PMATH 465 Notes

Smooth Manifolds

Fall 2025

Based on Professor Ruxandra Moraru's Lectures

Contents

1 Smooth Manifolds

2 Smooth Maps

3 Tangent and Cotangent Spaces

3.1 Derivative of a Function

Let M be a manifold and $a \in M$. Given a smooth map $f : M \to \mathbb{R}$, how can we define the derivative of f at a? We want to pass it to $\mathbb{R}^n$ using charts and use the derivative in $\mathbb{R}^n$ to define it.

Suppose $M = \mathbb{R}^n$ with coordinates $(x_1, \cdots, x_n)$. Recall that

$$\frac{\partial f}{\partial x_i}(a) := \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}$$

is the **partial derivative of f at** a with respect to $(x_1, \cdots, x_n)$. Then

$$Df(a) := \begin{bmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_n}(a) \end{bmatrix}$$

is called the **derivative of f at** a. Note that if $g : M \to \mathbb{R}$ is another map, then

$$Df(a) = Dg(a) \iff D(f - g)(a) = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}$$

Construction 3.1. Now let M be a manifold of dimension n. Let (U, φ) be a chart of M with $a \in U$. Then the map

$$f \circ \varphi^{-1} : \varphi(U) \subseteq \mathbb{R}^n \to \mathbb{R} \;\text{ is a smooth map}$$

in the sense of Calculus 3. Let $\varphi = (x_1, \cdots, x_n) : M \to \mathbb{R}^n$ with coordinates x_i. Then

$$\frac{\partial}{\partial x_i}(f \circ \varphi^{-1})(\varphi(a)) \;\text{ exist for all } i = 1, \cdots, n$$

But this value may depend on the choice of chart (U, φ). However, we will show the following claim.

Claim. If there exists a chart (U, φ) with $a \in U$ and $\varphi = (x_1, \cdots, x_n)$ and

$$\frac{\partial}{\partial x_i}(f \circ \varphi^{-1})(\varphi(a)) = 0 \;\text{ for all } i = 1, \cdots, n$$

Then for ANY other chart (V, ψ) with $a \in V$ and $\psi = (y_1, \cdots, y_n)$ we have

$$\frac{\partial}{\partial y_i}(f \circ \psi^{-1})(\psi(a)) = 0 \;\text{ for all } i = 1, \cdots, n$$

Hence the partial derivatives is well-defined up to a function with zero derivative.

Proof of Claim. Let (V, ψ) be another chart with $a \in V$. Let

$$g = f \circ \varphi^{-1} : \psi(V) \to \mathbb{R} \;\text{ and }\; h = f \circ \psi^{-1} : \varphi(U) \to \mathbb{R}$$

It follows that $h = f \circ \psi^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1}) = g \circ (\varphi \circ \psi^{-1})$ is a composition of smooth map g and diffeomorphism $\varphi \circ \psi^{-1}$. Therefore h is smooth. By Chain Rule we have

$$Dh(\psi(x)) = Dg(\varphi(x))D(\varphi \circ \psi^{-1})(\psi(x))$$

since $h(\psi(x)) = g((\varphi \circ \psi^{-1})(\psi(x))) = g(\varphi(x))$. Note that

$$F = \psi \circ \psi^{-1} : \psi(U \cap V) \to \psi(U \cap V)$$

is a diffeomorphism. So F is smooth and F^{-1} is smooth. Also

$$F \circ F^{-1} = \mathrm{id}_{\varphi(U \cap V)} \quad \text{and} \quad F^{-1} \circ F = \mathrm{id}_{\psi(U \cap V)}$$

Therefore $DF \circ DF^{-1} = D\mathrm{id}_{\varphi(U \cap V)} = I_n$. Therefore DF is invertible on $\psi(U \cap V)$ and

$$Dh(\psi(x)) = 0 \iff Dg(\varphi(x)) = 0$$

This completes the proof. This allows us to define derivative as an equivalence class.

