THE LOCAL STRUCTURE OF SMOOTH MAPS OF MANIFOLDS

By

Jonathan Michael Bloom

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
BACHELORS OF ARTS WITH HONORS
AT
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2004

Table of Contents

Acknowledgements

I am profoundly grateful to my advisor, Professor Peter Kronheimer. Without his skillful guidance, this thesis would not have been possible. Although my topic is not his primary field of study, he always knew the next important question to be asked and connection to be explored; time and again he saved me from losing my way. On a personal level, I am thankful for his compassionate supervision and the confidence he placed in me. I would also like to thank Professor Barry Mazur, who graciously met with me long ago to discuss possible thesis topics. I walked straight from that meeting to the library and buried myself in a stack of books on Singularity Theory. I could not have hoped for a more satisfying topic of study. I am indebted to Professor Wilfried Schmid for possessing the boldness to define objects the right way from the start and the pedagogical talent to do so effectively. After four years of thought, I can say that I understand most of what he had hoped to convey. I would never have had this opportunity in the first place had it not been for the inspirational efforts of the late Professor Arnold Ross, who encouraged me to think deeply of simply things.

I am deeply indebted to Peter Anderegg, Benjamin Bakker, and Gopal Sarma, each of whom selflessly devoted many hours to proofreading my drafts. I would also like to thank my parents and grandparents for their loving support and grand care packages; my roommates Chris, Christine, James, Jenny, Monica, Nico, Ryan, Wendy, and especially Daniela, who was a caring friend during the spring break; Ameera, George, and Jennifer for always checking in on me; Owen and Nick for seeing the big picture; Sasha and Dimitar for their thesis camaraderie; Dr. John Boller for spurring a torrent of writing; George and David for LaTeXsupport; Geraldine for providing me with endless electronica and trip-hop; and Olga for keeping life exciting.

Cambridge, MA — Jonathan Bloom
April 5, 2004

Introduction

Functions, just like human beings,

are characterized by their singularities.

- P. Montel.

The aim of this thesis is to provide a self-contained introduction to the modern study of the local structure of smooth maps of manifolds. By smooth, we mean differentiable to all orders. We consider two smooth maps to have the same local structure at a point if the maps are locally equivalent up to a change of coordinates. A critical point of a smooth map is a point of the domain at which the derivative does not have full rank. In §1 we show that, away from critical points, a smooth map exhibits a single, simple local structure (that of a linear projection). Thus the remaining chapters are concerned with the structure of smooth maps at their critical points, i.e. with the structure of singularities. In general, infinitely many local structures can occur at singularities. However, with certain restrictions on dimension, almost all maps exhibit at most a select handful of these. Our goal, then, is to discern which types of singularities can generically occur, and then to establish normal forms for the structure of these singularities.

The first three chapters are devoted to developing machinery to this end, with an emphasis on transversality. In §1, we provide the necessary background on smooth manifolds and prove the aforementioned result on the local structure of maps away from singularities. In §2, we prove Sard's Theorem, a fundamental piece of analysis that will power the more elaborate machinery of §3. In §3, we introduce the notion of transversality and prove a variety of elementary results. We then construct the jet bundle of smooth maps of manifolds and use it to place a topology on the space of smooth maps. Finally, we prove and then generalize our fundamental tool: the Thom Transversality Theorem. This theorem is a potent formalization of the intuition that almost all maps are transverse to a fixed submanifold.

The latter three chapters then use these tools to accomplish our goal in the three simplest situations, respectively. In §4, we look at smooth functions (by which we mean smooth maps of manifolds down to $\mathbb{R}$). This is the situation studied in Morse Theory. We will see that the generic singularities occur at so-called non-degenerate critical points and then establish normal quadratic forms for smooth functions at such points. In §5, we consider the opposite extreme, where the codomain has at least twice the dimension of the domain. We will see that in these dimensions there are no generic singularities, and then we will extend this result to prove the famous Whitney Embedding Theorem: Every smooth n-manifold can be embedded into $\mathbb{R}^{2n+1}$. Lastly, in §6, we look at smooth maps between 2-manifolds and show that generically the singularities consist of folds along curves with isolated cusps. We then establish normal forms for folds and simple cusps.

Chapter 1

Preliminaries: Smooth Manifolds

In §1.1, we present the necessary background on smooth manifolds and smooth maps. In §1.2, we give normal forms for the local structure of smooth maps at regular points.

1.1 Smooth Manifolds and Smooth Maps

The reader who is already familiar with smooth manifolds is advised to skim this section. Our treatment is adapted from Lang and the notation is standard [7]. We define manifolds as abstract objects having certain properties rather than as special subsets of Euclidean space. The latter approach, taken in [6] and [14], has the advantage of concreteness. We often think of manifolds this way, and it is perhaps best as a first exposure. In §5, we will see that the abstract definition of a manifold is in fact no more general: any manifold can be realized as a submanifold of some Euclidean space. Since the concrete definition is equivalent, one might wonder why we bother with the abstract definition in the first place.

In fact, the abstract approach echoes a general theme in higher mathematics: we define (or redefine) a category of object by its relevant properties and then prove that any object with these properties is in fact in the category. By proceeding abstractly, we can define the object with as little extraneous data as possible. The concrete approach to manifolds fails in this regard since it equips each manifold with a non-canonical embedding. To see why this excess baggage causes difficulty, consider the task of constructing new manifolds. When an abstract manifold does not arise naturally as a subset of Euclidean space, it is often far simpler to verify that it satisfies the properties of the abstract definition than to find a real embedding.

So without further ado, let X be set.

Definition 1.1.1. A *smooth atlas* on X is a collection of pairs (U_i, φ_i) satisfying the following conditions for some fixed n:

(1) Each U_i is a subset of X and the U_i cover X.

(2) Each φ_i is a bijection of U_i onto an open subset $\varphi_i(U_i)$ of $\mathbb{R}^n$ and $\varphi_i(U_i \bigcap U_j)$ is open in $\mathbb{R}^n$ for each pair i, j.

(3) The map

$$\varphi_j \cdot \varphi_i^{-1} : \varphi_i(U_i \bigcap U_j) \to \varphi_j(U_i \bigcap U_j)$$

is a smooth diffeomorphism for each pair i, j.

We can give X a topology in a unique way such that each U_i is open and each φ_i is a homeomorphism. This topology inherits the properties of local compactness and second countability from Euclidean space. In the theory of manifolds, we will also require this topology to be Hausdorff and, thus, paracompact.

Each pair (U_i, φ_i) is called a *chart* of the atlas. If $x \in U_i$ then we call (U_i, φ_i) a *chart at x*. Let U be an open subset of X and φ a homeomorphism from U onto some open subset of $\mathbb{R}^n$. Then the pair (U, φ) is *compatible* with the atlas (U_i, φ_i) if

$\varphi \cdot \varphi_i^{-1}$ is a diffeomorphism for each i such that $U \bigcap U_i \neq \emptyset$. Two atlases on X are *compatible* if every chart of the first is compatible with every chart of the second. It is straightforward to check that this gives an equivalence relation on the collection of atlases.

Definition 1.1.2. A *smooth manifold* is a set X together with an equivalence class of atlases on X, called its *differential structure*. A smooth manifold is n*-dimensional* if each chart maps into $\mathbb{R}^n$.

A chart $\varphi : U \to \mathbb{R}^n$ is given by n coordinate functions $\varphi_1, ..., \varphi_n$, often written $x_1, ..., x_n$. Under φ, a point $p \in U$ has image $(x_1(p), ..., x_n(p))$, or, abusing notation, simply $(x_1, ..., x_n)$. We call the functions $x_1, ..., x_n$ a set of *local coordinates* on U, and sometimes we will refer to U as a *coordinate neighborhood* (or just 'nbhd').

Given smooth manifolds X and Y with atlases (U_i, φ_i) and (V_i, ψ_i), respectively, we can equip the set $X \times Y$ with the atlas $(U_i \times V_j, \varphi_i \times \psi_j)$. This manifold is the *product* of X and Y.

Now let X be a smooth manifold and Z a subset of X. Suppose that for each point $z \in Z$ there exists a chart (U, φ) at z such that:

(1) φ gives a diffeomorphism of U with a product $V_1 \times V_2$ where V_1 is an open subset of $\mathbb{R}^m$ and V_2 is an open subset of $\mathbb{R}^p$, with m and p fixed.

(2) $\varphi(Z \bigcap U) = V_1 \times v_2$ for some $v_2 \in V_2$.

It is not difficult to verify that the collection of pairs $(Z \bigcup U, \varphi|_{Z \bigcup U})$ constitutes a smooth atlas for Z. The set Z equipped with this differential structure is called a *submanifold* of X. If follows that at every point of a submanifold Z, we can find a chart sending $Z \subset X$ diffeomorphically to $\mathbb{R}^k \subset \mathbb{R}^n$. Also note that an open subset of X can be given the structure of an open submanifold, and a submanifold of a strictly lower dimension is necessarily closed in X. The *codimension* of a submanifold Z of X is defined by the equation codim $Z = \dim X - \dim Z$.

We use the differential structure on a manifold to define the notion of a smooth map of manifolds.

Definition 1.1.3. Let X and Y be two smooth manifolds. Let $f : X \to Y$ be a map.

(1) f is *smooth* if, given $x \in X$, there exists a chart (U, φ) at x and and a chart (V, ψ) at $f(x)$ such that $f(U) \subset V$ and the map $\psi \cdot f \cdot \varphi^{-1} : \varphi(U) \to \psi(V)$ is smooth.

(2) By $\mathcal{C}^\infty(X, Y)$ we denote the space of smooth maps from X to Y.

(3) f is a *diffeomorphism* if it is smooth and has a smooth inverse.

In general, there is no canonical choice of local coordinates near a point of a manifold. The space of maps seems blurred, as distinct maps of manifolds can yield identical maps of Euclidean spaces when pulled back and pushed forward via distinct charts. We can make this notion of equivalence of maps precise up to a "change of coordinates" without invoking charts at all.

Definition 1.1.4. Let $f, g : X \to Y$ be smooth maps.

(1) f and g are *equivalent* if there exist diffeomorphisms h_1 of X and h_2 of Y such that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle h_1}\downarrow & & \downarrow{\scriptstyle h_2} \\ X & \xrightarrow[g]{} & Y \end{array}$$

(2) f and g are *locally equivalent* at a point $x \in X$ if there exists an open neighborhood U of x such that $f|_U$ and $g|_U$ are equivalent. Here U is considered as an open submanifold.

With f and g as above, let (U, φ) and (V, ψ) be charts at x and $g(x)$ respectively with $f(U) \subset V$. Then if f and g are locally equivalent at x, we sometimes say (even though it is technically not true) that f is locally equivalent at x to the map $\psi \cdot g \cdot \varphi^{-1}$ of Euclidean spaces. This second use of the term "local equivalence" proves convenient for statements about the normal forms of maps and should not cause confusion in practice. Note that both equivalence and local equivalence are equivalence relations on the space of smooth maps from X to Y, denoted $\mathcal{C}^\infty(X, Y)$.

We would like the derivative of a smooth map of manifolds to be a smooth map of manifolds as well. To this end, we introduce the notions of the tangent space and tangent bundle of a smooth manifold.

Definition 1.1.5. Let X be an n-dimensional smooth manifold and $x \in X$. Consider triples of the form (U, φ, v), where (U, φ) is a chart at x and $v \in \mathbb{R}^n$.

(1) Two triples (U, φ, v) and (V, ψ, w) are *equivalent* if

$$d(\psi \cdot \varphi^{-1})_{\varphi(x)} v = w.$$

This gives an equivalence relation on such triples.

(2) A *tangent vector* to X at x is an equivalence class of such triples.

(3) The *tangent space* to X at x, denoted $T_x X$, is the set of all tangent vectors to X at x. We can equip $T_x X$ with the topology of the vector space $\mathbb{R}^n$ via the bijection sending the equivalence class of (U, φ, v) to v, with (U, φ) a fixed chart at x.

(4) The *tangent bundle* of X is the set $TX = \bigsqcup_{x\in X} T_xX$ (disjoint union). We can canonically construct an equivalence class of atlases on TX in order to give TX the structure of a $2n$-dimensional smooth manifold.

Let $f : X \to Y$ be a smooth map of manifolds. Using charts we can interpret the derivative of f at x as the unique map $(df)_x : T_xX \to T_{f(x)}Y$ having the following property: If (U, φ) is a chart at x and (V, ψ) is a chart at $f(x)$ such that $f(U) \subset V$, and if $\vec{v} = (U, \varphi, v)$ is a tangent vector at x represented by v in the chart (U, φ), then $(df)_x(\vec{v})$ is the tangent vector at $f(x)$ represented by $d(\psi \cdot f \cdot \varphi^{-1})_{\varphi(x)}(v)$. Fixing these charts, we have the diagram

$$\begin{tikzcd}
T_xX \arrow[r, "(df)_x"] \arrow[d] & T_{f(x)}Y \arrow[d] \\
\mathbb{R}^n \arrow[r, "d(\psi \cdot f \cdot \varphi^{-1})_{\varphi(x)}"'] & \mathbb{R}^m
\end{tikzcd}$$

From the diagram it is clear that $(df)_x$ is linear. Recall that the corank of a linear map $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ is defined to be $\min\{n, m\} - \text{rank } \lambda$. If corank $\lambda = 0$ we say that λ has *full rank*.

Definition 1.1.6. Let $f : X \to Y$ be a smooth map of manifolds and $x \in X$.

The *rank* (*corank*) of f at x is the rank (corank) of $(df)_x$ as a linear map from T_xX to $T_{f(x)}Y$.

We will need one more elementary result on smooth manifolds [14, p.35].

Theorem 1.1.1 (Inverse Function Theorem). *Let X and Y be smooth manifolds of equal dimension, $f : X \to Y$ a smooth map, and $x \in X$ such that f has full rank at x. Then there exists an open nbhd U of x such that $f(U)$ is an open nbhd of $f(x)$ and $f|_U$ has a smooth inverse.*

The Inverse Function Theorem suggests that the rank of f at x encodes a good deal of information about the structure of f near x. We will begin to freely exploit this connection in the next section.

1.2 Linearization

The goal of this section is to completely determine the local structure of a smooth map at any point of full rank (up to local equivalence). At such a point, a smooth map must be at least one of the following types.

Definition 1.2.1. Let $f : X \to Y$ be a smooth map of manifolds with $q = f(p)$.

(1) f is an *immersion* at p if $(df)_p : T_pX \to T_qY$ is injective.

(2) f is a *submersion* at p if $(df)_p : T_pX \to T_qY$ is surjective.

(3) f is a *local diffeomorphism* at p if $(df)_p : T_pX \to T_qY$ is bijective.

(1′) f is an *immersion* if f is an immersion at p for every $p \in X$.

(2′) f is an *submersion* if f is an submersion at p for every $p \in X$.

(3′) f is an *local diffeomorphism* if f is a local diffeomorphism at p for every $p \in X$.

The following lemma immediately follows from the above definitions and the fact that $\dim X = \dim T_pX$.

Lemma 1.2.1. *Let $f : X \to Y$ be a smooth map of manifolds.*

(1) If f is an immersion at any point then $\dim X \leq \dim Y$.

(2) If f is a submersion at any point then $\dim X \geq \dim Y$.

(3) If f is a local diffeomorphism at any point then $\dim X = \dim Y$.

Lemma 1.2.1 suggests the logic behind the names immersion and submersion. The former places a smaller manifold into a larger one while the latter smoothly packs a larger manifold into a smaller one. It is important to note, however, that the image of either type of map need not be a submanifold. In §5, we will see examples of this and establish sufficient conditions for the image of a smooth map to be a submanifold.

We now construct normal forms for the local structure of immersions, submersions, and local diffeomorphisms. In fact, most of the work of this is encoded by the Inverse Function Theorem. To make this more explicit, we now express this theorem using the concept of local equivalence (in the sense following its definition in §1.1). The reader is encouraged to verify that the following theorem is simply a restatement of Theorem 1.1.1.

Theorem 1.2.2 (Inverse Function Theorem). *Let $f : X \to Y$ be a smooth map of manifolds. Suppose that f is a local diffeomorphism at a point $p \in X$. Then f is locally equivalent to the identity map $(x_1, ..., x_n) \mapsto (x_1, ..., x_n)$ at p.*

The Inverse Function Theorem allows us to locally linearize maps between manifolds of different dimension as well. We will first show this in the case of immersions (so necessarily $\dim X \leq \dim Y$).

Lemma 1.2.3 (Immersion Lemma). *Let $f : X \to Y$ be a smooth map of manifolds. Suppose f is an immersion at a point $p \in X$. Then f is locally equivalent to the map $(x_1, ..., x_n) \mapsto (x_1, ...x_n, 0, ..., 0)$ at p.*

Proof. Our strategy is to construct a map between equidimensional manifolds and then apply the Inverse Function Theorem. f is a immersion at p so locally we may assume $f : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^l$, where $l = m - n$. Then

$$(df)_p = \begin{bmatrix} (\frac{\partial f_1}{\partial x_1})_p & (\frac{\partial f_1}{\partial x_2})_p & \cdots & (\frac{\partial f_1}{\partial x_n})_p \\ \vdots & \vdots & & \vdots \\ (\frac{\partial f_m}{\partial x_1})_p & (\frac{\partial f_m}{\partial x_2})_p & \cdots & (\frac{\partial f_m}{\partial x_n})_p \end{bmatrix}$$

has rank n, where f_i is the ith coordinate function of f. Thus there exist n linearly independent rows of $(df)_p$. Without loss of generality we assume the first n rows are linearly independent (otherwise we can permute the coordinates on the range). Define $\overline{f} : \mathbb{R}^n \to \mathbb{R}^n$ by $\bar{f} = (f_1, ...f_n)$, so $\det(d\bar{f})_p \neq 0$.

