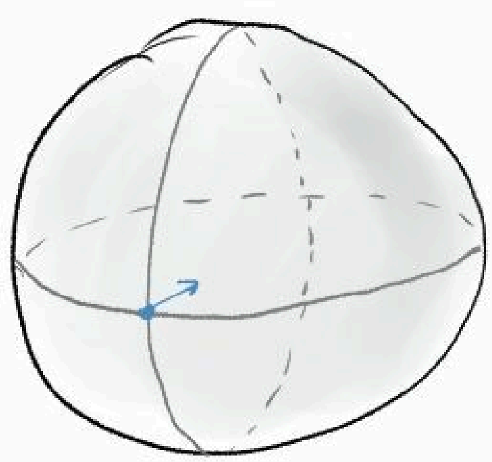
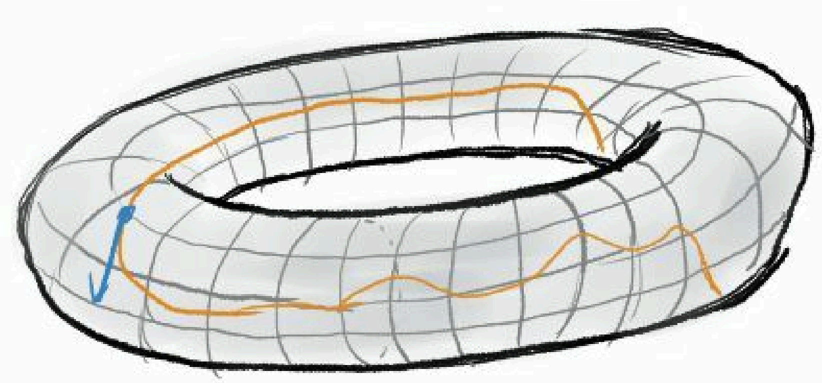
EleutherAI Math Reading Group

Differential Geometry I: Manifolds

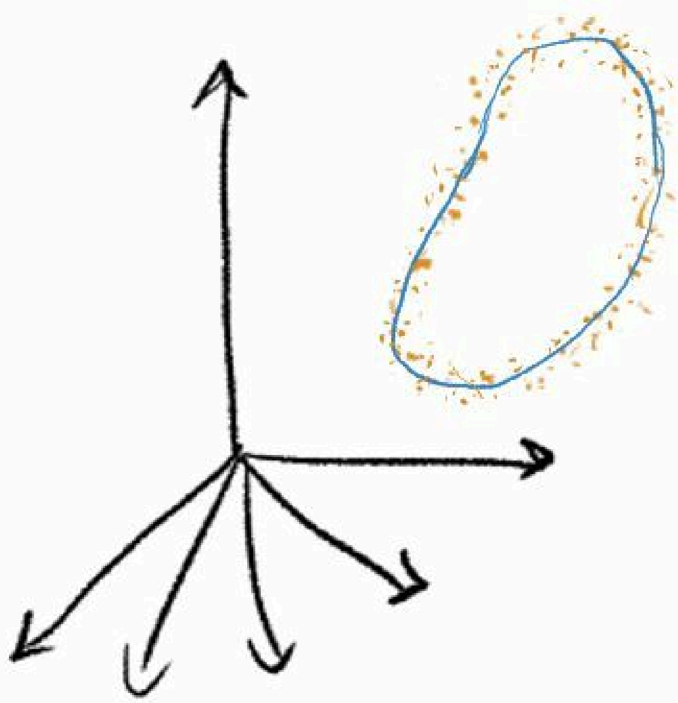




Motivation

Manifold Hypothesis

Low dimensional nonlinear space



High-dimensional ambient space

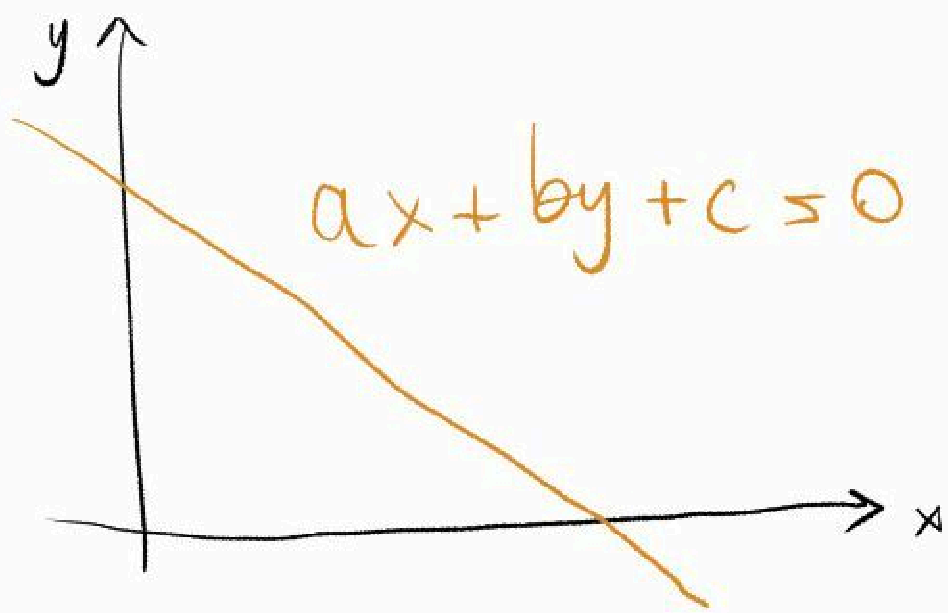
Latent coordinates

e.g. Autoencoders, GANs, ...

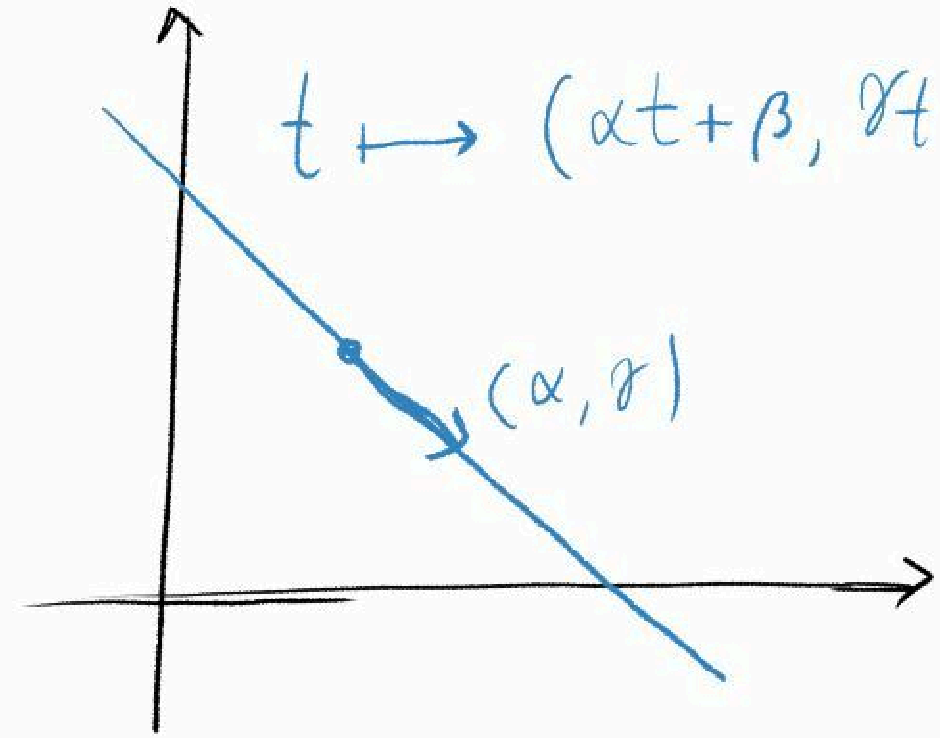
↳ Space that we can move around in

"Locally linear"

