The Implicit Function Theorem

16

§1. Manifolds (take II)

I want to start by reminding you of some things from Lecture 11.

There we characterized a k-dim'l manifold (or k-manifold) as a level set of a C^1 function

$$\vec{F} : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}.$$

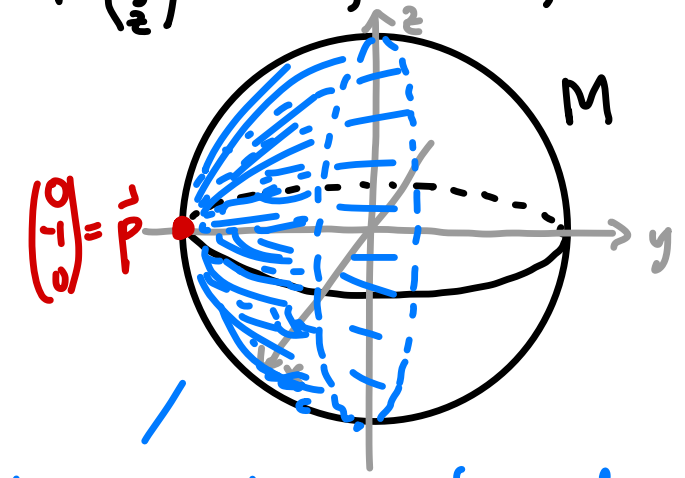
Taking $M := \vec{F}^{-1}(\vec{c})$, we made the provisional

Definition: (i) M is smooth at $\vec{a} \in M$ if $\operatorname{rank}(D\vec{F}(\vec{a})) = n-k$.

(ii) M is a manifold (or "smooth") if it is smooth at every point.

We also introduced the Implicit Function Theorem, which says that you can always write a manifold locally as the graph of a function $\begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \vec{f}\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ after relabeling coordinates.

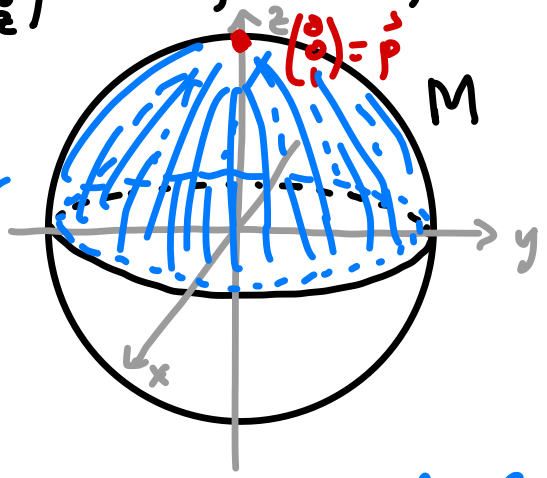
Ex/ $\vec{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2+y^2+z^2-1$, $c = 0$



we can write $M \cap \{y < 0\}$ as the graph of $y = -\sqrt{1-x^2-z^2}$ over the xz-plane

We also introduced the Implicit Function Theorem, which says that you can always write a manifold locally as the graph of a function $\begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \vec{f}\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ after relabeling coordinates.

Ex / $\vec{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2+y^2+z^2-1, \quad c = 0$



Similarly, we can write $M \cap \{z > 0\}$ as the graph of $z = +\sqrt{1-x^2-y^2}$ over the xy-plane.

//

Here is how we stated the Implicit Function Theorem in case $k = n-1$ (M a hypersurface):

Implicit Function Theorem (special case)

Let M be the level set $f^{-1}(0)$ of $f: \mathbb{R}^n \to \mathbb{R}$, and $\vec{a} \in M$ a point at which $\frac{\partial f}{\partial x_n}(\vec{a}) \neq 0$.

Writing $\vec{x} = \begin{pmatrix} \tilde{x} \\ x_n \end{pmatrix}$ and $\vec{a} = \begin{pmatrix} \tilde{a} \\ a_n \end{pmatrix}$, i.e. $\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}$ and $\begin{pmatrix} a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}$,