Now define $F : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n \times \mathbb{R}^l$ by $F(x, y) = f(x) + (0, y)$. Then

$$(dF)_p = \left[\begin{array}{c|c} (d\bar{f})_p & 0 \\ \hline * & I_l \end{array}\right]$$

so $\det(dF)_p = \det(d\bar{f})_p \neq 0$. By the Inverse Function Theorem, F has a local smooth inverse G near p. On a neighborhood of p we have $G \cdot f(x) = G \cdot F(x, 0) = (x, 0)$. Thus f is linearized by a change of coordinates on the range. □

Next, we establish a normal form for submersions (so necessarily $\dim X \geq \dim Y$).

Lemma 1.2.4 (Submersion Lemma). *Let $f : X \to Y$ be a smooth map of manifolds. Suppose f is a submersion at a point $p \in X$. Then f is locally equivalent to the map $(x_1, ..., x_n) \mapsto (x_1, ..., x_m)$ at p.*

Proof. Again our strategy is to construct a map between equidimensional manifolds and then apply the Inverse Function Theorem. f is an immersion at p so locally we may assume $f : \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^m$, where $l = n - m$. Then

$$(df)_p = \begin{bmatrix} (\frac{\partial f_1}{\partial x_1})_p & (\frac{\partial f_1}{\partial x_2})_p & \cdots & (\frac{\partial f_1}{\partial x_n})_p \\ \vdots & \vdots & & \vdots \\ (\frac{\partial f_m}{\partial x_1})_p & (\frac{\partial f_m}{\partial x_2})_p & \cdots & (\frac{\partial f_m}{\partial x_n})_p \end{bmatrix}$$

has rank m, where f_i is the ith coordinate function of f. Thus there exist m linearly independent columns of $(df)_p$. Without loss of generality we assume the first m columns are linearly independent (otherwise we can permute the coordinates on the domain). Define $\bar{f} : \mathbb{R}^m \to \mathbb{R}^m$ by $\bar{f}(x_1, ..., x_m) = f(x_1, ..., x_m, 0, ...0)$, so $\det(d\bar{f})_p \neq 0$.

Now define $F : \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^l$ by $F(x, y) = (f(x, y), y)$. Then

$$(dF)_p = \left[\begin{array}{c|c} (d\bar{f})_p & * \\ \hline 0 & I_l \end{array}\right]$$

so $\det(dF)_p = \det(d\bar{f})_p \neq 0$. By the Inverse Function Theorem, F is a diffeomorphism near p. Let $g : \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^m$ be the projection onto $\mathbb{R}^m$. Then on a neighborhood of p we have $g \cdot F(x, y) = g(f(x, y), y) = f(x, y)$. Thus f is linearized to g by a change of coordinates on the domain. $\square$

We can combine these last two results to give a complete characterization of the local structure of smooth maps at points of full rank.

Theorem 1.2.5 (Linearization Theorem). *Let $f : X \to Y$ be a smooth map of manifolds. Let $p \in X$ be a point such that $(df)_p$ has full rank. Then f is locally equivalent to any linear map of full rank at p.*

Such a point p is called a regular point. We have established that the local structure of smooth functions at regular points is as simple as we could have hoped. A natural question to ask is: Just how common are regular points? It turns out that this is not the “right” question to ask. In the next chapter, we will both formulate the “right” question and provide a very precise answer with Sard’s Theorem.

Chapter 2

Power: The Theorem of Sard

In §2.1, we show that smooth maps pull regular points back to submanifolds. In §2.2, we prove Sard's Theorem.

2.1 Critical Values and Regular Values

Definition 2.1.1. Let $f : X \to Y$ be a smooth map of manifolds.

(1) $p \in X$ is a *critical point* of f if corank $(df_p) > 0$.

(2) $p \in X$ is a *regular point* of f if corank $(df_p) = 0$.

(3) $q \in Y$ is a *critical value* of f if there exists a critical point p of f with $q = f(p)$.

(4) $q \in Y$ is a *regular value* of f if q is not a critical value of f.

Note that the regular values of f include all points of $Y - f(X)$.

We ended §1.2 with a question: How common are regular points? Our intuition from calculus suggests that most smooth maps have many regular points and few critical points. Still, there exists a very simple map for which every point is critical, namely a constant map. The key observation is that while a constant map has many critical points, it has only one critical value. Near a critical point, the degeneration of the derivative limits the range of the map in a way made precise by Taylor's Theorem. Informally, Sard's theorem states that the set of critical values of a smooth map is small. We now make precise this notion of "small".

Definition 2.1.2. (1) Let S be a subset of $\mathbb{R}^n$. Then S has *measure zero* if for every $\varepsilon > 0$, there exists a cover of S by a countable number of open cubes $C_1, C_2, \ldots$ such that $\sum_{i=1}^{\infty} \mathrm{vol}[C_i] < \varepsilon$.

(2) Let X be a smooth manifold and S a subset of X. Then S is of *measure zero* if there exists a countable open cover $U_1, U_2, \ldots$ of S and charts $\phi_i : U_i \to \mathbb{R}^n$ such that $\phi_i(U_i \bigcap S)$ has measure zero in $\mathbb{R}^n$.

Sard's Theorem says that the set of critical values of a smooth map of manifolds has measure zero in Y. We will prove Sard's Theorem in the next section. For now we establish the most useful property of regular values.

Theorem 2.1.1. *Let $f : X \to Y$ be a smooth map of manifolds, $q \in Y$ a regular value of f. Then $f^{-1}(q)$ is a smooth submanifold of X with* $\dim f^{-1}(q) = \dim X - \min\{\dim X, \dim Y\}$.

Proof. Let $q \in Y$ be a regular value of f, $W = f^{-1}(q)$, $n = \dim X$, and $m = \dim Y$. Suppose $p \in W$.

If $n \leq m$ then f is an immersion at p. By the Immersion Lemma, we can choose local coordinates on X and Y so that $p = q = 0$ and $f(x_1, \ldots, x_n) = (x_1, \ldots, x_n, 0, \ldots, 0)$. Thus p is an isolated point of $W \subset X$ (i.e., there exists a nbhd of p containing no other point of W). Since the topology of X is second countable, we conclude that W consists of countable number of isolated points and is therefore a smooth submanifold of dimension 0.

If $n \geq m$ then f is a submersion at p. By the Submersion Lemma, we can choose local coordinates on X and Y so that $p = q = 0$ and $f(x_1, ..., x_n) = (x_1, ..., x_m)$. Then locally $W = \{(0, ..., 0, x_{m+1}, ..., x_n)\}$ so $x_{m+1}, ...x_n$ serve as local coordinates for W near p, mapping a nbhd of p in W (under the subspace topology) diffeomorphically onto the subspace $\mathbb{R}^{n-m} \subset \mathbb{R}^n$. $\square$

Remark 2.1.1. If q is not in the image of f then $W = \emptyset$, which vacuously satisfies the definition of a manifold of any dimension.

Theorem 2.1.1 provides a useful tool for generating manifolds. For example, we can realize the n-sphere as a submanifold of $\mathbb{R}^{n+1}$ by considering the preimage of 0 under the map $(x_1, ..., x_{n+1}) \mapsto 1 - (x_1^2 + \cdots + x_{n+1}^2)$.

The same method used to prove Theorem 2.1.1 yields a related result that we will need in §3.1.

Proposition 2.1.2. *Let $f : X \to Y$ be a map of smooth manifolds, and W a submanifold of Y. If f is a submersion at each point in $f^{-1}(W)$, then $f^{-1}(W)$ is a submanifold of X with* $\operatorname{codim} f^{-1}(W) = \operatorname{codim} W$.

Proof. Let $l = \dim W$. Suppose $p \in X$ with $f(p) = q$. First choose local coordinates $y_1, ..., y_m$ on Y so that, near q, W corresponds to the first l coordinates. Next choose local coordinates $x_1, ..., x_n$ on X to linearize f near p to $f(x_1, ..., x_n) = (x_1, ..., x_m)$. Then locally $f^{-1}(W) = \{(x_1, ..., x_l, 0, ..., 0, x_{m+1}, ..., x_n)\}$ so $x_1, ..., x_l, x_{m+1}, ...x_n$ serve as local coordinates for $f^{-1}(W)$ near p, mapping a nbhd of p in $f^{-1}(W)$ (under the subspace topology) diffeomorphically onto the subspace $\mathbb{R}^{n-(m-l)} \subset \mathbb{R}^n$. $\square$

2.2 Proof of Sard's Theorem

Theorem 2.2.1 (Sard). *Let $f : X \to Y$ be a smooth map of manifolds. Then the set of critical values of f has measure zero in Y.*

It suffices to prove the following:

Proposition 2.2.2. *Let U be an open subset of $\mathbb{R}^n$, and let $f : U \to \mathbb{R}^m$ be a smooth map. Then the set of critical values of f has measure zero in $\mathbb{R}^m$.*

Proof of Theorem from the Proposition. Let C denote the set of critical points of f. Let $\{U_i\}$ be a countable open cover of X with the property that each $f(U_i)$ is contained in some coordinate nbhd V_i of Y with chart $\varphi_i : V_i \to \mathbb{R}^m$. By Proposition 2.2.2, the set of critical values of $\varphi_i \cdot f|_{U_i}$ has measure zero in $\mathbb{R}^m$. So $\varphi_i(V_i \bigcap f(C))$ has measure zero in $\mathbb{R}^m$. Therefore $f(C)$ is of measure zero in Y. □

In order to carry out the induction step in our proof of the Proposition 2.2.2, we will need the the following result from measure theory [15, p.51]:

Theorem 2.2.3 (Fubini). *Let S be a subset of $\mathbb{R}_n = \mathbb{R} \times \mathbb{R}^{n-1}$ such that the intersection of S with each hyperplane $x \times \mathbb{R}^{n-1}$ has measure zero in $\mathbb{R}^{n-1}$. Then S has measure zero in $\mathbb{R}^n$.*

We will also need the elementary fact that the countable union of measure zero sets has measure zero. Our proof follows that of Milnor [11, p.16].

Proof of Proposition 2.2.2. We proceed by induction on n. The case $n = 0$ holds trivially. Assume true for $n - 1$. Let C be the set of critical points of f. For each $i \geq 1$, let C_i be the set of points $x \in X$ such that all of the partial derivatives of f of order $\leq i$ vanish at x. Clearly the C_i form a descending chain

$$C \supset C_1 \supset C_2 \supset C_3 \cdots.$$

We divide the proof into parts:

Part A: $f(C - C_1)$ has measure zero.

Part B: $f(C_i - C_{i+1})$ has measure zero for every $i \geq 1$.

Part C: $f(C_k)$ has measure zero for some k.

Assuming the above, $f(C) = f(C - C_1) \bigcup \left(\bigcup_{i=1}^{k-1} f(C_i - C_{i+1}) \right) \bigcup f(C_k)$ is a countable union of measure zero sets. Therefore $f(C)$ has measure zero.

Proof of Part A: For each $\bar{x} \in (C - C_1)$, we will show that there is an open nbhd V of $\bar{x}$ such that $f(V \bigcap C)$ has measure zero. A countable number of such nbhds will cover $\mathbb{R}^n$, so we conclude that $f(C - C_1)$ is a countable union of measure zero sets.

So let $\bar{x} \in (C - C_1)$. Then the first order partial derivatives of f at $\bar{x}$ are not all zero, so we may assume without loss of generality that $\frac{\partial f_1}{\partial x_1}(\bar{x}) \neq 0$. Define the map $h : U \to \mathbb{R}^n$ by

$$h(x_1, ..., x_n) = (f_1(x_1, ..., x_n), x_2, ..., x_n).$$

By construction $(dh)_{\bar{x}}$ is invertible. By the Inverse Function Theorem, there exists an open nbhd $V \in U$ of $\bar{x}$ such that $h|_V$ is a diffeomorphism onto its image W. Now consider the map $g = f \cdot h^{-1} : W \to \mathbb{R}^m$. Let $C' \subset W$ be the set of critical points of g. Since h^{-1} is a diffeomorphism, $h^{-1}(C') = V \bigcap C$ and thus $g(C') = f(V \bigcap C)$. So it suffices to show that $g(C')$ has measure zero. Since $h^{-1} = (f_1^{-1}, ...)$, the first component function of $g = f \cdot h^{-1}$ is the identity map. Thus, for fixed $x_1 \in \mathbb{R}$, g maps the hyperplane $\{x_1\} \times \mathbb{R}^{n-1}$ into the hyperplane $\{x_1\} \times \mathbb{R}^{m-1}$. By induction, $g(C' \bigcap \{x_1\} \times \mathbb{R}^{n-1})$ has measure zero in $\{x_1\} \times \mathbb{R}^{m-1}$ for every $x_1 \in \mathbb{R}$. Therefore by Fubini's Theorem, $g(C')$ has measure zero.

Proof of Part B: Again it suffices to show that, for each $\bar{x} \in (C_i - C_{i+1})$, there is an open nbhd V of $\bar{x}$ such that $f(V \bigcap C)$ has measure zero.

Let $\bar{x} \in (C_i - C_{i+1})$. Then $\frac{\partial^{i+1} f_r}{\partial x_{s_1} \partial x_{s_2} \cdots \partial x_{s_{i+1}}}(\bar{x}) \neq 0$ for some $1 \leq s_1, ..., s_{i+1} \leq n$ and the map

$$w(x) = \frac{\partial^{i+1} f_r}{\partial x_{s_2} \cdots \partial x_{s_{i+1}}}(\bar{x})$$

vanishes on C_i but $\frac{\partial w}{\partial x_{s_1}} \neq 0$. We may assume $s_1 = 1$. Define $h : U \to \mathbb{R}^n$ by

$$h(x_1, ..., x_n) = (w_1(x_1, ..., x_n), x_2, ..., x_n).$$

Note that h carries C_i to the hyperplane $0 \times \mathbb{R}^{n-1}$. As before $(dh)_{\bar{x}}$ is invertible and there exists an open nbhd $V \in U$ of $\overline{x}$ such that $h|_V$ is a diffeomorphism onto its image W. Again consider the map $g = f \cdot h^{-1} : W \to \mathbb{R}^m$. By induction, the set of critical values of the restriction

$$\overline{g} : 0 \times \mathbb{R}^{n-1} \bigcap W \to \mathbb{R}^m$$

has measure zero in $\mathbb{R}^m$. But since h carries C_i to $0 \times \mathbb{R}^{n-1}$, we have $(C_i \bigcap V) \subset h^{-1}(0 \times \mathbb{R}^{n-1})$. Therefore the set of critical values of $\overline{g}$ contains $f(C_i \bigcap V)$, so $f(C_i \bigcap V)$ has measure zero as well.

Proof of Part C: Let $I^n \subset U$ be a cube of edge length δ. We will show that for k sufficiently large, $f(C_k \bigcap I^n)$ has measure zero in $\mathbb{R}^m$. Then, since we can cover C_k with a countable number of such cubes, $f(C_k)$ has measure zero as well.

By Taylor's Theorem, the compactness of I^n, and the definition of C_k, we see that

$$f(x + h) = f(x) + R(x, h)$$

where

$$\|R(x, h)\| \leq c \|h\|^{k+1}$$

for $x \in C_k \bigcap I^n$ and $x+h \in I^n$. Here $c \in \mathbb{R}$ depends only on f and I^n. Now subdivide I^n into r^n cubes of edge length $\frac{\delta}{r}$. Suppose I_1 is a sub-cube containing a point $x \in C_k$. It follows from induction on the Pythagorean Theorem that the diameter of I_1 is $\sqrt{n}\frac{\delta}{r}$. Therefore any point of I_1 can be expressed as $x+h$ with $\|h\| \leq \sqrt{n}\frac{\delta}{r}$. From the Taylor formula and the triangle inequality, it follows that $f(I_1)$ sits inside a cube with edge length $\frac{a}{r^{k+1}}$, where $a = 2(\sqrt{n}\delta)^{k+1}$. Thus $f(C_k \bigcap I^n)$ is contained in a union of at most r^n sub-cubes with total volume

$$V \leq r^n(\tfrac{a}{r^{k+1}})^m = a^m r^{n-(k+1)m}.$$

Hence if $k > \frac{n}{m} - 1$, we can cover $f(C_k \bigcap I^n)$ with a countable number of cubes of arbitrarily small total volume by choosing r to be sufficiently large (i.e., by taking a sufficiently fine partition of I^n into sub-cubes). Therefore $f(C_k \bigcap I^n)$ has measure zero. $\square$

It is a basic result of measure theory that a measure zero subset of $\mathbb{R}^n$ cannot contain a non-empty open set. This fact clearly extends to manifolds, so we have...

Corollary 2.2.4. *Let $f : X \to Y$ be a smooth map of manifolds. Then the regular values of f form a dense subset of Y.*

Chapter 3

Language: Transversality

In §3.1, we define the basic notions of the theory of transversality and show that transverse maps pull submanifolds back to submanifolds. In §3.2, we define the jet bundle and use it to topologize $\mathcal{C}^\infty(X, Y)$ in §3.3. In §3.4, we prove the Thom Transversality Theorem and generalize to the Multijet Transversality Theorem in §3.5.

3.1 The Basics

Transversality can be viewed as a far-reaching generalization of the notion of regular value in §2. Informally, two manifolds are transverse if their intersection occurs in the most general possible form. This is best conveyed through a series of images (see the next page). Notice that the intersections in each of the the non-transverse images can be remove or made transverse with only a slight adjustment to one of the manifolds. By contrast, in each of the images labeled transverse, the intersection cannot be easily removed. In this sense, transverse intersections are stable.

