \min\{\epsilon'/2, \epsilon'/6B\}$. Now consider the open neighbourhood

$$\mathcal{N}^1(\delta, \epsilon) = \bigcap_{j\in J}\mathcal{N}^1(\phi; (V_j, g_j), (W_{i(j)}, h_{i(j)}), \overline{X}_j, \delta) \times \bigcap_{i\in I}\mathcal{N}^1(y; (U_i, h_i), (\mathbb{R}, \mathrm{id}), \overline{W}_i, \epsilon).$$

To show that F_1 is continuous, we show that $F_1(\mathcal{N}^1(\delta, \epsilon)) \subset \mathcal{N}$: that is, if $(\hat{\phi}, \hat{y}) \in \mathcal{N}^1(\delta, \epsilon)$, then $F_1(\hat{\phi}, \hat{y}) \in \mathcal{N}$.

Let $(\hat{\phi}, \hat{y}) \in \mathcal{N}^1(\delta, \epsilon)$, $j \in J$, $x \in \overline{X}_j$ and $u = g_j x \in g_j \overline{X}_j$. Then

$$\begin{aligned} |\hat{y}\hat{\phi}g_j^{-1}(u) - y\phi g_j^{-1}(u)| &= |\hat{y}\hat{\phi}g_j^{-1}(u) - y\hat{\phi}g_j^{-1}(u) + y\hat{\phi}g_j^{-1}(u) - y\phi g_j^{-1}(u)| \\ &\leq |\hat{y}\hat{\phi}g_j^{-1}(u) - y\hat{\phi}g_j^{-1}(u)| + |y\hat{\phi}g_j^{-1}(u) - y\phi g_j^{-1}(u)| \end{aligned} \tag{3.11}$$

We recall that since $u \in g_j\overline{X}_j$, then $u \in g_j\hat{\phi}^{-1}\overline{W}_i$ and $u' = h_{i(j)}\hat{\phi}g_j^{-1}(u) \in h_{i(j)}\overline{W}_i$. Then,

$$|\hat{y}\hat{\phi}g_j^{-1}(u) - y\hat{\phi}g_j^{-1}(u)| = |\hat{y}h_{i(j)}^{-1}(u') - yh_{i(j)}^{-1}(u')| < \epsilon < \epsilon'/2, \tag{3.12}$$

because $\hat{y} \in \cap_i \mathcal{N}^1(y; (U_i, h_i), (\mathbb{R}, \mathrm{id}), \overline{W}_i, \epsilon)$. Also, if $u'' = h_{i(j)}\phi g_j^{-1}(u) \in h_{i(j)}\overline{W}_i$ then

$$|y\hat{\phi}g_j^{-1}(u) - y\phi g_j^{-1}(u)| = |yh_{i(j)}^{-1}(u') - yh_{i(j)}^{-1}(u'')| < \epsilon'/2, \tag{3.13}$$

since

$$\|u'' - u'\| = \|h_{i(j)}\phi g_j^{-1}(u) - h_{i(j)}\hat{\phi}g_j^{-1}(u)\| < \delta$$

by $\mathcal{N}^1(\delta, \epsilon)$. Hence, combining (3.12) and (3.13) in (3.11), we have

$$|\hat{y}\hat{\phi}g_j^{-1}(u) - y\phi g_j^{-1}(u)| < \epsilon'/2 + \epsilon'/2 = \epsilon'.$$

Now, we deal with the derivatives.

$$\begin{aligned} \|D\hat{y}\hat{\phi}g_j^{-1}(u) - Dy\phi g_j^{-1}(u)\| &= \|D\hat{y}\hat{\phi}g_j^{-1}(u) - Dy\hat{\phi}g_j^{-1}(u) \\ &\quad + Dy\hat{\phi}g_j^{-1}(u) - Dy\phi g_j^{-1}(u)\| \\ &\leq \|D\hat{y}\hat{\phi}g_j^{-1}(u) - Dy\hat{\phi}g_j^{-1}(u)\| \\ &\quad + \|Dy\hat{\phi}g_j^{-1}(u) - Dy\phi g_j^{-1}(u)\| \\ &= \|D\hat{y}h_i^{-1}h_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}h_i\hat{\phi}g_j^{-1}(u)\| \\ &\quad + \|Dyh_i^{-1}h_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}h_i\phi g_j^{-1}(u)\|, \end{aligned}$$

where $h_i = h_{i(j)}$. Let $u' = h_i\hat{\phi}g_j^{-1}(u)$ and $u'' = h_i\phi g_j^{-1}(u)$, as above. Using the Chain Rule and the triangle inequality,

$$\begin{aligned} \|D\hat{y}\hat{\phi}g_j^{-1}(u) - Dy\phi g_j^{-1}(u)\| &\leq \|D\hat{y}h_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u)\| \\ &\quad + \|Dyh_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}(u'')Dh_i\phi g_j^{-1}(u)\| \\ &\leq \|D\hat{y}h_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u)\| \\ &\quad + \|Dyh_i^{-1}(u')Dh_i\hat{\phi}g_j^{-1}(u) - Dyh_i^{-1}(u')Dh_i\phi g_j^{-1}(u)\| \\ &\quad + \|Dyh_i^{-1}(u')Dh_i\phi g_j^{-1}(u) - Dyh_i^{-1}(u'')Dh_i\phi g_j^{-1}(u)\| \\ &\leq \|D\hat{y}h_i^{-1}(u') - Dyh_i^{-1}(u')\| \cdot \|Dh_i\hat{\phi}g_j^{-1}(u)\| \\ &\quad + \|Dh_i\hat{\phi}g_j^{-1}(u) - Dh_i\phi g_j^{-1}(u)\| \cdot \|Dyh_i^{-1}(u')\| \\ &\quad + \|Dyh_i^{-1}(u') - Dyh_i^{-1}(u')\| \cdot \|Dh_i\phi g_j^{-1}(u)\|. \end{aligned}$$

We recall that:

(i) $\|D\hat{y}h_i^{-1}(u') - Dyh_i^{-1}(u')\| < \epsilon$, since $y \in \cap_i \mathcal{N}^1(y; (U_i, h_i), (\mathbb{R}, \mathrm{id}), \overline{W}_i, \epsilon)$.

(ii) $\|Dh_i\phi g_j^{-1}(u)\| < B$ and $\|Dyh^{-1}(u)\| < A$.

(iii) $\|Dh_i\hat{\phi}g_j^{-1}(u) - Dh_i\phi g_j^{-1}(u)\| < \delta$, since $\hat{\phi} \in \cap_j \mathcal{N}^1(\phi; (V_j, g_j), (W_i, h_i), \overline{X}_j, \delta)$ and

$$\begin{aligned} \|Dh_i\hat{\phi}g_j^{-1}(u)\| &< \|Dh_i\hat{\phi}g_j^{-1}(u) - DH_i\phi g_j^{-1}(u)\| + \|Dh_i\phi g_j^{-1}(u)\| \\ &< \delta + B < 2B. \end{aligned}$$

(iv) $\|Dyh_i^{-1}(u') - Dyh_i^{-1}(u'')\| < \frac{\epsilon'}{3B}$.

All in all, we have

$$\|D\hat{y}\hat{\phi}g_j^{-1}(u) - Dy\phi g_j^{-1}(u)\| < \epsilon \cdot 2B + \delta \cdot A + \frac{\epsilon'}{3B}B < \frac{\epsilon'}{6B}2B + \frac{\epsilon'}{3A}A + \frac{\epsilon'}{3} = \epsilon'.$$

Therefore, if $(\hat{\phi}, \hat{y}) \in \mathcal{N}^1(\delta, \epsilon)$, then $F_1(\hat{\phi}, \hat{y}) = \hat{y} \circ \hat{\phi} \in \mathcal{N}$. Thus, F_1 is continuous. □

Lemma 3.7. The function

$$\begin{aligned} F_n : \quad \mathrm{Dif}^1(M) \times \mathcal{C}^1(M, \mathbb{R}) &\to \mathcal{C}^1(M, \mathbb{R}) \\ (\phi, y) &\mapsto y \circ \phi^n \end{aligned}$$

is continuous, for $n \in \mathbb{N}$.

Proof. The two first cases ($F_0(\phi, y) = y$ and F_1 as above) are continuous. We suppose that F_n is continuous and we want to show that F_{n+1} is also continuous. Let $G(\phi, y) = (\phi, F_1(\phi, y))$. By the previous lemma, G is a continuous map.

$$F_{n+1}(\phi, y) = y \circ \phi^{n+1} = y \circ \phi \circ \phi^n = F_1(\phi, y) \circ \phi^n = F_n(\phi, F_1(\phi, y)) = F_n(G(\phi, y)).$$

Thus, $F_{n+1} = F_n \circ G$ is a composition of continuous functions; hence F_{n+1} is continuous and by induction F_n is continuous for all $n \in \mathbb{N}$. □

Corollary 3.6. The map $\mathcal{F}^{(2)}$ in (3.10) is continuous.

Proof. Since all the components of $\mathcal{F}^{(2)}$ are continuous, the map is continuous. □

The next proposition follows in the same way as Proposition 3.1 and 3.2. We take the same S as in (3.2). This implies the proof of the openness part and hence the proof of the Takens' Embedding Theorem.

Proposition 3.10. Let M be a compact manifold, and K be a compact subset of M. Then the set

$$\mathcal{Y}^1 = \{(\phi, y) \in \mathrm{Dif}^1(M) \times \mathcal{C}^1(M, \mathbb{R}) : \Phi_{(\phi, y; k)} \text{ embedding in } K\}$$

is open in $\mathrm{Dif}^1(M) \times \mathcal{C}^1(M, \mathbb{R})$.

3.9 Relaxing to $\mathcal{C}^1$ Condition

We note that the theorem needs that $\phi \in \text{Dif}^2(M)$ and $y \in \mathcal{C}^2(M, \mathbb{R})$. However, in the openess part (Proposition 3.10) we only use $\mathcal{C}^1$ maps. Moreover, it is also possible to relax the condition to $\mathcal{C}^1$ in the dense part. To do this, an easy way is understanding the $\mathcal{C}^2$ sets as dense sets of $\mathcal{C}^1$.

