NOTES ON DIFFERENTIABLE MANIFOLDS

Marc Culler and John Wood - Fall 2008
(errors are due to Marc Culler)

1. Smooth functions

We will use the notation D_i to denote the partial derivative of a real-valued function of several variables with respect to the i^{th} variable. That is, if U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is a function then $D_i f$ is the function given by

$$D_i f = \lim_{t \to 0} \frac{f(u + te_i) - f(u)}{t},$$

where e_i denotes the i^{th} standard basis vector of $\mathbb{R}^n$. The domain of $D_i f$ consists of all points of U where the limit exists. If $i_1, \ldots, i_k$ are integers between 1 and n then we will say that $D_{i_1} \cdots D_{i_k} f$ is an *order k partial derivative* of f. The partial derivative of order 0 is defined to be the function f itself.

Definition 1.1. Let U be an open set in $\mathbb{R}^n$. A real-valued function $f : U \to \mathbb{R}$ is said to be *smooth* if all of its partial derivatives of all orders are defined and continuous at each point of U. If K is an arbitrary subset of $\mathbb{R}^n$, a function $f : K \to \mathbb{R}$ is said to be smooth if every point of K has an open neighborhood U in $\mathbb{R}^n$ such that $f|_{K \cap U}$ extends to a smooth function defined on U.

It is important to realize that a smooth function need not have a power series expansion. If $U = (a, b) \subset \mathbb{R}$ and $f : U \to \mathbb{R}$ is smooth then the Taylor series centered at any point $x_0 \in (a, b)$ is defined, since all derivatives exist at x_0. But it may not converge to f in any open interval containing x_0.The following standard example illustrates this.

Example 1.2.

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-1/x} & \text{if } x > 0 \end{cases}$$

This function is smooth on all of $\mathbb{R}$, and its Taylor series centered at 0 is identically 0. But $f(x) > 0$ for all $x > 0$, so the Taylor series does not converge to f in any open interval containing 0.

Exercise 1.1. Verify that the function f defined above is smooth on $\mathbb{R}$.

Date: September 25, 2008.

Problem 1.2. Let $a < b < c < d$. Construct a smooth function $g(x)$ on $\mathbb{R}$ such that

- $0 \leq f(x) \leq 1$ for all $x \in \mathbb{R}$;
- $f(x) = 0$ if and only if $x \in (-\infty, a] \cup [d, \infty)$;
- $f(x) = 1$ if and only if $x \in [b, c]$.

Even though smooth functions may not have Taylor expansions, they are nonetheless well approximated by polynomial functions.

To explain this we first recall a basic result from calculus:

Theorem 1.3. *Suppose that f is a real-valued function whose partial derivatives of order 1 and 2 are defined and continuous on an open set $U \subset \mathbb{R}^2$. Then $D_1 D_2 f = D_2 D_1 f$.*

Proof. Fix an arbitrary point $u \in U$. Let $R = [a, b] \times [c, d]$ be any rectangle with $u \in R \subset U$. Since $D_2 D_1 f$ is continuous, we may compute its integral over R as an iterated integral:

$$\begin{aligned} \int_R D_2 D_1 f \, dA &= \int_a^b \int_c^d D_2 D_1 f(x, y) dy \, dx \\ &= \int_a^b (D_1 f(x, d) - D_1 f(x, c)) \, dx \\ &= f(b, d) - f(a, d) - f(b, c) + f(a, c). \end{aligned}$$

Similarly,

$$\begin{aligned} \int_R D_1 D_2 f \, dA &= \int_c^d \int_a^b D_1 D_2 f(x, y) dx \, dy \\ &= \int_c^d (D_2 f(b, y) - D_2 f(a, y)) \, dx \\ &= f(b, d) - f(b, c) - f(a, d) + f(a, c). \end{aligned}$$

In particular, the two functions $D_2 D_1 f$ and $D_1 D_2 f$ have the same average value on the rectangle R.

Next consider a sequence (R_n) of rectangles containing u whose diameters tend to 0. Let M_n and m_n denote the respective maximum and minimum values of $D_2 D_1 f$ on the compact set R_n. We have $M_n \to D_2 D_1 f(u)$ and $m_n \to D_2 D_1 f(u)$ as $n \to \infty$. But

$$m_n \leq \frac{1}{\text{Area } R_n} \int_{R_n} D_2 D_1 f \, dA \leq M_n.$$

Thus, by the "squeeze principle"

$$\frac{1}{\text{Area } R_n} \int_{R_n} D_2 D_1 f \, dA \quad \to \quad D_2 D_1 f(u) \text{ as } n \to \infty.$$

A symmetrical argument shows that

$$\frac{1}{\text{Area } R_n} \int_{R_n} D_1 D_2 f \, dA \quad \to \quad D_1 D_2 f(u) \text{ as } n \to \infty.$$

Since the two averages are equal, it follows that $D_2 D_1 f(u) = D_1 D_2 f(u)$. $\square$

Exercise 1.3. Generalize Theorem 1.3 to functions of n variables.

Exercise 1.4. Suppose U is an open set in $\mathbb{R}^n$ and f is smooth on U. If N is a non-negative integer and $u \in U$, show that there is a unique polynomial T of degree N such that partial derivatives of f and T of order at most N agree at the point u.

Definition 1.4. Let f be a smooth real-valued function on an open set $U \subset \mathbb{R}^n$. Let $u = (u_1, \dots, u_n) \in U$. The *degree N Taylor polynomial of f centered at u* is the unique polynomial P_u of degree N such that all of the partial derivatives of f and P_u of order at most N agree at the point u. Using traditional calculus notation, and applying Exercise 1.3, we may write

$$P_u(x_1, \dots, x_n) = \sum \frac{1}{k_1! \cdots k_n!} \frac{\partial^{k_1+\cdots+k_n} f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(u)\,(x_1 - u_1)^{k_1} \cdots (x_n - u_n)^{k_n}.$$

where the sum runs over all n-tuples $(k_1, \dots, k_n)$ of non-negative integers such that $k_1 + \cdots + k_n \le N$.

Lemma 1.5. *Suppose that f is a real-valued function which is smooth on the closed interval $[a, b] \subset \mathbb{R}$. Suppose that $f^{(i)}(a) = 0$ for $i = 0, \dots, k-1$ and that $|f^{(k)}(t)| \le M$ for all $t \in [a, b]$. Then*

$$f(t) \le \frac{M}{k!}(t-a)^k$$

for all $t \in [a, b]$.

