

1 Inverse Function Theorem, Tangent Spaces (May 21)

1.1 Diffeomorphisms and Coordinate Systems

By convention, any manifold labelled N will have dimension n and any labelled M will have dimension m .

Definition 1. A diffeomorphism $F : N \rightarrow M$ of smooth manifolds is a bijective smooth map with smooth inverse.

Proposition 1.1. Coordinate systems are diffeomorphisms.

Proof. Let M be a smooth manifold and (U, ϕ) a coordinate chart on M . Choose the atlases

$$\begin{aligned} \{(U, \phi)\} &\text{ on } U \\ \{\phi(U), \text{Id}\} &\text{ on } \phi(U). \end{aligned}$$

Then

$$\begin{aligned} \text{Id} \circ \phi \circ \phi^{-1} &: \phi(U) \rightarrow \phi(U) \\ \phi \circ \phi^{-1} \circ \text{Id}^{-1} &: \phi(U) \rightarrow \phi(U) \end{aligned}$$

are both smooth, implying that ϕ and ϕ^{-1} are smooth, respectively. \square

The converse is true; it uses maximality of the smooth structure.

Proposition 1.2. Diffeomorphisms from open subsets of manifolds to open subsets of Euclidean space are coordinate systems belonging to the manifold's smooth structure.

Proof. Was left as an exercise in class, so here's a solution. Let $F : U \rightarrow F(U)$ be a diffeomorphism of the open subset U of the smooth manifold M with an open subset $F(U) \subseteq \mathbb{R}^m$. Then (U, F) is a coordinate chart on M . Choose a coordinate chart (V, ψ) for M . If $U \cap V = \emptyset$ we are done, and otherwise, the transition maps are $F \circ \phi^{-1}$ and $\phi \circ F^{-1}$, both of which are clearly smooth. So the transition map is a diffeomorphism, and so (U, F) is a coordinate chart belonging to the smooth structure by maximality. \square

1.2 Coordinate Derivatives, Inverse Function Theorem

In calculus, we take derivatives. How can we take derivatives of functions on manifolds? The first thing we can try is differentiating with respect to local coordinates. If (U, ϕ) is a coordinate system on a manifold, we will write $(U, \phi) = (U, x^1, \dots, x^m)$ to mean that x^i is the i th component of ϕ . More specifically, if r^1, \dots, r^m are the coordinates on \mathbb{R}^m , then $x^i = r^i \circ \phi$.

Definition 2. Let $f : M \rightarrow \mathbb{R}$ be a smooth function on the smooth M . Let $p \in M$ and let $(U, \phi) = (U, x^1, \dots, x^m)$ be a coordinate chart around p . Define

$$\frac{\partial f}{\partial x^i} \Big|_p := \frac{\partial(f \circ \phi^{-1})}{\partial r^i} \Big|_{\phi(p)}$$

as the i th partial derivative of f at p with respect to the coordinates (U, x^1, \dots, x^m) .

What about maps between manifolds? We can do something similar. Let $F : N \rightarrow M$ be a smooth map between smooth manifolds. Let $(U, \phi) = (U, x^1, \dots, x^n)$ and $(V, \psi) = (V, y^1, \dots, y^m)$ be coordinate charts on N and M , respectively. Define the i th component of F with respect to the coordinates (V, y^1, \dots, y^m) by $F^i := y^i \circ F = r^i \circ \psi \circ F$. Then $F^i : N \rightarrow \mathbb{R}$, so by our previous definition we can look at

$$\frac{\partial F^i}{\partial x^j} \Big|_p = \frac{\partial(F^i \circ \phi^{-1})}{\partial r^j} \Big|_{\phi(p)} = \frac{\partial(r^i \circ \psi \circ F \circ \phi^{-1})}{\partial r^j} \Big|_{\phi(p)} = \frac{\partial(\psi \circ F \circ \phi^{-1})^i}{\partial r^j} \Big|_{\phi(p)}.$$

We will call the $m \times n$ matrix $\left[\frac{\partial F^i}{\partial x^j} \right]_p$ the *Jacobian of F at p (relative to the coordinates (U, x^1, \dots, x^n) and (V, y^1, \dots, y^m))*.

