Manifolds and Maps

J.W.R

September 6, 2002

Chapter 1

Review Of Calculus

1.1 Derivatives

Let $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ be open and $f : X \to Y$. Recall that for any $r = 0, 1, 2, \ldots, \infty$ the map f is called C^r iff all its partials of order $\leq r$ exist and are continuous. When f is C^1 there is a map

$$Df : X \to \mathbb{R}^{m \times n}$$

which assigns to each $x \in X$ its **derivative** [1] $Df(x)$ defined by

$$Df(x)v = \frac{d}{dt} f(x + tv)\bigg|_{t=0}.$$

By the chain rule $Df(x)$ is nothing more than the matrix of partials of f evaluated at x:

$$Df(x) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_m} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_m} \end{bmatrix}$$

where $f = (f_1, f_2, \ldots, f_n)$ and the partials are evaluated at $x = (x_1, x_2, \ldots, x_m)$.

The basic idea of differential calculus is that near $x_0 \in X$ the map f is well approximated by the affine map

$$x \mapsto f(x_0) + Df(x_0)(x - x_0).$$

The precise way of saying this is as follows:

[1] Sometimes called the *Jacobian matrix* of f at x

Proposition 1. *If f is C^1 and $x \in X$ then for every $\epsilon > 0$ there exists $\delta > 0$ such that*

$$|f(x+v) - f(x) - Df(x)v| \leq \epsilon |v|$$

whenever $|v| < \delta$.

Remark 2. Conversely, if for every $x \in X$ there is a (necessarily unique) matrix $Df(x)$ satisfying the condition of proposition 1 and if the map $Df : X \to \mathbb{R}^{m\times n}$ is continuous then the map f is C^1.

Exercise 3. Let $GL(n,\mathbb{R}) \subset \mathbb{R}^{n\times n}$ denote the set of invertible matrices. This is an open set since it is defined by the condition that $a \in GL(n,\mathbb{R}) \iff \det(A) \neq 0$ and $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is a continuous function. Let $\iota : GL(n,\mathbb{R}) \to GL(n.\mathbb{R})$ be the map which assigns to each matrix a its inverse:

$$\iota(a) = a^{-1}.$$

that

$$D\iota(a)B = -a^{-1}Ba^{-1}$$

for $a \in GL(n,\mathbb{R})$ and $B \in \mathbb{R}^{nsn}$. (Hint: $\iota(a)a = 1$.)

Exercise 4. Show that

$$D\det(a)B = \det(a)\, tr(a^{-1}B)$$

for $a \in GL(n,\mathbb{R})$ and $B \in \mathbb{R}^{n\times n}$.

1.2 The Chain Rule

For open subsets $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ and $r = 0, 1, 2 \ldots \infty$ denote by $C^r(X,Y)$ the set of C^r maps $f : X \to Y$. These form the morphisms of the **category of open subsets of euclidean space and C^r maps** whose objects are open subsets of some $\mathbb{R}^n$:

For an open subset $X \subset \mathbb{R}^n$ the identity map

$$\mathsf{id}_X : X \to X : x \mapsto \mathsf{id}_X(x) = x$$

is C^r;

For open subsets $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, and $Z \subset \mathbb{R}^p$, and C^r maps $f : X \to Y$ and $g : Y \to Z$ the composition $g \circ f : X \to Z$ is a C^r map.

The isomorphisms of this category are called C^r-**diffeomorphisms**: thus a map

$$f : X \to Y$$

is a C^r-diffeomorphism iff it is bijective and both it and its inverse

$$f^{-1} : Y \to X$$

are C^r. (Thus a C^0-diffeomorphism is a homeomorphism.)

Theorem 5 (The Chain Rule)**.** *If $r \geq 1$ the derivative of the composition $g \circ f$ is given by*

$$D(g \circ f)(x) = Dg(f(x))Df(x)$$

for $x \in X$.

Remark 6. For each open $X \subset \mathbb{R}^m$ define

$$TX = X \times \mathbb{R}^m$$

and for each C^{r+1} map

$$f : X \subset \mathbb{R}^m \to Y \subset \mathbb{R}^n$$

of open sets define a C^r map

$$Tf : TX \to TY : (x, v) \mapsto Tf(x, v) = (f(x), Df(x)v).$$

Then the chain rule can be written in the compact form

$$T(g \circ f) = (Tg) \circ (Tf).$$

We also have the formula

$$T\,\mathsf{id}_X = \mathsf{id}_{TX}\,.$$

Taken together these formulas say that the operation T is a functor from the C^{r+1}-category to the C^r-category. It is called the **tangent functor.**

Since the derivative of the identity map is the identity matrix it follows that if

$$f : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^m$$

is a C^r-diffeomorphism ($r \geq 1$) then the linear map

$$Df(x) : \mathbb{R}^m \to \mathbb{R}^n$$

is a linear isomorphism for each $x \in X$ and hence that $m = n$. Thus C^r-diffeomorphic open sets have the same dimeansion. This is also true for $r = 0$ (Brouwer's *invariance of domain* theorem) but is harder to prove.

1.3 The Lipschitz Category

Definition 7. Let X and Y be metric spaces with metrics d_X and d_Y respectively and let $f : X \to Y$. Then the **Lipschitz constant** of f is the constant $\mathsf{lip}(f) \in [0, \infty]$ given by

$$\mathsf{lip}(f) = \sup \frac{d_Y(f(x_1), f(x_2))}{d_X(x_1, x_2)}$$

where the supremum is over all $x_1, x_2 \in X$ with $x_1 \neq x_2$. The map f is called **Lipschitz** iff $\mathsf{lip}(f) < \infty$ and **locally Lipschitz** iff every point of X has a neighborhood U such that $f|U$ is Lipschitz. The map f is a **lipeomorphism** iff it is bijective and both f and f^{-1} are Lipschitz; f is a **local lipeomorphism** at a point $x_0 \in X$ iff there are neighborhoods U of x_0 in X and V of $T(x_0)$ in Y such that $f|U$ is a lipeomorphism from U onto V.

Example 8. A map $T : \mathbb{R}^m \to \mathbb{R}^n$ of form

$$T(x) = y_0 + A(x - x_0)$$

where $y_0 \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^m$, and $A \in \mathbb{R}^{n\times m}$ is Lipschitz with

$$\mathsf{lip}(T) = |A| = \sup_{|v|=1} |Av|.$$

If the matrix A is invertible (so that $m = n$), then T is a lipeomorhiphism with

$$T^{-1}(y) = x_0 + A^{-1}(y - y_0).$$

Remark 9. The Lipschitz maps form a category: the identity map $\mathsf{id}_X : X \to X$ is Lipschitz with $\mathsf{lip}(\mathsf{id}_X) = 1$ and a composition $g \circ f$ of Lipschitz maps is Lipschitz with

$$\mathsf{lip}(g \circ f) \le \mathsf{lip}(g)\,\mathsf{lip}(f).$$

The lipeomorphisms are the isomorphisms of this category.

Definition 10. Call a map $\Gamma : X \to X$ from a metric space X to itself a **contraction** iff $\mathsf{lip}(\Gamma) < 1$; i.e. iff there exists $\lambda < 1$ such that

$$d(\Gamma(x_1), \Gamma(x_2)) \le \lambda d(x_1, x_2)$$

for all $x_1, x_2 \in X$. An **attractive fixed point** for $\Gamma : X \to X$ is a (necessarily unique) point $x \in X$ such that

$$\lim_{n\to\infty} \Gamma^n(x_0) = x$$

for all $x_0 \in X$. Here $\Gamma^n(x_0)$ is the n-th iterate of Γ on x_0 defined inductively by

$$\Gamma^0(x_0) = x_0$$

and

$$\Gamma^{n+1}(x_0) = \Gamma(\Gamma^n(x_0)).$$

Lemma 11 (Banach's Contraction Principle)**.** *A contraction map $\Gamma : X \to X$ on a complete metric space has a unique fixed point. This fixed point is attractive.*

Proof. If $\Gamma(x) = x$ and $\Gamma(y) = y$ then the inequality $d(x, y) = d(\Gamma(x), \Gamma(y)) \le \lambda d(x, y)$ implies $d(x, y) = 0$ which gives uniqueness. For existence choose $x_0 \in X$ arbitrarily and let

$$x_n = \Gamma^n(x_0)$$

denote the n-th iterate of Γ on x_0. Then by induction on k

$$d(x_{k+1}, x_k) \le \lambda^k d(x_1, x_0)$$

for $n \geq m$:

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{k=m}^{n-1} d(x_{k+1}, x_k) \\ &\leq \sum_{k=m}^{\infty} \lambda^k d(x_1, x_0) \\ &= \frac{\lambda^m}{1-\lambda} d(x_1, x_0) \end{aligned}$$

so that $\lim_{m,n\to\infty} d(x_n, x_m) = 0$, i. e. the sequence x_n is Cauchy. If

$$x = \lim_{n\to\infty} x_n$$

then

$$\begin{aligned} \Gamma(x) &= \Gamma(\lim_{n\to\infty} x_n) \\ &= \lim_{n\to\infty} \Gamma(x_n) \\ &= \lim_{n\to\infty} x_{n+1} \\ &= x \end{aligned}$$

so that x is an attractive fixed point of Γ as required. □

Theorem 12 (Lipschitz inverse function theorem)**.** *Let $T : \mathbb{R}^m \to \mathbb{R}^m$ be a lipeomorphism, $X \subset \mathbb{R}^m$ be open, $x_0 \in X$, and $R : X \to \mathbb{R}^m$ satisfy*

$$\mathsf{lip}(T^{-1})\,\mathsf{lip}(R) < 1$$

and

$$R(x_0) = 0.$$

Then the map $f : X \to \mathbb{R}^m$ given by

$$f(x) = T(x) + R(x)$$

is a local lipeomorphism at x_0. In fact, there are neighborhoods U of x_0 in X and V of $y_0 = f(x_0)$ in $\mathbb{R}^m$ and a map $g : V \to U$ such that for $x \in U$ and $y \in V$ we have

$$y = f(x) \iff x = g(y)$$

and

$$\mathsf{lip}(f) \leq \mathsf{lip}(T) + \mathsf{lip}(R)$$

$$\mathsf{lip}(g) \leq \frac{\mathsf{lip}(T^{-1})}{1 - \mathsf{lip}(T^{-1})\,\mathsf{lip}(R)}$$

Proof. We shall find positive real numbers a and b and define U and V by

$$V = \{y \in \mathbb{R}^m : |y - y_0| < b\}$$

$$U = \{x \in X \cap f^{-1}(V) : |x - x_0| < a\}$$

and $y_0 = f(x_0)$. First choose $a > 0$ so small that $x \in X$ for $|x - x_0| < a$ and then choose $b > 0$ so small that $a_1 < a$ where

$$a_1 = \mathsf{lip}(T^{-1})b + \mathsf{lip}(T^{-1})\,\mathsf{lip}(R)a$$

Now for $y \in \mathbb{R}^m$ and $x \in X$ we have $y = f(x)$ if and only if $x = \Gamma_y(x)$ where

$$\Gamma_y(x) = T^{-1}(y - R(x)).$$

Let

$$B = \{x \in \mathbb{R}^m : |x - x_0| < a\}$$

and

$$\overline{B}_1 = \{x \in \mathbb{R}^m : |x - x_0| \leq a_1\};$$

Then the hypothesis on R gives

$$\mathsf{lip}(\Gamma_y) < 1$$

and the inequality $a_1 < a$ gives

$$\Gamma_y(B) \subset \overline{B}_1 \subset B.$$

Thus for $y \in V$ the Banach contraction principle gives a unique $x \in B$ with $y = f(x)$; since $U = B \cap f^{-1}(V)$ this point x lies in U. Thus we define $g : V \to U$ by

$$x = g(y) \iff x = \Gamma_y(x)$$

and g is the inverse to the bijection $f|U : U \to V$. To compute $\mathsf{lip}(g)$ choose $y_1, y_2 \in V$ and put $x_1 = g(y_1)$ and $x_2 = g(y_2)$. Then

$$\begin{aligned} |g(y_1) - g(y_2)| &= |x_1 - x_2| \\ &= |\Gamma_{y_1}(x_1) - \Gamma_{y_2}(x_2)| \\ &\leq \mathsf{lip}(T^{-1})(|y_1 - y_2| + \mathsf{lip}(R)|x_1 - x_2|) \\ &= \mathsf{lip}(T^{-1})|y_1 - y_2| + \mathsf{lip}(T^{-1})\,\mathsf{lip}(R)|g(y_1) - g(y_2)|) \end{aligned}$$

so that

$$|g(y_1) - g(y_2)| \leq \frac{\mathsf{lip}(T^{-1})|y_1 - y_2|}{1 - \mathsf{lip}(T^{-1})\,\mathsf{lip}(R)}$$

which establishes the inequality

$$\mathsf{lip}((g)) \leq \frac{\mathsf{lip}(T^{-1})}{1 - \mathsf{lip}(T^{-1})\,\mathsf{lip}(R)}$$

as required. □

Proposition 13 (Uniform Contraction Principle). *Suppose U is a topological space, X is a complete metric space with metric d, and for each $p \in U$*

$$\Gamma_p : X \to X$$

is a map. Assume

(1) *The map Γ_p is a uniform contraction: i.e. there is a number $\lambda < 1$ independent of $p \in U$ such that*

$$d(\Gamma_p(x_1), \Gamma_p(x_2)) \leq \lambda d(x_1, x_2)$$

for $x_1, x_2 \in X$;

(2) *The map*

$$U \times X \to X : (p, x) \mapsto \Gamma_p(x)$$

is continuous.

Then for $p \in U$ the equation $\Gamma_p(x) = x$ has a unique solution $x = x_p$ and the map

$$U \to X : p \mapsto x_p$$

is continuous.

1.4 The Inverse Function Theorem

Now suppose that $X \subset \mathbb{R}^m$ is open and convex, $x_0 \in X$, and $f : X \to \mathbb{R}^m$ is C^r with $r \geq 1$. Then by Taylor's formula we have

$$f(x) = T(x) + R(x)$$

where

$$T(X) = f(x_0) + Df(x_0)(x - x_0)$$

and both R and DR vanish at x_0. Since

$$\mathsf{lip}(R) = \sup_{x \in X} |DR(x)|$$

we may shrink X to achieve $\mathsf{lip}(R) < \epsilon$. This proves most of

Theorem 14 (Inverse Function Theorem). *If $Df(x_0)$ is invertible then there are neighborhoods U of x_0 in X and V of $f(x_0)$ in $\mathbb{R}^m$ such that f restricts to a C^r diffeomorphism from U onto V.*

Proof. By the Lipschitz Inverse Function Theorem we have U and V so that f restricts to a lipeomorphism; we will show that f^{-1} is C^1 and that

$$Df^{-1}(y) = Df(f^{-1}(y))^{-1}$$

for $y \in V$. For this we must choose $y \in V$ and $\epsilon > 0$ and find $\delta = \delta(\epsilon, y) > 0$ so that

$$|f^{-1}(y+w) - f^{-1}(y) - Df(f^{-1}(y))^{-1}w| \leq \epsilon|w|$$

whenever $|w| < \delta$. Let $x = f^{-1}(y)$ and $v = f^{-1}(y+w) - x$ so that $y = f(x)$ and $y + w = f(x+v)$; we must establish

$$|v - Df(x)^{-1}(f(x+v) - f(x))| \leq \epsilon|f(x+v) - f(x)|.$$

But since f is differentiable at x we have $\delta_1 > 0$ depending on $\epsilon_1 > 0$ such that

$$|f(x+v) - f(x) - Df(x)v| \leq \epsilon_1|v|$$

whenever $|v| < \delta_1$ so

$$\begin{aligned} |v - Df(x)^{-1}(f(x+v) - f(x))| &\leq |Df(x)^{-1}|\,|Df(x)v - f(x+v) - f(x)| \\ &\leq \epsilon_1|Df(x)^{-1}|\,|v| \\ &\leq \epsilon_1|Df(x)^{-1}|\,\mathsf{lip}(f^{-1})|w| \end{aligned}$$

if $|v| \leq \delta_1$. Since

$$|v| \leq \mathsf{lip}(f^{-1})|w|$$

we achieve the desired inequality with

$$\epsilon_1 = \epsilon\left(|Df(x)^{-1}|\ \mathsf{lip}(f^{-1})\right)^{-1}$$

and

$$\delta = \delta_1\ \mathsf{lip}(f^{-1})^{-1}.$$

This establishes the inverse function theorem in case $r = 1$. The case $r > 1$ follows immediately from the following

Remark 15. A C^r map which is a C^1 diffeomorphism is a C^r diffeomorphism.

Since $f \circ f^{-1} = \mathsf{id}$ the chain rule gives $Df(f^{-1}(y))Df^{-1}(y) = 1$ or $Df^{-1}(y) = Df(f^{-1}(y))^{-1}$ This formula (together with the chain rule) shows that Df^{-1} is C^{r-1} if Df and f^{-1} are $C^r{-}1$ since the map $GL_n \to GL_n : a \mapsto a^{-1}$ is C^∞ (each entry of a^{-1} is a rational function of the entries of a by Cramer's rule.) Hence remark 15 follows by induction on r. □

Note that remark 15 is false when $r = 0$: the map $\mathbb{R} \to \mathbb{R} : x \mapsto x^3$ is a C^∞ homeomorphism which is not a C^1 diffeomorphism.

1.5 The Implicit Function Theorem

Theorem 16 (Implicit Function Theorem)**.** *Let $W \subset \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ be open, $f : W \to \mathbb{R}^n$ be C^r with $r \geq 1$, and $(x_0, y_0) \in W$. Assume*

$$f(x_0, y_0) = 0$$

and

$$D_2f(x_0, y_0)) : \mathbb{R}^n \to \mathbb{R}^n \text{ is invertible.}$$

Then there are neighborhoods U of x_0 and V of y_0 and a C^r function $g : U \to V$ with $U \times V \subset W$ and

$$\mathsf{Graph}(g) = (U \times V) \cap f^{-1}(0).$$

Proof. Most texts prove the slightly weaker assertion that there is a C^r function g satisfying

$$f(x, g(x)) = 0;$$

the idea is that the equation $f(x, y) = 0$ implicitly defines y as a function of x. Our statement proves a little more, viz. that for $x \in U$ and $y \in V$ we have

$$f(x, y) = 0 \iff y = g(x).$$

The reason the theorem is true is that for each x near x_0 the mapping f_x defined by $f_x(y) = f(x, y)$ is a local diffeomorphism by the inverse function theorem and so we can take $g(x) = f_x^{-1}(0)$. It is a little tricky to make this idea into a proof (why is g C^r?) so we resort to the useful trick of adding an auxilliary variable.

Define $F : W \to \mathbb{R}^m \times \mathbb{R}^n$ by

$$F(x, y) = (x, f(x, y)).$$

Then

$$DF(x_0, y_0) = \begin{bmatrix} 1_{m\times m} & 0 \\ D_1f(x_0, y_0) & D_2f(x_0, y_0) \end{bmatrix}$$

where $1_{m\times m}$ is the identity map of $\mathbb{R}^m$. This matrix is invertible so F is a local diffeomorphism by the inverse function theorem; i.e there is are neighborhoods U_1 of x_0 and V of y_0 such that $U_1 \times V \subset W$, $F(U_1 \times V)$ is a neighborhood of $(0, 0)$ in $\mathbb{R}^m \times \mathbb{R}^n$ and

$$F : U_1 \times V \to F(U_1 \times V)$$

is a diffeormorphism. (We can always shrink a neighborhood of (x_0, y_0) to make it a product neighborhood.) Since $pr \circ F = pr$ where $pr(x, y) = x$ it follows that $pr = pr \circ F^{-1}$ so that F^{-1} has form

$$F^{-1}(x, y) = (x, h(x, y))$$

for a certain map $h : F(U_1 \times V) \to V$. Let U_2 be the neighborhood of x_0 defined by

$$U_2 = \{x : (x, 0) \in F(U_1 \times V)\}$$

($x_0 \in U_2$ as $(x_0, 0) = F(x_0, y_0)$) and define $g : U_2 \to V$ by

$$g(x) = h(x, 0)$$

(note that $g(x_0) = y_0$). Now for $x \in U_2$ we have $(x, 0) \in F(U_1 \times V)$ so $(x, g(x)) = F^{-1}(x, 0) \in U_1 \times V$ so $F(x, g(x)) = F(F^{-1}(x, 0)) = (x, 0)$ showing that

$$\{(x, g(x)) : x \in U_2\} \subset f^{-1}(0) \cap (U_1 \times V)\}.$$

(Note that this implies also that $U_2 \subset U_1$.) Also for $x \in U_1$, $y \in V$, and $f(x, y) = 0$ we have $F(x, y) = (x, 0) \in F(U_1 \times V)$ so that $x \in U_2$ whence $(x, y) = F^{-1}(x, 0) = (x, g(x))$ showing that

$$f^{-1}(0) \cap (U_1 \times V) \subset \{(x, g(x)) : x \in U_2\}.$$

Since we have shown that $U_2 \subset U_1$ we may take $U = U_2$ to satisfy the requirements of the theorem. □

The ideas in this proof can be used to choose local coordinates to give a map a simple form. First let us recall the canonical form for a linear map.

Proposition 17. *Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map of rank r. Then there exist invertible linear maps $a : \mathbb{R}^m \to \mathbb{R}^m$ and $b : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$bTa^{-1}(x, y) = (x, 0) \in \mathbb{R}^r \times \mathbb{R}^{n-r} = \mathbb{R}^n$$

for $(x, y) \in \mathbb{R}^r \times \mathbb{R}^{m-r} = \mathbb{R}^m$.

Now we try to come as close as we can to this normal form for a nonlinear map.

Theorem 18 (The Local Representative Theorem)**.** *Let $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ be open sets and $f : M \to N$ be a C^k map with $k \geq 1$. Suppose $p \in M$ and the rank of the linear map $Df(p) : \mathbb{R}^m \to \mathbb{R}^n$ is r. Then there exist neighborhoods U of p in M, V of $f(p)$ in N, X of 0 in $\mathbb{R}^r$, Y of 0 in $\mathbb{R}^{m-r}$ and Z of 0 in $\mathbb{R}^{n-r}$, C^k diffeomorphisms $\alpha : U \to X \times Y$ and $\beta : V \to X \times Z$ and a C^k map $g : X \times Y \to Z$ such that $\alpha(p) = (0, 0) \in X \times Y$, $\beta(f(p)) = (0, 0) \in X \times Z$,*

$$\beta \circ f \circ \alpha^{-1}(x, y) = (x, g(x, y))$$

for $(x, y) \in X \times Y$ with $g(0, 0) = 0$ and $Dg(0, 0) = 0$.