Proof. The proof is by induction on k. The base case $k = 0$ is immediate. For the induction step we suppose that the result holds for $k \ge 0$ and that f satisfies $f^{(i)}(a) = 0$ for $i = 0, \dots, k$ and $|f^{(k+1)}(t)| < M$ for all $t \in [a, b]$. If we set $g = f'$ then g satisfies:

- $g^{(i)}(a) = 0$ for $i = 0, \dots, k-1$; and
- $|g^k(t)| < M$ for all $t \in [a.b]$.

Thus, by the induction hypothesis,

$$|g(s)| < \frac{M}{k!}(s-a)^k \text{ for all } s \in [a, b].$$

For $t \in [a, b]$, since $f(a) = 0$, we have

$$|f(t)| = |\int_a^t g(s)\,ds| \le \int_a^t |g(s)|\,ds \le \int_a^t \frac{M}{k!}(s-a)^k\,ds = \frac{M}{(k+1)!}(t-a)^{k+1}$$

$\square$

Lemma 1.6. *Let U be a convex open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a smooth function. Suppose that $u_0 \in U$ is a point such that all of the partial derivatives of f of order at* *most k vanish at u_0. Suppose also that $sup_{u\in U}|g(u)| \leq M$ for each order $k+1$ partial derivative g of f. Then*

$$|f(u)| \leq \frac{Mn^{(k+1)/2}}{(k+1)!}||u-u_0||^{k+1}$$

for all $u \in U$.

Proof. For $u_0 \neq u \in U$, let $v = v_1 e_1 + \cdots + v_n e_n$ be the unit vector $\frac{1}{||u-u_0||}(u-u_0)$. Set $h(t) = f(u_0 + tv)$, for $t \in [0, ||u-u_0||]$. It follows from the chain rule for partial derivatives that

$$h^{(j)}(t) = \sum D_{i_1} \cdots D_{i_j} f(u_0 + tv) v_{i_1} \cdots v_{i_j},$$

where the sum is taken over all sequences $(i_1, \ldots, i_j)$ of integers between 1 and n. In particular we have $h^{(j)}(0) = 0$ for $j = 0, \ldots, k$. Moreover, indexing the sum as above, we have

$$\begin{aligned} |h^{(k+1)}(t)| &\leq M \sum |v_{i_1} \cdots v_{i_j}| \\ &= M(|v_1| + \cdots + |v_n|)^{k+1} \\ &\leq M(\sqrt{n})^{k+1} \end{aligned}$$

for all $t \in [0, ||u-u_0||]$. We now may apply Lemma 1.5 to finish the proof. □

Theorem 1.7. *Suppose that $U \subset \mathbb{R}^n$ is an open set and $f : U \to \mathbb{R}$ is smooth. Let P_u denote the degree k Taylor polynomial for f centered at $u \in U$. Then there exists a continuous function $(x, u) \mapsto \epsilon_u(x)$ on $U \times U$ such that $\epsilon_u(u) = 0$ and*

$$f(x) - P_u(x) = \epsilon_u(x)||x-u||^k$$

Proof. We define $\epsilon_u(u) = 0$ and $\epsilon_u(x) = (f(x) - P_u(x))/||x-u||^k$ for $u \neq x$. Since $P_u(x)$ is a polynomial in x with coefficients that are continous functions of u, it is clear that $(u, x) \mapsto \epsilon_u(x)$ is continuous on the complement of the diagonal in $U \times U$. So it suffices to show that $\lim_{(v,x)\to(u,u)} \epsilon_v(x) = 0$.

Let V be a neighborhood of (u, u) with compact closure. Then there exists M such that $sup_{v\in V}|g(v)| < M$ for all partial derivatives g of f with order $k+1$. The function $f(x) - P_v(x)$ has the property that all partial derivatives of order at most k vanish at v. Thus, by Lemma 1.6, for $v \in N$ and $u \neq x \in U$ we have

$$\epsilon_v(x) = (f(x) - P_v(x))/||x-v||^k \leq \frac{Mn^{(k+1)/2}}{(k+1)!}||x-v||.$$

In particular, $\epsilon_v(x) \to 0$ as $(v, x) \to (u, u)$, as required. □

It will also be important to know that the error in the degree k Taylor approximation can be expressed as a homogeneous polynomial of degree k whose coefficients are smooth functions.

Lemma 1.8. *Suppose that f is a real-valued function which is smooth on the closed interval $[a, b] \subset \mathbb{R}$. Suppose that $f^{(i)}(a) = 0$ for $i = 0, \ldots, k-1$. Then*

$$f(x) = \frac{(x-a)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} f^{(k)}((1-t)a + tx) dt$$

for all $x \in [a, b]$. In particular, $f(x) = (x-a)^k g(x)$ for some function g which is smooth on the interval $[a, b]$.

Proof. It suffices to give the proof for $a = 0$. The proof is by induction on n. The base case $k = 1$ is immediate. For the induction step, with $k \geq 2$, we integrate by parts:

$$\begin{aligned}
&\frac{x^k}{(k-1)!} \int_0^1 (1-t)^{k-1} f^{(k)}(tx) dt \\
&\quad = \frac{x^{k-1}}{(k-1)!} (1-t)^{k-1} f^{(k-1)}(tx) \Big|_0^1 + \frac{x^{k-1}}{(k-2)!} \int_0^1 (1-t)^{k-2} f^{(k-1)}(tx) dt \\
&\quad = 0 + f(x),
\end{aligned}$$

where the first term is 0 because $f^{(k-1)}(0) = 0$, and the second term equals $f(x)$ by the induction hypothesis.

The function

$$g(x) = \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} f^{(k)}(tx) dt$$

is seen to be smooth by differentiating under the integral sign. □

Lemma 1.9. *Let U be a convex open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a smooth function. Suppose that $u_0 \in U$ is a point such that all of the partial derivatives of f of order at most k vanish at u_0. then for all $u = u_0 + v \in U$, with $v = (v_1, \ldots, v_n)$, we have*

$$f(u_0 + v) = \sum \left(\frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} D_{i_1} \cdots D_{i_k} f(u_0 + tv) dt \right) v_{i_1} \cdots v_{i_j}$$

where the sum is taken over all sequences $(i_1, \ldots, i_k)$ of integers between 1 and n. In particular, f can be expressed as homogeneous polynomial of degree k whose coefficients are smooth functions on U.