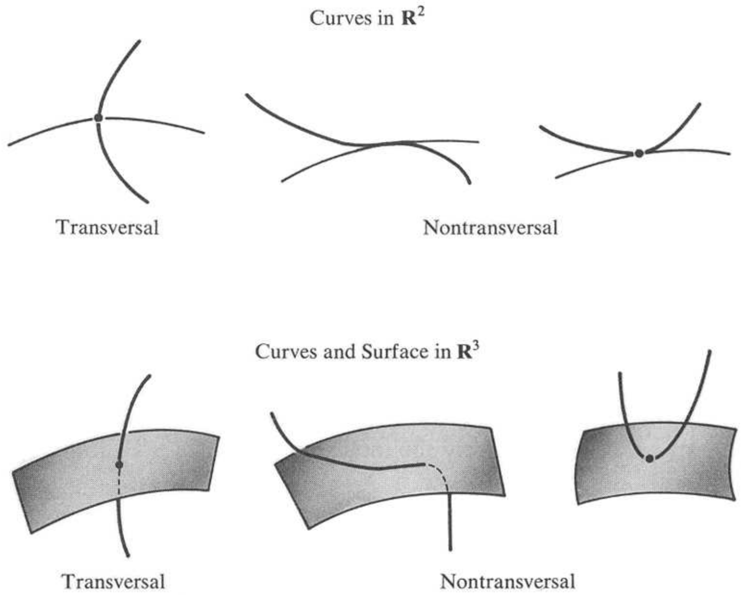

Figure 3.1: [6, p.30]

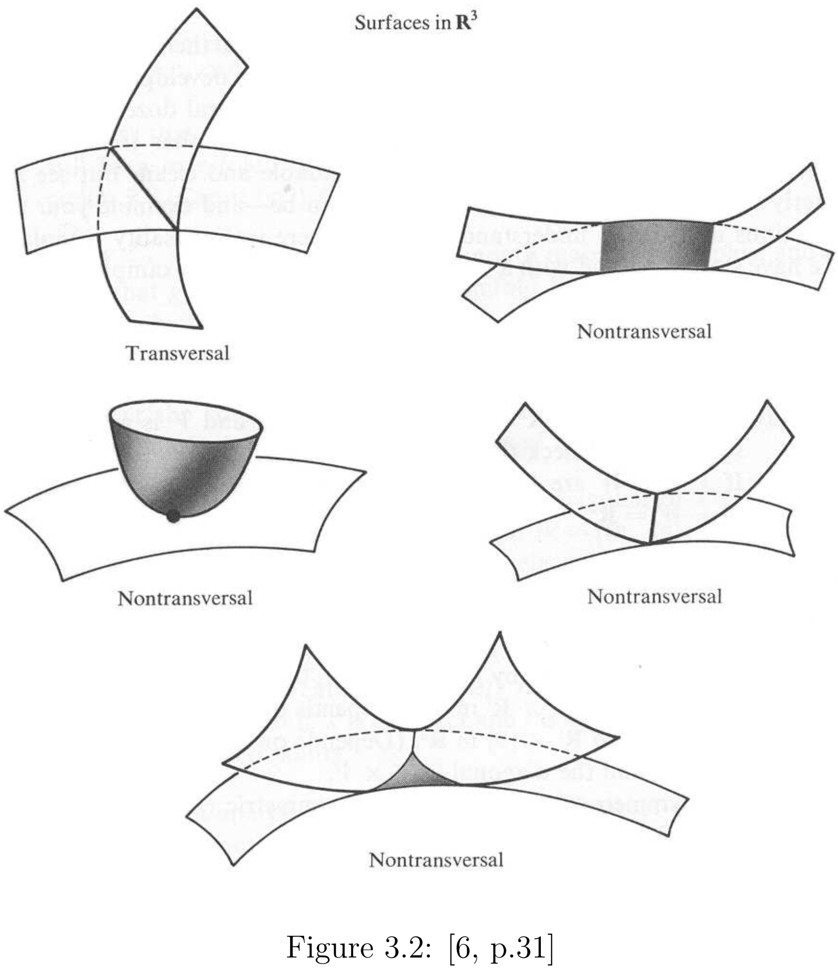

Figure 3.2: [6, p.31]

We can neatly define transversality of manifolds using tangent spaces.

Definition 3.1.1. Let X be a smooth manifold, W and Z submanifolds. Then W and Z are *transverse* at $p \in X$, denoted $W \pitchfork_p Z$, if either:

a) $p \notin W \bigcap Z$

b) $p \in W \bigcap Z$ and $T_pX = T_pW + T_pZ$.

W and Z are transverse, denoted $W \pitchfork Z$, if $W \pitchfork_p Z$ for every $p \in X$.

The reader is encouraged to review the figures with this definition in mind. In order to study the local structure of smooth maps, we must also define what it means for a smooth map to be transverse to a submanifold.

Definition 3.1.2. Let $f : X \to Y$ be a smooth map of manifolds, W a submanifold of Y. Then f is *transverse* to W at p, denoted $f \pitchfork_p W$, if either:

(a) $f(p) \notin W$

(b) $f(p) \in W$ and $T_{f(p)}Y = (df)_p(T_pX) + T_{f(p)}W$

f is *transverse* to W, denoted $f \pitchfork W$, if $f \pitchfork_p W$ for every $p \in X$. We may also write $f \pitchfork W$ on $U \subset X$ to convey that $f \pitchfork_p W$ for every $p \in U$.

Remark 3.1.1. (1) Suppose that $\dim X \geq \dim Y$ and that W consists of a single point $q \in Y$. Then $f \pitchfork W$ iff q is a regular value of f.

(2) It also follows immediately that submersions are transverse to every submanifold.

(3) Some texts define $f \pitchfork W$ iff graph$(f) \pitchfork X \times W$ [4, p.39]. The reader should verify that these two definitions are equivalent.

Imagine a smooth curve in $\mathbb{R}^2$ intersecting itself transversely at a point. With only two dimensions in which to move, it is impossible to remove this intersection through an arbitrarily small perturbation. However, if we now embed the curve in $\mathbb{R}^3$, we can remove the intersection with ease:

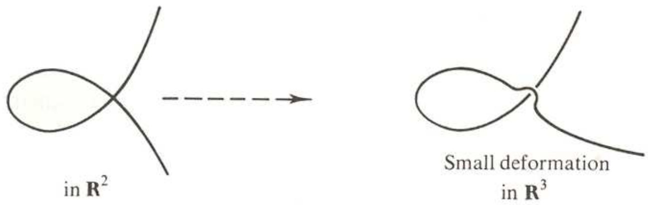

Figure 3.3: [6, p.50]

In $\mathbb{R}^3$ the intersection is not stable, so it cannot be transverse. Transverse curves in $\mathbb{R}^3$ must not intersect at all, and the next proposition generalizes this intuition to higher dimensional manifolds.

Proposition 3.1.1. *Let* $f : X \to Y$ *be a map of smooth manifolds,* W *a submanifold of* Y*. Suppose that* $\dim X + \dim W < \dim Y$*. Then* $f \pitchfork W$ *iff the image of* f *is disjoint from* W*.*

Proof. The latter condition clearly implies the former.

Now suppose there exists $p \in X$ with $f(p) \in W$. Then

$$\begin{aligned} \dim[(df)_p(T_pX) + T_{f(p)}W] &\leq \dim(df)_p(T_pX) + \dim T_{f(p)}W \\ &\leq \dim X + \dim W \\ &< \dim Y = \dim T_{f(p)}Y \end{aligned}$$

Therefore f is not transverse to W at p. □

In order to prove the main result of this section, we will need the following lemma.

Lemma 3.1.2. *Let $f : X \to Y$ be a map of smooth manifolds, W a submanifold of Y of codimension k, and $p \in X$ with $f(p) \in W$. Let U be a nbhd of $f(p)$ and $\phi : U \to \mathbb{R}^k$ a submersion such that $W \bigcap U = \phi^{-1}(0)$. Then $f \pitchfork W$ at p iff $\phi \cdot f$ is a submersion at p.*

Proof. Since ϕ is a submersion at $f(p)$, $\dim \operatorname{im}((d\phi)_{f(p)}) = k = \operatorname{codim} W$. Thus $\dim \ker(d\phi)_{f(p)} = \dim W = \dim T_{f(p)}W$. And $T_{f(p)}W \subset \ker(d\phi)_{f(p)}$ because ϕ is constant on W. Therefore $\ker(d\phi)_{f(p)} = T_{f(p)}W$.

Now $\phi \cdot f$ is a submersion at p iff $(d\phi \cdot f)_p$ is onto. Since $(d\phi)_{f(p)}$ is onto, the latter condition holds

iff $(df)_p(T_pX) + \ker((d\phi)_{f(p)}) = T_{f(p)}Y$

iff $(df)_p(T_pX) + T_{f(p)}W = T_{f(p)}Y$

iff $f \pitchfork W$ at p. □

By Remark 3.1.1, we can view the next theorem as a generalization of Theorem 2.1.1.

Theorem 3.1.3. *Let $f : X \to Y$ be a map of smooth manifolds, W a submanifold of Y. If $f \pitchfork W$ then $f^{-1}(W)$ is a submanifold of X with* $\operatorname{codim} f^{-1}(W) = \operatorname{codim} W$.

Proof. By the definition of submanifold, it is sufficient to show that every $p \in f^{-1}(W)$ has a neighborhood $V \subset X$ such that $f^{-1}(W) \bigcap V$ is a submanifold of X. Suppose $p \in X$ with $f(p) = q$. Since W is a submanifold, there is a neighborhood $U \subset Y$ of q and a submersion $\phi : U \to \mathbb{R}^k$ (with $k = \dim W$) such that $W \bigcap U = \phi^{-1}(0)$. By Lemma 3.1.2, $f \pitchfork W$ at p implies that $\phi \cdot f$ is a submersion at p. Therefore, by the Submersion Lemma, we can choose a sufficiently small neighborhood $V \subset X$ of p such that $f(V) \subset U$ and $\phi \cdot f$ a submersion on V. Near p we have $f^{-1}(W) \bigcap V = (\phi \cdot (f|_V))^{-1}(0)$ which is a submanifold of the stated dimension by Proposition 2.1.2. □

As an immediate corollary, we see that the intersection of transverse manifolds is itself a manifold.

Corollary 3.1.4. *Let X be a smooth manifold, W and Z submanifolds. If $W \pitchfork Z$ then $W \bigcap Z$ is a submanifold of X with* $\operatorname{codim}(W \bigcap Z) = \operatorname{codim} W + \operatorname{codim} Z$.

Proof. Let $i : Z \to X$ be the inclusion map. Then $(di)_p(T_pZ) = T_pZ$, so $W \pitchfork Z$ implies $i \pitchfork Z$. $W \bigcap Z = i^{-1}(Z)$ so we are done by Theorem 3.1.3. □

We can make any two manifolds transverse through an arbitrarily small perturbation of one of them. The following lemma formalizes this intuition and is the key to proving the more powerful transversality theorems of §3.4 and §3.5. Note how its proof translates the power of Sard's Theorem into a statement about transversality.

Lemma 3.1.5 (Transversality Lemma). *Let X, B, and Y be smooth manifolds, W a submanifold of Y. Let $j : B \to \mathcal{C}^\infty(X, Y)$ be a set map and define $\Phi : X \times B \to Y$ by $\Phi(x, b) = j(b)(x)$. If $\Phi \pitchfork W$ then the set $\{b \in B \mid j(b) \pitchfork W\}$ is dense in B.*

Proof. Let $W_\Phi = \Phi^{-1}(W)$. Since $\Phi \pitchfork W$, W_Φ is a submanifold of $X \times B$. Let $\pi : W_\Phi \to B$ be the restriction of the projection $X \times B \to B$. If $b \notin \pi(W_\Phi)$, then $j(b)(x) \notin W$ for every $x \in X$, so $j(b)(X) \bigcap W = \emptyset$ and $j(b) \pitchfork W$.

If $\dim W_\Phi < \dim B$, then $\pi(W_\Phi)$ has measure zero in B by Corollary 2.2.4. Therefore $j(b) \pitchfork W$ for every b in the dense set $B - \pi(W_\Phi)$.

For the case $\dim W_\Phi \geq \dim B$, it suffices to prove the following claim and apply Corollary 2.2.4.

Claim: If b is a regular value of π, then $j(b) \pitchfork W$.

Proof of Claim: Let b be a regular value of π and $x \in X$. We will show that $j(b) \pitchfork_x W$. If $(x, b) \notin W_\Phi$, then $j(b)(x) \notin W$, so $j(b) \pitchfork_x W$. So assume $(x, b) \in W_\Phi$. Since b

is a regular value of π and $\dim W_\Phi \geq \dim B$, we have that $(d\pi)_{(x,b)}(T_{(x,b)}W_\Phi) = T_bB$, or that $T_{(x,b)}W_\Phi$ contains $T_{(x,b)}(\{x\} \times B)$ as subsets of $T_{(x,b)}X \times B$. Thus

$$T_{(x,b)}(X \times B) = T_{(x,b)}W_\Phi + T_{(x,b)}(X \times \{b\})$$

Applying $(d\Phi)_{(x,b)}$ to both sides gives

$$(d\Phi)_{(x,b)}(T_{(x,b)}(X \times B)) = T_{j(b)(x))}W + (dj(b))_x(T_xX)$$

By assumption $\Phi \pitchfork W$ so

$$T_{\Phi(x,b)}Y = T_{\Phi(x,b)}W + (d\Phi)_{(x,b)}(T_{(x,b)}X \times B)$$

Combining gives

$$T_{j(b)(x)}Y = T_{j(b)(x)}W + (dj(b))_x(T_xX)$$

Therefore $j(b) \pitchfork W$ at x. $\square$

As our first application of the Transversality Lemma, we show that we can make a map transverse to a submanifold with an arbitrarily small translation.

Proposition 3.1.6. *Let $f : X \to \mathbb{R}^m$ be a smooth map of manifolds, W a submanifold of $\mathbb{R}^m$. Then $\{b \in \mathbb{R}^m \mid (f + b) \pitchfork W\}$ is a dense subset of $\mathbb{R}^m$.*

Proof. In the notation of the Transversality Lemma, let $B = Y = \mathbb{R}^m$ and define $j : \mathbb{R}^n \to \mathcal{C}^\infty(X, \mathbb{R}^n)$ by $j(b) = f + b$. Then $\Phi : X \times B \to B$ defined by $\Phi(x, b) = j(b)(x) = f(x) + b$ is clearly a submersion. Therefore $\Phi \pitchfork W$ and we are done by Proposition 2.1.2. $\square$

Once we have placed a topology on $\mathcal{C}^\infty(X, Y)$, we will be able to make further use of the Transversality Lemma to prove more powerful claims about the set of maps transverse to a submanifold. We develop this topology in the next two sections.

3.2 Jet Bundles

We can define the jet bundles of smooth maps of manifolds with coordinates or abstractly. In [13], Saunders describes the first jet bundle in terms of local coordinates in order to best study first order differential equations, which can be interpreted as closed embedded submanifolds of this bundle [13, p.103]. However, as one progresses to higher order jet bundles, the coordinate-based approach becomes decidedly unwieldy. For our purposes, it makes sense to follow the invariant treatment of jet bundles in [5]. Advantages include notational simplicity, increased utility, and greater overall beauty. Though the definition may appear strange at first, we will soon descend to local coordinates in order to develop an intuition for this new structure.

Definition 3.2.1. Let $f, g : X \to Y$ be smooth maps of manifolds. Let $p \in X$ with $f(p) = g(p) = q$.

(1) f has *first order contact* with g at p if $(df)_p = (dg)_p$ as maps of $T_pX \to T_qY$.

(2) f has *kth order contact* with g at p if $(df) : TX \to TY$ has $(k-1)$th order contact with (dg) at every point in T_pX. We denote this by $f \sim_k g$.

(3) Let $J^k(X,Y)_{p,q}$ denote the set of equivalence classes under "$\sim_k$ at p" of maps $f : X \to Y$ with $f(p) = q$.

(4) Let $J^k(X,Y) = \bigsqcup_{(p,q)\in X\times Y} J^k(X,Y)_{p,q}$. $J^k(X,Y)$ is the *kth jet bundle* of maps from $X \to Y$ and an element σ of $J^k(X,Y)$ is called a *k-jet* of maps from $X \to Y$.

(5) Let $\sigma \in J^k(X,Y)$. Then there exist unique $p \in X$ and $q \in Y$ with $\sigma \in J^k(X,Y)_{p,q}$. p is called the *source* of σ and q the *target* of σ. The map $\alpha : J^k(X,Y) \to X$ defined by $\sigma \mapsto$ (source of σ) is called the *source map*. The map $\beta : J^k(X,Y) \to Y$ defined by $\sigma \mapsto$ (target of σ) is called the *target map*.

To each smooth map $f : X \to Y$ we can assign a map $j^k f : X \to J^k(X,Y)$ sending a point x to the k-jet with source p by f, also called the k-jet of f at p. Note that $J^0(X,Y) \cong X \times Y$ and $j^0 f(X)$ is simply the graph of f.

The k-jet of f at p is just an invariant way of describing the Taylor expansion of f at p to kth order:

Proposition 3.2.1. *Let U be an open neighborhood of p in $\mathbb{R}^n$, $f, g : U \to \mathbb{R}^m$ smooth mappings, f_i and g_i the coordinate functions of f and g, respectively, and $x_1, ..., x_n$ coordinates on U. Then $f \sim_k g$ at p iff*

$$\tfrac{\partial^{|\alpha|} f_i}{\partial x^\alpha}(p) = \tfrac{\partial^{|\alpha|} g_i}{\partial x^\alpha}(p)$$

for every multi-index α with $|\alpha| \leq k$ and $1 \leq i \leq m$.

Proof. We proceed by induction on k. For $k = 1$, we have $f \sim_1 g$ at p iff $(df)_p = (dg)_p$ iff $\frac{\partial f_i}{\partial x_j}(p) = \frac{\partial g_i}{\partial x_j}(p)$ for every $1 \leq i, j \leq n$.

Assume the proposition is true for $k-1$. Let $y_1, y_2, ...$ be the coordinates of $\mathbb{R}^n$ in $U \times \mathbb{R}^n = TU$. Then $(df) : U \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^m = T\mathbb{R}^m$ is given by

$$(x, y) \mapsto (f(x), \overline{f}_1(x, y), ..., \overline{f}_m(x, y))$$

where

$$\overline{f}_i(x, y) = \textstyle\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) y_j$$

And similarly for (dg).

