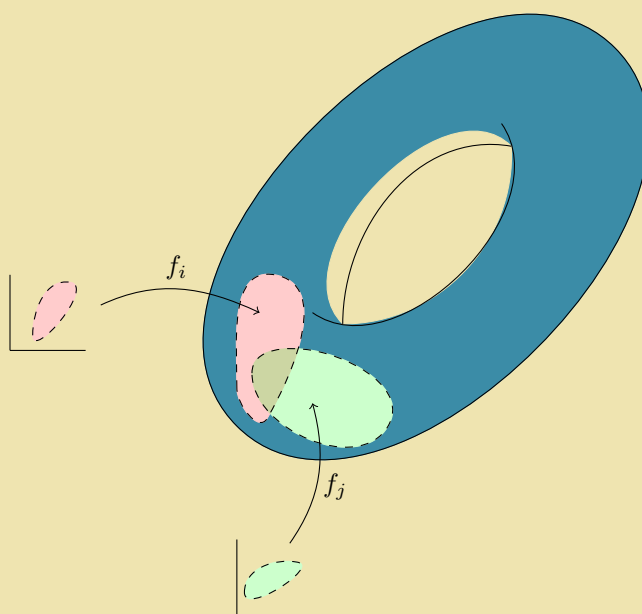# Geometric Calculus

Vasiliki Bitsouni
Anastasios Fragkos
Nikolaos Gialelis

f_i

f_j

The Copyright of the fonts belongs to the The Department of Mathematics of the University of the Aegean, Karlobasi, Samos, Greece. More information can be found here:

ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ

αβγδεζηθικλμνξοπρστυφχψω

ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ

αβγδεζηθικλμνξοπρστυφχψω

ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ

αβγδεζηθικλμνξοπρστυφχψω

ABCDEFGHIJKLMNOPQRSTUVWXYZ

abcdedghijklmnopqrstuvwxyz

BCDEFGHIJKLMNOPQRSTUVWXYZ

abcdedghijklmnopqrstuvwxyz

ABCDEFGHIJKLMNOPQRSTUVWXYZ

abcdedghijklmnopqrstuvwxyz

0123456789,./:""()[]{}<>-!#$%&*

*0123456789,./:""()[]<>{}-!#$%&**

0123456789,./:""()[]<>{}-!#$%&*

To …
V.B.

To …
A.F.

To …
N.G.

Preface of the first edition

T^est

Table of Contents

Prerequisites

1.1 Topology

1.1.1 Metric spaces and metric topologies

1.1.2 Some topological results

1.2 Analysis

1.2.1 Continuity and uniform continuity

1.3 Algebra

1.3.1 Isometric transformations of $\mathbb{R}^n$

We already saw (in [...]) the definition of isometries in metric spaces. Especially in spaces with inner product, those isometries have a very strong geometric "image". We will show that isometries in spaces with inner products are some combination of translations, reflections and rotations.

Definition 1.3.1 (Linear functions). *Let X, Y be two $\mathbb{R}-$vector spaces. We say a function $f : X \to Y$ is linear if for every $\lambda \in \mathbb{R}$ and $x_1, x_2 \in X$:*

$$f(x_1 + \lambda x_2) = f(x_1) + \lambda f(x_2)$$

We use the notation:

$$L(X \to Y) := \{f : X \to Y \mid f \text{ is linear}\}$$

Definition 1.3.2 (Orthogonal functions). *Let $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ be two spaces with inner product. We say a function $\tau : X \to Y$ is orthogonal if for every $x_1, x_2 \in X$:*

$$\langle x_1, x_2 \rangle_X = \langle \tau(x_1), \tau(x_2) \rangle_Y$$

We use the notation:

$$O(X \to Y) := \{\tau : X \to Y \mid \tau \text{ is orthogonal}\}$$

Corollary 1.3.1. *From Definition 1.3.2 follows that a function $\tau : X \to Y$ is orthogonal if and only if it preserves the forms of the inner products.*

$$\tau \in O(X \to Y) \Leftrightarrow T_X = T_Y \circ \tau$$

Also, having in mind the definition of the angle, every orthogonal function τ preserves the angles.

Lemma 1.3.1. *Let* $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ *be two spaces with inner product. Every* $\tau \in O(X \to Y)$ *is linear, thus* $O(X \to Y) \subseteq L(X \to Y)$.

Proof. Let $\tau \in O(X \to Y)$. We need to show that for every $\lambda \in \mathbb{R}$ and $x_1, x_2 \in X$ we have:

$$\tau(x_1 + \lambda x_2) = \tau(x_1) + \lambda \tau(x_2) \tag{1.1}$$

Instead, we will show that:

$$T_Y\big(\tau(x_1 + \lambda x_2) - \tau(x_1) - \lambda \tau(x_2)\big) = 0 \tag{1.2}$$

which equivalent to (1.1). T_Y is the form of the inner product $\langle \cdot, \cdot \rangle_Y$ of Y.

With many calculations to the left side of (1.2) we will find 0. Indicatively, we write:

$$\big\langle \tau(x_1 + \lambda x_2) - \tau(x_1) - \lambda \tau(x_2),\ \tau(x_1 + \lambda x_2) - \tau(x_1) - \lambda \tau(x_2) \big\rangle_Y =$$

$$\big\langle \tau(x_1 + \lambda x_2),\ \tau(x_1 + \lambda x_2) \big\rangle_Y - 2\big\langle \tau(x_1 + \lambda x_2),\ \tau(x_1) \big\rangle_Y + \big\langle \tau(x_1),\ \tau(x_1) \big\rangle_Y + \lambda^2 \big\langle \tau(x_2),\ \tau(x_2) \big\rangle_Y$$

and by the orthogonality of τ we get:

$$\langle x_1 + \lambda x_2,\ x_1 + \lambda x_2 \rangle_X - 2\langle x_1 + \lambda x_2,\ x_1 \rangle_X + \langle x_1, x_1 \rangle_X + \lambda^2 \langle x_2, x_2 \rangle_X$$

which is:

$$\langle x_1 + \lambda x_2 - x_1 - \lambda x_2,\ x_1 + \lambda x_2 - x_1 - \lambda x_2 \rangle_X = \langle 0_X, 0_X \rangle_X = 0$$

$\square$

Lemma 1.3.2. *Let* $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ *be two spaces with inner product. Linear isometries (*$\sigma \in I(X \to Y) \cap L(X \to Y)$*) are exactly those isometries* σ *for which* $\sigma(0_X) = 0_Y$.

$$\sigma \in I(X \to Y) \cap L(X \to Y) \Leftrightarrow \sigma \in I(X \to Y) \text{ and } \sigma(0_X) = 0_Y$$