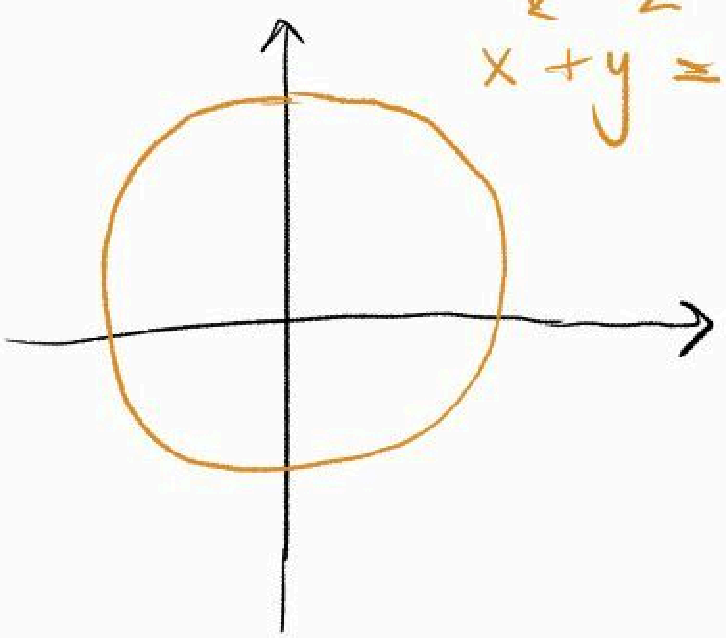
Implicit vs Explicit equations



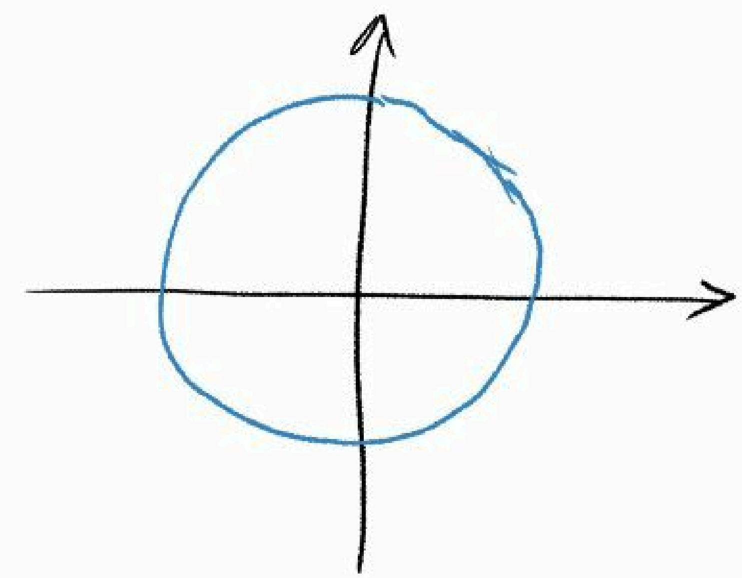
$$ax + by + c = 0$$



$$t \longmapsto (\alpha t + \beta, \gamma t + \delta)$$

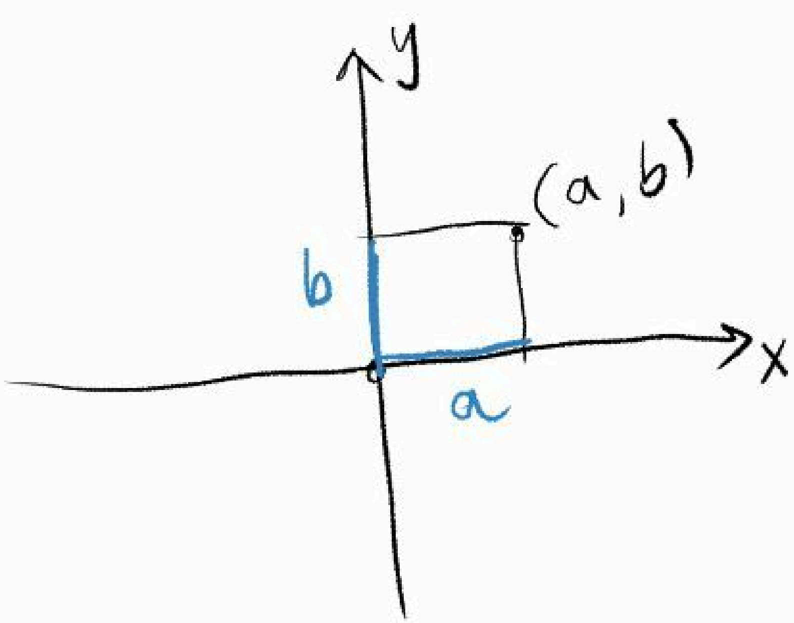


$$x^2 + y^2 = r^2$$

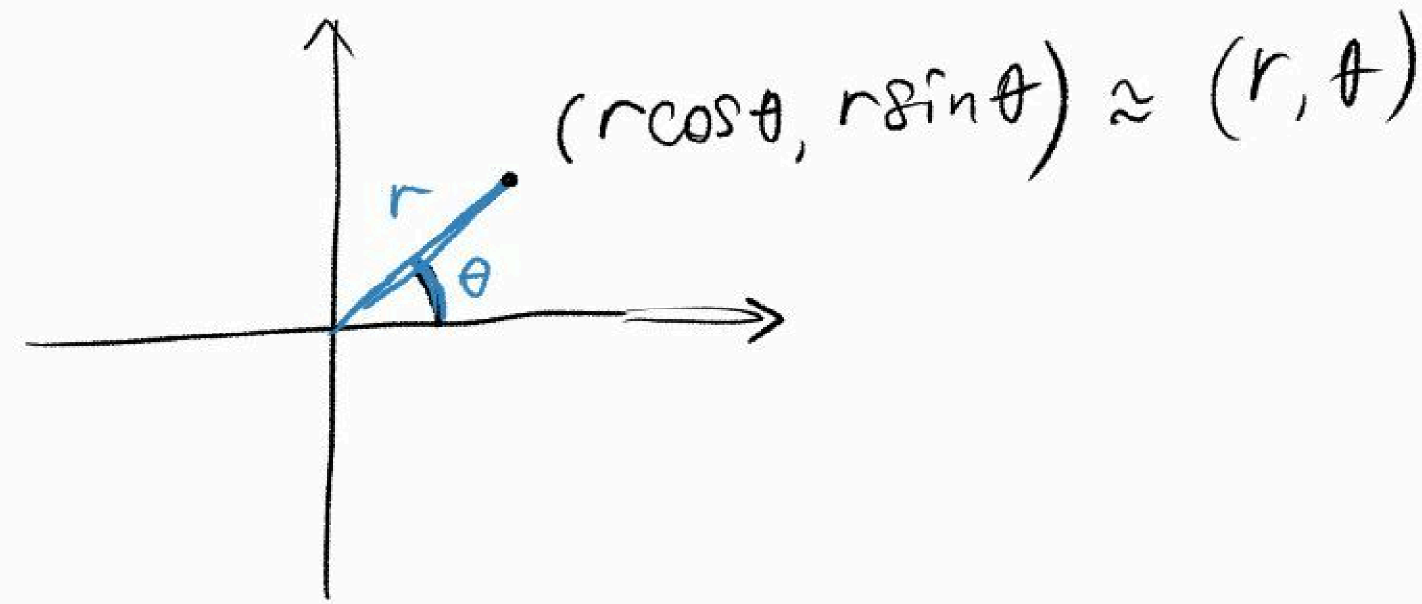


$$t \longmapsto (r\cos t, r\sin t)$$

Coordinate Systems



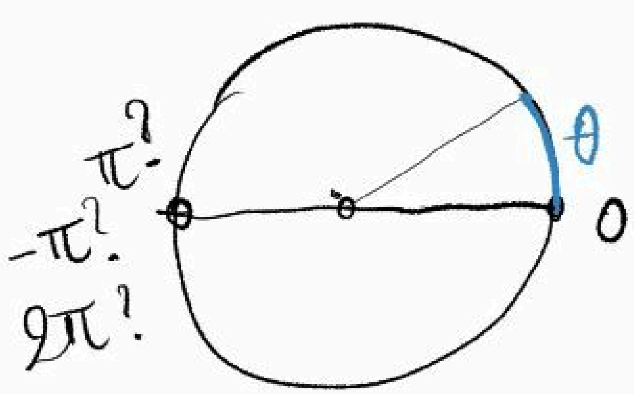
Cartesian coordinates



Polar coordinates

$$\begin{cases} r = \sqrt{x^2+y^2} \\ \theta = \operatorname{atan}\left(\frac{y}{x}\right) \end{cases} \quad [\operatorname{atan2}(y,x)]$$

⚠ θ undefined at $(0,0)$

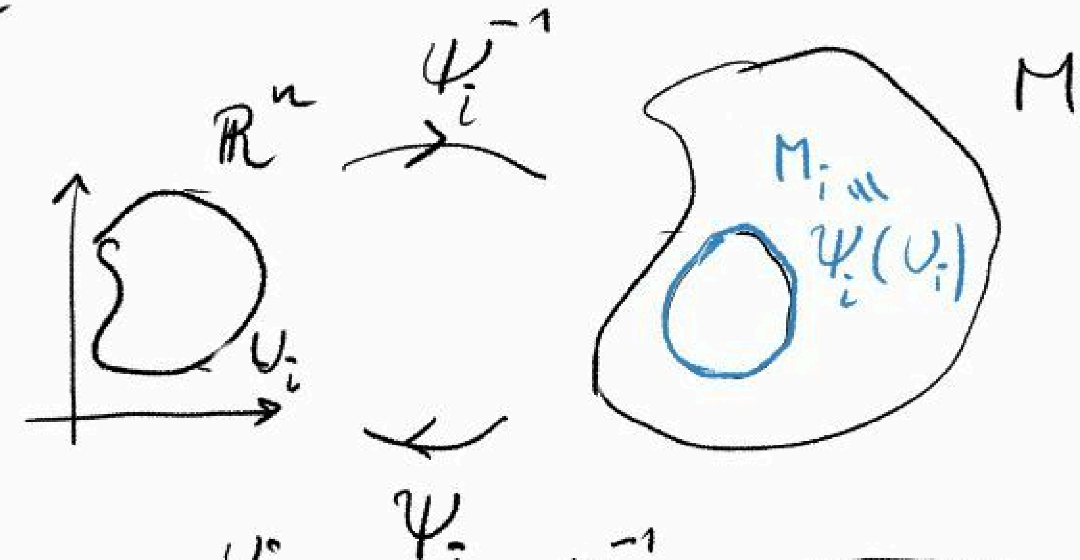


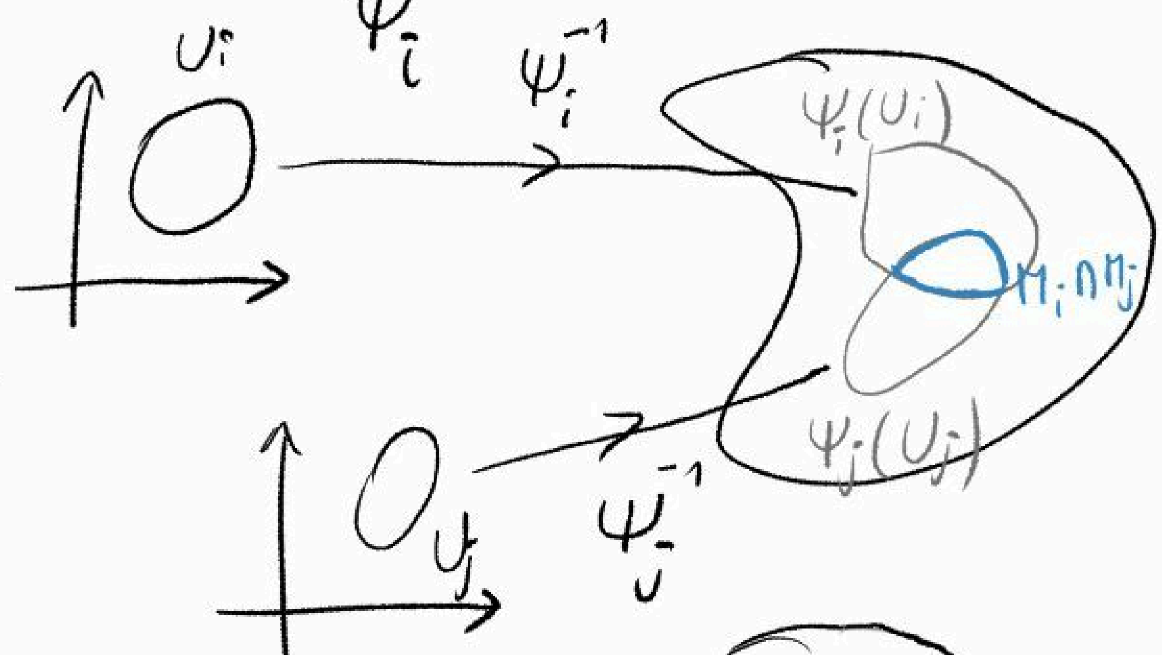
Most coordinate systems are only defined locally

Definition: An n-dimensional manifold M is a collection of points together with a collection of coordinate systems $\{\psi_i : M_i \to U_i\}_{i \in I}$ [Atlas]

[charts]

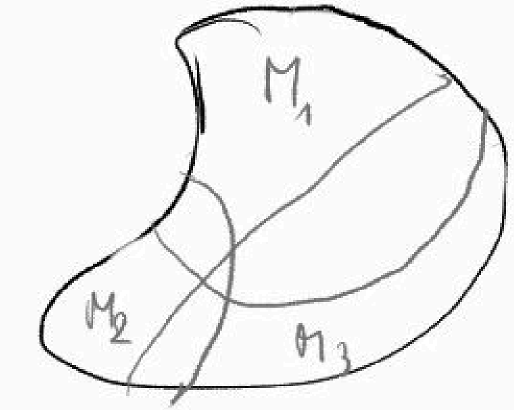
① Each coordinate system ψ_i is a bijective map between $U_i \subset \mathbb{R}^n$ and $M_i \subset M$





② The map $\psi_i \circ \psi_j^{-1}$ is differentiable on $\psi_j(M_i \cap M_j)$

③ Every point $p \in M$ is in some M_i $\simeq \bigcup_{i \in I} M_i = M$



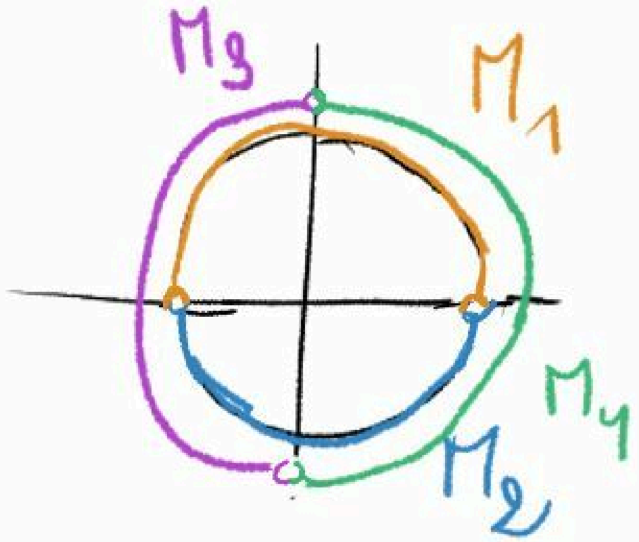
Example Circle $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$

- $\psi_1 : M_1 \to (-1,1)$
  $(x,y) \longmapsto x$

  $\psi_2 : M_2 \longrightarrow (-1,1)$
  $(x,y) \longmapsto x$

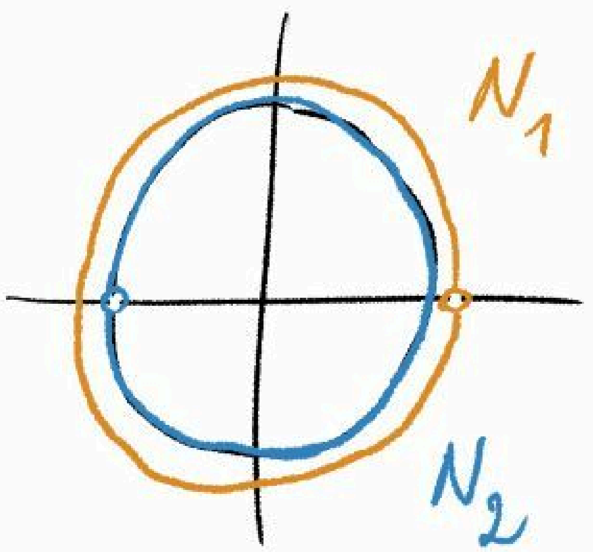
  $\psi_3 : M_3 \to (-1,1)$
  $(x,y) \mapsto y$

  $\psi_4 : M_4 \to (-1,1)$
  $(x,y) \mapsto y$



- $\phi_1 : N_1 \to (-\pi, \pi)$
  $(x,y) \longmapsto \operatorname{atan2}(y,x)$

  $\phi_2 : N_2 \longrightarrow (-\pi, \pi)$
  $(x,y) \longmapsto \operatorname{atan2}(y,-x)$



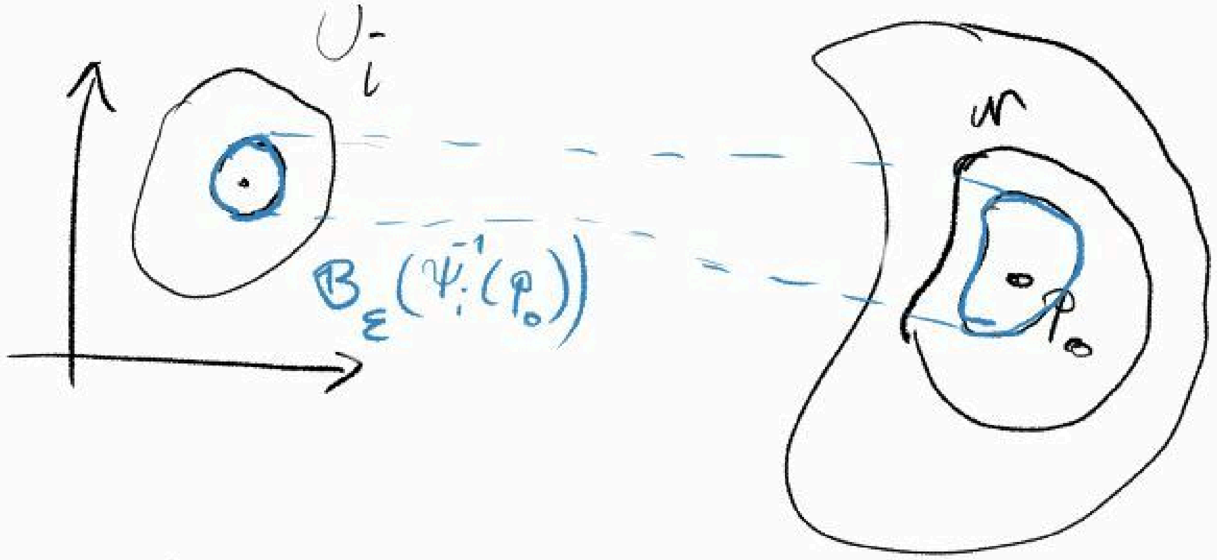
Manifolds are topological spaces

- A neighbourhood of $p_0 \in M$ is a subset $\mathcal{N} \subset M$ s.t

$$\forall p \in M, \quad \|\psi_i^{-1}(p) - \psi_i^{-1}(p_0)\| < \varepsilon \Rightarrow p \in \mathcal{N}$$

for some $\varepsilon > 0$

[NB: ε depends on the chart, but the property of being a neighborhood does not]



