BGSM/CRM

# AL&DNN

# Differential Geometry

## Background notions

S. Xambó

UPC & IMTech

2/11/2021

**Topics**. Group theory and differential geometry basics. Differential manifolds. Lie groups and Lie algebras.

References:
[1] (bronstein-bruna-cohen-velickovic-2021)
[2] (cohen-2021)
[3] (gallier-quaintance-2020)
[4] (xambo-2018)
[5] (lee-2013)
[6] (carne-2012)

Computations in *manifold learning*:
[7] (smirnov-2021)

# Topology

**Basic notions**
**Example: stereographic projection**
**Homotopies**
**Poincaré's fundamental group**
**Simply connected spaces**

With the exception of *projective spaces* and *Grassmannians*, to be introduced later, for our purposes we only need to consider *topological spaces* $X$ *that are subsets of some* $\mathbf{R}^n$ (which will simply be called *spaces*).

The topology of any such space $X \subseteq E$ is the topology induced by the standard topology of $E$ containing it, which is the topology induced by any *Euclidean norm* $\|x\|$ on $E$.

Thus an *open set* of $X \subseteq \mathbf{R}^n$ is any subset $U$ of $X$ of the form $U = V \cap X$, where $V$ is open in $\mathbf{R}^n$. The *closed sets* of $X$ are the complements of open sets.

A map $f : X \to X'$ between spaces is said to be *continuous* if $f^{-1}U'$ is an open set of $X$ for any open set $U'$ of $X'$. It is immediate to check that the composition of continuous maps is continuous. If $f$ is bijective and $f^{-1}$ is also continuous, we say that $f$ is a *homeomorphism*. This is equivalent to say that $U \subset X$ is open in $X$ if and only if $f(U)$ is open in $X'$.

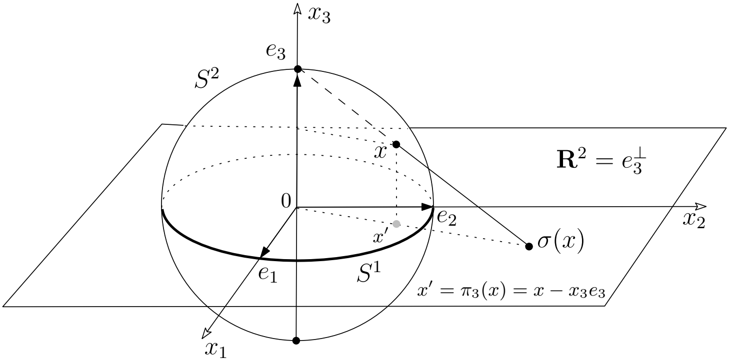


Figure 1.1: Stereographic projection of $S^2 - \{e_3\}$ to $\mathbf{R}^2$ from $e_3$. Analytically, $\sigma(x) = \lambda x'$, where $x' = x - x_3e_3$ (the ortogonal projecton of $x$ to $\mathbf{R}^2$) and $\lambda = 1/(1-x_3)$. Indeed, $\sigma(x) = e_3 + \lambda(x - e_3)$, for some $\lambda \in \mathbf{R}$, $\lambda \neq 0$, and $0 = e_3 \cdot \sigma(x) = 1 + \lambda(x_3 - 1)$. So $\lambda = 1/(1-x_3)$ and $e_3 + \lambda(x - e_3) = (x - x_3e_3)/(1-x_3)$ This map is defined, and is continuous, for all $x \in \mathbf{R}^3 - \{x_3 = 1\}$.

In general, consider the sphere $S^{n-1}$ of radius 1 in $\mathbf{R}^n$:

$$S^{n-1} = \{x \in \mathbf{R}^n \,|\, x^2 = 1\}.$$

Then $e_n \in S^{n-1}$ and the *stereographic projection* from $e_n$ is the map

$$\sigma : S^{n-1} - \{e_n\} \to \mathbf{R}^{n-1} = e_n^{\perp},$$

defined by requiring that $\sigma(x) \in \mathbf{R}^{n-1}$ be aligned with $e_n$ and $x$. By the same argument as for $n = 3$ we conclude that $\sigma(x) = (x - x_n e_n)/(1 - x_n)$, also defined and continuous for all $x \in \mathbf{R}^n - \{x_n = 1\}$.

The expression of the inverse map $\sigma^{-1} : \mathbf{R}^{n-1} \to S^{n-1} - \{e_n\}$ is

$$\sigma^{-1}(y) = \frac{2}{y^2+1} y + \frac{y^2-1}{y^2+1} e_n,$$

as this point is in the line joining $e_3$ and $y$ and belongs to $S^{n-1}$:

$$\sigma^{-1}(y)^2 = \frac{4y^2}{(y^2+1)^2} + \frac{(y^2-1)^2}{(y^2+1)^2} = 1.$$

Two continuous maps $f, g : X \to X'$ are said to be *homotopic*, and we write $f \simeq g$ to denote it, if there is a continuous map $H : I \times X \to X'$, where $I = [0, 1] \subset \mathbf{R}$, such that

$$H(0, x) = f(x) \text{ and } H(1, x) = g(x) \text{ forall } x \in X.$$

To see that this expresses the idea of *continuous deformation of $f$ into $g$* (or *homotopy*), consider the maps $h_s : X \to X'$, $s \in I$, defined by $h_s(x) = H(s, x)$. This is a continuously varying family $\{h_s\}_{s \in I}$ of continuous maps $h_s : X \to X'$ and by definition we have $h_0 = f$ and $h_1 = g$. The homotopy relation $\simeq$ turns out to be an *equivalence relation* in the set of continuous maps $X \to X'$, and the *homotopy class* of $f$, consisting of all continuous maps $X \to X'$ that are homotopic to $f$, is denoted by $[f]$.