Proof. The idea is that $\beta \circ f \circ \alpha^{-1}$ and f are equivalent in the sense that they differ by a change of co-ordinates; this is indeed an equivalence relation as

$$\beta_1 \circ (\beta_2 \circ f \circ \alpha_2^{-1}) \circ \alpha_1^{-1} = (\beta_1 \circ \beta_2) \circ f \circ (\alpha_1 \circ \alpha_2)^{-1}.$$

This means we may break our problem up into steps, first finding α_1, β_1 so that $f_1 = \beta_1 \circ f \circ \alpha_1^{-1}$ has a nice form, them replacing f by f_1 and finding α_2, β_2 etc. Hence we first make an affine change of variables so that $p = (0, 0) \in \mathbb{R}^r \times \mathbb{R}^{m-r}$, $f(p) = (0, 0) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$, and $Df(p)$ has the form of proposition 17:

$$Df(0, 0) = \begin{bmatrix} 1_{r \times r} & 0_{r \times (m-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (m-r)} \end{bmatrix}$$

Then
$$f(x,y) = (p(x,y), q(x,y)) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$$
for
$$(x,y) \in M \subset \mathbb{R}^r \times \mathbb{R}^{m-r}$$
with $p(0,0) = 0$, $q(0,0) = 0$, $D_1 p(0,0) = 1_{r\times r}$, $D_2 p(0,0) = 0_{r\times(m-r)}$ and $Dq(0,0) = 0_{(n-r)\times r}$. Now introduce the auxiliary map $\alpha : M \to R^m$ via
$$\alpha(x,y) = (p(x,y), y) \in \mathbb{R}^r \times \mathbb{R}^{m-r}$$
so that $D\alpha(0,0) = 1_{m\times m}$. By the inverse function theorem α restricts to a diffeomorphism $\alpha : U \to X \times Y$ on a suitable neighborhood U of p (we may shrink U if necessary to achieve that $\alpha(U)$ is a product) and thus achieve
$$f \circ \alpha^{-1}(x,y) = (x, g(x,y))$$
for $(x,y) \in X \times Y$ where $g = q \circ \alpha^{-1}$. We may shrink X and Y further to achieve $g(x,y) \in Z$ for some small neighborhood Z of 0 in $\mathbb{R}^{n-r}$ and then take β to be the identity. □

Several special cases of this theorem are of interest. [2]

Theorem 19 (The Submersion Theorem). *Let $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ be open sets and $f : M \to N$ be a C^k map with $k \geq 1$. Suppose $p \in M$ and that the linear map $Df(p) : \mathbb{R}^m \to \mathbb{R}^n$ is surjective (so $m \geq n$). Then there exist neighborhoods U of p in M, V of $f(p)$ in N, X of 0 in $\mathbb{R}^m$, Y of 0 in $\mathbb{R}^{m-n}$ and C^k diffeomorphisms*
$$\alpha : U \to X \times Y$$
and
$$\beta : V \to X$$
such that $\alpha(p) = (0,0) \in X \times Y$, $\beta(f(p)) = 0 \in X$, and
$$\beta \circ f \circ \alpha^{-1}(x,y) = x$$
for $(x,y) \in X \times Y$.

Proof. This is a special case of the local representative theorem with $m = r$. □

Theorem 20 (The Immersion Theorem). *Let $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ be open sets and $f : M \to N$ be a C^k map with $k \geq 1$. Suppose $p \in M$ and the the linear map $Df(p) : \mathbb{R}^m \to \mathbb{R}^n$ is injective (so $m \leq n$). Then there exist neighborhoods U of p in M, V of $f(p)$ in N, X of 0 in $\mathbb{R}^m$, and Z of 0 in $\mathbb{R}^{m-n}$, C^k diffeomorphisms*
$$\alpha : U \to X$$
and
$$\beta : V \to X \times Z$$

[2] The terms *immersion* and *submersion* are explained in 152 and 170

such that $\alpha(p) = 0 \in X$, $\beta(f(p)) = (0,0) \in X \times Z$, *and*

$$\beta \circ f \circ \alpha^{-1}(x,y) = (x,0)$$

for $(x,y) \in X$.

Proof. By the local representative theorem we may assume that f has the form

$$f(x) = (x, g(x))$$

for $x \in X_1$ where $g : X_1 \to Z_1 \subset \mathbb{R}^{n-m}$. Let $\beta : X_1 \times \mathbb{R}^{n-m} \to X_1 \times \mathbb{R}^{n-m}$ be given by

$$\beta(x,z) = (x, z - g(x)).$$

This map is clearly a diffeomorphism ($\beta^{-1}(x,z) = (x, z + g(x))$) and $\beta \circ f$ has the desired form. Now choose X and Z so that $(0,0) \in X \times Z \subset \beta(X_1 \times Z_1)$ and take $V = \beta^{-1}(X \times Z)$ (and $U = X$). □

Theorem 21 (The Rank Theorem)**.** *Let* $M \subset \mathbb{R}^m$ *and* $N \subset \mathbb{R}^n$ *be open sets and* $f : M \to N$ *be a* C^k *map with* $k \geq 1$. *Suppose the rank of the linear map* $Df(p) : \mathbb{R}^m \to \mathbb{R}^n$ *is a constant* r *independent of the choice of* $p \in M$. *Let* $p_0 \in M$ *Then there exist neighborhoods* U *of* p_0 *in* M, V *of* $f(p_0)$ *in* N, X *of* 0 *in* $\mathbb{R}^r$, Y *of* 0 *in* $\mathbb{R}^{m-r}$ *and* Z *of* 0 *in* $\mathbb{R}^{n-r}$, C^r *diffeomorphisms*

$$\alpha : U \to X \times Y$$

and

$$\beta : V \to X \times Z$$

such that $\alpha(p_0) = (0,0) \in X \times Y$, $\beta(f(p_0)) = (0,0) \in X \times Z$,

$$\beta \circ f \circ \alpha^{-1}(x,y) = (x,0)$$

for $(x,y) \in X \times Y$.

Proof. By the local representative theorem we may assume that f has the form

$$f(x,y) = (x, g(x,y))$$

where $x \in \mathbb{R}^r$, $p_0 = (0,0) \in \mathbb{R}^r \times \mathbb{R}^{m-r}$, $Dg(0,0) = 0$, etc. Note that

$$Df(x,y) = \begin{bmatrix} 1_{r\times r} & 0_{r\times(m-r)} \\ D_1g(x,y) & D_2g(x,y) \end{bmatrix}$$

so since the rank of $Df(x,y)$ is constant we must have $D_2g(x,y) = 0$ identically in (x,y). Hence $g(x,y) = g(x)$ is independent of y (at least locally) and we may argue as in the Immersion Theorem. □

Exercise 22. Here is another argument for achieving the set equality

$$\mathsf{Graph}(g) = f^{-1}(0) \cap (U \times V)$$

in the implicit function theorem. Once we have constructed g solving $(f(x, g(x)) = 0$ choose U and V so that $g(U) \subset V$, V is convex and

$$\int_0^1 D_2 f(x, y_1 + t(y_2 - y_1))dt$$

invertible for $y_1, y_2 \in V$ and $x \in U$. Then as

$$f(x, y_2) - f(x, y_1) = \int_0^1 D_2 f(x, y_1 + t(y_2 - y_1))dt(y_2 - y_1)$$

it follows that the solution $y \in V$ of $f(x, y) = 0$ is unique for each $x \in U$. (Supply the details.)

1.6 Manifolds

Definition 23. Let $r \geq 1$. A subset $M \subset \mathbb{R}^n$ is called a C^r**-manifold** (or more precisely a C^r-submanifold of $\mathbb{R}^n$) iff for every point $p_0 \in M$ the following equivalent conditions are satisfied:

(1) There are open sets $X \in \mathbb{R}^m$, $Y \in \mathbb{R}^{n-m}$, a neighborhood U of p_0 in $\mathbb{R}^n$, and a diffeomorphism

$$\Phi : U \to X \times Y$$

such that $\Phi(p_0) = (x_0, y_0) \in X \times Y$ and

$$\Phi(M \cap U) = X \times \{y_0\}.$$

(2) There is a neighborhood U of p_0 in $\mathbb{R}^n$, a point $y_0 \in \mathbb{R}^{n-m}$, and a C^r map

$$F : U \to \mathbb{R}^{n-m}$$

such that

$$M \cap U = F^{-1}(y_0)$$

and the derivative

$$DF(p) : \mathbb{R}^n \to \mathbb{R}^{n-m}$$

has maximal rank (i.e. is surjective) for $p \in U$;

(3) There is a neighborhood U of p_0 in $\mathbb{R}^n$, a point $x_0 \in \mathbb{R}^m$, a neighborhood X of x_0 in $\mathbb{R}^m$, and an injective C^r map

$$\phi : X \to U$$

such that $\phi(x_0) = p_0$,

$$\phi(X) = M \cap U,$$

and the derivative

$$D\phi(x) : \mathbb{R}^m \to \mathbb{R}^n$$

has maximal rank (i.e. is injective) for $x \in X$.

The diffeomorphism Φ is called a **submanifold chart** for M at p_0; the map F is called a **local defining equation** for M at p_0; the map ϕ is called a **local parameterization** of M at p_0; and the map

$$\phi^{-1} : M \cap U \to X$$

is called either a **co-ordinate chart** or a **system of local co-ordinates** for M at p_0. When M is connected, the integer m is independent of the choice of p_0 and is called the **dimension** of M; the number $n-m$ is called the **codimension** of M in $\mathbb{R}^n$.

Note the extreme cases M is a point (dimension $m = 0$) and $M \subset \mathbb{R}^n$ is open (codimension $n - m = 0$). More generally, note that a local defining equation $F(p) = y_0$ is really a system of $n - m$ equations (as $y_0 \in \mathbb{R}^{n-m}$); hence the principle that

> *the dimension of a manifold is the dimension of the ambient space minus the number of equations required to define it locally.*

Proof. Proof of (1) $\Longrightarrow$ (2) Assume (1) and write $\Phi : U \to X \times Y$ in the form

$$\Phi(p) = (G(p), F(p))$$

Then for $p \in U$ we have that $p \in M$ iff $\Phi(p) \in X \times \{y_0\}$ (by (1) iff $F(p) = y_0$. Also $DF(p)$ has maximal rank as it results from the invertible matrix $D\Phi(p)$ by eliminating the first m rows. This proves (2).

Proof. Proof of (2) $\Longrightarrow$ (1) Assume (2). Then by the submersion theorem 19 there is a diffeomorphism $\Phi : U \to X \times Y$ such that

$$F \circ \Phi^{-1}(x, y) = y$$

for $(x, y) \in X \times Y$. (It may be necessary to shrink the neighborhood U of p_0 to attain this.) But for $p \in U$ we have $p \in M$ iff $F(p) = y_0$ (by (2)) iff $y = F(\Phi^{-1}(x, y)) = y_0$ where $(x, y) = \Phi^{-1}(p)$ iff $p \in \Phi^{-1}(X \times \{y_0\})$. This proves (1).

Proof. (1) $\Longrightarrow$ (3) Assume (1) and define $\phi : X \to U$ by

$$\phi(x) = \Phi^{-1}(x, y_0)$$

for $x \in X$. Then ϕ is injective since it is a restiction of a bijection, $D\phi(x)$ is of maximal rank since it is obtained from the invertible matrix $D\Phi^{-1}(x, y_0)$ by discarding the last $n - m$ columns, and $\phi(X) = M \cap U$ since $\Phi^{-1}(X \times \{y_0\}) = \phi(X)$.

Proof. Proof of (3) $\Longrightarrow$ (1) Assume (3). Then by the Immersion Theorem there is are diffeomorphisms $\alpha : X \to X'$ and $\Phi : U \to X' \times Y$ such that $\Phi(\phi(x)) = (\alpha(x), 0)$ for $x \in X$. (It may be necessary to shrink U and X to achieve this.) Then $\phi(X) = \Phi^{-1}(X' \times \{0\}$ so that Φ is a submanifold chart as required.

Remark 24. There is a slightly more general definition of the notion of manifold which is more **intrinsic** in that it does not require M to be a subset of $\mathbb{R}^n$. By contrast, the current definition may be termed **extrinsic.**

1.7 Maps Between Manifolds

Let $M \subset \mathbb{R}^{m+k}$ be a C^r submanifold of dimension m, $N \subset \mathbb{R}^{n+l}$ be a C^r submanifold of dimension n and

$$f : M \to N.$$

Definition 25. The map f is said to be a C^r-**map** iff it satisfies the following equivalent conditions:

(1) There is a a neighborhood W of M in $\mathbb{R}^{m+k}$ and a C^r map $F : W \to \mathbb{R}^{n+k}$ which restricts to f:

$$f = F|M.$$

(2) For every $p \in M$ and every neighborhood V of $f(p)$ in $\mathbb{R}^{n+l}$ there is a neighborhood U of p in $\mathbb{R}^{m+k}$ and a C^r map $F : U \to V$ such that

$$F|M \cap U = f|M \cap U.$$

(3) For every $p \in M$, every C^r local parameterization $\psi : Y \to N$ of N at $f(p) \in N$ there is a C^r local parameterization $\phi : X \to M$ of M at $p \in M$ such that $f(\phi(X)) \subset \psi(Y)$ and the composition

$$\psi^{-1} \circ f \circ \phi : X \to Y$$

is a C^r map from the open subset $X \subset \mathbb{R}^m$ to the open subset $Y \subset \mathbb{R}^n$.

Remark 26. Since an open subset of $\mathbb{R}^m$ is a manifold we apparently have two definitions of what it means for f to be C^r; however in view of part (1) these coincide for we can take $W = M$,

Remark 27. When M is closed we can take $W = \mathbb{R}^{m+k}$ in part (1) of the definition.

1.8 Tangent Bundle

Let $M \subset \mathbb{R}^n$ be a C^{r+1} submanifold of dimension m.

Definition 28. For each point $p \in M$ define the **tangent space** T_pM to M at p by

$$T_pM = \{\dot{c}(0) : c \in C^1(\mathbb{R}, M), \ \ c(0) = p\}$$

where $\dot{c}$ denotes the derivative of c as a map $c : \mathbb{R} \to \mathbb{R}^n$:

$$\dot{c}(t_0) = \left.\frac{d}{dt}c(t)\right|_{t=t_0}.$$

The **tangent bundle** TM of M is the subset of $\mathbb{R}^n \times \mathbb{R}^n$ defined by

$$TM = \{(p, v) \in M \times \mathbb{R}^n : v \in T_pM\}.$$

Remark 29. When $M \subset \mathbb{R}^m$ is open ($k = 0$) the tangent space is the ambient space: $T_pM = \mathbb{R}^m$ and hence the tangent bundle is the product $TM = M \times \mathbb{R}^m$. (Compare with remark 6.)

Remark 30. Below we shall use an equality

$$T(X \times Y) = TX \times TY$$

where $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ are open. This is an abuse of notation, since

$$T(X \times Y) = X \times Y \times \mathbb{R}^m \times \mathbb{R}^n$$

and

$$TX \times TY = X \times \mathbb{R}^m \times Y \times \mathbb{R}^n$$

are set-theoretically distinct.

Proposition 31. *Each tangent space T_pM is a vector subspace of $\mathbb{R}^n$ of dimension m. The tangent bundle $TM \subset \mathbb{R}^n \times \mathbb{R}^n$ of the C^r submanifold $M \subset \mathbb{R}^n$ is a C^r submanifold of dimension $2m$.*

This is a corollary of each of the three following:

Proposition 32. *Let the C^{r+1} diffeomorphism $\Phi : U \to X \times Y$ be a submanifold chart for M as in part (1) of definition 23:*

$$M \cap U = \Phi^{-1}(X \times \{y_0\}).$$

Then

$$T_pM = D\Phi(p)^{-1}\left(\mathbb{R}^m \times \{0\}\right)$$

for $p \in M$. In fact, the map $T\Phi : TU \to TX \times TY$ defined by

$$T\Phi(p, w) = (\Phi(p), D\Phi(p)w)$$

for $(p, w) \in TU = U \times \mathbb{R}^n$ is a submanifold chart for $TM \subset \mathbb{R}^n \times \mathbb{R}^n$: It is a C^r-diffeomorphism and

$$TM \cap TU = T\Phi^{-1}\left(TX \times \{(y_0, 0)\}\right).$$

Proposition 33. *Let $F : U \to \mathbb{R}^{n-m}$ define M locally as in part (2) of definition 23:*

$$M \cap U = F^{-1}(y_0)$$

and $DF(p)$ is maximal rank for $p \in U$. Then the tangent space to M at $p \in M$ is the kernel of $DF(p)$:

$$T_pM = DF(p)^{-1}(0).$$

In fact, the map $TF : TU \times \mathbb{R}^n \to \mathbb{R}^{n-m} \times \mathbb{R}^{n-m}$ defined by

$$TF(p, w) = (F(p), DF(p)w)$$

for $(p, w) \in TU = U \times \mathbb{R}^n$ defines $TM \subset \mathbb{R}^n \times \mathbb{R}^n$ locally in the same sense:

$$TM \cap TU = (TF)^{-1}(y_0, 0)$$

and the matrix $D(TF)(p, w) \in \mathbb{R}^{2n \times 2(n-m)}$ defined by

$$D(TF)(p, w)(\hat{p}, \hat{w}) = (DF(p)\hat{p}, D^2F(p)\hat{p}w + DF(p)\hat{w})$$

for $(\hat{p}, \hat{w}) \in \mathbb{R}^n \times \mathbb{R}^n$ has maximal rank.

Proposition 34. *Let $\phi : X \to M \cap U$ be a local parameterization of M as in part (3) of definition 23:*

$$M \cap U = \phi(X).$$

Then the tangent space to M at $p = \phi(x) \in M$ is the image of $D\phi(x)$:

$$T_pM = D\phi(x)(\mathbb{R}^m)$$

for $x \in X \subset \mathbb{R}^m$ and $p = \phi(x) \in M$. In fact, the map

$$T\phi : TX \to TM : (x, \hat{x}) \mapsto T\phi(x, \hat{x}) = (\phi(x), D\phi(x)\hat{x})$$

is a C^r local parameterization of TM.

Example 35. The m-sphere S^m is the set of all $p \in \mathbb{R}^{m+1}$ whose distance from the origin is 1:

$$S^m = \{p \in \mathbb{R}^{m+1} : \|p\| = 1\}.$$

The tangent space to S^m at a point p is the set of all vectors $v \in \mathbb{R}^{m+1}$ which are perpendicular to p:

$$T_pS^m = \{v \in \mathbb{R}^{m+1} : \langle p, v \rangle = 0\}.$$

1.9 The Tangent Functor

Let $M \subset \mathbb{R}^{m+k}$ be a C^{r+1} submanifold of dimension m, $N \subset \mathbb{R}^{n+l}$ be a C^{r+1} submanifold of dimension n and

$$f : M \to N$$

be a C^{r+1} map.

Definition 36. For each $p \in M$ the **tangent map to f at p** is the linear map

$$T_p f : T_p M \to T_{f(p)} N$$

defined by

$$(T_p f)\, v = \left. \frac{d}{dt} f(c(t)) \right|_{t=0} .$$

for $v \in T_p M$ where $c : \mathbb{R} \to M$ is a C^1 curve satisfying $c(0) = p$ and $\dot{c}(0) = v$. The **tangent** of f is the map

$$Tf : TM \to TN$$

defined by

$$Tf(p, v) = (f(p), T_p f v)$$

for $(p, v) \in TM$.

Remark 37. In case $M \subset \mathbb{R}^m$ is open (i.e. $k = 0$) the map Tf has the form

$$Tf(p, v) = (f(p), Df(p)v)$$

so that

$$T_p f v = Df(p)v$$

in agreement with remark 6.

Proposition 38. *The operation which assigns to each C^{r+1} manifold M its tangent bundle TM and which assigns to each C^{r+1} map $f : M \to N$ its tangent map $Tf : TM \to TB$ is functorial. In other words,*

$$T\,\mathsf{id}_M = \mathsf{id}_{TM}$$

and

$$T(g \circ f) = (Tg) \circ (Tf)$$

for $f : M \to N$ and $g : N \to P$.

1.10 Monge Co-ordinates

Let $A : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ be linear, say

$$A(x, y) = Bx + Cy$$

for $x \in \mathbb{R}^m$, $y \in \mathbb{R}^k$, where $B : \mathbb{R}^m \to \mathbb{R}^k$ and $C : \mathbb{R}^k \to \mathbb{R}^k$ are linear. Let $H \subset \mathbb{R}^m \times \mathbb{R}^k$ be the kernel of A:

$$H = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^k : A(x, y) = 0\}.$$

Thus H is a vector subspace of $\mathbb{R}^m \times \mathbb{R}^k$ of dimension at least m. Let

$$\pi : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m : (x, y) \mapsto \pi(x, y) = x$$

denote projection onto the first factor. Let

$$V = \{0\} \times \mathbb{R}^k \subset \mathbb{R}^m \times \mathbb{R}^k$$

denote the kernel of π.

Proposition 39. *The following are equivalent:*

(1) *The linear map $C : \mathbb{R}^k \to \mathbb{R}^k$ is invertible.*

(2) *There is a linear map $S : \mathbb{R}^m \to \mathbb{R}^k$ such that the graph*

$$\mathsf{Graph}(S) = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^k : y = Sx\}.$$

of S is H;

(3) *The subspaces H and V intersect in zero:*

$$H \cap V = \{(0, 0)\}.$$

(4) *The subspaces H and V are complementary:*

$$\mathbb{R}^m \times \mathbb{R}^k = H \oplus V.$$

(5) *$\pi|H$ maps H isomorphically to $\mathbb{R}^m$.*

For each subset $I = \{i_1 < i_2 < \cdots < i_m\}$ of $\{1, 2, \ldots, n\}$ let

$$\pi_I : \mathbb{R}^n \to \mathbb{R}^m$$

denote the projection defined by

$$\pi_I(x) = (x_{i_1}, x_{i_2}, \ldots, x_{i_m})$$

for $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Lemma 40. *For each vector subspace $H \subset \mathbb{R}^n$ of dimension m there is a subset $I \subset \{1, 2, \ldots, n\}$ such that π_I maps H isomorphically to $\mathbb{R}^m$.*

Corollary 41. *Let $M \subset \mathbb{R}^n$ be a submanifold of dimension m and fix $p_0 \in M$. Then there is a set of indices $I \subset \{1, 2, \ldots, n\}$ such that $\pi_I|M$ is a local co-ordinate system for M near p_0.*

Coordinates of this form are called **Monge co-ordinates**. Another way to look at this situation ias that a manifold is locally a graph in the sense that some of its rectangular co-ordinates are functions of the others. Here's the precise statement:

Proposition 42. *Let $M \subset \mathbb{R}^n$ be as above. Then there is a map $f : X \to Y$ such that near p_0 the manifold M is defined by the equations*

$$x_J = f(x_I) \tag{$*$}$$

where $x_J = \pi_J(x)$, $x_I = \pi_I(x)$, and

$$J = \{1, 2, \ldots, n\} \setminus I$$

is the complementary set of indices to I. The tangent space T_xM to a manifold M having Monge co-ordinates $(*)$ is given by

$$T_xM = \{\hat{x} \in \mathbb{R}^n : \hat{x}_J = Df(x_I)\hat{x}_I\}.$$

Example 43. Monge co-ordinates show that any submanifold of $\mathbb{R}^n$ is locally the graph of a function. For example, the equation

$$x_1^2 + x_2^2 + \cdots + x_{m+1}^2 = 1$$

is the defining equation of the m-dimensional sphere S^m. In fact, in a neighborhood of any point where $x_i \neq 0$ the remaining co-ordinates $(x_1, \ldots x_{i-1}, x_{i+1}, \ldots x_n)$ are Monge co-ordinates for the sphere since the defining equation can be written in the form

$$x_i = \pm\sqrt{1 - x_1^2 - \cdots - x_{i-1}^2 - x_{i+1}^2 - \cdots - x_{m+1}^2}$$

near the point in question.

Chapter 2

Ordinary Differential Equations

2.1 Topological Flows

For any topological space M we denote by $\mathsf{Homeo}(M)$ the **homoemorphism group** of M; i. e. the set of all self-homeomorphisms $f : M \to M$. A **continuous flow** on a topological space M is a group homomorphism

$$\mathbb{R} \to \mathsf{Homeo}(M) : f \mapsto f^t$$

from the additive group of real numbers to the homeomorphism group of M such that the evaluation map

$$\mathbb{R} \times M \to M : (t, p) \mapsto f^t(p)$$

is continuous. The study of continuous flows is called *topological dynamics.*

Remark 44. For topological spaces U and M let $C^0(U, M)$ denote the space of continuous maps from U to M and let I be a set. We shall generally not distinguish [1] between the map

$$I \to C^0(U, M) : t \mapsto f^t$$

and the map

$$I \times U \to M : (t, p) \mapsto f^t(p)$$

and we denote the latter by f:

$$f(t, p) = f^t(p).$$

[1] This identification is called *Currying* (after the logician Haskell Curry) by computer scientists.