Proof. Suppose $u_0 + tv \in U$ and set $h(t) = f(u_0 + tv)$, for $t \in [0, 1]$. It follows from the chain rule for partial derivatives that

$$h^{(j)}(t) = \sum D_{i_1} \cdots D_{i_j} f(u_0 + tv) v_{i_1} \cdots v_{i_j},$$

where the sum is taken over all sequences $(i_1, \ldots, i_j)$ of integers between 1 and n. In particular we have $h^{(j)}(0) = 0$ for $j = 0, \ldots, k-1$. We now may apply Lemma 1.8 to finish the proof. □

Theorem 1.10. *Suppose that $U \subset \mathbb{R}^n$ is a convex open set and $f : U \to \mathbb{R}$ is smooth. Let P_u denote the degree k Taylor polynomial for f centered at $u \in U$. For each degree k monomial $x_1^{i_1} \cdots x_n^{i_n}$ there exists a smooth function $g_{(i_1,\dots,i_n)}$ such that, for $x = (x_1, \dots, x_n)$, we have*

$$f(x) - P_u(x) = \sum_{i_1+\cdots+i_n=k} g_{(i_1,\dots,i_k)}(x) x_1^{i_1} \cdots x_n^{i_n}.$$

Proof. This follows immediately from Lemma 1.9. □

It is straightforward to extend these notions to vector-valued functions.

Definition 1.11. Suppose that U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is a function. Let $\pi_i : \mathbb{R}^m \to \mathbb{R}$ be the projection on the i^{th} factor. The i^{th} *coordinate function* of f is the real-valued function $f_i \doteq \pi_i \circ f$. We say that f is smooth if f_i is smooth for each $i = 1, \dots, m$. As in the case of real-valued functions, if K is an arbitrary subset of $\mathbb{R}^n$, a function $f : K \to \mathbb{R}^m$ is said to be smooth if every point of K has an open neighborhood U in $\mathbb{R}^n$ such that $f|_{K\cap U}$ extends to a smooth function defined on U.

Suppose that U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is a smooth function. If $u \in U$ and k is a non-negative integer, the *degree k Taylor approximation of f centered at u* is the function P_u whose i^{th} coordinate function is the degree-k Taylor polynomial of f_i.

Theorem 1.7 immediately implies the following.

Theorem 1.12. *Suppose that $U \subset \mathbb{R}^n$ is an open set and $f : U \to \mathbb{R}^m$ is smooth. Let P_u denote the degree k Taylor approximation of f centered at $u \in U$. Then there exists a continuous function $(x, u) \mapsto \epsilon_u(x)$ on $U \times U$ such that $\epsilon_u(u) = 0$ and*

$$f(x) - P_u(x) = \epsilon_u(x)||x - u||^k$$

2. Derivatives

Definition 2.1. Let U be an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ be a function. We say that f is *differentiable* at the point $u \in U$ if there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\text{(D)} \qquad \lim_{h\to 0} \frac{1}{||h||}\Big(f(u+h) - f(u) - L(h)\Big) = 0.$$

Remark 2.2. If L is a linear transformation satisfying condition (D), we can compute the matrix of L with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ by simply taking $x = u + t\mathbf{e}_j$

where $\mathbf{e}_j$ is the j^{th} standard basis vector of $\mathbb{R}^n$. We have

$$\begin{aligned} L(\mathbf{e}_j) &= \lim_{t\to 0}\frac{1}{t}L(t\mathbf{e}_i) \\ &= \lim_{t\to 0}\frac{1}{t}(f(u+t\mathbf{e}_j)-f(u)) \\ &= D_j f(u). \end{aligned}$$

In particular, the j^{th} column of the matrix is the j^{th} partial derivative of f at u. It follows that the ij entry is $D_j f_i(u)$; thus the matrix of L with respect to the standard bases is the usual Jacobian matrix of f.

As a consequence we see that if there exists a linear transformation L satisfying (D), then it is unique. This justifies the following definition.

Definition 2.3. If U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable at $u \in U$, then the unique linear transformation satisfying equation (D) will be denoted Df_u.

Theorem 2.4. *If U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is smooth then f is differentiable at every point of U.*

Proof. This follows immediately from Theorem 1.12 applied to $f(x) - f(u)$ with $k = 1$, since the degree 1 Taylor approximmation of $f(x) - f(u)$ centered at u is a linear transformation. $\square$

Exercise 2.1. Verify the following:

- If $c : \mathbb{R}^n \to \mathbb{R}^n$ is constant then $Dc_u = 0$ for all $u \in \mathbb{R}^n$.
- If $l : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation then $dl_u = l$ for all $u \in \mathbb{R}^n$.
- Let U be an open set in $\mathbb{R}^n$. Suppose $f : U \to \mathbb{R}^m$ and $g : U \to \mathbb{R}^m$ are differentiable at $u \in U$. Then, for any real numbers a and b, $D(af + bg)_u = a\,Df_u + b\,Dg_u$.

The key formal property of Df is the chain rule. In deriving the chain rule, and in estimates that we make later on, it will be convenient to have the following

Definition 2.5. If L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ we define

$$||L|| = sup_{||x||=1}||L(x)||.$$

The supremum exists because it is the maximum value of a continuous function on the compact set $S^n = \{v \in \mathbb{R}^n : \ ||v|| = 1\}$.

Exercise 2.2. Verify the following properties of $||L||$:

- If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation then
$$||L(x)|| \leq ||L||\,||x||$$
for all $x \in \mathbb{R}^n$.
- The pair $(Hom(\mathbb{R}^n, \mathbb{R}^m), ||\cdot||)$ is a normed vector space, and the topology determined by the norm agrees with the topology obtained by using the standard bases to identify $Hom(\mathbb{R}^n, \mathbb{R}^m)$ with the vector space of $m \times n$-matrices. In particular, if L has matrix (a_{ij}) with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, then
$$\max |a_{ij}| \leq ||L|| \leq \sqrt{mn} \max |a_{ij}|.$$

Theorem 2.6 (Chain rule)**.** *Suppose that $U \subset \mathbb{R}^n$ is an open set and that $f : U \to \mathbb{R}^m$ is differentiable at $u \in U$. Suppose that $V \subset \mathbb{R}^m$ is an open set containing $f(u)$ and $g : V \to \mathbb{R}^k$ is differentiable at $f(u)$. Then $D(g \circ f)_u = (Dg_{f(u)}) \circ (Df_u)$.*