Theorem 3.7. Let M be a $\mathcal{C}^s$ manifold, $1 \leq s < \infty$. Then $\mathcal{C}^s(M, \mathbb{R}^n)$ is dense in $\mathcal{C}^0_s(M, \mathbb{R}^n)$.

Proof. Let $\{V_i\}_{i \in I}$ be a locally finite cover of M, and for every $i \in I$, let $\epsilon_i > 0$. Let $f : M \to \mathbb{R}^n$ be a continuous map. We seek some $g \in \mathcal{C}^\infty$ such that $|f - g| < \epsilon_i$ in V_i, for every $i \in I$. Let $x \in M$. Since we have a locally finite cover, we can take $W_x \subset M$ a neighborhood of x that intersects finitely many V_i's. And thus the minimum $\delta_x = \min\{\epsilon_i : x \in V_i\} > 0$ is achieved. Let $U_x = f^{-1}(B(f(x), \delta_x)) \cap W_x \subset W_x$. We define the constant maps

$$\begin{array}{rccl} g_x : & M & \to & \mathbb{R}^n \\ & y & \mapsto & g_x(y) = f(x). \end{array}$$

As M is second countable, it is possible to take a countable base of $\{U_x\}$ and we relabel these open covers as $\{U_j\}_{j \in J} = \mathcal{U}$ and the maps $\{g_j\}_{j \in J}$ associated to them. Consequently, for every $y \in U_j \cap V_i$, we have $|g_j(y) - f(y)| < \epsilon_i$. Let $\{\lambda_j\}_{j \in J}$ be a partition of unity subordinated to $\mathcal{U}$. We define

$$\begin{array}{rccl} g : & M & \to & \mathbb{R}^n \\ & y & \mapsto & g(y) = \sum_{j \in J} \lambda_j(y) g_j(y). \end{array}$$

Therefore, if $y \in V_i$ we have

$$\begin{aligned} |g(y) - f(y)| &= \Big| \sum_j \lambda_j(y) g_j(y) - \sum_j \lambda_j(y) f(y) \Big| = \Big| \sum_j \lambda_j(y)(g_j(y) - f(y)) \Big| \\ &\leq \sum_j |\lambda_j(y)||g_j(y) - f(y)| = \sum_j \lambda_j(y)|g_j(y) - f(y)| \\ &< \sum_j \lambda_j(y)\epsilon_i = \epsilon_i \sum_j \lambda_j(y) = \epsilon_i. \end{aligned}$$

□

We note that since $\mathcal{C}^2(M, \mathbb{R}) \subset \mathcal{C}^1(M, \mathbb{R}) \subset \mathcal{C}^0_S(M, \mathbb{R})$ and the first set is dense in the last one, in particular the first set is also dense in the second set. There is a similar result for $\mathrm{Dif}^2(M)$ and $\mathrm{Dif}^1(M)$. This appears as Theorem 2.7 in [13]. We do not prove it since it is a very long result and it falls outside our objectives.

Theorem 3.8. Let M be a $\mathcal{C}^s$ manifold, $1 \leq s < \infty$. Then $\mathrm{Dif}^s(M, \mathbb{R}^n)$ is dense in $\mathrm{Dif}^0(M, \mathbb{R}^n)$ in the strong topology.

Therefore, we have the chain of inclusions

$$\begin{aligned}\{(\phi, y) :\, \Phi_{(\phi,y)} \text{ is an embedding on } M\} &\subset \mathrm{Dif}^2(M) \times \mathcal{C}^2(M, \mathbb{R}) \\ &\subset \mathrm{Dif}^1(M) \times \mathcal{C}^1(M, \mathbb{R}).\end{aligned}$$

Each inclusion is dense and thus the first set is dense in the last set. Thus, we have an equivalent result of Theorem 3.5, but with $\mathcal{C}^1$ maps:

Theorem 3.9. Let M be a compact manifold of dimension m. For pairs (ϕ, y), with $\phi \in \mathrm{Dif}^1(M)$, $y \in \mathcal{C}^1(M, \mathbb{R})$, it is a generic property that the map $\Phi(\phi, y; k)$ is an embedding, for $k \geq 2m + 1$.

Chapter 4

Applications

In this chapter, we show some applications of the Takens' Embedding Theorem. We use the theorem for dynamical systems, as it was the field for which it was born. First of all, we apply it to the continuous dynamical systems. Secondly, we apply to others dynamical systems, such as discrete or abstract, as we have described previously. Finally, we use the theorem in real-time series to study their behavior and properties.

4.1 Continuous Dynamical Systems

We recall that given the initial condition problem (2.4),

$$\begin{cases} \dot{x} = F(x), \\ x(0) = x_0, \end{cases} \tag{4.1}$$

with $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz, a solution is a $\mathcal{C}^1$-function

$$\varphi(t; x_0) = (\varphi_{(1)}(t; x_0), \dots, \varphi_{(n)}(t; x_0))$$

such that satisfies (4.1). We suppose that $\varphi(t; x_0)$ is dense on the attractor set. Moreover, we suppose that this attractor set is also a manifold. Hence, it is a $\mathcal{C}^1$-manifold of dimension $m \leq n$ and it is given by the set

$$M = \{\overline{(\varphi_{(1)}(t), \dots, \varphi_{(n)}(t))} \in \mathbb{R}^n : t \in \mathbb{R}\}.$$

We fix a time-delay τ. The dynamical system

$$\begin{array}{rrcl} T : & \mathbb{R} \times M & \to & M \\ & (t, x_0) & \mapsto & T_t(x_0) = \varphi(t; x_0) \end{array}$$

gives rise to a function

$$\begin{array}{rccl} T_\tau : & M & \longrightarrow & M \\ & \varphi(t;x_0) & \mapsto & T_\tau(\varphi(t;x_0)) = \varphi(t+\tau;x_0). \end{array}$$

The map $T_\tau \in \mathrm{Dif}^1(M)$ and its powers are

$$T_\tau^k(\varphi(t;x_0)) = T_\tau \cdots T_\tau(\varphi(t;x_0)) = \varphi(t+k\tau;x_0).$$

Furthermore, we call an *observable* y a $\mathcal{C}^1$ function that goes from the manifold M to a real value. It can be whatever we think about. For example, it can be a projection or a linear combination of the different components, among many other possibilities. All in all, the map $\Phi_{(T_\tau,y)}$ of (3.1) is given by

$$\begin{array}{rccl} \Phi_{(T_\tau,y)} : & M & \to & \mathbb{R}^{2m+1} \\ & \varphi(t;x_0) & \mapsto & (y(\varphi(t;x_0)), y(\varphi(t+\tau;x_0)), \dots, y(\varphi(t+2m\tau;x_0))). \end{array}$$

Hence, the set of points

$$\{\Phi_{(T_\tau,y)}(\varphi(t;x_0)), \Phi_{(T_\tau,y)}(\varphi(t+\tau;x_0)), \dots\},$$

should return a discretization of the manifold.

Remark 4.1. We recall that it might happen that $2m+1 > n$. However, it is not a contradiction: Takens' Embedding Theorem says that the manifold is embedded generically by $\Phi_{(\phi,y;k)}$, for $k \geq 2m+1$: but it is not necessary; usually, we can embed the manifold in some real space of dimension $k < 2m+1$.

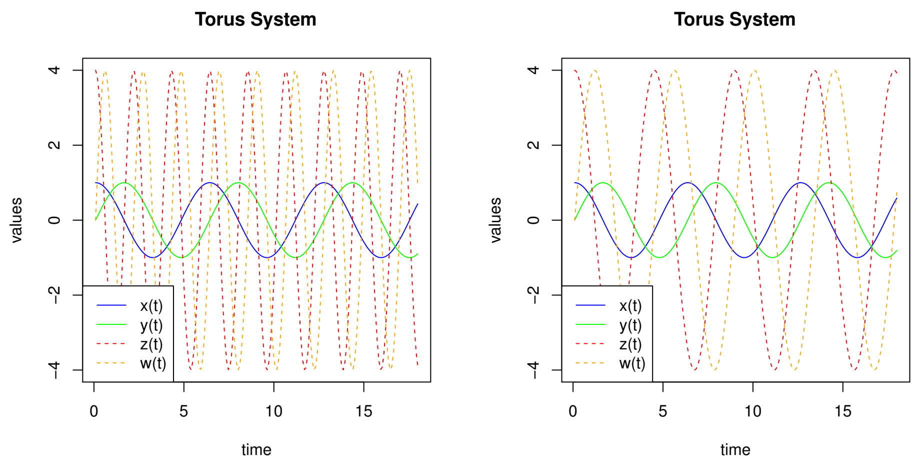
Remark 4.2. Finally, we note that it is a strong condition to suppose that we are in the attractor set. Nonetheless, if we are not on the attractor set, there is some time t_0 such that if M is the attractor set, then we have the difference $\|M - \varphi(t;x_0)\| < \epsilon$, for every $\epsilon > 0$ and $t \geq t_0$. Thus, there is some time where we are as close as we want from the attractor set. Hence, the behavior should be similar as if we start in the attractor set.

Example 4.1. In this case, we consider an ODE with a torus as an attractor set. The torus has been built as a product of two circumferences. Let us consider the system

$$\begin{cases} x' = -y + x(1 - \sqrt{x^2+y^2}), \\ y' = x + y(1 - \sqrt{x^2+y^2}), \\ z' = -kw + z(4 - \sqrt{z^2+w^2}), \\ w' = kz + w(4 - \sqrt{z^2+w^2}). \end{cases}$$

We have two decopled systems. In both cases, the attractor set is a circumference. In the plane (x, y) the circumference has radius 1 and in the plane (z, w) has radius 4. Moreover, we can see the torus as a fundamental polygon. Then, if we see the torus as a rectangle with the boundaries identified correctly, the parameter k is the slope of the solution in the rectangle. For example:

- If we take $k = 3$, the slope is rational. We have that the torus is covered by periodic orbits (Figure 4.1(a)).
- If we take $k = \sqrt{2}$, the slope is irrational. Thus, we have dense orbits on the torus (Figure 4.1(b)).



(a) (b)

Figure 4.1: Solutions of the torus. 4.1(a) With a rational slope. 4.1(b) With an irrational slope.

