Contents

Chapter 1

The $\mathbb{R}^n$ metric space

1.1 The $\mathbb{R}$ set

1.1.1 Properties of real numbers

Let us remember some of the most important properties of $\mathbb{R}$. We saw in [ ] that $(\mathbb{R}, +,)$ with the order relation $\leq$ forms an ordered field that satisfies the following conditions.

1. $x \leq y \Rightarrow x + z \leq y + z$
2. $0 \leq x, y \Rightarrow 0 \leq xy$

The order relation $\leq$ allows us to classify the real numbers between positive, negative and 0. With that, we defined the absolute value function.

$$x = \begin{cases} -x, x \geq 0 \\ x, x < 0 \end{cases} \tag{1.1}$$

Which satisfies:

1. $|x| \geq 0$
2. $|x|\, 0 \Leftrightarrow x = 0$
3. $|xy| = |x|\,|y|$
4. $|x + y| \leq |x| + |y|$

We also defined the distance between two numbers x and y as $d(x, y) = |x - y|$, such that it has the same properties as the absolute value.

1.1.2 Sequences

We studied to the sequences between of real numbers, and in particular, we defined the concepts of convergent and Cauchy sequences.

Definition 1.1.1. Let $\{a_n\}$ be a sequence of real numbers. We say that $\{a_n\}$ *is convergent to the real number* L (or *has limit* L) if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \mid |a_n - L| < \varepsilon, \forall n > n_0 \tag{1.2}$$

And we will write $\{a_n\} \to L$ or $\lim \{a_n\} = L$

Definition 1.1.2. Let $\{a_n\}$ be a sequence of real numbers. We say that $\{a_n\}$ *is a Cauchy sequence* if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \mid |a_m - a_n| < \varepsilon, \forall m, n > n_0 \tag{1.3}$$

From that, we saw other two important properties of real numbers. The first is that $\mathbb{R}$ is arquimedian, that is, $\mathbb{N}$ is not bounded in $\mathbb{R}$. The second one is that it forms a complete field, which means that satisfies the extreme property [ ] or, which is equivalent, the following condition.

$$\{a_n\} \text{ is a Cauchy sequence} \Leftrightarrow \{a_n\} \text{ is convergent} \tag{1.4}$$

1.1.3 Topological concepts

Definition 1.1.3. We define a *ball* B *of center* a *and radius* r as a subset of $\mathbb{R}$ determined as follows.

$$B(a, r) \coloneqq \{x \in \mathbb{R} \,|\, |x - a| < r\} \tag{1.5}$$

Definition 1.1.4. We define a *punctured ball* B^* *of center* a *and radius* r as a neighborhood of center a and radius r that does not contain a. It can also be expressed as follows.

$$B^*(a, r) \coloneqq \{x \in \mathbb{R} \,|\, 0 < |x - a| < r\} \tag{1.6}$$

Definition 1.1.5. Let $A \subseteq \mathbb{R}$ be a set. We say that a is an *interior point of* A if there is a neighborhood $\mathfrak{N}(a, r) \subset A$.

Definition 1.1.6. Let $A \subseteq \mathbb{R}$ be a set. We say that a is an *exterior point of* A if it is an interior point of A^c.

Definition 1.1.7. Let $A \subseteq \mathbb{R}$ be a set. We say that a is a *boundary point of* A if it is not interior or exterior.

Definition 1.1.8. Let $A \subseteq \mathbb{R}$ be a set. We say that a is an *accumulation point of* A if every neighborhood with center a contains points of A different to a.

Sets

Definition 1.1.9. Let $A \subseteq \mathbb{R}$ be a set. We say A *is an open set* if all its points are interior points.

Definition 1.1.10. Let $A \subseteq \mathbb{R}$ be a set. We say A *is a closed set* if it contains all its accumulation points (or, what is equivalent, if A^c is open).

Definition 1.1.11. Let $A \subseteq \mathbb{R}$ be a set. We say A *is a bounded set* if it is contained by some neighborhood of center 0.

Definition 1.1.12. Let $A \subseteq \mathbb{R}$ be a set. We say A *is a compact set* if every sequence of elements of A has some partial sequence convergent inside A (or, what is equivalent, if it is closed and bounded).

1.2 General concepts from metric spaces

1.2.1 Introduction

We have studied previously a great amount of concepts of real numbers. Some examples are open set, distances, convergence of sequences, and continuity of functions. Now we want to expand this knowledge and apply these notions to new sets like $\mathbb{R}^n$. This means we need to redefine each concept to $\mathbb{R}^n$, but if in other occasion we are interested in other set, we will have to redefine them again (and every time there is a new set. To simplify this, we can generalize this concepts by studying some kind of sets in general (called metric spaces), and then we will just need to apply the general definition to the case we are interested, in our case $\mathbb{R}$. Said that, let us start by defining a metric space.

Definition 1.2.1. A *metric space* $(\mathbb{M}, d)$ is a set $\mathbb{M}$ with a *metric* or *distance function* d of the form

$$\begin{aligned} d : \mathbb{M} \times \mathbb{M} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto d(x, y) \end{aligned} \tag{1.7}$$

That satisfies the following properties:

1. For any element $x \in \mathbb{M}$, we have $d(x, x) = 0$
2. Positivity: for any distinct $x, y \in \mathbb{M}$, we have $d(x, y) > 0$
3. Symmetry: for any $x, y \in \mathbb{M}$, we have $d(x, y) = d(y, x)$
4. Triangle inequality: for any $x, y, z \in \mathbb{M}$, we have $d(x, z) \leq d(x, y) + d(y, z)$

In some cases, if the metric d is evident, we just write the metric space as $\mathbb{M}$.
Sometimes the two first conditions of metric spaces are rephrased in one simple condition, which is that $d(x, y) = 0$ if and only if $x = y$, which is equivalent. The first condition is the same as saying $x = y \Rightarrow d(x, y) = 0$ (first implication of this new condition), and the second one $x \neq y \Rightarrow d(x, y) \neq 0 \equiv d(x, y) = 0 \Rightarrow x = y$ (second implication of this new condition). Now we have the definitions, we will see some examples of metrics [9].

Example 1.2.1. Let $\mathbb{R}$ be the set of real numbers and the function $d(x, y) := |x - y|$. This function satisfies all the conditions to be called distance and therefore makes $(\mathbb{R}, d)$ a metric space. We denote this distance as the *standard metric on* $\mathbb{R}$.

Example 1.2.2. Let $\mathbb{M}$ be a set (finite or infinite, not necessary with real numbers) and the following function.

$$d_{\text{disc}}(x, y) := \begin{cases} 0, x = y \\ 1, x \neq y \end{cases} \tag{1.8}$$

This function satisfies the condition of the first function, so $\mathbb{M}, d_{\text{disc}}$ is a metric space. We call this metric the *discrete metric* and we denote it by d_{disc}.