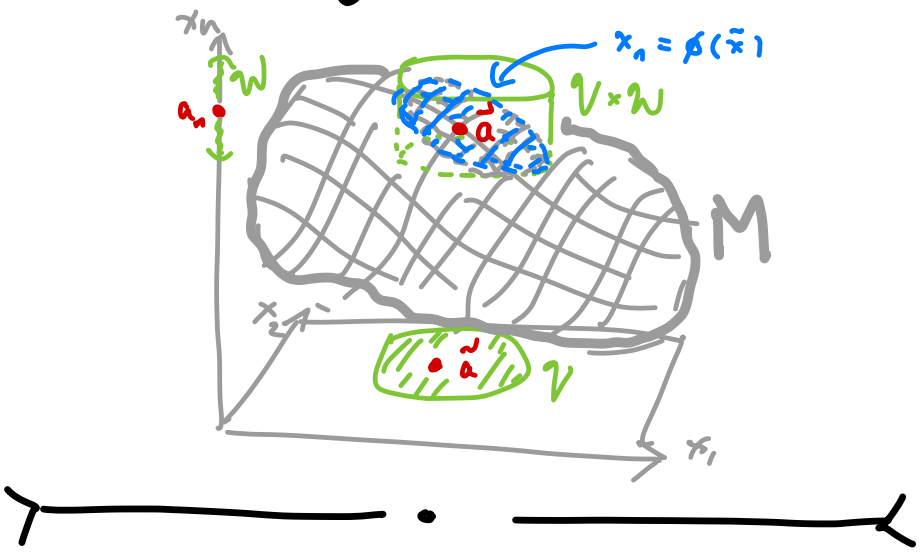
there are neighborhoods $\mathcal{V} \subset \mathbb{R}^{n-1}$ of $\tilde{a}$ and $\mathcal{W} \subset \mathbb{R}$ of a_n and a C^1 function $\phi: \mathcal{V} \to \mathcal{W}$ s.t. for $\tilde{x} \in \mathcal{V}$ and $x_n \in \mathcal{W}$,

$$\begin{pmatrix} \tilde{x} \\ x_n \end{pmatrix} \in M \iff x_n = \phi(\tilde{x}).$$

We also had in this case

$$\frac{\partial \phi}{\partial x_j}(\tilde{a}) = \frac{-\partial f / \partial x_j (\vec{a})}{\partial f / \partial x_n (\vec{a})}$$

and accompanying picture



Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ | image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$

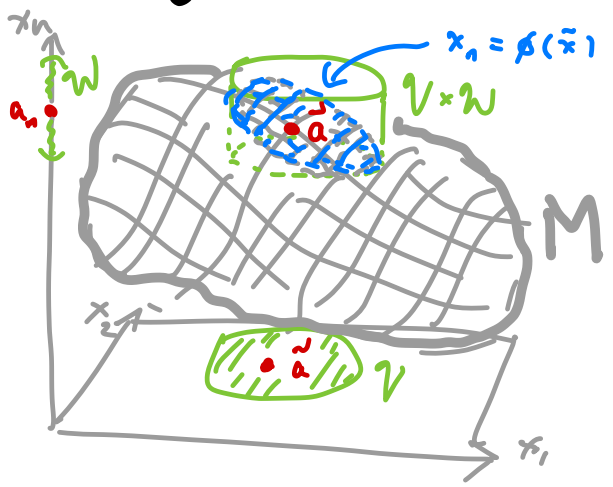
V open in $\mathbb{R}^k$ = coordinate k-plane

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

U open in $\mathbb{R}^k$

and accompanying picture



Returning to the definition of manifold, we are now ready for something less provisional. The functions here are going to be assumed smooth, i.e. C^∞ (infinitely differentiable).

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$)

V open in $\mathbb{R}^k$ = coordinate k-plane

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

U open in $\mathbb{R}^k$

this one is the Implicit Function Theorem!

Why are these characterizations equivalent?

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$)

V open in $\mathbb{R}^k$ = coordinate k-plane

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

U open in $\mathbb{R}^k$

this one is immediate because $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$ is a parametric description

Why are these characterizations equivalent?

- (ii) $\Rightarrow$ (i) is IFT

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$)
(V open in $\mathbb{R}^k$ = coordinate k-plane)

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.
(U open in $\mathbb{R}^k$)

What about this one?

Why are these characterizations equivalent?