Proof. The direction ($\Rightarrow$) holds. For the other direction ($\Leftarrow$), we notice that, for every $x_1, x_2 \in X$:

$$\sigma \in I(X \to Y) \Rightarrow ||x_1 - x_2||_X = ||\sigma(x_1 - x_2)||_Y \overset{\star}{\Rightarrow} T_X(x_1 - x_2) = T_Y\big(\sigma(x_1) - \sigma(x_2)\big)$$

where T_X, T_Y are the forms of the inner products in X and Y respectively, and $||\cdot||_X, ||\cdot||_Y$ are their induced normes. In implication ($\star$) the linearity of σ is used. Moreover, if we set $x_1 = 0_X$ or $x_2 = 0_X$:

$$T_X(x_1) = T_Y\big(\sigma(x_1)\big) \text{ and } T_X(x_2) = T_Y\big(\sigma(x_2)\big) \tag{1.3}$$

Now we write:

$$T_X(x_1 - x_2) = T_Y\big(\sigma(x_1) - \sigma(x_2)\big) \Leftrightarrow \langle x_1 - x_2, x_1 - x_2 \rangle_X = \big\langle \sigma(x_1) - \sigma(x_2),\ \sigma(x_1) - \sigma(x_2) \big\rangle_Y \Leftrightarrow$$

$$\Leftrightarrow \langle x_1, x_1 \rangle_X - 2\langle x_1, x_2 \rangle_X + \langle x_2, x_2 \rangle_X = \big\langle \sigma(x_1), \sigma(x_1) \big\rangle_Y - 2\big\langle \sigma(x_1), \sigma(x_2) \big\rangle_Y + \big\langle \sigma(x_2), \sigma(x_2) \big\rangle_Y$$

and from (1.3) we get:

$$\langle x_1, x_2 \rangle_X = \big\langle \sigma(x_1), \sigma(x_2) \big\rangle_Y$$

which shows that $\sigma \in O(X \to Y)$. Using Lemma 1.3.1 we prove that $\sigma \in L(X \to Y)$. $\square$

Theorem 1.3.1 (Structure theorem)**.** *Let* $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ *be two spaces with inner product. Then:*

$$I(X \to Y) \cap L(X \to Y) = O(X \to Y)$$

Proof. While proving Lemma 1.3.2 we showed that every $\sigma \in L(X \to Y)$ with $\sigma(0) = 0$ is orthogonal. So, using Lemma 1.3.2, we have shown that $I(X \to Y) \cap L(X \to Y) \subseteq O(X \to Y)$. It remains for inclusion "$\supseteq$" to be shown.

For this we will now prove that every orthogonal function τ is isometry with $\tau(0_X) = 0_Y$ (and

we will use, once more, Lemma 1.3.2). So let $\tau \in O(X \to Y)$. Because for every $x_1, x_2 \in X$, $\langle x_1 - x_2 \rangle_X = \langle \tau(x_1 - x_2) \rangle_Y$, the norms are equal:

$$\|x_1 - x_2\|_X = \|\tau(x_1 - x_2)\|_Y$$

(which are induced by our inner products). By the linearity of τ (which has been proven in Lemma 1.3.1) an equality of metrics follows:

$$\|x_1 - x_2\|_X = \|\tau(x_1) - \tau(x_2)\|_Y$$

which shows that τ is isometry. By letting $x_1 = x_2$ we also see that $\tau(0_X) = 0_Y$, and this concludes the proof. $\qquad\square$

Theorem 1.3.2 (Isometries as compositions of orthogonal functions and translations). *Let $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ be two spaces with inner products and $\sigma \in I(X \to Y)$. Then σ can be written in a unique way as a composition of an orthogonal function and a translation:*

$$\sigma = \mu \circ \tau$$

Here $\tau \in O(X \to Y)$ and $\mu : Y \to Y$, $\mu(y) = y + a$ (for some $a \in Y$).

Proof. First, notice that every translation is an isometry -so function $\tau(x) = \sigma(x) - \sigma(0_X)$ is an isometry, and especially $\tau(0_X) = 0_Y$. From Lemma 1.3.2 we have that $\tau \in I(X \to Y) \cap L(X \to Y)$, and then, from Theorem 1.3.1, $\tau \in O(X \to Y)$. If we now consider the translation $\mu(y) = y + \sigma(0_X)$, our previous analysis shows that $\mu \circ \tau = \sigma$.

We will prove the "unique" part of the theorem too. If we suppose that there exists some $\tau^* \in O(X \to Y)$ and translation $\mu^* : Y \to Y$, $\mu^*(y) = y + a^*$ such that $\sigma = \mu \circ \tau = \mu^* \circ \tau^*$, then -because τ, τ^* are linear (Lemma 1.3.1):

$$\sigma(0_X) = \mu(0_Y) = \mu^*(0_Y) \Rightarrow a = a^*$$

From this follows that $\mu = \mu^*$, and as an extension:

$$\sigma(x) = \tau(x) + a = \tau^*(x) + a \Rightarrow \tau = \tau^*$$

The proof is complete. $\qquad\square$

Lemma 1.3.3. *Let $(X, \langle \cdot, \cdot \rangle_X)$, $(Y, \langle \cdot, \cdot \rangle_Y)$ be two spaces of the same finite dimension, with inner products. Then every $\tau \in O(X \to Y)$ is a linear isomorphism.*

Proof. In Theorem 1.3.1 we saw that $O(X \to Y) = I(X \to Y) \cap L(X \to Y)$, so every $\tau \in O(X \to Y)$ is linear and isometric. Because it is isometric, we have also shown that is injective. Now, using the kernel-image dimension theorem -because of the same finite dimension of spaces X, Y- it follows that τ is surjective too. $\qquad\square$

Theorem 1.3.3 (Isometries of $\mathbb{R}^2$). *Let $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ be the Euclidean plane with the usual inner product. Every $\sigma \in I(\mathbb{R}^2 \to \mathbb{R}^2)$ can be written as:*

$$\sigma\big((x,y)\big)^T = \begin{cases} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \\ \qquad or \\ \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \end{cases}$$

where $\alpha^2 + \beta^2 = 1$ and $\gamma, \delta \in \mathbb{R}$.