Definition. Let $f : M \to \mathbb{R}$ be a smooth function. We say f has **zero derivative at** $a \in M$ if $D(f \circ \psi^{-1})(\varphi(a)) = 0$ for any chart (U, φ) of M with $a \in U$. Let

$$\mathcal{Z}_a := \{f \in \mathcal{C}^\infty(M) : f \text{ has zero derivatvie at } a\} \subseteq \mathcal{C}^\infty(M)$$

Definition. Let $a \in M$. The **cotagent space of** M **at** a is defined as

$$T_a^* M := \mathcal{C}^\infty(M)/\mathcal{Z}_a$$

as quotient of $\mathbb{R}$-vector spaces. This is clearly a vector space as $\mathcal{Z}_a$ is a subspace of $\mathcal{C}^\infty(M)$. For a smooth map $f \in \mathcal{C}^\infty(M)$ we denote

$$(df)_a := [f]_{\mathcal{Z}} = f + \mathcal{Z}_a = \text{equivalence class of } f \text{ in } T_a^* M$$

In particular for $f, g \in \mathcal{C}^\infty(M)$ we have

$$\begin{aligned}(df)_a = (dg)_a &\iff (f - g) \in \mathcal{Z}_a \\ &\iff f - g \text{ has zero derivative at } a \\ &\iff f = g + h \text{ for some } h \in \mathcal{Z}_a\end{aligned}$$

Definition. For $f \in \mathcal{C}^\infty(M)$, we call $(df)_a$ the **derivative of** f **at** a.

Proposition 3.2. Let M be an n-dimensional smooth manifold and $a \in M$, then

1. The cotangent space $T_a^* M$ is an n-dimensional $\mathbb{R}$-vector space.

2. If (U, φ) is any chart on M with $a \in U$ and $\varphi = (x_1, \cdots, x_n)$ then

$$\mathcal{B} = \{(dx_1)_a, \cdots, (dx_n)_a\}$$

is a basis for $T_a^* M$.

3. For all $f \in \mathcal{C}^\infty(M)$ we have

$$(df)_a = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(f \circ \varphi^{-1})(\varphi(a))(dx_i)_a$$

Note that by part 3, the vector representation of $(df)_a$ with respect to $\mathcal{B}$ is

$$\nabla(f \circ \varphi^{-1})(\varphi(a)) = \left(\frac{\partial}{\partial x_1}(f \circ \varphi^{-1})(\varphi(a)), \cdots, \frac{\partial}{\partial x_n}(f \circ \varphi^{-1})(\varphi(a))\right)$$

Proof.

Definition. Let $a \in M$. The **tangent space of** M **at** a is

$$T_a M := (T_a^* M)^* = \{\mathbb{R}\text{-linear maps } T_a^* M \to \mathbb{R}\}$$

the dual space of the cotagent space $T_a^* M$. Then $T_a M$ is an n-dimensional $\mathbb{R}$-vector space as well. Also, if (U, φ) is a chart with $a \in U$ and $\varphi = (x_1, \cdots, x_n)$ then

$$\left\{ \left(\frac{\partial}{\partial x_1} \right)_a, \cdots, \left(\frac{\partial}{\partial x_n} \right)_a \right\}$$

is the dual basis of $\{(dx_1)_a, \cdots, (dx_n)_a\}$ such that

$$\left(\frac{\partial}{\partial x_j} \right)_a ((dx_i)_a) = \delta_{ij}$$

for all $i, j \in \{1, \cdots, n\}$.

3.2 Derivations

Let M be a smooth manifold and let $a \in M$.

Definition. A **derivation at** a is an $\mathbb{R}$-linear map $X_a : \mathcal{C}^\infty(M) \to \mathbb{R}$ satisfying the **Leibniz rule**

$$X_a(fg) = g(a) X_a(f) + f(a) X_a(g)$$

for all $f, g \in \mathcal{C}^\infty(M)$.