Given a space $X$ and a point $x_0 \in X$, the elements of the *fundamental group* of $X$ with *base point* $x_0$, which is denoted by $\pi_1(X, x_0)$, are the homotopy classes $[\gamma]$ of *loops* on $X$ with base point $x_0$, by which we mean continuous maps $\gamma : I \to X$ such that $\gamma(0) = \gamma(1) = x_0$.

In this case, a homotopy $H : I \times I \to X$ is required to satisfy $H(s, 0) = x_0 = H(s, 1)$ for all $s \in I$, which means that all the paths $\gamma_s(t) = H(s, t)$ have to be loops on $X$ at $x_0$ (*loop homotopy*).

The group operation is defined by the rule$[\gamma][\gamma'] = [\gamma * \gamma']$, where $\gamma * \gamma'$ is the loop defined by

$$(\gamma * \gamma')(t) = \begin{cases} \gamma(2t) & \text{for} \quad 0 \leqslant t \leqslant \frac{1}{2}, \\ \gamma'(2t-1) & \text{for} \quad \frac{1}{2} \leqslant t \leqslant 1. \end{cases}$$

Note that this loop travels the whole loop $\gamma$ for $t \in [0, \frac{1}{2}]$ followed by traveling the whole loop $\gamma'$ for$t \in [\frac{1}{2}, 1]$. The composition $\gamma * \gamma'$ is not associative, but it becomes so at the level of homotopy classes.

Similarly, the constant loop $e : I \to X$, $e(t) = x_0$ for all $t$, is not a neutral element for the composition, but it is so for homotopy classes, namely $[e][\gamma] = [\gamma][e] = [\gamma]$; and the inverse loop $\gamma^{-1}$ defined by traveling $\gamma$ backwards, $\gamma^{-1}[t] = \gamma[1-t]$, satisfies $[\gamma][\gamma^{-1}] = [\gamma^{-1}][\gamma] = [e]$ although $\gamma * \gamma^{-1} \neq e$.

A continuous map $f : X \to X'$ induces a group homomorphism

$$\tilde{f} : \pi_1(X, x_0) \to \pi_1(X', x_0'), \text{ where } x_0' = f(x_0).$$

Actually if $\gamma$ is a loop on $X$ at $x_0$, then $\gamma' = f \circ \gamma$ is a loop on $X'$ at $x_0'$ and the homomorphism is defined by $[\gamma] \mapsto [\gamma']$. In particular we see that if $f$ is a homeomorphism, then $\tilde{f}$ is an isomorphism.

If $x_0, x_0' \in X$ are connected by a path $\delta$, then the map $\pi_1(X, x_0') \to \pi_1(X, x_0)$, $[\gamma] \mapsto [\delta][\gamma][\delta^{-1}]$ is an isomorphism of groups, with inverse the analogous map for $\delta^{-1}$.

In particular we see that for *path-connected spaces* the isomorphism class of $\pi_1(X, x_0)$ is the same for all points $x_0$. In such cases, we may simply write $\pi_1(X)$ to denote that isomorphism class.

This is especially apt when $X$ has some distinguished point, and of course also when $\pi_1(X) \simeq \{0\}$.

The space $X$ is *simply connected* if and only if it is connected and $\pi_1(X)$ is trivial.

A vector space $E$ is simply connected, as

$$H(s,t) = (1-s)\gamma(t)$$

is a loop homotopy of any given loop $\gamma$ on $E$ at $0$ to the constant loop at $0$.

The same argument works for *star-shaped* sets $X$, which be definition include, for some $p \in X$, the segment $px = \{p + t(x-p)\}_{0 \leqslant t \leqslant 1}$ for all $x \in X$.

The spheres $S^{n-1}$ are simply connected for $n \geqslant 3$, as in this case any loop on $S^{n-1}$ can be deformed to a loop that avoids $e_n$ and hence

$$\pi_1(S^{n-1}) = \pi_1(S^{n-1} - \{e_n\}) = \pi_1(\mathbf{R}^{n-1}) = \{0\}.$$

This last argument does not work for $S^1$ ($n = 2$), for any loop on $S^1$ going at least once round it cannot be deformed to avoid $e_2$.

Actually, in this case $\pi_1(S^1) \simeq \mathbf{Z}$, where the isomorphism is given by counting the number of times a loop on $S^1$ goes round $S^1$, with the sign $\pm$ determined by the *sense* (counterclockwise or clockwise) of the net number of turns.

# Topological groups

## Definition and examples

## Quaternions, $SU_2$ and $SO_3$

**Defintion**. A *topological group* is a group $G$ endowed with a topology such that the group operation $G \times G \to G$ and the inverse map $G \to G$, $g \mapsto g^{-1}$, are continuous.

**Examples**. The group $\mathrm{GL}_n$ of (real) invertible matrices of order $n$ is a topological group (*general linear group*). It is an open subset of $\mathbf{R}(n) \simeq \mathbf{R}^{n^2}$ and the expressions for the product of two matrices and for the inverse of a matrix show that they are continuous maps.

From this it follows that any subgroup of $\mathrm{GL}_n$ is a topological group with the induced topology. In particular, the following groups are topological groups:

- $\mathrm{SL}_n$ (*special linear group*): matrices of determinant 1.
- $\mathrm{O}_{r,s}$ (*orthogonal group of signature* $(r,s)$):

$$\{A \in \mathrm{GL}_n \mid A^T I_{r,s} A = I_{r,s}\},\ I_{r,s} = \mathrm{diag}(1, \overset{r)}{\ldots}, 1, -1, \overset{s)}{\ldots}, -1).$$