Proof. In Theorem 1.3.2 we proved that every $\sigma \in I(\mathbb{R}^2 \to \mathbb{R}^2)$ can be written in a unique way as a composition of an orthogonal function and a translation.

$$\sigma = \mu \circ \tau, \text{ where } \tau \in O(\mathbb{R}^2 \to \mathbb{R}^2) \text{ and } \mu : \mathbb{R}^2 \to \mathbb{R}^2 \text{ is a translation}$$

So first we will examine the orthogonal group $O(\mathbb{R}^2 \to \mathbb{R}^2)$. In Lemma 1.3.3 we saw that every $\tau \in O(\mathbb{R}^2 \to \mathbb{R}^2)$ is a linear isomorphism, therefore it is defined by a matrix:

$$M_\tau = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \text{ with non-zero determinant } a_1a_4 - a_2a_3 \neq 0$$

and τ can be written as:

$$\tau\big((x,y)\big)^T = M_\tau \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1x + a_2y \\ a_3x + a_4y \end{pmatrix}$$

Using Theorem 1.3.1 one can see that τ preserves the norms $||\cdot||_{\mathbb{R}^2} = |\cdot|$ (that are induced from their respective inner products) and thus:

$$|(x,y)| = |\tau(x,y)| \Rightarrow x^2 + y^2 = (a_1^2 + a_3^2)x^2 + 2(a_1a_2 + a_3a_4)xy + (a_2^2 + a_4^2)y^2 \Rightarrow \begin{cases} a_1^2 + a_3^2 = 1 \\ a_1a_2 + a_3a_4 = 0 \\ a_2^2 + a_4^2 = 1 \end{cases}$$

This way some restrictions on a_1, a_2, a_3, a_4 have been made:

   i. $a_1a_4 - a_2a_3 \neq 0$

   ii. $a_1^2 + a_3^2 = 1$

   iii. $a_1a_2 + a_3a_4 = 0$

   iv. $a_2^2 + a_4^2 = 1$

These restrictions will determine the α, β elements. Lets first assume that $a_1 = 0$. Then from i. and iii. $a_2a_3 \neq 0$ and $a_3a_4 = 0$, which shows that $a_1 = a_4 = 0$ and $a_2, a_3 \neq 0$. Moreover, from i. and iv., $a_2, a_3 \in \{\pm 1\}$, and matrix M_τ becomes:

$$M_\tau = \begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$$

If $a_1 \neq 0$, then from iii. we can write $a_2 = -a_3a_4/a_1$. If we now let $\lambda = a_4/a_1$, we notice that $a_2 = -\lambda a_3$ and $a_4 = \lambda a_1$. From ii. and iv., $a_1^2 + a_3^2 = 1 = \lambda^2(a_1^2 + a_3^2) \Rightarrow \lambda \in \{\pm 1\}$. Depending on the value of λ:

$$M_\tau = \begin{pmatrix} a_1 & -a_3 \\ a_3 & a_1 \end{pmatrix} \text{ if } \lambda = -1, \text{ and } M_\tau = \begin{pmatrix} a_1 & a_3 \\ a_3 & -a_1 \end{pmatrix} \text{ if } \lambda = -1$$

In any case, there exist $\alpha, \beta \in \mathbb{R}$ such that:

$$M_\tau = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \text{ or } M_\tau = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$$

with $\alpha^2 + \beta^2 = 1$. To end this proof we consider the translations:

$$\mu\big((x,y)\big)^T = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix}, \ \gamma, \delta \in \mathbb{R}$$

and from Theorem 1.3.2 follows that every $\sigma \in I(\mathbb{R}^2 \to \mathbb{R}^2)$ has the form:

$$\sigma\big((x,y)\big)^T = \begin{cases} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \\ \quad or \\ \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \end{cases}$$

where $\alpha^2 + \beta^2 = 1$, $\gamma, \delta \in \mathbb{R}$. $\qquad\qquad\square$

Corollary 1.3.2. *The following notation is as in Theorem 1.3.3. Matrix:*

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

can be written as:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for some angle θ. Also, matrix:

$$\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$$

can be written as:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

for some angle θ. So every isometry $\sigma \in I(\mathbb{R}^2 \to \mathbb{R}^2)$ is some combination of rotations, reflections and translations.

Proof. Is immediate from the fact that $a^2 + b^2 = 1$ and Theorem 1.3.3. $\qquad\square$

Corollary 1.3.2 shows what we wanted to show, for the two dimensional case: "every isometry is some combination of rotations, reflections and translations". What we will show next is that in every Euclidean space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ isometries are combinations of rotations, reflections and translations, if rotations are defined correctly.

In the following results $\mathbb{R}^n$ will be considered with its usual inner product. Also, T will be the form of said inner product.

Lemma 1.3.4. *Let $\tau \in O(\mathbb{R}^2 \to \mathbb{R}^n)$ and M_τ its matrix. If M_τ has a real eigenvalue λ, then $\lambda \in \{\pm 1\}$.*

Proof. Indeed, if v^T is an eigenvector to eigenvalue λ, then (by the orthogonality of τ):

$$\tau(v) = \lambda v^T \Rightarrow T(\tau(v)) = \lambda^2 T(v) = \lambda^2 T(\tau(v)) \Rightarrow \lambda^2 = 1 \text{ or } \lambda = 0$$

The $\lambda = 0$ case cannot hold: If $\mathscr{X}_{M_\tau}$ is the characteristic polynomial of M_τ, then $\mathscr{X}_{M_\tau}(0) = \det(M_\tau - 0 \cdot \mathrm{Id}_n) = \det M_\tau \neq 0$, so 0 cannot be an eigenvalue. The inequality holds because of Lemma 1.3.3. $\quad\square$

Lemma 1.3.5. *Let $\tau \in O(\mathbb{R}^n \to \mathbb{R}^n)$ and M_τ its matrix. If M_τ has a complex eigenvalue $\lambda = a + bi$ with eigenvector $v^T = x^T + iy^T$, then:*

$$\tau(x) = ax^T - by^T \text{ and } \tau(y) = bx^T + ay^T$$

Proof. Because λ is an eigenvalue of M_τ with eigenvector v:

$$\tau(v) = \lambda v^T \Rightarrow M_\tau \cdot v^T = \lambda v^T \Rightarrow (M_\tau - \lambda \cdot \mathrm{Id}_n) \cdot v^T = 0 \tag{1.4}$$

Now (1.4) can be seen as an linear system with the coordinates of v^T as unknowns. This system has determinant:

$$\mathscr{X}_{M_\tau}(\lambda) = \det(M_\tau - \lambda \cdot \mathrm{Id}_n) = 0$$

(because λ is an eigenvalue), therefore there exists non-zero eigenvector $v^T = x^T + iy^T$. So from (1.4) we have:

$$\left(M_\tau - (a + ib) \cdot \mathrm{Id}_n\right) \cdot (x^T + iy^T) = 0 \Rightarrow$$

$$\left[(M - a \cdot \mathrm{Id}_n) \cdot x^T + b \cdot \mathrm{Id}_n \cdot y^T\right] + i\left[(M - a \cdot \mathrm{Id}_n) \cdot y^T - b \cdot \mathrm{Id}_n \cdot x^T\right] = 0$$

which shows that both real and imaginary parts equal zero. Therefore:

$$\tau(x) = M_\tau \cdot x^T = ax^T - by^T \text{ and } \tau(y) = M_\tau \cdot y^T = bx^T + ay^T$$

$\square$

Lemma 1.3.6. *Let $\tau \in O(\mathbb{R}^n \to \mathbb{R}^n)$. There exists a $\tau-$invariant subspace V -with dimension 1 or 2- such that its orthocomplement $V^\perp$ is also $\tau-$invariant. Also, there exists an orthonormal base of V consisting of eigenvectors of M_τ, where M_τ is the matrix of τ.*

Proof. We first consider the case where the matrix M_τ of τ has a real eigenvalue $\lambda \in \mathbb{R}$. If v^T is an eigenvector to that eigenvalue, the vector $\omega^T = v^T/|v|$ is (trivially) an eigenvector of M_τ with length 1. Note that v can be chosen non-zero -throughout this proof we will choose non-zero eigenvectors without mentioning it every time. The set $\{\omega^T\}$ is an orthonormal base of $V = \text{span}\{\omega^T\}$ and V is $\tau-$invariant. We will now show that $V^\perp$ is also $\tau-$invariant: an orthogonal function τ preserves the angles (because it preserves the inner product), thus if u^T is an element of $V^\perp$, $u^T \perp \omega^T \Rightarrow \tau(u^T) \perp \tau(\omega^T)$, which shows that $\tau(u^T) \in V^\perp$. So in this case the result is proven.