In both cases, we start at the point $(1, 0, 4, 0)$, with a step $h = 0.1$. In Figure 4.2, we apply Takens' Theorem. We use a Runge-Kutta method to get the solutions. We now choose an observable to obtain the time-series $\{y\varphi(t_0 + kh) : k = 0, \ldots, N\}$. Since we have two decoupled systems, we shall mix both systems: otherwise, we would obtain a circumference as the attractor set. On the one hand, we have the variables x and y in one system. On the other hand, we have the variables z and w. In our case, we sum two variables, $x + z$ and we make the delay from these new signals.

In Example 4.1, we apply Takens' Theorem for $\mathcal{C}^2$-manifolds. We note that in Chapter 1 we build a $\mathcal{C}^1$-manifold and hence we may use the version of Theorem 3.9.

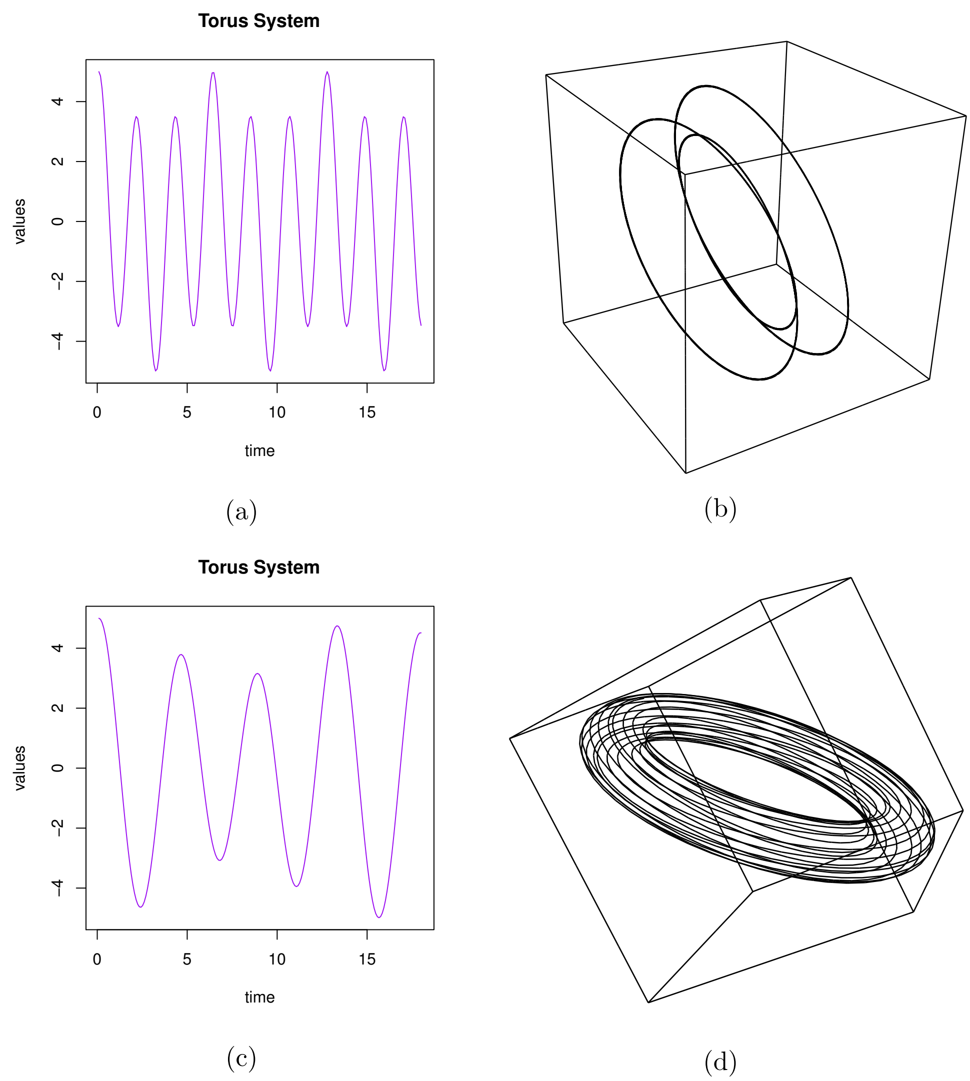


Figure 4.2: Takens' delay applied to signals of two manifolds. We sum the solutions x and z. 4.2(a) gives rise to a circumference and 4.2(b) is its Takens' delay with $\tau = 8$. 4.2(c) is dense in on the torus and 4.2(d) is its Takens' delay, with a slope $\tau = 12$.

4.2 Discrete Dynamical Systems

We recall that a discrete dynamical system is given by

$$\begin{array}{rccl} T: & G \times M & \to & M \\ & (n, x) & \mapsto & T_n(x) = f^n(x), \end{array}$$

where $G = \mathbb{Z}$ or $G = \mathbb{N}$. In this case, we have a recurrent relation

$$x_{l+1} = f(x_l) \in \mathbb{R}^n.$$

We get M by the graph $M = \overline{\{x_l : l \in G\}}$. Hence, we again consider an observable y of the set of points $\{x_l\}_{l \in G}$ such as in Section 4.1. The embedding map of Takens' Theorem is given by

$$\begin{array}{rccl} \Phi_{(T_\tau, y)} : & M & \to & \mathbb{R}^{2m+1} \\ & x_l & \mapsto & (y(x_l), y(x_{l+\tau}), \ldots, y(x_{l+2m\tau})). \end{array}$$

Hence, the set of points

$$\{\Phi_{(T_\tau, y)}(x_l), \Phi_{(T_\tau, y)}(x_{l+1}), \ldots\},$$

is again a discretization of the manifold. As well as in continuous dynamical systems, it is possible that $2m + 1 > n$, or even we are not on a manifold. On the first case, we recall that it is not necessary that the dimension of the space is $2m + 1$. Moreover, if a manifold is an attracting set of the system, for some l_0 we will be sufficiently close to the manifold, for all x_l, $l \geq l_0$.

Example 4.2. Consider the recurrence

$$x_{n+1} = \frac{2x_n}{\|x_n\| + 1}.$$

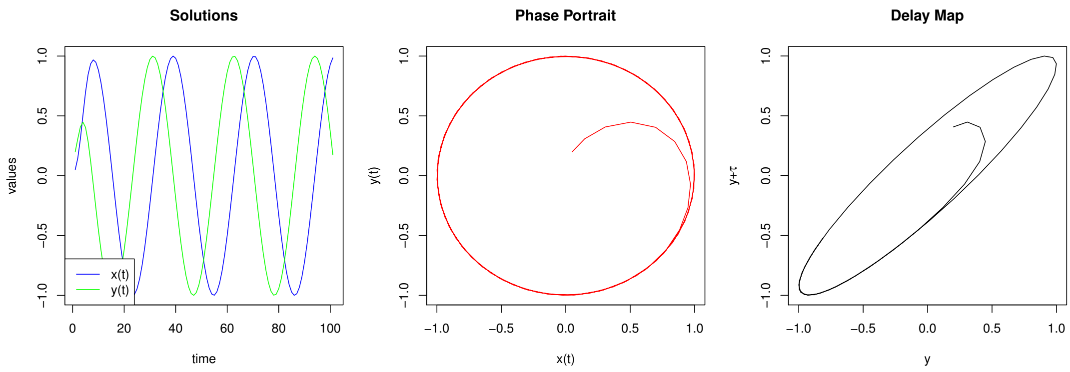
It has two limit points; 0 and $x_0/\|x_0\|$. Moreover, if $0 < \|x_0\| < 1$, we obtain an increasing recurrence in norm: and if $1 < \|x_0\|$, we obtain a decreasing recurrence in norm. Both recurrences tend to $x_0/\|x_0\|$. Therefore, if we consider the 2-dimensional discrete dynamical system

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \frac{2}{\sqrt{x_n^2 + y_n^2} + 1} \begin{pmatrix} \cos(k) & \sin(k) \\ -\sin(k) & \cos(k) \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \tag{4.2}$$

we have that the S^1 is a manifold that is an attracting set, for every initial condition $(x_0, y_0) \neq (0, 0)$. In Figure 4.3, we have the discrete-time solutions and the embedding, for the initial condition $(x_0, y_0) = (0.05, 0.2)$.

4.3 General Dynamical Systems

From the previous sections, we note that the argument is always the same: we suppose that for a given dynamical system there is some manifold that is an attractor set. At some point, we will be sufficiently close to this manifold and if it is sufficiently smooth



(a) (b) (c)

Figure 4.3: Discrete Dynamical System with the circumference as an attractor set. To see it properly, we joint the points with lines. We can observe a quick convergence. 4.3(a) Solutions. 4.3(b) Phase portrait. 4.3(c) Takens' delay, $\tau = 1$.

(that is a $\mathcal{C}^1$-manifold), by Takens' Embedding Theorem we can assure that $\Phi_{(\varphi,y)}$ embeds generically.

Furthermore, the definition of a dynamical system is very abstract. We may define some dynamical system over a system with a $\mathcal{C}^1$-manifold as an attractor set and we can try to obtain some observable of the system. If it is the case, then Takens' embedding Theorem works generically.

We recall the Example 2.12. If the set of coefficients of some time-series comes from an observable of the unity circumference, then an observable of the polynomial shall return the dynamics of the circumference.