- A subset of M is <u>open</u> if it contains a neighbourhood of each of its points

$\longrightarrow$ "open sets" define a topology on M

Product Manifold

let M, N be (m/n)-manifolds, then

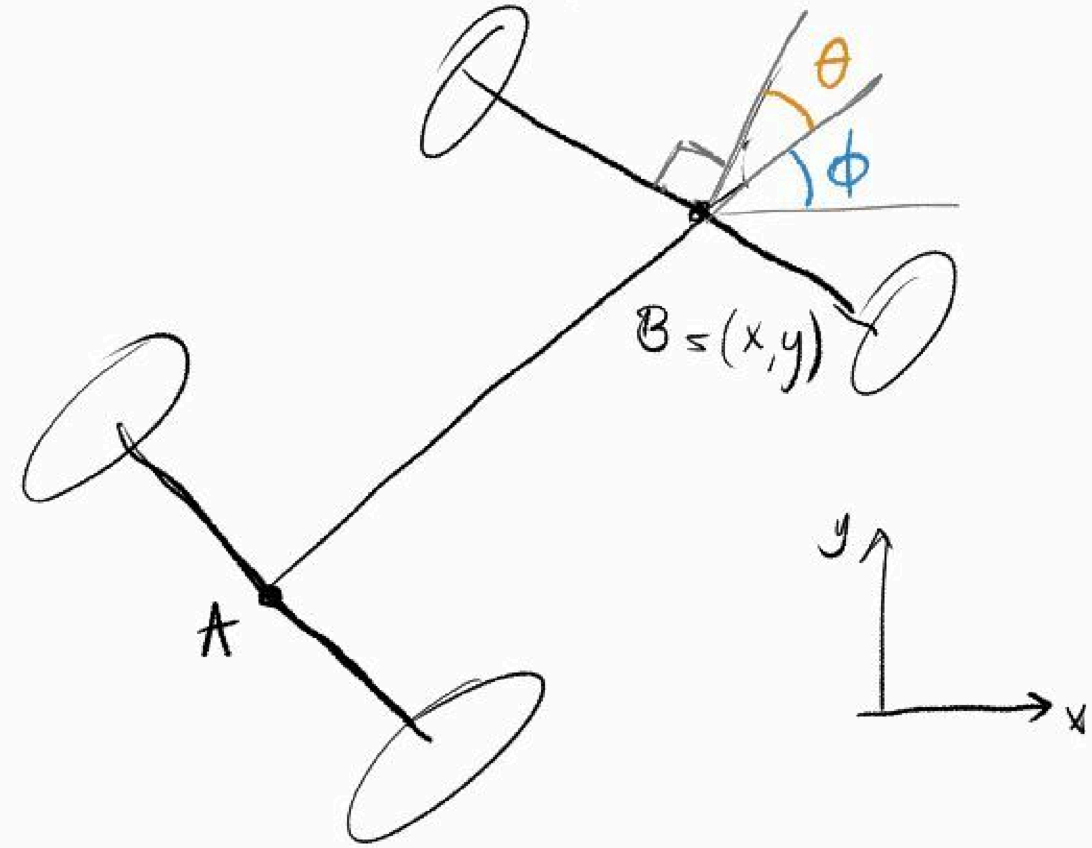
$M \times N = \{(p,q) \mid p \in M,\ q \in N\}$ is a manifold with the coordinate system $(x^1, \dots, x^m, y^1, \dots, y^n)$

with
$$(x^1, \dots, x^m)(p) = \psi(p) \qquad p \in M$$
$$(y^1, \dots, y^n)(q) = \phi(q) \qquad q \in N$$

Example:

$$S^1 \times S^1 \cong T^2$$

Example: Nelson's car



Configuration space of a car

$$M = \mathbb{R}^2 \times T^2$$

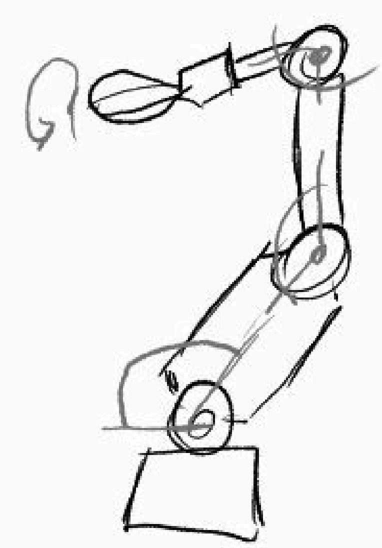
(x, y) position

(θ, ϕ)

θ: heading

ϕ: "absolute heading"

More generally, any kind of industrial robot



Tangent vectors

$$\varphi : [a, b] \to M$$

differentiable curve $\varphi(t)$ on a manifold M

\+ coordinate system (x^i) around $\varphi(t)$

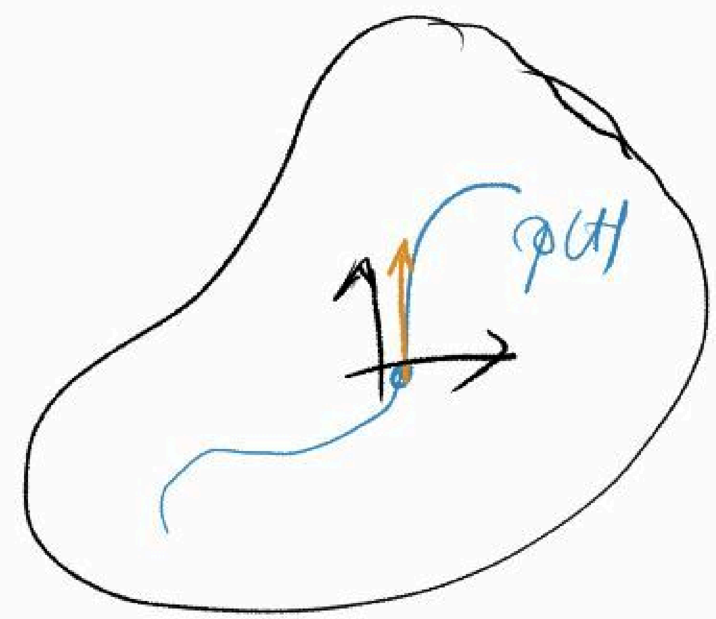
$$\to \quad (\xi^i) = (\dot{x}^i(t)) = \left(\frac{d}{dt} x^i \circ \varphi\right)(t)$$

coordinate tangent vector

change of coordinates: $\tilde{x} = f(x(p))$

$$\tilde{\xi}^j = \left(\frac{\partial \tilde{x}^j}{\partial x^i}\right)_p \xi^i \qquad \circledast$$

↳ Jacobian



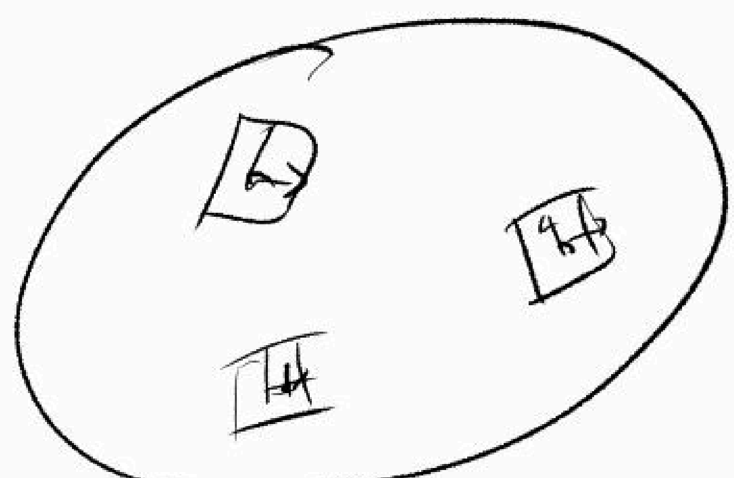
A <u>tangent vector</u> v at p is any object represented by some (ξ^i) which transforms according to $\circledast$

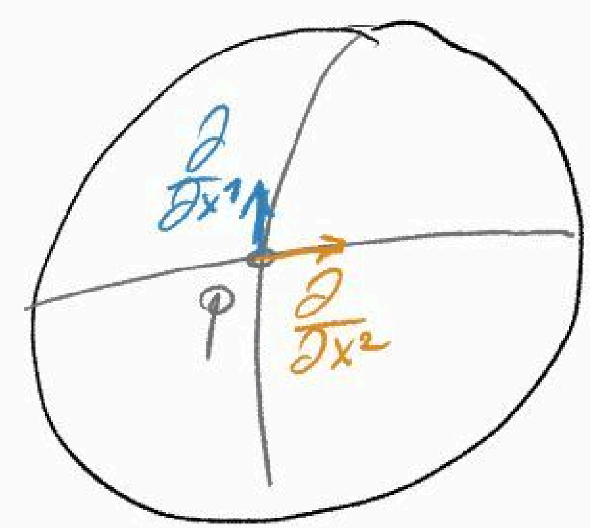
Tangent Space

The set of all vectors around φ is a vector space

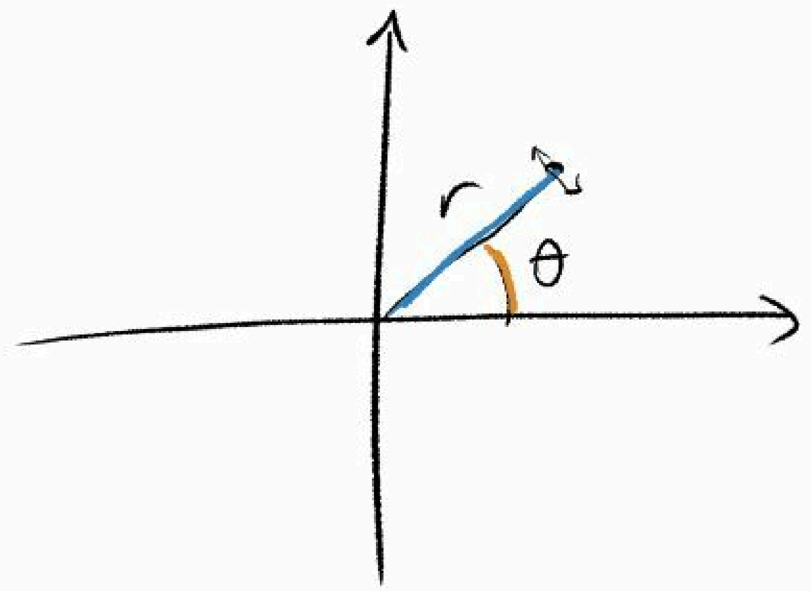
$T_\varphi(M)$ [$\sim$ infinitesimal displacements around φ]

Given a coordinate system (x^i) around φ, we can construct a basis of $T_\varphi M$ $\quad \xi^k = \left(\dfrac{\partial}{\partial x^k}\right)_\varphi$





Example : Polar coordinates



$$(x, y) = (r\cos\theta,\ r\sin\theta)$$

$$r = \sqrt[2]{x^2 + y^2}$$

$$\theta = \operatorname{atan2}(y, x) = \operatorname{atan}\left(\frac{y}{x}\right)$$

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y} = -r\sin\theta \frac{\partial}{\partial x} + r\cos\theta \frac{\partial}{\partial y}$$

