TMA 4190 Introduction to Topology
Lecturer: Gereon Quick
Lecture 05[1]

5. The Inverse Function Theorem and Immersions

The Inverse Function Theorem

For our quest to understand smooth manifolds, it can be smart to study maps between manifolds (even though it sounds like making things even more difficullt; but if we know something about X and about a map $f\colon X \to Y$ then we might be able to say something interesting about Y). Anyway, there are a lot of interesting problems than can be stated in terms of properties of maps.

We have learned about the derivative of a map as a linear transformation between tangent spaces. We may think of the derivative as the **best linear approximation** at a point.

So let $f\colon X \to Y$ be a smooth map between smooth manifolds. Remember that the derivative at $x \in X$, $df_x\colon T_xX \to T_{f(x)}Y$, is a linear map between vector spaces. Since it is easier to understand linear maps, it would be nice if we could classify maps like f by the behaviour of df_x (with x varying in X).

A natural question:

How much does df_x tell us about the map f?

For the behavior df_x, there are three basic cases:

- $\dim X = \dim Y$ in which case the nicest possbile behavior of f at x is that df_x an isomorphism.
- $\dim X < \dim Y$ in which case the nicest possbile behavior of f at x is that df_x one-to-one.
- $\dim X > \dim Y$ in which case the nicest possbile behavior of f at x is that df_x onto.

We are going to consider these cases separately.

[1] Following the books of Guillemin and Pollack: Differential Topology, and Milnor: Topology from the differentiable viewpoint.

First case: df_x is an isomorphism

We begin with the nicest case when df_x is an isomorphism. This implies in particular: $\dim X = \dim Y$.

Manifolds are characterized by the way they look in a neighborhood around any point (they look like Euclidean space). So let us **think locally**. In the nicest case, f sends a neighborhood of a point x diffeomorphically to a neighborhood of $y = f(x)$. In this case, f is called a **local diffeomorphism at x**.

If f is a diffeomorphism $U \to V$ between neighborhoods U arround $x \in X$ and $y = f(x) \in Y$, respectively, let f^{-1} be its smooth inverse. Then we have $f^{-1} \circ f = \mathrm{Id}_U$ and $f \circ f^{-1} = Id_V$. Then the chain rule implies

$$d(Id_U)_x = d(f^{-1})_y \circ df_x, \text{ and } d(Id_V)_y = df_x \circ d(f^{-1})_y.$$

But we obviously have $d(\mathrm{Id}_X) = \mathrm{Id}_{T_x(X)}$ for any manifold X and any point $x \in X$. Hence df_x is an isomorphism with inverse $d(f^{-1})_{f(x)}$.

Thus a **necessary condition** for f to be a local diffeomorphism at x is that its derivative $df_x \colon T_x(X) \to T_y(Y)$ is an isomorphism.

It is an important result that this is actually a **sufficient condition**.

In order to prove this, we recall the corresponding important result for Euclidean space from Calculus:

The Inverse Function Theorem in Calculus

Suppose that $f \colon \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in an open set containing a, and $\det df_a \neq 0$, i.e. df_a is an invertible linear map $\mathbb{R}^n \to \mathbb{R}^n$. Then there is an open set $V \subseteq \mathbb{R}^n$ containing a and an open set $W \subseteq \mathbb{R}^n$ containing $f(a)$ such that $f \colon V \to W$ has a continuous inverse $f^{-1} \colon W \to V$ which is differentiable and for all $y \in W$ satisfies

$$d(f^{-1})_y = \left(df_{f^{-1}(y)}\right)^{-1}.$$

Note that this is exactly the formula you are used to from Calculus 1 where we learned

$$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}.$$

(You may be used to this formula as $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$. But the fraction here is misleading, since $(f^{-1})'(y)$ is a linear map. The superscript "to the -1"

really means **take the inverse map**! In dimension 1, the inverse map happens to be given by multiplication by the inverse number. But for linear maps or matrices in dimensions > 1, we cannot write the inverse as a fraction.)