If M_τ has no real eigenvalue, then it has a complex eigenvalue $\lambda = a + ib \in \mathbb{C}\backslash\mathbb{R}$, with eigenvector $v^T = x^T + iy^T$. Using Lemma 1.3.5 we write:

$$\tau(x) = ax^T - by^T \text{ and } \tau(y) = bx^T + ay^T$$

and we notice that x^T, y^T must be linearly independent (otherwise, a real eigenvalue should exist). Lemma 1.3.5 also shows that subspace $V = \text{span}\{x^T, y^T\}$ is $\tau-$invariant and, in our case, has dimension 2. By the Gramm-Schmidt procedure we can orthonormalise the basis $\{x^T, y^T\}$ of V, finding a new base:

$$\left\{\omega^T = \frac{x^T}{|x|}, \ \psi^T = \left. y^T - \frac{\langle x, y\rangle \cdot x^T}{|x|} \middle/ \left| y^T - \frac{\langle x, y\rangle \cdot x^T}{|x|} \right| \right.\right\}$$

It remains to be shown that $V^\perp$ is $\tau-$invariant. Again, if $u \in V^\perp$, then $\tau(u) \perp \tau(\omega)$ and $\tau(u) \perp \tau(\psi)$. Because $\tau(\omega)$, $\tau(\psi)$ generate V, it follows that $\tau(u) \in V^\perp$. The Lemma has been proven. $\square$

We now prove the last Lemma needed for our "big theorem", which shows (in a way) the form of isometries we sought.

Lemma 1.3.7. *Let $\tau \in O(\mathbb{R}^n \to \mathbb{R}^n)$. We suppose an eigenvalue $\lambda \in \mathbb{C}\backslash\mathbb{R}$ with eigenvector $v^T = x^T + iy^T$ exists, and we consider a $2-$dimensional $\tau-$invariant subspace V with orthonormal basis $\{\omega^T, \psi^T\}$, as in Lemma 1.3.6. Then the function $\tau|_V : V \to \mathbb{R}$ is a $2-$dimensional rotation, meaning $\tau|_V$ has a matrix of the form:*

$$M_{\tau|_V} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for some basis of $\mathbb{R}^n$. Note that the $2-$dimensional space V can be considered (even if it is not) a subset of $\mathbb{R}^2$. If one does not want to make this assumption, they need to work with matrices of the form:

$$M_{\tau|_V} = \begin{pmatrix} 1 & & & & & & & & \mathbf{O} \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos\theta & -\sin\theta & & & \\ & & & \sin\theta & \cos\theta & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ \mathbf{O} & & & & & & & 1 \end{pmatrix}$$

which is more difficult to write.

Proof. Using Lemma 1.3.5 we can write:

$$\tau(x) = ax^T - by^T \text{ and } \tau(y) = bx^T + ay^T$$

so $\tau|_V$ has matrix:

$$M_{\tau|_V} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ with } \det M_\tau = a^2 + b^2 > 0$$

in basis $\{x, y\}$ of V. From the proof of Lemma 1.3.6 we observe that ω^T, v^T depend linearly on x^T, y^T, so in basis $\{\omega^T, \psi^T\}$ matrix M_τ has determinant $\det M_{\tau|_V} = a^2 + b^2 > 0$. From Corollary 1.3.2 follows that $M_{\tau|_V}$ is a rotation matrix. $\square$

Theorem 1.3.4. *(Isometries of $\mathbb{R}^n$) Let $\tau \in O(\mathbb{R}^n \to \mathbb{R}^n)$. There exists an orthonormal basis of $\mathbb{R}^n$ such that τ is defined by a matrix of the form:*

$$W = \begin{pmatrix} -1 & & & & & & \mathbf{O} \\ & 1 & & & & & \\ & & R(\theta_1) & & & & \\ & & & R(\theta_2) & & & \\ & & & & \ddots & & \\ \mathbf{O} & & & & & R(\theta_r) \end{pmatrix}$$

where $R(\theta_k)$ are rotation matrices, by angle θ_k. It is also possible some of the ± 1 to not appear, so $2r + 2 = n$ or $2r + 1 = n$ or $2r = n$. Using Theorem 1.3.2, every isometry $\sigma \in O(\mathbb{R}^n \to \mathbb{R}^n)$ can be written as:

$$\sigma\big((z_1, \cdots, z_n)\big) = \begin{pmatrix} -1 & & & & & & \mathbf{O} \\ & 1 & & & & & \\ & & R(\theta_1) & & & & \\ & & & R(\theta_2) & & & \\ & & & & \ddots & & \\ \mathbf{O} & & & & & R(\theta_r) \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \\ a_n \end{pmatrix}$$

in some orthonormal basis of $\mathbb{R}^n$, where $a_k \in \mathbb{R}$.

Proof. To prove this theorem we will show the first part of the theorem (the other, as we state, follows from Theorem 1.3.2). The proof is quite long, so we will work in steps. Let W_0 be the matrix of τ in some basis $\mathbb{B}_0$ of $\mathbb{R}^n$.

Step I: Let λ_1 be a real eigenvalue (if a real eigenvalue exists at all), which form Lemma 1.3.4 is ± 1. We consider an (non-zero) eigenvector v_1^T to that eigenvalue and we consider the subspace $V = \text{span}\{v_1^T\}$. From Lemma 1.3.6, V_1 is $\tau-$invariant, its orthocomplement $V_1^\perp$ is $\tau-$invariant, and there exists an orthonormal basis $\mathbb{B}_1 = \{\omega_1^T\}$ of V_1. If we consider $\mathbb{G}_1$ a basis of $V_1^\perp$, matrix W_0 in basis $\mathbb{B}_1 \cup \mathbb{G}_1$ becomes:

$$W_1 = \begin{pmatrix} \lambda_1 & & & & & \mathbf{O} \\ & {}^1\mu_{1,1} & {}^1\mu_{1,2} & \cdots & {}^1\mu_{1,n-1} \\ & {}^1\mu_{2,1} & {}^1\mu_{2,2} & \cdots & {}^1\mu_{2,n-1} \\ & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & {}^1\mu_{n-1,1} & {}^1\mu_{n-1,2} & \cdots & {}^1\mu_{n-1,n-1} \end{pmatrix}$$