- (ii) $\Rightarrow$ (i) is IFT
- (i) $\Rightarrow$ (iii) is easy

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$)

V open in $\mathbb{R}^k$ = coordinate k-plane

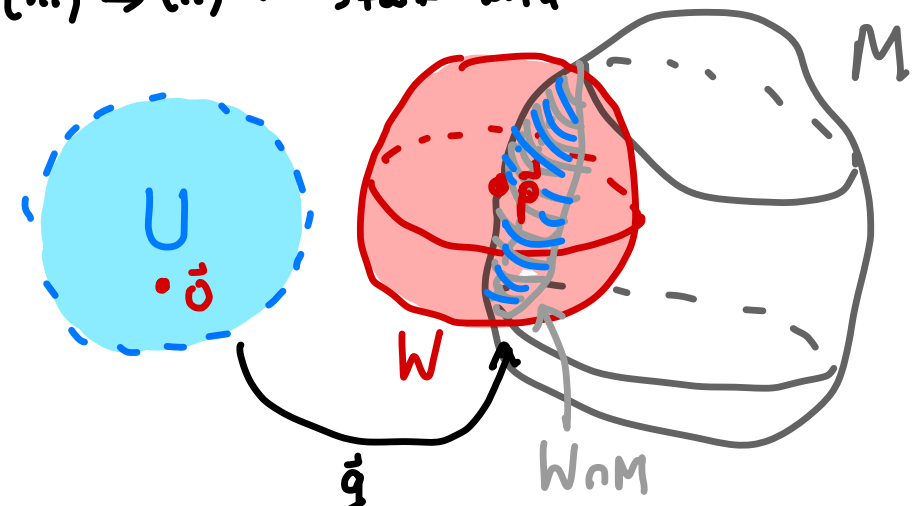
(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

U open in $\mathbb{R}^k$

Why are these characterizations equivalent?

- (ii) $\Rightarrow$ (i) is IFT
- (i) $\Rightarrow$ (iii) is easy
- (iii) $\Rightarrow$ (ii): start with



$$D\vec{g}(\vec{0}) = \begin{bmatrix} A \\ B \end{bmatrix} \begin{matrix} \updownarrow k \\ \updownarrow n-k \end{matrix} \quad \text{with } A \text{ invertible}$$

(A: $k \times k$)

$$\vec{G}\begin{pmatrix} \vec{u} \\ \vec{v} \end{pmatrix} := \vec{g}(\vec{u}) + \begin{pmatrix} \vec{0} \\ \vec{v} \end{pmatrix} \Rightarrow D\vec{G}\begin{pmatrix} \vec{0} \\ \vec{0} \end{pmatrix} = \left[\begin{array}{c|c} A & 0 \\ \hline B & I \end{array}\right] \quad (n \times n)$$

$\vec{G}$: map from $U \times \mathbb{R}^{n-k} \to W$ — invertible by HW!

The INVERSE FUNCTION THEOREM will an inverse $\vec{H} = \begin{pmatrix} \vec{H}_1 \\ \vec{H}_2 \end{pmatrix}: W \to U \times \mathbb{R}^{n-k}$ of $\vec{G}$.

($\vec{H}$: possibly shrinking this set)

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (V open in $\mathbb{R}^k$ = coordinate k-plane) / image of $\vec{x} \mapsto \begin{pmatrix} \vec{x} \\ \vec{f}(\vec{x}) \end{pmatrix}$

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ (U open in $\mathbb{R}^k$) with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

Why are these characterizations equivalent?

- (ii) $\Rightarrow$ (i) is IFT
- (i) $\Rightarrow$ (iii) is easy
- (iii) $\Rightarrow$ (ii): start with $\vec{g}: U \to M \cap W$ as in the defn.

$$D\vec{g}(\vec{0}) = \begin{bmatrix} A \\ B \end{bmatrix} \begin{matrix} \updownarrow k \\ \updownarrow n-k \end{matrix} \quad \text{with } A \text{ invertible}$$

(*)
$$\vec{G}\begin{pmatrix} \vec{u} \\ \vec{v} \end{pmatrix} := \vec{g}(\vec{u}) + \begin{pmatrix} \vec{0} \\ \vec{v} \end{pmatrix} \Rightarrow D\vec{G}\begin{pmatrix} \vec{0} \\ \vec{0} \end{pmatrix} = \left[\begin{array}{c|c} A & 0 \\ \hline B & I \end{array}\right] \quad (n \times n)$$

$\vec{G}$: map from $U \times \mathbb{R}^{n-k} \to W$ — invertible by HW!