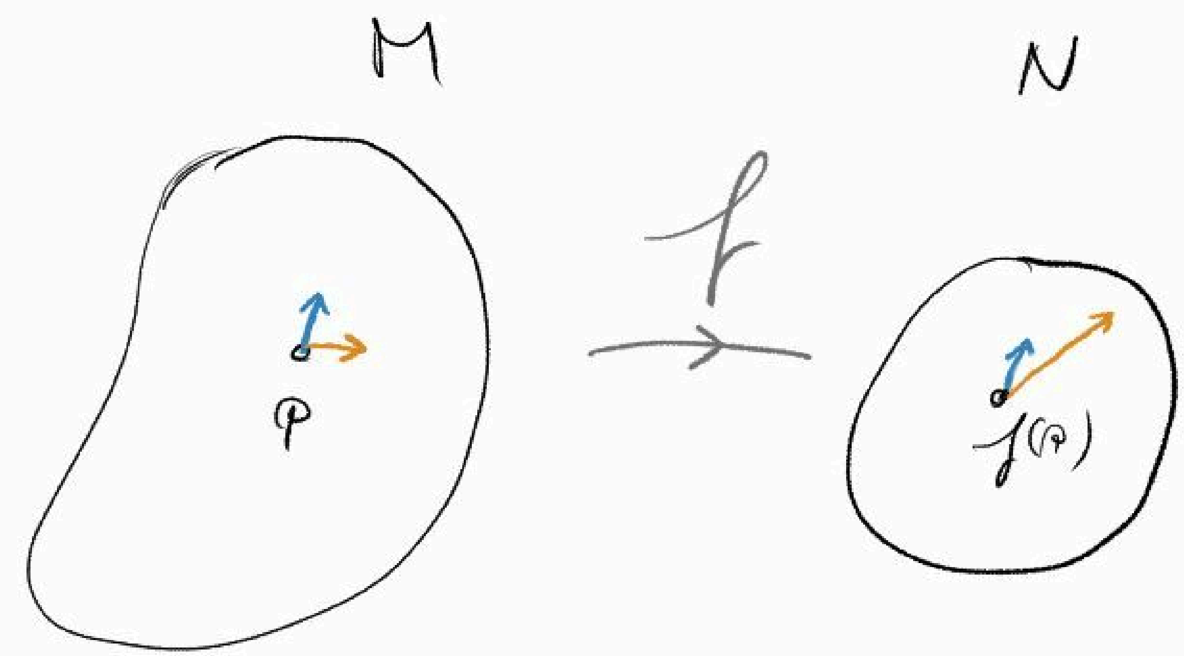
$$\rightsquigarrow \begin{bmatrix} \partial/\partial r \\ \partial/\partial \theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix} \begin{bmatrix} \partial/\partial x \\ \partial/\partial y \end{bmatrix}$$

Differentials

Let $f : M \to N$ be a differentiable map between manifolds

$p \in M$, (x^i) a coordinate system around p

(y^j) coordinates around $f(p)$



f induces a linear map $T_pM \to T_{f(p)}N$

$$\xi^i \frac{\partial}{\partial x^i} \longmapsto \eta^j \frac{\partial}{\partial y^j} \quad \text{with} \quad \eta^j = \underbrace{\left(\frac{\partial y^j}{\partial x^i}\right)}_{\text{differential / Jacobian of } f \text{ at } p} \xi^i = \underline{df_p}\, \xi^i$$

Moreover: this map is independent of the coordinate systems

Covectors & 1-forms

Let $f : M \to \mathbb{R}$ be a differentiable map, then

$df_p : T_pM \to \mathbb{R}$ is a linear form on T_pM

i.e. $df_p \in T_pM^*$ [cotangent space]

In particular, the coordinate differentials dx^i at p satisfy $dx^i \frac{\partial}{\partial x^j} = \delta^i_j$ $\Rightarrow$ $\underline{(dx^i)}$ is the dual basis of $\underline{\frac{\partial}{\partial x^i}}$

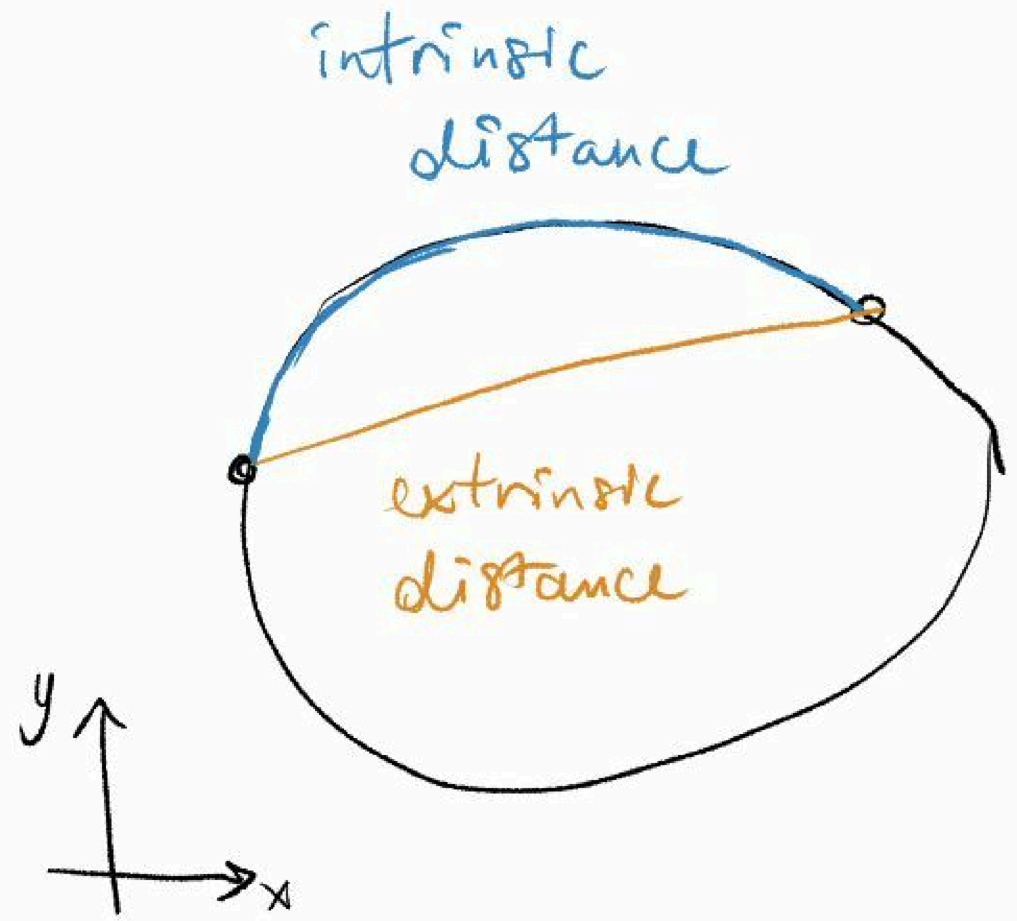
Change of coordinates $\omega = \eta_i dx^i$

$$\tilde{\eta}_i = \underbrace{\left(\frac{\partial x^i}{\partial \tilde{x}^j}\right)_p}_{\text{inverse Jacobian}} \eta_i$$

[covariant]

Riemannian metrics

Since manifolds are usually "embedded" in euclidian space we can inherit a metric on the manifold



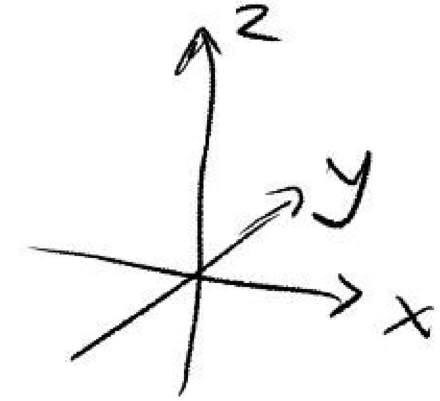
- length of a curve $p : [a,b] \to \mathbb{R}^n$: $\displaystyle\int_a^b \underbrace{\sqrt{\sum_{i=1}^{n}\left(\frac{dx^i}{dt}\right)^2}}_{\text{"length of tangent at curve"}} dt$

  $t \mapsto (x^1, \dots, x^n)$

- Intuition: take the shortest paths between points on the manifold

Examples and notations : $\mathbb{R}^3$

$$\underbrace{ds}_{\text{length element}} = \sqrt{dx^2 + dy^2 + dz^2} \rightsquigarrow ds^2 = dx^2 + dy^2 + dz^2$$



cylindrical coordinates : $ds^2 = dr^2 + r^2 d\theta^2 + dz^2$

spherical coordinates $ds^2 = d\rho^2 + \rho^2 \sin^2\phi \, d\theta^2 + d\phi^2$

Sphere : $ds^2 = R^2 \sin^2\phi \, d\theta^2 + d\phi^2$

Riemann metric

A Riemann metric on M is a map $g : M \to \left((T_p M)^2 \to \mathbb{R} \right)$

which associates to each point of M a non-degenerate symmetric bilinear form on $T_p M$

$$\to \quad ds^2 = g_{ij}\, dx^i dx^j$$

[quadratic form]

$$g_{ij} = g\left(\partial/\partial x^i , \frac{\partial}{\partial x^j} \right)$$

→ Can compute norms of tangent vectors!

N.B. This is a covariant tensor

[also inner-products]

$$\xi^i \frac{\partial}{\partial x^i} \qquad \eta^j \frac{\partial}{\partial x^j} \qquad g_{ij}\, \xi^i \eta^j = \langle \xi , \eta \rangle$$

Distance between two points

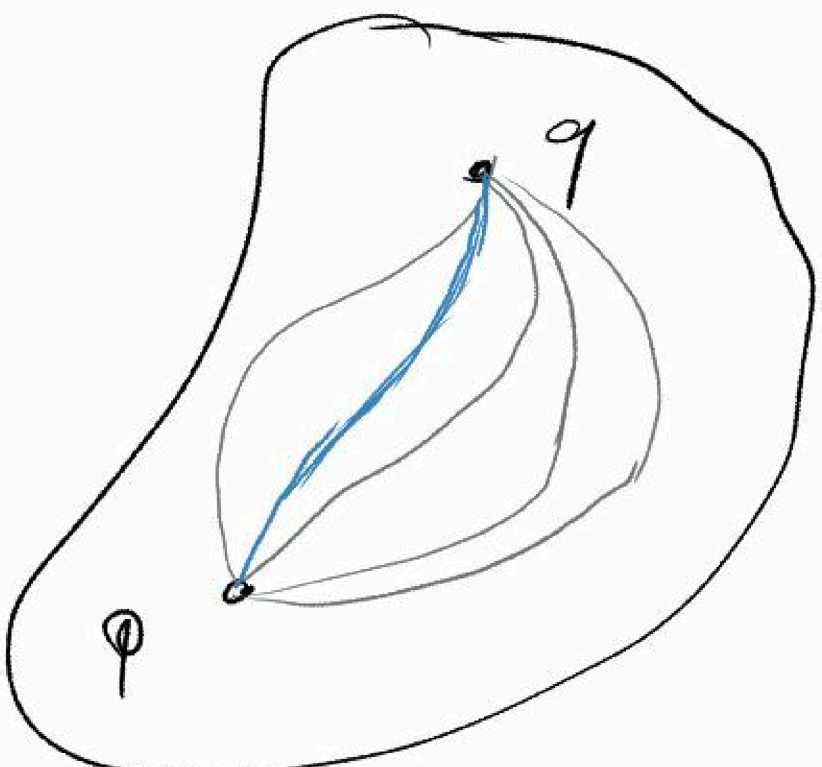
Let M be a manifold, $p, q \in M$

- a <u>path</u> between p and q is a curve $\gamma : [0,1] \to M$

$\gamma(0) = p$
$\gamma(1) = q$

- M is path connected if there is a path between any two points

- length of a path γ: $\int_0^1 L(\gamma, \dot{\gamma})\, dt$



where $L(\gamma, \dot{\gamma}) = \sqrt{g_{ij} \underbrace{\frac{dx^i}{dt} \frac{dx^j}{dt}}}$

$\gamma = (x^1, \ldots, x^n)$

- Distance between p and q $\equiv \inf_{\gamma} \int_0^1 L(\gamma, \dot{\gamma})\, dt$

Tangent vectors

$\varphi : [a, b] \to M$

differentiable curve $\varphi(t)$ on a manifold M

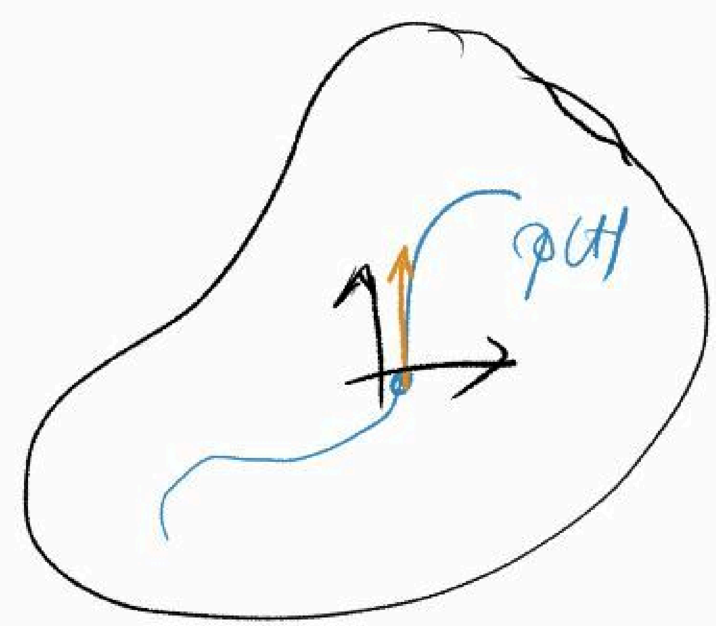
+ coordinate system (x^i) around $\varphi(t)$