In particular, we may make this identification for $I = \mathbb{R}$ and $U = M$ and define a flow to be a continuous map

$$f : \mathbb{R} \times M \to M$$

satisfying the identities

$$f(0, p) = p$$

and

$$f(t + s, p) = f(t, f(s, p))$$

for $p \in M$ and $t, s \in \mathbb{R}$.

2.2 Smooth Flows

For a C^r manifold denote by $\mathsf{Diff}^r(M)$ the C^r-**diffeomorphism group** of M; i.e. the set of all self-diffeomorphism $f : M \to M$ of class C^r. Thus $\mathsf{Diff}^0(M) = \mathsf{Homeo}(M)$. When $r = \infty$ the superscript is suppressed:

$$\mathsf{Diff}(M) = \mathsf{Diff}^\infty(M).$$

Definition 45. A C^r **flow** on a C^r-manifold M is a group homomorphism

$$\mathbb{R} \to \mathsf{Diff}^r(M) : f \mapsto f^t$$

from the additive group of real numbers to the C^r diffeomorphism group of M such that the evaluation map

$$\mathbb{R} \times M \to M : (t, p) \mapsto f^t(p)$$

is of class C^r.

2.3 Vector Fields

Let M be a C^{r+1} manifold. A **vector field** on M is a section of the tangent bundle $TM \to M$; i. e. a function ξ which assigns to each $p \in M$ a vector $\xi(p) \in T_pM$.

The vector field ξ is a C^r vector field or of **class** C^r iff the map

$$M \to TM : p \mapsto (p, \xi(p))$$

is a C^r map from the manifold M to the manifold TM. We denote by $\mathcal{X}^r(M)$ the vector space of C^r vector fields on M. When M is a submanifold of $\mathbb{R}^n$ this is a vector subspace of the vector space of C^r $\mathbb{R}^n$-valued functions on M:

$$\mathcal{X}^r(M) = \{\xi \in C^r(M, \mathbb{R}^n) : \xi(p) \in T_pM \,\forall p \in M\}$$

When $r = \infty$ the superscript is suppressed:

$$\mathcal{X}(M) = \mathcal{X}^\infty(M).$$

Remark 46. Note that for an open set $M \subset \mathbb{R}^m$ the tangent space

$$T_pM = \{\dot{\gamma}(0) \quad | \qquad \gamma : (\mathbb{R}, 0) \to (M, p)\}$$

is independent of $p \in M$:

$$T_pM = \mathbb{R}^m.$$

Thus the tangent bundle is the product

$$TM = M \times \mathbb{R}^m$$

and a vector field on M is a map of form

$$M \to TM = M \times \mathbb{R}^m : p \mapsto (p, \xi(p))$$

where $\xi : M \to \mathbb{R}^m$. The map $\xi : M \to \mathbb{R}^m$ is called the **principal part** of the vector field and is strictly speaking different from the vector field itself which is a map $M \to TM$. However, we shall generally not distinguish the two.

Note that for a flow $t \to f^t$ on M each $p \in M$ determines a curve

$$\mathbb{R} \to M : t \mapsto f^t(p)$$

through $f^0(p) = p$. Thus the tangent vector to this curve (at $t = 0$) lies in the tangent space T_pM.

Definition 47. The **infinitessimal generator** of a C^{r+1} flow

$$\mathbb{R} \to \mathsf{Diff}^{r+1}(M) : t \mapsto f^t$$

is the C^r vector field $\xi \in \mathcal{X}^r(M)$ defined by

$$\xi(p) = \left.\frac{d}{dt} f^t(p)\right|_{t=0}$$

for $p \in M$.

Definition 48. A C^1 curve $c : I \to M$ (where $I \subset \mathbb{R}$ is an interval) is called an **integral curve** of the vector field $\xi \in \mathcal{X}^0(M)$ iff it satisfies the equation

$$\dot{c}(t) = \xi(c(t))$$

for all $t \in I$. An **integral curve** of ξ **through the point** $p \in M$ is an integral curve c of ξ satisfying $c(0) = p$.

Proposition 49. *Let $\xi \in \mathcal{X}^0(M)$ be the infinitessimal generator of the C^1 flow $\mathbb{R} \to \mathsf{Diff}^1(M) : t \mapsto f^t$. Then for each $p \in M$ the curve*

$$\mathbb{R} \to M : t \mapsto f^t(p)$$

is an integral curve of ξ.

Proof. In other words

$$\frac{d}{dt} f^t(p)\bigg|_{t=t_0} = \xi(f^{t_0}(p)).$$

For $t = t_0$ this is the definition of the infinitessimal generator. In general write $t = s + t_0$ and use the definition of flow:

$$f^{s+t_0}(p) = f^s(f^{t_0}(p))$$

and the obvious identity

$$\frac{d}{dt} f^t(p)\bigg|_{t=t_0} = \frac{d}{ds} f^{s+t_0}(p)\bigg|_{s=0}.$$

□

2.4 Co-ordinates

Let M and N be a C^2 manifolds, ξ be a C^1 vector field on M, η be a C^1 vector field on N, and $f : M \to N$ be a C^1 map.

Definition 50. The map f **intertwines** the vector fields ξ and η iff the following two equivalent conditions hold:

(1) The map f carries integral curves of ξ to integral curves of η;

(2) We have the identity

$$(T_p f)\xi(p) = \eta(f(p))$$

for $p \in M$.

Some authors call ξ and η **f-related** when f interwines ξ and η.

To see the equivalence of the two parts of the definition note that the first part says that the curve $b = f \circ c$ solves $\dot{b} = \eta \circ b$ whenever c solves $\dot{c} = \xi \circ c$. By the Existence Theorem (proved below) given $p \in M$ we can always find an integral curve c through p. If we differentiate the equation $b(t) = f(c(t))$ and set $t = 0$ we obtain the equation in the second definition. The converse implication follows by the chain rule.

If $f : M \to N$ is a diffeomorphism, then f interwines ξ and η if and only if f^{-1} interwines η and ξ. Given the vector field ξ on M there is a unique $f_\sharp\xi$ on N such that f interwines ξ and $f_\sharp\xi$. It is defined by

$$(f_\sharp\xi)(q) = (T_p f)\,\xi(p)$$

for $q \in N$ where $p \in M$ is defined by $p = f^{-1}(q)$.

Proposition 51. *Let ξ be the infinitessimal generator of a C^1 flow $t \mapsto f^t$. Then for each $t \in \mathbb{R}$ the time t map $f^t \in \mathsf{Diff}^1(M)$ interwines ξ and ξ.*

Now let (α, U) be a C^{r+1} co-ordinate chart on M; i. e. $\alpha : U \to X$ is a C^{r+1} diffeomorphism from an open set $U \subset M$ to an open set $X \subset \mathbb{R}^m$. Then a vector field ξ on M restricts to a vector field $\xi|U$ on U and hence determines a vector field $\xi_\alpha = \alpha_\sharp(\xi|U)$ on X i.e. a function $\xi_\alpha : X \to \mathbb{R}^m$. This vector field is called the **local representative** of ξ with respect to α. In particular,

Proposition 52. *A curve $c : I \to X$ is an integral curve for ξ_α if and only if the curve $\alpha^{-1} \circ c : I \to M$ is an integral curve for ξ.*

Of course, if (V, β) is another co-ordinate chart the local representatives are related by the equation

$$\xi_\beta(y) = D\phi_{\beta\alpha}(\phi_{\alpha\beta}^{-1}(y))\xi_\alpha(\phi_{\alpha\beta}^{-1}(y))$$

for $y \in \beta(U \cap V)$ where $\phi_{\alpha\beta} : \alpha(U \cap V) \to \beta(U \cap V)$ is the diffeomorphism given by $\phi_{\alpha\beta}(x) = \beta \circ \alpha^{-1}(x)$ for $x \in \alpha(Y \cap V) \subset X$.

Exercise 53. Rewrite the differential equation

$$\dot{x} = u(x, y), \quad \dot{y} = v(x, y),$$

in polar co-ordinates:

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Exercise 54. Rewrite the differential equation

$$\dot{x} = u(x, y, z), \quad \dot{y} = v(x, y, z), \quad \dot{z} = w(x, y, z),$$

in spherical co-ordinates:

$$x = \rho\sin\phi\cos\theta, \quad y = \rho\sin\phi\sin\theta, \quad z = \rho\cos\phi.$$

2.5 Linear Flows

The group $GL(m, \mathbb{R})$ of all invertible $m \times m$ real matrices can be viewed as a subgroup of the diffeomorphism group $\mathsf{Diff}(\mathbb{R}^m)$ of $\mathbb{R}^m$: each matrix $a \in GL(m, \mathbb{R})$ determines the linear diffeomorphism $\mathbb{R}^m \to \mathbb{R}^m : x \mapsto ax$. Similarly, each matrix $A \in \mathbb{R}^{m\times m}$ determines a vector field $\mathbb{R}^m \to \mathbb{R}^m : x \mapsto Ax$; we call such a vector field a **linear vector field**. A flow

$$\mathbb{R} \to GL(m, \mathbb{R}) \subset \mathsf{Diff}(\mathbb{R}^m) : t \mapsto \Phi^t$$

is called a **linear flow**. Some authors call it a **one parameter subgroup** of $GL(m, \mathbb{R})$.

Theorem 55. **(1)** *Every continuous linear flow is C^∞.*

(2) *The infinitessimal generator of a linear flow is a linear vector field.*

(3) *The mapping which assigns to each linear flow its infinitessimal generator establishes a bijective correspondence beteween linear flows $t \mapsto \Phi^t$ and matrices $A \in \mathbb{R}^{m\times m}$. The inverse correspondence is given by the equation*

$$\Phi^t = \exp(tA)$$

for $t \in \mathbb{R}$ where exp *is the* **matrix exponential**.

Proof. The matrix exponential is defined by the series

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

This defines a C^∞ map

$$\exp : \mathbb{R}^{m\times m} \to GL(m, \mathbb{R}).$$

It satisfies the conditions

$$\exp(0) = e$$

the identity matrix and

$$\exp(A + B) = \exp(A)\exp(B) \text{ if } AB = BA.$$

This means that each map

$$\mathbb{R} \to GL(m, \mathbb{R}) : t \mapsto \exp(tA)$$

is a group homomorphism. The derivative of exp at the zero matrix is given by

$$D\exp(0)\hat{A} = \hat{A}$$

for $\hat{A} \in \mathbb{R}^{m\times m}$ which shows that (1) the generator of $t \mapsto \exp(tA)$ is A and (2) there are neighborhoods U of 0 in $\mathbb{R}^{m\times m}$ and V of e in $GL(m, \mathbb{R})$ such that exp restricts to a diffeomorphism (inverse function theorem). We shrink U if necessary to achieve $tA \in U$ whenever $A \in U$ and $|t| \le 1$.

Now let $t \mapsto \Phi^t$ be a linear flow and choose $\epsilon > 0$ so that $\Phi^t \in V$ for $t \le \epsilon$. There is a unique $A \in \epsilon^{-1}U$ such that

$$(*) \qquad \exp(tA) = \Phi^t$$

for $t = \epsilon$: we must show that this holds for all $t \in \mathbb{R}$. Since (*) holds for $t = n\tau$ whenever it holds for $t = \tau$ and n is an integer it is enough to show that it holds for all sufficiently small t. In fact, it is enough to show that it holds for $t = \tau$ whenever it holds for $t = 2\tau$, for then it holds for all dyadic multiples of ϵ and hence for all t by continuity. Hence assume (*) for $t = 2\tau$ where $2\tau \le \epsilon$. There is a unique solution $B \in U$ of $\exp(B) = \Phi^\tau$; squaring both sides gives $\exp(2B) = \Phi^{2\tau} = \exp(2\tau A)$ so $2B = 2\tau A$ whence $B = \tau A$ as required. □

Exercise 56. Show that

$$\exp(tJ) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

where

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Exercise 57. Prove in detail that exp is a C^∞ map.

Hint: The series converges uniformly because of the estimate

$$\|\exp(A)\| \le \sum_{n=0}^{\infty} \frac{\|A\|^n}{n!} = e^{\|A\|}$$

where $A \mapsto \|A\|$ is any operator norm on $\mathbb{R}^{m\times m}$; that is, any norm satisfying $\|AB\| \le \|A\|\|B\|$.

Exercise 58. Prove that $\exp(A+B) = \exp(A)\exp(B)$ when A and B commute.

Exercise 59. For $A \in \mathbb{R}^{m\times}$ define linear maps

$$\lambda(A),\ \rho(A),\ \mathsf{Ad}(A) : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$$

by

$$\lambda(A)B = AB, \quad \rho(A)B = BA,$$

and

$$\mathsf{Ad}(A) = AB - BA$$

for $B \in \mathbb{R}^{m\times m}$. Prove the two formulas

$$D\exp(A) = \lambda(\exp(A)) \sum_{n=0}^{\infty} \frac{\mathsf{Ad}(-A)^n}{(n+1)!}$$

and

$$D\exp(A) = \rho(\exp(A)) \sum_{n=0}^{\infty} \frac{\mathsf{Ad}(A)^n}{(n+1)!}$$

Note the special case $D\exp(A)B = \exp(A)B$ when $AB = BA$.

Hint: Note that

$$\mathsf{Ad}(A) = \lambda(A) - \rho(A)$$

and that

$$\lambda(A)\rho(A) = \rho(A)\lambda(A).$$

Now

$$D\exp(A) = \sum_{n=0}^{\infty} \frac{S_n(A)}{(n+1)!}$$

where

$$S_n(A) = \sum_{p+q=n} \lambda(A)^p \rho(A)^q.$$

Use the binomial theorem to conclude

$$S_n(A) = \sum_{r=0}^{n} \binom{n+1}{r+1} \lambda(A)^{n-r} \, \mathsf{Ad}(-A)^r.$$

Exercise 60. The formula for $\exp(A)$ is meaningful whenever A is a linear endomorphism of a finite dimensional vector space V and then defines a linear automorphism of V. We may thus define

$$\exp : gl(V) \to GL(V)$$

for any V where $gl(V)$ is the vector space of all linear transformations $A : V \to V$ (endomorphisms of V) and $GL(V)$ is the group of all invertible linear transformations $a : V \to V$ (automorphisms of V). Now $\mathsf{Ad}(A)$ is a linear transformation from $V = \mathbb{R}^{m\times m}$ to itself so the formula $\exp(\mathsf{Ad}(A))$ is meaninful. Show that

$$\exp(\mathsf{Ad}(A)) = \mathsf{ad}(\exp(A))$$

for $A \in \mathbb{R}^{m\times}$ where

$$\mathsf{ad} : GL(m, \mathbb{R}) \to GL(\mathbb{R}^{m\times m})$$

is defined by

$$\mathsf{ad}(b)A = bAb^{-1}$$

for $b \in GL(m, \mathbb{R}) = GL(\mathbb{R}^m)$ and $A \in gl(\mathbb{R}^m) = \mathbb{R}^{m\times m}$.

Exercise 61. Prove the formula

$$D \; \mathsf{ad}(e)A = \mathsf{Ad}(A)$$

where $e \in GL(\mathbb{R}^m)$ is the identity matrix and $A \in gl(\mathbb{R}^m)$

2.6 Existence and Uniqueness

Theorem 62. *Let M be a C^2 manifold and ξ be a C^1 vector field on M. Fix $t_0 \in \mathbb{R}$ and $p_0 \in M$. Then:*

(Existence) *There exists an open interval $I \subset \mathbb{R}$ about t_0 and an integral curve $c : I \to M$ for ξ with $c(t_0) = p_0$.*

(Uniqueness) *Any two integral curves $c_i : I_i \to M$ $(i = 1, 2)$ of ξ with $c_1(t_0) = c_2(t_0)$ agree on their common domain: $c_1(t) = c_2(t)$ for $t \in I_1 \cap I_2$.*

Proof. The existence theorem is local so we may choose co-ordinates (see proposition 52) and assume without loss of generality that M is an open subset of $\mathbb{R}^m$. The condition that c be an integral curve of ξ through p_0 can be expressed in the form $\dot{c}(t) = \xi(c(t))$ and $c(t_0) = p_0$ which by the fundamental theorem of calculus can be expressed in a single integral equation

$$c(t) = p_0 + \int_{t_0}^{t} \xi(c(\tau))\, d\tau.$$

This in turn is a fixed point equation:

$$\Gamma(c) = c$$

where

$$\Gamma(c)(t) = p_0 + \int_{t_0}^{t} \xi(c(\tau))\, d\tau.$$

Choose $\rho > 0$ so that small that

$$\bar{B}_\rho(p) = \{p \in \mathbb{R}^m : |p - p_0| \leq \rho\},$$

the closed ball of radius ρ centered at p_0, is a subset of M and so that the quantities

$$K = \sup\{|\xi(x)\| : x \in \bar{B}_\rho(p_0)\}$$

and

$$L = Lip(\xi|\bar{B}_\rho(p_0))$$

are finite (the former by the continuity of ξ the latter by the continuity of $D\xi$ and the mean value theorem) and then choose $\epsilon > 0$ so small that

$$\epsilon K < \rho$$

and

$$2\epsilon L < 1.$$

Let I be the open interval $(t_0 - \epsilon, t_0 + \epsilon)$ and let $\mathcal{C}$ be the set of all continuous curves $c : I \to \bar{B}_\rho(p_0)$ with $c(0) = p_0$. The estimate

$$\begin{aligned} \|\Gamma(c)(t) - p_0\| &\leq \int_0^t \|\xi(c(\tau))\|\, d\tau \\ &\leq \int_{t_0}^t K\, d\tau \\ &< \rho \end{aligned}$$

for $t \in I$ shows that $\Gamma(c) : I \to \bar{B}_\rho(p_0)$ so that we have a map

$$\Gamma : \mathcal{C} \to \mathcal{C}.$$

Equip $\mathcal{C}$ with the complete metric

$$\|c_1 - c_2\|_0 = \sup\{|c_1(t) - c_2(t)| \,:\, t \in I\}.$$

The estimate

$$\begin{aligned}\|\Gamma(c_1) - \Gamma(c_2)\|_0 &\leq \sup_{t\in I} \int_{t_0}^{t} |\xi(c_1(\tau)) - \xi(c_2(\tau))|\, d\tau \\ &\leq \int_{t_0-\epsilon}^{t_0+\epsilon} L|c_1(\tau) - c_2(\tau)|\, d\tau \\ &\leq 2\epsilon L\|c_1 - c_2\|_0\end{aligned}$$

shows that Γ is a contraction mapping on $\mathcal{C}$ and hence has a unique fixed point by the Banach contraction principle. This proves existence and uniqueness on I (local uniqueness).

For uniqueness note that the set

$$J = \{t \in I_1 \cap I_2 :\, c_1(t) = c_2(t)\}$$

is a closed subset of $I_1 \cap I_2$ since the curves c_i are continuous and M is Hausdorff. It is non-empty since $t_0 \in J$ by hypothesis and open by the local uniqueness (for general t_0) just proved. Hence $J = I_1 \cap I_2$ as required. □

2.7 Flow Boxes

Definition 63. A **flow box** for ξ is a map

$$f : I \times U \to M$$

where $I \subset \mathbb{R}$ is an open interval about 0, U is an open subset of M, and for each $p \in U$ the map $I \to M : t \mapsto f(t, p)$ is an integral curve through p.

Theorem 64. *Let ξ be a C^r vector field ($r \geq 1$) on a C^{r+1} manifold M. For every $p_0 \in M$ there is a C^r flow box for ξ with $p_0 \in U$.*

Proof. We first construct a C^0 flow box for ξ. The argument is a slight refinement of the proof of theorem 62. As before assume without loss of generality that M is an open subset of $\mathbb{R}^m$ and choose $\rho > 0$ so small that the closed ball $\bar{B}_\rho(p)$ is a subset of M. Choose numbers K and L, and a neighborhood U of p_0 in $\mathbb{R}^m$ such that

$$|\xi(q)| \leq K, \quad |\xi(q_1) - \xi_(q_2)| \leq L|q_1 - q_2|$$

for $p \in U$, $q, q_1, q_2 \in \bar{B}_\rho(p)$. Now shrink U and choose $\epsilon > 0$ so that

(*i*) $$q \in \bar{B}_\rho(p_0) \text{ if } p \in U \text{ and } |q - p| \leq \epsilon K$$

and

$$(ii). \qquad \epsilon L < 1$$

Let

$$I = (-\epsilon, \epsilon)$$

be the open interval of length 2ϵ centered at $0 \in \mathbb{R}$.

Now let $\mathcal{F}$ be the function space

$$\mathcal{F} = C^0(I \times U, \bar{B}_\rho(p_0))$$

of continous maps $f : I \times U \to \bar{B}_\rho(p_0)$ with the complete metric d defined by

$$d(f_1, f_2) = \sup\{|f_1(t,p) - f_2(t,p)| \,:\, (t,p) \in I \times U\}$$

and for $f \in \mathcal{F}$ defined $\Gamma(f) : I \times U \to \mathbb{R}^m$ by

$$\Gamma(f)(t,p) = p + \int_0^t \xi(f(\tau, p))\, d\tau.$$

Now $\Gamma : \mathcal{F} \to \mathcal{F}$ by (i) and $\mathsf{lip}(\Gamma) \leq 1$ by (ii) so Γ has a fixed point. The equation $\Gamma(f) = f$ is the integrated form of the differential equation which says that each map $t \mapsto f^t(p) = f(t,p)$ is an integral curve of ξ so the fixed point is a flow box as required. □

Next we show there is a C^1 flow box. To get the idea we assume that ξ is C^2. Write the differential equation in the form

$$(D0) \qquad \frac{d}{dt} f^t = \xi \circ f^t.$$

If we assume that f is C^1 we can apply the tangent functor and get another differential equation

$$(D1) \qquad \frac{d}{dt} F^t = (T\xi) \circ F^t, \quad F^t = Tf^t$$

of the same form. The method of solution via the Banach contaction principle will give a sequence of functions F^t_n which converge to a solution F^t of (D1) and direct calculation shows that $F^t_n = Tf^t_n$ (where f^t_n is the corresponding sequence of approxinmations for equation (D0)) provided this holds when $n = 0$. Thus since $F^t_n = Tf^t_n$ is converging uniformly we have that both f^t_n (the first component of Tf^t_n) and DF^t_n are converging uniformly. Hence the limit $f^t(p)$ is a C^1 function of p and $Df^t(p)$ is a continuous function of (t,p). But the differential equation (D0) shows that the partial derivative with respect to t is also continuous. Hence the partial derivatives of $f(t,p) = f^t(p)$ with respect to both t and p exist and are continuous and so f is C^1 as required.

This argument is essentially correct (though a bit sketchy) but assumes that ξ is C^2 rather than C^1. It turns out however that if we look at the argument a bit more carefully we see that the stronger hypothesis is not needed. Integrating (D1) and separating components we see that it can be rewritten as a fixed point equation

$$\Phi(f, A) = (f, A)$$

where

$$f : I \times U \to M \subset \mathbb{R}^m$$

as before and

$$A : I \times U \to \mathbb{R}^{m \times m}.$$

For the solution it will turn out that A is the derivative of f:

$$A(t, p) = Df^t(p).$$

The precise equations for Φ are

$$\Phi(f, A) = (\Gamma(f), \Phi_f(A))$$

where

$$\Gamma(f)(t, p) = p + \int_0^t \xi(f(\tau.p))\, d\tau$$

as before and $\Phi_f(A)$ is obtained from this equation by differentiating $\Gamma(f)(t, p)$ with respect to p and substituting $A(t, p)$ for $D_2 f(t, p) = Df^t(p)$:

$$\Phi_f(A)(t, p) = E + \int_0^t D\xi(f(\tau, p))A(\tau, p)\, d\tau.$$

(Here E is the identity matrix.)