The Jacobian itself is not independent of the coordinate systems, but since transition maps are diffeomorphisms, its rank is independent of the coordinate systems chosen. Precisely, if $(\tilde{U}, \tilde{\phi})$ and $(\tilde{V}, \tilde{\psi})$ are alternate coordinate charts around p and $F(p)$, respectively, then we have

$$\tilde{\psi} \circ F \circ \tilde{\phi}^{-1} = (\tilde{\psi} \circ \psi^{-1}) \circ (\psi \circ F \circ \phi^{-1}) \circ (\phi \circ \tilde{\phi}^{-1}),$$

implying

$$D(\tilde{\psi} \circ F \circ \tilde{\phi}^{-1})(\tilde{\phi}(p)) = \underbrace{D(\tilde{\psi} \circ \psi^{-1})(\psi(F(p)))}_{\in GL(m, \mathbb{R})} \cdot D(\psi \circ F \circ \phi^{-1})(\phi(p)) \cdot \underbrace{D(\phi \circ \tilde{\phi}^{-1})(\tilde{\phi}(p))}_{\in GL(n, \mathbb{R})},$$

and so linear algebra tells us that

$$\text{rank}(D(\psi \circ F \circ \phi^{-1})(\phi(p))) = \text{rank}(D(\tilde{\psi} \circ F \circ \tilde{\phi}^{-1})(\tilde{\phi}(p))).$$

Therefore "the rank of the Jacobian of F at p " is a well-defined quantity, independent of the coordinate charts. We state this as a proposition.

Proposition 1.3. If $F : N \rightarrow M$ is C^∞ at p , then the rank of the Jacobian of F at p is the same no matter what coordinate charts around p and $F(p)$ are used to calculate it.

In particular, if $m = n$, then we are led to a generalization of the inverse function theorem, as we can then speak of invertibility of the Jacobian.

Theorem 1.1. (*Inverse function theorem for manifolds*) Let $F : N \rightarrow M$ be a smooth map of smooth manifolds of the same dimension. If the Jacobian of F at $p \in N$ is invertible, then there is an open neighbourhood U of p in N and an open neighbourhood V of $F(p)$ in M such that $F|_U : U \rightarrow V$ is a diffeomorphism.

Proof. Was left as an exercise in class, so here's a solution. Choose coordinate charts $(U, \phi) = (U, x^1, \dots, x^n)$ at p and $(V, \psi) = (V, y^1, \dots, y^n)$ at $F(p)$. Then $\left[\frac{\partial F^i}{\partial x^j} \right]_p$ is invertible, but as we saw above, this is equivalent to saying $\left[\frac{\partial(\psi \circ F \circ \phi^{-1})^i}{\partial r^j} \right]_{\phi(p)}$ is invertible. By the inverse function theorem in \mathbb{R}^n , the map $\psi \circ F \circ \phi^{-1}$ is a diffeomorphism on a small neighbourhood of $\phi(p)$ in $\phi(U \cap F^{-1}(V))$. Since coordinate systems are diffeomorphisms, F is a diffeomorphism on a small neighbourhood of p . \square

(The following was not part of the lecture.) Note that the converse of the above theorem is true; if F restricts to a diffeomorphism in a neighbourhood of p , then the Jacobian with respect to any choices of coordinates is invertible. This can be seen by taking two coordinate charts (U, ϕ) around p and (V, ψ) around $F(p)$ and noting that since $\psi \circ F \circ \phi^{-1}$ is then a diffeomorphism of open sets of \mathbb{R}^n , the Jacobian of F with respect to these coordinate systems is invertible (and hence with respect to any coordinate systems). Therefore we have the following slightly stronger theorem:

Theorem 1.2. (*Stronger inverse function theorem for manifolds*) Let $F : N \rightarrow M$ be a smooth map of smooth manifolds of the same dimension. Then F is a local diffeomorphism at $p \in N$ if and only if the Jacobian of F at p is invertible.