$\to \quad (\xi^i) = (\dot{x}^i(t)) = \left(\frac{d}{dt} x^i \circ \varphi\right)(t)$ coordinate tangent vector

change of coordinates: $\tilde{x} = f(x(p))$

$$\tilde{\xi}^j = \left(\frac{\partial \tilde{x}^j}{\partial x^i}\right)_p \xi^i \qquad (*)$$

↳ Jacobian



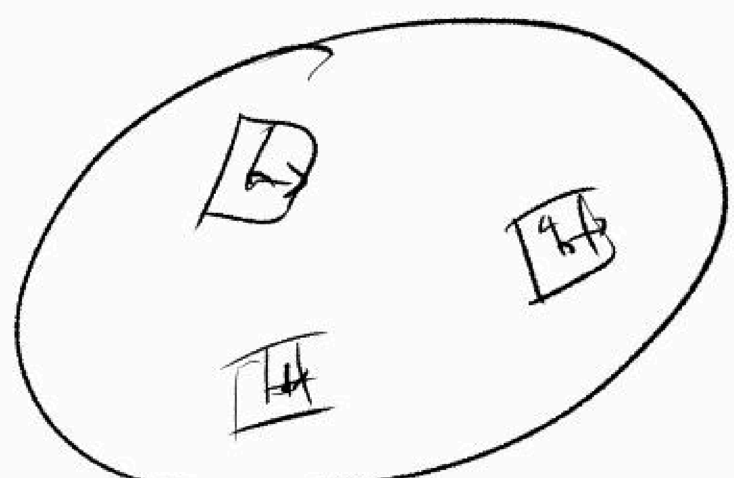
A tangent vector v at p is any object represented by some (ξ^i) which transforms according to $(*)$

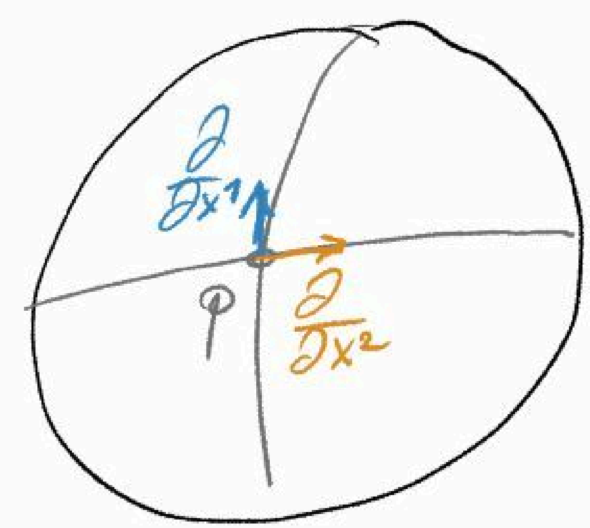
Tangent Space

The set of all vectors around φ is a vector space

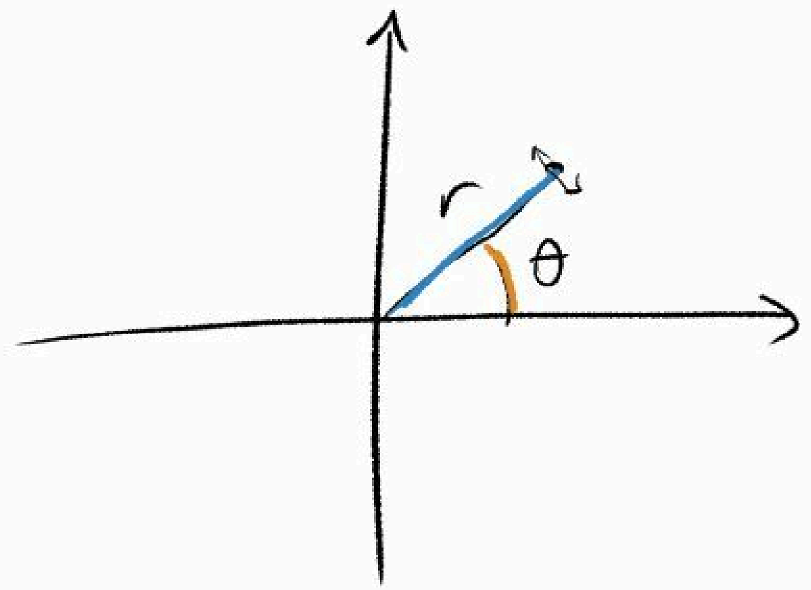
$T_\varphi(M)$ [$\sim$ infinitesimal displacements around φ]

Given a coordinate system (x^i) around φ, we can construct a basis of $T_\varphi M$ $\qquad \xi^k = \left(\frac{\partial}{\partial x^k}\right)_\varphi$





Example : Polar coordinates



$$(x, y) = (r\cos\theta, r\sin\theta)$$

$$r = \sqrt[2]{x^2+y^2}$$

$$\theta = \operatorname{atan2}(y, x) = \operatorname{atan}\left(\frac{y}{x}\right)$$

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y} = -r\sin\theta\frac{\partial}{\partial x} + r\cos\theta\frac{\partial}{\partial y}$$

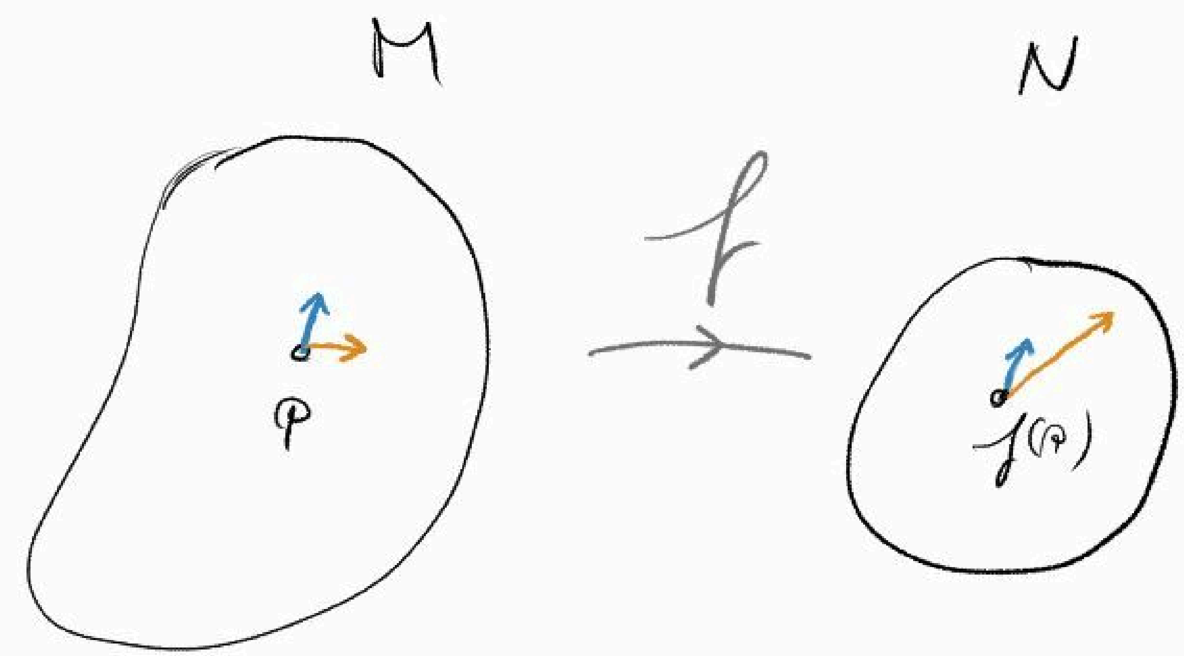
$$\rightsquigarrow \begin{bmatrix} \partial/\partial r \\ \partial/\partial\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix} \begin{bmatrix} \partial/\partial x \\ \partial/\partial y \end{bmatrix}$$

Differentials

Let $f : M \to N$ be a differentiable map between manifolds

$p \in M$, (x^i) a coordinate system around p

(y^j) coordinates around $f(p)$



f induces a linear map $T_pM \to T_{f(p)}N$

$$\xi^i \frac{\partial}{\partial x^i} \longmapsto \eta^j \frac{\partial}{\partial y^j} \quad \text{with} \quad \eta^j = \underbrace{\left(\frac{\partial y^j}{\partial x^i}\right)}_{\text{differential / Jacobian of } f \text{ at } p} \xi^i = \underline{df_p}\, \xi^i$$

Moreover: this map is independent of the coordinate systems

Covectors & 1-forms

Let $f: M \to \mathbb{R}$ be a differentiable map, then

$$df_p : T_pM \to \mathbb{R} \quad \text{is a linear form on } T_pM$$

i.e. $df_p \in T_pM^*$ [cotangent space]

In particular, the coordinate differentials dx^i at p satisfy

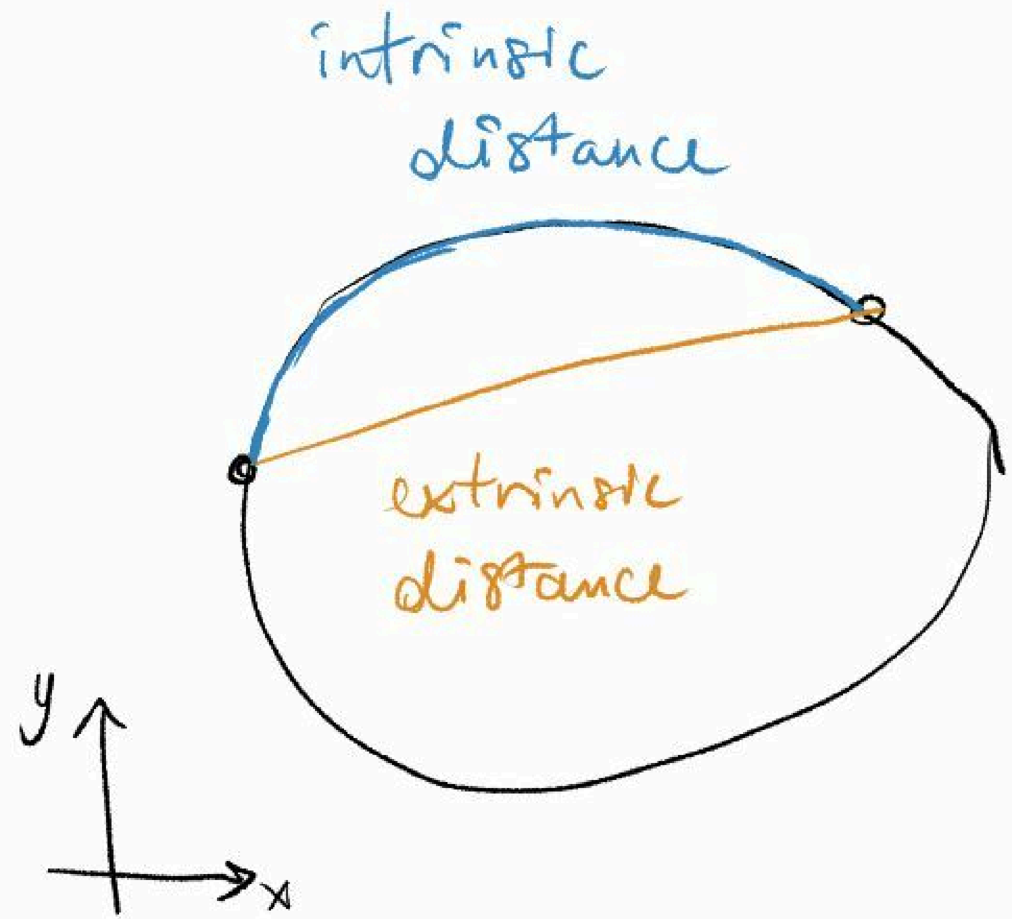
$$dx^i \frac{\partial}{\partial x^j} = \delta^i_j \quad \Rightarrow \underline{(dx^i)} \text{ is the dual basis of } \underline{\frac{\partial}{\partial x^i}}$$

Change of coordinates $\omega = \eta_i \, dx^i$

$$\tilde{\eta}_i = \underbrace{\left(\frac{\partial x^i}{\partial \tilde{x}^j}\right)_p}_{\text{inverse Jacobian}} \eta_i$$

[covariant]

Riemannian metrics



Since manifolds are usually "embedded" in euclidian space we can inherit a <u>metric</u> on the manifold

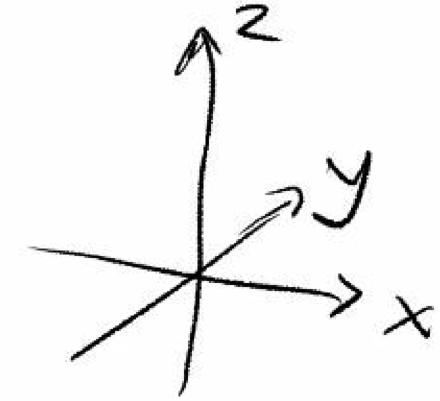
- length of a curve $p: [a,b] \to \mathbb{R}^n$: $\displaystyle\int_a^b \underbrace{\sqrt{\sum_{i=1}^{n} \left(\frac{dx^i}{dt}\right)^2}}_{\text{"length of tangent at curve"}} \, dt$

  $t \mapsto (x^1, \dots, x^n)$

- Intuition: take the <u>shortest</u> paths between points on the manifold

Examples and notations : $\mathbb{R}^3$

$$\underbrace{ds}_{\text{length element}} = \sqrt{dx^2+dy^2+dz^2} \rightsquigarrow ds^2 = dx^2+dy^2+dz^2$$



cylindrical coordinates : $ds^2 = dr^2 + r^2 d\theta^2 + dz^2$

spherical coordinates $ds^2 = d\rho^2 + \rho^2 \sin^2\phi \, d\theta^2 + d\phi^2$

Sphere : $ds^2 = R^2 \sin^2\phi \, d\theta^2 + d\phi^2$

Riemann metric

A Riemann metric on M is a map $g : M \to \left((T_p M)^2 \to \mathbb{R} \right)$

which associates to each point of M a non-degenerate symmetric bilinear form on $T_p M$

$\to \quad ds^2 = g_{ij}\, dx^i dx^j$ [quadratic form]

$g_{ij} = g\left(\partial/\partial x^i , \frac{\partial}{\partial x^j} \right)$

$\to$ Can compute norms of tangent vectors!