Now for $\mathbf{F} = \mathbb{R}^m$ or $\mathbf{F} = \mathbb{R}^{m \times m}$ (or any other complete normed vector space) $X = I \times U$ (or any other topological space) let $C^0(X, \mathbf{F})$ denote the space of continuous maps from X to $\mathbf{F}$ and let $BC^0(X, \mathbf{F})$ denote the space of *bounded* continuous maps from X to $\mathbf{F}$:

$$BC^0(X, \mathbf{F}) = \{f \in C^0(X, \mathbf{F}) : \|f\|_0 < \infty\}$$

where $\|f\|_0$ is the **sup norm**:

$$\|f\|_0 = \sup\{|f(x)| : x \in X\}.$$

Now $BC^0(X, \mathbf{F})$ is a Banach space (a complete normed space) since the uniform limit of continuous functions is continuous and the space

$$\mathcal{F} = \{f \in C^0(I \times U, \mathbb{R}^m) : f(I \times U) \subset \bar{B}_\rho(p_0)\}$$

is a closed subset of $BC^0(I \times U, \mathbb{R}^m)$ and is hence itself a complete metric space. As before $\Gamma : \mathcal{F} \to \mathcal{F}$ and $\mathsf{lip}(\Gamma) < 1$. Put

$$\mathcal{A} = BC^0(I \times U, \mathbb{R}^{m \times m})$$

so that $\Phi_f : \mathcal{A} \to \mathcal{A}$ and note the estimate

$$\mathsf{lip}(\Phi_f) \le \epsilon \| D\xi | \bar{B}_\rho(p_0) \|_0.$$

Hence by making ϵ smaller we may achieve $\mathsf{lip}(\Phi_f) \le \lambda < 1$ for all $f \in \mathcal{F}$. (Note however that even though $\mathsf{lip}(\Gamma) < 1$ we do not have that $\mathsf{lip}(\Phi) < 1$.) We now apply

Theorem 65 (The Fiber Contraction Principle)**.** *Suppose $\mathcal{F}$ and $\mathcal{A}$ are complete metric spaces, that*

$$\Phi : \mathcal{F} \times \mathcal{A} \to \mathcal{F} \times \mathcal{A}$$

is a map of form

$$\Phi(f, A) = (\Gamma(f), \Phi_f(A))$$

for $f \in \mathcal{F}$ and $A \in \mathcal{A}$ where $\Gamma : \mathcal{F} \to \mathcal{F}$. Assume

(1) *The map $\Gamma : \mathcal{F} \to \mathcal{F}$ has an attractive fixed point $f^* \in \mathcal{F}$; i. e.*

$$\lim_{n\to\infty} \Gamma^n(f_0) = f^*$$

for every $f \in \mathcal{F}$.

(2) *Each of the maps $\Phi_f : \mathcal{A} \to \mathcal{A}$ are contractions uniformly in f: i.e. there exists $\lambda < 1$ such that*

$$\mathsf{lip}(\Phi) \le \lambda < 1$$

for all $f \in \mathcal{F}$.

(3) *For each fixed $A \in \mathcal{A}$ the map*

$$\mathcal{F} \to \mathcal{A} : f \mapsto \Phi_f(A)$$

is continuous.

Then Φ has a (necessarily unique) attractive fixed point (f^*, A^*): i.e.

$$\Phi(f^*, A^*) = (f^*, A^*)$$

and

$$\lim_{n\to\infty} \Phi^n(f_0, A_0) = (f^*, A^*)$$

for all $(f_0, A_0) \in \mathcal{F} \times \mathcal{A}$.

Proof. It is clear that Φ has a unique fixed point (f^*, A^*) since $\Phi(f, A) = (f, A)$ implies $\Gamma(f) = f$ (and hence $f = f^*$) and $\Phi_{f^*}(A) = A = A^*$ the unique fixed point of the contraction Φ_{f^*}. The problem is to show that the fixed point is attractive.

Choose $(f_0, A_0) \in \mathcal{F} \times \mathcal{A}$ and defined $(f_n, A_n) = \Phi^n(f_0, A_0)$ to be the n-th iterate of Φ applied to (f_0, A_0):

$$A_{n+1} = \Phi_{f_n}(A_n).$$

Let d_n and ϵ_n be defined by

$$d_n = d(A_n, A^*)$$

and

$$\epsilon_n = d(\Phi_{f_n}(A^*), A^*).$$

Hypotheses (3) and (1) guarantee that $\lim_{n\to\infty} \epsilon_n = 0$; we must prove that $\lim_{n\to\infty} d_n = 0$. The triangle inequality gives

$$d_{n+1} \leq d(\Phi_{f_n}(A_n), \Phi_{f_n}(A^*)) + d(\Phi_{f_n}(A^*), \Phi_{f^*}(A^*))$$

so hypothesis (2) gives

$$d_{n+1} \leq \lambda d_n + \epsilon_n.$$

Suppose $\epsilon > 0$ is small and $\epsilon_n < \epsilon$ for $n > N = N(\epsilon)$. Then for $n > N(\epsilon)$ we obtain the inequality

$$\begin{aligned} d_n &\leq \epsilon + \epsilon\lambda + \cdots + \epsilon^{n-N+1} + \lambda^{n-N} d_N \\ &\leq \epsilon(1-\lambda)^{-1} + \lambda^{n-N} d_N. \end{aligned}$$

In particular taking $\epsilon = sup_n \epsilon$ ($< \infty$ since ϵ_n converges) and $N = 0$ gives that the d_n are bounded say by M so that $d_n \leq M$ and hence

$$d_n \leq \epsilon(1-\lambda)^{-1} + \lambda^{n-N} M$$

for $n > N = N(\epsilon)$. It follows that d_n tends to zero as required.

Now using the fiber contraction principle it is easy to see that the flow box f must be C^1. Indeed choose $f_0 \in \mathcal{F}$ arbitrarily say $f_0(t,p) = p$ define $f_n \in \mathcal{F}$ and $A_n \in \mathcal{A}$ by $\Phi^n(f_0, D_2 f_0) = (f_n, A_n)$. By induction on n we have $A_n = D_2 f_n$ so (as $D_2 f_n$ converges uniformly if follows that $D_2 f$ exists and is continuous. The differential equation itself show that $D_1 f$ is continuous. Hence f is C^1 as required.

Next we construct a C^r flow box by induction on r. Suppose ξ is C^{r+1} with $r \geq 1$. Then the vector field

$$M \times R^{m\times m} \to \mathbb{R}^m \times R^{m\times m} : (p, a) \mapsto (\xi(p), D\xi(p)a)$$

is C^r. The map

$$I \times U \times R^{m\times m} \to M \times R^{m\times m} : (t, p, a) \mapsto (f^t(p), F^t(p, a))$$

is a C^r flow box for this vector field. Then by the uniqueness Theorem

$$F^t(p, a) = Df^t(p)a.$$

Hence $(t,p) \mapsto Df^t(p) = D_2 f(t,p)$ is C^r and the differential equation

$$\frac{d}{dt} Df(t,p) = D\xi(f^t(p)) Df^t(p)$$

shows that $(t,p) \mapsto D_1 f(t,p)$ is C^r. Hence $(t,p) \mapsto f^t(p)$ is C^{r+1} as required.

To prove the theorem in case $r = \infty$ requires a bit more argument because we have not yet excluded the possibility that the domain of definition $I \times U$ of the C^r flow box depends on r. It follows from the work in the next section that *every* flow box is C^r: to prove this for $r = \infty$ it suffices to prove it for all $r < \infty$. (See corollary 68 below.) □

2.8 Partial Flows

Let ξ be a C^r vector field on a C^{r+1} manifold.

Definition 66. The **maximal partial flow generated by** ξ is the map

$$f : W \to M$$

where $W \subset \mathbb{R} \times M$ is the open set defined by

$$W = \bigcup_{\alpha} I_\alpha \times U_\alpha$$

with the union being over all domains $I_\alpha \times U_\alpha$ of C^1 flow boxes

$$f_\alpha : I_\alpha \times U_\alpha \to M$$

for ξ and the map f is defined [2] by

$$f_\alpha = f | I_\alpha \times U_\alpha.$$

Lemma 67. *Let $c : J \to M$ be an integral curve for ξ with $0 \in J$ and choose $t_0 \in J$. Then there is a C^r flow box $f : I \times U \to M$ with $t_0 \in I$ and $p_0 = c(0) \in U$.*

Corollary 68. *The maximal partial flow of a C^r vector field ($r \geq 1$) is C^r.*

Corollary 69. *Let $c : J \to M$ be an integral curve for ξ with $0 \in J$. Then $J \times \{c(0)\} \subset W$ where W is the domain of the maximal partial flow of f.*

Proof. Proof of lemma 67 We may assume that $r < \infty$ since than the case where $r = \infty$ follows by applying corollary 68 for every $r < \infty$. Let $\tilde{J}$ be the set of all $t \in J$ for which the theorem is true; that is, for which there is a C^r flow box $I \times U \to M$ with $(t, p_0) \in I \times U$. We must show that $\tilde{J} = J$. We do this by 'continuous induction', that is we show that $\tilde{J}$ is non-empty, open, and closed in J.

Now $0 \in \tilde{J}$ since there is a flow box at p_0 and $\tilde{J}$ is open since the domain of a flow box is. To see that $\tilde{J}$ is closed in J choose a limit point $\bar{t} \in J$ of $\tilde{J}$. We must find a flow box

$$f : I \times U \to M$$

[2]Well-defined by the uniqueness theorem.

with $(\bar{t}, p_0) \in I \times U$. Choose a flow box

$$f_1 : I_1 \times U_1 \to M$$

with $c(\bar{t}) \in U_1$. Now choose $\bar{s} \in \tilde{J}$ so close to $\bar{t}$ that $\bar{t} - \bar{s} \in I_1$. Then (as $\bar{s} \in \tilde{J}$) choose a flow box $f_0 : I_0 \times U_0 \to M$ with $(\bar{s}, p_0) \in I_0 \times U_0$. Now choose $U \subset U_0$ with $p_0 \in U$ and

$$f_0(\{\bar{s}\} \times U) \subset U_1,$$

let

$$I = I_0 \cup \bar{s} + I_1$$

and define f by

$$\begin{aligned} f(t, p) &= f_0(t, p) && \text{for } t \in I_0 \\ &= f_1(t - \bar{s}, f_0(\bar{s}, p)) && \text{for } t \in \bar{s} + I_1 \end{aligned}$$

where the two formulas agree on the overlap (by the uniqueness theorem). This gives the desired flow box. □

For each $p \in M$ there is a (necessarily unique) integral curve $c_p : I_p \to M$ for ξ through p which is maximal in the sense that if $c : I \to M$ is any other integral curve to ξ through p then $I \subset I_p$ (and $c = c_p|I$ by the uniqueness theorem). (For example, one could define I_p to be the union of all intervals I which are domains of an integral curve through p)

Proposition 70. *An integral curve $c : I \to M$ for ξ is maximal iff its graph $\{(t, c(t)) : t \in I\}$ of c is a closed subset of $\mathbb{R} \times M$.*

Corollary 71. *For $p \in M$ let $\alpha(p)$ and $\beta(p)$ be the end points of the interval I_p, domain of the maximal integral curve of ξ through p. Then the maps*

$$\alpha : M \to (0, \infty] \text{ and } \beta : M \to [-\infty, 0)$$

are respectively lower and upper semicontinuous.

Proof. This is because

$$\alpha(p) = \sup\{t : (t, p) \in W\}, \quad \beta(p) = \inf\{t : (t, p) \in W\},$$

and W is open. □

Example 72. Let $M = \mathbb{R}^2 \setminus \{(0, 0)\}$, $\xi(x, y) = (1, 0)$, so that the maximal partial flow is defined by

$$f(t, (x, y)) = (x + t, y)$$

for $(t, (x, y)) \in W$ where $(t, (x, y)) \in W$ iff either $y \neq 0$ or $x < 0$, $y = 0$, and $t < -x$ or $x > 0$, $y = 0$, and $-x < t$. Thus $\beta(x, y) = \infty$ unless $x < 0$ and $y = 0$ in which case $\beta(x, y) = -x$ and similarly $\alpha(x, y) = -\infty$ unless $x > 0$ and $y = 0$ in which case $\alpha(x, y) = -x$.

Exercise 73. Find the maximal partial flow when $M = \mathbb{R}$ and $\xi(x) = x^2$.

2.9 Complete Vector Fields

Let ξ be a C^r vector field ($r \geq 1$) on a C^{r+1} manifold M. The domain W of the maximal partial flow is a neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$ and the map $f : W \to M$ satisfies the identities

$$f(0, p) = p$$

and

$$f(t + s, p) = f(t, f(s, p)).$$

(This last identity is by the uniqueness theorem since both both curves $t \mapsto f(t + s, p)$ and $t \mapsto f(t, f(s, p))$ are integral curves to ξ through $f^s(p)$.) When $W = \mathbb{R} \times M$ we define (for each $t \in \mathbb{R}$) a map $f^t : M \to M$ by

$$f^t(p) = f(t, p).$$

Each f^t is a diffeomorphism (bijective since $f^{-t} \circ f^t = f^t \circ f^{-t} = f^0 = \mathsf{id}_M$) and the identities say that the map $t \mapsto f^t$ is a homomorphism of groups. Thus when $W = \mathbb{R} \times M$ the vectorfield ξ generates a flow. In thus case we call ξ complete:

Definition 74. The vector field ξ is **complete** iff it satisfies the three following equivalent conditions:

(1) There is a flow $f : \mathbb{R} \times M \to M$ having ξ as its infinitessimal generator.

(2) For every $p \in M$ there is an integral curve $c : \mathbb{R} \to M$ for ξ through p defined on all of $\mathbb{R}$.

(3) There is an open interval I about 0 in $\mathbb{R}$ such that for every $p \in M$ there is an integral curve $c : \mathbb{R} \to M$ for ξ through p defined on all of I.

Proof. Proof that the conditions are equivalent Clearly, (1) implies (2) imples (3). On the other hand if (3) holds then by corollaries 68 and 69 there is a C^r flow box of form $f : I \times M \to M$. Choose $h > 0$ so small that I contains a closed interval of length h and extend the flow box to all of $\mathbb{R} \times M$ by the formula

$$f^t(p) = f(t - nh, f_h^n(p))$$

where $f_h = f(h, \cdot)$ and where the integer n is chosen so that the 'remainder' $t - nh \in I$. The integer n is not unique for we can also have $t - (n + 1)h \in I$ but $f(t, p)$ is well defined by the uniqueness theorem:

$$f(t - (n + 1)h, f_h^{n+1}(p)) = f(t - nh, f_h^n(p))$$

since both sides are integral curves and equality holds for $t = (n + 1)h$. The formula shows that $(t, p) \mapsto f^t(p)$ is C^r and that each curve $t \mapsto f^t(p)$ is an integral curve of ξ. Thus $t \mapsto f^t$ is a flow with infinitessimal generator ξ as required. □

Corollary 75. *Every C^1 vector field on a compact C^2 manifold is complete.*

Proof. For every point $p \in$ there is a flow box with domain $U_p \times (-\epsilon_p, \epsilon_p)$ and $p \in U_p$. By compactness choose a finite collection of these so that the U_{p_i} cover M and let ϵ be the smallest of the corresponding ϵ_{p_i}. Then

$$M \times (-\epsilon, \epsilon) \subset \bigcup_i U_{p_i} \times (-\epsilon_{p_i}, \epsilon_{p_i})$$

so the vector field satisfies part (3) of the definition of completeness. □

Exercise 76. Let ξ be a bounded C^1 vector field on $\mathbb{R}^m$. Then ξ is complete.

(Hint: Show that the graph of each integral curve is closed in $\mathbb{R} \times \mathbb{R}^m$.)

Exercise 77. Let ξ be a vector field on $\mathbb{R}^m$ with $\mathsf{lip}(\xi) < \infty$. Then ξ is complete.

(Hint: Apply the Banach Contraction Princlple to $\Gamma_p : C^0(I, \mathbb{R}^n) \to C^0(I, \mathbb{R}^n)$ where the interval I about zero is independent of the initial point p of the desired integral curve. Conclude that the domain of the maximal partial flow contain $I \times \mathbb{R}^m$.)

Exercise 78. Let ξ be a vectorfield on $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ so that the system of equations $\dot{z} = \xi z$ takes the form

$$\dot{x}_i = \dot{y}_i$$

$$\dot{y}_i = -\partial_i V(x)$$

for $z = (x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ and where $V\mathbb{R}^n \to \mathbb{R}$ is a C^2 real valued fuction. Then the vector field ξ is complete if the function V is bounded below: $L \leq V(x)$ for some constant $L \in \mathbb{R}$ and all $x \in \mathbb{R}^n$.

In classical mechanics equations of this form arise frequently. The function V is called the *potential energy*. Show that that the *total energy*

$$H(x, y) = \|y\|^2/2 + V(x)$$

is constant along integral curves in that

$$\frac{d}{dt} H(x(t), y(t)) = 0$$

whenever $z = (x(t), y(t))$ is a solution of the system $\dot{z} = \xi(z)$. Obtain an estimate

$$|x(b) - x(a)| \leq K|b - a|$$

where the constant K involves the lower bound L and the energy $h = H(x(t), y(t))$ of the integral curve. Conclude that the graph of the integral curve $I \to \mathbb{R}^n \times \mathbb{R}^n$ can be closed in $\mathbb{R} \times (\mathbb{R}^n \times \mathbb{R}^n)$ only if its interval of definition I is the whole time axis: $I = \mathbb{R}$.

Exercise 79. Suppose $\xi : (0,\infty) \to (0,\infty)$ is a C^1 function on $(0,\infty)$, the open infinite half interval. Assume $\phi, \psi : [0,\infty) \to (0,\infty)$ satisfy

$$\phi(t) \le \xi(\phi(t))$$

$$\psi(t) = \xi(\psi(t))$$

and

$$\phi(0) = \psi(0).$$

Then

$$\phi(t) \le \psi(t) \text{ for } t/ge0.$$

Exercise 80. [Gronwall's Inequality] Suppose $b \ge 0$ and $\gamma : \mathbb{R} \to \mathbb{R}^n$ is a C^1 map satisfying

$$|\dot{\gamma(t)}| \le b|\gamma(t)|$$

for $t \in \mathbb{R}$. Then

$$|\gamma(t)| \le |\gamma(0)|e^{bt}$$

for $t \in \mathbb{R}$.

Exercise 81. Any C^1 vector field ξ on $\mathbb{R}^m$ which grows at most linearly (meaning that there exist constants a and b such that

$$|\xi(p)| \le a + b|p|$$

for all $p \in \mathbb{R}^m$) is complete. In particular, a bounded vector field is complete.

Exercise 82. Let ξ be a vector field on $\mathbb{R}^m$ with $\mathsf{lip}(\xi) < \infty$. Then ξ is complete.

Hint: One can argue that ξ grows at most linearly and apply the previous exercise. Alternatively, one can show that that the integral curves are defined on a interval independent of the initial condition. (This argument was sketched above.) Here is a sketch of a different argument. Construct the flow $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ is the uniform limit on compact sets of the sequence of maps $f_n : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ defined recursively by $f_0(t,p) = p$ and

$$f_{n+1}(t,p) = p + \int_0^t \xi(f_n(\tau,p))\,d\tau.$$

Induct on the estimate

$$|f_{n+1}(t,p) - f_n(t,p)| \le \int_0^t L|f_n(\tau,p) - f_{n-1}(\tau,p)|\,d\tau$$

where $L = \mathsf{lip}(\xi)$ to achieve as a bonus the estimate

$$|f(t,p) - p| \le L^{-1}e^{|t|L}|\xi(p)|.$$

2.10 Lie Brackets

Let ξ and η be smooth vector fields on a smooth manifold M and let

$$W_\xi \subset \mathbb{R} \times M \to M : (t,p) \mapsto f^t(p)$$

$$W_\eta \subset \mathbb{R} \times M \to M : (t,p) \mapsto g^t(p)$$

denote the maximal partial flows of ξ and η respectively.

Definition 83. The **Lie brackets** of the vector fields ξ and η is the vectorfield $[\xi, \eta]$ defined by

$$[\xi,\eta](p) = \left.\frac{d}{dt}\right|_{t=0+} f^{\sqrt{t}} \circ g^{\sqrt{t}} \circ f^{-\sqrt{t}} \circ g^{-\sqrt{t}}(p).$$

Theorem 84. *In case $M \subset \mathbb{R}^m$ we have*

$$[\xi,\eta](p) = D\xi(p)\eta(p) - D\eta(p)\xi(p).$$

Here we are identifying a vector field with its principal part.

Corollary 85. *The Lie brackets of two C^r vector fields is of class C^{r-1}.*

Remark 86. An important special case of the last formula is the formula

$$AB - BA = \left.\frac{d}{dt}\right|_{t=0+} \exp(\sqrt{t}A)\exp(\sqrt{t}B)\exp(-\sqrt{t}A)\exp(-\sqrt{t}B)$$

for square matrices A and B. It can be proved from the first few terms of the power series expansion

$$\exp(sA) = I + sA + \frac{s^2}{2}A^2 + O(s^3).$$

The formula shows that the Lie brackets of two linear vector fields is a linear vector field.

Proposition 87. *Let $f : M \to N$ be a C^1 map of C^2 manifolds, ξ_1, ξ_2 be C^1 vector fields on M, and η_1, η_2 be C^1 vector fields on N. Suppose that for $i = 1, 2$ the map f interwines the vector fields ξ_i and η_i. Then f interwines $[\xi_1, \xi_2]$ and $[\eta_1, \eta_2]$.*

2.11 Time Dependent Vector Fields

A **time-dependent vector field** on a manifold M is a map

$$\xi : \mathbb{R} \times M \to TM$$

such that

$$\xi(t,p) \in T_pM$$

for $t \in \mathbb{R}$ and $p \in M$. A **diffeotopy** of class C^r is a map

$$\mathbb{R} \to \mathsf{Diff}^r(M) : t \mapsto f_t$$

for which the evaluation map

$$\mathbb{R} \times M \to M : (t, p) \mapsto f_t(p)$$

is C^r. The **infinitessimal generator** of this diffeotopy is the time-dependent vector field defined by

$$\left.\frac{d}{ds} f_s(p)\right|_{s=t} = \xi(t, f_t(p)).$$

Theorem 88. *Suppose that M is compact and of class C^{r+1} where $r \geq 1$ and that ξ is a time-dependent vector field on M of class C^r. Then there is a (necessarily unique) C^r diffeotopy on M having ξ as its infinitessimal generator and satisfying the initial condition*

$$f_0 = \mathsf{id}_M .$$

Proof. Sketch of Proof Using a trick we reduce the problem to the time-independent case. Let $\tilde{M} = \mathbb{R} \times M$ and define a vector field $\tilde{\xi}$ on $\tilde{M}$ by

$$\tilde{\xi}(t, p) = (1, \xi(t, p)) \in T_{(t,p)}\tilde{M} = \mathbb{R} \times T_p M.$$

Then the flow $\tilde{f}$ of $\tilde{\xi}$ is related to the diffeotopy generated by ξ via the formula

$$\tilde{f}^t(0, p) = (t, f_t(p)).$$

Remark 89. Other terminology is sometimes used. A map

$$\mathbb{R}^2 \to \mathsf{Diff}(M) : (t, s) \mapsto f_s^t$$

satisfying

$$f_t^t = \mathsf{id}_M, \quad f_t^r \circ f_s^t = f_s^r$$

is sometimes called an **evolution system**. The formula

$$\tilde{f}^t(s, p) = (t + s, f_s^{t+s}(p))$$

establishes a correspondence between evolution systems on M and flows on $\mathbb{R} \times M$ which have generators of form $\tilde{\xi}$.

Chapter 3

MANIFOLDS

Throughout, **E**, **F**, **G** etc. denote finite dimensional vector spaces.