Proof. Set
$$\alpha(h) = f(u + h) - f(u) - Df_u(h)$$
and
$$\beta(h) = g(f(u) + h) - g(f(u)) - Dg_{f(u)}(h).$$
We have
$$\begin{aligned} g \circ f(u + h) &= g(f(u) + Df_u(h) + \alpha(h)) \\ &= g \circ f(u) + Dg_{f(u)} \circ Df_u(h) + Dg_{f(u)}(\alpha(h)) + \beta(Df_u(h) + \alpha(h)). \end{aligned}$$
Thus
$$\begin{aligned} &\frac{1}{||h||}\Big(g \circ f(u + h) - g \circ f(u) - Dg_{f(u)} \circ Df_u(h)\Big) \\ &\qquad = \frac{1}{||h||}\Big(Df_{f(u)}(\alpha(h)) + \beta(Df_u(h) + \alpha(h))\Big) \\ &\qquad = \frac{1}{||h||} Dg_{f(u)}(\alpha(h)) + \frac{1}{||h||}\beta(Df_u(h) + \alpha(h)) \\ &\qquad = Dg_{f(u)}(\frac{1}{||h||}\alpha(h)) + \frac{1}{||h||}\beta(Df_u(h) + \alpha(h)) \end{aligned}$$
To complete the proof we must show that the expression above tends to 0 as $h \to 0$. The differentiability of f implies directly that $\lim_{h\to 0} \frac{1}{||h||}\alpha(h) = 0$. Since the linear transformation $Dg_{f(u)}$ is continuous, this implies that $\lim_{h\to 0} Df_{f(u)}(\frac{1}{||h||}\alpha(h)) = 0$.

It is a little trickier to show that the second term tends to 0. For this we interpret the differentiablility of g slightly differently: we may write $\beta(h) = \epsilon(h)||h||$ where $\epsilon(h) \to 0$ as $h \to 0$. Since $\lim_{h\to 0} \frac{1}{||h||}\alpha(h) = 0$ there is a neighborhood V of 0 so that $||\alpha(h)|| \leq ||h||$

for $h \in V$. Thus $||Df_u(h)+\alpha(h)|| \leq (||Df_u||+1)||h||$ for $h \in V$. If we set $K = 1+||Df_u||$ then we have the following estimate for $0 \neq h \in V$:

$$\begin{aligned} ||\beta(Df_u(h)+\alpha(h))|| &= ||\epsilon(Df_u(h)+\alpha(h))||\ ||Df_u(h)+\alpha(h)|| \\ &\leq K||\epsilon(Df_u(h)+\alpha(h))||\ ||h||. \end{aligned}$$

This implies that for $h \in V$ we have

$$\frac{1}{||h||}||\beta(Df_u(h)+\alpha(h))|| \leq K||\epsilon(Df_u(h)+\alpha(h))||,$$

and hence

$$\lim_{h\to 0}\frac{1}{||h||}\beta(Df_u(h)+\alpha(h)) = 0.$$

□

3. The inverse and implicit function theorems

Theorem 3.1 (Inverse Function Theorem)**.** *Suppose that U is an open set in $\mathbb{R}^n$ and $f: U \to \mathbb{R}^n$ is smooth. Let $u \in U$. If Df_u is non-singular then there exist open sets V and W with $u \in V \subset U$ and $f(u) \in W \subset \mathbb{R}^n$ such that f maps V bijectively onto W. Moreover, the inverse of $f|_V$ is a smooth map from W to V.*

Proof. The main step of the proof is to use an iterative method to solve the equation $f(x) = y$, and to show that for each y sufficiently close to $f(u)$ there is exactly one solution near u. The iteration is a modification of Newton's method, designed to make it easier to analyze the convergence properties. Newton's method would use the following iteration: $x_0 = u$; $x_{n+1} = x_n + (Df_{x_n})^{-1}(y - f(x_n))$. Our simplified variant of Newton's iteration is: $x_0 = u$; $x_{n+1} = x_n + (Df_u)^{-1}(y - f(x_n))$. To analyze this, we define a function $G_y : U \to \mathbb{R}^n$ by

$$G_y(x) = x + (Df_u)^{-1}(y - f(x)).$$

Observe, for $x \in U$, that $G_y(x) = x \Leftrightarrow (Df_u)^{-1}(y - f(x)) = 0 \Leftrightarrow f(x) = y$. Thus, solving the equation $f(x) = y$ in U amounts to finding a fixed point of G_y.

We let $B_u(\delta)$ denote the ball $\{x \in \mathbb{R}^n : ||x-u|| < \delta\}$. We will make use of the continuous function $(v, w) \mapsto \epsilon_v(w)$ given by Theorem 1.12 in the case of a degree 1 Taylor approximation of f. Arguing backwards, as analysts are fond of doing, we start by choosing $\delta > 0$ with the following properties:

(1) $B_u(\delta) \subset U$;
(2) Df_v is invertible for all $v \in B_u(\delta)$;
(3) $||I - (Df_u)^{-1} \circ Df_v|| < \frac{1}{4}$ for all $v \in B_u(\delta)$;
(4) $||\epsilon_v(w)|| \leq \frac{1}{4||(Df_u)^{-1}||}$ for all $v, w \in B_u(\delta)$.

To see that such a number δ exists, we need four observations: (1) U is an open set containing u; (2) $\det Df_v$ is a continuous function of v that takes a positive value at $v = u$; (3) $||I - (Df_u)^{-1} \circ Df_v||$ is a continuous function of v that takes the value 0 at u; and (4) $||\epsilon_v(w)|| \to 0$ as $(v, w) \to (u, u)$.

Next we claim that if $v, w \in B_u(\delta)$ then $||G_y(w) - G_y(v)|| < \frac{1}{2}||w - v||$. We compute

$$
\begin{aligned}
||G_y(w) - G_y(v)|| &= ||w - v - (Df_u)^{-1}(f(w) - f(v))|| \quad \text{(the } y\text{'s cancel)} \\
&= ||w - v - (Df_u)^{-1}(Df_v(w - v) + \epsilon_v(w)||w - v||)|| \\
&\leq ||(I - (Df_u)^{-1} \circ Df_v)(w - v)|| + ||(Df_u)^{-1}(\epsilon_v(w))||\,||w - v|| \\
&\leq ||I - (Df_u)^{-1} \circ Df_v||\,||w - v|| + ||(Df_u)^{-1}||\,||\epsilon_v(w)||\,||w - v|| \\
&\leq \frac{1}{4}||w - v|| + ||(Df_u)^{-1}|| \left(\frac{||w - v||}{4||(Df_u)^{-1}||} \right) \\
&\leq \frac{1}{2}||w - v||.
\end{aligned}
$$