Example 4.3. We define the following power series

$$p(x) = \sum_{i=0}^{\infty} \cos(i) x^i.$$

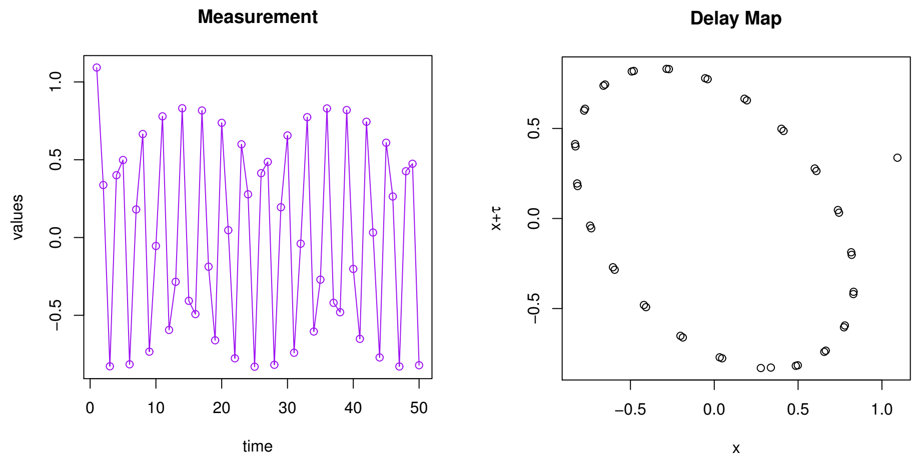
We note that the set $\{\cos(i)\}_{i \in \mathbb{N}}$ is a dense set in $[-1, 1]$. Furthermore, since $|\cos(i)| \leq 1$, the power series uniformly converges into the open $(-1, 1)$. Let T be the dynamical system as in Example 2.12. An observable of the series is given by $p(c)$, for $c \in (-1, 1)$. For example, we choose $c = 0.3$. If we only take the first 5 terms of the power series the error is

$$\epsilon = |\sum_{i=5}^{\infty} \cos i \cdot 0.3^i| \leq \sum_{i=5}^{\infty} |\cos i| \cdot 0.3^i \leq \sum_{i=5}^{\infty} 0.3^i = \frac{0.3^5}{0.7} < 10^{-2}.$$

Therefore, we set the dynamical system as $T_\sigma(p) = p^\sigma$, where

$$i \overset{\sigma}{\mapsto} \begin{cases} 0 & \text{if } i = 1, \\ i+2 & \text{if } i = 2k, \ \ k \in \mathbb{N}, \\ i-2 & \text{if } i = 2k+1, \ \ k \in \mathbb{N}. \end{cases}$$

For any polynomial of the dynamical system, we have a small error if we only take the first 5 terms.[1] From this, we build the time-series. In Figure 4.4(a), we can see the measurement time-series.



(a) (b)

Figure 4.4: A manifold reconstruction from a different dynamical system. 4.4(a) $p^\sigma(0.3)$. 4.4(b) Takens' delay, $\tau = 1$.

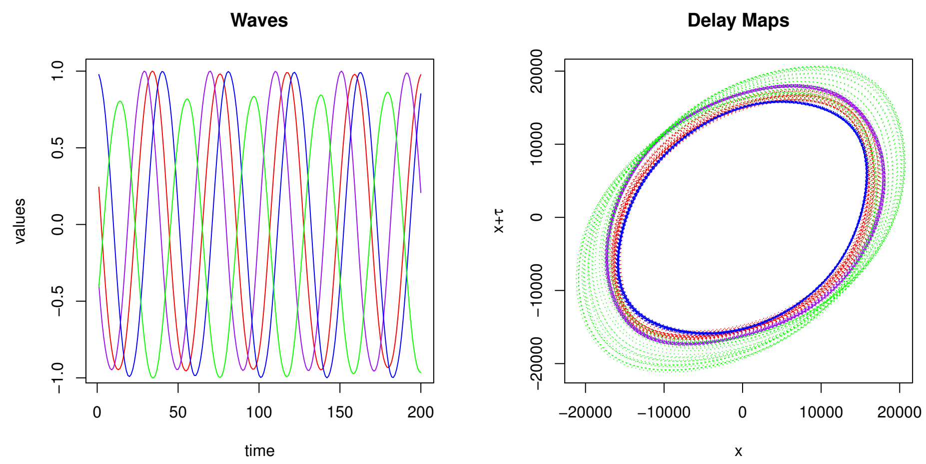
4.4 Signals

We have seen from the Takens' Embedding Theorem that we can obtain the attracting manifold of a time-series. However, we have only used time-series that we know which is the manifold they have as an attractor set. Actually, we usually have some signal from a measurement experiment and we want to find some properties about this signal. If the signal has some attractor set that is a manifold, we may reconstruct the manifold from the signal through Takens' Theorem. In this case, from the geometry of the attractor set we might obtain some properties that characterize the signal.

[1] We only prove it for the first polynomial of the dynamical system, but we can prove it similarly for the others.

For example, if we take some musical instrument and we play some note and we hold it, in suitable conditions the time-series will tend to some period signal and hence it gives rise to a circumference. From this analysis, we may try to find some characterization about the instrument: such as from which instrument is, or the tone color.

First of all, we try with different persons and the same instrument, the same note. In Figure 4.5(a), we can see four waves that represent a Do 4, played with a flute. The reason to use a flute is simple; it is the easiest instrument to play and all the testers are, in first sight, inexperts. In 4.5(a) the signals are normalized, to compare it better. Moreover, the time in the harmonic plots are in scale $1/44100Hz$. In this case, in 4.5(b) we do Takens' delay map in $\mathbb{R}^2$, since it is visually clearer.

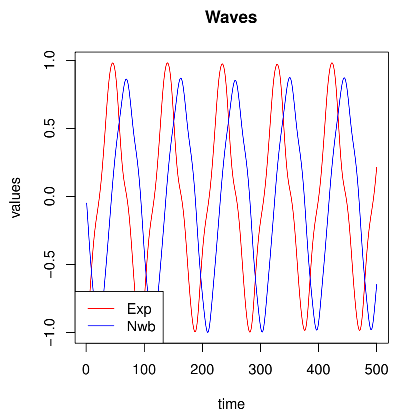


(a) (b)

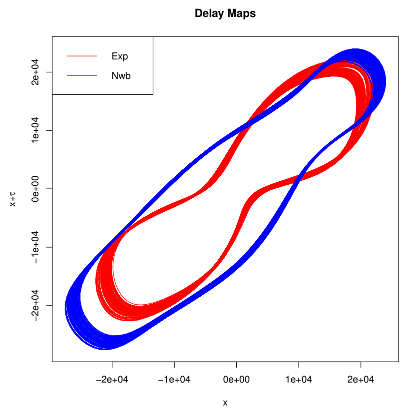
Figure 4.5: Experiments with a flute. 4.5(a) The signals, playing the Do 4 note. 4.5(b) Takens' delay, $\tau = 8$. The delay maps are all similar, and some of them more stable.

In the second experiment (Figure 4.6), two persons play with the same instrument and the same note, but one an habitual player and the other a newbie. We note that the geometry is a little different. It could be that there exists a difference between experts and newbie players.

In the third experiment, as we can see in Figure 4.7, we compare only one person playing various instruments. In this case, the subject plays the same note with the flute, the trumpet, the trumpet with a mute and the saxophone. In Figure 4.7, there exist differences in the geometry among instruments.



(a)



(b)

Figure 4.6: Experiments with a saxophone. An expert and a newbie. 4.6(a) The signals. 4.6(b) Takens' Delay, $\tau = 8$. At first sight, the geometry is different.

Furthermore, we can also try to apply Takens' Theorem to the harmony of the music. If we play two notes at the same time, we obtain a signal which is the union of these two sounds. When the notes are the same, the vibrations join and they amplify the sound. If the notes are different, we have a *musical interval*: that is, we have two signals with different frequency and then we can do their mathematical ratio. For example, some common ratios are

- *Octaves*, as $2:1$.
- *Perfect Fifth*, as $3:2$.
- *Major third*, as $5:4$.
- *Whole Tone*, as $9:8$.
- *Semitone*. This case is special, since the harmony is not perfect. In some cases it is defined as $2^{1/12}$, an irrational ratio, but closed to the rational ratio $16:15$.

Usually, when we play a note with an instrument, it appears some harmonics. Therefore, when we make a harmony with some instruments, the harmonics of each note are also included and this fact makes a more complex harmony. It is translated in some signal with all these properties: it contains the frequency ratio, the harmonics and the harmonics ratio. Hence, we can apply another time Takens' Theorem to the signal. In Figure 4.8(a), there are represented some of these ratios.

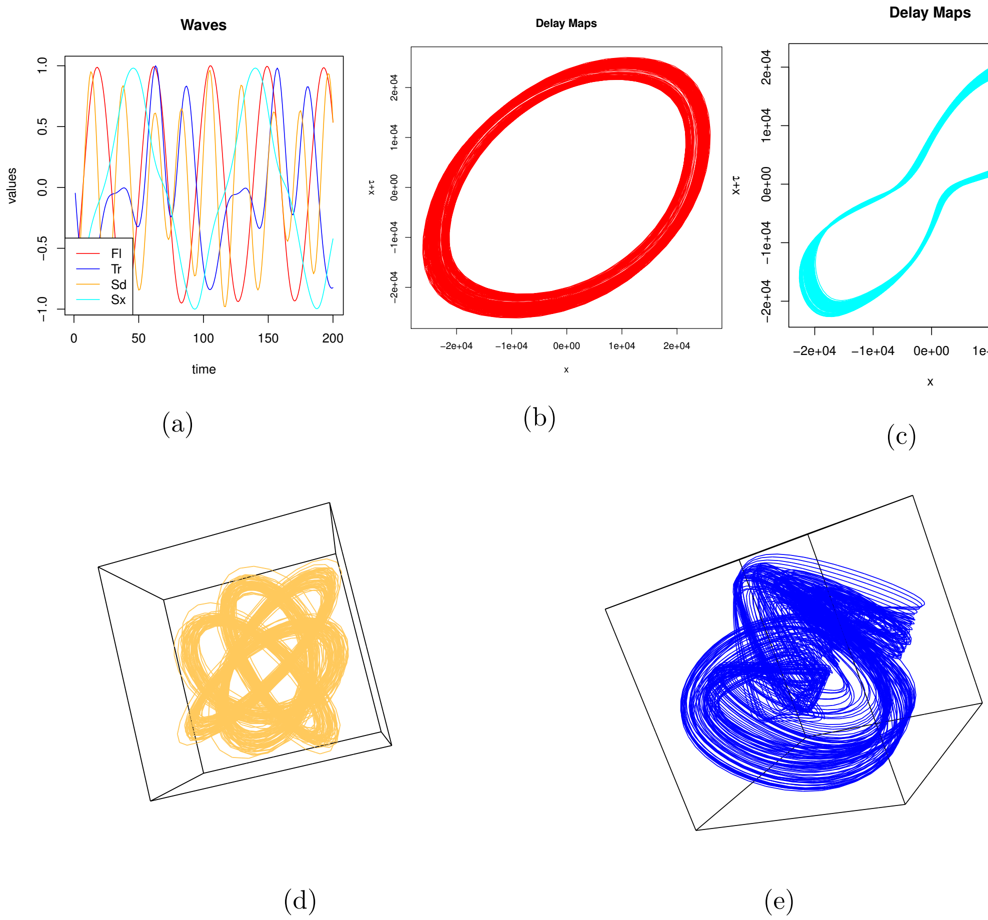


Figure 4.7: Comparison of the geometry of distinct instruments. 4.7(a) The signals. 4.7(b) Flute, with $\tau = 8$. 4.7(c) Saxophone, with $\tau = 8$. 4.7(d) Trumpet with mute, with $\tau = 8$. 4.7(e) Trumpet, with $\tau = 18$.