- $\mathrm{O}^+_{r,s} = \mathrm{SO}_{r,s}$ (*special orthogonal group of signature* $(r,s)$: subgroup of $\mathrm{O}_{r,s}$ of matrices $A$ such that $\det(A) = 1$. Note: $\mathrm{O}_{r,s} = \mathrm{O}^+_{r,s} \sqcup \mathrm{O}^-_{r,s}$.
- $\mathrm{O}^0_{r,s} = \mathrm{SO}^0_{r,s}$: The connected component of the identity of $\mathrm{SO}_{r,s}$.
- For the *Euclidean signature* $(n,0)$, we simply write $\mathrm{O}_n$ and $\mathrm{SO}_n$. In this case, $\mathrm{SO}^0_n = \mathrm{SO}_n$. So $\mathrm{O}_n = \{A \in \mathrm{GL}_n \mid A^T A = I_n\}$.
- $\mathrm{SO}_2 \simeq \mathrm{U}_1 = \{e^{i\theta} \mid 0 \leqslant \theta < 2\pi\}$ (*group of unit complex numbers*):

$$e^{i\theta} = \cos\theta + i\sin\theta \leftrightarrow \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

- $\mathrm{SE}_{r,s}$ ($\mathrm{SE}_n$ in the Euclidean case): the group of affine maps of $\mathbf{R}^n$, $x \mapsto xA + b$ with $A \in \mathrm{SO}_{r,s}$. In the Euclidean case, it is the *group of rigid motions*.

These maps can be identified with the matrices

$$\begin{pmatrix} A & 0 \\ b & 1 \end{pmatrix}, \quad (x,1)\begin{pmatrix} A & 0 \\ b & 1 \end{pmatrix} = (xA+b,1)$$

The composition is morphed into the matrix product

$$\begin{pmatrix} A & 0 \\ b & 1 \end{pmatrix}\begin{pmatrix} A' & 0 \\ b' & 1 \end{pmatrix} = \begin{pmatrix} AA' & 0 \\ bA'+b' & 1 \end{pmatrix}$$

and this shows that $\mathrm{SE}_{r,s}$ is a topological group.

Note that

$$\begin{pmatrix} A & 0 \\ b & 1 \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -bA^{-1} & 1 \end{pmatrix}.$$

Consider the injective $\mathbf{R}$-linear map $\mathbf{C}^2 \to \mathbf{C}(2)$,
$(z,w) \mapsto h = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$, and let $\mathbf{H}$ be its image.

It is easy to check that $\mathbf{H}$ is a subring of $\mathbf{C}(2)$.

Let $\tilde{h} = \begin{pmatrix} \bar{z} & -w \\ \bar{w} & z \end{pmatrix}$ (*conjugate-transpose*, or just *conjugate* of $h$).
Then $h\tilde{h} = (z\bar{z} + w\bar{w})I_2 = \det(h)I_2$. Since $\tilde{h} \in \mathbf{H}$, it follows that if $h \neq 0$, then $\frac{1}{\det(h)}\tilde{h} = h^{-1} \in \mathbf{H}$. So $\mathbf{H}$ is a field.
Notation: $\mathbf{H}^{\times} = \mathbf{H} - \{0\}$, the *multiplicative group* of $\mathbf{H}$.

Let $\mathbf{1} = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $\mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

These matrices satisfy *Hamilton's relations*: $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -\mathbf{1}$; and if $z = a + bi$, $w = c + di$, then $h = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. So $\mathbf{H}$ is isomorphic to *Hamilton's quaternion field*.

- Since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ have trace 0, we have $a = \frac{1}{2}\mathrm{tr}(h)$, which we will denote by $h_0$ (*scalar part of h*).
- Set $E = E_3 = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle = \{h \in \mathbf{H} \mid h_0 = 0\}$ (*vector quaternions*). The *vector part* of $h$ is $h_1 = h - h_0$.
- If $h' = a'\mathbf{1} + b'\mathbf{i} + c'\mathbf{j} + d'\mathbf{k}$, then

$$(\tilde{h}h')_0 = aa' + bb' + cc' + dd',$$

which is the Euclidean metric on $\mathbf{H}$ with orthonormal basis $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$. We will denote it by $h \cdot h'$. In particular, denoting by $|h|$ the norm $\|h\|$ of $h$ (often called the *modulus* of $h$),

$$|h|^2 = a^2 + b^2 + c^2 + d^2 = z\bar{z} + w\bar{w} = \det(h),$$

which implies that $|hh'| = |h||h'|$.

Restricted to $E_3$, the inner product $x \cdot x'$ is the Euclidean metric with orthonormal basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

- If $v, v' \in E_3$, then $vv' = -v \cdot v' + v \times v'$, where $v \times v'$ is the *cross product*. In fact, if $v = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $v' = v_1'\mathbf{i} + v_2'\mathbf{j} + v_3'\mathbf{k}$, a short computation shows that

$$vv' = -v \cdot v' + (v_2v_3' - v_3v_2')\mathbf{i} + (v_3v_1' - v_1v_3')\mathbf{j} + (v_1v_2' - v_2v_3')\mathbf{k}.$$

**Lemma**. For all $h, h' \in \mathbf{H}$, $(hh')^{\sim} = \tilde{h}'\tilde{h}$.

By definition, $\tilde{h} = \bar{h}^T$, where $\bar{h}$ is the complex-conjugate of $h$. Therefore $(hh')^{\sim} = (\bar{h}\bar{h}')^T = (\bar{h}')^T\bar{h}^T = \tilde{h}'\tilde{h}$. □

- For a given $h \in \mathbf{H}$, let $\underline{h} : \mathbf{H} \to \mathbf{H}$, $\underline{h}(x) = hx\tilde{h}$ (a real linear map, which belongs to $\mathrm{GL}(\mathbf{H})$ if $h \neq 0$).