Example 1.2.3. Let $d(\mathbb{M}, d)$ be a metric space and $\mathbb{L}$ a subset of $\mathbb{M}$. Then, the distance function restricted to $\mathbb{L} \times \mathbb{L}$ satisfies also the conditions of metric. We called it the metric on $\mathbb{L}$ inducted by the metric d on $\mathbb{M}$, and denote it by $d|_{\mathbb{L}\times\mathbb{L}}$. We say also that $(\mathbb{L}, d|_{\mathbb{L}\times\mathbb{L}})$ is a subspace of $\mathbb{M}, d$ induced by $\mathbb{L}$.

Example 1.2.4. Let $\mathbb{R}^n$ be the set of n-tuples of real numbers. We call the *taxicab metric or l^1 metric*, denoted by d_{l^1}, as the following distance function.

$$d_{l^1}\left(\left(x_1, \ldots, x_n\right), \left(y_1, \ldots, y_N\right)\right) := \sum_{i=1}^{n} |x_i - y_i| \tag{1.9}$$

This function satisfies the conditions of metric and hence $d\left(\mathbb{R}^n, d_{l^1}\right)$ is a metric space.

Example 1.2.5. Let $\mathbb{R}^n$ be the set of n-tuples of real numbers. As we have seen before [ ], it behaves as an affine space with a vector space $\mathbb{R}^n$ associated. In that case, we defined the euclidean distance $d(x, y)$ between two points x and y as the norm (using the standard dot product) of the vector $\overrightarrow{xy}$ that satisfied $y = x + \overrightarrow{xy}$. We can see that this satisfies the condition of a metric. Besides, using a Cartesian reference in the affine space [ ] with some origin O (usually the point 0), we can associate to each point x a vector $\vec{x}$ that satisfies $x = O + \vec{x}$. This means that we can write $\overrightarrow{xy}$ as $\vec{y} - \vec{x}$ and therefore, rewrite the distance function in another way.

$$d_{l^2}\left(x, y\right) = \sqrt{\langle \overrightarrow{xy}, \overrightarrow{xy} \rangle_I} := \|\overrightarrow{xy}\| = \|\vec{y} - \vec{x}\| \tag{1.10}$$

This metric is called the *l^2 metric or euclidean metric* (since the vector space is using an euclidean geometry as a bilineal form [ ]), and the space the *euclidean space*.

Example 1.2.6. Let $\mathbb{R}^n$ be the set of n-tuples of real numbers. The function

$$d_{l^\infty}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \sup\left\{|x_i - y_i| \mid 1 \leq i \leq n\right\} \tag{1.11}$$

satisfies all the condition of a distance function. The metric $(\mathbb{R}^n, d_{l^\infty})$ is called the *sup norm metric* or *l^∞ metric.*

1.2.2 Sequences

Definition 1.2.2. Let A be a set. A *sequence in A* is an application from natural numbers to A.

$$\begin{aligned} \mathbb{N} &\longrightarrow A \\ k &\longmapsto a_k \end{aligned} \tag{1.12}$$

We denote it by $\{a_1, a_2, \ldots\}$ or, more shortly, $\{a_k\}$.

Definition 1.2.3. Let $\{a_k\}$ be a sequence of points of o a metric space $(\mathbb{M}, d)$. We say *the sequence $\{a_k\}$ is bounded* is there exist a point a and a positive real number r such that $d(a_k, a) < r$.

Definition 1.2.4. Let $\{a_k\}$ be a sequence of points of o a metric space $(\mathbb{M}, d)$. We say that $\{a_k\}$ *is convergent to the point* $L \in (\mathbb{M}, d)$ (or *has limit* L) *with respect to the metric* d if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \mid d\left(a_k, L\right) < \varepsilon, \forall n > n_0 \tag{1.13}$$

In that case, we will write $\{a_k\} \to L$ or $\lim_{k\to\infty} \{a_k\} = L$. As before [], this is equivalent to say it is convergent to L if equivalent to say $\lim_{k\to\infty} d(x_k, L) = 0$.

Proposition 1.2.1. *Let* $(\mathbb{M}, d)$ *be a metric space and* $\{a_k\}$ *a sequence in* $\mathbb{M}$ *that is convergent to a point* L *with respect to the metric* d*. Then,* L *is unique.*

Proposition 1.2.2. *Let* $(\mathbb{M}, d)$ *be a metric space and* $\{a_k\}$ *a sequence in* $\mathbb{M}$ *that is convergent to a point* L *with respect to the metric* d*. Then,* $\{a_k\}$ *is bounded.*

Definition 1.2.5. Let $\{a_k\}$ be a sequence of points of $(\mathbb{M}, d)$. We say that $\{a_k\}$ *is a Cauchy sequence with respect to the metric* d if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^{+} \exists n_0 \in \mathbb{N} \mid d(x_k, x_l) < \varepsilon, \forall l, k > n_0 \tag{1.14}$$

Proposition 1.2.3. *Let* $(\mathbb{M}, d)$ *a metric space and* $\{a_k\}$ *a convergent sequence in* $(\mathbb{M}, d)$*. Then,* $\{a_k\}$ *is a Cauchy sequence.*

Definition 1.2.6. Let $(\mathbb{M}, d)$ a metric space. We say $(\mathbb{M}, d)$ is *complete* if every Cauchy sequence in $(\mathbb{M}, d)$ is a convergent sequence in $(\mathbb{M}, d)$.

1.2.3 Topological concepts

Balls

Definition 1.2.7. Let $(\mathbb{M}, d)$ be a metric space, a a point in $\mathbb{M}$, and r a positive real number. We define a *metric ball* $B_{(\mathbb{M},d)}$ *of center* a *and radius* r *in the metric* d as the set

$$B_{(\mathbb{M},d)}(a, r) \coloneqq \{x \in \mathbb{M} \,|\, d(x, a) < r\}. \tag{1.15}$$

If the metric space $(\mathbb{M}, d)$ is clear, we just write $B(a, r)$

Note that smaller the radius r, less elements contains the ball $B(a, r)$. Nevertheless, it always contains at least one element, a, since r is positive.

Definition 1.2.8. Let $(\mathbb{M}, d)$ be a metric space, a a point in $\mathbb{M}$, and r a positive real number. We define a *closed metric ball* $B^{\wr}_{(\mathbb{M},d)}$ *of center* a *and radius* r *in the metric* d as the set

$$B^{\wr}_{(\mathbb{M},d)}(a, r) \coloneqq \{x \in \mathbb{M} \,|\, d(x, a) \leq r\}. \tag{1.16}$$

We will see later why the closed ball is called *closed* [ ].

Definition 1.2.9. Let $(\mathbb{M}, d)$ be a metric space, a a point in $\mathbb{M}$, and r a positive real number. We define a *punctured metric ball* $B^{*}_{(\mathbb{M},d)}$ *of center* a *and radius* r *in the metric* d as the set

$$B^{*}_{(\mathbb{M},d)}(a, r) \coloneqq \{x \in \mathbb{M} \,|\, 0 < d(x, a) < r\}. \tag{1.17}$$

Now we have the generalized concepts of metric balls (also called simply *balls*), we can classify the points of $(\mathbb{M}, d)$ as we did with $\mathbb{R}$.