3.1 Manifolds defined

Definition 90. Let M be a set. A **chart** on M is a pair (α, U) where $U \subset M$, α is a bijection from U to an open subset of some finite dimensional vector space **E**; i.e. $\alpha(U) \subset \mathbf{E}$ is open and $\alpha : U \to \alpha(U)$ is a bijection. Two charts (α, U) and (β, V) are C^r **compatible** iff $\alpha(U \cap V)$ and $\beta(U \cap V)$ are both open (in their respective ambient spaces) and

$$\beta \circ \alpha^{-1} : \alpha(U \cap V) \to \beta(U \cap V)$$

is a C^r diffeomorphism. A C^r **atlas** on M is a collection $\mathfrak{A}$ of charts on M any two of which are C^r compatible and such that the sets U, as (α, U) ranges over $\mathfrak{A}$, cover M (i.e., for every $x \in M$ there is a chart $(\alpha, U) \in \mathfrak{A}$ with $x \in U$). A **maximal** C^r **atlas** is an atlas which contains every chart which is C^r compatible with each of its members. A maximal C^r atlas is also called a C^r **structure**.

Lemma 91. *If $\mathfrak{A}$ is an C^r atlas, then so is the collection of all charts C^r compatible with each member of $\mathfrak{A}$. In other words, every C^r atlas extends uniquely to a maximal C^r atlas.*

Definition 92. A C^r manifold is a pair consisting of a set M and a maximal C^r atlas $\mathfrak{A}$ on M.

Remark 93. In view of lemma 91, a C^r manifold is usually specified by giving its underlying set M and some C^r atlas on M. Generally, the notation for the atlas is suppressed and the manifold is denoted simply by M. The members of the atlas are called **admissible** C^r charts on M or simply charts on M.

Remark 94. In view of lemma 91, every C^r manifold is a C^q manifold for $q \leq r$.

Definition 95. The **manifold topology** of a C^r manifold M is the topology generated by the sets U as (α, U) ranges over the charts of M.

Example 96. The manifold topology need not be Hausdorff.

Example 97. The manifold topology need not be second countable.

Redefinition 98. Henceforth, unless otherwise stated, all manifolds will be assumed to be Hausdorff and second countable.

Definition 99. Let M and N be C^r manifolds and

$$f : M \to N$$

be a map. Given (admissible) charts (α, U) on M and (β, V) on N we define the **local representative**

$$f_{\beta\alpha} : \alpha(U \cap f^{-1}(V)) \to \beta(V)$$

of f with respect to (α, U) and (β, V) by

$$f_{\beta\alpha} = \beta \circ f \circ \alpha^{-1} | \alpha(U \cap f^{-1}(V)).$$

The map f is C^r iff *for all* C^r charts (β, V) on N and (α, U) on M the set $U \cap f^{-1}(V)$ is open in M and the local representative $f_{\beta\alpha}$ is C^r. Note that a map is C^r if for every $x \in M$ *there exist* charts C^r charts (β, V) and (α, U) with $x \in U \cap f^{-1}(V)$, $U \cap f^{-1}(V)$ open, and $f_{\beta\alpha}$ is C^r.

Definition 100. A map $f : M \to N$ between C^r-manifolds is a C^r **diffeomorphism** iff f is bijective and both f and f^{-1} are C^r. Two C^r manifolds M and N are C^r **diffeomorphic** iff there exists a C^r diffeomorphism from M onto N.

Proposition 101. *The identity map on a C^r manifold is a C^r map. The composition of C^r maps is C^r.*

This proposition says that the C^r manifolds and C^r maps form a category. The C^r diffeomorphisms are the isomorphisms of this category.

Proposition 102. *A C^r map is C^s for $s \leq r$. A map between C^r manifolds is a C^0 map if and only if it is continuous. If $r \geq 1$ then a C^r map which is a C^1 diffeomorphism is a C^r diffeomorphism, but a C^∞ homeomorphism need not be a C^1 diffeomorphism.*

Definition 103. Let M be a manifold. Then M has **dimension** m iff the ambient vector spaces of the open sets $\alpha(U)$ as (α, U) ranges over the charts of M all have dimension m. Some authors call a manifold of dimension m an m-manifold.

Remark 104. Suppose $U \subset \mathbf{E}$ and $V \subset \mathbf{F}$ are are C^r diffeomorphic open subsets of vector spaces. Then $\mathbf{E}$ and $\mathbf{F}$ have the same dimension. For $r \geq 1$ this is quite easy to see, for if $\phi : U \to V$ is a diffeomorphism then $D\phi(x) : \mathbf{E} \to \mathbf{F}$ is a vector space isomorphism for $x \in U$. For $r = 0$ (i.e. ϕ a homeomorphism) this is Brouwer's famous *invariance of domain* theorem. It follows that a connected manifold has dimension m for some m.

Example 105. Let M_1 be the C^r manifold having the real line $\mathbb{R}$ as its underlying subset and having as a C^r atlas the set consisting of the single chart (α, U) where $U = M_1 = \mathbb{R}$ and $\alpha(x) = x$ for $x \in U$ and let M_2 be the C^r manifold having the real line $\mathbb{R}$ as its underlying subset and having as a C^r atlas the set consisting of the single chart (β, V) where $V = M_2 = \mathbb{R}$ and $\beta(y) = y^3$ for $y \in V$. If $r = 0$, these manifolds are the same (i.e., the corresponding maximal atlas are set theoretically the same) while if $1 \leq r \leq \infty$ they are distinct. In any case, they are C^r diffeomorphic; a diffeomorphism given by $f : M_2 \to M_1 : f(z) = z^3$.

Remark 106. More generally, two C^r manifolds having the same underlying set are the same (i.e., the corresponding maximal C^r atlas are identical) if and only if the identity map (considered as a map from the manifold to the other) is a C^r diffeomorphism. Even if the identity map is not a C^r diffeomorphism the manifolds may still be diffeomorphic via another map.

Milnor's famous example shows that two C^∞ manifolds can be homeomorphic but not C^∞ diffeomorphic. On the other hand, for $r \geq 1$ C^r manifolds are C^r diffeomorphic if and only if they are C^1 diffeomorphic. (See Munkres for example.)

Proposition 107. *Let $\mathbb{R}$ have its usual C^r structure (i.e., $\mathbb{R}$ is the manifold M_1 of example 105). Then any C^r manifold which is homeomorphic to $\mathbb{R}$ is C^r diffeomorphic to $\mathbb{R}$.*

3.2 Examples

Example 108. An open subset $M \subset \mathbf{E}$ of a finite-dimensional vector space is canonically a C^∞ manifold; an atlas consists of the single chart (id_M, M). The notion of C^r map defined in definition 99 agrees with the usual notion of C^r map for such manifolds.

Example 109. Any open subset V of a C^r manifold M is again a C^r manifold; the atlas consists of all charts $(\alpha | U \cap V, U \cap V)$ as (α, U) ranges over the charts of M. If (α, U) is a C^r chart on M, then $\alpha : U \to \alpha(U)$ is a C^r diffeomorphism.

Remark 110. A manifold is locally connected. Thus its connected components are open subsets and hence connected manifolds. Thus one may view a C^0 manifold as a (Hausdorff and second countable) topological space each component of which is locally homeomorphic to some $\mathbb{R}^n$.

Example 111. The n-sphere S^n is a C^∞ manifold. The underlying set is given by

$$S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$$

where $\|\cdot\|$ denotes the usual norm on $\mathbb{R}^n$. A C^∞ atlas is given by the $2n+2$ charts $(\alpha_{i\sigma}, U_{i\sigma})$ where $i = 1, \ldots, n+1, \sigma = \pm$ defined by

$$U_{i\pm} = \{(x_1, \ldots, x_{n+1}) \in S^{n+1} : \pm x_i > 0\}$$

$$\alpha_{i\pm}(x_1, \ldots, x_{n+1}) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \in \mathbb{R}^n.$$

An atlas consisting of two charts and compatible with the above atlas consists of (α_+, U_+), (α_-, U_-) where

$$U_\pm = S^n \setminus \{(\pm 1, 0, 0, \ldots, 0)\}$$

$$\alpha_\pm(x) = y \in \mathbb{R}^n$$

where $(0, y) \in \mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$ is the unique intersection of the hyperplane $0 \times \mathbb{R}^n \subset \mathbb{R}^{n+1}$ with the line through $(\pm 1, 0, \ldots, 0)$ and x. (Stereographic projection). Note that the transition map

$$\gamma = \alpha_- \circ (\alpha_+)^{-1} : \mathbb{R}^n \setminus 0 \to \mathbb{R}^n \setminus 0$$

is given by

$$\gamma(y) = \| y \|^{-2} y.$$

The inclusion map $S^n \to \mathbb{R}^{n+1}$ is C^∞.

Definition 112. The **direct product** of two C^r manifold M and N is a C^r manifold. The atlas is given by the set of all pairs $(\alpha \times \beta, U \times V)$ as (α, U) range over all C^r charts on M and (β, V) range over all C^r charts on N. The projections $M \times N \to M$ and $M \times N \to N$ are C^r. (More generally a finite product of C^r manifolds is a C^r manifold.)

Example 113. The space $\mathbb{R}^2$ may be canonically identified with the field of complex numbers $\mathbb{C}$. Then

$$S^1 = \{z \in \mathbb{C} : z\overline{z} = 1\}.$$

The map

$$S^1 \times S^1 \to S^1 : (z_1, z_2) \to z_1 z_2$$

is C^∞ and gives S^1 the structure of an abelian group. The inverse map

$$S^1 \to S^1 : z \to z^{-1} = \overline{z}$$

is also C^∞. The map

$$e : \mathbb{R} \to S^1$$

given by

$$e(t) = e^{2\pi i t} = \cos(2\pi t) + i \sin(2\pi t)$$

is a C^∞ homomorphism of groups. It is moreover a C^∞ covering (see definition 120 below).

Example 114. The n-dimensional **torus** T^n is the C^∞ manifold

$$T^n = (S^1)^n = \underbrace{S^1 \times S^1 \times \cdots \times S^1}_{n}.$$

The group operations on T^n are C^∞. The map

$$\mathbb{R}^n \to T^n : (t_1, t_2, \cdots, t_n) \to (e(t_1), e(t_2), \cdots, e(t_n))$$

is C^∞ and a homomorphism of groups. The inclusion

$$T^n \to \mathbb{C}^n \simeq \mathbb{R}^{2n}$$

is C^∞.

Example 115. Real n-dimensional projective space is denoted by $\mathbb{R}P^n$ or $P^n(\mathbb{R})$ or simply P^n:

$$\mathbb{R}P^n = P^n(\mathbb{R}) = P^n$$

It is defined to be the set of all lines through the origin in $\mathbb{R}^{n+1}$. Equivalently it is the set of equivalence classes on $\mathbb{R}^{n+1} \setminus \{0\}$ under the equivalence relation $x \equiv y$ iff $x = ty$ for some $t \in \mathbb{R}$. For $x \in \mathbb{R}^{n+1}$ let $[x] \in \mathbb{R}P^n$ denoted its equivalence class. A C^∞ atlas for $\mathbb{R}P^n$ is given by $n+1$ charts $(\alpha_1, U_1), \ldots, (\alpha_{n+1}, U_{n+1})$ where

$$U_k = \alpha_k^{-1}(\mathbb{R}^n)$$

and

$$\alpha_k^{-1}(x_1, \ldots, x_n) = [x_1, \ldots, x_{k-1}, 1, x_k, \ldots, x_n].$$

The map $S^n \to \mathbb{R}P^n : x \to [x]$ is C^∞ double covering. (See definition 120 below.)

Example 116. Complex projective n-space is denoted by $\mathbb{C}P^n$ or sometimes $\mathbb{C}P_n$ or $P^n(\mathbb{C})$:

$$\mathbb{C}P^n = \mathbb{C}P_n = P^n(\mathbb{C})$$

It is a manifold of dimension $2n$. [1] It is defined to be the set of complex 1-dimensional subspaces of the complex vector space $\mathbb{C}^{n+1}$. (Warning: a vector space of complex dimension k has real dimension $2k$.) Equivalently $\mathbb{C}P^n$ is the set of equivalence classes on $\mathbb{C}^{n+1} \setminus 0$ under the equivalence relation $u \equiv v$ iff $u = \lambda v$ for some $\lambda \in \mathbb{C}$. For $v \in \mathbb{C}^{n+1}$ let $[v]$ denote its equivalence class. A C^∞ atlas for $\mathbb{C}P^n$ consists of $n+1$ charts (α_k, U_k) for $k = 1, \ldots, n+1$ where

$$\alpha_k(U_k) = \mathbb{C}^n \simeq \mathbb{R}^{2n}$$

and

$$\alpha_k^{-1}(z) = [z_1, \ldots, z_{k-1}, 1, z_k, \ldots, z_m]$$

[1] This is an example of a complex manifold, i.e. the overlap maps are holomorphic. Some authors indicate the real dimension of a manifold as a superscript writing e.g. M^m for a manifold of dimension m. For a complex manifold it is customary to indicate the complex dimension as a subscript.

for $z = (z_1, \dots, \_n) \in \mathbb{C}^n$. The map

$$\mathbb{R}^{2n+2} \setminus 0 = \mathbb{C}^{n+1} \setminus 0 \to \mathbb{C}P^n : v \to [v]$$

is C^∞ as well as the map

$$S^{2n+1} \to \mathbb{C}P^n : v \to [v]$$

(Note: $S^{2n+1} \subset \mathbb{R}^{2n+2}$.) This last map is surjective; hence $\mathbb{C}P^n$ is compact.

The manifold $\mathbb{C}P^1$ is C^∞ diffeomorphic to S^2 and the special case $n = 1$ of the above map

$$S^3 \to \mathbb{C}P^1 \simeq S^2$$

is called the **Hopf Map**.

Example 117. The **Grassman manifold** of unoriented k-planes in $\mathbb{R}^n$ is denoted by $G_{k,n}$ or [2] sometimes $G_{k,n}(\mathbb{R})$ or $G_k(\mathbb{R}^n)$:

$$G_{k,n} = G_{k,n}(\mathbb{R}) = G_k(\mathbb{R}^n).$$

It has as its underlying subset the set of all k-dimensional (vector) subspaces of $\mathbb{R}^n$) (defined for $0 < k < n$). A C^∞ atlas for $G_{k,n}$ consists of all pairs (α, U) determined as follows: Let $\mathbf{E} \in G_{k,n}$ and let $\mathbf{F}$ be any (vector space) complement to $\mathbf{E}$ in $\mathbb{R}^n$; i.e., $\mathbb{R}^n = \mathbf{E} \oplus \mathbf{F}$. The pair $(\mathbf{E}, \mathbf{F})$ determines $\alpha, U)$ via

$$U = \alpha^{-1}(L(\mathbf{E}, \mathbf{F}))$$

$$\alpha^{-1}(A) = \mathsf{Graph}(A)$$

for $A \in L(\mathbf{E}, \mathbf{F})$ where

$$\mathsf{Graph}(A) = \{x + Ax : x \in \mathbf{E}\}.$$

The natural identification $G_{1,n+1} \to RP^n$ is a C^∞ diffeomorphism (RP^n has the C^∞ structure given by definition 115). The map

$$G_{k,n} \to G_{n-k,n} : \mathbf{E} \to \mathbf{E}^\perp$$

(where $\mathbf{E}^\perp$ denotes the orthogonal complement to $\mathbf{E}$ with respect to the usual inner product on $\mathbb{R}^n$) is a C^∞ diffeomorphism. The map

$$GL(n, \mathbb{R}) \to G_{k,n} : a \to a(\mathbb{R}^k \times 0)$$

is a C^∞ surjective map. The fibers of this map (i.e., inverse images of points) are left cosets $a \cdot GL(n, k; \mathbb{R})$ where $GL(n, k; \mathbb{R})$ is the subgroup of $GL(n, \mathbb{R})$ consisting of all $b \in GL(n, \mathbb{R})$ such that $b(\mathbb{R}^k \times 0) = \mathbb{R}^k \times 0$. Thus we have a canonical correspondence between $G_{k,n}$ and the space of left cosets $GL(n, \mathbb{R})/Gl(n, k; \mathbb{R})$:

$$G_{k,n} = GL(n, \mathbb{R})/GL(n, k; \mathbb{R}).$$

(Warning: $GL(n, k; \mathbb{R})$ is not a normal subgroup of $GL(n, \mathbb{R})$. Thus $G_{k,n}$ is not a group in any natural fashion.)

[2] More generally, $G_k(\mathbf{H})$ denotes the manifold of k-dimensional subspaces of the vector space $\mathbf{H}$.

Example 118. The **Grassman manifold** of oriented k-planes in $\mathbb{R}^n$ is denoted by $G^+_{k,n}$ and has as its underlying set the set of all oriented vector spaces $(\mathbf{E}, o)$ where $\mathbf{E} \in G_{k,n}$. An element $(\mathbf{E}, o) \in G^+_{k,n}$ and a complement $\mathbf{F}$ to $\mathbf{E}$ in $\mathbb{R}^n$ determine a chart (α, U) by

$$\alpha(U) = L(\mathbf{E}, \mathbf{F})$$

$$\alpha^{-1}(A) = (\mathsf{Graph}(A), o')$$

where o' is the orientation on $\mathsf{Graph}(A)$ such that

$$o = \overline{A}^* o'$$

and $\overline{A} : \mathbf{E} \to \mathsf{Graph}(A)$ is the linear isomorphism defined by

$$\overline{A}x = x + Ax$$

for $x \in \mathbf{E}$.

The map

$$G^+_{k,n} \to G_{k,n} : (\mathbf{E}, o) \to \mathbf{E}$$

is a C^∞ double covering. $G^+_{k,n}$ is connected and compact. As in example 117, there is a natural bijective correspondence with the left coset space:

$$G^+_{k,n} = GL^+(n, \mathbb{R})/GL^+(n, k; \mathbb{R})$$

where $GL^+(h, k; \mathbb{R})$ is the subgroup of all $b \in GL^+(n, \mathbb{R}) \cap GL(n, k; \mathbb{R})$ such that $b|\mathbb{R}^k \in GL^+(k, \mathbb{R})$.

Example 119. The **Grassman manifold** of complex k-planes in $\mathbb{C}^n$ is denoted by $G_{k,n}(\mathbb{C})$ and has as its underlying set the collection of all complex k-dimensional subspaces (real dimension $2k$) of $\mathbb{C}^n$. The charts are defined exactly as for $G_{k,n}$ (see example 117) except that all subspaces and maps are taken to be complex linear. [3] As in example 117 we have a natural bijective correspondence:

$$G_{k,n}(\mathbb{C}) = GL(n, \mathbb{C})/GL(n, k; \mathbb{C}).$$

The inclusion

$$G_{k,n}(\mathbb{C}) \to G_{2k,2n}(\mathbb{R})$$

is $\mathbb{C}^\infty$. (Warning: This map is not surjective.)

Definition 120. Let M and N be C^r manifolds and $\pi : M \to N$. Then π is a C^r **covering** iff every $y \in N$ has a open neighborhood V such that π maps each component of $\pi^{-1}(V)$ C^r-diffeomorphically onto V. If in addition, $\pi^{-1}(y)$ consists of exactly n points, then π is called an n-fold covering. A **double covering** is a 2-fold covering.

[3] More generally, $G_k(\mathbf{H}, \mathbb{C})$ denotes the manifold of complex k-dimensional subspaces of the complex vector space $\mathbf{H}$.

Example 121. The map $\mathbb{R}^n \to T^n$ of example 114 is a C^∞ covering.

Example 122. The maps $S^n \to RP^n$ of example 115 and $G^+_{k,n} \to G_{k,n}$ of example 118 are C^∞ double coverings.

Definition 123. Let G be a collection of maps from a space M to itself. For $x \in M$ the **orbit** of x by G is the set

$$Gx = \{f(x) : f \in G\}.$$

The **orbit space** of M by G is the set

$$M/G = \{Gx : x \in M\}.$$

The map

$$M \to M/G : x \to Gx$$

is called the **canonical projection**.

Theorem 124. *Let M be a C^r manifold and G be a finite group of C^r diffeomorphisms from M onto itself. Suppose each $f \in G$ distinct from the identity is fixed point free (i.e., for $f \in G$ either $f(x) = x$ for all $x \in M$ of $f(x) \neq x$ for all $x \in M$). Then there is a unique C^r structure on M/G such that the canonical projection $M \to M/G$ is a C^r covering.*

Example 125. Let $M = S^n$ and $G = \{id, f\}$ where f is the antipodal map (i.e., $f(x) = -x$). Then M/G is projective space RP^n.

Example 126. Let $M = S^1 \times \mathbb{R}$ and $G = \{id, f\}$ where $f(z, x) = (-z, -x)$ ($S^1 \subset \mathbb{R}^2$). The orbit space M/G is called the Mobius strip. Note that the map f is orientation reversing: the map $S^1 \to S^1 : z \mapsto -z$ preserves orientation and the map $\mathbb{R} \to \mathbb{R} : x \mapsto -x$ reverses orientation.

Example 127. Let $M = T^2 = S^1 \times S^1$ and $G = \{id, f\}$ where $f(z_1, z_2) = (-z_1, \bar{z}_2)$. The orbit space M/G is called the **Klein bottle**.

Example 128. Let $M = S^3 = \{(v, w) \in \mathbb{C}^2 : v\overline{v} + w\overline{w} = 1\}$. Let p and q be relatively prime integers and $G = \{\mathsf{id}, f, f^2, \ldots, f^{p-1}\}$ where

$$f(v, w) = (e^{2\pi i/p}v, e^{2\pi iq/p}w).$$

The orbit space M/G is denoted by $L(p, q)$ and is called a **Lens space**.

3.3 The Tangent Space

Throughout M, N are C^r manifolds with $r \geq 1$.

Definition 129. Two C^1 curves $c_i : I_i \to M$ (where $I_i \subset \mathbb{R}$ is an interval about zero) are said to be **tangent** at $x \in M$ iff $c_1(0) = c_2(0) = x$ and $(\alpha \circ c_1)'(0) = (\alpha \circ c_2)'(0)$ for some chart (α, U) with $x \in U$. This definition is independent of the choice of the chart (α, U); i.e. if $c_1(0) = c_2(0) = x$ then $(\alpha \circ c_1)'(0) = (\alpha \circ c_2)'(0)$ for *some* chart (α, U) with $x \in U$ if and only if this is true for *every* chart (α, U) with $x \in U$. Tangency at x is an equivalence relation on the set of all C^1 curves $c : I \to M$ such that $c(0) = x$.

Definition 130. If $c : I \to M$ is C^1 we denote by $\dot{c}(0)$ the equivalence of c. The **tangent space** to M at $x \in M$ is denoted by M_x or $T_x M$ and defined to be the set of all equivalence classes $\dot{c}(0)$ with $c(0) = x$. If $c : I \to M$ is C^1 we define $\dot{c}(t) \in T_{c(t)} M$ by

$$\dot{c}(t) = \dot{c}_t(0)$$

for $t \in I$ where $c_t(s) = c(t + s)$.

Definition 131. Let $f : M \to N$ be C^1. For $x \in M$ the **tangent map** of f at x is the map

$$T_x f : T_x M \to T_{f(x)} N$$

defined by

$$T_x f(\dot{c}(0)) = \dot{\widehat{f \circ c}}(0),$$

where $c : I \to M$ is any C^1-curve in M thru x. One easily checks that $T_x f$ is well defined; i.e. if $c_1(0) = c_2(0) = x$ and $\dot{c}_1(0) = \dot{c}_2(0)$ then $\dot{\widehat{f \circ c}}(0) = \dot{\widehat{f \circ c}}(0)$.

Remark 132. Let $M \subset \mathbf{E}$ be an open subset of a vector space an $x \in M$. Then $T_x M$ and $\mathbf{E}$ are in natural correspondence via

$$v = \dot{c}(0)$$

for $v \in \mathbf{E}$ where

$$c(t) = x + tv.$$

We henceforth identify $T_x M$ with $\mathbf{E}$. If $N \subset \mathbf{F}$ is open in a vector space and $f : M \to N$ is C^1 then $T_x f : T_x M = \mathbf{E} \to T_{f(x)} N = \mathbf{F}$ is the map defined by

$$(T_x f)v = Df(x)v = \frac{d}{dt} f(x + tv)\bigg|_{t=0}$$

for $x \in M$ and $v \in \mathbf{E} = T_x N$.