Example. If $M = \mathbb{R}^n$ with coordinates $(x_1, \cdots, x_n)$, then the partial derivatives $\frac{\partial}{\partial x_i}|_a$ at a are derivations at a. Indeed, partial differentiation is $\mathbb{R}$-linear and

$$\frac{\partial}{\partial x_i}\bigg|_a (fg) = g(a) \frac{\partial}{\partial x_i}\bigg|_a (f) + f(a) \frac{\partial}{\partial x_i}\bigg|_a (g)$$

for all $f, g \in \mathcal{C}^\infty(M)$.

Theorem 3.3. Let M and N be smooth manifolds and $F : M \to N$ be a smooth map. Let $c \in N$ and $F^{-1}(c) = \{a \in M : F(a) = c\}$. If $DF_a : T_aM \to T_cN$ is surjective for all $a \in F^{-1}(c)$, then $F^{-1}(c)$ is a smooth manifold of dimension $\dim M - \dim N$.

Corollary 3.4. For all $a \in F^{-1}(c)$, we have $T_aF^{-1}(c) = \ker DF_a$.

Proof. Let $X_a \in T_aF^{-1}(c)$. Then we have $X_a = \gamma'(0)$ for some smooth curve $\gamma : (-\epsilon, \epsilon) \to F^{-1}(c)$ with $\gamma(0) = a$. Recall, by definition $\gamma'(0) = \gamma_{*,0}\left(\frac{d}{dt}|_{t=0}\right)$. So we have

$$\begin{aligned} DF_a(X_a) = DF_a(\gamma'(0)) &= F_{*,a}\left(\gamma_{*,0}\left(\frac{d}{dt}\bigg|_{t=0}\right)\right) \\ &= F_{*,a} \circ \gamma_{*,0}\left(\frac{d}{dt}\bigg|_{t=0}\right) = (F \circ \gamma)_{*,0}\left(\frac{d}{dt}\bigg|_{t=0}\right) \end{aligned}$$

Note that $\tilde{\gamma} := F \circ \gamma : (-\epsilon, \epsilon) \to N$ via $t \mapsto F(\gamma(t)) = c$ is a constant curve so that $\tilde{\gamma}'(0) = 0$. Hence

$$DF_a(X_a) = \tilde{\gamma}_{*,0}\left(\frac{d}{dt}\bigg|_{t=0}\right) = 0$$

This implies that $T_aF^{-1}(c) \subseteq \ker DF_a$. By the theorem $T_aF^{-1}(c)$ has dimension $(\dim M - \dim N)$. Since DF_a is surjective, we know (by rank-nullity)

$$\dim \ker DF_a = \dim T_aM - \dim T_cN = \dim T_aM - \dim N$$

Since $\dim T_aF^{-1}(c) = \dim \ker DF_a$, it follows that $T_aF^{-1}(c) = \ker DF_a$. $\square$

Example. Let $S^n = \{x \in \mathbb{R}^{n+1} : \|x\|^2 = 1\} = \{(x_1, \cdots, x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + \cdots + x_{n+1}^2 = 1\}$. Note that $S^n = F^{-1}(1)$ is the level set of the smooth function

$$F : \mathbb{R}^{n+1} \to \mathbb{R} \quad \text{by} \quad F(x) = \|x\|^2 = x_1^2 + \cdots + x_{n+1}^2$$

Note that $DF = (2x_1, \cdots, 2x_{n+1})$. At $a = (a_1, \cdots, a_{n+1}) \in S^n$ we have $DF_a = 2a$.

$$DF_a : T_a\mathbb{R}^{n+1} \to T_1\mathbb{R} \quad \text{by} \quad v \mapsto DF_a(v) = 2a \cdot v$$

Then DF_a is surjective because it is nonzero and $\dim T_1\mathbb{R} = 1$. Hence S^n is a smooth manifold of dimension $(n+1) - 1 = n$. Also,

$$\begin{aligned} T_aS^n &= \{v \in \mathbb{R}^{n+1} : DF_a(v) = 2a \cdot v = 0\} \\ &= \{v \in \mathbb{R}^{n+1} : a \cdot v = 0\} \end{aligned}$$

So the tangent space at a is all vectors in $\mathbb{R}^{n+1}$ that is orthogonal to a.