If $f \sim_k g$ at p then $(df)_p \sim_{k-1} (dg)_p$ on T_pX. By the induction hypothesis, (df) and (dg) have the equal partial derivatives to $(k-1)$st order at each $(p, v) \in p \times \mathbb{R}^n$. So for each multi-index α with $|\alpha| \leq k-1$, we have

$$\tfrac{\partial^{|\alpha|} \overline{f}_i}{\partial x^\alpha}(p, v) = \tfrac{\partial^{|\alpha|} \overline{g}_i}{\partial x^\alpha}(p, v)$$

Evaluating at $v = (0, ..., 1, ...0)$ (1 is the jth position) gives

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha}\frac{\partial f_i}{\partial x_j}(p) = \frac{\partial^{|\alpha|}}{\partial x^\alpha}\frac{\partial g_i}{x_j}(p)$$

All partial derivatives of f and g order $\leq k$ are obtained this way.

Conversely, suppose that f and g have equal partial derivatives of order $\leq k$ at p. Then (df) and (dg) have equal partial derivatives of order $\leq k-1$ at p, so by the induction hypothesis we have $(df)_p \sim_{k-1} (dg)_p$. Therefore $f \sim_k g$ at p. □

Corollary 3.2.2. *$f \sim_k g$ at p iff the Taylor expansions of f and g up to (and including) order k are identical at p*

The composition of maps corresponds to the composition of jets in a nice way:

Lemma 3.2.3. *Let U be an open subset of $\mathbb{R}^n$ and V an open subset of $\mathbb{R}^m$. Let $f_1, f_2 : U \to V$ and $g_1, g_2 : V \to \mathbb{R}^l$ be smooth maps. If $f_1 \sim_k f_2$ at p and $g_1 \sim_k g_2$ at $f(p)$, then $g_1 \cdot f_1 \sim_k g_2 \cdot f_2$ at p.*

Proof. We proceed by induction on k. For $k = 1$, this is the chain rule:

$$d(g_1 \cdot f_1)_p = (dg_1)_{f(p)} \cdot (df_1)_p = (dg_2)_{f(p)} \cdot (df_2)_p = d(g_2 \cdot f_2)_p$$

Assume true for $k-1$. By hypothesis, $(df_1) \sim_{k-1} (df_2)$ on T_pU and $(dg_1) \sim_{k-1} (dg_2)$ on $T_{f(p)}V$. Applying the induction hypothesis point-wise, we have $d(g_1 \cdot f_1) \sim_{k-1} d(g_2 \cdot f_2)$ on T_pU. Therefore $g_1 \cdot f_1 \sim_k g_2 \cdot f_2$ at p. □

Proposition 3.2.1 and Lemma 3.2.3 are enough to prove that $J^k(X, Y)$ is a manifold. The approach is a merely an elaboration on that taken with the tangent bundle. First, we pull back charts on $X \times Y$ to charts on $J^k(X, Y)$ via the map $\alpha \times \beta$. Then we verify that these latter charts are compatible, forming a smooth atlas.

We first need to establish some notation. Let A_n^k be vector space of polynomials in n variables of degree $\leq k$ which have their constant term equal to zero. Choose as coordinates for A_n^k the coefficients of the polynomials. Let $B_{n,m}^k = \bigoplus_{i=1}^m A_n^k$ and define the map $\tau_{n,m}^k : \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^m) \to B_{n,m}^k$ to send each smooth map $f = (f_1, ..., f_m)$ to the coefficients (except the constant term) of the kth-order Taylor polynomial of each f_i in the obvious way.

We can now describe the aforementioned chart (W, η) of $J^k(X, Y)$ coming from the charts (U, φ) of X and (V, ψ) of Y. We set $W = (\alpha \times \beta)^{-1}(U \times V)$ and define the map $\eta : W \to \varphi(U) \times \psi(V) \times B_{n,m}^k$ by

$$\eta(\sigma) = \big(\varphi(\alpha(\sigma)), \psi(\beta(\sigma)), \tau^k(\psi \cdot f \cdot \varphi^{-1})\big)$$

where f represents σ. Note that η is well-defined by Corollary 3.2.2.

To verify that such (W, η) form a smooth atlas on $J^k(X, Y)$, we must check that these charts are compatible. Essentially, this reduces to the claim that a change of basis on $\mathbb{R}^n$ smoothly sends Taylor series in the first basis to Taylor series in the second basis. We now state our conclusions and refer the unconvinced reader to [5, p.37].

Theorem 3.2.4. *Let X and Y be smooth manifolds with $n = \dim X$ and $m = \dim Y$. Then*

(1) $J^k(X, Y)$ is a smooth manifold with

$$\dim J^k(X, Y) = n + m + \dim(B_{n,m}^k)$$

(2) $\alpha : J^k(X, Y) \to X$, $\beta : J^k(X, Y) \to X$, and $\alpha \times \beta : J^k(X, Y) \to X \times Y$ are smooth submersions.

(3) If $f : X \to Y$ is smooth, then $j^k f : X \to J^k(X, Y)$ is smooth.

Since $J^k(X,Y)$ is a manifold, we can adapt the Transversality Lemma to a statement about jets.

Proposition 3.2.5. *Let X, B, and Y be smooth manifolds, W a submanifold of $J^k(X,Y)$. Let $G : X \times B \to Y$ be a smooth map and define $\Phi : X \times B \to J^k(X,Y)$ by $\Phi(x,b) = j^k G_b(x)$. If $\Phi \pitchfork W$ then $\{b \in B \mid j^k G_b \pitchfork W\}$ is a dense subset of B.*

Proof. If G is smooth than Φ is smooth by part (3) of Theorem 3.2.4. Define $j : B \to J^k(X,Y)$ by $j(b) = j^k G_b$. Then $\Phi(x,b) = j(b)(x)$ and we are done by the Transversality Lemma. $\square$

In §3.3, we will use the topology of $J^k(X,Y)$ to define a topology on $\mathcal{C}^\infty(X,Y)$. In §3.4, we will use Proposition 3.2.5 to prove the Thom Transversality Theorem. The remainder of this section is devoted to further study of the differential structure on $J^k(X,Y)$. In particular, we establish a collection of submanifolds of the jet bundle that will be quite valuable in §4, §5, and §6.

Definition 3.2.2. Let $\sigma \in J^1(X,Y)$ with source p. Let f represent σ.

(1) The *rank* of σ is defined to be the rank of $(df)_p$.

(2) The *corank* of σ is defined to be the corank of $(df)_p$.

(3) The *singularity set of corank* r is defined to be

$$\mathcal{S}_r = \{\sigma \in J^1(X,Y) \mid \text{corank } (\sigma) = r\}$$

Note that if f and g both represent σ, then $f \sim_1 g$ and thus $(df)_p = (dg)_p$. Therefore the rank and corank of 1-jets are well-defined. In order to prove that $\mathcal{S}_r$ is a manifold, we must first do some linear algebra.

Let V and W be real vector spaces with $n = \dim V$, $m = \dim W$, and $q = \min\{n,m\}$. Define $L^r(V,W) = \{S \in \mathrm{Hom}(X,Y) \mid \text{corank } S = r\}$.

Proposition 3.2.6. *$L^r(V,W)$ is a submanifold of* $\mathrm{Hom}(V,W)$ *with* $\operatorname{codim} L^r(X,Y) = (n-q+r)(m-q+r)$.

Proof. Let $S \in L^r(V,W)$ and $k = q - r = \operatorname{rank} S$. Choose bases on V and W such that S takes the form $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with A a $k \times k$ invertible matrix. Let U be a coordinate nbhd of S in $\mathrm{Hom}(X,Y)$ such that if $S' = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \in U$ then A' is invertible. We can choose such a U because the determinant map is continuous. Define $\phi : U \to \mathrm{Hom}(\mathbb{R}^{n-k}, \mathbb{R}^{m-k})$ by $\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \mapsto D' - C'A'^{-1}B$. Fixing A', B',and C' restricts ϕ to a translation. Therefore ϕ is a submersion. The following lemma implies that $L^r(V,W) \bigcap U = \phi^{-1}(0)$. Then by Theorem 2.1.1, $L^r(V,W)$ is a manifold and $\operatorname{codim} L^r(V,W) = \dim \mathrm{Hom}(\mathbb{R}^{n-k}, \mathbb{R}^{m-k}) = (n-k)(m-k)$. $\square$

Lemma 3.2.7. *Let $S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ be an $m \times n$ matrix with A a $k \times k$ invertible matrix. Then* $\operatorname{rank} S = k$ *iff $D - CA^{-1}B = 0$.*

Proof. The matrix

$$T = \begin{bmatrix} I_k & 0 \\ -CA^{-1} & I_{m-k} \end{bmatrix}$$

is an $m \times m$ invertible matrix. So

$$\operatorname{rank}(S) = \operatorname{rank}(TS) = \operatorname{rank} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

This latter matrix has rank k iff $D - CA^{-1}B = 0$ $\square$

Theorem 3.2.8. *Let X and Y be smooth manifolds with $n = \dim X$, $m = \dim Y$, and $q = min\{n, m\}$. Then $\mathcal{S}_r$ is a submanifold of $J^1(X, Y)$ with* codim $\mathcal{S}_r = (n-q+r)(m-q+r)$.

Proof. Let $\sigma \in \mathcal{S}_r$ with source p and target q. Let U and V be coordinate nbhds of p and q respectively. Then $J^1(X, Y)_{U\times V} \cong U \times V \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ and under this isomorphism $\mathcal{S}_r \cong U \times V \times L^r(\mathbb{R}^n, \mathbb{R}^m)$. Now apply Proposition 3.2.6. □

3.3 The Whitney $\mathcal{C}^\infty$ Topology

With the jet bundle in hand, we are ready to define a topology on $\mathcal{C}^\infty(X, Y)$.

Definition 3.3.1. Let X and Y be smooth manifolds. Fix a non-negative integer k. To each subset U of $J^k(X, Y)$, we associate a subset $M(U)$ of $\mathcal{C}^\infty(X, Y)$ defined by

$$M(U) = \{f \in \mathcal{C}^\infty(X, Y) \mid j^k f(X) \in U\}$$

The *Whitney $\mathcal{C}^k$ topology* on $\mathcal{C}^\infty(X, Y)$ is the topology whose basis is the family of sets $\{M(U)\}$ where U is an open subset of $J^k(X, Y)$. Denote by W_k the open subsets in the Whitney $\mathcal{C}^k$ topology. The *Whitney $\mathcal{C}^\infty$ topology* on $\mathcal{C}^\infty(X, Y)$ is the topology whose basis is $W = \bigcup_{k=0}^\infty W_k$. We sometimes abbreviate this name to the $\mathcal{C}^\infty$ *topology.*

Proposition 3.3.1. *The Whitney $\mathcal{C}^k$ and $\mathcal{C}^\infty$ topologies are well-defined.*

Proof. The family of sets $\{M(U) \mid U \subset J^k(X, Y) \text{ is open}\}$ forms a topological basis because $M(\emptyset) = \emptyset$, $M(J^k(X, Y)) = \mathcal{C}^\infty(X, Y)$, and $M(U) \bigcap M(V) = M(U \bigcap V)$. So the Whitney $\mathcal{C}^k$ topology is well-defined.

To extend this result to the $\mathcal{C}^\infty$ case, it suffices to verify that $W_k \subset W_l$ when $k \leq l$. For each such k and l, define the map $\pi^l_k : J^l(X, Y) \to J^k(X, Y)$ as follows: Let σ

be an $l - jet$ with source p, and let f represent σ. Then $\pi^l_k(\sigma)$ is the unique k-jet σ' with source p such that f represents σ'. It is clear from Proposition 3.2.1 that π^l_k is well-defined and continuous. Note $M(U) = M((\pi^l_k)^{-1}(U))$ and if U is an open subset of $J^k(X,Y)$ then $(\pi^l_k)^{-1}(U)$ is an open subset of $J^l(X,Y)$. Therefore $W_k \subset W_l$. $\square$

In order to get a feel for this topology, we will exhibit a neighborhood basis of a smooth map f. Let d be a metric on $J^k(X,Y)$ compatible with its topology (In §5 we will see that such a metric exists). For each continuous map $\delta : X \to \mathbb{R}^+$, define

$$B_\delta(f) = \{g \in \mathcal{C}^\infty(X,Y) \mid d(j^k f(x), j^k g(x)) < \delta(x) \text{ for every } x \in X\}$$

Proposition 3.3.2. *The $B_\delta(f)$ form a neighborhood basis at f in the $\mathcal{C}^k$ topology on $\mathcal{C}^\infty(X,Y)$.*

Proof. We divide the proof into three parts.

Part A: $B_\delta(f)$ is open.

Part B: If W is a nbhd of $f \in \mathcal{C}^\infty(X,Y)$, then there exists $\delta : X \to \mathbb{R}^+$ such that $B_\delta(f) \subset W$.

Part C: Given continuous maps $\delta, \gamma : X \to \mathbb{R}^+$, there exists a continuous map $\eta : X \to \mathbb{R}^+$ such that $B_\delta(f) \bigcap B_\gamma(f) = B_\eta(f)$.

Proof of Part A: Define the map $\Delta : J^k(X,Y) \to \mathbb{R}$ by $\sigma \mapsto \delta(\alpha(\sigma)) - d(j^k f(\alpha(\sigma)), \sigma)$. Then Δ is continuous, $U = \Delta^{-1}(0,\infty)$ is open, and $B_\delta(f) = M(U)$.

Proof of Part B: Let W be a nbhd of $f \in \mathcal{C}^\infty(X,Y)$. Then there exists an open subset V of $J^k(X,Y)$ with $f \in M(V) \subset W$. Define $\nu : X \to \mathbb{R}^+$ by

$$\nu(x) = \inf\{d(\sigma, j^k f(x)) \mid \sigma \in \alpha^{-1}(x) \bigcap (J^k(X,Y) - V)\}$$

with $\nu(x) = \infty$ if $\alpha^{-1}(x) \subset V$. ν is well-defined because $\alpha^{-1}(x) \bigcap (J^k(X,Y) - V)$ is closed and $j^k f(x) \in V$. Also ν is bounded below by a positive constant on any compact subset of X. Thus we can construct a continuous map $\delta : X \to \mathbb{R}^+$ such that $\delta(x) < \nu(x)$ for every $x \in X$ using a partition of unity (see §5.1 for the definition of this tool). Then $B_\delta(f) \subset M(V) \subset W$.

Proof of Part C: Define $\eta : X \to \mathbb{R}^+$ by $\eta(x) = \min\{\delta(x), \gamma(x)\}$. η is continuous and it is clear from the definition of $B_\delta(f)$ that $B_\delta(f) \bigcap B_\gamma(f) = B_\eta(f)$. $\square$

We may think of $B_\delta(f)$ as consisting of those smooth maps whose first k partial derivatives are all δ-close to f. If X is compact, we can find a countable nbhd basis by taking the collection of $B_{\delta_n}(f)$ where $\delta_n \equiv \frac{1}{n}$. So in this case, the $\mathcal{C}^\infty$ topology satisfies the first axiom of countability, and one may prove straightforwardly that a sequence of functions f_n in $\mathcal{C}^\infty(X, Y)$ converges to f (in the $\mathcal{C}^\infty$ topology) iff $j^k f_n$ converges uniformly to $j^k f$. One can further check that $j^k : \mathcal{C}^\infty(X, Y) \to \mathcal{C}^\infty(X, J^k(X, Y))$ is continuous with respect to the $\mathcal{C}^\infty$ topology on the domain and codomain [5, p.46].

We will need two more topological notions in the sections that follow:

Definition 3.3.2. Let F be a topological space. A subspace G of F is *residual* if it is the countable intersection of open dense subsets. F is a *Baire* space if all residual subsets are dense.

Lemma 3.3.3 (Baire Lemma). *Let X and Y be smooth manifolds. Then $\mathcal{C}^\infty(X, Y)$ with the $\mathcal{C}^\infty$ topology is a Baire space.*

A proof of the Baire Lemma may be found in [5, p.44]. Note that, in general, the properties open, dense, and residual are completely unordered in terms of strength. However, the Baire Lemma implies that all residual subsets of $\mathcal{C}^\infty(X, Y)$ are dense in $\mathcal{C}^\infty$ topology.

3.4 The Thom Transversality Theorem

We will need the following lemma:

Lemma 3.4.1. *Let X and Y be smooth manifolds, W a submanifold of Y. Let*

$$T_W = \{f \in \mathcal{C}^\infty(X,Y) \mid f \pitchfork W\}.$$

If W is closed then T_W is an open subset of $\mathcal{C}^\infty(X,Y)$ (in the $\mathcal{C}^1$, and thus, $\mathcal{C}^\infty$, topology).

Proof. Define a subset U of $J^1(X,Y)$ as follows. Let $\sigma \in J^1(X,Y)$ with source $p \in X$ and target $q \in Y$. Let f represent σ. Then $\sigma \in U$ iff either

(i) $q \notin W$

(ii) $q \in W$ with $T_qY = T_qW + (df)_p(T_pX)$.

It follows that $j^1f(p) \in U$ iff $j^1f \pitchfork_p W$. Thus

$$\begin{aligned} T_W &= \{f \in \mathcal{C}^\infty(X,Y) \mid j^1f \pitchfork W\} \\ &= \{f \in \mathcal{C}^\infty(X,Y) \mid j^1f(X) \subset U\} \\ &= M(U). \end{aligned}$$

Recall that the collection $\{M(U) \mid U \subset J^1(X,Y) \text{ is open}\}$ forms a basis for the $\mathcal{C}^1$ topology. Thus it suffices to show that U is open, or equivalently that $V = J^1(X,Y) - U$ is closed.