Points

Definition 1.2.10. Let A be a subset of a metric space $(\mathbb{M}, d)$ and a a point in $\mathbb{M}$. We say that a is an *interior point of* A if there is a ball $B_{(\mathbb{M},d)}(a, r) \subset A$.

Definition 1.2.11. Let A be a subset of a metric space $(\mathbb{M}, d)$ and a a point in $\mathbb{M}$. We say that a is an *exterior point of* A if there is a ball such that $B_{(\mathbb{M},d)}(a, r) \cup A = \emptyset$.

Definition 1.2.12. Let A be a subset of a metric space $(\mathbb{M}, d)$ and a a point in $\mathbb{M}$. We say that a is a *boundary point of* A if it is not interior or exterior or, which is equivalent, if every ball $B_{(\mathbb{M},d)}(a, r)$ contains elements of A and A^c.

Definition 1.2.13. Let A be a subset of a metric space $(\mathbb{M}, d)$ and a a point in $\mathbb{M}$. We say that a is an *accumulation point of* A if every ball with center a contains points of A different to a. In other words, every punctured ball satisfies $B^{*}_{(\mathbb{M},d)}(a, r) \cup A \neq \emptyset$.

Components of a set

Definition 1.2.14. Let A be a subset of a metric space $(\mathbb{M}, d)$. We define *the interior of* A as the set of all interior points of A, and we denote it by $\text{int}(A)$.

Definition 1.2.15. Let A be a subset of a metric space $(\mathbb{M}, d)$. We define *the exterior of* A as the set of all exterior points of A, and we denote it by $\text{ext}(A)$.

Definition 1.2.16. Let A be a subset of a metric space $(\mathbb{M}, d)$. We define *the boundary of* A as the set of all boundary points of A, and we denote it by ∂A.

Definition 1.2.17. Let A be a subset of a metric space $(\mathbb{M}, d)$. We define *the closure of* A as the set of all accumulation points of A, and we denote it by $\bar{A}$.

If a is an interior point of A, then it is an element of A, since $a \in B_{(\mathbb{M},d)}(a, r) \subset A$. If a is an exterior point of A, it can not be an element of A, since then we would have that $B_{(\mathbb{M},d)}(a, r) \cup A \neq \emptyset$. Therefore, a can not be both interior and exterior of A at the same time. We should not confuse this by thinking that is a point is not interior then is exterior, because is not true. In case of boundary points, they can be elements or not of A. In order to see more clearly this, we will see some examples and then some results.

Proposition 1.2.4. *Let A be a subset of a metric space $(\mathbb{M}, d)$ and a an exterior point of A. Then, it is an interior point of A^c.*

Proof. Let us suppose a is not interior of A^c. Then, does not exist any ball strictly contained in A^c. There for, for every ball there is always at least one element that is not in A^c, so this element will be in A. Hence, for every ball we will have that $B^*_{(\mathbb{M},d)}(a, r) \cup A \neq \emptyset$, contradicting the fact that is an exterior point of A. Then our assumption was wrong and we conclude a is interior of A^c. $\square$

Proposition 1.2.5. *Let A be a subset of a metric space $(\mathbb{M}, d)$ and a an exterior point of A. Then, the following statements are equivalent [9].*

1. *a is an accumulation point of A.*

2. *a is either an interior point or a boundary point of A.*

3. *There exists a sequence $\{a_k\}$ in A which converges to a with respect to the metric a.*

Corollary 1.2.6. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, $\bar{A} = int(A) \cap \partial A = \mathbb{M} \setminus ext(A)$.*

As we have said before, boundary points of a set A can be or not elements of A. This will serve us as criterion to classify sets in a metric space.

Definition 1.2.18. Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. We say A is *an open set* if it contains none of its boundary points, that is, if $\partial A \cap A = \emptyset$.

Definition 1.2.19. Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. We say A is *a closed set* if it contains all its boundary points, that is, if $\partial A \subseteq A$.

Note that a set can be both open and closed, if it has no boundary.

Definition 1.2.20. Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. We say A is *a bounded set* if there exist a point $a \in \mathbb{M}$ and a positive real number r such that the ball $B_{(\mathbb{M},d)}(a, r)$ contains A.

Definition 1.2.21. Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. We say A is *a compact set* if it a bounded and closed set.

Proposition 1.2.7. *Let $(\mathbb{M}, d)$ be a metric space and A an open subset of $\mathbb{M}$. Then, all its points are interior points of A.*

1.2.4 Basic properties of sets

We will see now some basic properties of sets [9].

Proposition 1.2.8. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, A is open if and only if $A = int(E)$. In other words, A is open if and only if for every $a \in A$, there exists an $r \in \mathbb{R}^+$ such that $B(a, r) \subseteq A$.*

Proposition 1.2.9. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, A is closed if and only if A contains all its accumulation points, that is, $\hat{A} = A$. In other words, A is closed if and only if for every convergent sequence $\{a_k\}$ in A, the limit $\lim_{k\to\infty} \{a_k\}$ of that sequence also lies in A.*

Proposition 1.2.10. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, for any $a \in A$ and $r \in \mathbb{R}^+$ the ball $B_{(\mathbb{M},d)}(a, r)$ is an open set and the closed ball $B^{\wr}_{(\mathbb{M},d)}(a, r)$ is a closed set.*

Proposition 1.2.11. *Let $(\mathbb{M}, d)$ be a metric space and a point a of $\mathbb{M}$. Then, the singleton set $\{a\}$ is a closed set.*

Proposition 1.2.12. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, A is open if and only if A^c is closed.*

Proposition 1.2.13. *Let $(\mathbb{M}, d)$ be a metric space, $A_1, \ldots, A_n$ a finite collection of open sets in $\mathbb{M}$, and $B_1, \ldots, B_n$ a finite collection of closed sets in $\mathbb{M}$. Then, $A_1 \cap \ldots \cap A_n$ is an open set and $F_1 \cup \ldots \cup B_n$ is a closed set.*

Proposition 1.2.14. *Let $(\mathbb{M}, d)$ be a metric space, $\{A_\alpha\}_{\alpha\in I}$ a collection of open sets in $\mathbb{M}$ and $\{B_\alpha\_\alpha \in I\}$ (where the index set I could be finite, countable, or uncountable). Then, the $\bigcup_{\alpha\in I} E_\alpha$ is an open set and $\bigcap_{\alpha\in I} B_\alpha$ is a closed set.*

Proposition 1.2.15. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, $int(A)$ is the largest open set which is contained in A; in other words, $int(A)$ is open, and given any other open set $B \subseteq A$, we have $A \subseteq int(A)$. Similarly $\bar{A}$ is the smallest closed set which contain A; in other words, $\bar{A}$ is closed, and given any other closed set $C \supset A, C \supset \bar{A}$.*

1.2.5 Other properties

Proposition 1.2.16. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, ∂A is a closed set and A is closed if and only if $\partial A \subseteq A$.*

Proposition 1.2.17. *Let $(\mathbb{M}, d)$ be a metric space and A a subset of $\mathbb{M}$. Then, $\partial A = \hat{A}$ $int(A)$.*