4.5 Chaos and Time-Series

Finally, in most cases the signals come from a chaotic system, and hence the possible attractor needs not to be a manifold. In this case, Takens' Embedding Theorem fails even though the attractor set is compact. Nonetheless, there are some extensions of Takens' Embedding Theorem that can be applied in the situations. One of them comes from Sauer, Yorke and Casdagli [7]. They state an equivalent version to Takens' Embedding Theorem for fractal sets. However, there is always a catch: we must change the *generic* word to *prevalent*. In finite dimensional spaces, a set is prevalent if its complement has measure zero (see [7]). We do not give a proof of this extension, since it is beyond of our objectives.

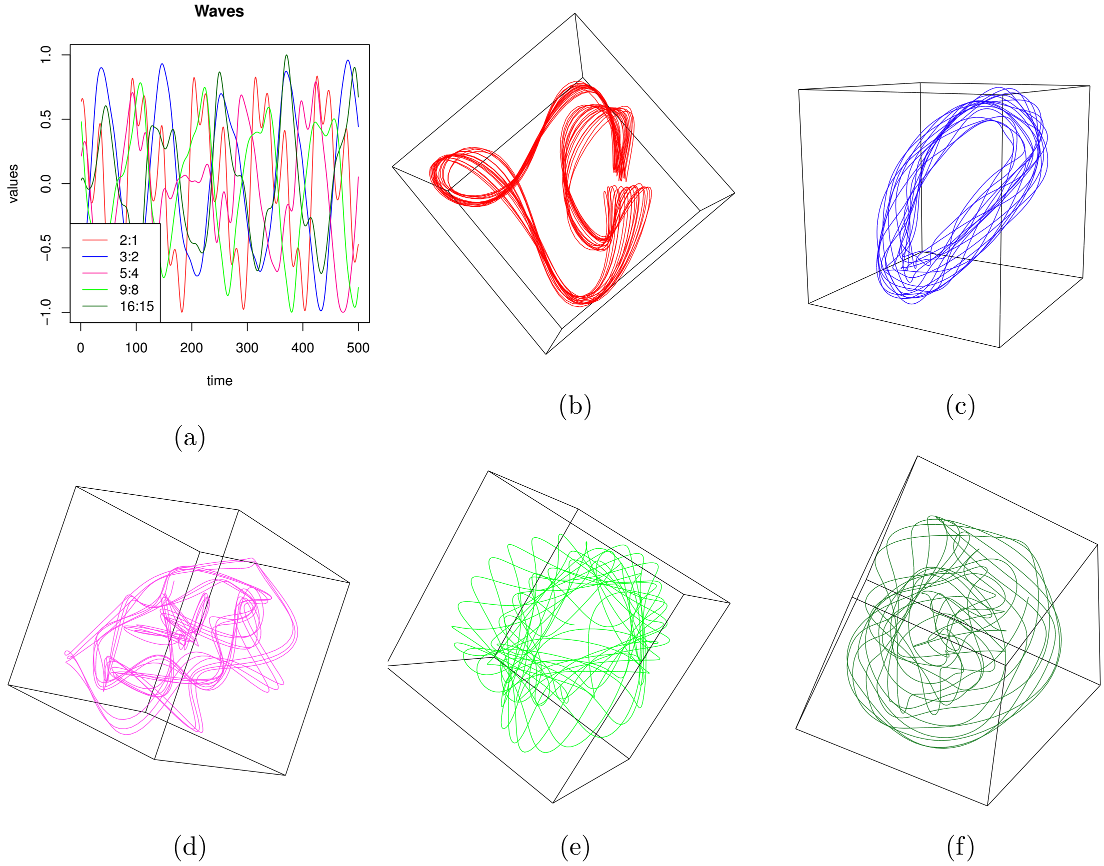


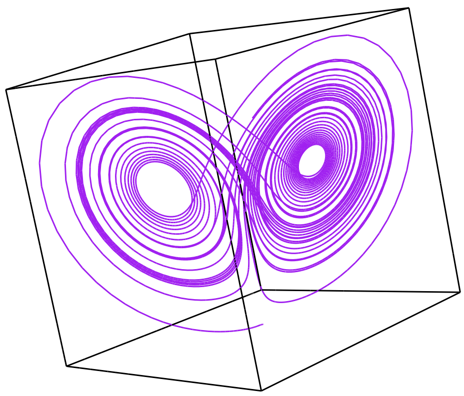
Figure 4.8: Experiments with two musical instruments. 4.8(a) The wave signals. 4.8(b) The octave, $\tau = 11$. 4.8(c) The Fifth, $\tau = 20$. 4.8(d) The third, $\tau = 27$. 4.8(e) Second, $\tau = 27$. 4.8(f) Semitone, $\tau = 27$.

In this case, we can apply the Takens' delay map with dynamical systems with a fractal set as an attractor set and even with any type of signals. For example, we may consider the Lorenz system, with the classical parameters. In Figure 4.9 we can see all the procedure.

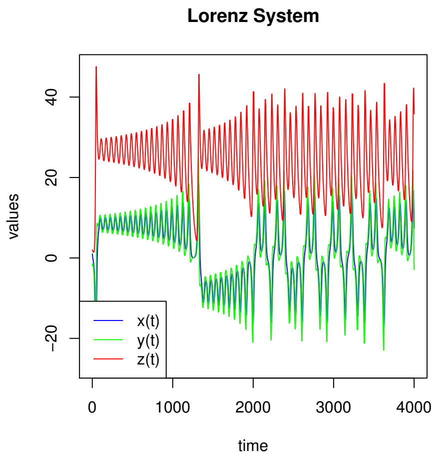
In addition, we can try to apply Takens' delay with time-series from an electroencephalography, an economic signal and a temperature measurement. We make all of them in $\mathbb{R}^3$, as we can see in Figure 4.10. In the worst scenario, the signals come from a stochastic process and hence the embedding fails.[2] [3]

[2]`https://es.finance.yahoo.com/quote/ACX.MC/history?period1=946854000&period2=1535407200&interval=1d&filter=history&frequency=1d&guccounter=1`.

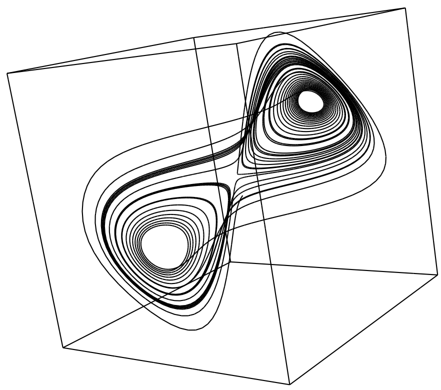
[3]`https://datahub.io/core/global-temp#resource-monthly`.



(a)



(b)



(c)

Figure 4.9: Takens' delay map applied to the Lorenz System. 4.9(a) Phase portrait. 4.9(b) Time-series solutions. 4.9(c) Takens' delay map applied to $x(t)$, with $\tau = 11$. The similitude between 4.9(a) and 4.9(c) is clear.