Of course, *local diffeomorphism* at p means that F restricts to a diffeomorphism on an open neighbourhood of p .

We would like a "coordinate-free" derivative. In MAT257, the derivative of a map $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ at p was thought of as the map $F_* : T_p \mathbb{R}^n \rightarrow T_{F(p)} \mathbb{R}^m$ defined by $F_*(v_p) = (DF(p)v)_{F(p)}$, where the subscript indicates the tangent space in which the vector lies. The difficulty in generalizing this to manifolds lies in defining the tangent space of an abstract manifold.

1.3 Abstracting the Tangent Space

If M is a submanifold in \mathbb{R}^n in the MAT257 sense, then we can define the tangent space as follows. Suppose $p \in M$. Then there is an open neighbourhood V of p in \mathbb{R}^n , an open set $U \subseteq \mathbb{R}^k$, and a C^∞ homeomorphism $\phi : U \rightarrow V \cap M$ with $\text{rank}(D\phi(q)) = k$ for each $q \in U$. If $q = \phi^{-1}(p)$, then let $T_p U$ be the "set of all vectors in \mathbb{R}^k thought of as pointing from q ". Then we define $T_p M := D\phi(q)(T_p U)$. Since the derivative has rank k at q , the space $T_p M$ will be a vector subspace of $T_p \mathbb{R}^n$ of dimension k . It is not hard to see that this is independent of the "parametrization" chosen near p .

The problem with this is that it doesn't generalize to abstract manifolds. We'd like to modify the definition so that it abstracts.

We will first attempt to do so by using curves. Let $p \in \mathbb{R}^n$ and $v \in T_p \mathbb{R}^n$. If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is smooth, then we can speak of its Jacobian at a point $p \in \mathbb{R}^n$. If $\gamma : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ is a smooth curve with $\gamma(0) = p$ and $\gamma'(0) = v$, then

$$\frac{d}{dt} \Big|_{t=0} F \circ \gamma = DF(p) \cdot v,$$

so we can think of $DF(p)$ as a map sending tangent vectors to tangent vectors. (A picture would really help here.)

Define $A = \{\text{smooth curves with } \gamma(0) = p\}$. Define \sim on A by $\gamma \sim \tilde{\gamma}$ if and only if $\gamma'(0) = \tilde{\gamma}'(0)$. Then we can think of a vector $v \in T_p \mathbb{R}^n$ as the equivalence class $[\gamma]$ of a curve γ with $\gamma(0) = p$, and we can think of $T_{F(p)} \mathbb{R}^m$ as A/\sim .

This generalizes to manifolds, since we know what a smooth curve is on a manifold. But who wants to work with equivalence classes? We don't.

1.4 Germs and Derivations

Introduce a smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. We can speak of the directional derivative of f . We have

$$D_v f = \frac{d}{dt} \Big|_{t=0} f \circ \gamma = \nabla f(p) \cdot v,$$

where the quantity $\nabla f(p) \cdot v$ is independent of γ . We can therefore choose to identify $v \in T_p \mathbb{R}^n$ with the map $D_v : C^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$. But the value of $D_v f$ only depends on "the local behaviour of f at p ", and so we would like to consider two inputs of D_v to be equivalent if they are equal on a smaller neighbourhood of p . For this, we develop germs.

Definition 3. Let $f : U \rightarrow \mathbb{R}$ and $g : V \rightarrow \mathbb{R}$ be smooth functions defined on open neighbourhoods of p . We will say that $f \sim g$ if and only if $f|_W = g|_W$ for some open neighbourhood $W \subseteq U \cap V$ of p . Denote by $C_p^\infty(\mathbb{R}^n)$ the set of all such equivalence classes. The equivalence class $[f]$ is called the germ of f at p .