The estimate above has the peculiar feature that it is independent of y. However, we will show that it does imply that if all of the iterates $x_0 = u, x_1 = G_y(u), x_2 = G_y^2(u), \ldots$ are defined and contained in $B_u(\delta)$, then they converge to a fixed point of G_y. In fact, for $m < n$, we have

$$
\begin{aligned}
||x_n - x_m|| &\leq ||x_{m+1} - x_m|| + \cdots + ||x_n - x_{n-1}|| \\
&< \frac{1}{2^m}||x_1 - x_0|| + \cdots \frac{1}{2^n}||x_1 - x_0|| \\
&\leq \frac{1}{2^{m-1}}||x_1 - x_0||,
\end{aligned}
$$

so the sequence (x_n) is Cauchy, and has a limit x_∞. Moreover, since G_y is continuous,

$$
G_y(x_\infty) = G_y(\lim_{n \to \infty} G_y^n(u)) = \lim_{n \to \infty} G_y^{n+1}(u) = x_\infty.
$$

The estimate also implies that G_y has at most one fixed point in $B_u(\delta)$, since if $G_y(v) = v$ and $G_y(w) = w$ then $||w - v|| = ||G_y(w) - G_y(v)|| \leq \frac{1}{2}||w - v||$, which implies $v = w$.

In order to ensure that all of the iterates lie in $B_u(\delta)$ we must take y close to $f(u)$. Since the estimate above shows $||x_n - x_0|| \leq 2||x_1 - x_0||$ for all $n \geq 1$ it suffices to show that if y is close enough to $f(u)$ then $||x_1 - x_0|| = ||G_y(u) - u|| < \delta/2$. Set $\delta' = \frac{\delta}{3||(Df_u)^{-1}||}$. Then for $y \in B_{f(u)}(\delta')$ we have

$$
||G_y(u) - u|| = ||u + (Df_u)^{-1}(y - f(u)) - u|| \leq ||(Df_u)^{-1}||\,||y - f(u))|| \leq \delta/3 < \delta/2.
$$

At this point we have established that for each $y \in B_{f(u)}(\delta')$ there exists a unique $x \in B_u(\delta)$ such that $f(x) = y$. If we set $W = B_{f(u)}(\delta')$ and $V = B_u(\delta) \cap f^{-1}(W)$ then V and W are open sets with $u \in V \subset U$ and $f(u) \in W \subset \mathbb{R}^n$ such that f maps V bijectively onto W.

It remains to show that the inverse function g of $f|_V$ is a smooth map from W to V. The first step is to show that g is differentiable at each point of W, and hence that the partial derivatives of g are defined on W. Of course, if $y = f(x)$ and if g is differentiable at y then the chain rule implies that Dg_y must equal $(Df_x)^{-1}$. (NOTE: this is the step where we use property (3) of the number δ!) However, we still must check that the linear transformation $(Df_x)^{-1}$ satisfies condition (D) in Definition 2.1. For any two points $f(v)$ and $f(w)$ in W we have

$$
\begin{aligned}
g(f(w)) - g(f(v)) - (Df_v)^{-1}(f(w) - f(v)) \\
&= w - v - (Df_v)^{-1}(Df_v(w - v) - \epsilon_v(w)||w - v||) \\
&= ||w - v||(Df_v)^{-1}(\epsilon_v(w)).
\end{aligned}
$$

Since $(Df_v)^{-1}(\epsilon_v(w)) \to 0$ as $w \to v$, this verifies condition (D) and shows that $Dg_{f(v)} = (Df_v)^{-1}$ at an arbitrary point $f(v) \in W$.

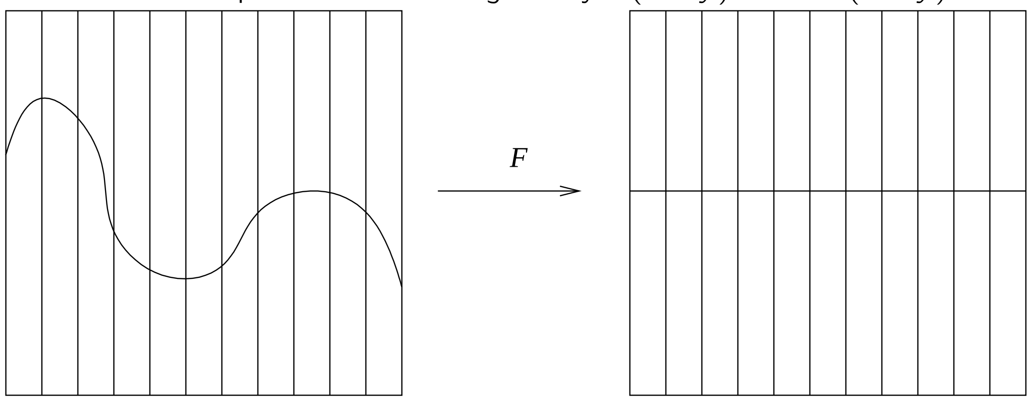
We prove that g is smooth by induction, using a magical boot-strapping argument. The idea is to prove that if g has continuous partials up to order k, then the first partials of g also have continuous partials up to order k. This implies that g in fact has continuous partials up to order $k+1$ and hence, by induction, of all orders. (The base case of the induction is the statement that g is differentiable.)

For the bootstrapping statement, consider the matrix of Df_v, with respect to the standard basis, as a matrix $A = A(v)$ whose entries are functions of v. These entries, being partial derivatives of the coordinate functions of the smooth function f, are themselves smooth. Since the determinant of $A(v)$ is non-zero for all $v \in V$, we may use the formula $A^{-1} = \frac{1}{\det a} \operatorname{adj} A$ to express the coefficients of $A(v)^{-1}$ as rational functions in the entries of $A(v)$, whose denominators are non-zero for all $v \in V$. In particular, the entries of the matrix $A(v)^{-1}$ are smooth functions of v. Next consider the entries of $A(g(w))^{-1}$, viewed as functions of $w \in W$. Each of these entries has the form $h \circ g$ where h is some entry of $A(v)^{-1}$ and is therefore smooth. Since g has continuous partials up to order k, so $h \circ g$. Thus the entries of $A(g(w))^{-1}$ Now $A(g(w))^{-1}$ is just the matrix of Dg_w, viewed as a matrix of functions on W, and hence its entries are just the first order partial derivatives of the coordinate functions of g. This proves that if g has continuous partials up to order k, then so do each of the first order partials of g. Therefore g is smooth. ☐

Exercise 3.1. Show that if the hypothesis of the Inverse Function Theorem is weakened to assume only that f has continuous partials up to order k, then the boot-strapping argument will prove that g also has continuous partials up to order k.