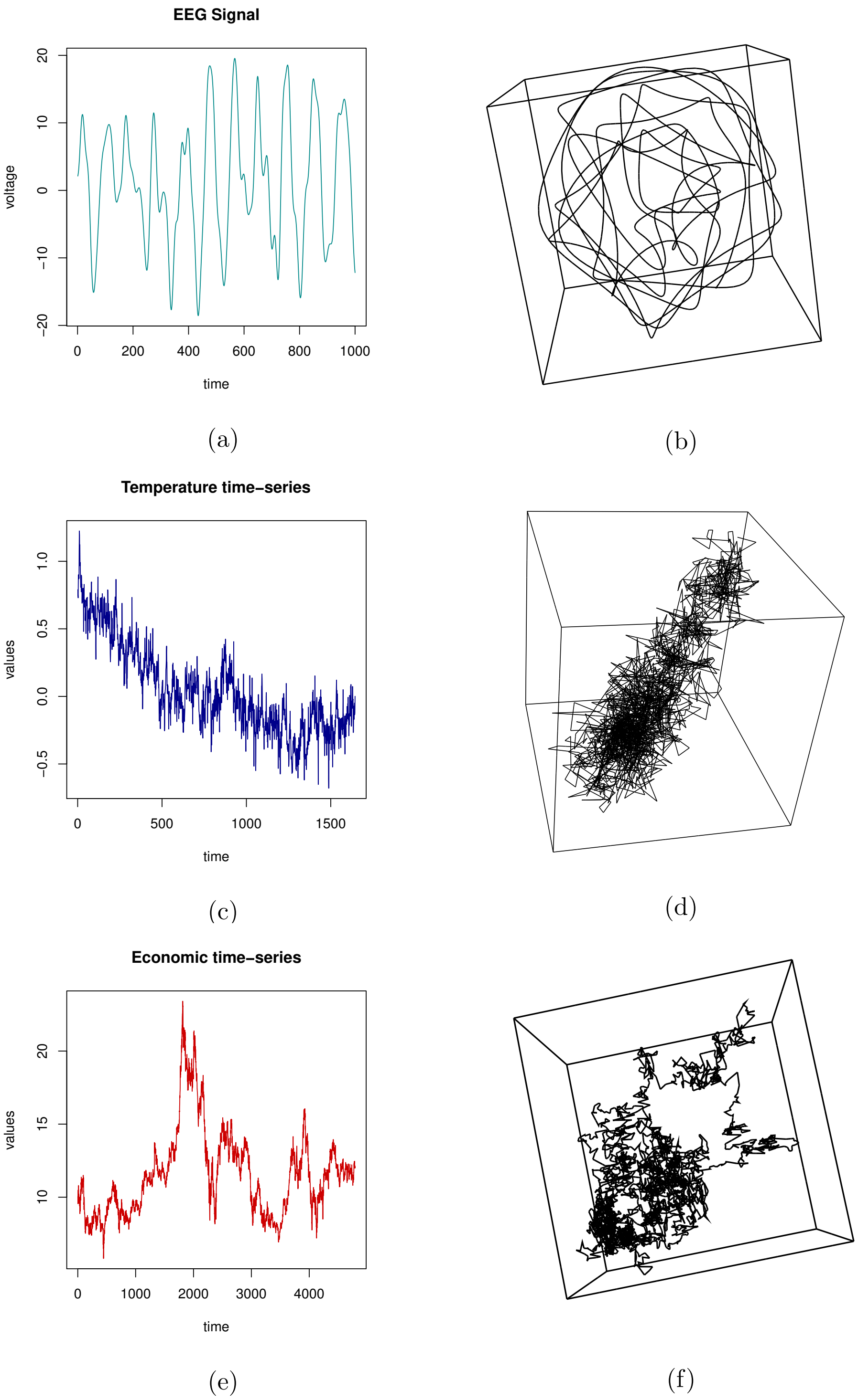


Figure 4.10: Takens' Embedding Theorem, map applied to: a voltage signal 4.10(a) 4.10(b), with delay $\tau = 21$; a temperature measurement 4.10(c) 4.10(d), with $\tau = 18$; and some economic time-series 4.10(e) 4.10(f), with $\tau = 200$.

Chapter 5

Conclusions and Future Work

In this manuscript, we prove the Takens' Embedding Theorem and we give a plotline to understand it. This theorem allows us to recover the attractor set of the dynamical system from a signal. The attractor set contains information about the system, such as if it is a manifold or a fractal set. In this work we focus mainly in the proof: however, there are a lot of applications. Among them, we compare harmonic signals and we apply it to some complex signals, such as temperature measurement or voltage.

The next stage might be some statistic studies. For example, we may make a contrast about the geometry of the attractor set between expert and newbie players of some instrument. In addition, we could look for instrument characterizations. For example, if there exists some characterization of some instruments, we could model new instruments using three-dimensional circumferences or knots.

In case of fractal sets, we may study their properties, such as the fractal dimension or Lyapunov exponents. To perfom the Takens' Delay, I'm developing a Grafical User Interface (GUI) with R. Through this interface, one can calculate the fractal dimension, estimate the delay or even the embedding dimension. Moreover, if the attractor set is low-dimensional, one can try to plot it in a 2-dimensional or 3-dimensional plot.

Appendices

Appendix A

In this appendix we prove the Lemma 2.6, appearing in Chapter 3 of the lecture [13]. This lemma appears often in Takens' Theorem proof. It needs some details that we introduce here.

Definition A.1. A *n-cube* $C \subset \mathbb{R}^n$ of *edge* $\lambda > 0$ is a product of closed intervals of length λ:

$$C = I_1 \times \cdots \times I_n \subset \mathbb{R}^n, \quad |I_i| = \lambda.$$

The *measure* of C is

$$\mu(C) = \lambda^n.$$

Now we introduce the concept of measure zero. To introduce correctly this concept we should do a course on Lebesgue's Measure. However, we only use the following definition.

Definition A.2. A subset $X \subset \mathbb{R}^n$ has *measure zero* if for every $\epsilon > 0$, we can cover X by a countable family of n-cubes $\{C_j\}_{j=1}^{\infty}$, the sum of whose measures is less than ϵ. That is,

$$\sum_{j=1}^{\infty} \mu(C_j) < \epsilon.$$

We have some basic properties about sets of measure zero.

Proposition A.1. Let $X \subset \mathbb{R}^n$ be a set of measure zero. If $Y \subset X$, then Y has measure zero too.

Proof. If $\{C_j\}_{j \in J}$ is a family of cubes that covers X and the sum of their measures is less than ϵ, then this same family covers also Y. $\square$

Proposition A.2. A countable union of sets of measure zero has measure zero.

Proof. In fact, if we have a countable family $\{X_i\}_{i\in I}$ of sets of measure zero, then for every X_i there exists a countable family of n-cubes $\{C_j\}_{j\in J}$ satisfying

$$\sum_{j\in J}\mu(C_{ij}) < \epsilon_i.$$

Hence, we choose $\epsilon_i = \frac{\epsilon}{2^i} > 0$. Therefore,

$$\sum_{i\in I}\sum_{j\in J}\mu(C_{ij}) = \sum_{i\in I}\frac{\epsilon}{2^i} \leq \epsilon.$$

□

The following lemma tells us how to preserve measure zero sets by applications.

Lemma A.1. Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}^n$, $f \in \mathcal{C}^1$. If $X \subset U$ has measure zero, then $f(X)$ has also measure zero.

Proof. By the Mean Value Theorem, for every point $x \in X$, there exists a neighborhood B such that the point is uniformly bounded on B. Then,

$$|f(y) - f(z)| \leq k|y - z|,$$

for $y, z \in B$ and some $k > 0$. Moreover, if $C \subset B$ is an n-cube of edge $\lambda > 0$, we have $\mu(C) = \lambda^n$. Hence for every $y, z \in C$,

$$|y - z| \leq \sqrt{\lambda^2 + \cdots + \lambda^2} = \sqrt{n\lambda^2} = \sqrt{n}\lambda.$$

Thus, $f(C)$ is contained in an n-cube C' such that $\mu(C') \leq n^{n/2}k^n\lambda^n = n^{n/2}k^n\mu(C)$.

As X is a subset of $\mathbb{R}^n$, we can cover X by a countable number of compact sets $\{X_j\}_{j\in\mathbb{Z}^+}$:

$$X = \bigcup_{j=1}^{\infty} X_j,$$

where X_j is contained in an open ball B_j. Hence, for every B_j we have an n-cube $C_j \subset B_j$ such that $f(C_j) \subseteq C_j'$ and $\mu(C_j') \leq n^{n/2}k_j^n\mu(C_j) = L^n\mu(C_j)$, where $L = \sqrt{n}k_j$. For each $\epsilon > 0$, since X has measure zero, X_j has measure zero and $X_j \subset \cup_{k\in K}C_{j,k}$ where each $C_{j,k}$ is an n-cube and $\sum\mu(C_{j,k}) < \epsilon$. Therefore, $f(X_j) \subset \cup_{k\in K}C'_{j,k}$ and

$$\sum_{k\in K}\mu(C'_{j,k}) < \sum_{k\in K}L^n\mu(C_{j,k}) < L^n\epsilon.$$

Hence, with a good choice of ϵ, we check that $f(X_j)$ has measure zero and since $\cup_{j\in J}f(X_j) = f(X)$, $f(X)$ is a countable union of measure zero sets and therefore it has measure zero. □

Let M be a m-dimensional manifold, $\mathcal{C}^r$-differentiable, $r \geq 1$. A subset $X \subset M$ has *measure zero* if for every local chart (U, h), the set $h(U \cap X) \subset \mathbb{R}^m$ has measure zero.

Proposition A.3. An n-cube in $\mathbb{R}^n$ does not have measure zero.

Proof. Let C be an n-cube of edge $\lambda > 0$, thus $\mu(C) = \lambda^n$. Suppose that there exists an n-cube cover $\cup_{k\in K} D_k$ such that $\sum_{k\in K} \mu(D_k) < \epsilon$, for all $\epsilon > 0$. In particular, for any $0 < \epsilon < \lambda^n$. However,

$$\mu(\bigcup_{k\in K} D_k) \leq \sum_{k\in K} \mu(D_k) \Longrightarrow \lambda^n \leq \sum_{k\in K} \mu(D_k) < \epsilon < \lambda^n,$$

hence we have a contradiction. □

Since the cubes do not have measure 0 and the subsets of measure zero have measure zero, a cube cannot be contained in a measure zero set. Hence, the interior of a measure zero set is empty (if not, it would mean that there exists some ball contained in the set, but in every ball we have some compact cube).

Let $X \subset \mathbb{R}^n$ be a closed set with measure zero. Then, $\overset{\circ}{\overline{X}} = \overset{\circ}{X} = \emptyset$. Hence $\overline{X} \neq \mathbb{R}^n$ and it is not dense. However, its complement X^c is dense, since there is a property in a metric space that $\overline{(X^c)} = (\overset{\circ}{X})^c$. Thus, $\overline{(X^c)} = \mathbb{R}^n$ and this implies that X is dense.

Let $X \subset M$ be a closed subset of a manifold that has measure zero. Therefore, for every local chart (U, h), the set $h(U \cap X) \subset \mathbb{R}^n$ has measure zero and

$$\begin{aligned} \overline{\overset{\circ}{h(U \cap X)}} &= h(\overline{\overset{\circ}{U \cap X}}) \subseteq h(\overbrace{\overline{U} \cap \overline{X}}^{\circ}) \\ &\subseteq h(\overset{\circ}{\overline{U}} \cap \overset{\circ}{\overline{X}}) \subseteq h(U \cap \overset{\circ}{X}) = \emptyset. \end{aligned}$$

Hence, $\overset{\circ}{X}$ does not intersect any local chart. In addition, $h(U \cap (\overset{\circ}{X})^c) = h(U)$ and then $h(U \cap \overline{X^c}) = h(U)$ for every local chart. Thus, X^c is dense. Therefore, if we want to check if a set is dense, we could try to prove that its complement has measure zero. Now, we can prove the main statement:

Lemma A.2. Let M and N be manifolds with dimensions m and n respectively, $m < n$. If $f : M \to N$ is a $\mathcal{C}^1$ function, then $N \setminus f(M)$ is dense in N.