Example. Let $O(n) = \{A \in M_n(\mathbb{R}) : AA^T = I\}$ be the orthogonal group. Then $O(n)$ is a manifold of dimension $\frac{n^2-n}{2}$ and for all $A \in O(n)$ we have

$$T_A O(n) = \{H \in M_n(\mathbb{R}) : HA^T + AH^T = 0\}$$

Proof. Let $F : M_n(\mathbb{R}) = \mathbb{R}^{n^2} \to \mathrm{Sym}(\mathbb{R}) = \mathbb{R}^{\frac{n^2+n}{2}}$ by $A \mapsto AA^T$, where

$$\mathrm{Sym}(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A^T = A\}$$

Clearly F is a smooth map (because it is defined by polynomials in the entries of A). Then

$$O(n) = F^{-1}(I)$$

Pick $A \in F^{-1}(I)$ so that $F(A) = I$. Let's compute $DF_A : T_A M_n(\mathbb{R}) \to T_I \mathrm{Sym}(\mathbb{R})$. Recall that $T_A M_n(\mathbb{R}) = M_n(\mathbb{R})$ because for any $H \in M_n(\mathbb{R})$ we can write $H = \gamma'(0)$ where $\gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ by $\gamma(t) = A + tH$ so that $\gamma(0) = A$. Also, for all $H \in T_A M_n(\mathbb{R})$

$$DF_A(H) = D_A(\gamma'(0)) = DF_A\left(\gamma_{*,0}\left(\frac{d}{dt}\Big|_{t=0}\right)\right) = (F \circ \gamma)_{*,0}\left(\frac{d}{dt}\Big|_{t=0}\right)$$

Set $\tilde{\gamma} = F \circ \gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ with

$$\tilde{\gamma}(t) = F(\gamma(t)) = F(A + tH) = (A + tH)(A + tH)^T$$

Hence we have

$$\begin{aligned} DF_A(H) = \tilde{\gamma}_{*,0}\left(\frac{d}{dt}\Big|_{t=0}\right) = \tilde{\gamma}'(0) = \frac{d}{dt}\Big|_{t=0}(\tilde{\gamma}(t)) \\ &= \frac{d}{dt}\Big|_{t=0}\left((A + tH)(A + tH)^T\right) \\ &= \frac{d}{dt}\Big|_{t=0}\left((A + tH)(A^T + tH^T)\right) \\ &= H(A^T + tH^T) + (A + tH)H^T\Big|_{t=0} \\ &= HA^T + AH^T \end{aligned}$$

Hence $DF_A : T_A M_n(\mathbb{R}) \to T_I \mathrm{Sym}(\mathbb{R})$ maps $H \mapsto HA^T + AH^T$. [Exericse: $T_I \mathrm{Sym}(\mathbb{R}) = \mathrm{Sym}(\mathbb{R})$]. We claim that DF_A is surjective for all $A \in O(n)$. Let $B \in T_I \mathrm{Sym}(\mathbb{R})$ so that $B^T = B$. It's trivial to YING COU that

$$DF_A\left(\frac{BA}{2}\right) = \left(\frac{BA}{2}\right)A^T + A\left(\frac{BA}{2}\right)^T = \frac{B}{2}AA^T + AA^T\left(\frac{B^T}{2}\right) = \frac{B}{2} + \frac{B^T}{2} = B$$

Hence DF_A is surjecitve for $A \in O(n)$. Therefore $O(n)$ is a manifold of dimension $n^2 - \frac{n^2+n}{2} = \frac{n^2-n}{2}$ and for all $A \in O(n)$ we have

$$T_A O(n) = \ker DF_A = \{H \in M_n(\mathbb{R}) : DF_A = HA^T + AH^T = 0\}$$

In particular when $A = I$ we have $T_I O(n) = \{H \in M_n(\mathbb{R}) : H + H^T = 0\}$, the set of all skew-symmetric matrices.