NB. This is a covariant tensor

[also inner-products]

$\xi^i \frac{\partial}{\partial x^i} \qquad \eta^j \frac{\partial}{\partial x^j} \qquad g_{ij}\, \xi^i \eta^j = \langle \xi, \eta \rangle$

① Distance between two points

Let M be a manifold, $p, q \in M$

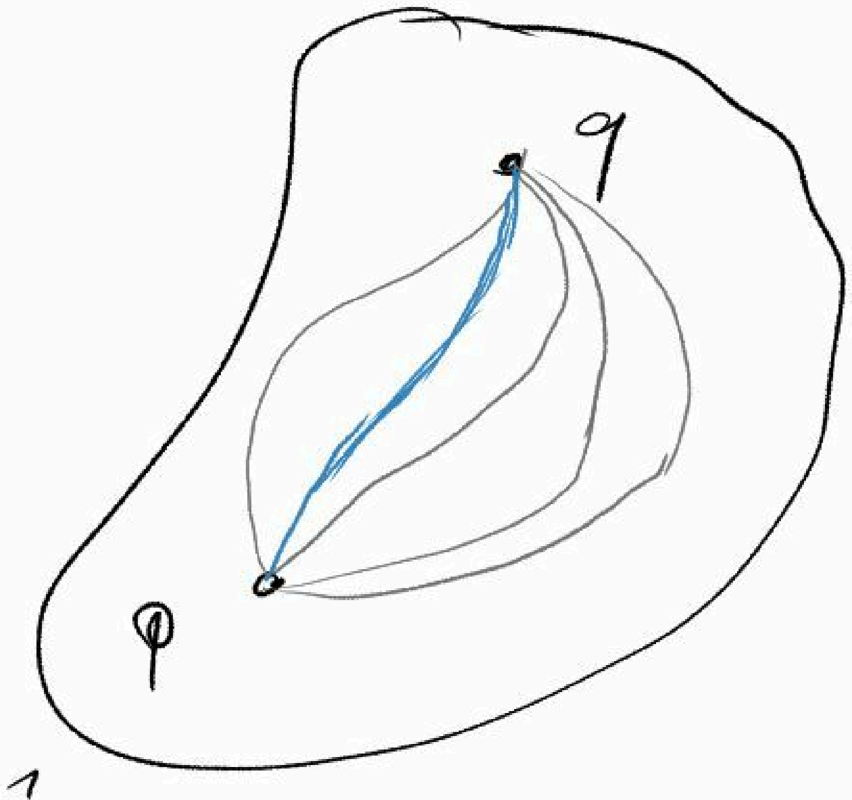
- a **path** between p and q is a curve $\gamma : [0,1] \to M$
$$\gamma(0) = p$$
$$\gamma(1) = q$$

- M is path connected if there is a path between any two points

- length of a path γ: $\displaystyle\int_0^1 L(\gamma, \dot\gamma)\, dt$

where $L(\gamma, \dot\gamma) = \sqrt{g_{ij} \underbrace{\frac{dx^i}{dt} \frac{dx^j}{dt}}}$

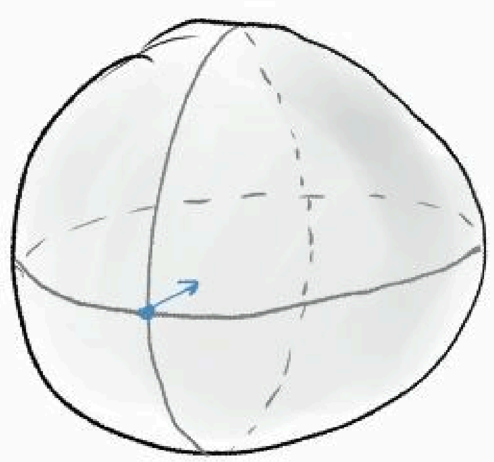
$\gamma = (x^1, \dots, x^n)$

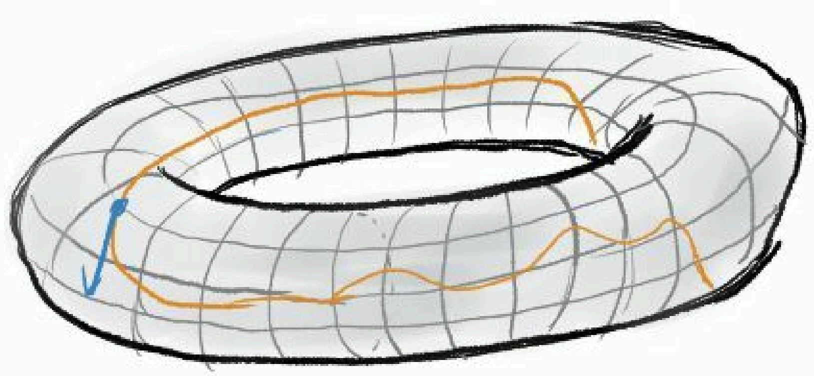


- Distance between p and q $=$ $\displaystyle\inf_{\gamma} \int_0^1 L(\gamma, \dot\gamma)\, dt$

EleutherAI Math Reading Group

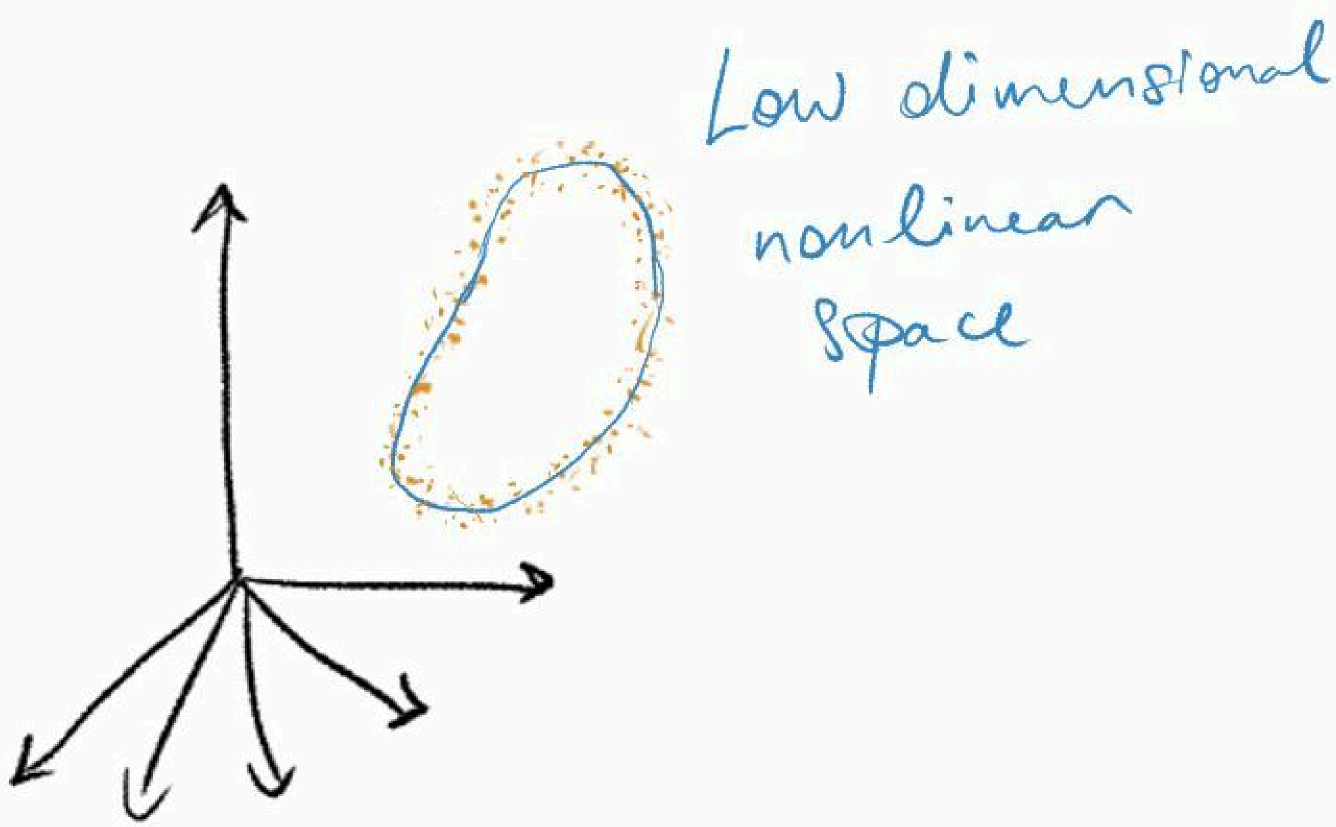
Differential Geometry I: Manifolds





Motivation

Manifold Hypothesis



High-dimensional ambient space

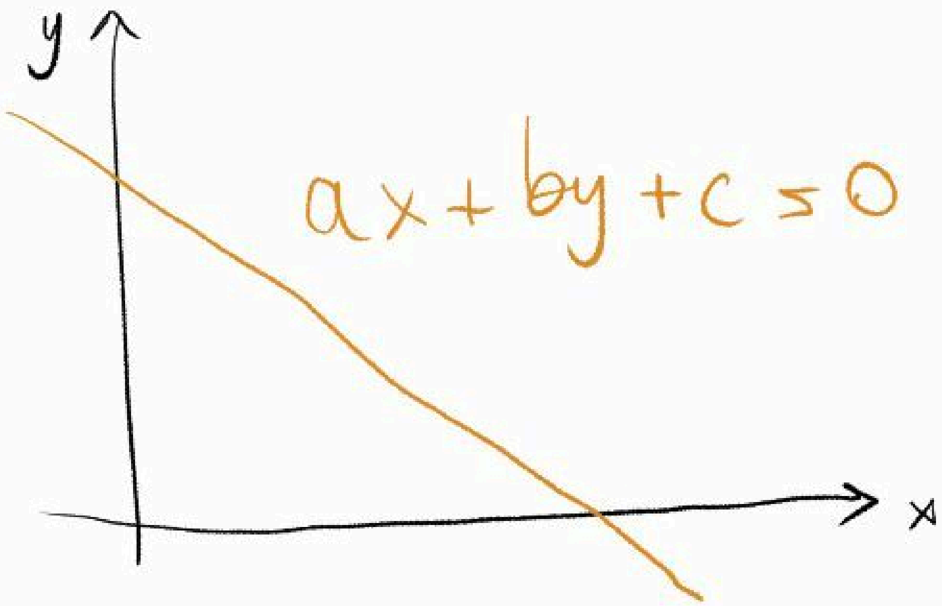
Latent coordinates

e.g. Autoencoders, GANs, ...