Theorem 3.2 (Implicit Function Theorem)**.** *Suppose U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^k$ is a smooth function for $k < n$. Write $\mathbb{R}^n = X \oplus Y$ where X is the $n - k$-dimensional subspace of spanned by the first $n - k$ standard basis elements and Y is the subspace* *spanned by the last k. Suppose that $u \in U$ is a point such that $f(u) = 0$ and the linear transformation $D_u f$ restricts to an injection on Y. Then there exists a neighborhood $V = V_x \oplus V_y$ of u, and a smooth function $g : V_x \to V_y$ so that $V \cap f^{-1}(0)$ is the graph of g.*

Proof. Consider the map $F : U \to X \oplus Y$ given by $F(x \oplus y) = x \oplus f(x \oplus y)$.



The matrix of DF_u with respect to the standard basis has the form

$$\begin{bmatrix} I_{n-k} & * \\ 0 & A \end{bmatrix}$$

where A is the matrix of $D_u f|_Y$. In particular DF_u is nonsingular. Let V and W be open sets in X and Y such that $u \in V \oplus W$ and such that F restricts to a bijection with smooth inverse on $V \oplus W$. Let G denote the inverse function. We have

$$V \oplus W \cap f^{-1}(0) = V \oplus W \cap G(V \oplus \{0\}).$$

Let π denote the natural projection from $X \oplus Y$ to Y and set $g(v) = \pi(G(v \oplus 0))$ for v in V. We then have $G(v \oplus 0) = v \oplus g(v)$, and hence $V \oplus W \cap f^{-1}(0)$ is the graph of g. Since g is a composition of smooth functions, g is smooth. □

4. Smooth manifolds

A smooth manifold is a topological manifold with extra structure which allows for a notion of a smooth function.

Definition 4.1. By a (topological) *n-manifold M* we shall mean a topological space such that

- every point of M has an open neighborhood which is homeomorphic to $\mathbb{R}^n$;
- M is Hausdorff; and
- there is a countable basis for the topology of M.

Remark 4.2. The definition ignores the question of whether an n-manifold could also be an m-manifold for $m \neq n$. The answer to this question (no) is contained in Brouwer's theorem on Invariance of Domain, which is proved using algebraic topology.

Here are two examples of spaces which satisfy the first condition in the definition, but not all three of the conditions. These will not be considered to be manifolds in the context of these notes.

Example 4.3. Let $M = \mathbb{R} \cap \{o\}$, where $o \notin \mathbb{R}$, and take as a basis for the topology the collection of all sets U such that either $U \subset \mathbb{R}$ is open in the standard topology on $\mathbb{R}$; or $U = \{o\} \cup V - \{0\}$ where $V \subset \mathbb{R}$ is an open neighborhood of 0 in the standard topology on $\mathbb{R}$. This is a non-Hausdorff space in which every point has an open neighborhood homeomorphic to $\mathbb{R}$.

Example 4.4. Let Ω be an uncountable well-ordered set in which every point has countably many predecessors, and let o be the first element of Ω. Let $M = \Omega \times [0, 1) - \{(o, 0)\}$ and give M the order topology determined by the dictionary order. In this space, known as the "long ray", every point has an open neighborhood homeomorphic to $\mathbb{R}$, but there is an uncountable family of pairwise disjoint non-empty open subsets.

Definition 4.5. Let M be a topological n-manifold. A *chart* for M is a pair (U, ϕ) where U is an open set in M and $\phi : U \to \mathbb{R}^n$ is a map such that $\phi(U)$ is open and ϕ is a homeomorphism when regarded as a map from U to $\phi(U)$. If (U, α) is a chart then U is called its *domain*.

Two charts (U_1, ϕ_1) and (U_2, ϕ_2) on an n-manifold M will be said to be *compatible* if the two maps

$$\phi_1 \circ \phi_2^{-1}|_{\phi_2(U_1 \cap U_2)} : \phi_2(U_1 \cap U_2) \to \mathbb{R}^n$$

and

$$\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2)} : \phi_1(U_1 \cap U_2) \to \mathbb{R}^n$$

are both smooth. (In particular, if $U_1 \cap U_2 = \emptyset$ then (U_1, ϕ_1) and (U_2, ϕ_2) are compatible.)

Definition 4.6. A collection $\mathcal{A}$ of charts on an n-manifold M is called an *atlas* for M if the domains of the charts in $\mathcal{A}$ cover M.

Exercise 4.1. Suppose that $\mathcal{A}$ is an atlas for M. Let $\hat{\mathcal{A}}$ denote the collection of all charts on M that are compatible with every chart in $\mathcal{A}$. Show that $\hat{\mathcal{A}}$ is an atlas for M. Conclude that any atlas for M is contained in a unique maximal atlas.

Definition 4.7. A *differentiable structure* on an n-manifold M is a non-empty maximal atlas. According to Exercise 4.1, an atlas on M uniquely determines a differentiable structure on M. A *smooth n-manifold* is a pair $(M, \mathcal{A})$, where M is a topological n-manifold and $\mathcal{A}$ is a differentiable structure on M. Usually, the differentiable structure will be implicit and the pair $(M, \mathcal{A})$ will be denoted M. By the *standard* differentiable structure on $\mathbb{R}^n$ will be the one determined by the atlas $\{(\mathbb{R}^n, \mathrm{id})\}$.

Definition 4.8. Suppose that $(M, \mathcal{A})$ is a smooth m-manifold and $(N, \mathcal{B})$ is a smooth n-manifold. A map $f : M \to N$ will be said to be *smooth at a point* $p \in M$ if it satisfies the following condition

- if $(U, \phi) \in \mathcal{A}$ is a chart with $p \in U$ and $(V, \psi) \in \mathcal{B}$ is a chart with $f(p) \in V$ then the composition $\psi \circ f \circ \phi^{-1} : W \to \mathbb{R}^n$ is a smooth map, where W denotes the (open) set on which the composition is defined.

The map will be called *smooth* if it is smooth at every point of M. If f is a bijection, and if f^{-1} is also smooth then f will be called a *diffeomorphism*.

Exercise 4.2. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$\phi(x) = \begin{cases} x & \text{for } x \leq 0; \\ 2x & \text{for } x > 0 \end{cases}.$$

Let $\mathcal{A}$ denote the atlas $(\mathbb{R}, \phi)$ on the 1-manifold $\mathbb{R}$. Characterize the smooth maps from $(\mathbb{R}, \mathcal{A})$ to $(\mathbb{R}, \mathcal{S})$, where $\mathcal{S}$ denotes the standard differentiable structure. Prove that these two smooth manifolds are diffeomeorphic.