1.3 The $\mathbb{R}^n$ space

1.3.1 Introduction

We have seen there are different ways to define a distance in $\mathbb{R}^n$, for example with the metric d_{l^1}, d_{l^2} or even d_{disc}. From now we will continue studying the concepts in $\mathbb{R}^n$ using the euclidean metric because it is the metric that best expresses the real world. However, it is important to remember that next topics we will discuss are not restricted to this metric but they can be treated with every other metric. For example, concepts like continuity of functions can also be defined with other metrics, but several results can change when using another metric.
The action of taking the euclidean metric for $\mathbb{R}^n$ requires for an important clarification in order to be sure we are studying the topic rigorously. As we said before, we define the distance between two points x and y with the norm of the vector $\overrightarrow{xy}$. We are not taking the norm of $x - y$ because there is not an established operation for the affine space, only for its associated vector space. Similarly, we can not talk about the distance between two associated vectors $\vec{x}$ and $\vec{y}$ (the ones we define when we defined the euclidean metric), again, because they are not points. In general, we should not confuse vectors with points, because have different behaviors. Therefore, the unique ways to

write de distance between two points will be the ones we presented when we define the euclidean distance. Any other way is incorrect because makes confusion between points and vectors.

$$d_{l^2}(x,y) = \sqrt{\langle \overrightarrow{xy}, \overrightarrow{xy} \rangle_I} := \|\overrightarrow{xy}\| = \|\vec{y} - \vec{x}\| \tag{1.18}$$

From the last expression we should have into account another detail. A point of $\mathbb{R}^n$ is a n-tuple of the form $(x_1, \dots, x_n)$ and its associated vector $\vec{x}$ is also a n-tuple, but its components not are necessary the same numbers as these of x. This is because the origin O of the Cartesian frame can be any point of $\mathbb{R}^n$. What we will do from now is to select O as the origin of $\mathbb{R}^n$, and then the components of x and $\vec{x}$ will coincide. This will make the calculation easier because we will just need to know the components of one, but also the notation. We will denote the components of vectors and points equal, so for example, if we write the addition of two associated vectors we will write $\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_n + y_n)$. But again, with expressions like this we should not think ot is an addition of two points, it is only that components are equally expressed. If we think in components individually the same is applied, the sum of components of vectors are not the same as the sum of components of points (which is not defined), although it looks the same. Every time we see an operation we are referring to to associated vectors and not the points themselves. Said that, the expression of the distance can be expressed (thinking about operations of vectors) as

$$d(x,y) = \sum_{i=1}^{n} (x_i - y_i)^2. \tag{1.19}$$

Finally, we need to establish the basis of the vector space, because will determine the expressions of the coordinates of vectors but also points. A point $(x_1, \dots, x_n)$ will change its components if we change of basis, so we will define a unique basis that will be used always from now. We will choose the standard basis. We choose this for two reasons. The first one is that it is the unique bases where a vector (and point) has the same coordinates as its actual values. In other words, is the unique basis that satisfies

$$x = (x_1, \dots, x_n) = \sum_{i=1}^{n} x_i \vec{e}_i. \tag{1.20}$$

The second one is that is an orthonormal basis, and therefore it has a lot of useful properties, among which is that their dot product generates the Kronecker delta (and that simplifies a lot the calculations).

$$\langle \vec{e}_i, \vec{e}_j \rangle_I = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \tag{1.21}$$

Now we have clarified all the details related to the expression of points, vectors, and the metric of $\mathbb{R}^n$, we can start studying the first concepts.

1.3.2 Sequences

Since we have defined the general sequence in metric spaces, we just need to specify that now sequences are applications that map natural numbers to n-tuples of real numbers. However, as a specific trait of $\mathbb{R}^n$, we need to clarify the notation. We usually write x_i to denote the component i of the point or vector, but in sequence this can be misunderstood by the ith term of the sequence. In order to solve that, always we deal with sequences we will use sub-index to denote the element of the sequence and super-index to denote the component of the point or vector. With that, a sequence could be expressed as follows.

$$a_k = \{a_1, a_2, \dots\} = \left\{ \begin{pmatrix} x_1^1 \\ \vdots \\ x_1^n \end{pmatrix}, \begin{pmatrix} x_2^1 \\ \vdots \\ x_2^n \end{pmatrix}, \dots \right\} \tag{1.22}$$

The concepts of convergent sequence and Cauchy sequence are also applicable to the metric $(\mathbb{M}, d_{l^2})$. In particular, if we think about the distance of two points, since we are dealing with n-tuples, we can wonder about what happens to their components in particular. It seems reasonable that each component will converge to the each component of a point, but we need to express this rigorously. That is what we will do now.

Proposition 1.3.1. *Let $\{a_k\}$ be a sequence of points in $\mathbb{R}^n$. Then, $\{a_k\}$ has the limit L if and only if each succession $\{a_k^i\}$ of coordinates $i = 1, \ldots, n$ has as a limit the correspondent coordinate L^j of L.*

Proof. As we have seen before [ ], we know that

$$0 \leq \left|a_k^i - L^i\right| \leq d\left(a_k, L\right) \leq \sum_{i=1}^{n} \left|a_k^i - L^i\right|$$

First, we will prove that the convergence of components implies the convergence of the whole sequence and then te opposite implication. Since $\lim_{k\to\infty} a_k^i = L^i, \forall i$, we get

$$\lim_{k\to\infty} a_k^i = L^i \Rightarrow \lim_{k\to\infty} a_k^i - L^i = 0 \Rightarrow \lim_{k\to\infty} \left|a_k^i - L^i\right| = 0 \Rightarrow$$
$$\lim_{k\to\infty} \sum_{i=1}^{n} \left|a_k^i - L^i\right| = \sum_{i=1}^{n} \lim_{k\to\infty} \left|a_k^i - L^i\right| = \sum_{i=1}^{n} 0 = 0$$

Therefore, by the Squeeze Theorem, we have that $\lim_{k\to\infty} d\left(a_k, L\right) = 0$ or, which is equivalent, that $\lim_{k\to\infty} a_k = L$.
Let us suppose now that $\lim_{k\to\infty} a_k = L$, or what is equivalent, that $\lim_{k\to\infty} d(a_k, L) = 0$. By the first inequality, we have that

$$0 \leq \left|a_k^i - L^i\right| \leq d\left(a_k, L\right) \Rightarrow \lim_{k\to\infty} 0 \leq \lim_{k\to\infty} \left|a_k^i - L^i\right| \leq \lim_{k\to\infty} d\left(a_k, L\right) \tag{1.23}$$

Both sides of the new inequality converge to 0 and, again for the Squeeze $\left|a_k^i - L^i\right|$ too, or which is equivalent, $\lim_{k\to\infty} a_k^i = L^i, \forall i$. □

We have a similar result about bounded sequences.