Step II: The function with matrix $({}^1\mu_{i,j})_{i,j \leqslant n-1}$ is orthogonal, because τ is orthogonal and $V_1^\perp$ is $\tau-$invariant. So we can continue the process described in Step I, until after some iterations matrix $({}^\ell\mu_{i,j})_{i,j \leqslant n-\ell}$ does not have any real eigenvalues left. In a base $\big(\bigcup_{k=1}^\ell \mathbb{B}_k\big) \cup \mathbb{G}_\ell$ the matrix of W_1 will become:

$$W_\ell = \begin{pmatrix} \lambda_1 & & & & & & & \mathbf{O} \\ & \ddots & & & & & & \\ & & \lambda_\ell & & & & & \\ & & & {}^\ell\mu_{1,1} & {}^\ell\mu_{1,2} & \cdots & {}^\ell\mu_{1,n-\ell} \\ & & & {}^\ell\mu_{2,1} & {}^\ell\mu_{2,2} & \cdots & {}^\ell\mu_{2,n-\ell} \\ & & & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & & & {}^\ell\mu_{n-\ell,1} & {}^\ell\mu_{n-\ell,2} & \cdots & {}^\ell\mu_{n-\ell,n-\ell} \end{pmatrix}$$

Step III: After Step II no real eigenvalues remain. So we consider the complex eigenvalues -let $\zeta_1 = \varepsilon_1 + i\epsilon_1 \in \mathbb{C}\backslash\mathbb{R}$ be a complex eigenvalue with eigenvector $u_1 = x_1 + iy_1$. From the proof of Lemma 1.3.6, x_1, y_1 are linearly independent, so $U_1 = \operatorname{span}\{x_1, y_1\}$ is a two dimensional subspace. Also, U_1, $V_\ell^\perp \cap U_1^\perp$ are $(^\ell\mu_{i,j})_{i,j\leqslant n-\ell}-$ invariant and there exists an orthonormal basis $\mathbb{E}_1 = \{\chi_1, \psi_1\}$ of U_1. If $\mathbb{G}_{\ell+1}$ is a basis of $V_\ell^\perp \cap U_1^\perp$, then in basis $\left(\bigcup_{k=1}^\ell \mathbb{B}_k\right) \cup \mathbb{E}_1 \cup \mathbb{G}_{\ell+1}$ matrix W_ℓ becomes:

$$W_{\ell+1} = \begin{pmatrix} \lambda_1 & & & & & & & & \mathbf{O} \\ & \ddots & & & & & & & \\ & & \lambda_\ell & & & & & & \\ & & & {}^1\xi_{1,1} & {}^1\xi_{1,2} & & & & \\ & & & {}^1\xi_{2,1} & {}^1\xi_{2,2} & & & & \\ & & & & & {}^{\ell+1}\mu_{1,1} & {}^{\ell+1}\mu_{1,2} & \cdots & {}^{\ell+1}\mu_{1,n-\ell-1} \\ & & & & & {}^{\ell+1}\mu_{2,1} & {}^{\ell+1}\mu_{2,2} & \cdots & {}^{\ell+1}\mu_{2,n-\ell-1} \\ & & & & & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & & & & & {}^{\ell+1}\mu_{n-\ell-1,1} & {}^{\ell+1}\mu_{n-\ell-1,2} & \cdots & {}^{\ell+1}\mu_{n-\ell-1,n-\ell-1} \end{pmatrix}$$

Using Lemma 1.3.7, matrix $({}^1\xi_{i,j})_{i,j\leqslant 2}$ is a 2-dimensional rotation, by angle φ_1. We denote this matrix by $R(\varphi_1)$ and we write:

$$W_{\ell+1} = \begin{pmatrix} \lambda_1 & & & & & & & \mathbf{O} \\ & \ddots & & & & & & \\ & & \lambda_\ell & & & & & \\ & & & R(\varphi_1) & & & & \\ & & & & {}^{\ell+1}\mu_{1,1} & {}^{\ell+1}\mu_{1,2} & \cdots & {}^{\ell+1}\mu_{1,n-\ell-1} \\ & & & & {}^{\ell+1}\mu_{2,1} & {}^{\ell+1}\mu_{2,2} & \cdots & {}^{\ell+1}\mu_{2,n-\ell-1} \\ & & & & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & & & & {}^{\ell+1}\mu_{n-\ell-1,1} & {}^{\ell+1}\mu_{n-\ell-1,2} & \cdots & {}^{\ell+1}\mu_{n-\ell-1,n-\ell-1} \end{pmatrix}$$

Step IV: We continue the process described in Step III until no eigenvalues remain. After some amount s of iterations, we get, in basis $\mathscr{B} = \left(\bigcup_{k=1}^\ell \mathbb{B}_k\right) \cup \left(\bigcup_{k=1}^s \mathbb{E}_k\right)$, the matrix:

$$W_{\ell+s} = \begin{pmatrix} \lambda_1 & & & & & \mathbf{O} \\ & \ddots & & & & \\ & & \lambda_\ell & & & \\ & & & R(\varphi_1) & & \\ & & & & \ddots & \\ \mathbf{O} & & & & & R(\varphi_s) \end{pmatrix}$$

with $\lambda_k \in \{\pm 1\}$, $R(\varphi_k)$ are $2-$dimensional rotation matrices by φ_k, and $\ell + 2s = n$. $W_{\ell+s}$ is very "close" to what we want to find.

Step V: By permuting the elements of basis $\mathscr{B}$ we can bring "together", in pairs, the 1 and the -1 elements of $W_{\ell+s}$. With the correct permutations, matrices in block form:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

appear in $W_{\ell+s}$. These matrices are rotation matrices $R(0)$ and $R(\pi)$. If some ± 1 remain, we permute once again basis $\mathscr{B}$ so they move to $(1,1)$ and $(2,2)$ positions of $W_{\ell+s}$. Finally, we change our angle

notation from 0, π, φ_k to θ_k, and in the permuted basis $\mathscr{B}$ matrix $W_{\ell+s}$ becomes:
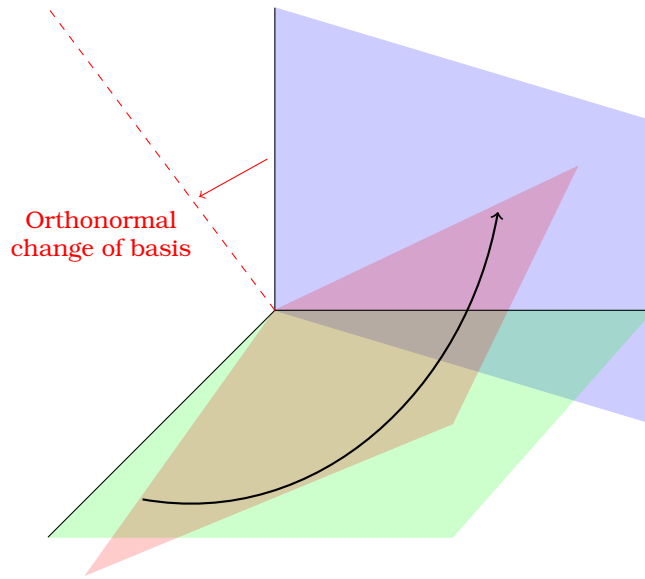
$$W = \begin{pmatrix} -1 & & & & & \mathbf{O} \\ & 1 & & & & \\ & & R(\theta_1) & & & \\ & & & R(\theta_2) & & \\ & & & & \ddots & \\ \mathbf{O} & & & & & R(\theta_r) \end{pmatrix}$$