**Lemma**. The map $\mathbf{H}^\times \to \mathrm{GL}(\mathbf{H})$ is a group homomorphism.

If $h, h' \in \mathbf{H}^\times$, then $\underline{hh'}(x) = hh'x(hh')^{\sim} = hh'x\tilde{h}'\tilde{h} = \underline{h}(\underline{h'}(x))$. □

**Lemma**. The map $\underline{h}$ is *linear similarity* of *ratio* $|h|^2$.

Indeed, $|\underline{h}(x)|^2 = (hx\tilde{h})(hx\tilde{h})^\sim = hx\tilde{h}h\tilde{x}\tilde{h}$, and the claim follows because $\tilde{h}h = |h|^2$, $x\tilde{x} = |x|^2$, and $h\tilde{h} = |h|^2$, so that $|\underline{h}(x)|^2 = |x|^2|h|^4$ and hence $|\underline{h}(x)| = |h|^2|x|$. □

**Lemma**. If $h \neq 0$, $\underline{h}$ induces a linear similarity of $E_3$ of ratio $|h|^2$.

It is enough to show that $(hx\tilde{h})_0 = 0$ if $x_0 = 0$. This is a consequence of the formula $h_0 = \frac{1}{2}\mathrm{tr}(h)$, for all $h \in \mathbf{H}$:

$$(hx\tilde{h})_0 = \tfrac{1}{2}\mathrm{tr}(hx\tilde{h}) = \tfrac{1}{2}\mathrm{tr}(\tilde{h}hx) = \tfrac{1}{2}|h|^2\mathrm{tr}(x) = 0.$$ □

We have used that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, for all $A, B \in \mathbf{R}(n)$.

- $SU_2 = \{h \in \mathbf{H} : |h| = 1\} = S^3(\mathbf{H})$. In particular, $SU_2$ is simply connected.

**Corollary**. If $h \in SU_2$, then $\underline{h} \in SO_3$ and the map $SU_2 \to SO_3$, $h \mapsto \underline{h}$ is a group homomorphism. □

**Lemma**. The kernel of the homomorphism $\mathrm{SU}_2 \to \mathrm{SO}_3$ is $\pm\mathbf{1}$.

If $h$ is in the kernel, then $hv = vh$ for any $v \in E_3$. In particular, we have $h\mathbf{i} = \mathbf{i}h$, which implies $w = 0$, hence $h = \begin{pmatrix} z & 0 \\ 0 & \bar{z} \end{pmatrix}$ and $z\bar{z} = 1$. Now $h\mathbf{j} = \mathbf{j}h$ yields that $z = \pm 1$, hence $h = \pm\mathbf{1}$. □.

**Theorem**. The homomorphism $\mathrm{SU}_2 \to \mathrm{SO}_3$ is surjective.

Let $v \in S^2(E)$ be a unit vector. Then $v^2 = -v \cdot v = -1$. Given any $\theta \in \mathbf{R}$, $h = e^{\theta v} = \cos\theta + v \sin\theta \in \mathrm{SU}_2$. Since $v$ commutes with $h$, $\underline{h}(v) = e^{\theta v} v e^{-\theta v} = v$. This means that $\underline{h}$ is a rotation about the axis $\langle v \rangle$. Now, if $w \in v^{\perp}$, then $vw = v \times w = -w \times v = -wv$, and therefore $\underline{h}(w) = e^{\theta v} w e^{-\theta v} = e^{2\theta v} w = (\cos 2\theta + v \sin 2\theta) w =$ $w\, cos 2\theta + (v \times w) \sin 2\theta$, which implies (take $w$ of unit length) that $\underline{h}$ induces a rotation of amplitude $2\theta$ in $v^{\perp}$. In sum, $\underline{h}$ is the rotation of amplitude $2\theta$ about the axis $\langle v \rangle$. Thus the rotation $R_{v,\alpha}$ of amplitude $\alpha$ about $v$ is equal to $\underline{h}$, where $h = \cos\alpha/2 + v \sin\alpha/2$. □

# The differential realm

## Differencials, directional derivatives and gradients

## Manifolds

## Tangent spaces

## Inverse function theorem

## Implicit function theorem

## Projective spaces

## Grassmannians

## Tangent bundle and vector fields

Let $U \subseteq \mathbf{R}^n$ be an open set and $f : U \to \mathbf{R}^m$ a map.

We say that $f$ is *differentiable* at $x \in U$ if there is a linear function $\ell_x : \mathbf{R}^n \to \mathbf{R}^m$ that approximates the increment $(\Delta_x f)(v) = f(x+v) - f(x)$, as a function of $v$, up to second order terms. More formally,

$$f(x+v) - f(x) = \ell_x(v) + o(v), \text{ where } o(v)/\|v\| \to 0 \text{ when } v \to 0.$$

If $\ell_x$ exists, it is unique, is denoted by $d_x f$, is called the *differential of $f$* at $x$, and $f$ is said to be *differentiable* at $x$.

In that case, for any $v \in \mathbf{R}^n$ the directional derivative $\partial_v f(x) = D_v f(x) = \frac{df(x+tv)}{dt}\big|_{t=0}$ exists, and $\partial_v f(x) = d_x f(v)$. The partial derivatives $\partial_i f(x) = \partial_{e_i} f(x)$ exist, so also exists $\nabla f(x)$, and $d_x f(v) = \nabla f(x) \cdot v$, defined as $(\nabla f_1(x) \cdot v, \ldots, \nabla f_m(x) \cdot v)$, where $f = (f_1, \ldots, f_m)$.

If $d_x f$ exists for any $x \in U$, $f$ is said to be *differentiable* in $U$. In this case, the partial derivatives $\partial_i f(x)$ exist for all $x \in U$.

$\nabla f(x)$ is also called the *Jacobian matrix* of $f$ at $x$. Its entries are $\partial_i f_j$ $(i \in [n], j \in [m])$.

The function $f$ is *smooth*, or of *class* $\mathcal{C}^\infty$, if $f$ has continuous partial derivatives of all orders at any point of $U$. The vector space of smooth functions $U \to \mathbf{R}^m$ is denoted $\mathcal{C}^\infty(U, \mathbf{R}^m)$. For $m = 1$, it is an algebra that we denote simply by $\mathcal{C}^\infty(U)$.