Proposition 1.3.2. *Let $\{a_k\}$ be a sequence of points in $\mathbb{R}^n$. Then, $\{a_k\}$ is bounded if and only if each succession $\{a_k^i\}$ of coordinates $i = 1, \ldots, n$ is bounded.*

Example 1.3.1. Let be the following sequence.

$$\{a_k\} = \left\{\left(1 - \frac{1}{k}, -2\sin\left(\frac{k+1}{k}\pi\right)\right)\right\}$$

By the proposition 1.3.1, we can just study the limits of each component to know the global limit.

$$\lim_{k\to\infty} 1 - \frac{1}{k} = 1$$
$$\lim_{k\to\infty} -2\sin\left(\frac{k+1}{k}\pi\right) = -2\sin\left(\pi\right) = 0$$

Therefore, the limit is $L = (1, 0)$.

Example 1.3.2. Let be the following sequence.

$$\{a_k\} = \left\{\left(\sin k, (-1)^k, \frac{1}{k}\right)\right\}$$

We can see it is a bounded sequence because every component is bounded (Prop 1.3.2). However, by the proposition 1.3.1, since not every component converges the sequence is not convergent.

Example 1.3.3. Let be the following sequence.

$$\{a_k\} = \{k, k^2, \ldots, k^n\}$$

Where the super-indexes represent the degree of the exponent. This sequence is not bounded and therefore not convergent (Prop 1.2.2).

Some properties

Proposition 1.3.3. *Let a_k and a_k two convergent sequence with limits L_a and L_b respectively, and a real number λ. Then,*

$$\{a_k + b_k\} \to L_a + L_b \qquad \{\lambda a_k\} \to \lambda L_a \tag{1.24}$$

We can do this operations because we considered it before [ ].
Now we have discussed the convergence of sequences, we can also study when a sequence in this metric is a Cauchy sequence. In fact, we have the following result.

Proposition 1.3.4. *Let $\{a_k\}$ be a sequence of points in $\mathbb{R}^n$. Then, $\{a_k\}$ is a Cauchy sequence if and only if each succession $\{a_k^i\}$ of coordinates $i = 1, \dots, n$ has is a Cauchy sequence.*

Proposition 1.3.5. *The set $\mathbb{R}^n$ with the metric d_{l^2} is complete.*

Proof. By the proposition 1.3.1, a sequence $\{a_k\}$ in the metric space $(\mathbb{R}^n, d_{l^2})$ is convergent if and only if its coordinates are convergent. As we have seen before, the set $\mathbb{R}$ is complete, so this is equivalent to say that the coordinates are Cauchy sequences. Finally, by the proposition 1.3.4, the coordinates of the sequence are convergent if and only if the $\{a_k\}$ is a Cauchy sequence. □

1.3.3 Some sets

Definition 1.3.1. Let $a_i, b_i \in \mathbb{R}$, $i = 1, \dots, n$ and $a_i < b_i$ for all $i = 1, \dots n$. Then, we define the *n-dimensional open interval I* as follows

$$I = \{\vec{x} \in \mathbb{R}^n | a_i < x_i < b_i, i = 1, \dots, n\} = (a_1, b_1) \times \cdots \times (a_n, b_n). \tag{1.25}$$

Definition 1.3.2. Let $a_i, b_i \in \mathbb{R}$, $i = 1, \dots, n$ and $a_i < b_i$ for all $i = 1, \dots n$. Then, we define the *n-dimensional open interval I* as follows

$$I = \{\vec{x} \in \mathbb{R}^n | a_i \leq x_i \leq b_i, i = 1, \dots, n\} = [a_1, b_1] \times \cdots \times [a_n, b_n]. \tag{1.26}$$

Proposition 1.3.6. *Let $I \in \mathbb{R}^n$ be an open interval. Then, I is open.*

Proof. Done in the following examples. □

Proposition 1.3.7. *Let $I \in \mathbb{R}^n$ be a closed interval. Then, I is closed.*

1.3.4 Examples of topological concepts

Let us see some examples.

Example 1.3.4. [10]. In the set $[0, 1] \cup \{2\}$, the point $a = 2$ is not an accumulation point. In the set $\left\{x \in \mathbb{R}^n \,\middle|\, x^i > 0\right\} \cup \{0, \dots, 0, -1\}$, the point $a = (0, \dots, 0, -1)$ is not an accumulation point.

Example 1.3.5. [10]. In the set $C = (0, 1)$, the points $a = 0$ and $a = 1$ are accumulations points and are not in the C. In the set $D = \{x \in \mathbb{R}^n \mid \|x\| < 1\}$, the points $a = (0, \dots, 0, 1)$ and every point such that $\|a\| = 1$ are accumulation points.

Example 1.3.6. [1] We have seen previously that in
$mathbbR$ the open intervals are open sets and closed intervals are closed sets []. From these sets we can construct new ones using the Cartesian product. In particular, if we take two open intervals $I_!$ and I_2, the Cartesian product $I_1 \times I_2$ is an open set of $\mathbb{R}^2$.
To see that the product of two open sets is an open set, we will take some point $a = (a_1, a_2)$ of $I_1 \times I_2$ and see it is interior. Since I_1 and I_2 are open sets, there is a ball $B(a_1, r_1)_1 \subset I_1$ and a ball $B(a_2, r_2) \subset I_2$, both of one dimension. Let be $r = \min\{r_1, r_2\}$. If we see the ball $B(a, r)$ is completely contained in $I_1 \times I_2$, it will satisfy the condition of interior point.
Let be a point $x \in B(a, r)$. By the definition 1.2.7, this means that $d(x, a) < r$, and by the equivalence of the equation 1.19, that $(x_1 - a_1)^2 + (x_2 - a_2)^2 < r^2$. Therefore,

$$(x_1 - a_1)^2 < r^2 \Rightarrow |x_1 - a_1| < r < r_1 \Rightarrow |x_1 - a_1| < r_1$$
$$(x_2 - a_2)^2 < r^2 \Rightarrow |x_2 - a_2| < r < r_2 \Rightarrow |x_2 - a_2| < r_2.$$

This shows that $a_1 \in B(a_1, r_1) \subset I_1$ and that $a_1 \in B(a_2, r_2) \subset I_2$, so a_1 is interior of I_2 and a_2 interior of I_2. Consequently, (a_1, a_2) is an interior point of $I_1 \times I_2$.

Example 1.3.7. In the set $F = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$, every point a of F is an accumulation point.

Let us see some examples of sets

Example 1.3.8. The sets $\mathbb{R}^n$ and $\emptyset$ are open set.

Example 1.3.9. The sets $\mathbb{R}^n$ and $\emptyset$ are closed sets.

As in case of real numbers, closed sets are characterized as these sets such that, if there is a sequence with element in the sets and has a limit, then the it is in the set. Besides, the finite union and intersection of closed sets are closed, and a set is open if and only if its complementary is closed. As a consequence, the finite intersection of open sets is open [3].
We have the same property of compact sets as in $\mathbb{R}$.

Proposition 1.3.8. *A set $A \subseteq \mathbb{R}^n$ is compact if and only if is closed and bounded.*

Bibliography

[1] Tom M Apostol. *Calculus, vol. 2.* Reverté, 1985.