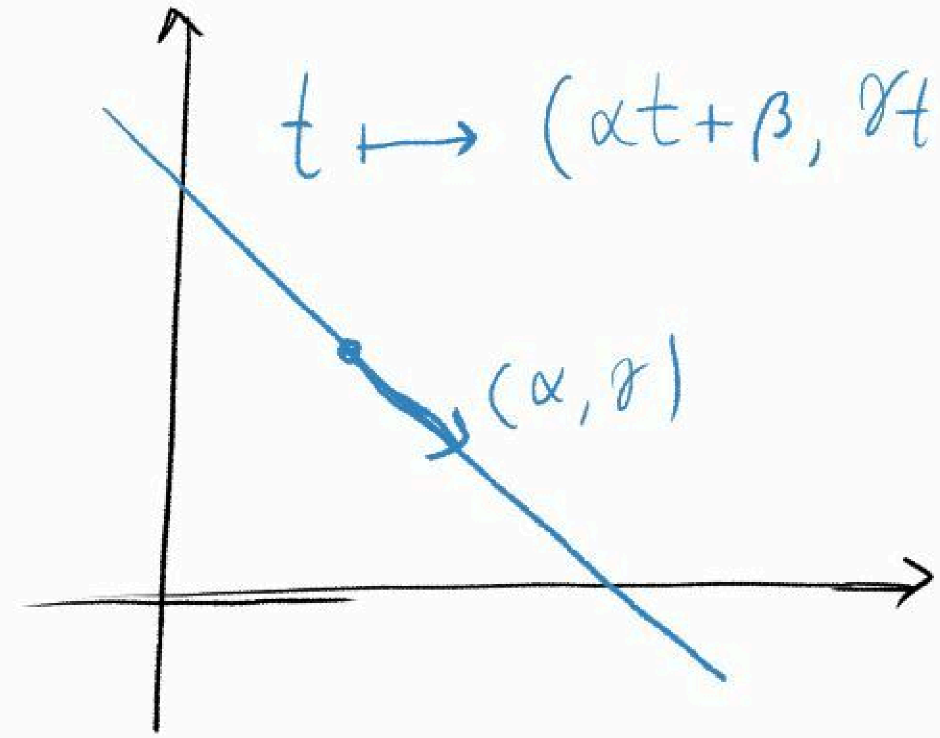
↳ Space that we can move around in

"Locally linear"

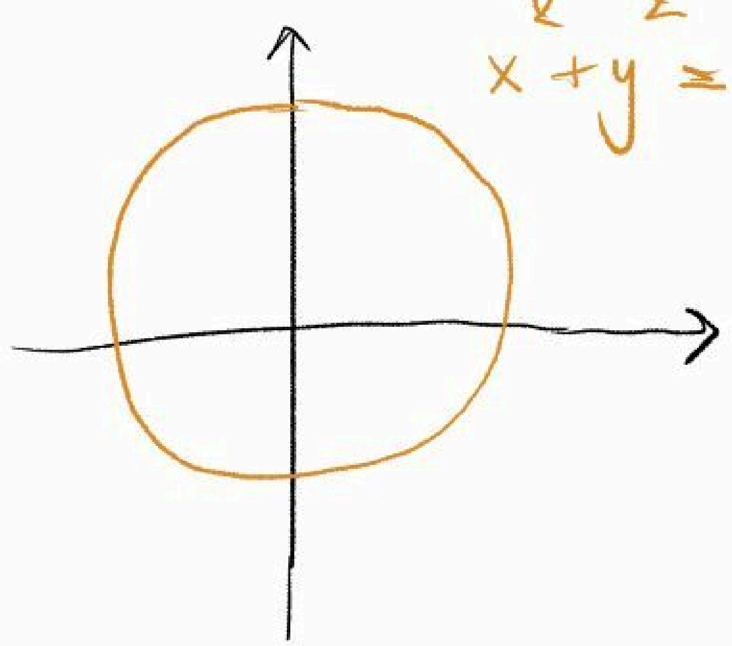
Implicit vs Explicit equations



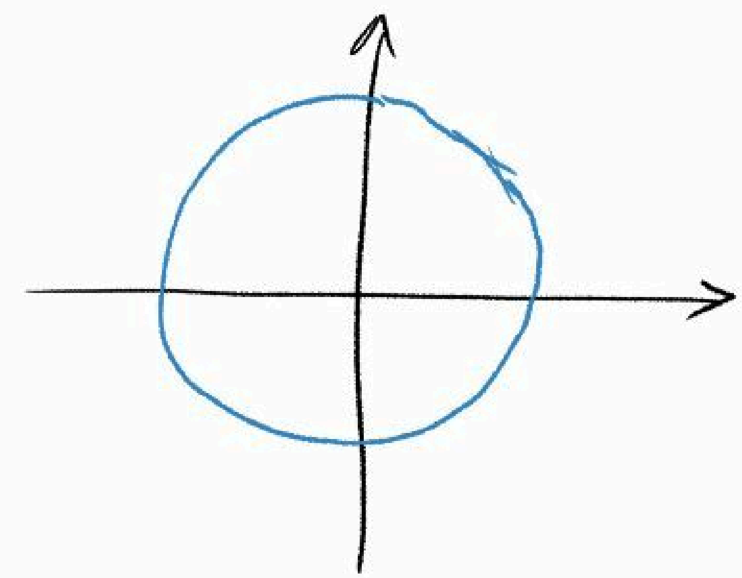
$ax + by + c = 0$



$t \mapsto (\alpha t + \beta, \gamma t + \delta)$

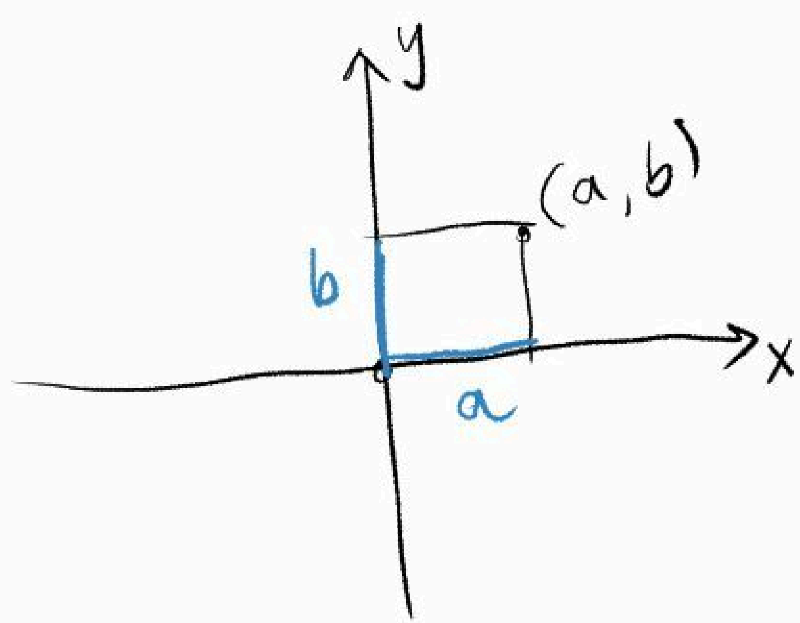


$x^2 + y^2 = r^2$

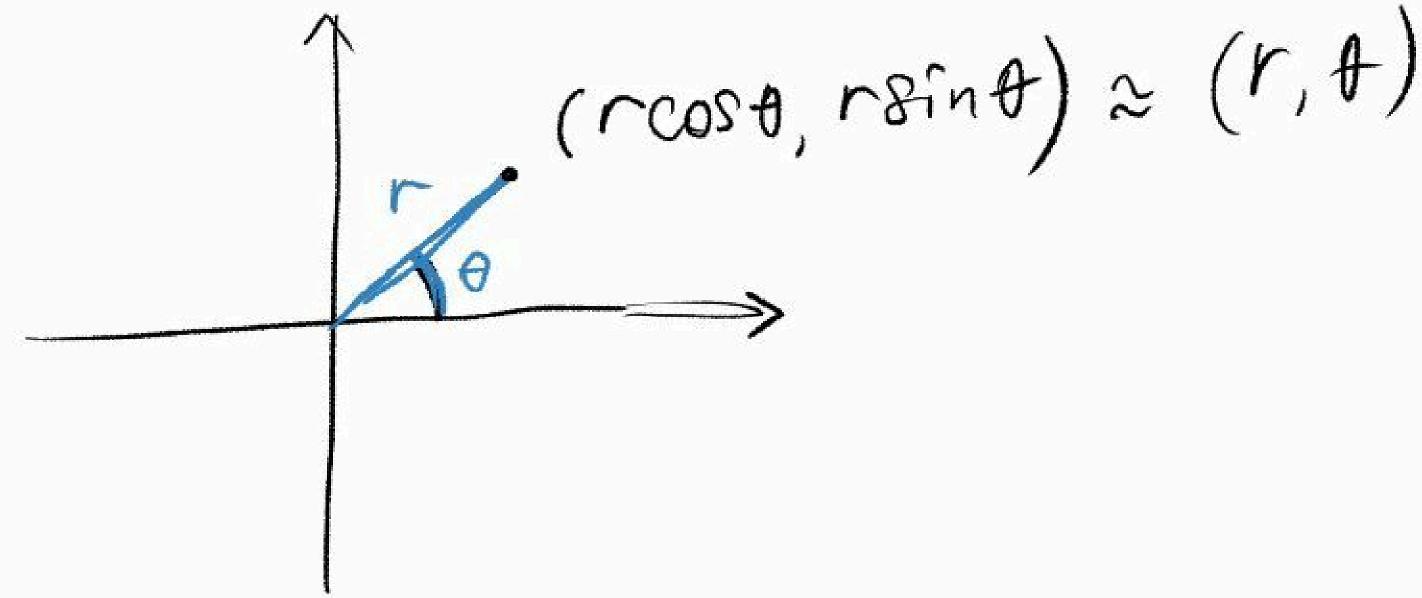


$t \mapsto (r\cos t, r\sin t)$

Coordinate Systems



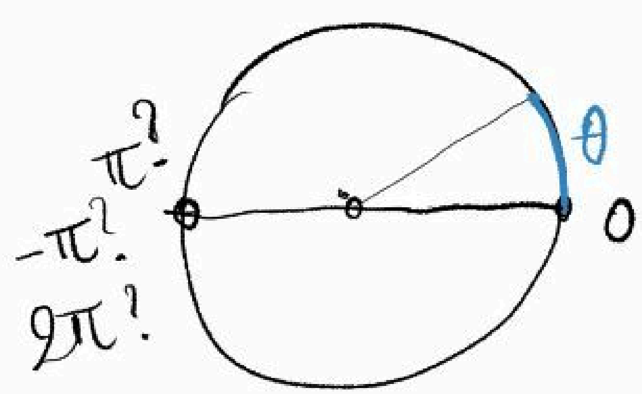
Cartesian coordinates



Polar coordinates

$$\begin{cases} r = \sqrt{x^2+y^2} \\ \theta = \operatorname{atan}\left(\frac{y}{x}\right) \end{cases} \quad [\operatorname{atan2}(y,x)]$$

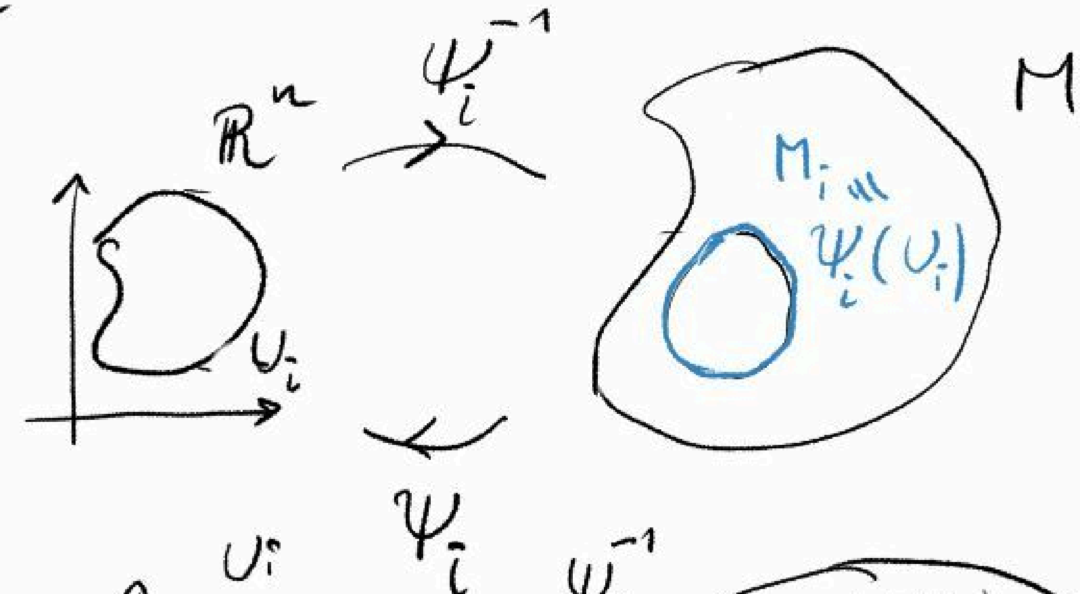
⚠ θ undefined at $(0,0)$

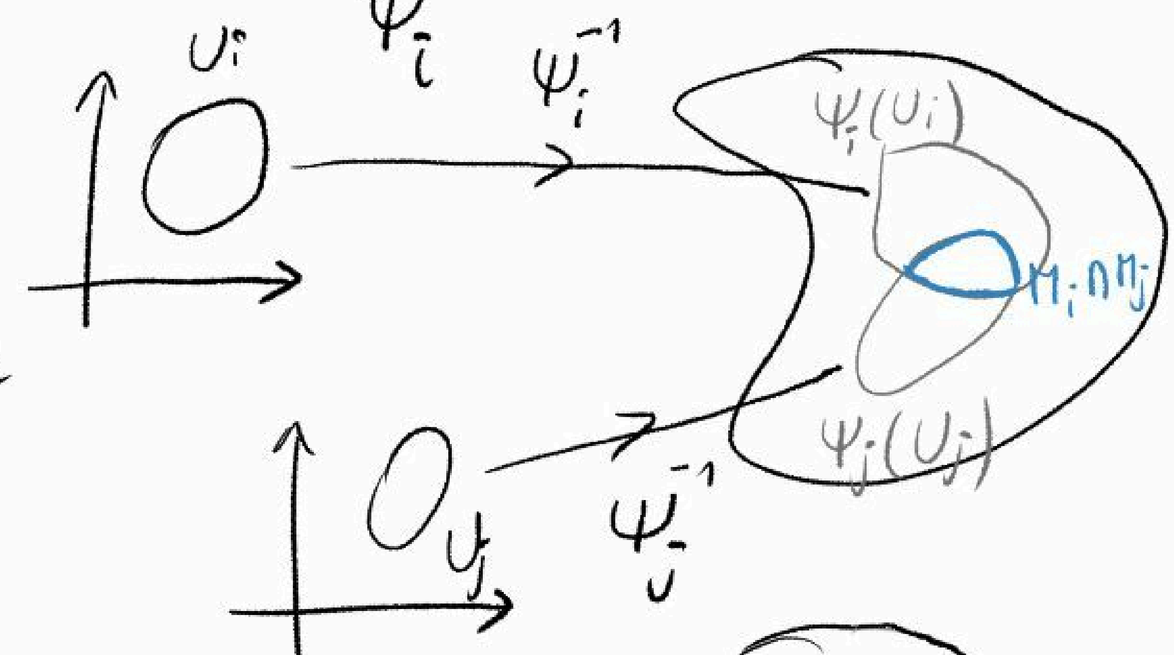


Most coordinate systems are only defined <u>locally</u>

① <u>Definition</u>: An n-dimensional manifold M is a collection of points together with a collection of coordinate systems $\{\psi_i : M_i \to U_i\}_{i \in I}$ [Atlas]
[charts]