More generally, if $Y \subseteq \mathbf{R}^n$, a map $f : Y \to \mathbf{R}^m$ is said to be *differentialbe* (respectively *smooth*) if for any point $y \in Y$ there is an open set $U_y \subseteq \mathbf{R}^n$ that contains $y$ and a differentiable (smooth) function $\varphi_y : U_y \to \mathbf{R}^m$ such that $f(x) = \varphi_y(x)$ for all $x \in U_y \cap Y$.

If $f : Y \to \mathbf{R}^m$ is smooth and $Z = f(Y)$, we say that $f : Y \to Z$ is a *diffeomorphism* if $f$ is bijective and $f^{-1} : Z \to Y$ is smooth.

For example, the stereographic projection $\sigma : S^{n-1} - \{\boldsymbol{e}_n\} \to \mathbf{R}^{n-1}$ is a diffeormorphism.

Indeed, the expression $\sigma(x) = (x - x_n e_n)/(1 - x_n)$ for $\sigma$ shows that it makes sense, and is smooth, for any point not on the hyperplane $x_n = 1$, while $\sigma^{-1} : \mathbf{R}^{n-1} \to \mathbf{R}^n$, $y \mapsto (2y + (y^2 - 1)e_n)/(y^2 + 1)$, is also smooth and its image is $S^{n-1} - \{e_n\}$.

A space $Y$ is said to be a *manifold* of *dimension* $d$ if each point $y \in Y$ has an open neighborhood (in $Y$) that is diffeomorphic to an open set of $\mathbf{R}^d$. The dimension $d$ of $Y$ is denoted by $\dim(Y)$.

Example. Any non-empty open set $U$ of $\mathbf{R}^n$ is a manifold and $\dim U = n$.

Example. What we have said about the stereographic projection shows that $S^{n-1}$ is a manifold of dimension $n-1$ for any $n \geqslant 1$.

If $Y \subseteq E_n$ is a manifold, and $y \in Y$, a vector $v \in E_n$ is said to be *tangent* to $Y$ at $y$ if there is a smooth function $\gamma : (-\varepsilon, \varepsilon) \to Y$, $(-\varepsilon, \varepsilon) \subset \mathbf{R}$, such that $\gamma(0) = y$ and $\dot\gamma(0) = v$.

We will write $T_y Y$ to denote the set of vectors tangent to $Y$ at $y$, and we will say that it is the *tangent space* to $Y$ at $y$.

For example, $T_y E_n = E_n$ for any point $y \in E_n$, because if $\gamma(t) = y + tv$, $v \in E_n$, then we have $\dot\gamma(t) = v$ for any $t$.

Since $\mathrm{GL}(E_n) \subset \mathrm{End}(E_n)$ is open,

$$T_{\mathrm{Id}}\mathrm{GL}(E_n) = T_{\mathrm{Id}}\mathrm{End}(E_n) = \mathrm{End}(E_n).$$

In general, $T_y Y$ is a linear subspace of $E_n$ and $\dim T_y Y = \dim Y$.

Let $E$ and $F$ be vector spaces, $U$ a non-empty open set of $E$ and $f: U \to F$ a smooth function.

**Theorem**. If $u \in U$ is such that $d_u f: E \to F$ is injective, then there exists an open set $U' \subseteq U$, $u \in U'$, such that $f: U' \to f(U')$ is a diffeomorfism.

This means that $f(U)$ is a manifold of dimension $\dim(E)$ near $f(u)$. Moreover, $T_u(f(U)) = (d_u f)(E)$.

See, for example, [8, §5.3, Th. 3]. □

Let $E$ and $F$ be vector spaces, $U$ a non-empty open set of $E$ and $f : U \to F$ a smooth function.

**Theorem**. Set $Z = \{z \in U \,|\, f(z) = 0\}$. If $z \in Z$ is such that $d_z f : E \to F$ is *surjective*, then there exists an open set $U' \subseteq U$, $z \in U'$, such that $Z' = Z \cap U'$ is a manifold of dimension $d = \dim(E) - \dim(F)$ and $T_z Z' = \ker(d_z f)$.

See, for example, [8, §5.3, Th. 4]. □

If $F = \mathbf{R}^m$ and $f = (f_1, \ldots, f_m)$, then

$$Z = Z(f) = Z(f_1) \cap \cdots \cap Z(f_m)$$

and the theorem implies that $Z$ is a manifold around a point $z$ if $d_z f_1, \ldots, d_z f_m$ are linearly independent, and in this case $T_z Z = \ker d_z f = \cap_j \ker d_z f_j$ (cf. 21-05b-Opt, *classical Lagrange multipliers*).

## Example

Although we know, via the stereographic projection, that $S^{n-1}$ is a manifold of dimension $n-1$, it is instructive to prove it again using the implicit function theorem.

Consider the function $f : E_n \to \mathbf{R}$ given by $f(x) = x^2$, so that $S^{n-1} = Z(f-1)$.

To apply the theorem, let us find $d_y f$ at a point $y \in S^{n-1}$.

For any vector $v \in E_n$,

$$(d_y f)(v) = \tfrac{d}{dt} f(y+tv)|_{t=0} = \tfrac{d}{dt}(y+tv)^2|_{t=0} = 2y \cdot v.$$

Now for any non-zero $y$, in particular for any $y \in S^{n-1}$, the map $E_n \to \mathbf{R}$, $v \mapsto 2y \cdot v$ is surjective. Therefore $S^{n-1}$ is a manifold of dimension $n-1$ around anyone of its points $y$, and $T_y S^{n-1} = y^\perp$.

## Example

Consider the group $\mathrm{SL}(E) \subset \mathrm{GL}(E)$, which by definition can be represented as $Z(\det -1)$.