[2] Tom M Apostol. *Análisis matemático.* Reverté, 2020.

[3] Joaquín M Ortega Aramburu. *Introducció a l'anàlisi matemàtica*, volume 28. Univ. Autònoma de Barcelona, 2002.

[4] R. Courant and F. John. *Introducción al cálculo y al análisis matemático, vol. 2.* Limusa, 1999.

[5] Seán Dineen. *Multivariate Calculus and Geometry.* Springer, 2020.

[6] J.Rogawski. *Calculus.* Ruth Baruth, 2012.

[7] Jerrold E Marsden and Anthony Tromba. *Vector calculus.* Macmillan, 2003.

[8] A. Méndez. Càlcul en diverses variables, 2014.

[9] Terence Tao. Analysis II, Texts and Readings in Mathematics, vol. 38. *Hindustan Book Agency, New Delhi*, 2009.

[10] J. Verdera. Càlcul en diverses variables, 2015.

Chapter 2

Functions in $\mathbb{R}^n$. Limits and continuity

2.1 Definition

We have defined a *distance function* in the set of $\mathbb{R}^n$ to establish a metric in its points. However, since we will work more with the associated vector $\vec{x}$ of a point x than with the point itself, from now we will mention in general the vectors and the expression $\|\vec{x}-\vec{y}\|$ rather than the points and their distance.

Definition 2.1.1. Let $\mathbb{R}^n$ be the set of n-tuples of real numbers and $\mathbb{R}$ the set of real numbers, both with the metric d_{l^2}. We define a *scalar field* as a function that maps $\mathbb{R}^n$ to $\mathbb{R}$.

$$\begin{aligned} f(\vec{x}) : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \vec{x} &\longmapsto f(\vec{x}) \end{aligned} \tag{2.1}$$

Definition 2.1.2. Let $\mathbb{R}^n$ be the set of n-tuples of real numbers and $\mathbb{R}^m$ be the set of m-tuples of real numbers, both with the metric d_{l^2}. We define a *vector field* as a function that maps $\mathbb{R}^n$ to $\mathbb{R}^m$.

$$\begin{aligned} \vec{f}(\vec{x}) : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \vec{x} &\longmapsto \vec{f}(\vec{x}) \end{aligned} \tag{2.2}$$

Note we are applying functions to the associated vectors to some points, so it is the same that applying these functions to the points themselves. Another detail about these functions is that we are now denoting with arrows the vectors that have more than one dimension, since uni-dimensional vectors act like scalars. Finally, we clarified that we would use the euclidean metric space in $\mathbb{R}^n$, but now we are studying more than one space and one of them could have another metric. In these previous definitions we stated what metrics we use, but from now we will not tell it explicitly. From now we will use the euclidean metric for all the possible sets and it will be considered that we know it.

2.2 Limits

2.2.1 Introduction

Definition 2.2.1. Let be a function $\vec{f}(\vec{x}) : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function. We say that $\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{L}$ if it satisfies

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta(\varepsilon) \mid 0 < \|\vec{x}-\vec{a}\| < \delta \Rightarrow \left\|\vec{f}(\vec{x}) - \vec{L}\right\| < \varepsilon \tag{2.3}$$

in all possible ways.

There is another way to express this concept. If we define $\vec{h} := \vec{x}-\vec{a}$ then, instead of having $\vec{x}\to\vec{a}$ we will have $\vec{x}-\vec{a}\to\vec{0} \Rightarrow \vec{h}\to\vec{0}$. Since the distance function of the points is a non-degenerate form, the unique point whose distance to origin point is the origin point (or what is the same, $\vec{h} = 0 \Leftrightarrow \|\vec{h}\| = 0$). Therefore we could just write $\|\vec{h}\| \to 0$ as an equivalence to the first part of the definition of limit. Besides, since $\vec{h} = \vec{x}-\vec{a} \Leftrightarrow \vec{h}+\vec{a} = \vec{x}$, we can just substitute $\vec{f}(\vec{x})$ by $\vec{f}(\vec{a}+\vec{h})$. Said that, we can denote the limit by these three expressions.

$$\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{L} \qquad \lim_{\vec{h}\to\vec{0}} \vec{f}(\vec{a}+\vec{h}) = \vec{L} \qquad \lim_{\|\vec{h}\|\to 0} \vec{f}(\vec{a}+\vec{h}) = \vec{L} \tag{2.4}$$

Apart from the diversity in the rephrasing the limit, there are two important details we must consider. The first one is the definition talks about what happens when a point x approaches to a without being equal, and therefore it could be that a is not in Ω. As happened with one variable functions, the only criterion is that a needs to be an accumulation point, since it guarantees that there is always a point x inside a ball $B^*(a,\delta)$ that is also in Ω.
The second detail is that in the definition we said "in all possible ways". Since we are dealing with n-tuples of points, each component can change in a different rate than the other, and the result is an arbitrary curve that approaches to the point a. Let se see some examples to illustrate this.

Example 2.2.1. Let be the following piecewise function.

$$f(x,y)=\begin{cases}\frac{2xy}{x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases}$$

And let us see what happens when x and y approach to 0. First, we will try by using the path $x=x, y=0$.

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{y=0,x\to 0}\frac{2xy}{x^2+y^2}=\lim_{y=0,x\to 0}\frac{2x\cdot 0}{x^2+0}=\lim_{y=0,x\to 0}\frac{0}{x^2}=0$$

Let us try now the path $x=0, y=y$.

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{x=0,y\to 0}\frac{2xy}{x^2+y^2}=\lim_{x=0,y\to 0}\frac{2\cdot 0\cdot y}{0+y^2}=\lim_{x=0,y\to 0}\frac{0}{y^2}=0$$

We see that by going through the x axis and through the y axis the function converges to 0. Let us now try a combination of these forms, using the path $y=mx, m\neq 0$.

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{x=\to,y=mx}\frac{2xy}{x^2+y^2}=\lim_{x\to 0}\frac{2xmx}{x^2+m^2x^2}=\lim_{x\to 0}\frac{2mx^2}{x^2+m^2x^2}=\frac{2m}{1+m^2}\neq 0$$

Where we could cancel x because does not have the value of 0 (is just approaching to it). With that, we can see the path we choose to go to a point a influences the result of the limit, and since we needed it to converge through all paths, we can conclude this function is not convergent. However, we have only saw right paths, and it would be interesting to see what happens if we take a curved path, like $(x,y))=(t\cos t, t\sin t)$ and make t approach to 0.

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{(x,y)\to(0,0)}\frac{2xy}{x^2+y^2}=\lim_{t\to 0}\frac{2\cdot t\cos t\cdot t\sin t}{t^2\cos^2 t+t^2\sin^2 t}=\lim_{t\to 0}\frac{t\sin(2t)}{t^2}=$$
$$\lim_{t\to 0}\frac{\sin(2t)}{t}=\frac{0}{0}\overset{L'H}{\Rightarrow}\lim_{t\to 0}\frac{2\cos 2t}{1}=2$$

And another time we get a different result.