The INVERSE FUNCTION THEOREM will an inverse $\vec{H} = \begin{pmatrix} \vec{H}_1 \\ \vec{H}_2 \end{pmatrix}: W \to U \times \mathbb{R}^{n-k}$ of $\vec{G}$. For each $\vec{x} \in W$, we write

$$\vec{x} = \vec{G}\begin{pmatrix} \vec{u} \\ \vec{v} \end{pmatrix} = \vec{G}\begin{pmatrix} \vec{H}_1(\vec{x}) \\ \vec{H}_2(\vec{x}) \end{pmatrix}.$$

Then $\vec{x} \in M \Leftrightarrow \vec{x} = \vec{g}(\vec{u})$ for some $\vec{u} \in U$

$\overset{(*)}{\Leftrightarrow} \vec{v} = \vec{0} \Leftrightarrow \vec{H}_2(\vec{x}) = \vec{0}$. So

$M \cap W = \vec{H}_2^{-1}(\vec{0})$ is a level set. $\square$

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (V open in $\mathbb{R}^k$ = coordinate k-plane) / image of $\vec{x} \mapsto (\vec{x}, \vec{f}(\vec{x}))$

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$. (U open in $\mathbb{R}^k$)

Ex/ $\vec{G}\begin{pmatrix} u \\ v_1 \\ v_2 \end{pmatrix} = \vec{g}(u) + \begin{pmatrix} 0 \\ v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u \\ u^2+v_1 \\ u^3+v_2 \end{pmatrix}$

has inverse $\vec{H}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y-x^2 \\ z-x^3 \end{pmatrix}$ ($\vec{H}_1$, $\vec{H}_2$)

(i) $\vec{f}(x) = \begin{pmatrix} x^2 \\ x^3 \end{pmatrix}$

$\vec{f}: \mathbb{R} \to \mathbb{R}^2$

(ii) $\vec{0} = \vec{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y-x^2 \\ z-x^3 \end{pmatrix}$

$\vec{F}: \mathbb{R}^3 \to \mathbb{R}^2$

(iii) $\vec{g}(u) = \begin{pmatrix} u \\ u^2 \\ u^3 \end{pmatrix}$ "twisted cubic curve"

$\vec{g}: \mathbb{R} \to \mathbb{R}^3$

//

Definition: A subset $M \subset \mathbb{R}^n$ is a k-manifold if every point $\vec{p} \in M$ has a neighborhood $W \subset \mathbb{R}^n$ s.t. we can describe $M \cap W$ in one of the following 3 equivalent ways:

(i) **EXPLICIT:** as the graph of a C^∞ $\vec{f}: V \to \mathbb{R}^{n-k}$ (image of $\vec{x} \mapsto \begin{pmatrix} \vec{x} \\ \vec{f}(\vec{x}) \end{pmatrix}$)

(V open in $\mathbb{R}^k$ = coordinate k-plane)

(ii) **IMPLICIT:** as the level set $\vec{F}^{-1}(\vec{0})$ of a C^∞ $\vec{F}: W \to \mathbb{R}^{n-k}$ with $\operatorname{rank}(D\vec{F}(\vec{x})) = n-k$ $(\forall \vec{x} \in M \cap W)$

(iii) **PARAMETRIC:** as the image of a C^∞ $\vec{g}: U \to \mathbb{R}^n$ with continuous g^{-1} and $\operatorname{rank}(D\vec{g}(\vec{u})) = k$ $(\forall \vec{u} \in U)$.

(U open in $\mathbb{R}^k$)

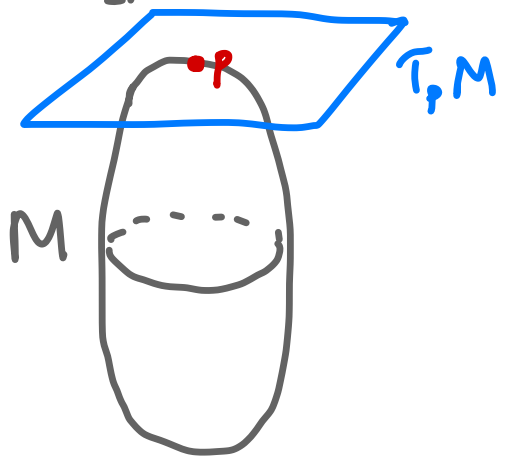
PROBLEM: Given the parametric surface M described by $\begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} \sin(u)\cos(v) \\ \sin(u)\sin(v) \\ 2\cos(u) \end{pmatrix}$ $u \in (0, \pi)$, $v \in (0, 2\pi)$

give an implicit description of M and an explicit description in a neighborhood of $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$. (Can you find its tangent plane there?)