The Inverse Function Theorem

Let X and Y be smooth manifolds. Suppose that $f\colon X \to Y$ is a smooth map whose derivative

$$df_x\colon T_x(X) \to T_{f(x)}(Y)$$

at a point $x \in X$ is an isomorphism. Then f is a **local diffeomorphism at x**.

The great thing about the IFT is that it tells us that in order to check that f is a diffeomorphism in a neighborhood of a point x, we just need to check that a **single number**, the determinant of df_x, is **nonzero**.

Idea of Proof: We can assume that X and Y are subsets in $\mathbb{R}^N$ for some large N. Let $\phi\colon U \to X$ be a local parametrization around $x \in X$, and $\psi\colon W \to Y$ a local parametrization around $y = f(x) \in Y$ with $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^n$ open and $\phi(0) = x$ and $\psi(0) = y$. (The dimension has to be the same when the tangent spaces are isomorphic.)

We define the map $\theta\colon U \to W$ as in the following diagram:

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ U \arrow[u, "\phi"] \arrow[r, "\theta=\psi^{-1}\circ f\circ\phi"'] & W. \arrow[u, "\psi"'] \end{tikzcd}$$

Then recall that df_x is defined such that the following diagram commutes

$$\begin{tikzcd} T_x(X) \arrow[r, dotted, "df_x"] & T_y(Y) \\ \mathbb{R}^k \arrow[u, "d\phi_0"] \arrow[r, "d\theta_0"'] & \mathbb{R}^l. \arrow[u, "d\psi_0"'] \end{tikzcd}$$

Our assumption is that df_x is an isomorphism which implies that **$d\theta_0$ is an isomorphism**. By the **IFT in Calculus**, this implies that there is

- an open neighborhood $V \subseteq U$ around 0 and
- an open neighborhood $V' \subseteq W$ aroud 0 such that

- $\theta_{|V}: V \to V'$ is a diffeomorphism.

Since ϕ and ψ are diffeomorphisms, $\phi(V) \subseteq X$ and $\psi(V') \subseteq Y$ are open neighborhoods of x and y, respectively. Moreover, $\phi_{|V}$ and $\psi_{|V'}$ are local parametrizations around x and y, respectively, and

$$f_{|\phi(V)}: \phi(V) \to \psi(V')$$

is a diffeomorphism. **QED**

Note that this is a **local statement**, i.e. if df_x is invertible, it only tells us that f is invertible in a **neighborhood of** x. Even if df_x is invertible for every $x \in X$, one cannot conclude that $f: X \to Y$ is globally a diffeomorphism. But such an f is a **local** diffeomorphism for every point $x \in X$. We call such a map a **local diffeomorphism** (without having to refer to a point).

Example 1: A global diffeomorphism

The map

$$(-\pi/2,\pi/2) \to \mathbb{R},\ t \mapsto \tan t$$

is a global diffeomorphism.

Example 2: A local but not global diffeomorphism

A standard example of a local diffeomorphism which is **not** a global diffeomorphism is the map

$$f: \mathbb{R}^1 \to S^1 \subset \mathbb{R}^2,\ t \mapsto (\cos t, \sin t)$$

that we have already met in Lecture 2. Let us check how this example works:

First, f is not a global diffeomorphism because it is not injective. And in Lecutre 2 we have seen that f is not a homeomorphism even when we restrict it to $[0,2\pi) \to S^1$. But anyway, S^1 **is compact** and $\mathbb{R}$ **is not**, so there is no chance of finding a diffeomorphism between them.

Second, the **IFT** tells us that f is indeed a **local diffeomorphism**, since df_t is an isomorphism for every $t \in \mathbb{R}$. For, let $t_0 \in \mathbb{R}$ such that $\cos(t_0) < 0$ (for other points the argument is similar, we just want to be able to choose a parametrization), and consider the local parametrization

$$\psi: (-1,1) \to V,\ y \mapsto (-\sqrt{1-y^2},y)$$

of S^1 around $f(t_0)$ with $V = \{(x,y) \in S^1 : x < 0\}$.