We will see that $d_{\mathrm{Id}} \det = \mathrm{tr}$, from which it follows, since $\mathrm{tr} : \mathrm{End}(E) \to \mathbf{R}$ is surjective, that $\mathrm{SL}(E)$ is a manifold near $\mathrm{Id}$ of dimension $n^2 - 1$ ($n = \dim E$) and

$$T_{\mathrm{Id}}\mathrm{SL}(E) = \{h \in \mathrm{End}(E) \,|\, \mathrm{tr}(h) = 0\} = \mathrm{End}_0(E).$$

To prove the claim, note that for any $h \in \mathrm{End}(E)$ we have

$$(d_{\mathrm{Id}} \det)(h) = \tfrac{d}{dt} \det(\mathrm{Id} + th)|_{t=0} = \tfrac{d}{dt}(1 + \mathrm{tr}(th) + \cdots)|_{t=0} = \mathrm{tr}(h).$$

Finally note that $\mathrm{SL}(E)$ is a manifold of dimension $n^2 - 1$ near any $g \in \mathrm{SL}(E)$ because the map $L_g : \mathrm{SL}(E) \to \mathrm{SL}(E)$, $f \mapsto gf$, is a diffeomorfism and $L_g(\mathrm{Id}) = g$.

The notion of manifold given on page 27 needs a broadening that liberates it from having to be a subset of some vector space (see, for instance, [8, §5.1], or [9, §1.2b]).

The definition of an abstract manifold is quite natural, as it is based on reflecting that it looks like an open set of a vector space in the neighborhood of each of its points, with differentiable transitions between overlapping neighborhoods.

For example, if we identify antipodal points on the sphere $S^{n-1}$, $P^{n-1} = S^{n-1}/\{\pm 1\}$, we have a manifold in the abstract sense. Indeed, any open set of $S^{n-1}$ that does not contain pairs of antipodal points is mapped injectively into $P^{n-1}$, which means that locally $P^{n-1}$ looks like the manifold $S^{n-1}$. Since $P^{n-1} \simeq \mathbf{P}(\mathbf{R}^n)$, $[x] \mapsto [x/\|x\|]$ we may conclude that the projective space $\mathbf{P}(\mathbf{R}^n)$ is a manifold of dimension $n-1$.

This can also be concluded by means of the *coordinates* $x_1, \ldots, x_n$ in $\mathbf{R}^n$: $\mathbf{P}^n - \{x_j = 0\} \leftrightarrow \mathbf{R}^{n-1}$, $[x_1, \ldots, x_n] \mapsto [x_1/x_j, x_{j-1}/x_j, x_{j+1}/x_j, \ldots, x_n/x_j]$.

Given a $k$-dimensional linear subspace $L$ of the vector space $E$, let $g(L) \in \mathbf{P}(\wedge^k E)$ be defined as $[x_1 \wedge \cdots \wedge x_k]$, where $x_1, \ldots, x_k$ is any basis of $L$. The point $g(L)$ only depends on $L$, for the exterior product of two basis are proportional.

Moreover, $L \mapsto g(L)$ is injective, as the vectors $x \in L$ are precisely those satisfying $x \wedge x_1 \wedge \cdots \wedge x_k = 0$.

Let $\mathrm{Gr}_k(E) \subset \mathbf{P}(\wedge^k E)$ be the image of $g$. It turns out that this is a submanifold of dimension $(k+1)(n-k)$ of $\mathbf{P}(\wedge^k E)$. Such manifolds are called *Grassmann manifolds*, popularly *Grassmannians*, [9, §17.2b]. The projective space $\mathbf{P}(E)$ is the special case $\mathrm{Gr}_1(E)$.

The *tangent bundle* $TM$ of a manifold $M$ of dimension $n$ is manifold of dimension $2n$ endowed with a differentiable map $\pi : TM \to M$ with the property that $\pi^{-1}(x) \simeq T_x M$.

- For an open set $U \subseteq E$, $TU = U \times E$, with $\pi$ the projection map.
- $TS^{n-1} = \{(y, v) \in S^{n-1} \times \mathbf{R}^n : y \cdot v = 0\}$.

The *cotangent bundle* has a similar meaning, but with $T_x(M)$ replaced by $T_x^* M$ (the *dual space* of $T_x M$).

A *vector field* $v$ on $X$ assigns a tangent vector $v_x \in T_x M$ for any $x \in M$ in such a way that the map $M \to TM$, $x \mapsto v_x$, is differentiable.

*Vector bundles* are a generalization of the tangent and cotangent bundles. They are *locally trivial* families of *vector spaces*. The dimension of these spaces is the *rank* of the vector bundle.

*Example*: $V = \{(x, v) \in S^{n-1} \times \mathbf{R}^n : v \in \langle x \rangle\}$. Its rank is 1 (*a line bundle*).

# Lie groups and algebras

## Definition and examples

## Remarks on $\mathrm{O}_{r,s}$

## Lie algebras

We have seen that the groups $\mathrm{GL}_n$ and $\mathrm{SL}_n$ are at the same time *topological groups* and manifolds, and that in fact the *multiplication and inversion* maps are *smooth*. In other words, they are *Lie groups*. Their dimensions are $n^2$ and $n^2-1$, respectively.

**Example**. $\mathrm{O}_{r,s}$ is a Lie group of dimension $\binom{n}{2}$, $n=r+s$.

Let $\gamma:(-\epsilon,\epsilon)\to\mathrm{O}_{r,s}$ be a differentiable path with $\gamma(0)=\mathrm{Id}$ and let $B=\dot\gamma(0)\in\mathrm{M}_n=\mathbf{R}(n)$. Since $\gamma(t)^T I_{r,s}\gamma(t)=I_{r,s}$, on taking the derivative with respect to $t$, at $t=0$, we get $B^T I_{r,s}+I_{r,s}B=0$. This shows that $T_{\mathrm{Id}}\mathrm{O}_{r,s}\subseteq\mathfrak{so}_{r,s}=\{B\in\mathrm{M}_n : B^T I_{r,s}=-I_{r,s}B\}$.