Let $\sigma_1, \sigma_2, ...$ be a convergent sequence with $\sigma_i \in V$ for every i and $\lim_{i\to\infty} \sigma_i = \sigma$. We will show that $\sigma \in V$. Let $p = \alpha(\sigma)$ and $q = \beta(\sigma)$, and let f represent σ. Since $\sigma_i \notin U$ and W is closed, we conclude that $q = lim_{i\to\infty}\beta(\sigma_i)$ lies in W. Choose charts (U,φ) at p in X and (V,ψ) at q in Y such that $f(U) \subset V$. Via these charts we may assume that $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, and $W = \mathbb{R}^k \subset \mathbb{R}^m$. Let $\pi : \mathbb{R}^m \to \mathbb{R}^m/\mathbb{R}^k = \mathbb{R}^{m-k}$ be projection. By Lemma 3.1.2, $f \pitchfork W$ iff $\pi \cdot f$ is a submersion at 0 iff $\pi \cdot (df)_0 \notin F$ where

$$F = \{A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \mid \mathrm{rank}\ A < m - k\}.$$

Define the map

$$\eta : \mathbb{R}^n \times W \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \subset J^1(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^{m-k})$$

by $\eta(x, w, B) = \pi \cdot B$. η is continuous and F is closed (as follows from Proposition 3.2.6), so $\eta^{-1}(F)$ is closed in $\mathbb{R}^n \times W \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ which, in turn, is closed in $J^1(\mathbb{R}^n, \mathbb{R}^m)$. Moreover V is precisely $\eta^{-1}(F)$ since $(x, y, (dg)_0) \in V \Leftrightarrow g$ is not transverse to W at $0 \Leftrightarrow \eta(x, y, g) = \pi \cdot (dg)_0 \in F$. Since V is closed in the local situation, we conclude that σ is in V. $\square$

Recall Proposition 3.1.6, in which we showed that we can make any map $X \to Y$ transverse to a fixed submanifold of Y by an arbitrarily small translation of the map. The Thom Transversality Theorem makes an even stronger claim: we can make the k-jet of any map $X \to Y$ transverse to a fixed submanifold of $J^k(X, Y)$ by an arbitrarily small perturbation of the map. A translation will not suffice here since it would leave the partial derivatives of the map fixed. Instead, we will need to perturb the map locally by a polynomial. To realize this local perturbation smoothly, we need the help of cut-off functions. Namely, if $U \subset X$ is open and $V \subset X$ is compact with $V \subset U$, then there exists a smooth function $\rho : X \to \mathbb{R}$ (a *cut-off function*) such that

$$\rho = \begin{cases} 1 \text{ on a neighborhood of V} \\ 0 \text{ off U} \end{cases}$$

See [7, p.36] for a proof of this standard fact.

Theorem 3.4.2 (Thom Transversality Theorem). *Let X and Y be smooth manifolds and W a submanifold of $J^k(X, Y)$. Then the set*

$$T_W = \{f \in \mathcal{C}^\infty(X, Y) \mid j^k f \pitchfork W\}$$

is a residual subset of $\mathcal{C}^\infty(X, Y)$ in the $\mathcal{C}^\infty$ topology. Moreover, T_W is open if W is closed.

Proof. Our proof follows [5, p.54]. We must show that T_W is the countable intersection of open dense subsets. Choose a countable covering of W by open subsets $W_1, W_2, ...$ in W such that each W_r satisfies:

a) The closure of W_r in $J^k(X, Y)$ is contained in W.

b) $\overline{W}_r$ is compact.

c) There exist coordinate neighborhoods $U_r \subset X$ and $V_r \subset Y$ such that $\pi(W_r) \subset U_r \times V_r$, where $\pi : J^k(X, Y) \to X \times Y$ is the projection $\alpha \times \beta$.

d) $\overline{U}_r$ is compact.

We can do this by first collecting such an open nbhd at each point of W. Since W is second countable, we can then extract a countable subcover. Now let

$$T_{W_r} = \{f \in \mathcal{C}^\infty(X, Y) \mid j^k f \pitchfork W \text{ on } \overline{W}_r\}.$$

Then $T_W = \bigcap_{r=1}^\infty T_{W_r}$, so we have reduced the proof to showing that each T_{W_r} is open and dense.

Open: Let $T_r = \{g \in \mathcal{C}^\infty(X, J^k(X, Y)) \mid g \pitchfork W \text{ on } \overline{W}_r\}$. T_r is open by Lemma 3.4.1. $j^k : \mathcal{C}^\infty(X, Y) \to \mathcal{C}^\infty(X, J^k(X, Y))$ is continuous. Therefore $T_{W_r} = (j^k)^{-1}(T_r)$ is open. Note that the same reasoning shows that T_W is open if W is closed.

Dense: Choose charts $\psi : U \to \mathbb{R}^n$ and $\eta : V \to \mathbb{R}^m$ and smooth functions $\rho : \mathbb{R}^n \to [0, 1] \subset \mathbb{R}$ and $\rho' : \mathbb{R}^m \to [0, 1] \subset \mathbb{R}$ such that

$$\rho = \begin{cases} 1 \text{ on a neighborhood of } \psi \cdot \alpha(\overline{W}_r) \\ 0 \text{ off } \psi(U_r) \end{cases} \quad \rho' = \begin{cases} 1 \text{ on a neighborhood of } \eta \cdot \beta(\overline{W}_r) \\ 0 \text{ off } \eta(V_r) \end{cases}$$

These cut-off functions exist because $\overline{W}_r$ is compact.

Let B' be the space of degree k polynomial maps $\mathbb{R}^n \to \mathbb{R}^m$. For $b \in B'$, define $g_b : X \to Y$ by

$$g_b(x) = \begin{cases} f(x) \text{ if } x \notin U_r \text{ or } f(x) \notin V_r. \\ \eta^{-1}(\rho(\psi(x))\rho'(\eta f(x))b(\psi(x)) + \eta f(x)) \text{ otherwise.} \end{cases}$$

By inspection the map $(x, b) \mapsto g_b(x)$ is smooth.

Now define $\Phi : X \times B' \to J^k(X, Y)$ by $\Phi(x, b) = j^k g_b(x)$. We claim that we can construct an open nbhd B of 0 in B' such that $\Phi|_{X \times B}$ will be transverse to W on some nbhd of $\overline{W}_r$. Assuming this, we can apply Proposition 3.2.5 on $X \times B$ to obtain $b_1, b_2, ...$ in B converging to 0 such that $j^k g_{b_i} \pitchfork W$ on $\overline{W}_r$. Since $g_0 = f$ and $g_b = f$ off U_r, we have $\lim_{i \to \infty} g_{b_i} = f$. Therefore T_{W_r} is a dense subset of $\mathcal{C}^\infty(X, Y)$.

We now construct a nbhd B of 0 in B' such that $\Phi|_{B \times X} : B \times X \to J^k(X, Y)$ is a local diffeomorphism and thus transverse to every submanifold of $J^k(X, Y)$. Let $\varepsilon = \frac{1}{2} \min\{d(\text{supp } \rho, \mathbb{R}^m - \eta(V_r)), d(\eta\beta(\overline{W}_r), (\rho')^{-1}[0, 1))\}$. Since $\Phi(x, b) \in \overline{W}_r$, we have that $x \in \alpha(\overline{W}_r)$ and $g_b(x) \in \beta(\overline{W}_r)$. Then $s = d(\eta f(x), \eta g_b(x)) < \varepsilon$ because

$$\eta g_b(x) = \rho(\psi(x))\rho'(\eta f(x))b(\psi(x)) + \eta f(x).$$

So

$$s = |\rho\psi(x)\rho'\eta f(x)b\psi(x)| \begin{cases} \leq |b\psi(x)| < \varepsilon & \text{if } \psi x \in \text{supp } \rho \\ = 0 & \text{if } \psi x \notin \text{supp } \rho. \end{cases}$$

Now $g_b(x)$ is in $\beta(\overline{W}_r)$, so using the definition of ε we conclude that $\eta f(x)$ is in the interior of $(\rho')^{-1}(1)$. Thus our choice of ε sufficiently limits the polynomial perturbation b so that $\eta f(x)$ lies within the region on which $\rho' \equiv 1$ when $\Phi(x, b) \in \overline{W}_r$. Recall as well that $\rho \equiv 1$ on a nbhd of $\psi\alpha(\overline{W}_r)$. Taken together, we have that $\eta g_b(x) = b\psi(x) + \eta f(x)$ and that $g_b(x') = \eta^{-1}(b\psi + \eta f)(x')$ for all x' in a nbhd of x. This argument also holds for all b' in some nbhd of b. We can now define a smooth inverse $\sigma \mapsto (x', b')$ of Φ near (x, b) as follows. For $\sigma \in J^k(X, Y)$ near $\Phi(x, b)$, let $x' = \alpha(\sigma)$, and let b' be the unique polynomial map of degree $\leq k$ such that $\sigma = j^k(\eta^{-1}(b'\psi + \eta f))(x')$. Therefore, as desired, $\Phi|_{B\times X}$ is a local diffeomorphism at (x, b). □

We can immediately generalize the Thom Transversality Theorem to deal with multiple submanifolds of the jet bundle.

Corollary 3.4.3. *Let X and Y be smooth manifolds, $\{W_i\}$ be a countable collection of submanifolds of $J^k(X, Y)$. Let*

$$T_W = \{f \in \mathcal{C}^\infty(X, Y) \mid j^k f \pitchfork W_i \text{ for every } 1 \leq i \leq k\}$$

Then T_W is a residual subset of $\mathcal{C}^\infty(X, Y)$ in the $\mathcal{C}^\infty$ topology. If the number of W_i is finite and each W_i is closed, then T_W is open as well.

Before generalizing this result further in §3.5, we ought to at least mention a more elementary transversality theorem which, though intuitive, suffers from not being particularly useful.

Theorem 3.4.4 (Elementary Transversality Theorem). *Let X and Y be smooth manifolds, W a submanifold of Y. Then the set $\{f \in \mathcal{C}^\infty(X, Y) \mid f \pitchfork W\}$ is a residual subset of $\mathcal{C}^\infty(X, Y)$ in the $\mathcal{C}^\infty$ topology. Moreover, this set is open if W is closed.*

The Elementary Transversality Theorem can be proven directly by the same method we used to prove the Thom Transversality Theorem. Alternatively, it can be deduced rather quickly from Thom by noting that $J^0(X,Y) = X \times Y$ and $j^0 f(x) = (x, f(x))$. We leave either approach to the industrious reader.

3.5 The Multijet Transversality Theorem

For the reader who fears the Thom Transversality Theorem to be insufficiently abstract, we now generalize still one step further. In particular, we will need the Multijet Transversality Theorem to prove claims about the generality of certain global properties of smooth maps, such as injectivity in §5.2.

Definition 3.5.1. Let X and Y be smooth manifolds. Define

$$\begin{aligned} X^s &= X \times \cdots \times X \quad (s \text{ times}) \\ X^{(s)} &= \{(x_1, ..., x_s) \in X^s \mid x_i \neq x_j \text{ for every } 1 \leq i < j \leq s\} \end{aligned}$$

Define the multijet source map $\alpha^s : (J^k(X,Y))^s \to X^s$ to be $\alpha \times \cdots \times \alpha$ (s times). Then

(1) $J^k_s(X,Y) = (\alpha^s)^{-1}(J^k(X,Y))$ is the *s-fold k-jet bundle* of maps from X to Y

or equivalently

(1′) $J^k_s(X,Y)$ is the *multijet bundle* of all s-tuples of k-jets with distinct sources.

Note that $X^{(s)}$ is an open subset of X^s and therefore a submanifold. In addition, α^s is continuous so $J^k_s(X,Y)$ is an open subset of $(J^k(X,Y))^s$ and a submanifold as well. Finally, we define $j^k_s : \mathcal{C}^\infty(X,Y) \to \mathcal{C}^\infty(X^{(s)}, J^k_s(X,Y))$ by $j^k_s f(x_1, ..., x_s) = (j^k f(x_1), ..., j^k f(x_s))$.

Theorem 3.5.1 (Multijet Transversality Theorem). *Let X and Y be smooth manifolds, W a submanifold of Js. Then the set*

$$T_W = \{f \in \mathcal{C}^\infty(X, Y) \mid j_s^k f \pitchfork W\}$$

is a residual subset of $\mathcal{C}^\infty(X, Y)$ in the $\mathcal{C}^\infty$ topology.

Proof. The proof follows the same strategy as the Thom Transversality Theorem, by proving a multijet analog to Lemma 3.4.1, then using the Transversality Lemma to prove a multijet analog to Proposition 3.2.5. The only complication is that we must now effect local polynomial perturbations of f on s regions simultaneously. This can be done (with care) because the sources of an s-fold k-jet are distinct. □

One can go further to show that if W is compact, than T_W is open [5, p.57], but we will not need this result.

In §1, §2, and §3, we developed powerful theorems on transversality.

In §4, §5, and §6, we use them to study the singularities of smooth maps.

Chapter 4

High to Low: Morse Theory

In §4.1, we give normal forms for non-degenerate critical points. In §4.2, we show that non-degenerate critical points are the only generic singularities of smooth functions.

4.1 The Morse Lemma

Let X be a smooth manifold, $f : X \to \mathbb{R}$ a smooth function. Since $\mathbb{R}$ is one dimensional as a manifold, the derivative of f must have rank zero or one at each $p \in X$. Thus a critical point p of f is simply a point for which all the partial derivatives of f vanish. Relative to any coordinate system we have:

$$\left(\frac{\partial f}{\partial x_1}\right)_p = ... = \left(\frac{\partial f}{\partial x_n}\right)_p = 0$$

However, not all critical points are created equal. The following tool encodes the critical information that we will use to construct normal forms for the structure of functions near most critical points.

Definition 4.1.1. Let $f : X \to \mathbb{R}$ be a smooth function.

(1) The *Hessian* of f at p, with respect to local coordinates $x_1, ..., x_n$, is the matrix $H_p(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_p$ of second order partial derivatives:

$$\begin{bmatrix} \left(\frac{\partial^2 f}{\partial x_1^2}\right)_p & \cdots & \left(\frac{\partial^2 f}{\partial x_1 \partial x_n}\right)_p \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial f^2}{\partial x_n \partial x_1}\right)_p & \cdots & \left(\frac{\partial^2 f}{\partial x_n^2}\right)_p \end{bmatrix}$$

(2) A critical point p of f is *degenerate* if $\det H_p(f) = 0$. Otherwise, p is *non-degenerate*.

(3) The *index* of f at a non-degenerate critical point p is the maximum dimension of a vector subspace of $\mathbb{R}^n$ on which $H_p(f)$ is negative definite.

Remark 4.1.1. $H_p(f)$ is negative definite on V if the corresponding bilinear form $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is negative definite, i.e. $H(v, v) < 0$ for every non-zero $v \in V$. Equivalently, the index can be viewed as the number of negative eigenvalues of the non-singular Hessian matrix.

Note that we have defined the Hessian of f at p in a way that depends on the particular chart chosen at p. There also exists an invariant formulation of the Hessian using the concept of intrinsic derivative [5, 64]. While we have avoided the latter approach for simplicity, we must now do a little work to verify that the degeneracy and index of a function at a point are well-defined notions.

Proposition 4.1.1. *The degeneracy and index of f at p do not depend on the coordinates chosen on X.*

Proof. Let $A = H_p(f)$ be the Hessian matrix with respect to the coordinates $x_1, ..., x_n$ given by a chart (U, φ) of X at p. Let $\psi : \varphi(U) \to \varphi(U)$ be a change of coordinates defined by $\psi(x_1, ..., x_n) = (y_1(x_1, ..., x_n), ..., y_n(x_1, ..., x_n))$. Then the matrix $P = (d\psi)_0$ is non-singular, and the matrix of the Hessian of f at p with respect to the coordinates $y_1, ..., y_n$ is given by $(P^{-1})^T A P^{-1}$. The latter claim is an exercise in quadratic forms, namely that a change of coordinates replaces a quadratic form with matrix A by a quadratic form with matrix $B^T AB$, where B is non-singular. Clearly A is singular iff BAB^T is singular. And if A is non-singular, than A and BAB^T have the same index by Sylvester's Law [1, p.245]. $\square$

By the Submersion Lemma, a smooth function is locally equivalent at a regular point to projection onto the first coordinate. The Morse Lemma provides normal forms for the local behavior of smooth functions at non-degenerate critical points.

Theorem 4.1.2 (Morse Lemma). *Let $f : X \to \mathbb{R}$ be a smooth function, $p \in X$ a non-degenerate critical point of f, and λ the index of f at p. Then near p, f is equivalent to the map $(x_1, ..., x_n) \mapsto f(p) - x_1^2 - ... - x_\lambda^2 + x_{\lambda+1}^2 + ... + x_n^2$.*

Our proof of the Morse Lemma fleshes out the sketch given by Milnor [10, p.6] and will require the following calculus result.

Lemma 4.1.3. *Let f be a smooth function on some convex region $V \subset \mathbb{R}^n$, with $f(0) = 0$. Then there exist smooth functions $g_1, ..., g_n$ on V with*

$$f(x_1, ..., x_n) = \Sigma_{i=1}^n x_i g_i(x_1, ..., x_n)$$

and $g_i(0) = \frac{\partial f}{\partial x_i}(0)$ for every $1 \leq i \leq n$.

Proof.

$$f(x_1, ..., x_n) = \int_0^1 \frac{df}{dt}(x_1 t, ..., x_n t) dt = \int_0^1 \left(\sum_{i=0}^n \frac{\partial f}{\partial x_i}(x_1 t, ..., x_n t) x_i \right) dt$$

by the fundamental theorem of calculus and the chain rule. Note that convexity guarantees that the above integral is defined. So it suffices to set

$$g_i(x_1, ..., x_n) = \int_0^1 \frac{\partial f}{\partial x_i}(x_1 t, ..., x_n t) dt$$

where $g_i(0) = \frac{\partial f}{\partial x_i}(0)$ again follows from the fundamental theorem of calculus. $\square$

Proof of the Morse Lemma. In *Part A*, we will prove the existence of a change of coordinates on the domain which yields the diagonalized quadratic form $f(p) \pm x_1^2 \pm ... \pm x_n^2$. In *Part B* we will show that the index of $f(p) - x_1^2 - ... - x_\lambda^2 + x_{\lambda+1}^2 + ... + x_n^2$ at 0 is λ.