For an $\epsilon > 0$ such that both $\cos(t_0 - \epsilon) < 0$ and $\cos(t_0 + \epsilon) < 0$, we let $\phi\colon U = (t_0 - \epsilon, t_0\epsilon) \to \mathbb{R}$ be the local parametrization around t_0 given by the identity (we don't shift U to be centered around 0). Then the map $\theta\colon U \to W$ (see proof of the IFT) is defined as

$$\theta = \psi^{-1} \circ f \circ \phi, t \mapsto \sin t.$$

Then we get

$$d\theta_t\colon \mathbb{R} \to \mathbb{R},\ z \mapsto (\cos t) \cdot z$$

and

$$d\psi_t\colon \mathbb{R} \to \mathbb{R}^2,\ z \mapsto (-\frac{y}{\sqrt{1-y^2}}, 1) \cdot z.$$

Since ϕ is the identity, we have

$$df_{t_0} = d\psi_{\sin t_0} \circ d\theta_{t_0}$$

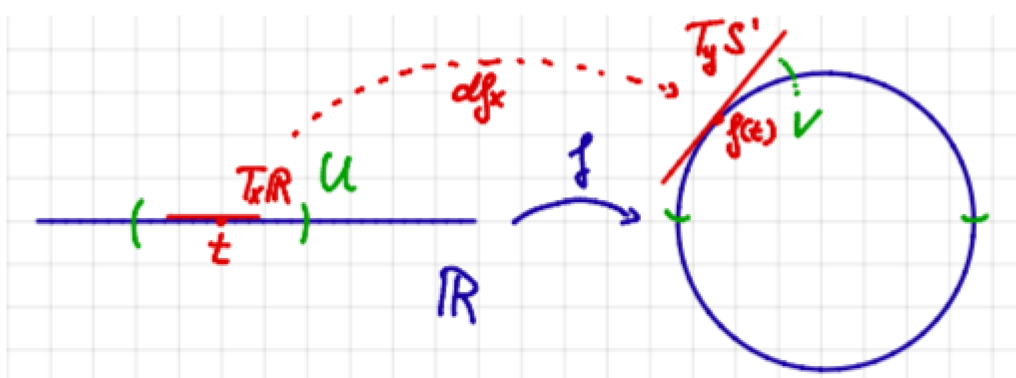
and hence

$$\begin{aligned} df_{t_0}(z) =& (-\frac{\sin t_0}{\cos t_0}, 1)(\cos t_0) \cdot z \\ =& (-\sin t_0, \cos t_0) \cdot z. \end{aligned}$$

Summarizing we have

$$\begin{aligned} df_{t_0}\colon T_{t_0}\mathbb{R} =& \mathbb{R} \to T_{f(t_0)}S^1 = d\psi(\mathbb{R}) = (-\sin(t_0), \cos(t_0)) \cdot \mathbb{R}^2, \\ & z \mapsto (-\sin(t_0), \cos(t_0)) \cdot z \end{aligned}$$

which is an isomorphism (when $\cos(t_0) \neq 0$).
For any other point in $\mathbb{R}$, there is a similar argument.



We close this first case, with an observation and some new terminology (way of speaking).

In some situations it would be nice if we could assume that the linear isomorphism df_x was the identity. This is usually not the case of course. But our freedom of choosing local parametrizations allows us to do the following. Assume that df_x is an isomorphism as in the IFT. Then, after possibly shrinkning U, we can assume $U = V$ and find a diffeomorphism $\gamma\colon U \to U$ such that $d\theta_0$ composed with $d\alpha_0$ becomes the identity.

df_x looks like the identity

If df_x is an isomorphism, we can choose local parametrizations $\phi\colon U \to X$ and $\psi\colon U \to Y$ around x and $f(x)$, respectively, with the same open domain $U \subset \mathbb{R}^n$, such that the diagram commutes:

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ U \arrow[u, "\phi"] \arrow[r, "\mathrm{Id}_U"'] & U. \arrow[u, "\psi"'] \end{tikzcd}$$

For example, in Example 2 above, we would replace
- $(-1,1)$ with $U = (t_0 - \epsilon, t_0 + \epsilon)$ and
- ψ with
$$\psi \circ \theta\colon t \mapsto (-\sqrt{1 - \sin^2 t}, \sin t) = (\cos t, \sin t)$$
(remember $\cos t < 0$ for $t \in (t_0 - \epsilon, t_0 + \epsilon)$ by our choice of t_0 and ϵ).