① Each coordinate system ψ_i is a bijective map between $U_i \subset \mathbb{R}^n$ and $M_i \subset M$

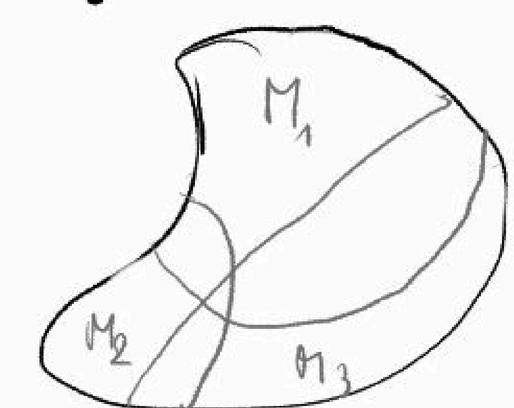




② The map $\psi_i \circ \psi_j^{-1}$ is differentiable on $\psi_j(M_i \cap M_j)$

③ Every point $p \in M$ is in some M_i

$\simeq \bigcup_{i \in I} M_i = M$



Example

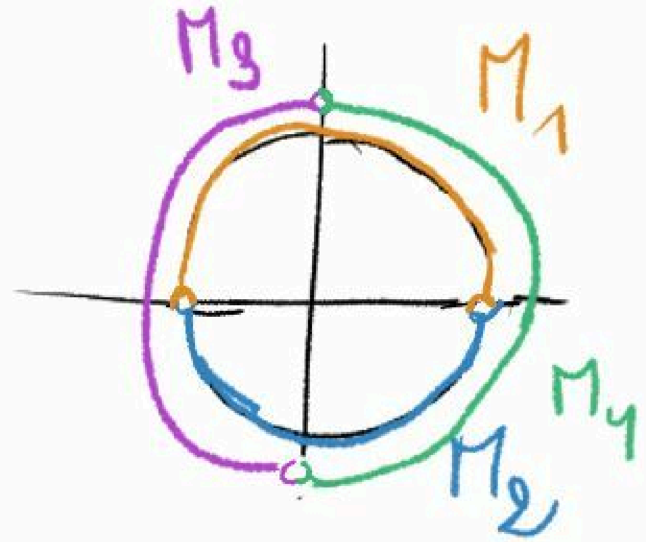
Circle $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$

- $\psi_1 : M_1 \to (-1, 1)$
  $(x,y) \longmapsto x$

  $\psi_2 : M_2 \to (-1, 1)$
  $(x,y) \longmapsto x$

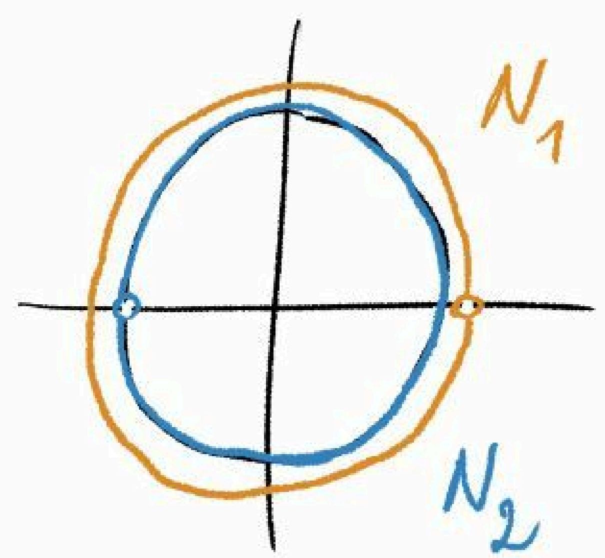
  $\psi_3 : M_3 \to (-1, 1)$
  $(x,y) \longmapsto y$

  $\psi_4 : M_4 \to (-1, 1)$
  $(x,y) \longmapsto y$



- $\phi_1 : N_1 \to (-\pi, \pi)$
  $(x,y) \longmapsto \mathrm{atan2}(y, x)$

  $\phi_2 : N_2 \to (-\pi, \pi)$
  $(x,y) \longmapsto \mathrm{atan2}(y, -x)$



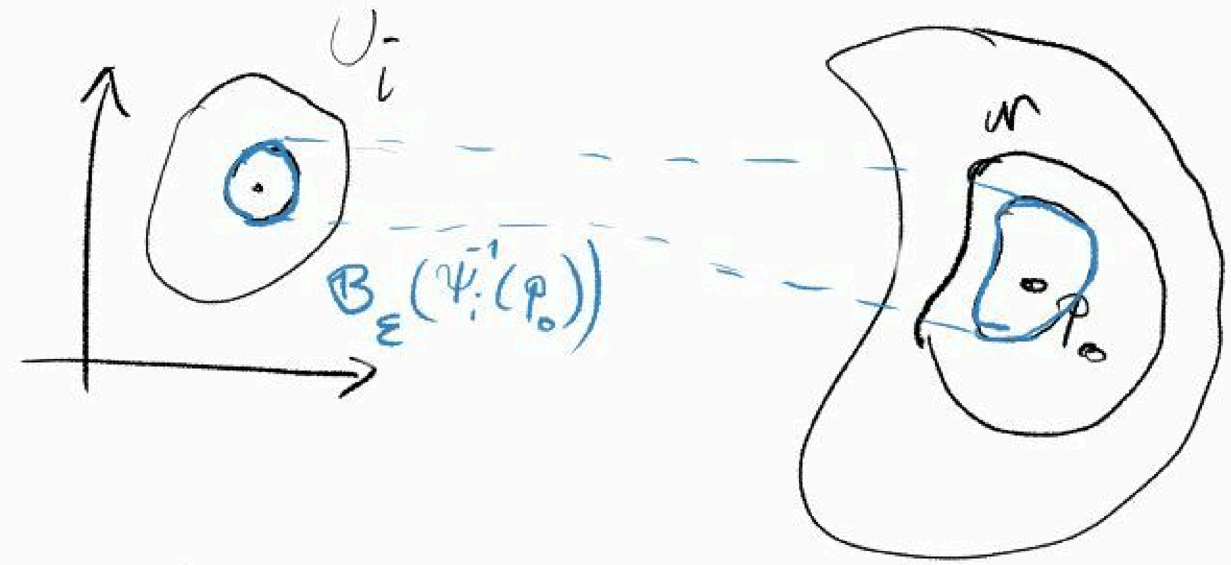
Manifolds are topological spaces

- A neighbourhood of $p_0 \in M$ is a subset $N \subset M$ s.t

$$\forall p \in M, \quad \|\psi_i^{-1}(p) - \psi_i^{-1}(p_0)\| < \varepsilon \Rightarrow p \in N$$

for some $\varepsilon > 0$

[NB: ε depends on the chart, but the property of being a neighborhood does not]



- A subset of M is <u>open</u> if it contains a neighbourhood of each of its points

$\longrightarrow$ "open sets" define a topology on M

Product Manifold

let M, N be (m/n)-manifolds, then

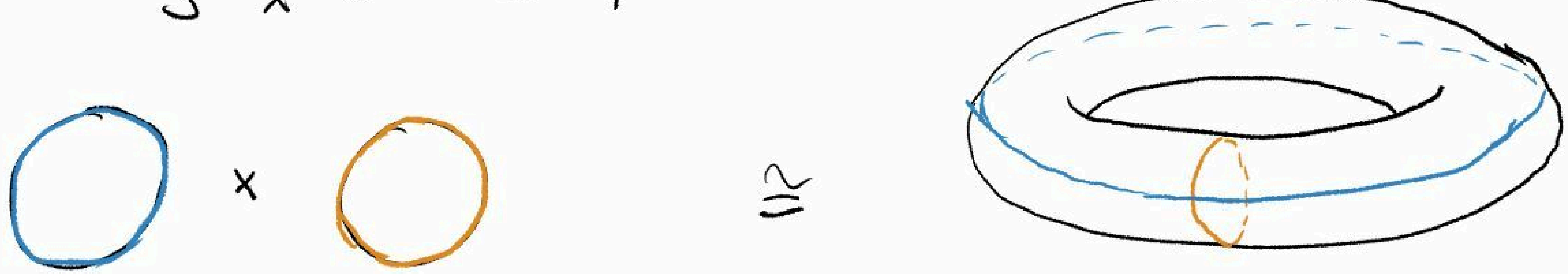
$M \times N = \{(p,q) \mid p \in M,\ q \in N\}$ is a manifold with the coordinate system $(x^1, \dots, x^m, y^1, \dots, y^n)$

with $(x^1, \dots, x^m)(p) = \psi(p) \qquad p \in M$

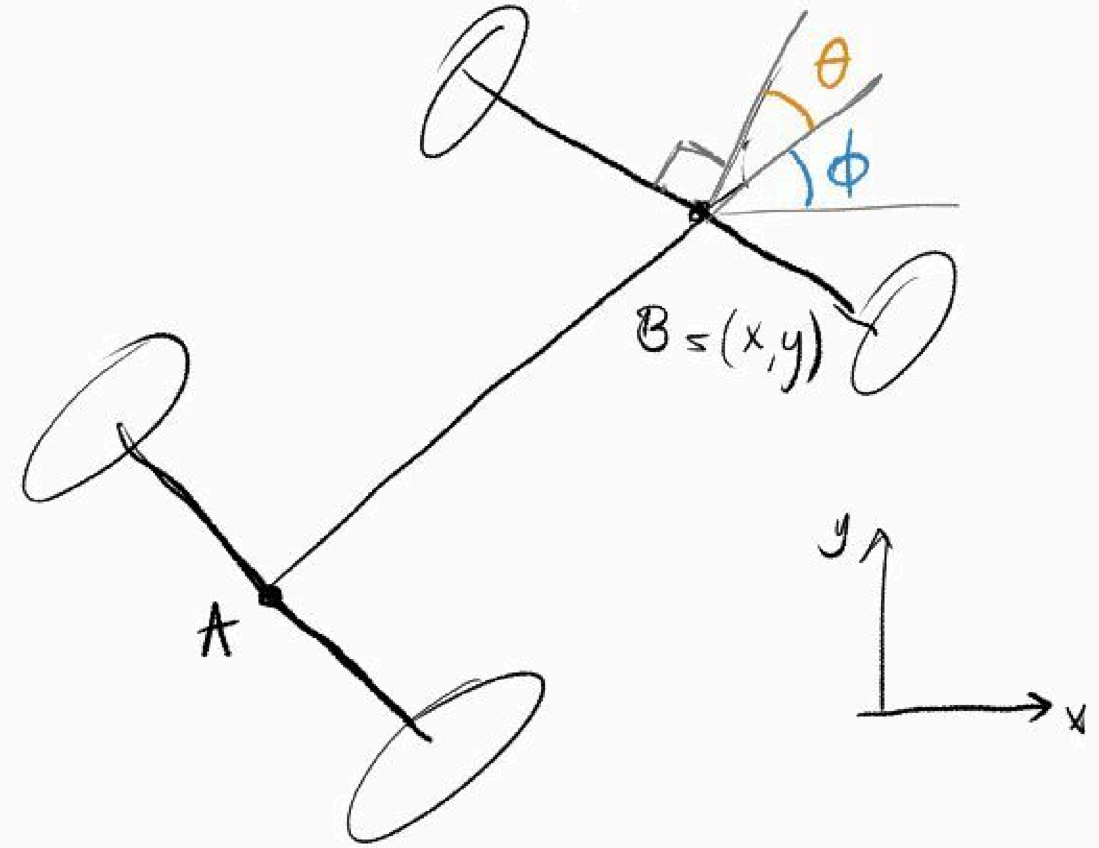
$(y^1, \dots, y^n)(q) = \phi(q) \qquad q \in N$

Example:

$$S^1 \times S^1 \cong T^2$$



Example: Nelson's car



Configuration space of a car

$$M = \mathbb{R}^2 \times T^2$$

(x,y) position

(θ, ϕ)

heading

"absolute heading"

More generally, any kind of industrial robot