5. Constructing smooth manifolds

Proposition 5.1. *An open subset V of a smooth n-manifold $(M, \mathcal{A})$ has a unique differentiable structure such that the inclusion map $V \to M$ is smooth.*

Proof. It is immediate that $B = \{(U \cap V, \phi|_V) : (U, \phi) \in \mathcal{A}\}$ is an atlas on V. Let $\mathcal{B}$ denote the differentiable structure containing B. It is clear from the definitions that the inclusion map is a smooth map from $(V, \mathcal{B})$ to $(M, \mathcal{A})$, and that the atlas B must be contained in any differentiable structure for which the inclusion map is smooth.. $\square$

Example 5.2. The algebra $M_n(R)$ of $n \times n$-matrices can be identified, as a vector space, with $\mathbb{R}^{n^2}$, and therefore naturally acquires a differentiable structure. The determinant defines a smooth (in fact, polynomial) map $\det : M_n(\mathbb{R}) \to \mathbb{R}$. The general linear group $GL(n, \mathbb{R})$, consisting of invertible $n \times n$ matrices, is an open set in $M_n(R)$ since it is the preimage of $\mathbb{R} - \{0\}$ under the map det. Thus $GL(n, \mathbb{R})$ has a natural differentiable structure for which the inclusion $GL(n, \mathbb{R}) \to M_n(\mathbb{R})$ is smooth. We will take this to be the standard differentiable structure on $GL(n, \mathbb{R})$.

Next we describe some important examples for which an atlas can be constructed explicitly.

Example 5.3 (The n-sphere). For $n \geq 0$ we define the n-sphere to be

$$S^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$

We give S^n the subspace topology it inherits as a subset of $\mathbb{R}^{n+1}$. (Thus S^n is Hausdorff and has a countable basis, being a subspace of $\mathbb{R}^{n+1}$. Moreover, being closed and bounded, S^n is compact.)

Let $N = (0, \ldots, 0, 1) \in S^n$ and $S = (0, \ldots, 0, -1) \in S^n$. Let $\phi : S^n - \{N\} \to \mathbb{R}^n$ denote the stereographic projection:

$$\phi(x_1, \ldots, x_{n+1}) = \frac{1}{1 - x_{n+1}}(x_1, \ldots, x_n),$$

and set $\phi'(x_1, \ldots, x_{n+1}) = (\phi(x_1, \ldots, -x_{n+1})$. Then S^n has an atlas consisting of the two charts: $(S^n - \{N\}, \phi)$ and $(S^n - \{S\}, \phi')$.

Using the formula

$$\phi^{-1}(x_1, \ldots, x_n) = \frac{1}{1 + x_1^2 + \cdots + x_n^2}(2x_1, \ldots, 2x_n, x_1^2 + \cdots x_n^2 - 1)$$

it is easily checked that the two charts are compatible.

Example 5.4 (Projective space). The n-dimensional (real) projective space $\mathbb{RP}^n$ is a topological space whose "points" are the lines in $\mathbb{R}^{n+1}$. We may construct it as a quotient space $\mathbb{RP}^n = \mathbb{R}^{n+1} - \{0\}/\sim$ where $\sim$ is the equivalence relation $x \sim y \Leftrightarrow x = \lambda y$ for some $\lambda \in \mathbb{R}$. Thus each equivalence class is the intersection of a line in $\mathbb{R}^{n+1}$ with $\mathbb{R}^{n+1} - \{0\}$. The equivalence classes can also be viewed as the orbits of the action of the multiplicative group $\mathbb{R}^\times$ on $\mathbb{R}^{n+1} - \{0\}$.

Here one must confront the issue that the Hausdorff property does not behave well with respect to the quotient space construction. It is easy to construct non-Hausdorff quotients of a nice Hausdorff space, such as an open subset of $\mathbb{R}^{n+1}$. For this reason, and also to see that $\mathbb{RP}^n$ is compact, it is useful to realize $\mathbb{RP}^n$ as a quotient of a compact space, namely S^n. If we restrict the equivalence relation $\sim$ to S^n we see that the resulting equivalence classes each consist of a pair of antipodal points. (Thus the restricted equivalence relation agrees with the orbit relation for the action of the group of order two generated by the antipodal map $x \mapsto -x$.) Moreover, each equivalence class in $\mathbb{R}^{n+1}$ meets S^n, and each saturated open set in S^n is the intersection of S^n with a unique saturated open set in $\mathbb{R}^{n+1}$. In fact, if $r : \mathbb{R}^n - \{0\} \to S^n$ is the map defined by $r(v) = \frac{v}{||v||}$, and if W is a saturated open set in S^n, then $r^{-1}(W)$ is a saturated open set in $\mathbb{R}^n - \{0\}$ such that $W = r^{-1}(W) \cap S^n$.

It follows that the inclusion $S^n \to \mathbb{R}^{n+1}$ induces a homeomorphism from $S^n/\sim$ to $\mathbb{R}^{n+1}/\sim$. Given two distinct equivalence classes $\{\pm x\}$ and $\{\pm y\}$, it is easy to construct disjoint saturated open sets U and V with $\{\pm x\} \subset U$ and $\{\pm y\} \subset V$. So we see from this that $\mathbb{RP}^n$ is, in fact, a compact Hausdorff space.

Points of projective space are usually represented by their *homogeneous coordinates*: the symbol $(x_1 : \ldots : x_{n+1})$ (where not all x_i can be 0) denotes the equivalence class of the point $(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1}$. In particular, $(x_1 : \ldots : x_{n+1}) = (\lambda x_1 : \ldots : \lambda x_{n+1})$ for any non-zero number λ.

For $i = 1, \ldots, n+1$, consider the saturated open sets

$$\widehat{U}_i = \{(x_1, \ldots, x_{n+1}) : x_i \neq 0\} \subset \mathbb{R}^{n+1}.$$

Let $U_i \subset \mathbb{RP}^n$ denote the open set which is the image of $\widehat{U}_i$ under the quotient mapping. We will define an atlas for $\mathbb{RP}^n$ having these open sets as the domains of the charts.