The map D_v is constant on germs, so it induces a map $D_v : C_p^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$ (note the notational abuse). The set of germs at p has some nice algebraic properties. That it is an "algebra" was not covered in lecture.

Proposition 1.4. $C_p^\infty(\mathbb{R}^n)$ is a vector space over \mathbb{R} . It can be made into a ring with multiplication of germs, and it can be made into an "algebra over \mathbb{R} "; a ring which is also a vector space over \mathbb{R} with the vector space scalar multiplication satisfying the homogeneity condition $a(vw) = (av) \cdot w = v \cdot (aw)$.

Proof. Was left as an exercise in class, so here's a solution. We define three operations on $C_p^\infty(\mathbb{R}^n)$:

- Vector addition: $[f] + [g] := [f + g]$.
- Vector scaling: $a[f] := [af]$.

- Ring multiplication: $[f] \cdot [g] := [fg]$.

We must first check that these operations are well defined. If $[f] = [\tilde{f}]$ and $[g] = [\tilde{g}]$, then $f = \tilde{f}$ on a neighbourhood of p and $g = \tilde{g}$ on another neighbourhood of p . It then follows that $f + g = \tilde{f} + \tilde{g}$, $af = a\tilde{f}$, and $fg = \tilde{f}\tilde{g}$ on the intersections of these neighbourhoods. By definition we have $[f + g] = [\tilde{f} + \tilde{g}]$, $[af] = [a\tilde{f}]$, and $[fg] = [\tilde{f}\tilde{g}]$, implying that our three operations are well-defined.

$C_p^\infty(\mathbb{R}^n)$ is clearly a vector space over \mathbb{R} under the first two operations, and is also a ring over the first and last operation. Homogeneity of the ring multiplication with respect to vector scaling follows from the corresponding assertion for $C^\infty(\mathbb{R}^n)$. (At this point it is just definition pushing.) \square

Note that the ring $C_p^\infty(\mathbb{R}^n)$ is commutative and has unity - the identity element of the multiplication is the germ $[x \mapsto 1]$. From now on we will abuse notation (even more) and let f denote its germ $[f] \in C_p^\infty(\mathbb{R}^n)$.

We note two properties of our map $D_v : C_p^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$:

1. D_v is linear.

2. D_v satisfies the "Leibnitz rule"

$$D_v(fg) = f(p)D_v(g) + D_v(f)g(p).$$

Any map $D : C_p^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$ satisfying the above properties is called a *derivation at p*. The set of all derivations at p is denoted \mathcal{D}_p . It turns out that this view of the tangent space is what generalizes to manifolds. Before we prove the "identification theorem", we need two lemmas.

Lemma 1.3. *If f is C^∞ on an open ball U centred at p , then there are smooth $g_i \in C^\infty(U)$ such that $g_i(p) = \frac{\partial f}{\partial x_i}(p)$ and*

$$f(x) = f(p) + \sum_{i=1}^n (x^i - p^i)g_i(x).$$

Proof. Define $\gamma(t) = p + t(x - p)$. Then

$$f(x) - f(p) = \int_0^1 \frac{d}{dt} f(\gamma(t)) dt = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x^i} \Big|_{\gamma(t)} (x^i - p^i) dt = \sum_{i=1}^n (x^i - p^i) \underbrace{\int_0^1 \frac{\partial f}{\partial x^i} (p + t(x - p)) dt}_{g_i(x)}.$$

\square

Lemma 1.4. *Derivations of constants are zero.*

Proof. Let $c \in \mathbb{R}$. Then

$$\begin{aligned} D(c) &= c \cdot D(1) = c \cdot D(1 \cdot 1) \\ &= c \cdot 1 \cdot D(1) + c \cdot D(1) \cdot 1 \\ &= 2c \cdot D(1) \\ &= 2D(c), \end{aligned}$$

implying $D(c) = 0$. \square

Theorem 1.5. *We can identify $T_p\mathbb{R}^n$ with \mathcal{D} . More specifically,*

1. \mathcal{D}_p is a vector space over \mathbb{R} .
2. The map $\Phi : T_p\mathbb{R}^n \rightarrow \mathcal{D}_p$ sending v to D_v is a vector space isomorphism.