Definition 133. Let M be any manifold and $x \in M$. We shall give each tangent space $T_x M$ the structure of a vector space. Let (α, U) be a chart at x. $T_x \alpha : T_x M \to T_{\alpha(x)} U = \mathbf{E}$ is bijective; we introduce the vector space structure in $T_x M$ by defining the vector space operations so that $T_x \alpha$ is a linear isomorphism. This vector space structure is well-defined i.e., independent of the choice of (α, U). This is because if (β, V) is another chart at x then the map

$$(T_x \beta) \circ (T_x \alpha)^{-1} = D(\beta \circ \alpha^{-1})(\alpha(x))$$

is a linear isomorphism.

Proposition 134. *If $f : M \to N$ is C^1, then $T_x f : T_x M \to T_{f(x)} N$ is linear for each $x \in M$.*

Definition 135. The **tangent bundle** of the manifold M is the map

$$\tau_M : TM \to M$$

defined by

$$TM = \{(x, v) : \; x \in M, \;\; v \in T_x M\}$$

and

$$\tau_M(x, v) = x.$$

Given a C^1-map

$$f : M \to N$$

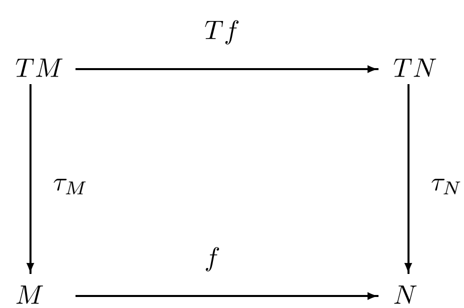
the **tangent map** to f is the map

$$Tf : TM \to TN$$

defined by

$$Tf(x, v) = (f(x), (T_x f)v).$$

Clearly, $\tau_N \circ Tf = f \circ \tau_M$, i.e. the diagram



commutes. When $M \subset \mathbf{E}$ and $N \subset \mathbf{F}$ are open subsets of vector spaces the fomula for

$$Tf : TM = M \times \mathbf{E} \to TN = N \times \mathbf{F}$$

is

$$Tf(x, v) = (f(x), \, Df(x)v)$$

for $x \in M$ and $v \in \mathbf{E}$.

The charts $(T\alpha, TU)$ as (α, U) runs over the charts on M form a C^{r-1} atlas on TM which makes TM a C^{r-1} manifold and τ_M a C^{r-1} map. Charts on TM of form $(T\alpha, TU)$ where (α, U) is a chart on M are called **natural charts** on

TM. The local representative of $Tf : TM \to TN$ with repect to natural charts $(T\alpha, TU)$ on TM and $(T\beta, TV)$ on TN is

$$(Tf)_{T\beta,T\alpha} = Tf_{\beta\alpha}$$

It follows that Tf is C^{r-1} if f is C^r.

Proposition 136. *The tangent map to the identity map of M is the identity map of TM:*

$$T\,id_M = \mathsf{id}_{TM}\ .$$

The tangent map of a composition is the composition of the tangent maps:

$$T(g \circ f) = (Tg) \circ (Tf)$$

The last formula is nothing more than the chain rule for differentiation. This proposition says that T defines a functor, called appropriately the **tangent functor**, from the category of C^r manifolds to the category of C^{r-1} manifolds. (Compare proposition 101).

3.4 Submanifolds

Definition 137. Let M be a C^r manifold $r \geq 1$ and $W \subset M$ be a subset. A **submanifold chart** for W in M is a chart (α, U) on M such that $\alpha(U) = U_1 \times U_2$ (where U_i is an open set in a vector space $\mathbf{E}_i$) and

$$\alpha(U \cap W) = U_1 \times \{y_2\}$$

for some $y_2 \in U_2$. The subset W is a C^r **submanifold** of M iff for every $x \in W$ there is a C^r-submanifold chart (α, U) for W in M at x (i.e. $x \in U$). A **closed submanifold** of M is a submanifold or M which is a closed subset of M.

Let $W \subset M$ be a C^r submanifold of a C^r manifold. Then the collection of all pairs $(\alpha|U \cap W, U \cap W)$ as (α, U) ranges over the C^r charts having the submanifold property for W is a C^r atlas for W. (We have identified $U_1 \times \{y_2\}$ with the open set $U_1 \subset \mathbf{E}$.) A C^r submanifold is a C^r manifold; the C^r structure is the one given by this atlas.

Proposition 138. *Let $W \subset M$ be a C^r submanifold of a C^r manifold. Then*

(1) *the inclusion $\iota : W \to M$ is a C^r map;*

(2) *the manifold topology of W is the topology it inherits as a subset of M;*

(3) *for each $x \in W$, the linear map $T_x\iota : T_xW \to T_xM$ is injective.*

Remark 139. It is customary to abuse language and use the linear injection $T_x\iota : T_xW \to T_xM$ to identify T_xW with a vector subspace of T_xM: henceforth [4]

$$T_xW \subset T_xM$$

[4] This is an abuse since the tangent space was defined as a set of equivalence classes.

when W is a submanifold of M. In case $M \subset \mathbb{R}^m$ we combine this identification with the earlier identification $T_x\mathbb{R}^m = \mathbb{R}^m$ for $x \in \mathbb{R}^m$ we arrive at the equation

$$T_xW = \{c'(0) :\ c \in C^1(\mathbb{R}, \mathbb{R}^m),\ c(0) = x,\ c(\mathbb{R}) \subset W\}$$

for $W \subset \mathbb{R}^m$ a C^1 submanifold. Here $C^1(\mathbb{R}, \mathbb{R}^m)$ denotes the space of all C^1-curves in $\mathbb{R}^m$.

Definition 140. A subset W of a topological space M is **locally closed** iff every $x \in W$ has a neighborhood U in M such that $U \cap W$ is closed in U. Note that W is closed if and only if every $x \in M$ has a neighborhood U in M such that $U \cap W$ is closed in U.

Proposition 141. *A submanifold is locally closed (in the ambient manifold) but not necessarily closed.*

Example 142. If $\mathbf{F} \subset \mathbf{E}$ is a vector subspace of a finite dimensional vector space and $U \subset \mathbf{E}$ is open, then $U \cap \mathbf{F}$ is a C^∞ submanifold of U.

Example 143. S^n is a C^∞ submanifold of $\mathbb{R}^{n+1}$; the C^∞ structures on S^n given by example 137 and example 111 agree.

Example 144. If $f : M \to N$ is a C^r map, then

$$\mathsf{Graph}(f) = \{(x, y) \in M \times N : y = f(x)\}$$

is a C^r submanifold of $M \times N$ which is C^r diffeomorphic to M. In particular, for each $y \in N$, $M \times \{y\}$ is a C^r submanifold of $M \times N$.

Definition 145. Let $f : M \to N$ be $C^r(r \geq 1)$. A point $x \in M$ is a **regular point** of f iff $T_xf : T_xM \to T_{f(x)}N$ is surjective; otherwise x is a **critical point** of f. A point $y \in N$ is a **critical value** of f iff $f^{-1}(y)$ contains a critical point; otherwise it is a **regular value**.

Definition 146. Let W be a submanifold of a manifold M. Suppose W has dimension p and M has dimension m. Then the **codimension** of W in M is definied to be $m - p$.

Remark 147. An open subset of M is a submanifold of codimension zero (often called an open submanifold of M). The C^∞ structures given by example 137 and example 109 agree.

Similarly, a zero-dimensional manifold is a discrete point set. The codimension of a zero-dimensional submanifold is the dimension of the ambient manifold.

Proposition 148. *Let $f : M \to N$ be $C^r(r \geq 1)$ and let $y \in N$ be a regular value. Then $f^{-1}(y)$ is a C^r submanifold of M; its codimension is the dimension of N. The tangent space to $f^{-1}(y)$ at a point $x \in f^{-1}(y)$ is the kernal of the linear map $T_xf : T_xM \to T_{f(x)}N$:*

$$T_xf^{-1}(y) = \{v \in T_xM : T_xfv = 0\}.$$

Example 149. Take $M = \mathbb{R}^{n+1}$, $N = \mathbb{R}$, and $f(x) = \langle x, x \rangle$. Then 1 is a regular value of f so $f^{-1}(1) = S^n$ is a C^∞ submanifold of $\mathbb{R}^{n+1}$.

Example 150. The set of all linear maps $A \in L(\mathbb{R}^n, \mathbb{R}^p)$ of rank k is a C^∞ submanifold of $L(\mathbb{R}^n, \mathbb{R}^p)$.

Example 151. Let $O(n)$ denote **orthogonal group** in n dimensions, i.e. the set of all matrices $a \in GL(n, \mathbb{R})$ such that $\langle ax, ay \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{R}^n$. Define

$$f : GL(n, \mathbb{R}) \to L^2_s(\mathbb{R}^n, \mathbb{R})$$

by [5]

$$f(a)xy = \langle ax, ay \rangle - \langle x, y \rangle$$

for $x, y \in \mathbb{R}^n$. Then f is C^∞, $O(n) = f^{-1}(0)$, and 0 is a regular value of f. Hence $O(n)$ is a C^∞ submanifold of $GL(n, \mathbb{R})$. The dimension of $O(n)$ is $n(n-1)/2$. The tangent space to $O(n)$ at the identity matrix e is the space $so(n)$ of skew-symmetric matrices:

$$so(n) = T_e O(n) = \{A \in gl(n, \mathbb{R}) : A + A^* = 0\}.$$

The manifold $O(n)$ has two components; the identity component (i.e. the component containing e) consists of all $a \in O(n)$ having determinant 1:

$$SO(n) = \{a \in O(n) : det(a) = 1\}$$

(All matrices in $a \in O(n)$ have $det(a) = \pm 1$.)

3.5 Immersions and embeddings

Definition 152. Let $f : W \to M$ be a C^1 map of C^1 manifolds, and $x \in W$. We say that f is an *immersion at x* iff the linear map

$$T_x f : T_x W \to T_{f(x)} M$$

is injective. The map f is an **immersion** iff f is an immersion at x for every $x \in W$.

Proposition 153. *Let $f : W \to M$ be a C^r map of C^r manifolds and $x \in W$. Then f is an immersion at x if and only if there exist C^r charts (α, U) at x in W and (β, V) at $f(x)$ in M with $f(U) \subset V$ such that*

$$f_{\beta\alpha}(x_1) = (x_1, 0)$$

for all $x_1 \in \alpha(U)$. Here $\mathbf{E}_1$ is the ambient vector space of $\alpha(U)$, $\mathbf{E}_1 \times \mathbf{E}_2$ is the ambient vector space of $\beta(V)$, $\alpha(x) = 0$, $\beta(f(x)) = (0, 0)$, and $f_{\beta\alpha} = \beta \circ f \circ \alpha^{-1}$.

[5] $L^2_s(\mathbb{R}^n, \mathbb{R})$ is the vector space of all symmetric bilinear forms $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$; it has dimension $n(n+1)/2$.

Definition 154. Let W and M be topological spaces and $f : W \to M$ be continuous. Call f is **proper** iff it satisfies the following equivalent conditions:

(1) for every compact $K \subset M, f^{-1}(K) \subset W$ is compact;

(2) every $y \in M$ has a neighborhood V in M such that $f^{-1}(K)$ is compact for every compact $K \subset V$.

Call $f : W \to M$ is **locally proper** iff

(3) every $y \in f(W)$ has a neighborhood V in M such that $f^{-1}(K)$ is compact for every compact $K \subset V$.

Thus f is proper implies f is locally proper but not conversely. Note that if W is a compact Hausdorff space, any continuous $f : W \to M$ is proper.

Lemma 155. *Let W and M be C^r manifolds $(r \geq 1)$ and $f : W \to M$ a C^r bijective immersion. Then f is a C^r diffeomorphism.*

Remark 156. Lemma 155 is false if we do not require W to be second countable.

Definition 157. Let W and M be C^r manifolds and $f : W \to M$. The map f is a C^r **embedding** iff f is a C^r immersion and $f : W \to f(W)$ is a homeomorphism (where $f(W)$ has the topology it inherits as a subset of M). The map f is a C^r **closed embedding** iff f is a C^r embedding and a closed map (i.e., maps closed subsets of W to closed subsets of M).

Proposition 158. *Let $f : W \to M$ be an injective C^r immersion of C^r manifolds $(r \leq 1)$. Then the following are equivalent:*

(1) *f is a C^r embedding;*

(2) *f is locally proper;*

(3) *$f(W) \subset M$ is a C^r submanifold.*

Also the following are equivalent:

(4) *f is a C^r closed embedding;*

(5) *f is proper;*

(6) *$f(W) \subset M$ is a closed C^r submanifold.*

Corollary 159. *Let $f : W \to M$ be an injective C^r immersion. If W is compact, then f is a closed embedding.*

Example 160. There is an injective C^∞ immersion $f : \mathbb{R} \to T^2$ such that $f(\mathbb{R})$ is dense in T^2.

Example 161. There is an injective C^∞ immersion $f : \mathbb{R} \to \mathbb{R}^2$ such that $f(\mathbb{R})$ is closed in $\mathbb{R}^2$ but f is not an embedding.

Definition 162. Let $f_i : W_i \to M$ $(i = 1, 2)$ be injective C^r immersions $(r \geq 1)$ of C^r manifolds. The maps f_1 and f_2 are **immersion-equivalent** iff $f_1(W_1) = f_2(W_2)$ and $f_2 \circ f_1^{-1} : W_1 \to W_2$ is continuous.

Proposition 163. *If f_1 and f_2 are immersion-equivalent C^r injective immersions, then $f_2 \circ f_1^{-1}$ is a C^r diffeomorphism.*

Proposition 164. *Let $W \subset M$ be a C^r submanifold of a C^r manifold $(r \geq 1)$. Then*

(1) *The inclusion $W \to M$ is a C^r embedding.*

(2) *Any C^r injective immersion with image W is immersion-equivalent to the inclusion.*

Corollary 165. *Any injective immersion which is immersion-equivalent to some embedding is itself an embedding.*

Example 166. There exists C^∞ injective immersions $f_i : \mathbb{R} \to \mathbb{R}^2$ (for $i = 1, 2$) such that $f_1(\mathbb{R}) = f_2(\mathbb{R})$ but f_1 and f_2 are not immersion-equivalent.

Example 167. The map

$$f : T^2 \to \mathbb{R}^3$$

given by

$$f(e^{i\theta}, e^{i\phi}) = (\cos(\theta)(b + a\cos(\phi)),\ \sin(\theta)(b + a\cos\phi)), \pm a\sin(\phi))$$

(where $0 < a < b$) is a closed embedding.

Example 168. The map

$$S^2 \to \mathbb{R}^4 : (x, y, z) \to (x^2 - y^2, xy, xz, yz)$$

factors through an embedding $RP^2 \to \mathbb{R}^4$.

Remark 169. This terminology is common but not standard. Some authors use the term *imbedding* or *embedding* for injective immersion and the term *regular embedding* or *locally proper embedding* for embedding (see proposition 158). These authors also use the term *submanifold* for the image of an injective immersion. The example in example 166 shows the danger of this latter terminology.

3.6 Submersions and locally trivial maps

Definition 170. Let $q : P \to M$ be C^1 and $z \in P$. We say that q is a *submersion at z* iff z is a regular point of q, that is, iff the linear map

$$T_z q : T_z P \to T_{f(z)} M$$

is surjective. The map q is called a **submersion** iff it is a submersion at z for every $z \in P$.

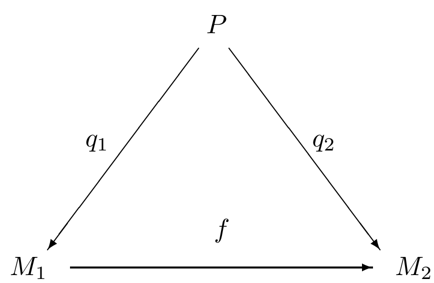
Proposition 171. *Let $q : P \to M$ be C^r ($r \geq 1$) and $z \in P$. Then q is a submersion at z if and only if there exist C^r charts (β, V) at z in P and (α, U) at $q(z)$ in M such that*

$$q_{\alpha\beta}(x, y) = x$$

for $(x, y) \in \beta(V)$. Here $\beta(V) \subset \mathbf{E} \times \mathbf{F}$, $\alpha(U) \subset \mathbf{E}$, $\beta(z) = (x_0, y_0)$, $\alpha(q(z)) = x_0$, and $q_{\alpha\beta} = \alpha \circ q \circ \beta^{-1}$.

Corollary 172. *A submersion is an open mapping (i.e., maps open sets to open sets). Hence, if $q : P \to M$ is a surjective submersion M has the topology induced from P by q (i.e. $U \subset M$ is open if and only if $q^{-1}(U) \subset P$ is open).*

Proposition 173. *Let P, M_1, and M_2 be C^r manifolds ($r \geq 1$) and $q_i : P \to M_i$ ($i = 1, 2$) be C^r surjective submersions. Suppose $f : M_1 \to M_2$ is a bijection such that $f \circ q_1 = q_2$; i.e. the diagram*



commutes. Then f is a C^r diffeomorphism.

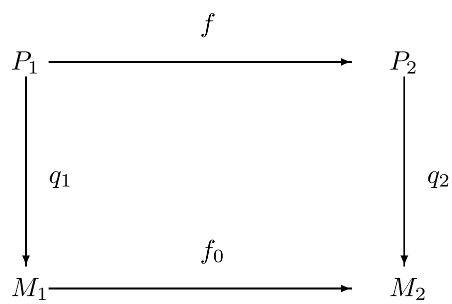
Remark 174. Proposition 173 says that if P is a C^r manifold, M is a set and $q : P \to M$ is a surjection, then there is at most one C^r structure on M such that q is a C^r submersion.

Definition 175. Let $q : P \to M$ and $x \in M$. The **fiber** of q over x is the set $q^{-1}(x)$. It is sometimes denoted by P_x or $P(x)$. Often P is called the **total space** of q and M is called the **base space** of q. A **section** of q is a map $\xi : M \to P$ such that $q \circ \xi = id_M$. A **local section** of q is a section of $q|q^{-1}(U) : q^{-1}(U) \to U$ where $U \subset M$ is open.

Proposition 176. *The fibers of a C^r submersion ($r \geq 1$) are closed C^r submanifolds of its total space.*

Proposition 177. *A C^r section of a C^r submersion is a closed embedding.*

Definition 178. Let P_i, M_i (for $i = 1, 2$) be C^r manifolds and $q_i : P_i \to M_i$. The maps q_1 and q_2 are C^r **equivalent** iff there exist C^r diffeomorphisms $f : P_1 \to P_2$ and $f_0 : M_1 \to M_2$ such that



Definition 179. Let P and M be C^r manifolds and $q : P \to M$. q is C^r **trivial** iff there exists a (nonempty) C^r manifold F such that $q : P \to M$ is C^r equivalent to the map

$$M \times F \to M : (x, v) \to x.$$

q is C^r **locally trivial** iff every $x \in M$ has an open neighborhood U in M such that $q^{-1} \mid q^{-1}(U) : q^{-1}(U) \to U$ is trivial.

Remark 180. Some authors call a projection $M \times F \to M$ *trivial* and a map equivalent to it *trivializable*. A locally trivial map is sometimes called a *fiber bundle* but usually this term is used in a more restrictive sense.

Proposition 181. *A trivial map is locally trivial.*

Proposition 182. *A C^r locally trivial map is a C^r surjective submersion.*

Proposition 183. *If $q : P \to M$ is C^r locally trivial and M is connected, then any two fibers of q are C^r diffeomorphic.*

Example 184. Let $P = \{(x, y) \in \mathbb{R}^2 :\mid y \mid > x\}, M = \mathbb{R}$, and $q : P \to M : q(x, y) = x$. Then q is a C^∞ surjective submersion which is not C^0 locally trivial.

Remark 185. A C^r covering (see definition 120) is a C^r locally trivial map with zero-dimensional fiber.

Proposition 186. *If M is a C^r manifold $(r \geq 1)$, then $\tau_M : TM \to M$ is C^{r-1} locally trivial.*

Remark 187. A C^∞ manifold is called **parallelizable** iff $\tau_M : TM \to M$ is C^∞ trivial. It is known that S^n is parallelizable if and only if $n = 1, 3, 7$.

Example 188. The map

$$S^{2n+1} \to \mathbb{C}P^n$$

(see example 111 and example 116) is C^∞ locally trivial. The fiber is (diffeomorphic to) S^1.

Example 189. Let $\mathbf{E}$ be a vector space. A k-**frame** in $\mathbf{E}$ is a linear injection from $\mathbb{R}^k$ to $\mathbf{E}$; i.e., an element of $L_{inj}(\mathbb{R}^k, \mathbf{E})$. If $k = dim(\mathbf{E})$ a k-frame is called a **frame**. $L_{inj}(\mathbb{R}^k, \mathbb{R}^n)$ is open in $L(\mathbb{R}^k, \mathbb{R}^n)$ and the map

$$L_{inj}(\mathbb{R}^k, \mathbb{R}^n) \to G_{k,n} : u \mapsto u(\mathbb{R}^k)$$

is C^∞ locally trivial. Here $G_{k,n}$ is the Grassman manifold of k-planes in $\mathbb{R}^n$ (see example 117).

Example 190. An **orthonormal** k-frame in $\mathbb{R}^n$ is a linear map $u : \mathbb{R}^k \to \mathbb{R}^n$ such that

$$\langle ux, uy \rangle = \langle x, y \rangle$$

for $x, y \in \mathbb{R}^k$ (Standard inner product on both $\mathbb{R}^k$ and $\mathbb{R}^n$). The set of all orthonormal k-frames in $\mathbb{R}^n$ is usually denoted by $V_k(\mathbb{R}^n)$ and is called the **Stiefel variety** of (orthonormal) k- frames in $\mathbb{R}^n$. $V_k(\mathbb{R}^n)$ is a C^∞ submanifold of $L_{inj}(\mathbb{R}^k, \mathbb{R}^n)$ and the map

$$V_k(\mathbb{R}^n) \to G_{k,n} : u \mapsto u(\mathbb{R}^k)$$

is C^∞ locally trivial. Moreover, the map

$$V_k(\mathbb{R}^n) \to (S^{n-1})^k : u \mapsto (ue_1, ue_2, \ldots, ue_k)$$

where for $i = 1, 2, \ldots, k$

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^k$$

(1 in the i-th position) is a closed embedding. Thus $V_k(\mathbb{R}^n)$ is often defined to be the image of this embedding.

Example 191. Let

$$E_{k,n} = \{(\mathbf{F}, v) \in G_{k,n} \times \mathbb{R}^n : v \in \mathbf{F}\}.$$

Then $E_{k,n}$ is a C^∞ submanifold of $G_{k,n} \times \mathbb{R}^n$ and the map

$$E_{k,n} \to G_{k,n} : (\mathbf{F}, v) \to \mathbf{F}$$

is C^∞ locally trivial.

Theorem 192. *A proper C^r surjective submersion is C^r locally trivial. In particular, a C^r surjective submersion with a compact total space is C^r locally trivial.*

3.7 Partitions of Unity

Throughout, M denotes a topological space and Λ is a set of continuous real valued functions on M. We shall assume that Λ is **admissible** by which we mean that it satisfies the following three conditions:

(1) if $f_1, \ldots, f_n \in \Lambda$ and $g : \mathbb{R}^n \to \mathbb{R}$ is C^∞, then $g \circ (f_1, \ldots, f_n) \in \Lambda$;

(2) if $f : M \to \mathbb{R}$ is such that for every $x \in M$ there is a neighborhood U of x in M and a function $g \in \Lambda$ with $f \mid U = g \mid U$, then $f \in \Lambda$;

(3) Λ induces the topology of M.

Remark 193. We recall the definition of the topology induced by a set of functions: Let X be a set and Γ a set of functions f with source X and target a topological space Y_f (i.e., $f : X \to Y_f$). Note that the target may vary. The topology on X **induced** by Γ is the topology which has the sets $f^{-1}(V)$ as f ranges over Γ and V ranges over open subsets of Y_f as a subbasis. It is the weakest topology (least open sets) in which all the maps $f \in \Gamma$ are continuous. For example, the Tychonoff topology on a product of topological spaces is the topology induced by the projections onto the factors.