(ii) $x^2 + y^2 + \frac{1}{4}z^2 = 1$

(i) $z = 2\sqrt{1 - x^2 - y^2}$

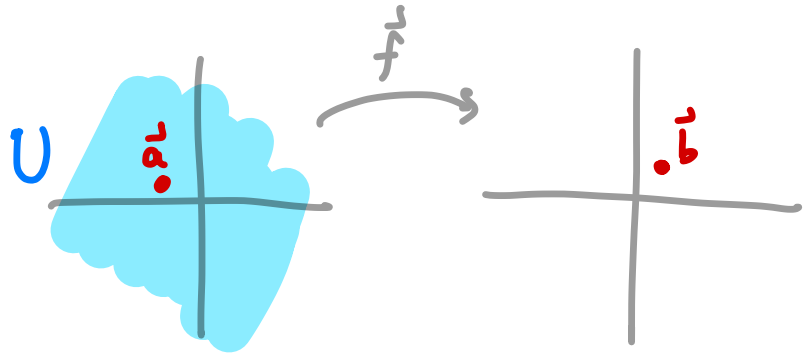
Using (ii), $T_{\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}} M = \{z = 2\}$.



§ 2. Inverse Function Theorem

This is the result we used above. It says that if the derivative of a function is invertible somewhere, it has a local inverse.

Theorem: Let $U \subset \mathbb{R}^n$ be open, and $\vec{f}: U \to \mathbb{R}^n$ a C^1 map with $\vec{b} = \vec{f}(\vec{a})$ and $D\vec{f}(\vec{a})$ invertible.

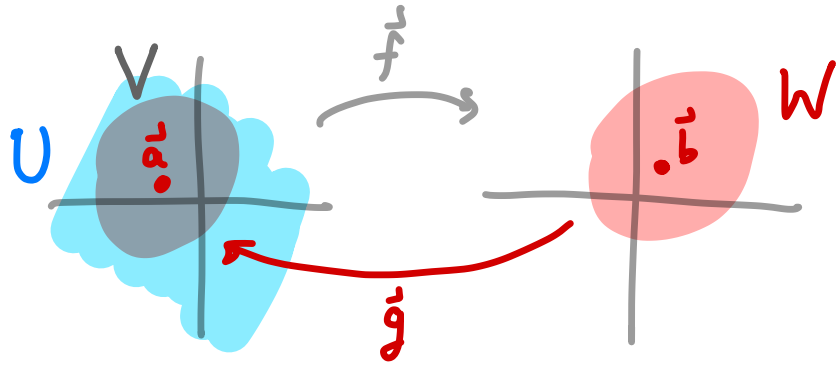


§ 2. Inverse Function Theorem

This is the result we used above. It says that if the derivative of a function is invertible somewhere, it has a local inverse.

Theorem: Let $U \subset \mathbb{R}^n$ be open, and $\vec{f}: U \to \mathbb{R}^n$ a C^1 map with $\vec{b} = \vec{f}(\vec{a})$ and $D\vec{f}(\vec{a})$ invertible.

Then there exist neighborhoods V of $\vec{a}$ & W of $\vec{b}$, and a C^1 map $\vec{g}: W \to V$ such that $g \circ f$ & $f \circ g$ are the identity.



Note that by the Chain Rule we then get

$$D\vec{f}(\vec{g}(\vec{y}))\, D\vec{g}(\vec{y}) = I \Rightarrow \boxed{D\vec{g}(\vec{y}) = Df(g(y))^{-1}}$$

which generalizes the 1-variable formula

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

§ 2. Inverse Function Theorem

This is the result we used above. It says that if the derivative of a function is invertible somewhere, it has a local inverse.

Theorem: Let $U \subset \mathbb{R}^n$ be open, and $\vec{f}: U \to \mathbb{R}^n$ a C^1 map with $\vec{b} = \vec{f}(\vec{a})$ and $D\vec{f}(\vec{a})$ invertible.

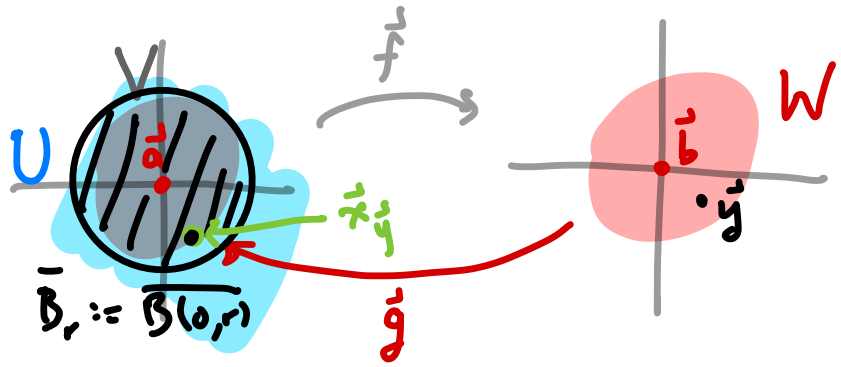
Then there exist neighborhoods V of $\vec{a}$ & W of $\vec{b}$, and a C^1 map $\vec{g}: W \to V$ such that $g \circ f$ & $f \circ g$ are the identity.