Proof. 1. Was left as an exercise in class, so here's a proof. We must check that if $a \in \mathbb{R}$ and $D_1, D_2 \in \mathcal{D}_p$, then the function $aD_1 + D_2$ is a derivation. It is linear as a sum of linear functions. If $f, g \in C_p^\infty(\mathbb{R}^n)$, then

$$\begin{aligned} (aD_1 + D_2)(fg) &= aD_1(fg) + D_2(fg) \\ &= a[f(p)D_1(g) + D_1(f)g(p)] + f(p)D_2(g) + D_2(f)g(p) \\ &= f(p)[aD_1(g) + D_2(g)] + [aD_1(f) + D_2(f)]g(p) \\ &= f(p)(aD_1 + D_2)(g) + (aD_1 + D_2)(f)g(p), \end{aligned}$$

so $aD_1 + D_2$ satisfies the Leibnitz rule and is thus a derivation at p . Therefore \mathcal{D}_p is a vector space over \mathbb{R} .

2. We check linearity, injectivity, and surjectivity.

- Linearity: if $a \in \mathbb{R}$ and $v_1, v_2 \in T_p\mathbb{R}^n$, then for $f \in C_p^\infty(\mathbb{R}^n)$ we have

$$\Phi(av_1 + v_2)(f) = D_{av_1 + v_2}(f) = Df(p)(av_1 + v_2) = aDf(p)v_1 + Df(p)v_2 = (a\Phi(v_1) + \Phi(v_2))(f).$$

- Injectivity: suppose $D_v(f) = 0$ for all $f \in C_p^\infty(\mathbb{R}^n)$. In particular, $D_vx^i = 0$ for the i th coordinate map $x^i \in C_p^\infty(\mathbb{R}^n)$. Expanded, this says

$$0 = D_vx^i = Dx^i(p)v = e_i^T v = v^i,$$

where e_i is the i th standard basis vector of \mathbb{R}^n . So $v = 0$ if $\Phi(v) = 0$.

- Surjectivity: suppose $D \in \mathcal{D}_p$. For any $f \in C_p^\infty(\mathbb{R}^n)$, we have, by the two lemmas,

$$\begin{aligned} Df &= D \left(f(p) + \sum_{i=1}^n (x^i - p^i)g_i \right) \\ &= \sum_{i=1}^n [(p^i - p^i)Dg_i + D(x^i - p^i)g_i(p)] \\ &= \sum_{i=1}^n Dx^i \frac{\partial f}{\partial x^i}(p), \end{aligned}$$

so if we take $v = (Dx^1, \dots, Dx^n)$ then $Df = D_v f$ for all $f \in C_p^\infty(\mathbb{R}^n)$. Therefore $\Phi(v) = D$.

So Φ is a vector space isomorphism $T_p \mathbb{R}^n \xrightarrow{\sim} \mathcal{D}_p$.

□

We can finally define the tangent space to a point on an abstract manifold. The space of germs $C_p^\infty(M)$ for $p \in M$ is defined in the exact same way as in \mathbb{R}^n .

Definition 4. Let M be a smooth manifold. We say $v : C_p^\infty(M) \rightarrow \mathbb{R}$ is a derivation at $p \in M$ if v is linear and satisfies the Leibnitz rule. We define the tangent space $T_p M$ to M at p to be the set of all derivations at p .

(The following was not part of the lecture and is included for completeness.) Now we can finally define the derivative of a map between manifolds.

Definition 5. Let $F : N \rightarrow M$ be a smooth map of smooth manifolds and let $p \in N$. The map F induces a linear map $F_* : T_p N \rightarrow T_{F(p)} M$ defined by

$$(F_* X_p)(f) = X_p(f \circ F),$$

where $X_p \in T_p N$ is a derivation at p and $f \in C_{F(p)}^\infty(M)$.