Lemma 194. *. Let Λ be admissible. Then the sets $f^{-1}(0, \infty)$ where f ranges over Λ form a basis for the topology of M.*

Proposition 195. *If M is a C^r manifold $(0 \le r \le \infty)$, then $\Lambda = C^r(M, \mathbb{R})$ is an admissible set.*

Lemma 196. *A manifold is normal and Lindelöf.*

Definition 197. Let $f : M \to \mathbb{R}$. The **support** of f is denoted by $\mathsf{Supp}(f)$ and defined by

$$\mathsf{Supp}(f) = \mathsf{Clos}\{x \in M : f(x) \neq 0\}$$

where for $X \subset M$, $\mathsf{Clos}(X)$ denotes the closure of X.

Definition 198. A **partition of unity** on M is a collection $\{g_i\}_i$ of continuous real valued functions on M such that

(1) $g_i \ge 0$ for each i;

(2) every $x \in M$ has a neighborhood U such that $U \cap \mathsf{Supp}(g_i) = \phi$ for all but finitely many of the g_i;

(3) for each $x \in M$

$$\sum_i g_i(x) = 1.$$

Note that by (2), the sum in (3) is finite. If, in addition, each $g_\in \Lambda$ then $\{g_i\}_i$ is called a Λ-*partition of unity*.

Definition 199. A partition of unity $\{g_i\}_i$ on M is **subordinate** to an open cover of M iff for each g_i there is an element U of the cover such that $\mathsf{Supp}(g_i) \subset U$. The space M **admits** Λ-partitions of unity iff for every open cover of M there is a Λ-partition of unity subordinate to the cover.

Remark 200. A Hausdorff space admits continuous partitions of unity if and only if it is paracompact.

Theorem 201. *Let M be normal and Lindelöf, and let Λ be admissible. Then M admits Λ-partitions of unity.*

Proof. Suppose an open cover is given. Choose $f_n \in \Lambda$ $(1 \leq n < \infty)$ so that $f_n : M \to [0, \infty)$

$$U_n = \{x \in M : f_n(x) > 0\}$$

is a cover and $\{\overline{U}_n\}_n$ refines the given cover. Let

$$V_n = \{x \in U_n : f_k(x) < \frac{1}{n} \text{ for } 1 \leq k < n\}.$$

Refine $\{V_n\}$ to $\{W_n\}$ where

$$W_n = \{x \in M : h_n(x) > 0\}$$

and $h_n \in \Lambda$, $h_n : M \to [0, \infty)$. Let

$$g_n(x) = h_n(x) / \sum_k h_k(x).$$

Then $\{g_n\}$ is the required partition of unity. □

Corollary 202. *A C^r manifold admits C^r partitions of unity.*

Corollary 203. *A manifold is paracompact.*

Corollary 204 (C^r Urysohn's Lemma)**.** *Let X and Y be disjoint closed subsets of a C^r manifold M. Then there is a C^r function $g : M \to \mathbb{R}$ such that $0 \leq g(x) \leq 1$ for all $x \in M$, $g(x) = 0$ for $x \in X$ and $g(x) = 1$ for $x \in Y$.*

Corollary 205. *Any closed subset of a C^r manifold is the zero set of a C^r real-valued function.*

3.8 Embeddings

Definition 206. A topological space X has **covering dimension at most** m iff every open cover of X has a refinement with the property that each point of X is contained in at most $m + 1$ members of the refinement. The **covering dimension** of X is the least m such that X has covering dimension at most m; it is ∞ if there is no such m.

Theorem 207. *An m-dimensional manifold has covering dimension m.*

Proposition 208. *Let X be a paracompact space with covering dimension m. Then every open cover of X has a refinement which is the union of collections $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_m$ such that any of the elements of each $\mathcal{C}_i$ $(i = 0, 1, \ldots, m)$ are pairwise disjoint.*

Corollary 209. *An m-dimensional manifold can be covered by $m + 1$ charts.*

Theorem 210. *Let M be an m-dimensional C^r manifold $(r \geq 0)$ and let $k = (m+1)^2$. Then there exists a C^r closed embedding $f : M \to R^k$.*

Corollary 211. *A manifold is metrizable.*

Sketch of proof of theorem 210. Let $(\alpha_i, U_i)(0 < i \leq m)$ be a cover of M by charts. Choose a cover $V_0, \ldots, V_m$ of M with $\overline{V}_i \subset U_i$ and $g_i : M \to \mathbb{R}$ with $0 \leq g_i \leq 1$, $g_i \mid M \backslash U_i \equiv 0$, $g_i \mid \overline{V}_i \equiv 1$. Define $g_i \alpha_i : M \to \mathbb{R}^m$ by

$$\begin{aligned} g_i \alpha_i(x) &= g_i(x)\alpha_i(x) && \text{for } x \in U_i \\ &= 0 && \text{for } x \notin U_i. \end{aligned}$$

Define $f : M \to \mathbb{R}^k$ by

$$f = (g_0,\, g_0\alpha_0,\; g_1,\, g_1\alpha_1, \ldots, g_m,\, g_m\alpha_m).$$

f is a closed embedding. For 208 and 209, see Munkres.

Chapter 4

TRANSVERSALITY

4.1 Function Space Topologies

Recall that a **subbase** for a topological space C is a collection $\mathcal{S}$ of subsets of C such that the open subsets of C are precisely those sets expressible as arbitrary unions of finite intersections of members of $\mathcal{S}$. Every collection $\mathcal{S}$ of subsets of a set C is a subbase for a unique topology on C (viz. that topology whose open sets are the arbitrary unions of finite intersections of S) but it is often helpful to know several subbases for a topological space.

For example, let M and N be topological spaces let $C(M, N)$ be the set of all continuous maps $f : M \to N$. The set $C(M, N)$ has two interesting topologies, the **Whitney** topology having as subbase the collection of all sets

$$\{f \in C(M, N) : \mathsf{Graph}(f) \subset Z\}$$

where $Z \subset M \times N$ is open and the **compact open** topology having as subbase the collection of all sets

$$\{f \in C(M, N) : f(K) \subset V\}$$

where K is a compact subset of M and V is an open subset of N. [1] In this section we write

$$C(M, N)_{wh}$$

is used to denote the space $C(M, N)$ with the Whitney topology and

$$C(M, N)_{co}$$

is used to denote the space $C(M, N)$ with the compact open topology. However, for us the only important topology on $C(M, N)$ is the Whitney topology and in subsequent sections we shall always assume that $C(M, N)$ has this topology unless the contrary is stated.

[1] The Whitney topology is also called the **fine** topology or the **strong** topology and the compact open topology is also called the **coarse** topology or the **weak topology**.

Proposition 212. *The Whitney topology is stronger than the compact-open topology; i.e. the identity map*

$$C(M,N)_{wh} \to C(M,N)_{co}$$

is continuous.

Proof. Given K and V we have

$$f(K) \subset V \iff \mathsf{Graph}(f) \subset (M \setminus K) \times V$$

for all $f \in C(M,N)$. □

Proposition 213. *If M is compact and Hausdorff, then the Whitney and compact-open topologies on $C(M,N)$ are the same; i.e. the identity map*

$$C(M,N)_{wh} \to C(M,N)_{co}$$

is a homeomorphism.

Proof. Given $f \in C(M,N)$ and $Z \subset M \times N$ satisfying $\mathsf{Graph}(f) \subset Z$ we must find compact subsets $K_1, K_2, \ldots K_l$ of M and open subsets $V_1, V_2, \ldots V_l$ of N such that for any $g \in C(M,N)$ the conditions $g(K_i) \subset V_i$ for $i = 1, 2, \ldots l$ imply that $\mathsf{Graph}(f) \subset Z$. For every $x \in M$ choose a compact neighborhood K_x of x in M and V_x of $f(x)$ in N so that

$$K_x \times V_x \subset Z$$

and extract from the open cover $\{int(K_x)\}_{x \in M}$ a finite subcover. For any $x' \in M$ we have $x' \in K_x$ for some K_x of the finite subcover, whence $g(x') \in V_x$ if $g(K_x) \subset V_x$, whence $(x, g(x)) \subset K_x \times V_x \subset Z$. This shows that $\mathsf{Graph}(g) \subset Z$ for any g satisfying $g(K_x) \subset V_x$ for all K_x of the finite subcover (as required). □

Proposition 214. *If $L \subset N$ is closed then the set*

$$\{f \in C(M,N)_{wh} : f(M) \cap L = \emptyset\}$$

is open in $C(M,N)_{wh}$.

Proof. $f(M) \cap L = \emptyset$ if and only if $\mathsf{Graph}(f) \subset M \times (N \setminus L)$. (It is because this poroposition fails for $C(M,N)_{co}$ that the compact open topology is of little use for us.) □

In case M is locally compact and N is metrizable we have alternate descriptions of these topologies. Fix a metric d giving the topology of N and given a collection

$$\mathcal{P} = \{(K_i, \epsilon_i) : i \in I\}$$

where $\{K_i\}_{i \in I}$ forms a locally finite collection of compact subsets of M and each ϵ_i is a positive real number and given a map $f : M \to N$ let

$$\mathcal{W}(f, \mathcal{P}) = \{g \in C(M,N) : d(g(x), f(x)) < \epsilon_i \ \ \forall x \in K_i \ \forall i \in I\}.$$

Proposition 215. *The collection of all sets $\mathcal{W}(f, \mathcal{P})$ forms a basis for the topology $C(M,N)_{wh}$; the collection of all sets $\mathcal{W}(f, \mathcal{P})$ with I finite forms a basis for the topology $C(M,N)_{co}$.*

4.2 Jets

Given positive integers m and n and a non-negative integer r denote by $J^r(m,n)$ the space of all polynomial maps $p : \mathbb{R}^m \to \mathbb{R}^n$ of degree $\leq r$; i.e.

$$p \in J^r(m,n) \iff p(v) = \sum_{|\kappa| \leq r} p_\kappa v^\kappa$$

for $v \in \mathbb{R}^m$ where the coefficients p_κ lie in $\mathbb{R}^n$ and we have used multi-index notation

$$v^\kappa = v_1^{\kappa_1} v_2^{\kappa_2} \ldots v_m^{\kappa_m}$$
$$|\kappa| = \kappa_1 + \kappa_2 + \ldots + \kappa_m$$

for an m-tuple

$$\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$$

of non-negative integers. Clearly $J^r(m,n)$ is a vector space and a clever argument (due to Paul Ehrenfest) shows that its dimension is given by

$$dim(J^r(m,n)) = n\binom{r+m}{r}.$$

We shall often use the identification

$$J^r(m,n) = \prod_{k=0}^{r} L^k_{sym}(\mathbb{R}^m, \mathbb{R}^n).$$

Here the symbol $\prod$ denotes Cartesian product and $L^k_{sym}(\mathbb{R}^m, \mathbb{R}^n)$ denotes the space of symmetric k-linear maps from $(\mathbb{R}^m)^k$ to $\mathbb{R}^n$. With this identification the polynomial p above would be written:

$$p(v) = \sum_{k=0}^{r} p_k v^k$$

where

$$p_k \in L^k_{sym}(\mathbb{R}^m, \mathbb{R}^n)$$

and notation $p_k v^k$ simply denotes the value of the k-linear map p_k for k identical inputs v. Evidently

$$p_k v^k = \sum_{|\kappa|=k} p_\kappa v^\kappa.$$

Given open sets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ define an open subset $J^r(U,V) \subset \mathbb{R}^m \times J^r(m,n)$ by

$$J^r(U,V) = \{(x,p) \in U \times J^r(m,n) : p_0 \in V\}.$$

Given a C^r map $f : U \to V$ define a continuous map

$$j^r f : U \to J^r(U,V)$$

by

$$j^r f(x)(v) = (x, \sum_{|\kappa| \le r} \frac{\partial^\kappa f(x) v^\kappa}{\kappa!})$$

where the multi-index notation for partial derivatives is given by

$$\partial^\kappa = \partial_1^{\kappa_1} \partial_2^{\kappa_2} \dots \partial_m^{\kappa_m}$$

and the factorial notation is defined by

$$\kappa! = \kappa_1! \, \kappa_2! \, \dots \kappa_m!.$$

Thus first co-ordinate of $j^r f(x)$ is x and the second is the Taylor polynomial of f of order r evaluated at x. In the alternate notation

$$j^r f(x) = (x, f(x), Df(x), \frac{1}{2} D^2 f(x), \dots, \frac{1}{r!} D^r f(x)).$$

The set $J^r(U, V)$ is called the space of r-**jets** from U to V, the point $j^r f(x)$ is called the r-**jet** of f at x, and the map $j^r f$ is called the r-**jet extension** of f. Since a map with prescribed derivatives at a single point can easily be constructed, we have that every point $(x, p) \in J^r(U, V)$ has form $(x, p) = j^r f(x)$ for some $f \in C^\infty(U, V)$.

Proposition 216. *Given open sets $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$, and $W \subset \mathbb{R}^k$ there is a unique polynomial map*

$$\mathcal{E}^r : J^r(U, V, W) \to J^r(U, W)$$

such that

$$j^r(g \circ f)(x) = \mathcal{E}^r(j^r g(y), j^r f(x))$$

for any $f \in C^r(U, V)$ and $g \in C^r(V, W)$. Here the set $J^r(U, V, W)$ is defined by

$$J^r(U, V, W) = \{((y, q), (x, p) \in J^r(V, W) \times J^r(U, V) : p_0 = y\}$$

Proof. This is nothing more than the chain rule (for higher order derivatives). The exact formula for $\mathcal{E}^r$ is called Faa di Bruno's formula. When $r = 1$ it is given by

$$\mathcal{E}^1((y, z, B), (x, y, A)) = (x, z, BA)$$

for $x \in U$, $y \in V$, $z \in W$, $A \in L(\mathbb{R}^m, \mathbb{R}^n)$, and $B \in L(R^n, \mathbb{R}^k)$. □

Corollary 217. (Change of variables formaula) *Let $U_1, U_2 \subset \mathbb{R}^m$ and $V_1, V_2 \subset \mathbb{R}^n$ be open and let*

$$\phi : U_1 \to U_2$$

and

$$\psi : V_1 \to V_2$$

be C^s diffeomorphisms. Then for $r \leq s$ there is a unique map

$$j^r_{\psi\phi} : J^r(U_1, V_1) \to J^r(U_2, V_2)$$

such that

$$j^r_{\psi\phi}(j^r g_1(x_1)) = j^r g_2(x_2)$$

whenever $x_i \in U_i$ and $g_i \in C^r(U_i, V_i)$ satisfy

$$x_2 = \phi(x_1)$$

and

$$g_2 = \psi \circ g_1 \circ \phi^{-1};$$

the map $j^r_{\psi\phi}$ is in fact a C^{s-r} diffeomorphism.

Remark 218. In the special case $r = 1$ the map

$$j^1_{\psi\phi} : J^1(U_1, V_1) \to J^1(U_2, V_2)$$

has form

$$j^1_{\psi\phi}(x, y, A) = (\phi^{-1}(x), \psi(y), D\psi(y)\, A\, D\phi^{-1}(x))$$

where we have made the identification

$$J^1(U, V) = U \times V \times L(\mathbb{R}^m, \mathbb{R}^n).$$

Now let M and N be C^s manifolds with $s \geq r$ and call two pairs $(x_1, f_1), (x_2, f_2) \in M \times C^r(M, N)$ *equivalent* iff $x_1 = x_2$, $f_1(x_1) = f_2(x_2)$ and for some charts (α, U) on M at $x_1 = x_2$ and (β, V) on N at $f_1(x_1) = f_2(x_2)$ we have

$$j^r(\beta \circ f_1 \circ \alpha^{-1})(x_1) = j^r(\beta \circ f_2 \circ \alpha^{-1})(x_2).$$

This is indeed an equivalence relation and by the change of variables formula it is independent of the choice of charts. We denote the equivalence class of the pair (x, f) by $j^r f(x)$ and the set of all equivalence classes by $J^r(M, N)$. As before $j^r f(x)$ is called the r-**jet** of f at x, $J^r(M, N)$ is called the space of r-**jets** from M to N, and for $f \in C^r(M, N)$ the map $j^r f \in C^0(M, J^r(M, N))$ is called the r-**jet extension** of f.

Given charts (α, U) on M and (β, V) on N we have the natural inclusion $J^r(U, V) \subset J^r(M, N)$ and a bijection

$$j^r_{\beta\alpha} : J^r(U, V) \to J^r(\alpha(U), \beta(V))$$

given by

$$j^r_{\beta\alpha}(j^r f(x)) = j^r(\beta \circ f \circ \alpha^{-1})(\alpha(x))$$

for $x \in U$, $f(x) \in V$. Any two of these charts are C^{s-r}-compatible (change of variables formula again) and so constitute a C^{s-r} atlas on $J^r(M, N)$. Henceforth we denote by $J^r(M, N)$ the C^{s-r} manifold determined by this atlas.

For $x \in M$ denote by $J^r(M,N)_x$ the set of all r=jets based at x:

$$J^r(M,N)_x = \{j^r f(x) : f \in C^r(M,N)\}$$

and for $(x,y) \in M \times N$ denote by $J^r(M,N)_{(x,y)}$ the set of all r-jets based at x with target y:

$$J^r(M,N)_{(x,y)} = \{j^r f(x) : f \in C^r(M,N),\ f(x) = y\}.$$

There is a natural bijection

$$J^1(M,N)_{(x,y)} \longleftrightarrow L(T_xM, T_yN)$$

determined by

$$j^1 f(x) \longleftrightarrow T_x f$$

for $f : M \to N$ with $y = f(x)$. (See remark 218.)

For $s \geq r$ there is a unique map called the **natural projection**

$$\pi^r_s : J^s(M,N) \to J^r(M,N)$$

such that

$$(\pi^r_s) \circ (j^s f) = j^r f$$

for every $f : M \to N$. Of course in local co-ordinates the map (π^r_s) is nothing more then the projection which discards terms of order $> r$. However, there is no natural way to define an inclusion of $J^r(M,N)$ into $J^s(M,N)$. In local coordinates we can of course simply extend by 0 – i. e. put the coeficients $p_\alpha = 0$ for $r < |\alpha| \leq s$ but this operation does not commute with changes of variables.

4.3 The C^r Whitney Topology

Let M and N be C^r manifolds and note that the r-jet extension

$$j^r : C^r(M,N) \to C^0(M, J^r(M,N))$$

is injective. (It is never surjective; e.g. in local coordinates with $r = 1$ a map $F(x) = (F_1(x), F_2(x), F_3(x)) \in U \times V \times L(\mathbb{R}^m, \mathbb{R}^n)$ satisfies $F = j^1 f$ for some f iff $F_1(x) = x$, and $F_3(x) = DF_2(x)$.) for all x.) The image $j^r(C^r(M,N))$ of this map is a closed subset of $C^0(M, J^r(M,N))$ because if a sequence of functions together with their derivatives of order $\leq r$ converge uniformly on compact subsets then the limit function is of class C^r and the limit of the derivatives is the derivative of the limit.

Definition 219. For $0 \leq r < \infty$ the C^r topology [2] on $C^r(M,N)$ is the topology induced from the (Whitney) topology on $C^0(M, J^r(M,N)$ by r-jet extension. The C^∞ topology on $C^\infty(M,N)$ is the union of the C^r topologies for $r < \infty$.

[2] More precisely, the C^r Whitney topology

Hence by definition the map

$$j^r : C^r(M,N) \to C^0(M, J^r(M,N))$$

is a closed embedding.

An alternate description of the Whitney topology of proposition 215 has an analog here which is useful for reducing arguments to local coordinates. Fix $f \in C^r(M,N)$ and let

$$\mathcal{P} = \{(\alpha_i, U_i), (\beta_i, V_i), K_i, \epsilon_i\}_{i \in I} \tag{4.1}$$

denote a system consisting of a collection $\{(\beta_i, V_i)\}_{i \in I}$ of charts on N; a collection $\{(\alpha_i, U_i)\}_{i \in I}$ of charts on on M satisfying $f(U_i) \subset V_i$ for $i \in I$; a locally finite collection $\{K_i\}_{i \in I}$ of compact compact subsets of M satisfying $K_i \subset U_i$ for $i \in I$; a collection $\{\epsilon_i\}_{i \in I}$ of positive real numbers. For this map f and system $\mathcal{P}$ define a subset $\mathcal{W}^r(f, \mathcal{P}) \subset C^r(M,N)$ by

$$g \in \mathcal{W}^r(f,\mathcal{P}) \iff \begin{cases} g(K_i) \subset V_i, & |D^k g_i(x) - D^k f_i(x)| < \epsilon_i \\ & \forall k \leq r,\, i \in I,\, x \in \alpha_i(K_i). \end{cases}$$

Here g_i and f_i denote the local representatives of f and g:

$$f_i(x) = \beta_i \circ f \circ \alpha_i^{-1}(x)$$

$$g_i(x) = \beta_i \circ g \circ \alpha_i^{-1}(x)$$

for $x \in \alpha(K_i)$. Note that when M is compact the index set I must be finite since the collection $\{K_i\}_{i \in I}$ is locally finite.

Proposition 220. *The sets $\mathcal{W}^r(f, \mathcal{P}C^r(M,N)$ form a basis for the C^r topology $C^r(M,N)$.*

Recall that a subset of a topological space is called **residual** [3] iff it is a countable intersection of open dense sets (archetypal example: irrational numbers are a residual subset of the real numbers) and a topological space is called a **Baire space** if every residual set is dense. According to the *Baire Category Theorem* a complete metric space is a Baire space.

Proposition 221. *The space $C^r(M,N)$ is a Baire space.*

The idea is that (in Baire spaces at least) residual sets are very large. One often uses the terminology "for generic f in C the property $P(f)$ holds" to mean "the set of f in C for which the property $P(f)$ holds is a residual subset of C". Another way of saying a set is large (at least a subset of Euclidean space) is to say that its complement has measure zero. These two notions are not the same [4] but at least a closed set of measure zero has open dense (hence residual) complement.

---

[3] Second category in older parlance.

[4] The ultimate counter example is constructed from a Cantor set of nearly full measure. An increasing union of such sets will have full measure while its complement is residual.

4.4 Sard's Theorem

Let M be a C^r manifold of dimension m, N be a C^r manifold of dimension n, and $f : M \to N$ a C^r map with $r \geq 1$. A point $x \in M$ is called a **critical point** of f if the derivative

$$T_x f : T_x M \to T_{f(x)} N$$

is not surjective; otherwise it is called a **regular point** of f. A point $y \in N$ is called a **critical value** of f iff $y = f(x)$ for some critical point x of f; otherwise it is called a **regular value** of f.

We denote the critical points, regular points, critical values, and regular values of f by $\mathcal{CP}(f)$, $\mathcal{RP}(f)$, $\mathcal{CV}(f)$ and $\mathcal{RV}(f)$ respectively so that for $x \in M$ we have

$$x \in \mathcal{CV}(f) \iff (T_x f)(T_x M) \neq T_{f(x)} N$$

and

$$\mathcal{RP}(f) = M \setminus \mathcal{CP}(f),$$
$$\mathcal{CV}(f) = f(\mathcal{CP}(f)),$$
$$\mathcal{RV}(f) = N \setminus \mathcal{CV}(f).$$

Proposition 222. *The set $\mathcal{CV}(f) \subset M$ of critical points of $f : M \to N$ is closed; hence if M is compact then the set $\mathcal{CV}(f))$ of critical values is also closed. If the dimension m of M is less than the dimension n of N, then every point is a critical point: $\mathcal{CV}(f) = M$.*

Theorem 223 (Sard's Theorem)**.** *Let*

$$\rho(m, n) = max\left(0, 1 + \frac{(m-n)^2}{2}\right)$$

and $r > \rho(m, n)$. Then for any C^r map $f : M \to N$ from an m-dimensional manifold M to an n-dimensional manifold N, the set $N \setminus \mathcal{RV}(f))$ of regular values is residual.