In fact we now proceed to show that $T_{\mathrm{Id}}\mathrm{O}_{r,s}=\mathfrak{so}_{r,s}$.

Let $B\in\mathfrak{so}_{r,s}$, and consider the map $\gamma:\mathfrak{so}_{r,s}\to\mathrm{GL}_n$ defined by $\gamma(t)=e^{tB}$. As we will see in a moment, we actually have $\gamma(t)\in\mathrm{O}_{r,s}$, with $\gamma(0)=\mathrm{Id}$, and clearly $\dot\gamma(0)=B$, so $B\in T_{\mathrm{Id}}\mathrm{O}_{r,s}$.

Let us check that $\gamma(t) \in \mathrm{O}_{r,s}$ for all $t$.

Using that $(B^T)^k I_{r,s} = I_{r,s}(-1)^k B^k$, which follows from $B^T I_{r,s} = -I_{r,s} B$ by induction on $k$, we infer that the claim holds:

$$(e^{tB})^T I_{r,s} e^{tB} = I_{r,s} e^{-tB} e^{tB} = I_{r,s}.$$

That $O_{r,s}$ is a manifold of dimension $\binom{n}{2}$ is a nice application of the inverse function theorem.

Consider the map $\exp : \mathfrak{so}_{r,s} \to O_{r,s}$, $B \mapsto e^B$. Then $d_0 \exp$ is a linear map from $T_0\,\mathfrak{so}_{r,s} = \mathfrak{so}_{r,s}$ to $T_{\mathrm{Id}}\mathrm{O}_{r,s} = \mathfrak{so}_{r,s}$, and this map is the identity: $d_0 \exp(B) = (D_B \exp)(0) = (de^{tB}/dt)|_{t=0} = B$.

It follows that $\exp$ induces a diffeomorphism of an open neighborhood of $0$ in $\mathfrak{so}_{r,s}$ and an open neighborhood of $\mathrm{Id}$ in $\mathrm{O}_{r,s}$ and this implies that $\mathrm{O}_{r,s}$ is a manifold, hence a Lie group, of dimension $\binom{n}{2}$

(1) For any $(r,s)$, $\mathrm{O}_{r,s} = \mathrm{O}^+_{r,s} \sqcup \mathrm{O}^-_{r,s}$, $\mathrm{O}^+_{r,s} = \mathrm{SO}_{r,s}$ and $\mathrm{O}^-_{r,s} = \alpha\mathrm{SO}_{r,s}$ for any given $\alpha \in \mathrm{O}^-_{r,s}$ (as $\alpha$ we can take the orthogonal reflection $m_u$ with respect to a non-isotropic vector $u$: $m_u(x) = x$ if $x \in u^\perp$ and $m_u(u) = -u$).

(2) If $(r,s) = (n,0)$ (*Euclidean case*) or $(r,s) = (0,n)$ (*anti-Euclidean case*), then $\mathrm{SO}_{r,s}$ is connected and hence $O_{r,s}$ has two connected components.

(3) If $r, s \geqslant 1$, then $\mathrm{SO}_{r,s} = \mathrm{SO}^0_{r,s} \sqcup m_u m_{\bar{u}}\mathrm{SO}^0_{r,s}$, where $u, \bar{u}$ are any non-isotropic vectors of oposite signatures ($u^2\bar{u}^2 < 0$). It follows that in this case $O_{r,s}$ has 4 connected components.

(4) Example. $\mathrm{O}_{1,3}$ is the *general Lorentz group*, $\mathrm{O}^+_{1,3} = \mathrm{SO}_{1,3}$ is the *proper* Lorentz group, and $\mathrm{SO}^0_{1,3}$ is the *orthochronous* or *restricted* Lorentz group (proper Lorentz transformations that *preserve the time orientation*).

Let $G$ be any of the Lie groups considered so far, and write $\mathfrak{lie}(G)$ to denote its tangent space at the identity element of $G$. More specifically, we have:

$$\mathfrak{lie}(\mathrm{GL}(E)) = \mathrm{End}(E)$$

$$\mathfrak{lie}(\mathrm{SL}(E)) = \mathrm{End}_0(E) \text{ (the traceless endomorphisms of } E\text{)}$$

$$\mathfrak{lie}(\mathrm{O}_{r,s}) = \mathfrak{lie}(\mathrm{SO}_{r,s}) = \mathfrak{lie}(\mathrm{SO}^0_{r,s}) = \mathfrak{so}_{r,s}$$

In all cases, $\mathfrak{lie}(G)$ is closed under the *commutator bracket* ($[A, A'] = AA' - A'A$) and hence it is a *Lie algebra*. This claim is clear for $\mathfrak{lie}(\mathrm{GL}(E))$. The case of $\mathfrak{lie}(\mathrm{SL}(E))$ is an immediate consequence of the fact that $\mathrm{tr}([A, B]) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0$. The case of $\mathfrak{so}_{r,s}$ is checked with the following computation, where $B, C \in \mathfrak{so}_{r,s}$:

$$\begin{aligned}[B, C]^T I_{r,s} &= (C^T B^T - B^T C^T) I_{r,s} = -C^T I_{r,s} B + B^T I_{r,s} C \\ &= I_{r,s} CB - I_{r,s} BC = -I_{r,s}[B, C].\end{aligned}$$

We have seen that $\mathrm{SE}_{r,s}$ (in particular $\mathrm{SE}_n$) is a topological group.

By inspecting its multiplication and inverse maps, page 17, we see that it is a Lie group.

Its Lie algebra $\mathfrak{se}_{r,s}$ (tangent space at $\mathrm{Id}$) can be determined as for $\mathrm{SO}_{r,s}$, and the result is that it is the Lia algebra of matrices of the form

$$\begin{pmatrix} B & 0 \\ v & 0 \end{pmatrix}, \quad B \in \mathfrak{so}_{r,s},\ v \in \mathbf{R}^n.$$

The argument with the exponential can be adapted to this case and the outcome is that $\mathrm{SE}_{r,s}$ is a Lie group of dimension $\binom{n+1}{2}$, $n = r + s$.