This problem in fact is not only for multivariable functions. With functions of one variable we can also approach the variable x to a in different ways apart from going directly from the right or from the left. For example, we could do it with the sequence $\{x_k\}=a+(-1)^k$, which tends to a but in an alternative way.

Example 2.2.2. Let be the following piecewise function.

$$f(x,y)=\begin{cases}x, (x,y)\neq(0,0)\\ y, (x,y)=(0,0)\end{cases}$$

And let us see what happens when x and y approach to 0. At first sight we can think it converges to 0, so we will try to prove it. For that, we will take $\delta(\varepsilon)=\varepsilon$. With that,

$$\|\vec{x}-\vec{a}\|<\delta\Rightarrow\left\|\vec{x}-\vec{0}\right\|<\delta\Rightarrow\|\vec{x}\|<\delta=\varepsilon\Rightarrow\varepsilon>\sqrt{x^2+y^2}=\sqrt{x^2}=|x|\Rightarrow$$
$$\varepsilon>|f(\vec{x})-0|\Rightarrow|f(\vec{x})-L|<\varepsilon,$$

which satisfies the condition of limit

2.2.2 Properties

Proposition 2.2.1. *Let $\vec{f}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ and $\vec{g}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be two functions such that $\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})=\vec{L}_f$ and $\lim_{\vec{x}\to\vec{a}}\vec{g}(\vec{x})=\vec{L}_g$. Then, the limit of $\vec{f}(\vec{x})+\vec{g}(\vec{x})$ exists and it satisfies*

$$\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})+\vec{g}(\vec{x})=\vec{L}_f+\vec{L}_g. \tag{2.5}$$

Proof. Since the limit of both functions exist, they satisfy the condition 2.2.1 for any expression of ε and, in particular, there exist two numbers δ_f and δ_g such that

$$\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|<\frac{\varepsilon}{2},\quad \left\|\vec{g}(\vec{x})-\vec{L}_g\right\|<\frac{\varepsilon}{2}.$$

We want to be sure both conditions are satisfied at the same time and, as we did with the study of limits in $\mathbb{R}$, we know that selecting a smaller delta will guarantee that. In particular, we will choose our delta as $\delta_0 ? \min\{\delta_f,\delta_g\}$. And now we know these two inequalities work, we get

$$\frac{\varepsilon}{2}+\frac{\varepsilon}{2}>\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|+\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|\geq\left\|\vec{f}(\vec{x})-\vec{L}_f+\vec{g}(\vec{x})-\vec{L}_g\right\|\Rightarrow$$
$$\varepsilon>\left\|\left(\vec{f}(\vec{x})+\vec{g}(\vec{x})\right)-\left(\vec{L}_f+\vec{L}_g\right)\right\|,$$

which satisfies the definition of limit. □

Proposition 2.2.2. *Let $\vec{f}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be a function such that such that $\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})=\vec{L}_f$ and λ a real number. Then, the limit of $\lambda\vec{f}(\vec{x})$ exists and it satisfies*

$$\lim_{\vec{x}\to\vec{a}}\lambda\vec{f}(\vec{x})=\lambda\vec{L}_f. \tag{2.6}$$

Proof. Since the limit of $\vec{f}(\vec{x})$ exists, it satisfies the condition of 2.2.1 for any expression of ε and, in particular, there exists a δ such that

$$\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|<\frac{\varepsilon}{|\lambda|},$$

for some $\lambda\in\mathbb{R}$. Then, for all $\vec{x}$ such that $\|\vec{x}-\vec{a}\|<\delta$ we have

$$\left\|\lambda\vec{f}(\vec{x})-\lambda\vec{L}_f\right\|=|\lambda|\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|<|\lambda|\frac{\varepsilon}{|\lambda|}=\varepsilon\Rightarrow\left\|\lambda\vec{f}(\vec{x})-\lambda\vec{L}_f\right\|<\varepsilon,$$

which satisfies de definition of limit. □

Proposition 2.2.3. *Let $\vec{f}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ and $\vec{g}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be two functions such that $\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})=\vec{L}_f$ and $\lim_{\vec{x}\to\vec{a}}\vec{g}(\vec{x})=\vec{L}_g$. Then, the limit of $\langle\vec{f}(\vec{x},\vec{g}(\vec{x})\rangle_I$ exists and it satisfies*

$$\lim_{\vec{x}\to\vec{a}}\langle\vec{f}(\vec{x}),\vec{g}(\vec{x})\rangle_I=\langle\vec{L}_f,\vec{L}_g\rangle_I. \tag{2.7}$$

Proof. Since the limits of $\vec{f}(\vec{x})$ and $\vec{g}(\vec{x})$ exist, they satisfy the condition 2.2.1 for any expression of ε and, in particular, there exist two numbers δ_f and δ_g such that

$$\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|<\min\left\{\frac{\left\|\vec{L}_f\right\|}{\left(2\left\|\vec{L}_f\right\|+1\right)\left\|\vec{L}_g\right\|}\varepsilon,1\right\}\qquad\left\|\vec{g}(\vec{x})-\vec{L}_f\right\|<\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}$$

We can do it because the norms of the limits exist (they are finite). As before, in order to be sure both condition are true at the same time, we will select as our delta the number $\delta=\min\{\delta_1,\delta_2\}$. Before starting using our conditions, first we will bound our expression with other known properties.

$$\left|\langle\vec{f}(\vec{x}),\vec{g}(\vec{x})\rangle_I-\langle\vec{L}_f,\vec{L}_g\rangle_I\right|=$$
$$\left|\langle\vec{f}(\vec{x}),\vec{g}(\vec{x})\rangle_I-\langle\vec{f}(\vec{x}),\vec{L}_g\rangle_I-\langle\vec{L}_f,\vec{g}(\vec{x})\rangle_I+\langle\vec{L}_f,\vec{L}_g\rangle_I+\langle\vec{L}_f,\vec{g}(\vec{x})\rangle_I-\langle\vec{L}_f,\vec{L}_g\rangle_I+\langle\vec{L}_g,\vec{f}(\vec{x})\rangle_I-\langle\vec{L}_f,\vec{L}_g\rangle_I\right|=$$
$$\left|\langle\vec{f}(\vec{x})-\vec{L}_f,\vec{g}(\vec{x})-\vec{L}_g\rangle_I+\langle\vec{L}_f,\vec{g}(\vec{x})-\vec{L}_g\rangle_I+\langle\vec{L}_g,\vec{f}(\vec{x})-\vec{L}_f\rangle_I\right|\leq$$
$$\left|\langle\vec{f}(\vec{x})-\vec{L}_f,\vec{g}(\vec{x})-\vec{L}_g\rangle_I\right|+\left|\langle\vec{L}_f,\vec{g}(\vec{x})-\vec{L}_g\rangle_I\right|+\left|\langle\vec{L}_g,\vec{f}(\vec{x})-\vec{L}_g\rangle_I\right|\leq$$
$$\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_f\right\|\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_g\right\|\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|$$