Definition. The **Lie algebra** of a Lie group G is $T_e G$, where e is the identity of G. The Lie algebra of G is denoted by $\mathfrak{g}$ or $\mathrm{Lie}(G)$.

Example. We know $O(n)$ is a Lie group with identity I and the Lie algebra is

$$\mathfrak{o}(n) = \mathrm{Lie}(O(n)) = T_I O(n) = \{H \in M_n(\mathbb{R}) : H^T = H\}$$

by the example above.

3.3 Tangent and Cotangent Bundles

So far we are looking at $T_a M$ for each $a \in M$ separately. We want to put them all together.

Definition. Let M be a smooth manifold. Let TM be the disjoint union of $T_a M$ for $a \in M$

$$TM := \bigsqcup_{a \in M} T_a M$$

This is called the **tangent bundle** of M. Similalry define

$$T^*M = \bigsqcup_{a \in M} T_a^* M$$

to be the **cotangent bundle** of M.

Theorem 3.5. Let M be a smooth manifold of dimension n. Then both TM and T^*M are smooth manifolds of dimension $2n$.

Proof. Let $\{(U_\alpha, \varphi_\alpha)\}$ be any smooth atlas on M so that

$$\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) \subseteq \mathbb{R}^n \text{ by } a \mapsto \varphi_\alpha(a) = (x_1^\alpha(a), \cdots, x_n^\alpha(a))$$

Then for all $a \in U_\alpha$ we have

$$T_a M = \mathrm{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x_i^\alpha}\right|_a : i = 1, \cdots, n \right\} \text{ and } T_a^* M = \mathrm{span}_{\mathbb{R}} \{(dx_i^\alpha)_a : i = 1, \cdots, n\}$$

Let $W_\alpha = \bigcup_{a \in U_\alpha} T_a M$ be the union of all tangent spaces at $a \in U_\alpha$. Similarly $W_\alpha^* = \bigcup_{a \in U_\alpha} T_a^* M$. Then we have

$$\bigcup_\alpha W_\alpha = TM \text{ and } \bigcup_\alpha W_\alpha^* = T^*M$$

Also we let $\psi_\alpha : W_\alpha = \bigcup_{a\in U_\alpha} T_aM \to \varphi_\alpha(U_\alpha)\times\mathbb{R}^n \subseteq \mathbb{R}^{2n}$ by

$$X_a = \sum_{i=1}^n c_i \frac{\partial}{\partial x_i}\Big|_a \mapsto (\varphi_\alpha(a),(c_1,\cdots,c_n))$$

Similarly define $\chi_\alpha : W_\alpha^* = \bigcup_{a\in U_\alpha} T_a^*M \to \varphi_\alpha(U_\alpha)\times\mathbb{R}^n \subseteq \mathbb{R}^{2n}$ by

$$(df)_a = \sum_{i=1}^n c_i(dx_i)_a \mapsto (\varphi_\alpha(a),(c_1,\cdots,c_n))$$

4 Immersions, Embeddings, Submersions and Submanifolds

4.1 Immersions, Embeddings, Submersions

Definition. Let M and N be smooth manifolds of dimension m and n, respectively. Conisder a smooth map $F : M \to N$. Then for all $p \in M$, the map $DF_p : T_pM \to T_{F(p)}N$ is an $\mathbb{R}$-linear map.

1. If DF_p is injective, we say F is an **immersion at** p.
2. If DF_p is surjecitve, we say F is a **submersion at** p.

If F is a immersion/submersion at every point on an open set $U \subseteq M$, we say F is an immersion/submersion on M. In particular, if $U = M$ then we simply say F is an immersion/submersion.

Remark. If F is an immersion at p, then $\dim T_pM \leq \dim T_{F(p)}N$. Hence $\dim M \leq \dim N$. Similarly if F is a submersion at p we have $\dim M \geq \dim N$.

Example. Diffeomorphisms are both immersions and submersions on M because DF_p is an isomorphism for all $p \in M$.