This agrees with the intuition that the *degres of freedom* a rigid motion are $n$ for the translation plus the degrees of freedom (dimension) of a rotation.

# Appendix

## Two properties of the stereographic projection $\sigma$

**Lemma**. The section $S'$ of the hyperplane $\Pi : u \cdot x = \delta$ ($u \in \mathbf{R}^n$ unitary, $\delta \in \mathbf{R}_+$) with the unit sphere $S^{n-1}$ is empty if $\delta > 1$, the point $u$ if $\delta = 1$ and the sphere with center at $\delta u$ and radius $\rho = \sqrt{1-\delta^2}$ if $\delta < 1$

The plane $\Pi$ cuts the line $\{\lambda u\}_{\lambda \in \mathbf{R}}$ at $\delta u$. For any $x \in S'$, we have $1 = x^2 = (x - \delta u)^2 + (\delta u)^2 \geqslant \delta^2$. Hence the intersection is empty unless $\delta \leqslant 1$. For $\delta = 1$, the only solution is $x = u$ (and $\Pi$ is the tangent hyperplane to $S^{n-1}$ at $u$). If $\delta < 1$, then any $x$ in the intersection satisfies, writing $\rho = \|x - \delta u\|$, $1 = \rho^2 + \delta^2$, which shows that $S'$ is the sphere in $\Pi$ with center $\delta u$ and radius $\rho$. □

Note: for $\delta = 0$, the section $S^{n-2}$ has radius 1, the greatest possible (*equatorial spheres*).

Let $S^{n-2} \subset S^{n-1}$ be the section with the hyperplane $\Pi : u \cdot x = \delta$, $u \in \mathbf{R}^n$ a unit vector and $\delta \in \mathbf{R}_+$.

The $y = \sigma(x) \in \sigma(S^{n-2})$ iff and only if $x = \frac{2y+(y^2-1)e_n}{y^2+1}$ belongs to $\Pi$, namely,

$$2(u \cdot y) + c(y^2 - 1) = \delta(y^2 + 1),$$

where $c = u \cdot e_n$ (the cosine of the angle $\widehat{u, e_n}$).

Letting $\bar{u}$ be the orthogonal projection of $u$ to $\mathbf{R}^{n-1}$, it is equivalent to

$$(\delta - c)y^2 - 2(\bar{u} \cdot y) + \delta + c = 0.$$

The condition $\delta = c$ means that $\Pi$ passes through $e_n$, and in this case $\sigma(S^{n-2})$ is the hyperplane $\bar{u} \cdot y = \delta$ of $\mathbf{R}^{n-2}$, that is $\Pi \cap \mathbf{R}^{n-2}$. This conclusion clearly matches the geometric intuition of the case.

If $\delta \neq c$, then $\sigma(S^{n-2})$ is the $\mathbf{R}^{n-1}$ sphere with center $u'$ and radius $\rho'$, where $u' = \bar{u}/(\delta - c)$ and ${\rho'}^2 = {u'}^2 - (\delta + c)/(\delta - c)$.

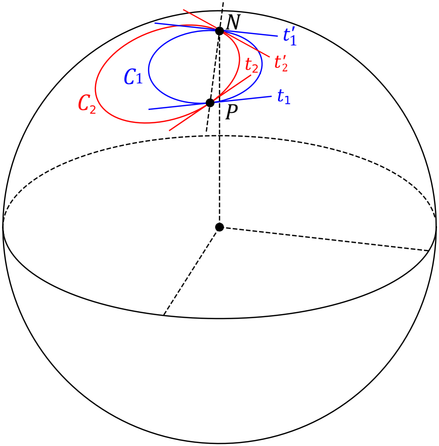


Figure 9.1: Let $N, P \in S^2$, $P \neq N$. Let $t_1$ and $t_2$ be lines tangent to $S^2$ at $P$. The planes $\Pi_i = [N, t_i]$ ($i = 1, 2$) cut $S^2$ along the circles $C_i$ that pass through $N$ and $P$ and which touch $t_i$ at $P$. If we let $t'_i$ denote the tangents to the $C_i$ at $N$, then $\angle t'_1 t'_2 = \angle t_1 t_2$. Notice that $t'_i$ is the intersection of $\Pi_i$ with the tangent plane to $S^2$ at $N$. This implies that $\angle t_1 t_2 = \angle t''_1 t''_2$, where $t''_i$ is the tangent to $\sigma(C_i)$ at $\sigma(P)$.

# References I

[1] M. M. Bronstein, J. Bruna, T. Cohen, and P. Velickovic, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," 2021.
arXiv:2104.13478.

[2] T. Cohen, *Equivariant Convolutional Networks*.
PhD thesis, University of Amsterdam, 2021.

[3] J. Gallier and J. Quaintance, *Differential geometry and Lie groups: a computational perspective*, vol. 12 of *Geometry and Computing*.
Springer, 2020.
xv+777p.

# References II

[4] S. Xambó-Descamps, *Real Spinorial Groups: A Short Mathematical Introduction*.

Springer, 2018.

[5] M. Lee John, *Introduction to smooth manifolds*, vol. 218 of *Graduate Texts in Mathematics*.

Springer, 2013.

xvi+708p.

[6] T. K. Carne, "Geometry and groups," 2012.

Electronic publication.

# References III

[7] O. Smirnov, "TensorFlow RiemOpt: a library for optimization on Riemannian manifolds," 2021.

https://arxiv.org/pdf/2105.13921.

[8] I. M. Singer and J. A. Thorpe, *Lecture notes on elementary topology and geometry*.

Scott, Foresman and Company, 1967.

[9] T. Frankel, *The geometry of physics: an introduction*.

Cambridge University Press, 2011.