Part A. We can assume without loss of generality that $0 = p = f(p)$ and $X = \mathbb{R}^n$, since we are only concerned with local equivalence. By Lemma 4.1.3, there exist smooth functions $g_1, ..., g_n$ on $\mathbb{R}^n$ with

$$f(x_1, ..., x_n) = \Sigma_{i=1}^n x_i g_i(x_1, ..., x_n)$$

and $g_i(0) = \frac{\partial f}{\partial x_i}(0)$. Since $0 \in \mathbb{R}^n$ is a critical point, we have $\frac{\partial f}{\partial x_i}(0) = 0$ for every $1 \leq i \leq n$. Therefore we can apply Lemma 4.1.3 again, this time to each of the g_i. So there exist smooth functions h_{ij}, $1 \leq i, j \leq n$, such that

$$g_i(x_1, ..., x_n) = \Sigma_{i=1}^n x_i h_{ij}(x_1, ..., x_n)$$

Substitution gives

$$f(x_1, ..., x_n) = \Sigma_{i,j=1}^n x_i x_j h_{ij}(x_1, ..., x_n)$$

Furthermore, we can assume $h_{ij} = h_{ji}$ (otherwise replace each h_{ij} with $\frac{1}{2}(h_{ij} + h_{ji})$). Differentiating gives $\frac{\partial^2 f}{\partial x_i \partial x_j}(0) = 2h_{ij}(0)$, so the matrix $(h_{ij}(0)) = (\frac{1}{2}(\frac{\partial^2 f}{\partial x_i \partial x_j}(0)))$. By hypothesis, 0 is a non-degenerate critical point of f, so we conclude that $(h_{ij}(0))$ is non-singular.

We now precede as in the proof of the diagonalization of quadratic forms [2, p.286]. Suppose by induction that there exist coordinates $u_1, ..., u_n$ on a neighborhood U_1 of 0 such that

$$f(u_1, ..., u_n) = \pm u_1^2 \pm ... \pm u_{r-1}^2 + \textstyle\sum_{i,j=r}^n u_1 u_2 H_{ij}(u_1, ..., u_n)$$

on U_1, where the H_{ij} are smooth functions with $H_{ij} = H_{ji}$ and the matrix $(H_{ij}(0))$ non-singular. We have already established the base case $r = 0$.

For the induction step, we first show that we can make $H_{rr}(0) \neq 0$ by a non-singular linear transformation on the last $n - r + 1$ coordinates. The proof works the same for any r, so for simplicity let r = 1. If we have $H_{ii}(0) \neq 0$ for some $1 \leq i \leq n$ then we are done by transposing u_1 and u_i. Otherwise, since $(H_{ij}(0))$ is non-singular, there exists some $H_{ij}(0) \neq 0$ with $i \neq j$. Through a pair of transpositions, we can assume $H_{11}(0) = 0$ and $H_{12}(0) = H_{21}(0) \neq 0$. We define a new set of coordinates $u_1', ..., u_n'$ on U_1 by

$u_1' = \frac{1}{2}(u_1 + u_2)$ $\qquad$ $u_2' = \frac{1}{2}(u_1 - u_2)$ $\qquad$ $u_i' = u_i$ for $i > 2$

This linear transformation is invertible with inverse given by

$u_1 = (u_1' - u_2')$ $\qquad$ $u_2 = (u_1' + u_2')$ $\qquad$ $u_i = u_i'$ for $i > 2$

Substituting in these new coordinates and regrouping terms, we have $f(0) = \Sigma_{i,j=1}^n u_i' u_j' H_{ij}'(0)$ with $H_{11}'(0) = H_{12}(0) + H_{21}(0) = 2H_{12} \neq 0$.

So without loss of generality we assume $H_{rr}(0) > 0$ (sending u_r to $-u_r$ if necessary). Then there exists a neighborhood $U_2 \subset U_1$ of 0 on which H_{rr} is positive. We now define a new set of coordinates $v_1, ..., v_n$ by

$v_i = u_i$ for $i \neq r$.

$v_r = \sqrt{H_{rr}(u_1, ..., u_n)}\left[u_r + \sum_{i>r} u_i H_{ir}(u_1, ..., u_n)/H_{rr}(u_1, ..., u_n)\right]$

Note v_r is well-defined and smooth on U_2. A simple calculation shows $\frac{\partial v_r}{\partial u_r} = \sqrt{H_{rr}}$ so $\frac{\partial v_r}{\partial u_r}(0) \neq 0$. It follows from the Inverse Function Theorem that the change of coordinates map ϕ defined by $(u_1, ..., u_n) \mapsto (v_1(u_1, ..., u_n), ..., v_n(u_1, ..., u_n))$ is a diffeomorphism in some sufficiently small neighborhood $U_3 \subset U_2$ of 0. Then

$$
\begin{aligned}
f &= \pm u_1^2 \pm \dots \pm u_{r-1}^2 + \sum_{i,j=r}^{n} u_i u_j H_{ij} \\
&= \pm u_1^2 \pm \dots \pm u_{r-1}^2 + \left[u_r^2 H_{rr} + 2u_r \sum_{i>r} u_i H_{ri} + \sum_{i>r} u_i^2 H_{ii}/H_{rr} \right] \\
&\quad + \sum_{i>r} u_i^2 (H_{ii} - H_{ii}/H_{rr}) + \sum_{i,j>r, i\neq j} u_i u_j H_{ij}
\end{aligned}
$$

The term in brackets is v_r^2 so it is clear that we can choose smooth functions $H'_{ij}(v_1, ..., v_n)$ for $i > r$ so that

$$f(v_1, ..., v_n) = \textstyle\sum_{i=1}^{r} \pm v_i^2 + \sum_{i,j>r}^{n} v_i v_j H'_{ij}(v_1, ..., v_n).$$

with $H'_{ij} = H'_{ji}$. Furthermore, $(H'_{ij}(0)) = ((d\phi)_0^{-1})^T (H_{ij}(0))(d\phi)_0^{-1}$ is non-singular. This completes the induction step and the first part of the proof.

Part B. Define $g : \mathbb{R}^n \to \mathbb{R}$ by $g(x_1, ..., x_n) = g(p) - x_1^2 - ... - x_\lambda^2 + x_{\lambda+1}^2 + ... + x_n^2$. Computing partial derivatives we have

$$
H_p(g) = \begin{bmatrix}
-2 & & & & & \\
& \ddots & & & & \\
& & -2 & & & \\
& & & 2 & & \\
& & & & \ddots & \\
& & & & & 2
\end{bmatrix}
$$

The first λ basis vectors span a subspace $V \subset \mathbb{R}(n)$ on which $H_p(g)$ is negative definite, so the index of g at p is at least λ. The latter basis vectors span a subspace $W \subset \mathbb{R}^n$ of dimension $n - \lambda$ on which $H_p(g)$ is positive definite. If there exists a subspace V' of dimension greater than λ on which $H_p(g)$ is positive definite, then V'

and W would intersect nontrivially, a contradiction. Therefore, the index of g at p equals λ. $\square$

A function $f : X \to \mathbb{R}$ is called *Morse* if all of its critical points are non-degenerate. Between the Submersion Lemma and the Morse Lemma, we have completely determined the local structure of Morse functions. In addition, we have the following corollary.

Corollary 4.1.4. *Non-degenerate critical points are isolated.*

Proof. By the Morse Lemma, we need only consider the function $f(x_1, ..., x_n) = -x_1^2 - ... - x_\lambda^2 + x_{\lambda+1}^2 + ... + x_n^2$. Clearly the partial derivatives $\frac{\partial f}{\partial x_i}(p) = \pm 2x_i$ do not simultaneously vanish away from the origin. $\square$

4.2 Morse Functions are Generic

We have seen that Morse functions have simple local behavior, but this result would be of little use if few functions satisfied the Morse property. In this section, we will see that in fact almost all functions are Morse. First, we had better make "almost all" precise.

Definition 4.2.1. A property P is *generic* if the set $\{f \in \mathcal{C}^\infty(X, Y) | f \text{ has property } P\}$ is a residual subset of $\mathcal{C}^\infty(X, Y)$.

The goal of this section is to show that the quality of being Morse is a generic property of smooth functions. Our strategy will be to translate non-degeneracy into a transversality condition on jets and apply the Thom Transversality Theorem. Recall from §3.2 that $\mathcal{S}_r$ is the smooth submanifold of $J^1(X, \mathbb{R})$ consisting of those jets which drop rank by r. For a smooth function $f : X \to \mathbb{R}$, only $\mathcal{S}_0$ and $\mathcal{S}_1$ can be non-empty. Moreover, $j^1 f : X \to J^1(X, \mathbb{R})$ maps critical points to $\mathcal{S}_1$ and regular points to $\mathcal{S}_0$. The following proposition provides the key link between non-degeneracy and transversality.

Lemma 4.2.1. *Let $f : X \to \mathbb{R}$ be a smooth function with a critical point at p. Then p is non-degenerate iff $j^1 f \pitchfork \mathcal{S}_1$ at p.*

Proof. Degeneracy is a local property, so we assume $X = \mathbb{R}^n$. Now $J^1(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n \times \mathbb{R} \times \operatorname{Hom}(\mathbb{R}^n, \mathbb{R})$. Let $\pi : J^1(\mathbb{R}^n, \mathbb{R}) \to \operatorname{Hom}(\mathbb{R}^n, \mathbb{R})$ be the projection. By Lemma 3.1.2, $j^1 f \pitchfork \mathcal{S}_1$ at p iff $\pi \cdot j^1 f : \mathbb{R}^n \to \operatorname{Hom}(\mathbb{R}^n, \mathbb{R})$ is a submersion at p. But $\pi \cdot j^1 f$ is simply the gradient of f. And $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is a submersion iff $(d\nabla f)_p = H_p(f)$ is non-singular. □

Theorem 4.2.2. *Let X be a smooth manifold. The set of Morse functions is an open dense subset of $\mathcal{C}^\infty(X, \mathbb{R})$ in the $\mathcal{C}^\infty$ topology.*

Proof. Note that codim $\mathcal{S}_1 > 0$ by Theorem 3.2.8, so Ss_1 is a closed submanifold of $J^k(X, Y)$. Thus the Thom Transversality Theorem and Baire Lemma imply that $\{f \in \mathcal{C}^\infty(X, \mathbb{R}) \mid j_1 f \pitchfork \mathcal{S}_1\}$ is an open dense subset of $\mathcal{C}^\infty(X, \mathbb{R})$. By Lemma 4.2.1, this set is exactly the set of Morse functions. □

In fact, using the Multijet Transversality Theorem we can make a statement about the critical values of smooth functions. We say that a function f has *distinct critical values* if no two critical point of f have the same image.

Lemma 4.2.3. *Let X be a smooth manifold. Then the set of smooth functions with distinct critical values is a residual subset of $\mathcal{C}^\infty(X, \mathbb{R})$.*

Proof. Let $\Delta\mathbb{R}^n$ be the diagonal $\{(x,x) \in \mathbb{R}^n \times \mathbb{R}^n\}$. We claim that the set $S = (\mathcal{S}_1 \times \mathcal{S}_1) \bigcap (\beta^2)^{-1}(\Delta\mathbb{R})$ is a submanifold of the multijet bundle $J^1_2(X,\mathbb{R})$. Let U be an open coordinate nbhd in X diffeomorphic to $\mathbb{R}^n$. In local coordinates

$$\begin{aligned} J^1_2(U,\mathbb{R}) &\cong (\mathbb{R}^n \times \mathbb{R}^n - \Delta\mathbb{R}^n) \times (\mathbb{R} \times \mathbb{R}) \times \mathrm{Hom}(\mathbb{R}^n,\mathbb{R})^2 \\ S &\cong (\mathbb{R}^n \times \mathbb{R}^n - \Delta\mathbb{R}^n) \times (\Delta\mathbb{R}) \times \mathrm{Hom}(\mathbb{R}^n,0)^2 \end{aligned}$$

so S is clearly a submanifold. Comparing the dimensions of the above factors, we see codim $S = 2n+1$.

Since S is a submanifold, we can apply the Multijet Transversality Theorem to conclude that the set of maps $f : X \to \mathbb{R}^n$ such that $j^1_2 f \pitchfork S$ is residual. Since $\dim X^{(2)} = 2n < 2n+1 = \text{codim } S$, by Proposition 3.1.1 we know that $j^1_2 f \pitchfork S$ implies $j^1_2 f(X^{(2)}) \bigcap S = \emptyset$. If p and q are distinct critical points of such an f, than certainly $j^1_2 f(p,q) = (j^1 f(p), j^1 f(q)) \in (\mathcal{S}_1 \times \mathcal{S}_1)$. Thus $(j^1 f(p), j^1 f(q)) \in (\beta^2)^{-1}(\Delta\mathbb{R})$ and we conclude $f(p) \neq f(q)$. ☐

The next proposition follows immediately from Lemma 4.2.3 and Theorem 4.2.2.

Proposition 4.2.4. *Let X be a smooth manifold. The set of Morse functions with distinct critical values is a residual subset of $\mathcal{C}^\infty(X,\mathbb{R})$.*

In order to elucidate the importance of this theorem, we conclude the chapter with a brief excursion into the theory of stable maps. Such maps are a central object of study in higher dimensional Singularity Theory and Catastrophe Theory. In fact, the latter field grew out of Thom's and others complete classification of the singularities of stable maps in dimension ≤ 4 (the so-called seven elementary catastrophes, two of which are the topic of §6).

Definition 4.2.2. Let $f : X \to Y$ be a smooth map of manifolds. f is *stable* if the equivalence class of f is open in $\mathcal{C}^\infty(X, Y)$ with the $\mathcal{C}^\infty$ topology.

Informally, f is stable if all nearby maps look like f. Note that if f is stable, then all of its differentially invariant properties are unchanged by sufficiently small perturbations of f.

Theorem 4.2.5. *Let $f : X \to \mathbb{R}$ be a smooth function. Then f is stable iff f is a Morse function with distinct critical values.*

We can easily prove the forward direction using Proposition 4.2.4, for it suffices to note that any function equivalent to a Morse function with distinct critical points is itself such a function. We invite the reader to fill in the details.

On the other hand, it is hard to show that Morse functions with distinct critical values are stable. The difficulty is a more general one: there is no good direct method for verifying the global stability of a map. V. Arnold and John Mather confronted this problem by developing a local condition called infinitesimal stability which is generally far easier to verify. At first glance, infinitesimal stability appears to be a weaker condition than global stability, but in 1969 Mather succeeded in proving that these properties are entirely equivalent [9]. This breakthrough precipitated a torrent of major advances in the decade that followed.

Chapter 5

Low to High: The Whitney Embedding Theorem

In §5.1, we prove the that 1:1 proper immersions are embeddings and that proper maps form a non-empty open subset of $\mathcal{C}^\infty(X, \mathbb{R}^m)$. In §5.2, we show that there are no generic singularities of maps from n-dimensional manifolds to m-dimensional manifolds when $m \geq 2n$. We then prove the Whitney Embedding Theorem.

5.1 Embeddings and Proper Maps

Definition 5.1.1. Let $f : X \to Y$ be a smooth map of manifolds. f is an *embedding* of X into Y if f is a diffeomorphism of X onto a submanifold of Y. In this case, we say f *embeds* X into Y.

The figures below illustrate the existence of 1:1 immersions which are not embeddings. In Figure 5.1, the real line is mapped into the plane to form a figure eight. The image fails to be a submanifold at the central point. In Figure 5.1, the real line is wrapped around the torus with an irrational ratio of equatorial to meridional loops so as to be an injective. Yet the image fails to be a submanifold near all points.

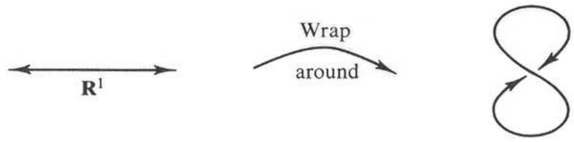

Figure 5.1: [6, p.16]

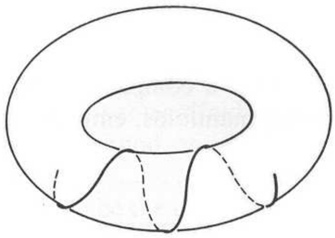

Figure 5.2: [6, p.17]

It seems that we run into trouble when many points near infinity are mapped close together. The following property gives us the additional control at infinity, preventing us from packing in too many distant points into a compact set in the image.

Definition 5.1.2. Let $f : F \to G$ be a continuous maps of topological spaces. Then f is *proper* if the preimage of every compact subset of G is compact in F.

Proposition 5.1.1. *Let $f : X \to Y$ be a smooth map of manifolds. If f is a proper $1 : 1$ immersion, then f embeds X into Y.*

Proof. By the Immersion Lemma, if the image of a 1:1 immersion is a submanifold then the inverse map is smooth. We show in two steps that if a 1:1 immersion is also proper, then this image is in fact a submanifold.

Step 1: f is a homeomorphism onto $f(X)$.

Step 2: $f(X)$ is a submanifold.

Proof of Step 1: Let U be an open subset of X. We will show that $f(U)$ is open in $f(X)$. Suppose by contradiction that $f(U)$ is not open. Then there exists a convergent sequence $y_1, y_2, ...$ in $f(X) - f(U)$ with $\lim_{i\to\infty} y_i = y \in f(U)$. Let $x_1, x_2, ...$ be the sequence of unique preimages of $y_1, y_2, ...$ and let x be the preimage of y. Then $x \in U$. On the other hand, $\{y, y_i\}$ is a compact subset of Y and f is proper, so $\{x_i\}$ is compact in X. Thus we have a convergent subsequence $x_{i_1}, x_{i_2}, ...$ in $X - U$. Mapping this subsequence forward, we see that its limit must be x as well. But since $X - U$ is closed, we arrive at the contraction $x \in X - U$.