which is what we wanted to prove. $\qquad\square$

Corollary 1.3.3. *It is intuitive (for example in the $3-$dimensional case) to imagine rotation in many dimensions as sequential $2-$dimensional rotations. So one can call every function with matrix:*

$$\begin{pmatrix} 1 & & & & \mathbf{O} \\ & R(\theta_1) & & & \\ & & R(\theta_2) & & \\ & & & \ddots & \\ \mathbf{O} & & & & R(\theta_r) \end{pmatrix}$$

a rotation in many dimensions (the 1 element can appear or not).

A $3-$dimensional rotation $\begin{pmatrix} 1 & \mathbf{O} \\ \mathbf{O} & R(\theta) \end{pmatrix}$, done on the red plane.

Definition 1.3.3 (Rigid motions). *We call rigid motion every isometry $\sigma \in I(\mathbb{R}^n \to \mathbb{R}^n)$ that does not contain any kind of reflection. According to Theorem 1.3.4, rigid motions are some combination of rotations and translations.*

Differentiation

2.1 Derivatives, directed and partial derivatives

The study of derivatives in many dimensions is, in its core, a problem of finding a "good" definition. The generalisation of single variable derivative in many dimensions faces obstacles that are integrated to the definition of the single variable case, so a "special treatment" is needed. Fist, lets remember the known derivative definition.

Definition 2.1.1 (Derivative - Single variable). *Let $f : A \to \mathbb{R}$ be a real valued function defined on the open set $A \subseteq \mathbb{R}$, and $x_0 \in A$ be a point of A. We say that f is differentiable at x_0 if the limit:*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists and is a real number. We also define the derivative of f at x_0 be the before-mentioned limit (we use the symbol $f'(x_0)$).

Of course one usually generalises this definition by noticing that A does not need to be open. The derivative is a local property of functions and, thus, it suffices that x_0 is located in an "open region" of A, or that it can be reached through A.

Now, for the many dimensions, note that the transition cannot be done so easily. Even in $\mathbb{R}^2$, the quotient:

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

is not defined, as the usual algebraic structure of $\mathbb{R}^2$ does not have multiplicative inverses. It is possible to enrich the $\mathbb{R}^2$ structure so that division is possible. By writing every $(x, y) \in \mathbb{R}^2$ in its complex form $x + iy$ and using the structure of $\mathbb{C}$, it is possible to define the derivative of $f : \mathbb{C} \to \mathbb{C}$ at x_0 using the limit:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}, \quad \text{where } x_0, h \in \mathbb{C}$$

The study of such differentiable functions $f : \mathbb{C} \to \mathbb{C}$ of the complex plane (of the "holomorphic funtions") is interesting but it does not give a "good" definition, that can be extended in more dimensions.

One constantly needs extensions of the real numbers (as complex numbers are). In just four dimensions, the extensions of the reals -the quaternions $\mathbb{H}$- have lost some of their structure -they are a division ring and an integer domain, but not a field. So this "dream" of extending the derivative definition this way fails.

Here is where analysis and geometry come. First, a geometer might have given a physical meaning to the derivative, making it possible to extend the definition in curves. So, let $\gamma : I \to \mathbb{R}^n$ be an n-dimensional curve. By writing $\gamma = (\gamma_1, \cdots, \gamma_n)$, where $\gamma_k : I \to \mathbb{R}$, $k \in [n]$ are the "coordinate"

functions of γ, we can say that γ has "speed" to a point $t_0 \in I$ if its coordinates have "speed". This notion of speed is widely used in geometry and physics. Then, if for every $k \in [n]$ the limit:

$$\lim_{h \to 0} \frac{\gamma_k(t_0 + h) - \gamma_k(t_0)}{h}$$

exists (in the reals), the derivative of γ at t_0 can be defined as:

$$\gamma'(t_0) = \left(\lim_{h \to 0} \frac{\gamma_k(t_0 + h) - \gamma_k(t_0)}{h} \right)_{k \in [n]}$$

or equivalently:

$$\gamma'(t_0) = \lim_{h \to 0} \left[\frac{1}{h} \big(\gamma(t_0 + h) - \gamma(t_0) \big) \right]$$

This definition in curves also has another geometrically intuitive property:

Proposition 2.1.1. *Let $\gamma : I \to \mathbb{R}^n$ be a differentiable curve at $t_0 \in I$. Then the tangent line at t_0 is the line:*

$$\varepsilon : \gamma(t_0) + s \cdot \gamma'(t_0), \quad \text{as } s \in \mathbb{R}$$

This means that the "limit position" of lines passing through $\gamma(t_0)$, $\gamma(t_0 + h)$ (as $h \to 0$) is the before-mentioned line.

Furthermore, if γ has continuous derivative at t_0, the before-mentioned line is also the line that is the limit of lines approximating γ, which are passing through any two (different) points $\gamma(t_0 - \ell)$, $\gamma(t_0 + \ell)$ ($h, \ell \to 0$).

Proof. Lets approximate the tangent line by the line passing through $\gamma(t_0)$, $\gamma(t_0 + h)$ for sufficiently small h. In this proof we will first approximate the tangent with these kind of lines.