Note that by the Chain Rule we then get

$$D\vec{f}(\vec{g}(\vec{y}))\, D\vec{g}(\vec{y}) = I \Rightarrow \boxed{D\vec{g}(\vec{y}) = Df(g(y))^{-1}}$$

which generalizes the 1-variable formula

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



Proof: We may assume $\vec{a} = \vec{b} = \vec{0}$ & $D\vec{f}(\vec{0}) = I_n$

$\vec{f}$ $C^1 \Rightarrow D\vec{f}$ continuous $\Rightarrow \exists\, r > 0$ s.t.

$\|D\vec{\phi}(\vec{x})\| = \|Df(\vec{x}) - I\| \le \frac{1}{2}$ for $\|\vec{x}\| < r$.

Take $W = B(\vec{0}, r/2)$ and $V = \vec{f}^{-1}(W) \cap B(\vec{0}, r)$.

Define $\vec{\phi}(\vec{x}) := \vec{x} - \vec{f}(\vec{x}) + \vec{y}$ for any fixed $\vec{y} \in W$.

$$\|\vec{\phi}(\vec{x})\| \le \|\vec{\phi}(\vec{x}) - \vec{\phi}(\vec{0})\| + \|\vec{\phi}(\vec{0})\|$$

$$\overset{\text{MVT}}{\le} \left(\max_{x \in \bar{B}_r} \|D\vec{\phi}(\vec{x})\|\right)\|\vec{x} - \vec{0}\| + \|\vec{y}\|$$

$$\le \frac{1}{2}\|\vec{x}\| + \|\vec{y}\| < \frac{r}{2} + \frac{r}{2} = r$$

$\Rightarrow \vec{\phi}(\bar{B}_r) \subset B_r \left(\in \bar{B}_r\right)$.

$\|\vec{\phi}(\vec{x}) - \vec{\phi}(\vec{y})\| \overset{\text{MVT}}{\le} \frac{1}{2}\|\vec{x} - \vec{y}\| \Rightarrow \vec{\phi}$ is a contraction mapping, so has a unique fixed point $\vec{x}_{\vec{y}} \in B_r$.

Theorem: Let $U \subset \mathbb{R}^n$ be open, and $\vec{f}: U \to \mathbb{R}^n$ a C^1 map with $\vec{b} = \vec{f}(\vec{a})$ and $D\vec{f}(\vec{a})$ invertible.

Then there exist neighborhoods V of $\vec{a}$ & W of $\vec{b}$, and a C^1 map $\vec{g}: W \to V$ such that $g \circ f$ & $f \circ g$ are the identity. ($\Rightarrow D\vec{g}(\vec{y}) = Df(g(\vec{y}))^{-1}$)

$\vec{x}_{\vec{y}} = \vec{\phi}(\vec{x}_{\vec{y}}) = \vec{x}_{\vec{y}} - \vec{f}(\vec{x}_{\vec{y}}) + \vec{y}$ (fixed pt. condition)

$\Rightarrow f(\vec{x}_{\vec{y}}) = \vec{y}$!!!

Put $\vec{g}(\vec{y}) := \vec{x}_{\vec{y}}$. Is this C^1?

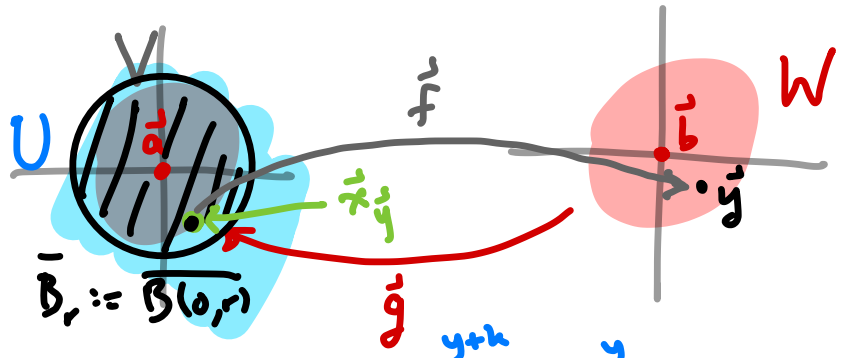
First, we MVT to check it is C^0 (easy but a bit tedious, see Shifrin).