Proof of Step 2: Let $p \in X$, $q = f(p)$. Since f is a homeomorphism onto its image, we can transport each chart (U, φ) of X at p to a chart $(f(U), \varphi f^{-1})$ of $f(X)$ at q. By the Immersion Lemma, it is clear that the resulting differential structure on each $f(U)$ is locally compatible with that of Y. Thus $f(X)$ is a submanifold of Y. □

Consider the space $\mathcal{C}^\infty(X, Y)$ of smooth maps from X to Y. If X is compact, then all maps in $\mathcal{C}^\infty(X, Y)$ are proper. However, if X is non-compact, then it is not immediately obvious that a single proper smooth map exists. We will see that proper maps always exist by first considering the space $\mathcal{C}^\infty(X, \mathbb{R})$. In order to construct a proper map in $\mathcal{C}^\infty(X, \mathbb{R})$, we will need the following standard tool on smooth manifolds.

Definition 5.1.3. A *partition of unity* of a smooth manifold X is an open covering $\{U_i\}$ of X and a system of smooth functions $\psi_i : X \to \mathbb{R}$ satisfying the following conditions:

(1) For every $x \in X$ we have $0 \leq \psi_i \leq 1$.

(2) The support of ψ_i is contained in U_i.

(3) The covering is locally finite.

(4) For every $x \in X$ we have $\sum_i \psi_i(x) = 1$. Note that the sum is taken over all i but is finite by (2) and (3).

A partition of unity $\{\psi_i, U_i\}$ is *subordinate* to an open cover $\{V_j\}$ of X if each U_i is contained in some V_j.

The main theorem on partitions of unity is their existence for paracompact spaces: Given any open cover $\{V_j\}$ of a smooth manifold, there exists a partition of unity $\{\psi_i, U_i\}$ subordinate to $\{V_j\}$ [7, p.32].

Lemma 5.1.2. *Let X be a smooth manifold. Then there exists a smooth proper function $\rho : X \to \mathbb{R}$.*

Proof. Let $\{\psi_i, U_i\}_{i=1,2,\ldots}$ be a partition of unity subordinate to a countable cover of X by open sets with compact closure. Define $\rho : X \to \mathbb{R}$ by $\rho(x) = \sum_{i=1}^{\infty} i\psi_i(x)$. ρ is well-defined and smooth because partitions of unity are locally finite. Now if $\rho(x) \leq j$ then by (4) we must have $\psi_i(x) \neq 0$ for some $i \leq j$. Therefore $\rho^{-1}([-j, j]) \subset \bigcup_{i=1}^{j} \{x | \psi_i(x) \neq 0\}$, the latter set having compact closure. So $\rho^{-1}([-j, j])$ is a closed subset of a compact set and therefore compact. To finish, note that every compact subset of $\mathbb{R}$ is contained in $[-j, j]$ for some j. $\square$

Having found a proper map in $\mathcal{C}^\infty(X, \mathbb{R})$, we now extend our result to $\mathcal{C}^\infty(X, Y)$.

Proposition 5.1.3. *Let X be a smooth manifold. Then the set of proper maps $X \to \mathbb{R}^m$ is a non-empty open subset of $\mathcal{C}^\infty(X, Y)$.*

Proof. Non-empty: By Lemma 5.1.2, there exists a smooth proper function $\rho : X \to \mathbb{R}$. Compose ρ with a linear injection $\mathbb{R} \to \mathbb{R}^m$ to obtain a smooth proper map $X \to \mathbb{R}^m$.

Open: Let $f : X \to \mathbb{R}^m$ be proper and let $V_x = \{y \in \mathbb{R}^m \mid d(f(x), y) < 1\}$. Let $V = \bigcup_{x \in X} x \times V_X$ in $J^0(X, Y) = X \times \mathbb{R}^m$. V is open because f is continuous. Therefore $M(V)$ is an open in $\mathcal{C}^\infty(X, Y)$ and clearly $f \in M(V)$. We will show that if $g \in M(V)$ then g is proper. Let $B_j = \{y \in \mathbb{R}^m \mid d(y, 0) \leq j\}$. Then by the triangle inequality we have $g^{-1}(B_j) \subset f^{-1}(B_{j+1})$. Thus $g^{-1}(B_j)$ is a closed subset of a compact set, and thus compact. To finish, note that every compact subset of $\mathbb{R}^m$ is contained in B_j for some j. □

5.2 Proof of the Whitney Embedding Theorem

Let X be a smooth manifold of dimension n. The Whitney Embedding Theorem states that X embeds into $\mathbb{R}^{2n+1}$. To see intuitively why $2n + 1$ dimensions suffice, trace a smooth curve in $\mathbb{R}^3$ with your fingertip. The drawn curve may intersect itself, but with the slightest of adjustments we can push the curve off of itself entirely.

In this section, we will formalize this intuition using the tools developed in §3. But first, it is interesting to note a more elementary approach which proves the Whitney Embedding Theorem for manifolds that are already embedded in $\mathbb{R}^N$ for some N. The idea is to show that if $N > 2n+1$, then there exists a linear projection $\pi : \mathbb{R}^N \to \mathbb{R}^{N-1}$ such that $\pi \cdot \rho : X \to \mathbb{R}^{N-1}$ is a 1:1 proper immersion. Thus by induction we can reduce N to $2n + 1$. This approach is taken by Guillemin and Pollack without loss of generality because they define manifolds concretely as subsets of some $\mathbb{R}^N$ at the outset [6, p.51].

With our abstract definition, this approach reduces the Whitney Embedding Theorem to showing that X can be realized as a subset of some $\mathbb{R}^N$. If X is compact, we can use a cover of X by a finite number of charts to construct such an embedding [12, p.223]. However, a constructive argument seems impossible in the case that X is non-compact. Instead, we will prove existence using a more powerful approach based on the Thom Transversality Theorem.

The following obvious lemma provides the key link between immersions $X \mapsto Y$ and submanifolds of $J^1(X, Y)$.

Lemma 5.2.1. *$f : X \to Y$ is an immersion iff $j^1 f(X) \bigcap \left(\bigcup_{r \neq 0} \mathcal{S}_r\right) = \emptyset$.*

Proposition 5.2.2 (Whitney Immersion Theorem). *Let X and Y be smooth manifolds with $\dim Y \geq 2 \dim X$. Then the set of immersions $X \to Y$ is an open dense subset of $\mathcal{C}^\infty(X, Y)$.*

Proof. Open: By Theorem 3.2.8, $\mathcal{S}_0$ an open submanifold of $J^1(X, Y)$. Thus, the set of immersions $X \to Y$, which equals $M(\mathcal{S}_0)$, is open in $\mathcal{C}^\infty(X, Y)$.

Dense: By Lemma 3.2.8, codim $\mathcal{S}_r = (n - q + r)(m - q + r)$ where $n = \dim X$, $m = \dim Y$, and $q = min\{n, m\}$. Since $\dim Y \geq 2 \dim X$, we have $m \geq 2n$ and $q = n$, so for $r \geq 1$ we have

$$\text{codim } \mathcal{S}_r = r(m - n + r) \geq m - n + r \geq n + r > n.$$

Therefore, $j^1 f \pitchfork \mathcal{S}_r$ for every $r \geq 1$ iff $j^1 f(X) \bigcap \left(\bigcup_{r \neq 0} \mathcal{S}_r\right) = \emptyset$ iff f is an immersion, by Proposition 3.1.1 and Lemma 5.2.1. We are done by Corollary 3.4.3 to the Thom Transversality Theorem. □

Note where the dimension requirement is used: we must have $m \geq 2n$ in order to enforce codim $\mathcal{S}_1 =$ codim $L^1(\mathbb{R}^n, \mathbb{R}^m) > n$ and in turn apply Thom via Proposition 3.1.1. We have shown that there are no generic singularities of maps from

n-dimensional manifolds to m-dimensional manifolds when $m \geq 2n$. We now turn to the global property of injectivity and formalize the intuition that most maps from a small space to a large space are 1:1.

Lemma 5.2.3. *Let X and Y be smooth manifolds with $\dim Y \geq 2\dim X + 1$. Then the set of 1:1 maps $X \to Y$ is a residual subset of $\mathcal{C}^\infty(X, Y)$.*

Proof. Let $n = \dim X$ and $m = \dim Y$. The diagonal $\Delta Y = \{(y, y) \in Y \times Y\}$ is diffeomorphic to Y and thus a submanifold of $Y \times Y$ dimension m. Define $W = (\beta^2)^{-1}(\Delta Y)$, a submanifold of $J_2^0(X, Y)$ of codimension m by Theorem 2.1.2 and Theorem 3.2.4. Let $f : X \to Y$ be a smooth map. Then f is 1:1 iff the image of $j_2^0 f : X^{(2)} \to J_2^0(X, Y)$ does not intersect W. Since $\operatorname{codim} W = m > 2n = \dim X^{(2)}$, by Proposition 3.1.1 we conclude that f is 1:1 iff $j_2^0 f \pitchfork W$. We are done by the Multijet Transversality Theorem. $\square$

Theorem 5.2.4. *Let X and Y be smooth manifolds with $\dim Y \geq 2\dim X + 1$. Then the set of 1:1 immersions $X \to Y$ is a residual (and thus dense) subset of $\mathcal{C}^\infty(X, Y)$.*

Proof. The result is immediate from the Whitney Immersion Theorem, Lemma 5.2.3, and the Baire Lemma. $\square$

In fact, the set of 1:1 immersions in these relative dimensions is open. The proof of this fact does not require any new tools, but it is tedious [5, p.61]. At any rate, we do not need it to prove our main theorem.

Theorem 5.2.5 (Whitney Embedding Theorem). *Let X be a smooth manifold of dimension n. Then X embeds into $\mathbb{R}^{2n+1}$.*

Proof. The set of proper maps $X \to \mathbb{R}^{2n+1}$ is non-empty and open by Proposition 5.1.3. The set of 1:1 immersions $X \to \mathbb{R}^{2n+1}$ is dense by Theorem 5.2.4. Therefore there exists a proper 1:1 immersion $X \to \mathbb{R}^{2n+1}$, which is an embedding by Proposition 5.1.1. □

For the curious, in these relative dimensions the stable maps are exactly the 1:1 immersions [5, p.81].

Whitney proved Theorems 5.2.2 and 5.2.5 in 1936 [17]. Eight years later, using more difficult techniques in self-intersection theory and algebraic topology, Whitney was able to improve the dimensions of each theorem by one: Any mapping of an n-manifold ($n \geq 2$) into $\mathbb{R}^{2n-1}$ may be perturbed slightly to an immersion [19], and any n-manifold may be embedded in $\mathbb{R}^{2n}$ [18]. At the time, Whitney remarked,"It is a highly difficult problem to see if the imbedding and immersion theorems of the preceding paper and the present one can be improved upon" [19, p.1]. However, as algebraic topology evolved, it became clear that the first obstructions to immersing an n-manifold X are the Stiefel-Whitney characteristic classes $\omega_i(X)$ of the stable normal bundle. In 1960, W. Massey proved that $\omega_i(X) = 0$ for $i > n - \nu(n)$, where $\nu(n)$ is the number of ones in the base-2 expansion of n (so $n = 2^{i_1} + \cdots + 2^{i_{\nu(n)}}$ with $i_1 < \cdots < \nu(n)$). This result is best possible since $\omega_{n-\nu(n)}(\mathbb{R}\mathbf{P}^{2^{i_1}} \times \cdots \mathbb{R}\mathbf{P}^{2^{i_{\nu(n)}}}) \neq 0$. In particular, $\mathbb{R}\mathbf{P}^{2^{i_1}} \times \cdots \times \mathbb{R}\mathbf{P}^{2^{i_{\nu(n)}}}$ cannot be immersed in $\mathbb{R}^N$ for $N < n - \nu(n)$.

It was thus conjectured that any n-manifold can be immersed in $\mathbb{R}^{n-\nu(n)}$. By 1963 E. H. Brown and F.P. Beterson had strengthened Massey's algebraic results, and in 1977 they proposed a program for proving this conjecture. In 1985, Cohen completed this program [3].

Chapter 6

Two to Two: Maps from the Plane to the Plane

In §6.1 we prove that smooth maps (of 2-manifolds) with only folds and cusps form an open dense set. We then give a normal form for folds. In §6.2 we give a normal form for simple cusps.

6.1 Folds

In §4, we established normal forms for a generic set of maps from an n-manifold down to $\mathbb{R}$. Then in §5, we showed that almost all maps of an n-manifold into a $2n$-manifold have no singularities at all. We now turn our attention to the next simplest case: maps from the plane to the plane. We will see that, generically, the singularities of such maps consist of general fold curves with isolated cusp points.

Recall the singularity set $\mathcal{S}_r \subset J^k(X, Y)$ of k-jets of corank r. In Theorem 3.2.8, we saw that $\mathcal{S}_r$ is a submanifold of codimension $(n-q+r)(m-q+r)$, where $n = \dim X$, $m = \dim Y$, and $q = min\{n, m\}$. For a smooth map f, let $\mathcal{S}_r(f) = (j^k f)^{-1}(\mathcal{S}_r) \subset X$ denote set of the points of X at which the k-jet of f has corank r. If $j^k f \pitchfork \mathcal{S}_r$ then

$\mathcal{S}_r(f)$ is a submanifold X of codimension $(n-q+r)(m-q+r)$ as well. Here we are using Theorems 3.2.4 and 3.1.3.

Now let $f : X \to Y$ be a map of 2-manifolds with $j^k f \pitchfork \mathcal{S}_1$. Since we are concerned with local structure of f at singularities, we may take $X = Y = \mathbb{R}^2$. From the above remarks, we compute

$$\begin{aligned} \dim J^1(\mathbb{R}^2, \mathbb{R}^2) &= \dim(\mathbb{R}^2 \times \mathbb{R}^2 \times B^1_{2,2}) = 10 \\ \text{codim } \mathcal{S}_1 = 1 \quad &\Longrightarrow \quad \dim \mathcal{S}_1 = 9 \\ \text{codim } \mathcal{S}_1(f) = 1 \quad &\Longrightarrow \quad \dim \mathcal{S}_1(f) = 1. \end{aligned}$$

Thus $\mathcal{S}_1(f)$ is a collection of smooth curves in X, called the the *general fold* of f. At each $p \in \mathcal{S}_1(f)$, exactly one of the following two situations must occur.

$$\begin{cases} (a) \quad T_p\mathcal{S}_1(f) \bigoplus \operatorname{Ker}(df)_p = T_pX; \\ (b) \quad T_p\mathcal{S}_1(f) = \operatorname{Ker}(df)_p. \end{cases}$$

Definition 6.1.1. Let $f : X \to Y$ be a map of 2-manifolds with $j^1 f \pitchfork S_1$. Then $p \in \mathcal{S}_1(f)$ satisfying (a) is called a *fold point* of f. And $p \in \mathcal{S}_1(f)$ satisfying (b) is called a *cusp point* of f.

The property that f has only folds and cusps is generic.

Proposition 6.1.1. *Let X and Y be smooth 2-manifolds. Then the set of maps $X \to Y$ with only fold and cusp singularities is an open dense subset of $\mathcal{C}^\infty(X, Y)$.*

Proof. We show that if f satisfies $j^1 f \pitchfork \mathcal{S}_1$ and $j^1 f \pitchfork \mathcal{S}_2$, then all critical points of f are folds and cusps. By Thom and the Baire Lemma, such f form an open dense subset of $\mathcal{C}^\infty(X, Y)$ (noting that $\mathcal{S}_1$ and $\mathcal{S}_2$ are closed submanifold).

Suppose $j^1 f \pitchfork \mathcal{S}_2$ and $j^1 f \pitchfork \mathcal{S}_1$. A critical point p of f is either in $\mathcal{S}_1(f)$ or $\mathcal{S}_2(f)$. But a non-empty $\mathcal{S}_2(f)$ would be a submanifold of codimension 4 in X. So p is in $\mathcal{S}_1(f)$ and, thus, a fold or cusp. $\square$

We now give a normal form for the structure of a fold singularity.

Theorem 6.1.2. *At a fold point, f is locally equivalent to the map $(x_1, x_2) \mapsto (x_1, \pm x_2^2)$.*

Proof. Let p be a fold point of f. Then $j^1 f \pitchfork S_1$ and we can assume $p = f(p) = 0$. Consider the restriction map $f|_{\mathcal{S}_1(f)}$ from the general fold of f to Y. As a fold point, 0 satisfies $T_0\mathcal{S}_1(f) \bigoplus \mathrm{Ker}(df)_0 = T_0X$. f has rank 1 at p, so $(df)_0$ must map $T_0\mathcal{S}_1(f)$ injectively into T_0Y. Thus $d(f|_{\mathcal{S}_1(f)})_0 = (df)_0|_{T_0\mathcal{S}_1(f)}$ is injective, and we conclude $f|_{\mathcal{S}_1(f)}$ is an immersion at 0. By the Immersion Lemma, the local image of $\mathcal{S}_1(f)$ is a smooth curve (this will not be the case at a cusp). In fact, since $f|_{\mathcal{S}_1(f)}$ is a local diffeomorphism onto its image, we have that f is locally equivalent to a smooth map $(x_1, x_2) \mapsto (x_1, g(x_1, x_2))$. Computing $(df)_0$, it is clear that $\mathcal{S}_1(f)$ is exactly the set on which $\frac{\partial g}{\partial x_2}$ vanishes. Thus $g(0) = \frac{\partial g}{\partial x_2}(0) = 0$. Applying Lemma 4.1.3 twice as in the proof of the Morse Lemma gives $g(x_1, x_2) = x_2^2 h(x_1, x_2)$ for some smooth function h.

If $h(0) = 0$, then a simple computation shows that f would have the same 2-jet at 0 as the map $\tilde{f}$ given by $(x_1, x_2) \mapsto (x_1, 0)$. Now the condition that $j^1 f \pitchfork_0 \mathcal{S}_1$ is a condition on $j^1 f(0)$ and $(dj^1 f)_0$, i.e. on the 2-jet of f at 0. So we would have $j^1\tilde{f} \pitchfork_0 \mathcal{S}_1$ as well and thus a submanifold $\mathcal{S}_1(\tilde{f})$ of X of codimension 1. This contradicts the fact that $\tilde{f}$ has rank 1 at every point of X.