In general, we are going to explain how to choose suitable parametrizations in the next section.

We would like to reformulate the IFT by saying that f is equivalent to the identity. To make this precise, we introduce the following terminology:

Equivalence of maps

We say that two maps $f\colon X \to Y$ and $g\colon X' \to Y'$ are **equivalent** if there exist diffeomorphisms α and β completing a commutative square

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ X' \arrow[u, "\alpha"] \arrow[r, "g"'] & Y'. \arrow[u, "\beta"'] \end{tikzcd}$$

(One might also want to say that f and g are the same up to diffeomorphism.)

Hence the IFT says that if df_x is an isomorphism, then f is **locally equivalent at x** to the identity. Since a linear map is equivalent to the identity if and only if it is an isomorphism, we get:

IFT revisited

f is locally equivalent to the identity precisely when df_x is.

Immersions

We move on to the next case:

Second case: df_x is injective

Let us now assume $\dim X < \dim Y$. Then the best possible behavior of df_x is that

$$df_x \colon T_x(X) \to T_{f(x)}(Y)$$

is an **injective linear map**.

Let us introduce some terminology for this case.

Immersions

If df_x is injective, we say that f is an **immersion at x**. If f is an immersion at every point, we say that f is an **immersion**.

The **canonical immersion** is the standard inclusion for $n \leq m$:

$$\mathbb{R}^n \hookrightarrow \mathbb{R}^m,\ (a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n, 0, \ldots, 0).$$

Following our previous observation (i.e. up to diffeomorphism), the canonical immersion is **locally** the only immersion:

Local Immersion Theorem

Suppose that $f \colon X \to Y$ is an **immersion at x**, and $y = f(x)$. Then there exist local coordinates around x and y such that

$$f(x_1, \ldots, x_n) = (x_1, \ldots, x_n, 0, \ldots, 0).$$

In other words, f is **locally equivalent to the canonical immersion**.

How to read the Local Immersion Theorem

We should read the statement in the LIT as follows: We can choose local parametrizations $\phi\colon U \to X$ around x and $\psi\colon V \to Y$ around y such that in the commutative diagram

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ U \arrow[u, "\phi"] \arrow[r, "\theta=\psi^{-1}\circ f\circ\phi"'] & V \arrow[u, "\psi"'] \end{tikzcd}$$

the map θ is the **canonical immersion** restricted to U.

Proof of the Local Immersion Theorem:
We start by choosing any local parametrization $\phi\colon U \to X$ with $\phi(0) = x$ and $\psi\colon V \to Y$ with $\psi(0) = y$:

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ U \arrow[u, "\phi"] \arrow[r, "\theta=\psi^{-1}\circ f\circ\phi"'] & V \arrow[u, "\psi"'] \end{tikzcd}$$

Now the plan is to manipulate ϕ and ψ such that g becomes the canonical immersion.

By the assumption, we know $d\theta_0\colon \mathbb{R}^n \to \mathbb{R}^m$ is injective. Now recall that we can represent a linear map between the vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ by an $m \times n$-matrix. In order to do that we have to choose a basis for the vector spaces.
(For $\mathbb{R}^n$ we usually use the standard basis. That's why we often don't think about bases when we look at a linear map $\mathbb{R}^n \to \mathbb{R}^m$.)
By choosing a suitable basis for $\mathbb{R}^m$, we can assume that $d\theta_0$ is the matrix

$$M \begin{pmatrix} I_n \\ \hline 0 \end{pmatrix}$$

which consists of the $n \times n$-identity matrix sitting in the first n rows, and the zero $(m - n) \times n$-matrix occupying the remaining rows.