Proof. We prove that its complement $f(M)$ has measure zero. Firstly, we show that if $U \subset \mathbb{R}^m$ is open and $g : U \to \mathbb{R}^n$ is $\mathcal{C}^1$, with $m < n$, then $g(U) \subset \mathbb{R}^n$ is dense. We can identify U as a subset of $\mathbb{R}^n$:

$$U \cong U \times \{0\} \subset U \times \mathbb{R}^{n-m} \subset \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n.$$

Hence, $U \times \{0\}$ has measure zero in $U \times \mathbb{R}^{n-m}$. We have an application:

$$g\pi :, \mathbb{R}^n \to U \to \mathbb{R}^n$$

where π is the projection of the first m components (hence $\pi \in \mathcal{C}^1$) and if $g \in \mathcal{C}^1$, the composition $g\pi$ is $\mathcal{C}^1$. Therefore, the fact that $U \times \{0\}$ has measure zero implies that $g(U)$ has measure zero too, by Lemma A.1.

Finally, let (U, h) be a local chart in M and (V, g) a local chart in N. We must show that $g(V \cap f(U))$ has measure 0. The real map $gfh^{-1} : h(U) \to g(V)$ is $\mathcal{C}^1$, since it is a composition of $\mathcal{C}^1$ maps. $h(U)$ is an open set of dimension m. Thus, the image $g(f(U))$ has measure zero, and since $g(V \cap f(U)) \subset g(f(U))$, $g(V \cap f(U))$ has measure zero. $\square$

Appendix B

In this appendix we prove the *Preimage Theorem* (it is our Lemma 2.7). We follow the argument given by [17], but it is a common theorem that appears in a lot of books of Differential Topology.

To prove this, we need a definition and a previous theorem.

Definition B.1. The *canonical submersion* between two real spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ is the standard projection of $\mathbb{R}^m$ into $\mathbb{R}^n$ for $m \geq n$, by the first n coordinates:

$$(x_1, \ldots, x_n, \ldots, x_m) \to (x_1, \ldots, x_n).$$

We usually write the map by cs_n.

Given two real spaces, the canonical submersion is well defined. Therefore, we could say $\mathrm{cs}_n = \mathrm{cs}$ to simplify notation. However, we consider that the notation cs_n helps to follow the plot.

Theorem B.1 (Local Submersion Theorem)**.** *Let M, N be two manifolds, $\dim M = m$ and $\dim N = n$. Suppose that $f : M \to N$ is a submersion at $x \in M$, and $y = f(x)$. Then there exist local charts around x, (U, h) and around y, (V, g) such that $gfh^{-1}(h(z)) = \mathrm{cs}_n(h(z))$.*

Proof. Let $h^{-1}(0) = x$ and $g^{-1}(0) = y$. Consider the following diagram

$$\begin{array}{ccc} M & \xrightarrow{\quad f \quad} & N \\ {\scriptstyle h}\big\downarrow & & \big\downarrow{\scriptstyle g} \\ \mathbb{R}^m & \xrightarrow{\quad gfh^{-1} \quad} & \mathbb{R}^n \end{array}$$

We shall choose $\hat{h}$ such that the map $gf\hat{h}^{-1}$ is a canonical submersion. Let $\rho = gfh^{-1}$. Since f is a submersion at x, the derivative $d\rho(h(x)) = d\rho(0)$ is surjective. Thus, the rank of $d\rho(0)$ is maximum, $\mathrm{rn}(d\rho(0)) = l$. Therefore, we can make a linear change of coordinates in $\mathbb{R}^k$ and we can assume that $d\rho(0)$ is a matrix $\begin{pmatrix}\mathrm{Id}_l & 0\end{pmatrix}_{(l\times k)}$, where Id_l is the identity matrix. Let $a \in U$, $a = (a_1, \dots, a_m)$. We define

$$\begin{array}{rccl} G: & U & \to & \mathbb{R}^m \\ & a & \mapsto & (\rho(a), a_{n+1}, \dots, a_m). \end{array}$$

The derivative at 0 is

$$dG(0) = \Big(\frac{dG_0}{a_1}\Big|\dots\Big|\frac{dG_0}{a_m}\Big).$$

The first n columns are the derivative of $\rho(0)$, hence

$$dG(0) = \begin{pmatrix} d\rho(0) & 0 \\ 0 & I_{k-l} \end{pmatrix} = I_k.$$

Thus, G is a local diffeomorphism at 0. Therefore, there exists a neighborhood U' of 0 such that the function G^{-1} exists, by the Inverse Function Theorem. Observe that

$$\begin{array}{rccccl} \mathrm{cs}_n \circ G: & U & \to & \mathbb{R}^k & \to & \mathbb{R}^l \\ & a & \mapsto & (g(a), \dots, a_m) & \mapsto & g(a). \end{array}$$

Hence, $\mathrm{cs}_n \circ G = g$ and then $g \circ G^{-1} = \mathrm{cs}_n$. Let $\hat{h} = G \circ h$. We choose the local chart $(U', \hat{h})$. Thus, the diagram is

$$\begin{array}{ccc} M & \xrightarrow{\ f\ } & N \\ \hat{h}\downarrow & & \downarrow g \\ \mathbb{R}^m & \xrightarrow{\ gf\hat{h}^{-1}\ } & \mathbb{R}^n \end{array}$$

In this case, $gf\hat{h}^{-1}$ is a canonical submersion, since

$$\begin{aligned} gf\hat{h}^{-1}(\hat{h}(x)) &= gfh^{-1}G^{-1}(\hat{h}(x)) = \rho\,(G^{-1}(\hat{h}(x))) \\ &= \rho\,(G^{-1}(x_1, \dots, x_n, \dots, x_m)) = \rho\,(\rho^{-1}(x_1, \dots, x_n)) \\ &= (x_1, \dots, x_n) = \mathrm{cs}_n(h(x)). \end{aligned}$$

□

The main result follows from the previous fact.

Lemma B.1 (Preimage Theorem)**.** Let M and N be manifolds with dimensions m and n respectively, $m > n$, and $f : M \to N$ be a $\mathcal{C}^1$ function. Consider $q \in N$. If f is submersive at every p such that $f(p) = q$, then the set $f^{-1}(q)$ is a submanifold of M, of dimension $m - n$.

Proof. Suppose f is a submersion at a point $p \in f^{-1}(q)$. Let (U, h) be a local chart around p and (V, g) another local chart around q, such that $gfh^{-1}(h(x)) = \mathrm{cs}_n(h(x))$ for all $x \in U$ and $g(q) = 0$ (by Theorem B.1). Then $f^{-1}(q) \cap U$ is the set of points where $x_1 = 0, \dots, x_n = 0$, since

$$\mathrm{cs}_n^{-1} g(q) = \mathrm{cs}_n^{-1}(0, \dots, 0) = (0, \dots, 0, x_{n+1}, \dots, x_m).$$

The functions $x_{n+1}, \dots, x_m$ are diffeomorphisms between the restrictions of the real spaces. Hence, $f^{-1}(q) = \{0\} \times \mathbb{R}^k$ is a submanifold. The local charts are given by $(\hat{U}, \hat{h})$, where $\hat{U} = f^{-1}(q) \cap U$ and $\hat{h} = \pi_{(2)} \circ h$, where

$$\begin{array}{rccc} \pi_{(2)}: & \to & \mathbb{R}^n \times \mathbb{R}^{m-n} \\ & (a, b) & \mapsto & b \end{array}$$

is an homeomorphism. Therefore, the dimension is $m - n$. □

Appendix C

In this appendix, we prove the *Lebesgue's Lemma*. This lemma needs some concepts of metric spaces, such as the diameter of a set. These concepts we take for granted and we only give the proof. The proof appears in the lecture [18].

Lemma C.1 (Lebesgue's Lemma)**.** Let $\mathcal{X}$ be a compact metric space and let $\mathcal{U}$ be an open cover of $\mathcal{X}$. Then there exists a real number $\delta > 0$ such that any subset of $\mathcal{X}$ of diameter less than δ is contained in some member of $\mathcal{U}$. δ is called the *Lebesgue number* of $\mathcal{U}$.

Proof. We suppose the opposite; then, there exists some sequence of subsets of $\mathcal{X}$, say $\{A_n\}_{n\in\mathbb{N}}$ such that any $A_n \subset U$, $U \in \mathcal{U}$ and furthermore $\lim_{n\to\infty} \operatorname{diam}(A_n) = 0$. For every $n \in \mathbb{N}$, we choose $x_n \in A_n$. Then, since it is compact, it is sequentially compact and it has some convergent subsequence. We can suppose that $\lim_{n\to\infty}\{x_n\} = x$. Since $\mathcal{U}$ is a cover set of $\mathcal{X}$, we can take $x \in U \in \mathcal{U}$, for some U. U is an open set, hence there is an open ball $B(x,\epsilon) \subseteq U$, for some $\epsilon > 0$. We choose $N > 0$ such that

- $\operatorname{diam}(A_N) < \epsilon/2$, since the diameters tend to zero. Then, $d(a, x_N) < \epsilon/2$, for all $a \in A_N$. [1]
- $x_n \in B(x, \epsilon/2)$ for some n, since $\{x_n\}_{n\in\mathbb{N}} \to x$. Then, $d(x_N, x) < \epsilon/2$.

By the properties of the distance functions, if $a \in A_N$, then

$$d(a, x) \leq d(a, x_N) + d(x_N, x) = \epsilon/2 + \epsilon/2 = \epsilon.$$

Therefore, $a \in U$ and thus $A_N \subset U$. This contradicts the fact that A_N is not contained in any U. □

[1] We denote by $d(x, y)$ the distance between $x, y \in \mathcal{X}$.

Appendix D

In this appendix we prove Lemma 2.9. We only require the equivalence of compactness and sequentially compactness, as in Lebesgue's Lemma. Thus we prove it directly. The arguments of the proof follows the ones given by [13].

Lemma D.1. Let $U \subset \mathbb{R}^m$ be an open set and $W \subset U$ an open set with compact closure $\overline{W} \subset U$. Let $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^1$ embedding. There exists $\epsilon > 0$ such that if $g : U \to \mathbb{R}^n$ is $\mathcal{C}^1$ and

$$\|Dg(x) - Df(x)\| < \epsilon \text{ and } |g(x) - f(x)| < \epsilon$$

for all $x \in W$, then $g|_W$ is an embedding.