Thus we have $g(x_1, x_2) = \pm x_2^2 h(x_1, x_2)$ with $h(0) > 0$, and the following is valid change of coordinates on some nbhd of 0.

$$\bar{x}_1 = x_1 \qquad \bar{x}_2 = x_2^2\sqrt{h(x_1, x_2)}$$

In these coordinates, f is the map $(\bar{x}_1, \bar{x}_2) \mapsto (\bar{x}_1, \pm\bar{x}_2^2)$. $\square$

Figure 6.1 illustrates why we call such a point a fold. We can visualize the map

$(x_1, x_2) \mapsto (x_1, x_2^2)$ by first mapping $\mathbb{R}^2$ to a parabolic cylinder in $\mathbb{R}^3$ via $(x_1, x_2) \mapsto (x_1, x_2, x_2^2)$ and then projecting onto the (x_1, x_3) plane. The projection creases the parabolic cylinder along the general fold $S_1(f)$.

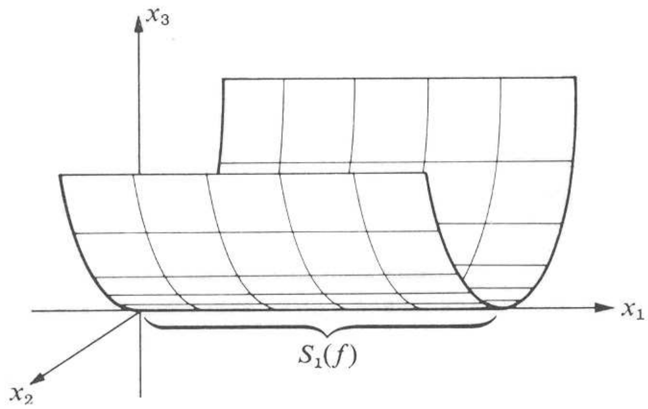

Figure 6.1: [5, p.89]

6.2 Cusps

Let $f : X \to Y$ be a map of 2-manifolds with $j^1 f \pitchfork S_1$. Recall that a cusp point p of f satisfies the cusp condition $T_p\mathcal{S}_1(f) = \mathrm{Ker}(df)_p$. We would like to find a normal form for cusps. This case is much more difficult than that of folds in §6.1. In fact, without a further restriction on p there are infinitely many locally non-equivalent forms that f could take at p. Fortunately, all but one of these forms almost never occur. In order to distinguish the generic case, consider a non-vanishing vector field ξ along $\mathcal{S}_1(f)$ (near p) such that at each point of $\mathcal{S}_1(f)$ ξ is in the kernel of (df). This is always possible locally. By the cusp condition, ξ is tangent to $\mathcal{S}_1(f)$ at p. The nature of the cusp at p depends on what order of contact ξ has with $\mathcal{S}_1(f)$ at p. To be precise, let

k be a smooth function on X with $k = 0$ on $\mathcal{S}_1(f)$ and $(dk)_p \neq 0$. Since $k = 0$ on $\mathcal{S}_1(f)$ and ξ is tangent to $\mathcal{S}_1(f)$ at p, the function $(dk)(\xi) : \mathcal{S}_1(f) \to \mathbb{R}$ has a zero at p.

Definition 6.2.1. We say p is a *simple cusp* if this zero is a simple zero.

Theorem 6.2.1. *At a simple cusp, f is locally equivalent to the map $(x_1, x_2) \mapsto x_1, x_1^3 + x_1x_2$.*

As with fold points, we can visualize simple cusps via a map to $\mathbb{R}^3$ and a projection:

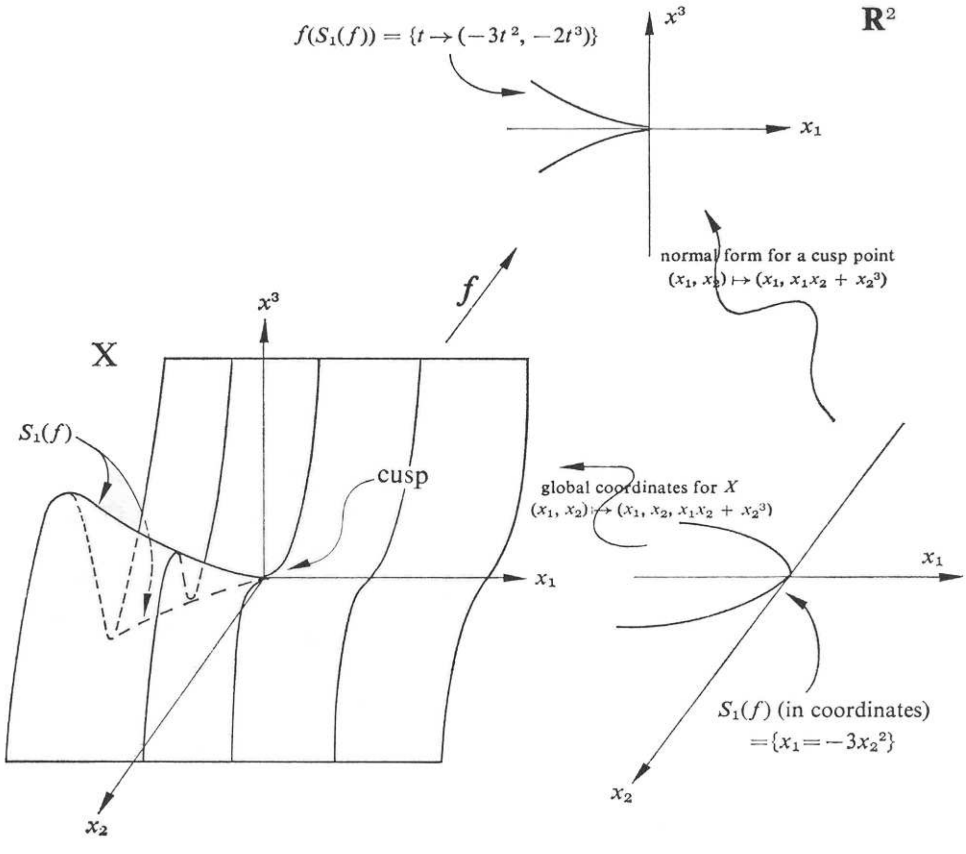

Figure 6.2: [5, p.147]

The general fold is a parabola, and the image of the cusp point is indeed a cusp of the image of the general fold. Whitney's original proof of Theorem 6.2.1 involves many lemmas followed by six pages of coordinate changes [20]. This approach, though successful for the cusp, degenerates into an intractable mess for higher dimensional singularities. In order to prove Theorem 6.2.1 in a clean and generalizable fashion, we will call upon a powerful extension of the Weierstrass Preparation Theorem of complex analysis (though it will be expressed in the language of commutative algebra). It is perhaps not entirely surprising that complex analysis comes into play, as a major obstacle to establishing normal forms for smooth real maps is that these maps are not necessarily real analytic (we cannot replace them with their Taylor series). This problem evaporates over the complex numbers, as smooth complex functions are necessarily complex analytic.

We will need some terminology. Let X be a smooth manifold. By $\mathcal{C}_p^\infty(X)$ we denote the ring of smooth germs of functions on X at p. $\mathcal{C}_p^\infty(X)$ is a local ring with maximal ideal $M_p(X)$ consisting of germs taking the value 0 at p. Note that $\mathbb{R} \cong \mathcal{C}_p^\infty(X)/M_p(X)$. Thus if A is a finite-dimensional $\mathcal{C}_p^\infty(X)$-module, then $A/M_p(X)A$ is a real vector space.

Theorem 6.2.2 (Generalized Malgrange Preparation Theorem). *Let X and Y be smooth manifolds and $f : X \to Y$ a smooth map with $f(p) = q$. Let A be a finitely generated $\mathcal{C}_p^\infty(X)$-module. Then A is a finitely generated $\mathcal{C}_q^\infty(Y)$-module iff $A/M_q(Y)A$ is a finite dimensional vector space over $\mathbb{R}$.*

Our proof of Theorem 6.2.1 will use Theorem 6.2.2 in the form of a more easily applicable corollary, though we will not deduce the corollary here. Define inductively a sequence of ideals $M_p^k(X)$ in $\mathcal{C}_p^\infty(X)$ by letting $M_p^1(X)$ be $M_p(X)$, and $M_p^k(X)$ be the vector space generated by germs of the form fg with f in $M_p(X)$ and g in $M_p^{k-1}(X)$. Using Lemma 4.1.3, it is not hard to show by induction that $M_0^k(X)$ consists precisely of those germs of smooth functions f whose Taylor series at 0 begin with terms of degree k.

Corollary 6.2.3. *If the projections of $e_1, ..., e_k$ form a spanning set of vectors in the vector space $A/(M_p^{k+1}(X)A + M_q(Y)A)$, then $e_1, ..., e_k$ form a set of generators for A as a $\mathcal{C}_q^\infty(Y)$-module.*

Proof of Theorem 6.2.1. Our proof follows that of Golubitsky and Guillemin [5, p.147]. Since f has rank 1 at p, through a linear change of coordinates on the domain we can assume $\frac{\partial f_1}{\partial x_1} \neq 0$ and then apply the Immersion Theorem to force f into the form $(x_1, h(x_1, x_2))$ with h smooth. We can also assume $(df)_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ in this coordinate system, i.e. $\frac{\partial h}{\partial x_1} = \frac{\partial h}{\partial x_2} \neq 0$.

We now show that $d(\frac{\partial h}{\partial x_2})_0 \neq 0$. For suppose otherwise, i.e. that

$$\frac{\partial}{\partial x_1}\frac{\partial h}{\partial x_2}(0) = \frac{\partial}{\partial x_2}\frac{\partial h}{\partial x_2}(0) = 0.$$

Then f has the same 2-jet as the map $(x_1, x_2) \mapsto (x_1, \delta x_2)$ where $\delta = \frac{1}{2}\frac{\partial^2 h}{\partial x_1^2}(0)$. As in the proof of Theorem 6.1.2, we note that the latter map is of rank 1 everywhere and thus its 1-jet is not transverse to $\mathcal{S}_1$ at 0. But this condition depends only on the 2-jet of the map, contradicting that $j^1 f \pitchfork_0 \mathcal{S}_1$.

As in the fold case, $\mathcal{S}_1(f)$ is defined by the equation $\frac{\partial h}{\partial x_2} = 0$, and so at each point of $\mathcal{S}_1(X)$ the kernel of (df) is spanned by $\frac{\partial h}{\partial x_2}$. Thus $\frac{\partial h}{\partial x_2}$ meets the requirement of our vector field ξ in the definition of simple cusp, while $\frac{\partial h}{\partial x_2}$ suffices for our function k because $\frac{\partial h}{\partial x_2}(0) = 0$ and $d(\frac{\partial h}{\partial x_2})_0 \neq 0$. The cusp condition is $\frac{\partial^2 h}{\partial x_2^2}(0) = 0$, while the simple cusp condition is $\frac{\partial^3 h}{\partial x_2^3}(0) \neq 0$. Therefore, at the origin, we have

$$h = \frac{\partial h}{\partial x_2} = \frac{\partial^2 h}{\partial x_2^2} = 0 \text{ and } \frac{\partial^3 h}{\partial x_2^3} \neq 0.$$

We are now in a position to apply the corollary to the Malgrange Preparation Theorem.

Recall that f is given by $f(x_1, x_2) = (x_1, h(x_1, x_2))$. Via f, $\mathcal{C}_0^\infty(\mathbb{R}^2)$ becomes a module over itself: $a \cdot b(x_1, x_2) = a(f(x_1, x_2))b(x_1, x_2)$ where a is in the ring $\mathcal{C}_0^\infty(\mathbb{R}^2)$ and b is in the module $\mathcal{C}_0^\infty(\mathbb{R}^2)$. By Corollary 6.2.3, this module is generated by 1, x_2, and x_2^2 if the vector space $\mathcal{C}_0^\infty(\mathbb{R}^2)/(M_0^4(\mathbb{R}^2) + (x_1, h))$ is generated by 1, x_2, and x_2^2. (In the corollary's notation, $A = \mathcal{C}_0^\infty(\mathbb{R}^2)$, $M_p(X) = (x_1, h)$, and $M_q(X) = (x_1, x_2)$). The conditions on h guarantee that $(x_1, h) \supset M_0^3(\mathbb{R}^2)$. (This step makes evident why it is necessary for the cusp to be simple in order to apply our corollary and arrive at a simple normal form). Thus $M_0^3(\mathbb{R}^2) \supset (M_0^4(\mathbb{R}^2) + (x_1, h))$ and the the vector space $\mathcal{C}_0^\infty(\mathbb{R}^2)/(M_0^4(\mathbb{R}^2) + (x_1, h))$ is indeed spanned by 1, x_2, and x_2^2.

Since our module is generated by 1, x_2, and x_2^2, we can express x_2^3 as

$$x_2^3 = 3a_2(x_1, h)x_2^2 + a_1(x_1, h)x_2 + a_0(x_1, h).$$

where the a_i are smooth functions of $(y_1, y_2) = (x_1, h)$ vanishing at 0. We can rewrite this equation in the form

$$(x_2 - a)^3 + b(x_2 - a) = c$$

with $a = a_2$, and b and c new functions of (y_1, y_2) vanishing at 0. If we set $x_1 = 0$ in this equation, the left hand side takes the form $x_2^3 + ...$, the dots indicating terms of order > 3 in x_2. Since $h(0, x_2) = x_2^3 + ...$ as well, the right and left hand sides of the equation can only be equal if $\frac{\partial c}{\partial y_2}(y_1, y_2) \neq 0$ at 0. Now note that the leading term of the Taylor series of h is a non-zero multiple of x_1x_2; so comparing the linear terms on either side of the equation gives $\frac{\partial b}{\partial y_1} \neq 0$ at 0, while comparing the quadratic terms gives $\frac{\partial c}{\partial y_1} = 0$ at 0. Taken together, these results show that the following are legitimate coordinate changes:

$$\begin{cases} \bar{x}_1 = b(x_1, h) \\ \bar{x_2} = x_2 - a(x_1, h) \end{cases} \qquad \begin{cases} \bar{y}_1 = b(y_1, y_2) \\ \bar{y_2} = c(y_1, y_2) \end{cases}$$

In these coordinates f takes the normal form $(\bar{x}_1, \bar{x}_2) \mapsto (\bar{x}_1, \bar{x}_1^3 + \bar{x}_1\bar{x}_2)$. □

It is evident from Figure 6.2 that simple cusps are isolated. More rigorously, we saw that with $h = x_2^3 - x_1x_2$, the cusp condition on $f = (x_1, h)$ requires $\frac{\partial^2 h}{\partial x_2^2}(0) = 0$, while the simple cusp condition requires $\frac{\partial^3 h}{\partial x_2^3}(0) \neq 0$. Thus f has no other cusps in a sufficiently small nbhd of a simple cusp.

With more effort, we could extend Proposition 6.1.1 to the following claim: the set of smooth maps with only folds and simple cusps is open and dense. We will refrain however, as a proper attack would lead us into the territory of higher order analogs of the singularity sets $\mathcal{S}_r$. In any case, the reader who has made it this far ought to be convinced of the truth of the claim.

Bibliography

[1] Michael Artin, *Algebra*, Prentice-Hall, Upper Saddle River, New Jersey, 1991.

[2] Garrett Birkhoff and Saunders Mac Lane, *A Survey of Modern Algebra*, 4th Ed., MacMillan Publishing Co., New York, 1979. (First Ed. 1941).

[3] Ralph L. Cohen, *The Immersion Conjecture for Differentiable Manifolds*, Annals of Math., 2nd Ser., Vol. 122, No. 2 (Sep., 1985), 237–328.

[4] C. G. Gibson, *Singular Points of Smooth Mappings*, Pitman Publishing, London, 1979.

[5] Martin Golubitsky and Victor Guillemin, *Stable Mappings and Their Singularities*, Springer-Verlag, New York, 1973. (Second Printing 1980).

[6] Victor Guillemin and Alan Pollack, *Differential Topology*, Prentice Hall, Englewood Cliffs, New Jersey, 1974.

[7] Serge Lang, *Differential Manifolds*, Addison Wesley Publishing Co., Reading, Massachusetts, 1972.

[8] Yung-Chen Lu, *Singularity Theory and an Introduction to Catastrophe Theory*, Springer-Verlag, New York, 1976.

[9] John N. Mather, *Stability of $\mathcal{C}^\infty$ Mappings: II. Infinitesimal Stability Implies Stability*, Annals of Math., 2nd Ser., Vol. 89, No. 2 (Mar., 1969), 254–291.

[10] John Milnor, *Morse Theory*, Princeton University Press, Princeton, New Jersey, 1963. (Second Printing 1965).

[11] ______, *Topology from the Differential Viewpoint*, University of Virgina Press, Charlottesville, Virginia, 1965. (Fourth Printing 1976).

[12] James R. Munkres, *Topology: A First Course*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[13] D. J. Saunders, *The Geometry of Jet Bundles*, London Math. Soc. Lecture Note Series 142, Cambridge University Press, Cambridge, 1989.

[14] Michael Spivak, *Calculus on Manifolds*, Addison-Wesley Publishing Co., New York, 1965.

[15] Shlomo Sternberg, *Lectures on Differential Geometry*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

[16] René Thom, *Structural Stability and Morphogenesis*, W. A. Benjamin, Inc., Reading, Massachusetts, 1975.

[17] Hassler Whitney, *Differentiable Manifolds*, Annals of Math., 2nd Ser., Vol. 37, No. 3 (Jul., 1963), 645–680.

[18] ______, *The Self-Intersection of a Smooth n-Manifold in $2n$-Space*, Annals of Math., 2nd Ser., Vol. 45, No. 2 (Apr., 1944), 220–246.

[19] ______, *The Singularities of a Smooth n-Manifold in $2n-1$-Space*, Annals of Math., 2nd Ser., Vol. 45, No. 2 (Apr., 1944), 247–293.

[20] ______, *On Singularities of Mappings of Euclidean Spaces. I. Mappings of the Plane into the Plane*, Annals of Math., 2nd Ser., Vol. 62, No. 3 (Nov., 1955), 374–410.