Example. Let $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$ with $m < n$. Let $\iota : \mathbb{R}^n \to \mathbb{R}^m$ be the inclusion map.

$$D\iota = \begin{bmatrix} I_m \\ \mathcal{O}_{n-m} \end{bmatrix}$$

In this case, the rank of $D\iota$ is m, so $D\iota$ is injective (full column rank), so $D\iota$ is injective.

Example. Let $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$ with $m > n$. Let $\pi : \mathbb{R}^m \to \mathbb{R}^n$ be the projection map onto the first n coordinates. Then

$$D\pi = \begin{bmatrix} I_n & \mathcal{O}_{m-n} \end{bmatrix}$$

It has rank n (full row rank), so $D\pi$ is surjecitve.

Theorem 4.1 (Canonical Immersion Theorem). Let $F : M \to N$ be an immersion at $p \in M$. Then $\dim M \leq \dim N$ and there exist charts (U, φ) of M and (V, ψ) of N with $p \in U$ and $F(p) \in V$ such that

$$\psi \circ F \circ \varphi^{-1} = \iota|_{\varphi(U)}$$

where $\iota : \mathbb{R}^m \to \mathbb{R}^n$ is the inclusion map.

Theorem 4.2 (Canonical Submersion Theorem). Let $F : M \to N$ be an submersion at $p \in M$. Then $\dim M \geq \dim N$ and there exist charts (U, φ) of M and (V, ψ) of N with $p \in U$ and $F(p) \in V$ such that

$$\psi \circ F \circ \varphi^{-1} = \pi|_{\varphi(U)}$$

where $\pi : \mathbb{R}^m \to \mathbb{R}^n$ is the projection map.

These two theorems follow from the *Constant Rank Theorem*.

Definition. A smooth map $F : M \to N$ is said to have **constant rank at** $p \in M$ if there is a open neighborhood U with $p \in U$ such that $\text{rank}(DF_q) = \text{rank}(DF_r)$ for all $q, r \in U$. In other words, the derivative matrix at points of U all have the same rank.

Example. Diffeomorphisms $M \to N$ have constant rank everywhere because DF_p is an isomorphism for all $p \in M$, so $\text{rank}(DF_p) = \dim M$ for all $p \in M$. Take $U = M$ as in the definition.

Example. Let $m < n$. The canonical inclusion $\iota : \mathbb{R}^m \to \mathbb{R}^n$ has constant rank m and the canonical projection $\iota : \mathbb{R}^n \to \mathbb{R}^m$ has constant rank m.

Proposition 4.3. If $F : M \to N$ is an immersion at $p \in M$, then DF has constant rank $m = \dim M$ in an open neighborhood W of p. Hence F has constant rank m at p.

Proof. Suppose F is an immersion at p. Then $DF_p : T_pM \to T_{F(p)}N$ is injective. Let (U, φ) and (V, ψ) be charts of M, N respectively with $p \in U$ and $F(p) \in V$. Then

$$J := \text{Jac}(\psi \circ F \circ \varphi^{-1})(\varphi(p))$$

is the Jacobian matrix of F at p, which is the matrix representation of DF_p with respect to the bases of T_pM and $T_{F(p)}N$

$$\left\{ \frac{\partial}{\partial x_i}\bigg|_p : 1 \leq i \leq m \right\} \quad \text{and} \quad \left\{ \frac{\partial}{\partial y_j}\bigg|_{F(p)} : 1 \leq j \leq n \right\}$$

where $\varphi = (x_1, \cdots, x_m)$ and $\psi = (y_1, \cdots, y_n)$. Since DF_p is injective, we know that

$$\text{rank}(\text{Jac}(\psi \circ F \circ \varphi^{-1})(\varphi(p))) = \dim T_pM = m$$

Hence there exists a $m \times m$ manior A of J such that $\det A(\varphi(p)) \neq 0$. But $\det A$ is a smooth function on $\varphi(U)$. So, since $\det A(\varphi(p)) \neq 0$ then $\det A \neq 0$ on an open neighborhood $\tilde{W}$ of $\varphi(p)$ in $\varphi(U) \subseteq \mathbb{R}^m$. Set $W = \varphi^{-1}(\tilde{W}) \subseteq M$. Then $p \in W$ and DF has rank m on W. □

Proposition 4.4. If $F : M \to N$ is an submersion at $p \in M$, then DF has constant rank $n = \dim N$ in an open neighborhood W of p. Hence F has constant rank n at p.