Proof. Similarly to Theorem 2.4, there exists $\epsilon_0 > 0$ sufficiently small such that if $g \in \mathcal{C}^1(U, \mathbb{R}^k)$ and $\|Dg(x) - Df(x)\| < \epsilon_0$, for all $x \in W$, then $g|_W$ is an immersion. We suppose that the lemma is false. Since the lemma is false for W, it is false for $\overline{W}$. Hence there exists a sequence of immersive functions $\{g_n\}_{n\in\Lambda}$, $g_n \in \mathcal{C}^1(U, \mathbb{R}^n)$ such that

$$\|Dg_n(x) - Df(x)\| \to 0, \tag{D.1}$$

$$|g_n(x) - f(x)| \to 0, \tag{D.2}$$

but it is not an embedding in $\overline{W}$ and thus by Proposition 2.9, the g_n is not injective. Consequently, there exists $a_n, b_n \in W$ such that $g_n(a_n) = g_n(b_n)$, with $a_n \neq b_n$. Since $\overline{W}$ is compact and since we use the real topology, it is also sequentially compact and hence we can choose $a_n \to a \in U$ and $b_n \to b \in U$. Hence, by (D.2), $f(a) = f(b)$ and since f is injective, $a = b$. As, $a_n \neq b_n$, we define

$$v_n = \frac{a_n - b_n}{|a_n - b_n|}.$$

Thus, $v_n \in \mathcal{S}^{m-1}$ and by continuity, $v_n \to v \in \mathcal{S}^{m-1}$. Therefore, by the Mean value theorem for vector-valued functions,

$$g_n(a_n) - g_n(b_n) = \int_0^1 Df(b_n + (a_n - b_n)t)\, dt \cdot (a_n - b_n).$$

Hence, if

$$C_n = \frac{|g_n(a_n) - g_n(b_n) - Dg_n(b)(a_n - b_n)|}{|a_n - b_n|},$$

then

$$\begin{aligned} C_n &= \frac{|\int_0^1 Df(b_n + (a_n - b_n)t)\, dt \cdot (a_n - b_n) - Dg_n(b)(a_n - b_n)|}{|a_n - b_n|} \\ &= |\textstyle\int_0^1 (Df(b_n + (a_n - b_n)t) - Dg_n(b))\, dt| \cdot |a_n - b_n| \frac{1}{|a_n - b_n|} \\ &= |\textstyle\int_0^1 (Df(b_n + (a_n - b_n)t) - Dg_n(b))\, dt| \\ &\leq \textstyle\int_0^1 \|Df(b_n + (a_n - b_n)t) - Dg_n(b)\|\, dt. \end{aligned}$$

We note by (D.1) that

$$\int_0^1 \|Df(b_n + (a_n - b_n)t) - Dg_n(b)\|\, dt \to 0.$$

Hence, $C_n \to 0$. Since

$$\frac{|g_n(a_n) - g_n(b_n)|}{|a_n - b_n|} \to 0,$$

we have

$$\frac{Dg_n(b)(a_n - b_n)}{|a_n - b_n|} \to 0,$$

that is, $Dg_n(b)v_n \to 0$. However, $v_n \to v \neq 0$, and thus $Dg_n(b) \to 0$. Consequently, g_n is not an immersion at $b \in U$: which is a contradiction. Therefore, g is an embedding in $\overline{W}$ and moreover it is also an embedding in W. $\square$

Appendix E

In this appendix, we establish and prove Whitney Embedding Theorem for compact manifolds. There are two versions of the Whitney Theorem; a general version and the explicit one. We prove the general version and state the other one. We follow the proof given by [12], but it can be also found in a basic lecture of differential topology: for example in [13].

Theorem E.1 (General Compact Whitney embedding)**.** Any compact manifold can be embedded in $\mathbb{R}^N$, for sufficiently large N.

Proof. Let m be the dimension of the manifold. Recall that we can always take a regular covering, by 2.8. Moreover, since the manifold is compact, we can take a finite regular covering. Let $\{(U, h_i)\}_{i=1}^k$ be the regular recovering, with $\{V_i = h_i^{-1}(B(0,1))\}$, $V_i \subset U_i$. Choose a partition of unity $\{f_1, \dots, f_k\}$ subordinate to the regular covering, as in Theorem 2.2. We define the maps

$$\begin{array}{rccl} \phi_i : & M & \to & \mathbb{R}^m \\ & x & \mapsto & \phi_i(x) = \begin{cases} f_i(x)h_i(x) & x \in U_i, \\ 0 & x \notin U_i. \end{cases} \end{array}$$

We note that ϕ_i is a $\mathcal{C}^1$ function. The only problem should be when we change from U_i to its complementary. However, since $\operatorname{supp} f_i \subset h_i^{-1}\overline{B(0,2)}$, we have that $U_i^c \subset (h_i^{-1}\overline{B(0,2)})^c$ and this implies that in the boundary it is locally constant 0. Then we define the continuous map

$$\begin{array}{rccl} \rho : & M & \to & \mathbb{R}^{k(m+1)} \\ & x & \mapsto & \rho(x) = (\phi_1(x), \dots, \phi_k(x), f_1(x), \dots, f_k(x)). \end{array}$$

We note that every ϕ_i has m components and we have k of them: thus we have $k \cdot m$ components plus the k components by f. Hence, it is well-defined. Since $\sum f_i = 1$, we note that $\rho(x) = \rho(x')$ implies that $f_i(x) = f_i(x')$, for every $i = 1, \ldots, k$. Since $\sum f_i = 1$, there exists i_0 such that $f_{i_0}(x) = f_{i_0}(x') \neq 0$ Hence, $x, x' \in U_i$. The coordinate $m \cdot i$ is $h_i(x)f_i(x) = h_i(x')f_i(x')$ and thus $h_i(x) = h_i(x')$. Since h_i is injective (is an homeomorphism), we have $x = x'$. Therefore, ρ is injective.

It remains to prove that $D\rho$ is injective, because an embedding in a compact set is the same as an injective immersion and we have just seen that it is injective. Let $x \in M$, then the differential $D\rho$ maps $v \in T_xM$ to

$$\Big(Df_1(v)h_1(x) + f_1(x)Dh_1(v), \ldots, Df_k(v)h_k(x) + f_k(x)Dh_k(v), Df_1(v), \ldots, Df_k(v)\Big).$$

Suppose that for some $v \neq 0$, we have the matrix equal to 0. Therefore, every component is zero. Thus, $Df_i(v) = 0$, for all $i = 1, \ldots, k$ and then $f_i(x)Dh_i(v) = 0$, for all $i = 1, \ldots, k$. However, since h_i is a smooth and injective function, we have $Dh_i(v) \neq 0$, and moreover there exists some i such that $f_i(x) \neq 0$. Therefore, we have some i such that $f_i(x)Dh_i(v) \neq 0$: it contradicts the fact that there is some $v \neq 0$ such that $D\rho(v) = 0$. Hence, it is immersive and thus an embedding. □

Whitney's Theorem establishes that every compact manifold can be embedded in some $\mathbb{R}^N$, for N sufficiently large. This allows to understand a compact manifold as a subset of a real space. In particular, we can inherit the metric of the real space and we obtain the following corollary:

Corollary E.1. Every compact manifold is metrizable.

We now recall the explicit version of the previous theorem.

Theorem E.2 (Compact Whitney Embedding Theorem)**.** Every compact manifold of dimension m can be embedded in $\mathbb{R}^{2m+1}$.

Furthermore, we can formulate en equivalent version for immersions, since immersions allow auto-intersections.

Theorem E.3 (Compact Whitney Immersivity Theorem)**.** Every compact manifold of dimension m can be immersed in $\mathbb{R}^{2m}$.

Index

Bibliography

[1] J. P. Huke. Embedding nonlinear dynamical systems: A guide to takens' theorem. *University of Manchester*, 2006.

[2] F. Takens. Detecting strange attractors in turbulence, 1981. Lecture notes.

[3] Xavier Bornas. *Psicopatologia i Caos*. Bubok Publishing S.L.

[4] GeLi Wang, PeiCai Yang, JianChun Bian, and XiuJi Zhou. A novel approach in predicting non-stationary time series by combining external forces. *Chinese Science Bulletin*, 56(28-29):3053–3056, 2011.

[5] Hao Ye, Richard J. Beamish, Sarah M. Glaser, Sue C. H. Grant, Chih-hao Hsieh, Laura J. Richards, Jon T. Schnute, and George Sugihara. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proceedings of the National Academy of Sciences*, 112(13):E1569–E1576, 2015.

[6] Frank J. Fabozzi Sergio M. Focard. *The mathematics of financial modeling & investment management*. Wiley Finance, 2004.

[7] J. A. Yorke T. Sauer and M. Casdagli. Embedology. *J. of Statistical Physics*, 65:579–616, 1991.

[8] Eugene M. Izhikevich. *Dynamical Systems in Neuroscience*. Massachusetts Institute of Technology, 2007.

[9] K. D. Joshi. *Introduction to General Topology*. Wiley Eastern Limited, 1983.

[10] C. Kosniowski. *A First Course in Algebraic Topology*. University of Cambridge, 1980.

[11] N. Batle and F. Rosselló. *Topologia General*. Universitat de les Illes Balears, 2000.

[12] M. Gualtieri. Geometry and topology, 1300y. `http://www.math.toronto.edu/mgualt/MAT1300/1300%20Lecture%20notes.pdf`, 2008. Teaching notes.

[13] Morris W. Hirsch. *Differential Topology.* Springer-Verlag New York Inc., 1976.

[14] G. Teschl. *Ordinary Differential Equations and Dynamical Systems.* American Mathematical Society, 2012.

[15] L. Noakes. The takens embedding theorem. *International Journal of Bifurcation and Chaos*, 1991.

[16] V. Guillemin M. Golubitsky. *Stable Mappings and Their Singularities.* Springer-Verlag New York Inc., 1973.

[17] A. Pollack V. Guillemin. *Differential Topology.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

[18] M. A. Armstrong. *Basic Topology.* Springer-Verlag, New York, 1983.

[19] Joan C. Artés Freddy Dumortier, Jaume Llibre. *Qualitative Theory of Planar Differential Systems.* Springer-Verlag, New York, 2006.

[20] Giuseppe Iurato. On the history of differentiable manifolds. *International Mathematical Forum*, Vol. 7:477–514, 2012.