Actually, one can improve this result to

$$\rho(m, n) = max(0, m - n)$$

but this requires more work. The properties of $\rho(m, n)$ required to make the proof presented here go through are

$$q \leq \rho(m, n) \tag{4.2}$$

$$\rho(m-1, n) \leq \rho(m, n) - q + 1 \tag{4.3}$$

$$\rho(m-1, n-1) \leq \rho(m, n) \tag{4.4}$$

where [5]

$$q = [m/n] + 1.$$

[5] $[m/n]$ is the greatest integer in m/n

As a corollary of 223 the set of non-values $N \setminus f(M)$ is residual when $m < n$ but this is false when $r = 0$ as there are space filling curves, i.e, surjective continuous maps $f : \mathbb{R} \to \mathbb{R}^2$. Whitney (1935) has constructed an example of a C^1 map $f : \mathbb{R}^2 \to \mathbb{R}$ and a C^0 curve $c : \mathbb{R} \to \mathbb{R}^2$ such that $Df(c(t)) = 0$ for $t \in \mathbb{R}$ but $t \mapsto c(f(t))$ is not constant. If c is C^1 the chain rule shows that this cannot happen, but also if f is C^r with $r > \rho(2,1)$ (and c only C^0) Sard's theorem 223 shows that this cannot happen. (For if $f(c(t))$ takes two values and $c : \mathbb{R} \to \mathcal{CP}(f)$ then $\mathcal{CV}(f)) \subset \mathbb{R}$ contains an interval and therefore cannot have residual complement – a residual set is dense.)

Proof. Proof of 223 As a countable intersection of residual sets is residual it is enough to prove for every point $x \in M$ there are neighborhoods U of x in M and V of $f(x)$ in N with $f(U) \subset V$ and for which the regular values of $f|U$ intersect V in a residual subset of V. Indeed, if this local assertion is true we can find a countable cover $\{K_i\}_{i \in \mathbb{N}}$ of M by compact sets so that $K_i \subset U_i \subset M$, $f(U_i) \subset V_i \subset N$ and $V_i \setminus f(\mathcal{CV}(f) \bigcap U_i))$ is residual (and hence dense) in V_i. Hence the larger set $N \setminus f(\mathcal{CV}(f) \cap K_i)$ is open and dense in N so the countable intersection

$$\mathcal{RV}(f) = N \setminus f(\mathcal{CV}(f)) = \bigcap_{i \in N} (N \setminus f(\mathcal{CV}(f) \cap K_i))$$

is residual. These considerations show that we may use co-ordinate charts to localize the theorem; i. e. that we may (and will) assume [6] M is an open subset of $\mathbb{R}^m$ and N is an open subset of $\mathbb{R}^n$. We will prove that $\mathcal{CV}(f)$ is of measure zero from which (as it is a countable union of compact sets) it follows that its complement is residual.

We can write M as a finite union

$$M = Z_q \cup \bigcup_{1 \le |\kappa| < q} \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} Z_{\kappa,i,j} \cup \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} Z_{i,j}$$

where

$$Z_q = \{x \in M : \partial^\kappa f_j(x) = 0 \ (1 \le j \le n, \ \ 1 \le |\kappa| \le q)\}$$

$$Z_{\kappa,i,j} = \{x \in M : \partial^\kappa f_j(x) = 0, \ \partial_i \partial^\kappa f_j(x) \neq 0\}$$

$$Z_{i,j} = \{x \in M : \partial_i f_j(x) \neq 0\}$$

Here $f_1, f_2, \ldots, f_n$ are the components of f. Note that $Z_q \subset \mathcal{CP}(f)$, that $Z_{\kappa,i,j}$ is a submanifold of M of codimension 1 (dimension $m-1$) and of class $C^{r-|\kappa|}$ and that $Z_{i,j}$ is an open subset of M.

Our proof will be by induction on m. We divide the proof into three steps:

Step (1) $f(Z_q)$ has measure zero.

Step (2) $f(\mathcal{CP}(f) \cap Z_{\kappa,i,j})$ has measure zero.

[6] We will use the localization argument again however.

Step (3) $f(\mathcal{CP}(f) \cap Z_{i,j})$ has measure zero.

Proof. Step (1) By our localization argument it is enough to show that $f(I^m \cap Z_q)$ has measure zero where

$$I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_m, b_m]$$

is a cube of edge length L; viz. a product of intervals of equal length

$$L = b_1 - a_1 = b_2 = a_2 = \ldots = b_m - a_m.$$

We write I as a union of cubes I_l ($l = 1, 2, \ldots k^m$) of edge L/k so that

$$f(\mathbb{Z}_q) \subset \bigcup_{l=1}^{k^m} f(\mathbb{Z}_q \cap I_l)$$

If $\mathbb{Z}_q \cap I_l = \emptyset$ then its image $f(\mathbb{Z}_q \cap I_l)$ is also empty and hence certainly of measure zero. Hence assume $\mathbb{Z}_q \cap I_l \neq \emptyset$ and choose $x_0 \in \mathbb{Z}_q \cap I_l \neq \emptyset$. Then by Taylor's formula we have an inequality

$$|f(x) - f(x_0)| \leq C|x - x_0|^{q+1}$$

where

$$C = \sup_{x \in I} |D^{q+1} f(x)|.$$

The cube I_i is contained in a ball of radius $\sqrt{m}L/k$ so our inequality implies that the set $f(I_l)$ is contained in a ball centered at $f(x_0)$ and with radius $C(\sqrt{m}L/k)^{q+1}$ and hence in a cube with edge $2C(\sqrt{m}L/k)^{q+1}$. This proves

$$\mathsf{meas}(f(I_l)) \leq (2C(\sqrt{m}L/k)^{q+1})^n$$

As $f(Z_q \cap I_l \subset f(I_l)$ and there are at most k^m cubes I_l with $Z_q \cap I_l$ non-empty this gives

$$\mathsf{meas}(f(Z_q)) \leq C' k^{m-(q+1)n}$$

where $C' = (2C(\sqrt{m}L)^{q+1})^n$. As the exponent $m - (q+1)n$ is negative by 4.2 we let $k \to \infty$ to achieve $\mathsf{meas}(f(Z_q)) = 0$.

Proof. Step (2) For any submanifold $Z \subset M$ we have

$$f(\mathcal{CP}(f) \cap Z) \subset f(\mathcal{CP}(f|Z))$$

for if $y \in f(\mathcal{CP}(f) \cap Z)$ there is an $x \in \mathcal{CP}(f) \cap Z$ with $y = f(x)$ and hence a $\hat{y} \in T_y N$ such that the linear equation $(T_x f)\hat{x} = \hat{y}$ has no solution $\hat{x} \in T_x M$. Hence the linear equation certainly has no solution $\hat{x} \in T_x W \subset T_x M$ i.e. $x \in \mathcal{CV}(f|Z)$ as required. As the manifold $Z_{\kappa,i,j}$ is of class $C^{r-|\kappa|}$ and with $|\kappa| < q$ it is of class C^{r-q+1}. By 4.3 the induction hypothesis applies to $f|Z_{\kappa,i,j}$ as required.

Proof. Step (3) Choose $x_0 \in Z_{i,j}$ and use the inverse function theorem to change co-ordinates so that

$$f_j(x_1, x_2, \ldots, x_m) = x_i$$

for $x = (x_1, x_2, \ldots, x_m)$ near x_0. Then permute the co-ordinates so $i = j = 1$ to bring f to the form

$$f(t, u) = (t, g(t, u))$$

where $t = x_1$, $u = (x_2, \ldots, x_m)$ and $g : M \to R^{n-1}$. (We have used the localization argument to assume without loss of generality that $M = Z_{1,1}$.) Let

$$M_t = \{u \in \mathbb{R}^{m-1} : (t, u) \in M\}$$

$$N_t = \{v \in \mathbb{R}^{n-1} : (t, v) \in N\}$$

and

$$g_t : M_t \to N_t$$

be given by

$$g_t(u) = g(t, u).$$

Since

$$Df(t, u) = \begin{bmatrix} 1 & 0 \\ \star & Dg_t(u) \end{bmatrix}$$

we have that

$$\mathcal{CP}(f) = \{(t, u) \in M : u \in \mathcal{CP}(g_t)\}$$

and hence that

$$\mathcal{CV}(f) = \{(t, v) \in N : v \in \mathcal{CV}(g_t)\}$$

i.e.

$$f(\mathcal{CV}(f)) \cap N_t = g_t(\mathcal{CV}(g_t))$$

By 4.4 the induction hypothesis applies to each g_t so that $\mathcal{CV}(f)$ intersects each slice N_t in a set of measure zero (in $\mathbb{R}^{n-1}$). Now the proof is complete via Fubini's theorem. □

Remark 224. There is a generalization of Sard's theorem to Banach manifolds (infinite dimensional) due to Smale. It goes as follows. One assumes that $f : M \to N$ is a **Fredholm** map which means that the dimensions of the kernel and cokernel or the linear map $T_x f : T_x M \to T_x N$ are finite dimensional for every $x \in M$. Then one defines the **Fredholm index** of f by

$$\mathsf{index}(f) = \dim(\ker(T_x f)) - \dim(\mathsf{coker}(T_x f))$$

(it follows that this index is locally constant). Smale's theorem is that the regular values of f form a residual subset of N provided that f is C^r with $r > \max(0, \mathsf{index}(f))$.

4.5 The Submanifold Theorem

Let M and N be C^r manifolds ($r \geq 1$), $f : M \to N$ a C^r map, and $W \subset N$ a C^r submanifold.

Definition 225. We say that f is **transversal to** W **at** the point $x \in M$ (notation: $f \pitchfork_x W$) iff either $f(x) \notin W$ or else $f(x) \in W$ and every vector in $T_{f(x)}N$ can be written as the sum of a vector in the image of $T_x f : T_x M \to T_{f(x)}N$ and a vector tangent to W. Thus if $f(x) \in W$ we have

$$f \pitchfork_x W \iff T_{f(x)}N = (T_x f)(T_x M) + T_{f(x)}W.$$

For a subset $K \subset M$ we say f is **transverse to** W **on** K (notation: $f \pitchfork_K W$) iff f is transverse to W at every $x \in K$:

$$f \pitchfork_K W \iff f \pitchfork_x W \ \ \forall x \in K$$

Finally we say f is **transverse to** W (notation: $f \pitchfork W$) iff f is transverse to W on M:

$$f \pitchfork W \iff f \pitchfork_M W.$$

Theorem 226. *If f is transverse to W then $f^{-1}(W)$ is a C^r submanifold of M. Moreover, the codimension is preserved:*

$$\mathsf{codim}(f^{-1}(W), M) = \mathsf{codim}(W, N).$$

Proof. Choose $x_0 \in f^{-1}(W)$ and choose a submanifold chart (β, V) for W in N at $f(x_0)$. Thus $\beta(V) = V_1 \times V_2 \subset \mathbb{R}^{n-k} \times \mathbb{R}^k = \mathbb{R}^n$ and $\beta(V \cap W) = V_1 \times \{0\}$. Choose a chart (α, U) at x_0 with $f(U) \subset V$. The the local reprsentative $f_{\beta\alpha} = \beta \circ f \circ \alpha^{-1}$ has form

$$f_{\beta\alpha}(x) = (f_1(x), f_2(x))$$

where $f_i : \alpha(U) \to V_i$ for $i = 1, 2$. Clearly then

$$\alpha(U \cap W) = \{x \in \alpha(U) : f_2(x) = 0\}$$

which will give a non-degenerate local defining equation for $f^{-1}(W)$ if $Df_2(x) : \mathbb{R}^m \to \mathbb{R}^k$ is surjective. But the tangent space to W is given by

$$T_y\beta(V \cap W) = \mathbb{R}^{n-k} \times \{0\} \subset \mathbb{R}^n = T_y\beta(V)$$

so the definition of transversality says that (when $f_2(x) = 0$) given any $\hat{y} = (\hat{y_1}, \hat{y_2}) \in \mathbb{R}^{n-1} \times \mathbb{R}^k$ we can solve the equation

$$\hat{y} = Df(x)\hat{x} + (\hat{w}, 0)$$

for $\hat{x} \in \mathbb{R}^m$ and $\hat{w} \in \mathbb{R}^{n-k}$. This equation can be written as two equations

$$\hat{y_1} = Df_1(x)\hat{x} + \hat{w}$$

$$\hat{y_2} = Df_2(x)\hat{x}$$

which can (always) be solved iff $Df_2(x)$ is surjective as claimed. □

4.6 The Openness Theorem

Let M and N be C^r manifolds ($r \geq 1$), $f : M \to N$ a C^r map, and $W \subset N$ a C^r submanifold.

Theorem 227. *Assume that W is closed in N. Then the set*

$$\pitchfork^r (M, N, W) = \{f \in C^r(M, N) : f \pitchfork W\}$$

is open in $C^r(M, N)$.

Proof. Recall the identification

$$J^1(M, N) = \{(x, y, A) : x \in M,\ y \in N,\ A \in L(T_x M, T_y N)\}$$

of 4.2. With this identification define

$$Z = \{(x, y, A) : y \in W,\ T_y N \neq A(T_x M) + T_y W\}$$

so that the subset $Z \subset J^1(M, N)$ has the property that

$$f \pitchfork W \iff (j^1 f)(M) \cap Z = \emptyset.$$

Moreover Z is closed in $J^1(M, N)$ (since W is and by the next lemma) so that in the case $r = 1$ theorem follows from proposition 214. The general case $r > 1$ follows trivially since the inclusion map $C^r(M, N) \to C^1(M, N)$ is continuous. □

Lemma 228. *Let $\mathbf{E}$ and $\mathbf{F}$ be finite dimensional vector spaces and $\mathbf{H}$ be a subspace of $\mathbf{F}$. Then the set of all linear maps $A : \mathbf{E} \to \mathbf{F}$ such that*

$$\mathbf{F} = A\mathbf{E} + \mathbf{H}$$

is open in $L(\mathbf{E}, \mathbf{F})$.

The theorem fails when W is not closed. For example let $M = \mathbb{R}$, $N = \mathbb{R}^2$, $W = (0, \infty) \times \{0\}$ (the positive x-axis) and $f(x) = (x - a, (x - a)^2)$. The f is transverse to W for $a = 0$ (as the image of f and W do not intersect) but not for $a > 0$. (Exercise: "interchange" the parabola and the ray: let f be given by $f(x) = a + e^x$ and $W = \{(x, y) : y = x^2\}$. Why doesn't this contradict the theorem?) The above proof however does have something to say even when W is not closed.

Proposition 229. *Drop the assumption that W is closed. Let $K \subset M$ be closed in M and $L \subset W$ be closed in N. Then the set*

$$\pitchfork^r (M, N, W, K, L) = \{f \in C^r(M, N) : f \pitchfork_{K \cap F^{-1}(L)} W\}$$

is open in $C^r(M, N)$.

4.7 The Abstract Density Theorem

Theorem 230. *Let $\mathcal{A}$, M, N be C^r manifolds ($r \geq 1$), $W \subset N$ a C^r submanifold and*

$$f : \mathcal{A} \times M \to N$$

be C^r. For each $a \in \mathcal{A}$ define

$$f_a : M \to N$$

by

$$f_a(x) = f(a, x).$$

Let m be the dimension of M and k be the codimension of W in N. Assume

(1) $f \pitchfork W$;

(2) $r > n - k$.

Then the set

$$\mathcal{A}_W = \{a \in \mathcal{A} : f_a \pitchfork W\}$$

is residual in $\mathcal{A}$.

Proof. We form the submanifold

$$f^{-1}(W) \subset \mathcal{A} \times M$$

and the map

$$\pi : f^{-1}(W) \to \mathcal{A} : (a, x) \mapsto \pi(a, x) = a$$

and observe that

$$\mathcal{RV}(\pi) = \mathcal{A}_W.$$

□

Remark 231. Using Smale's generalization of Sard's theorem explained in 224 the proof goes through even when $\mathcal{A}$ is an (infinite dimensional) Banach manifold (e.g. $\mathcal{A} = C^r(M, N)$ when M is compact.)

4.8 The Jet Transversality Theorem

Theorem 232. *Let $s > r \geq 0$, M and N be C^∞ manifolds, $W \subset J^r(M, N)$ a C^∞ submanifold. Then the set*

$$T^s(M, N, W) = \{f \in C^r(M, N) : f \pitchfork W\}$$

is residual in $C^s(M, N)$.

Proof. The neatest proof (at least when M is compact) is via the theory of Banach manifolds. The function space $C^s(M,N)$ is a Banach manifold and the map

$$C^s(M,N)\times M\to J^r(M,N):(f,x)\mapsto j^rf(x)$$

is a C^{s-r} submersion. (We'll essentialy prove that the derivative is onto below.) A submersion is trivially transverse to any submanifold so the transversality hypothesis of the abstract transversality theorem holds. We will give a proof which avoids the theory of Banach manifolds.

Choose charts (α,U) on M (β,V) on N a compact set $K\subset U$ and a compact set $L\subset W\cap J^r(U,V)$. Define an open subset $C^s((M,K),(N,V))$ of $C^s(M,N)$ by

$$C^s((M,K),(N,V))=\{f\in C^s(M,N):f(K)\subset V\}.$$

We first show that the set

$$T^s_{K,V,L}=\{f\in C^s((M,K),(N,V)):j^rf\pitchfork_{K\cap(j^rf)^{-1}(L)}W\}$$

is open dense in $C^s((M,K),(N,V))$.

The set $T^s_{K,V,L}$ is open by 227 and the continuity of $j^r:C^s(M,N)\to C^1(M,J^r(M,N))$ For density choose $f\in C^s((M,K),(N,V))$; we must approximate f by $f_a\in T^s_{K,V,L}$. Choose a C^s function $h:M\to\mathbb{R}$ supported in U and identically one on a neighborhood U' of K:

$$K\subset U'\subset h^{-1}(1)\subset \mathsf{Supp}(h)\subset U.$$

Let m and n be the dimensions of M and N respectively and $\mathcal{A}$ be a small neighborhood of 0 in the vector space $J^r(m,n)$. For $a\in\mathcal{A}$ define $f_a:M\to N$ by

$$\beta\circ f_a(x)=\beta\circ f(x)+h(x)a(\alpha(x))$$

for $x\in U$ and $f_a(x)=f(x)$ otherwise. (For the neighborhood $\mathcal{A}$ sufficiently small the right hand side lies in V). The map

$$F:\mathcal{A}\times U'\to J^r(M,N)$$

defined by

$$F(a,x)=j^rf_a(x)$$

is a submersion of class C^{s-r}. (To see this note that in local co-ordinates

$$F(a,x)=(x,f(x)+a_0,\,Df(x)+a_1,\,\ldots,\,D^rf(x)+a_r)$$

where we have suppressed the notations for α and β. Thus

$$DF(a,x)(\hat{x},\hat{a})=(\hat{x},+Df(x)\hat{x}+\hat{a}_0,\ldots,D^{r+1}f(x)\hat{x}+\hat{a}_r)$$

and the equation

$$DF(a,x)(\hat{x},\hat{a})=(\hat{x}_1,\hat{a}_1)$$

can be solved for $(\hat{x}, \hat{a})$ for any choice of $(\hat{x}_1, \hat{a}_1)$.) A submersion is trivially transverse to every submanifold so that we may apply the abstract transversality theorem 230 to conclude that for a residual (and hence dense) set of $a \in \mathcal{A}$ we have that $j^r f_a \pitchfork W$. Since $K \subset U'$ and $L \subset W$ any such f_a lies in $T^s_{K,V,L}$ *a fortiori*. (Note that hyposthesis in that s must be large enough causes no difficulty here; as we already know that $T^s_{K,V,L}$ is open in $\{f \in C^s((M,K),(N,V))$ and the C^∞ maps are dense.)

Next we show that each $f_0 \in C^s(M,N)$ has a neighborhood Q in which $T^s(M,N,W)$ is dense. (This shows that $T^s(M,N,W)$ is dense in $C^r(M,N)$.) For this choose a (countable) locally finite cover $\{(\alpha_i, U_i)\}_{i\in I}$ of M by charts so small that there is a family $\{(\beta_i, V_i\}_{i\in I}$ of charts on N with $f_0(U_i) \subset V_i$. Choose for each index $i \in I$ a compact set $K_i \subset U_i$ so that the collection $\{K_i\}_{i\in I}$ covers M and take

$$Q = \{f \in C^r(M,N) : f(K_i) \subset U_i,\ \forall i \in I\}.$$

For each i choose a countable cover $\{L_{i,n}\}_{\{}n \in \mathbb{N}\}$ of $W \cap J^r(U_i, V_i)$ by compact sets. Then

$$Q \cap T^s(M,N,W) = \bigcap_{i,n} T^s_{K_i,V_i,L_{i,n}}$$

is residual in Q and hence dense in Q.

Finally write W as a countable union

$$W = \bigcup_n L_n$$

of closed sets. Then each of the sets

$$T^s_n = \{f \in C^r(M,N) : f \pitchfork_{f^{-1}(L_n)} L_n f^{-1}(L_n)\}$$

is open by 229 and dense (in $C^r(M,N)$) as it contains $T^s(M,N,W)$. Hence the intersection

$$T^s(M,N,W) = \bigcap_n T^s_n$$

is residual in $C^r(M,N)$ as required. □

Corollary 233. *Let $f : M \to N$ be a C^s map of C^s manifolds $s \geq 1$ and $W \subset N$ a C^s submanifold. Then the set*

$$T^s(M,N,W) = \{f \in C^s(M,N) : f \pitchfork W\}$$

is residual in $C^s(M,N)$.

Proof. This is essentially a special case ($r = 0$) of the jet transversality theorem. We use the identification

$$J^0(M,N) = M \times N, \quad j^0 f(x) = (x, f(x))$$

and put

$$W' = M \times W \subset J^0(M,N).$$

Then $j^0 f \pitchfork W' \longleftrightarrow f \pitchfork W$. □

4.9 The Transversality Isoptopy Theorem

Recall that a **homotopy** of maps from M to N is a map

$$J \to C^0(M, N) : t \mapsto f_t$$

for which the evaluation map

$$J \times M \to N : (t, x) \mapsto f_t(x)$$

is continuous. Here $J \subset \mathbb{R}$ is an interval, possibly all of $\mathbb{R}$. When M and N are manifolds we say that the homotopy is *smooth* or of class C^r whene the evaluation map is. An **isotopy** is a homotopy where each f_t is a homeomorphism and a **diffeotopy** is an isootopy where each f_t is a diffeomorphism.

Proposition 234. *Let M and N be smooth manifolds. Assume that M is compact and $r \geq 0$. Then $C^r(M, N)$ is locally smoothly path connected. This means that every $f_0 \in C^r$ has a neighborhood Q in $C^r(M, N)$ such that for $f_1 \in Q$ there is a C^r homotopy*

$$f : \mathbb{R} \times M \to N$$

with $f(0, x) = f_0(x)$, $f(1, x) = f_0(1)$, and $f_t \in Q$ for $x \in M$ and $t \in \mathbb{R}$ (and where $f_t(x) = f(t, x)$.

Theorem 235. *Let M and N be smooth manifolds an $W \subset N$ a C^∞ submanifold. Assume that M is compact and W is closed. Let*

$$f : \mathbb{R} \times M \to N$$

be a smooth homotopy such that each f_t is transverse to W:

$$f_t \pitchfork W$$

for all $t \in \mathbb{R}$ (where $f_t(x) = f(t, x)$. Then the various submanifolds $f_t^{-1}(W) \subset M$ are all ambient isotopic. More precisely there is a smooth diffeotopy

$$\mathbb{R} \to \mathsf{Diff}(M) : t \mapsto \phi_t$$

such that

$$\phi_t(f_t^{-1}(W)) = f_0^{-1}(W)$$

for all $t \in \mathbb{R}$.