Proof. Same as above. □

Theorem 4.5 (Constant Rank Theorem). Suppose that $F : M^m \to N^n$ is a smooth map that has constant rank r on an open neighborhood of U of $p \in M$ so that $DF_q : T_qM \to T_{F(q)}N$ has rank r for all $q \in U$. Then there exists charts (U, φ) of M and (V, ψ) of N with $p \in U$ and $F(p) \in V$ such that the map $\psi \circ F \circ \varphi^{-1} : \varphi(U) \subseteq \mathbb{R}^m \to \psi(V) \subseteq \mathbb{R}^n$ sends

$$(x_1, \cdots, x_r, x_{r+1}, \cdots, x_m) \mapsto (x_1, \cdots, x_r, 0, \cdots, 0)$$

where $\varphi = \varphi(x_1, \cdots, x_n)$. Note that $r \leq m$ because $\dim T_pM = m$.

In other words, this means we can always choose charts so that $\psi \circ F \circ \varphi^{-1}$ looks like a projection map onto the first r coordinates! If F is immersion then $r = m$ so this is exactly the inclusion. If F is a submersion then $r = n$ so this is the projection!

Proof. Let us first assume that $M \subseteq \mathbb{R}^m$ and $N \subseteq \mathbb{R}^n$ are open. Hence

$$F : M \subseteq \mathbb{R}^m \to N \subseteq \mathbb{R}^n \text{ by } x = (x_1, \cdots, x_m) \mapsto (F_1(x), \cdots, F_n(x)) = (y_1, \cdots, y_n)$$

and we know the Jacobian matrix $DF = \left[\frac{\partial F_i}{\partial x_j}\right]$ has rank r near p. This means there are r linearly independent rows. After possibily permuting the variables $y_1, \cdots, y_n$, we can assume the first r rows are linearly independent near p. Hence

$$S := \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial F_r}{\partial x_1} & \cdots & \frac{\partial F_r}{\partial x_n} \end{bmatrix} \text{ has rank } r$$

Plus, after possibly permuting the $x_1, \cdots, x_n$ we can assume that the first columns of S are linearly independent. Define the map $\varphi : \mathbb{R}^m \to \mathbb{R}^m$ by $(x_1, \cdots, x_m) \mapsto (F_1(x), \cdots, F_r(x), x_{r+1}, \cdots, x_m)$. Then we have that

$$D\varphi = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \cdots & \frac{\partial F_1}{\partial x_{r+1}} & \cdots & \cdots & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \frac{\partial F_r}{\partial x_1} & \cdots & \cdots & \frac{\partial F_r}{\partial x_{r+1}} & \cdots & \cdots & \cdots & \frac{\partial F_r}{\partial x_n} \\ 0 & \cdots & 0 & 1 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix} = \begin{bmatrix} & S & \\ \mathcal{O}_{m-r\times r} & & I_{m-r\times n-r} \end{bmatrix}$$

By the Inverse Function Theorem, φ is locally invertible with smooth inverse φ^{-1}. Set

$$F \circ \varphi^{-1}(x) = (x_1, \cdots, x_r, B(x_1, \cdots, x_r))$$

and set $\psi : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\psi(\underbrace{y_1, \cdots, y_r}_{u}, \underbrace{y_{r+1}, \cdots, y_n}_{v}) = (u, v - B(u))$$

Then we have $\psi \circ F \circ \varphi^{-1}(x) = (x_1, \cdots, x_r, 0, \cdots, 0)$. $\square$