The approximation that passes through $\gamma(t_0)$, $\gamma(t_0 + h)$ is a line of the form:

$$\gamma(t_0) + s \cdot \big(\gamma(t_0 + h) - \gamma(t_0) \big), \quad \text{where } s \in \mathbb{R}$$

because it is parallel to vector $\gamma(t_0 + h) - \gamma(t_0)$. One might want to take limits to find the tangent line, thought this is not possible to do so, as $\gamma(t_0 + h) - \gamma(t_0) \to 0_{\mathbb{R}^3}$. So instead we write:

$$\gamma(t_0) + s \cdot \frac{1}{h} \big(\gamma(t_0 + h) - \gamma(t_0) \big), \quad \text{where } s \in \mathbb{R}$$

because it is also parallel to vector $1/h \cdot \gamma(t_0 + h) - \gamma(t_0)$. Now, by taking limits as $h \to 0$:

$$\gamma(t_0) + s \cdot \gamma'(t_0), \quad \text{where } s \in \mathbb{R}$$

We will also see the more general case where the lines pass through $\gamma(t_0 - \ell)$, $\gamma(t_0 + h)$. We again approximate our tangent by the lines:

$$\gamma(t_0 - \ell) + s \cdot \frac{1}{h + \ell} \big(\gamma(t_0 + h) - \gamma(t_0 - \ell) \big), \quad \text{where } s \in \mathbb{R}$$

or equivalently:

$$\gamma(t_0 - \ell) + s \cdot \left(\frac{\gamma_k(t + h) - \gamma_k(t - \ell)}{h + \ell} \right)_{k \in [n]}$$

Notice that here we cannot just take limits as in our previous case. Instead, we use the mean value theorem (of one variable), finding $\xi_k \in (\min\{t_0 - \ell, t_0 + h\}, \max\{t_0 - \ell, t_0 + h\})$ such that:

$$\gamma'_k(\xi_k) = \frac{\gamma_k(t + h) - \gamma_k(t - \ell)}{h + \ell}$$

Our approximation of the tangent line now becomes:

$$\gamma(t_0 - \ell) + s \cdot \gamma'(\xi), \quad \text{where } s \in \mathbb{R} \text{ and } \xi = (\xi_k)_{k \in [n]}$$

These ξ, even if it is not "labeled", depend on h, ℓ -do not forget that its coordinates ξ_k belong to an interval that depends on h, ℓ. Now we let $h, \ell \to 0$, $h \neq -\ell$ and we get $\xi \to t_0$ (by definition). Then, by the continuity of the derivative, the approximations have the limit:

$$\gamma(t_0) + s \cdot \gamma'(t_0), \quad \text{where } s \in \mathbb{R}$$

Thus the tangent line in this case is the before-mentioned line, and the proof is complete. $\square$

This idea of "tangent lines", or generally "tangent shapes", will in fact give our definition of the derivative. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$, which can be visualised as a surface in $\mathbb{R}^3$ (a strange surface, in most cases). If we assume f has some kind of differentiability, it should get approximated by some "multidimensional" lines. These "lines" -in our case- are planes (affine linear spaces).
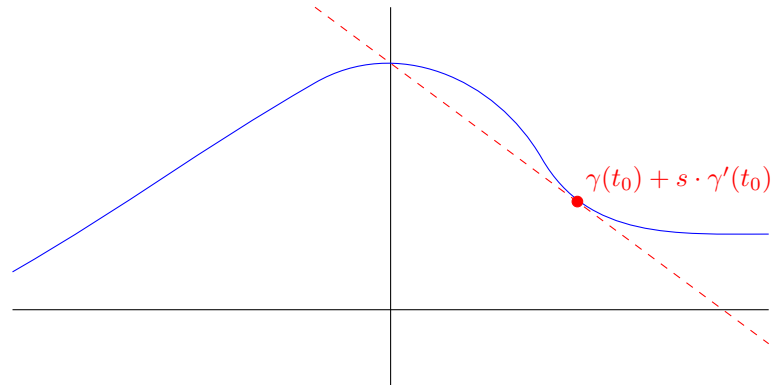
Tangent line on a graph of $\gamma(t) = (t, f(t))$ (where $t \in \mathbb{R}$).

$$\left\{ \left(x + x_0, f(x_0) + M_{f,x_0} x \right) \mid x \in \mathbb{R}^2 \right\}$$
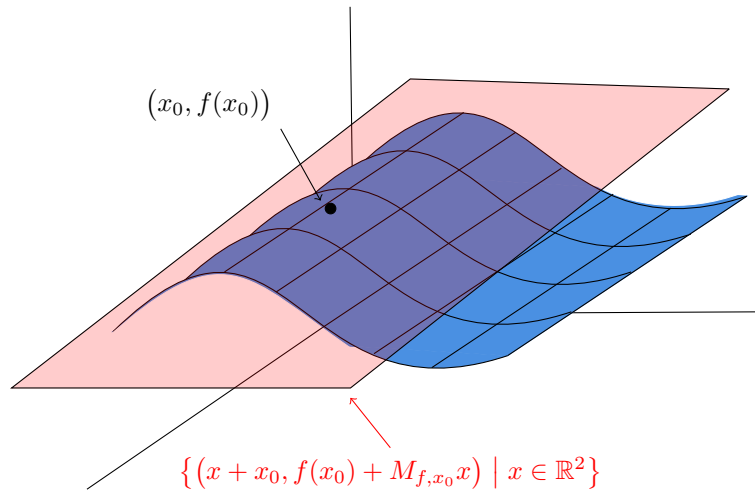
Tangent plane to a surface $\left(x, f(x) \right)$ (where $x \in \mathbb{R}^2$).

We give the following definition:

Definition 2.1.2 (Derivative - Many variables). *Let $f : A \to \mathbb{R}^m$ be a function, $A \subseteq \mathbb{R}^n$ an open set and $x_0 \in A$. We say that f is differentiable at x_0 if there exists a linear function M_{f,x_0} such that:*

$$\lim_{h \to 0_{\mathbb{R}^n}} \frac{|f(x_0 + h) - f(x_0) - M_{f,x_0} h|}{|h|} = 0$$

In this case, $f(x + x_0)$ is approximated by $f(x_0) + M_{f,x_0} x$, and the graph of f is approximated by the affine linear space $\left\{ \left(x + x_0, f(x_0) + M_{f,x_0} x \right) \mid x \in \mathbb{R}^n \right\}$, near x_0. The linear function M_{f,x_0} is called the derivative of f at x_0. We also say that f is differentiable if for every $x_0 \in A$ is differentiable.

Here M_{f,x_0} is linear and thus we use, instead of $M_{f,x_0}(x)$, $M_{f,x_0} x$ (in a way, we use the symbol $M_{f,x_0} x$ as if M_{f,x_0} was a matrix).

The derivative of f at x_0 (as in Definition 2.1.2) is a linear function. If we gather all of these

M_{f,x_0}, $x_0 \in A$, we can make another function Df, such that $(Df)(x_0) = M_{f,x_0}$, which is called the derivative of f (notice that the "at x_0" is not used). Now, in the one variable case, we remembered that the derivative f' can give us an operator in the differentiable functions. We saw that the function:

$$\frac{d}{dx} : \mathcal{D}(\mathbb{R} \rightharpoonup \mathbb{R}) \to (\mathbb{R} \rightharpoonup \mathbb{R}), \text{ with } \frac{d}{dx}(f) = \frac{df}{dx} = f'$$

is a linear operator. $\mathcal{D}(\mathbb{R} \rightharpoonup \mathbb{R})$ is the set of all differentiable functions f with domain $A \subseteq \mathbb{R}$ and image $f(A) \subseteq \mathbb{R}$, and $(\mathbb{R} \rightharpoonup \mathbb{R})$ is the set of all functions g with domain $B \subseteq \mathbb{R}$ and image $g(B) \subseteq \mathbb{R}$. We will also show that:

$$D : \mathcal{D}(\mathbb{R}^n \rightharpoonup \mathbb{R}^m) \to (\mathbb{R}^n \rightharpoonup \mathbb{R}^m), \text{ with } D(f) = Df$$

is a linear operator (which also gives meaning to the symbol Df). $\mathcal{D}(\mathbb{R}^n \rightharpoonup \mathbb{R}^m)$ is the set of all differentiable functions f with domain $A \subseteq \mathbb{R}^n$ and image $f(A) \subseteq \mathbb{R}^m$, and $(\mathbb{R}^n \rightharpoonup \mathbb{R}^m)$ is the set of all functions g with domain $B \subseteq \mathbb{R}^n$ and image $g(B) \subseteq \mathbb{R}^m$.