Choosing a basis

Recall that choosing a suitable basis works as follows:
Let $e_1^n, \ldots, e_n^n \in \mathbb{R}^n$ be the standard basis, and let $b_1 = d\theta_0(e_1^n), \ldots, b_n =$

$d\theta_0(e_n^n) \in \mathbb{R}^m$ be the images under $d\theta_0$. In terms of the standard basis of $\mathbb{R}^m$, the matrix for $d\theta_0$ is given by the $m\times n$-matrix A with b_i as ith column vector.
Since $d\theta_0$ is injective, the vectors $b_1,\ldots,b_n$ are linearly independent. We would to extend these vectors to a suitable basis of $\mathbb{R}^m$. Let $\operatorname{span}(b_1,\ldots,b_n)$ be the image of dg_0 in $\mathbb{R}^m$, and let $\operatorname{span}(b_1,\ldots,b_n)^\perp$ be its orthogonal complement in $\mathbb{R}^m$. Let $c_{n+1},\ldots,c_m$ be a basis for $\operatorname{span}(b_1,\ldots,b_n)^\perp$. (You learned in Matte 3 how to find such a basis: $\operatorname{span}(b_1,\ldots,b_n)^\perp$ is the null space or kernel of the matrix A above.) We define a new basis for $\mathbb{R}^m$ as $b_1,\ldots,b_n,c_{n+1},\ldots,c_m \in \mathbb{R}^m$.
In terms of this basis, the matrix of $d\theta_0$ is exactly $M\left(\frac{I_n}{0}\right)$. (Recall also that, in order to switch from the standard basis of $\mathbb{R}^m$ to that new basis, we apply the $m\times m$-matrix B whose first n columns are $b_1,\ldots,b_n$ and remaining $m-n$ columns are $c_{n+1},\ldots,c_m$. Again, since $d\theta_0$ is injective, B is an invertible matrix which sends the standard basis $e_1^m,\ldots,e_m^m \in \mathbb{R}^m$ to the basis $b_1,\ldots,b_n,c_{n+1},\ldots,c_m \in \mathbb{R}^m$.)

Back to the proof: We define a new map

$$\Theta\colon U\times\mathbb{R}^{m-n} \to \mathbb{R}^m, \text{ by } \Theta(x,z) = \theta(x) + (0,z).$$

It is related to θ by the picture

$$\begin{tikzcd} U \arrow[rr, "\text{canonical}", "\text{immersion}"'] \arrow[dr, "\theta"'] & & U\times\mathbb{R}^{m-n} \arrow[dl, "\Theta"] \\ & \mathbb{R}^m. & \end{tikzcd}$$

Since θ is a local diffeomorphism at 0, we can choose U and V small enough such that θ sends open sets to open sets. Moreover, the matrix representing $d\Theta_0$ (in our chosen basis) is just the $m\times m$-identity matrix I_m (it's $M\left(\frac{I_n}{0}\right)$ with the zero replaced with the remainind standard basis vectors $e_{n+1}^m,\ldots,e_m^m$). By the Inverse Function Theorem, this implies that Θ is a local diffeomorphism of $\mathbb{R}^m$ of itself at 0. Since ψ and Θ are local diffeomorphisms at 0, so is the composition $\psi\circ\Theta$. Hence we can use $\psi\circ\Theta$ as a local parametrization around y. Finally, after possibly shrinking U and V we get the desired commutative diagram

$$\begin{tikzcd} X \arrow[r, "f"] & Y \\ U \arrow[u, "\phi"] \arrow[r, "\text{canonical}", "\text{immersion}"'] & V \arrow[u, "\psi\circ\Theta"'] \end{tikzcd}$$

which proves the theorem. **QED**

Still an immersion in a neighborhood

We observe from the proof of the theorem that if $f\colon X \to Y$ is an immersion at x, then it is also an **immersion for all points in a neighborhood of x**. For, local parametrization $\phi\colon U \to X$ of the proof also parametrizes any point in the image of ϕ which is an open subset around x (open because ϕ is a diffeomorphism onto its image).

Local nature

To be an immersion is a **local condition**. For example, if $\dim X = \dim Y$, then being an immersion means being a local diffeomorohism. Hence in order to say more about f we need to add some (more global) topological properites to the local differential data.
For example, for a local diffeomorphism to be a global one, it has to be one-to-one and onto.