To define a map $\phi_i : U_i \to \mathbb{R}^n$ it suffices to define a map $\widehat{\phi}_i : \widehat{U}_i \to \mathbb{R}^n$ such that $\widehat{\phi}_i$ is constant on equivalence classes. We do this as follows

$$\widehat{\phi}_i(x_1, \ldots, x_n) = (\frac{x_1}{x_i}, \ldots \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{n+1}}{x_i}).$$

That is, we divide all coordinates by x_i, which is non-zero for all points of $\widehat{U}_i$, and then delete the 1 in the i^{th} position. It is clear that $\phi : U_i \to \mathbb{R}^n$ is continuous, bijective and open. We illustrate the proof that the charts are compatible in the case of (U_1, ϕ_1) and (U_2, ϕ_2). Here one checks that $\phi_1(U_1 \cap U_2) = \{(x_1, \ldots, x_n) : x_1 \neq 0\}$ and that $\phi_2 \circ \phi_1^{-1}$ is given by

$$(x_1, \ldots, x_n) \mapsto (1 : x_1 : \ldots : x_n) \mapsto (\frac{1}{x_1}, \frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1})$$

which is smooth where $x_1 \neq 0$.

Example 5.5 (Grassmann manifolds)**.** Next we generalize the construction of $\mathbb{RP}^n$ by constructing a compact manifold $G_{k,n}$ whose "points" are k-dimensional subspaces of $\mathbb{R}^n$. Let $M_{n,k}$ denote the vector space of $n \times k$ matrices and let $X \subset M_{n,k}$ be the subset consisting of matrices with rank k. We will need some notation for dealing with submatrices. Suppose $1 \leq i_1 < i_2 < \cdots < i_m \leq n$, and $I = (i_1, \ldots, i_n)$. If $A \in M_{n,k}$ then we will write A_I to denote the submatrix of A consisting of the rows indexed by elements of I. Now, a matrix in $M_{n,k}$ has rank k if and only if $\det A_I \neq 0$ for some k-tuple $I = (i_1, \ldots, i_k)$. It follows that the set X is an open set in $M_{n,k}$.

We will construct $G_{k,n}$ as a quotient space of X. For $A \in M_{n,k}$ let $\operatorname{col} A \subset \mathbb{R}^n$ denote the column space of A. For $A \in X$, the subspace $\operatorname{col} A$ has dimension k, and every k-dimensional subspace arises as $\operatorname{col} A$ for some $A \in X$. Define an equivalence relation on X by $A \sim B \Leftrightarrow \operatorname{col} A = \operatorname{col} B$. We define $G_{k,n}$ to be the quotient space $X/\sim$. Observe that $A \sim B \Leftrightarrow B = AG$ for some nonsingular $k \times k$-matrix G.

Once again, we must deal with the issue of showing that $G_{k,n}$ is Hausdorff. For this we consider the subset $Y \subset X$ consisting of matrices with orthonormal columns. In other words, $A \in Y$ if and only if $A^T A = I_k$. Note that Y is closed and bounded, and hence compact. Suppose that A and B are elements of Y, with $A \sim B$. Let G be a $k \times k$ non-singular matrix such that $B = AG$. We then have

$$I_k = B^T B = G^T A^T AG = G^T G,$$

so the matrix G must be an orthogonal matrix. This shows that the restriction of $\sim$ to Y is characterized as $A \sim B \Leftrightarrow B = AG$, where G is a $k \times k$ orthogonal matrix. Since the set O_k of $k \times k$ orthogonal matrices is compact, this implies that the equivalence classes for the restriction of $\sim$ to Y are compact sets.

There is a natural map $r : X \to Y$ such that $r(A) = A$ for all $A \in Y$, defined as follows. Regard the columns of A as an ordered linearly independent subset of $\mathbb{R}^n$, and apply the Gramm-Schmidt orthogonalization process to produce an ordered orthonormal set of vectors which span the same subspace. Take this orthonormal set to be the columns of $r(A) \in Y$. The map r is continuous, and if W is a saturated open set in Y then $r^{-1}(W)$ is a saturated open set in X such that $Y \cap r^{-1}(W) = W$. This shows that $G_{k,n} = Y/\sim$.

To show that Y/sim is Hausdorff, if suffices to show that if A and B are inequivalent elements of Y then there exist open neighborhoods U of A and V of B such that the saturated open sets $U \cdot O_k$ and $V \cdot O_k$. For this we recall from Exercise 2.2 that the topology on Y agrees with the metric topology defined by the metric $d(A, B) = ||A - B||$. Moreover, since $\{Av : ||v|| = 1\} = \{AGv : ||v = 1\}$ for any $G \in O_k$, right multiplication by an element of O_k defines an isometry of Y. Since the sets $A \cdot O_k$ and $B \cdot O_k$ are disjoint compact sets, there exists $\epsilon > 0$ such $d(AG, BG') > 2\epsilon$ for all $G, G' \in O_k$. Thus we may take $U = U_0 \cdot O_k$ and $V = V_0 \cdot O_k$, where U_0 and V_0 are the balls of radius $epsilon$ about A and B respectively, defined in terms of the metric d. Note that, since Y is compact, we also have that $G_{k,n}$ is compact.

Now we are ready to construct an atlas for $G_{k,n}$. For each k-tuple $I = (i_1, \ldots, i_k)$ with $1 \le i_1 < \cdots < i_k \le n$ let $\widehat{U}_I = \{A \in X : \det A_i \neq 0\} \subset X$. The sets U_I are open in X, and are saturated since $(AG)_I = (A_I)G$ for any $k \times k$ matrix G. We define $U_I \subset G_{n,k}$ to be the image of $\widehat{U}_I$ under the quotient map. These are open sets which cover $G_{k,n}$, and will be the domains of our charts.

Consider the map $\widehat{\phi}_I : \widehat{U}_I \to M_{n-k,k}$ defined as follows:

$$\widehat{\phi}_I(A) = (AA_I^{-1})_{I^c}$$

where I^c is the $(n-k)$-tuple obtained by removing the indices in I from $(1, 2, \ldots, n)$. In other words, $\widehat{\phi}_I(A)$ is obtained by multiplying A by the inverse of the submatrix A_I and then removing the $k \times k$ identity matrix that appears as the I-submatrix of the result. Since $\widehat{\phi}_I$ is constant on equivalence classes, it descends to a map $\phi_I : U_I \to M_{n-k,k}$. It is easy to see that ϕ_I is a bijection, and that the coordinates functions of ϕ_I and ϕ_I^{-1} are rational functions. Since rational functions are smooth on their domain, this implies that the collection of charts (U_I, ϕ_I) form an atlas for $G_{k,n}$, giving it the structure of a smooth $(nk - k^2)$-manifold.

Next we describe three general constructions of smooth manifolds.