First, we need a definition.

Definition 2.1.3 (Sets of differentiable functions). *We define:*

- $\mathcal{D}(S \to T) := \{f : S \to T \mid f \text{ is differentiable}\}$

- $\mathcal{D}^k(S \to T) := \{f : S \to T \mid f \text{ is differentiable } k - times\}, k \in \overline{\overline{\mathbb{N}}}$

- $\mathcal{D}(S \rightharpoonup T) := \{f : A \to B \mid f \text{ is differentiable and } A \subseteq S, \ B \subseteq T\}$

- $\mathcal{D}^k(S \rightharpoonup T) := \{f : A \to B \mid f \text{ is differentiable } k - times \text{ and } A \subseteq S, \ B \subseteq T\}, k \in \overline{\overline{\mathbb{N}}}$

S and T are subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

It is easy to check that all of these sets are $\mathbb{R}-$vector spaces. For example $\mathcal{D}(S \to T)$ is an $\mathbb{R}-$vector space. To show this, lets show that for every $f, g \in \mathcal{D}(S \to T)$ we have $f + g \in \mathcal{D}(S \to T)$ (all the other axioms can be also be proved relatively easily).

If Df and Dg are the derivatives of f and g, then:

$$0 = \lim_{h \to 0_{\mathbb{R}^n}} \frac{|f(x_0 + h) - f(x_0) - (Df)(x_0)h|}{|h|} + \lim_{h \to 0_{\mathbb{R}^n}} \frac{|g(x_0 + h) - g(x_0) - (Dg)(x_0)h|}{|h|} =$$

$$= \lim_{h \to 0_{\mathbb{R}^n}} \frac{|(f + g)(x_0 + h) - (f + g)(x_0) - (Df + Dg)(x_0)h|}{|h|}$$

So $f + g$ is differentiable and $D(f + g) = Df + Dg$ (as we have shown $\big(D(f + g)\big)(x_0) = (Df)(x_0) + (Dg)(x_0)$ for some random $x_0 \in S$).

The previous proof also shows that:

Proposition 2.1.2. *The operator:*

$$D : \mathcal{D}(\mathbb{R}^n \rightharpoonup \mathbb{R}^m) \to (\mathbb{R}^n \rightharpoonup \mathbb{R}^m), \text{ with } D(f) = Df$$

is linear. This, as we said previously, gives meaning to the symbol Df we use.

The differentiability of a function $\big(x, f(x)\big)$, $x \in \mathbb{R}^n$, at a point x_0 is interlinked (in a way) with finding a tangent plane. One has to find a matrix:

$$(a_1, \cdots, a_n)$$

so n unknowns have to be determined. Is there an easy way to find those a_k, $k \in [n]$?

Lets consider the case of a function $\big(x, f(x)\big)$, where $x \in \mathbb{R}^2$ and $x_0 = (\underset{1}{\pi}x_0, \underset{2}{\pi}x_0) \in \mathbb{R}^2$ is a point. A tangent plane is defined by exactly to (different) lines that are subsets of it. If we restrict our

function on the lines $\{(t, \underset{2}{\pi}x_0) \mid t \in \mathbb{R}\}$, $\{(\underset{1}{\pi}x_0, t) \mid t \in \mathbb{R}\}$ we get two curves that meet at x_0. If the plane we seek is tangent to our function, it should also be tangent to our curves, therefore the tangent lines of these curves are subsets of the plane. It seems like we can find the tangent plane by finding the derivatives of "simple", functions of one variable, and in that case:

$$(a_1, a_2) = \left(\frac{df(\cdot, \underset{2}{\pi}x_0)}{d\underset{1}{\pi}x}(\underset{1}{\pi}x_0), \ \frac{df(\underset{1}{\pi}x_0, \cdot)}{d\underset{2}{\pi}x}(\underset{2}{\pi}x_0) \right)$$

(where we can write $x = (\underset{1}{\pi}x, \ \underset{2}{\pi}x)$). The first function is function f where its second variable is considered constant, and the second function is function f where the first variable is considered constant. We will explore this idea further later, first we need the idea of "partial derivatives".

$$\frac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_m)}, \ \frac{\partial \underset{i_1 \cdots i_m}{\widehat{\pi}^n} f}{\partial \underset{i_1 \cdots i_m}{\widehat{\pi}^n} x}, \ \amalg_{i_1 \cdots i_m}^n f, \ \pi_I f, \ \prod, \ \coprod, \ d\pi_{i_1 \cdots i_m}^n f$$

$$\underset{i_1 \cdots i_m}{\pi^b} f, \ \underset{i_1 \cdots i_m}{\widehat{\pi}^k}, \ \underset{a}{\amalg^b}, \ \underset{i_1 \cdots i_m}{\widehat{\amalg}^n}$$

2.2 Exercises

Exercise 2.1. _Show that the "continuous derivative at t_0" hypothesis is needed in Proposition 2.1.1._
The problem lies when we try to approximate the tangent line by not having point $\gamma(t_0)$ as one point of
the approximation lines. If that hypothesis is not taken into account, tangent lines get approximated by
lines that pass through $\gamma(t_0)$ and some other point on γ.

Hint: Consider the function:

$$\gamma(t) = \begin{cases} x^2 \sin(1/x), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

which has discontinuous derivative:

$$\gamma'(t) = \begin{cases} 2x \sin(1/x) - \cos(1/x), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

Notation

Bibliography

In English

[KP] Krantz G. S., Parks R. H.: *The Implicit Function Theorem - History, Theory, and Applications*, Birkhäuser, 2003.

[Mu] Munkres R. J.: *Analysis on Manifolds*, Addison-Wesley Publishing Company, 1991.

[Ru] Rudin W.: *Principles of Mathematical Analysis*, McGrawl-Hill Inc, 1976.

[Se] Seeley R. T.: *Extension of C^∞ functions defined in a half space*, American Mathematical Society, 1963.

In Greek

[BG] Bitsouni V., Gialelis N.: *Geometric Analysis*, National and Kapodestrian University of Athens, 2021 (Lecture Notes).

[Fr] Fragkos A.: *Introduction to Plane Hyperbolic Geometry*, Academia.edu, 2021.

Index