To see that it is differentiable, write

$\vec{x} = g(\vec{y})$, $\vec{x} + \vec{h} = \vec{g}(\vec{y} + \vec{k})$, $A = D\vec{f}(\vec{x})$:

$$\lim_{\vec{k} \to \vec{0}} \frac{\vec{g}(\vec{y}+\vec{k}) - \vec{g}(\vec{y}) - A^{-1}\vec{k}}{\|\vec{k}\|} = \dots$$

($\vec{g}(\vec{y}+\vec{k}) = \vec{x}+\vec{h}$, $\vec{g}(\vec{y}) = \vec{x}$)



$$\dots = \lim_{\vec{k} \to \vec{0}} \frac{\vec{h} - A^{-1}\left(\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x})\right)}{\|\vec{k}\|}$$

($\vec{y}+\vec{k}$, $\vec{y}$)

$$= \lim_{\vec{k} \to \vec{0}} -A^{-1} \frac{\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - A\vec{h}}{\|\vec{h}\|} \cdot \frac{\|\vec{h}\|}{\|\vec{k}\|}$$

(by uniform continuity of $\vec{g}$: $\leq C\|\vec{k}\|$)

$-A^{-1}$: constant

$\frac{\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - A\vec{h}}{\|\vec{h}\|}$: goes to 0 by differentiability of $\vec{f}$

$\frac{\|\vec{h}\|}{\|\vec{k}\|}$: bounded

$= \vec{0}$

Finally, to get that $\vec{g}$ is C^1, i.e. that $D\vec{g}(\vec{y})$ is continuous in $\vec{y}$, write it as the composition

$$y \xrightarrow[g]{C^0} g(y) \xrightarrow[Df]{C^0} DF(g(\vec{y})) \xrightarrow[(\cdot)^{-1}]{C^0} Df(\vec{g}(\vec{y}))^{-1}$$

(by your HW for next week)

of continuous functions. □

Theorem: Let $U \subset \mathbb{R}^n$ be open, and $\vec{f}: U \to \mathbb{R}^n$ a C^1 map with $\vec{b} = \vec{f}(\vec{a})$ and $D\vec{f}(\vec{a})$ invertible.

Then there exist neighborhoods V of $\vec{a}$ & W of $\vec{b}$, and a C^1 map $\vec{g}: W \to V$ such that $g \circ f$ & $f \circ g$ are the identity. ($\Rightarrow D\vec{g}(\vec{y}) = Df(g(y))^{-1}$)

EXERCISE: At what points does $\vec{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + e^y \\ y + e^x \end{pmatrix}$ have a local C^1 inverse $\vec{g}$? Calculate $D\vec{g}\left(\vec{f}\begin{pmatrix} x \\ y \end{pmatrix}\right)$.

↖ invertibility condition

$$0 \neq \det \begin{pmatrix} 1 & e^y \\ e^x & 1 \end{pmatrix} = 1 - e^{x+y} \iff$$

$$e^{x+y} \neq 1 \iff x + y \neq 0 \iff x \neq -y.$$

$$D\vec{g}(\vec{f}(\vec{x})) = D\vec{f}(\vec{x})^{-1} = \begin{pmatrix} 1 & e^y \\ e^x & 1 \end{pmatrix}^{-1} = \frac{1}{1 - e^{x+y}} \begin{pmatrix} 1 & -e^y \\ -e^x & 1 \end{pmatrix}$$

§3. Implicit Function Theorem

Theorem: Let $U \subset \mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$ be open, $\vec{F} : U \to \mathbb{R}^m$ be a C^1 map, and $M := \left\{ \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \in U \;\middle|\; \vec{F}\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \vec{0} \right\}$.

Writing $D\vec{F} = \left[\frac{\partial \vec{F}}{\partial \vec{x}} \;\middle|\; \frac{\partial \vec{F}}{\partial \vec{y}} \right]$, assume $\frac{\partial \vec{F}}{\partial \vec{y}} \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}$ is invertible.

Notation:

$$\frac{\partial \vec{F}}{\partial \vec{x}} = \left[\frac{\partial F_i}{\partial x_j} \right]_{\substack{i=1,\dots,m \\ j=1,\dots,n-m}}$$

$$\frac{\partial \vec{F}}{\partial \vec{y}} = \left[\frac{\partial F_i}{\partial y_j} \right]_{i,j=1,\dots,m} \quad (\text{square } m \times m)$$

Invertibility of $\partial \vec{F} / \partial \vec{y}$ is one way to have $D\vec{F}$ of maximal rank ($= m$).

§ 3. Implicit Function Theorem

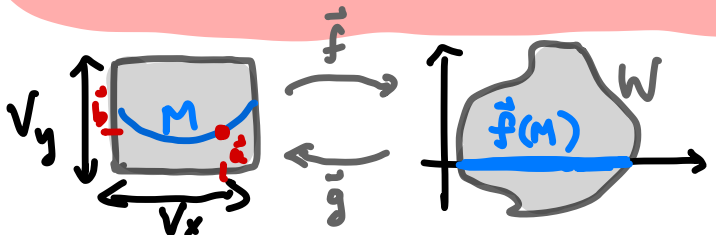
Theorem: Let $U \subset \mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$ be open, $\vec{F}: U \to \mathbb{R}^m$ be a C^1 map, and $M := \left\{ \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \in U \mid \vec{F}\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \vec{0} \right\}$.

Writing $D\vec{F} = \left[\frac{\partial \vec{F}}{\partial \vec{x}} \mid \frac{\partial \vec{F}}{\partial \vec{y}} \right]$, assume $\frac{\partial \vec{F}}{\partial \vec{y}} \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}$ is invertible. Then

- $\exists$ neighborhoods $\begin{cases} V_x \subset \mathbb{R}^{n-m} \text{ of } \vec{a} \\ V_y \subset \mathbb{R}^m \text{ of } \vec{b} \end{cases}$ and a C^1 function $\vec{\phi}: V_x \to V_y$ s.t.

  $$\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \in M \iff \vec{y} = \vec{\phi}(\vec{x}).$$

- $$D\vec{\phi}(\vec{x}) = -\frac{\partial \vec{F}}{\partial \vec{y}}\begin{pmatrix} \vec{x} \\ \vec{\phi}(\vec{x}) \end{pmatrix}^{-1} \frac{\partial \vec{F}}{\partial \vec{x}}\begin{pmatrix} \vec{x} \\ \vec{\phi}(\vec{x}) \end{pmatrix}.$$



Proof: Define $\vec{f}: U \to \mathbb{R}^n$ by $\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \mapsto \begin{pmatrix} \vec{x} \\ \vec{F}\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \end{pmatrix}$.

again, by old HW exercise

Then $D\vec{f} = \left[\begin{array}{c|c} I & 0 \\ \hline \frac{\partial \vec{F}}{\partial \vec{x}} & \frac{\partial \vec{F}}{\partial \vec{y}} \end{array}\right]$ is invertible at $\begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}$.

Apply the Inverse Function Theorem to get $\vec{g}: W \to V \ (= V_x \times V_y)$ inverse to $\vec{f}$, then define $\vec{\phi}: V_x \to V_y$ by $\begin{pmatrix} \vec{x} \\ \vec{\phi}(\vec{x}) \end{pmatrix} := \vec{g}\begin{pmatrix} \vec{x} \\ \vec{0} \end{pmatrix}$.

Now
$$\begin{aligned} \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \in M &\iff \vec{F}\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \vec{0} \\ &\iff \vec{f}\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \begin{pmatrix} \vec{x} \\ \vec{0} \end{pmatrix} \\ &\iff \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \vec{g}\begin{pmatrix} \vec{x} \\ \vec{0} \end{pmatrix} \left(= \begin{pmatrix} \vec{x} \\ \vec{\phi}(\vec{x}) \end{pmatrix} \right) \\ &\iff \vec{y} = \vec{\phi}(\vec{x}). \end{aligned}$$

Since the composition $V_x \xrightarrow{\begin{pmatrix} id \\ \phi \end{pmatrix}} V_x \times V_y \xrightarrow{\vec{F}} \mathbb{R}^m$ is zero,

$$\begin{aligned} 0 = D\left(F \circ \begin{pmatrix} id \\ \phi \end{pmatrix}\right) &= DF\begin{pmatrix} x \\ \phi(x) \end{pmatrix} \begin{pmatrix} I \\ D\phi(x) \end{pmatrix} \\ &= \left(\frac{\partial \vec{F}}{\partial \vec{x}} \mid \frac{\partial \vec{F}}{\partial \vec{y}} \right) \begin{pmatrix} I \\ D\phi(x) \end{pmatrix} = \frac{\partial \vec{F}}{\partial \vec{x}}\begin{pmatrix} x \\ \phi(x) \end{pmatrix} + \frac{\partial \vec{F}}{\partial \vec{y}}\begin{pmatrix} x \\ \phi(x) \end{pmatrix} D\phi(x) \end{aligned}$$

yields the formula shown. □