Where in the last step we have used the Cauchy-Swartz inequality. Now, we can use the fact that both functions individually can satisfy the condition we presented above. With that,

$$\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_f\right\|\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_g\right\|\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|<$$

$$1\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_f\right\|\left\|\vec{g}(\vec{x})-\vec{L}_g\right\|+\left\|\vec{L}_g\right\|\left\|\vec{f}(\vec{x})-\vec{L}_f\right\|\leq$$

$$\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}+\left\|\vec{L}_f\right\|\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}+\left\|\vec{L}_g\right\|\frac{\left\|\vec{L}_f\right\|\varepsilon}{\left(2\left\|\vec{L}_f\right\|+1\right)\left\|\vec{L}_g\right\|}=$$

$$\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}+\left\|\vec{L}_f\right\|\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}+\left\|\vec{L}_f\right\|\frac{\varepsilon}{2\left\|\vec{L}_f\right\|+1}=\frac{2\left\|\vec{L}_f\right\|+1}{2\left\|\vec{L}_f\right\|+1}\varepsilon=\varepsilon$$

Which satisfies the condition of limit. □

Corollary 2.2.4. *Let $\vec{f}:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be a function such that such that $\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})=\vec{L}_f$. Then, the limit of $\left\|\vec{f}(\vec{x})\right\|$ exists and it satisfies*

$$\lim_{\vec{x}\to\vec{a}}\left\|\vec{f}(\vec{x})\right\|=\left\|\vec{L}_f\right\|. \tag{2.8}$$

Corollary 2.2.5. *Let $\vec{f}:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be a function and a an accumulation point of Ω. Then, it converges in a to the point L_f if and only if each function $f_i(\vec{x})$ of coordinates $i=1,\dots,m$ converges to the correspondent coordinates L_i of L.*

Proof. Let us suppose each coordinate converges, that is,

$$\lim_{\vec{x}\to\vec{a}}f_i(\vec{x})=L_i\Leftrightarrow\lim_{\vec{x}\to\vec{a}}|f_i(\vec{x})-L_i|=0.$$

In this case,

$$\lim_{\vec{x}\to\vec{a}}\vec{f}(\vec{x})=\lim_{\vec{x}\to\vec{a}}\sum_{i=1}nf_i(\vec{x})\vec{e}_i=\sum_{i=1}n\left(\lim_{\vec{x}\to\vec{a}}f_i(\vec{x})\right)\vec{e}_i=\sum_{i=1}^{n}L_i\vec{e}_i=\vec{L}.$$

Let us suppose now we know the global function converges. Then, by the proposition 2.2.3, we have

$$\lim_{\vec{x}\to\vec{a}}f_i(\vec{x})=\lim_{\vec{x}\to\vec{a}}\langle\vec{f}(\vec{x}),\vec{e}_i\rangle_I=\langle\vec{L},\vec{e}_i\rangle_I=L_i,\forall i,$$

which implies the convergence of each coordinate. □

For scalar fields

Proposition 2.2.6. *Let $\vec{f}:\mathbb{R}^n\longrightarrow\mathbb{R}$ and $\vec{g}:\mathbb{R}^n\longrightarrow\mathbb{R}$ be two functions such that $\lim_{\vec{x}\to\vec{a}}f(\vec{x})=L_f$ and $\lim_{\vec{x}\to\vec{a}}g(\vec{x})=L_g$. Then, the limit of $f(\vec{x})/g(\vec{x})$ exists and it satisfies*

$$\lim_{\vec{x}\to\vec{a}}\frac{f(\vec{x})}{g(\vec{x})}=\frac{L_f}{L_g} \tag{2.9}$$

Proposition 2.2.7. *[1] If $\lim_{(x,y)\to(a,b)}f(x,y)=L$ and exist the following uni-dimensional limits*

$$\lim_{x\to a}f(x,y)\qquad\lim_{y\to bf(x,y)} \tag{2.10}$$

then the iterated limits exist and coincide, that is,

$$\lim_{x\to a}\lim_{y\to b}f(x,y)=\lim_{y\to b}\lim_{x\to a}f(x,y)=L. \tag{2.11}$$

The reciprocal is not always true.

It is important to note that, if the iterated limits do not exist, we cannot be sure that the general limit neither.

2.2.3 Methods to prove or disprove limits

To prove limits in scalar functions:

- Sandwich theorem
- If $\vec{x} \to \vec{0}$, we can write (in 2D) $x = \rho \cos\theta$ and $y = \rho \sin\theta$. If it exists for any value of θ then exists.
- If $\vec{x} \to \vec{a} \neq \vec{0}$. We can write (in 2D) $\Delta x = h_x \coloneqq \rho\cos\theta$ and we will have the same as the previous case.

To disprove limits in scalar functions:

- If we make the polar coordinates substitution and the final expression depends on the angle, the limit does not exist. It works for $\vec{x} \to \vec{0}$ and for $\vec{x} \to \vec{a} \neq \vec{0}$.
- If the iterated limits have different values, the limit does not exist.
- If we follow the path $y = y(x)$ and the final result depends on other parameters, the limits does not exist. Some typical substitutions are $y = ax$ and $y = ax^n$.

2.3 Continuity

2.3.1 Introduction

Definition 2.3.1. Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a point of Ω. We say $\vec{f}(\vec{x})$ *is continuous in the point* a if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta \mid \|\vec{x} - \vec{a}\| < \delta \Rightarrow \left\|\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right\| < \varepsilon \tag{2.12}$$

Note this definition is not the statement of a limit for two reasons. The first one is that a is required now to be a point of Ω, while in the definition of limit it is only required for a to be an accumulation point. The second one is that $\|\vec{x} - \vec{a}\|$ now can reach the value of 0 while in the limit it has to be greater. Even though these nuances, we can relate this concept with limits.

Proposition 2.3.1. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a point of Ω. If a is an accumulation point of Ω, then $\vec{f}$ is continuous in a if and only if*

$$\lim_{\vec{x} \to \vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a}) \tag{2.13}$$

Proposition 2.3.2. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a point of Ω. If a is an isolated point of Ω, then $\vec{f}$ is continuous in a.*

2.3.2 Basic properties

Sum of continuous, product by a scalar, dot product, and components.

2.3.3 Continuity in a set

Definition 2.3.2. Let $\vec{f} : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and $A \subseteq \mathbb{R}^n$ be a set. We say $\vec{f}(\vec{x})$ *is continue in* D if it is continuous in every point x of A. In other words, if it satisfies that

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta(\varepsilon, \vec{x}, \vec{x}') \mid \|\vec{x}' - \vec{x}\| < \delta \Rightarrow \left\|\vec{f}(\vec{x}') - \vec{f}(\vec{x})\right\| < \varepsilon, \forall \vec{x}, \vec{x}' \in A \tag{2.14}$$

Sum of continuous, product by a scalar, dot product, and components.

2.3.4 Topological properties

Proposition 2.3.3. *Let $\vec{f} : K \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function, with K a compact set. If f is continuous, then $\vec{f}(K)$ is compact.*

Proposition 2.3.4. *Let $\vec{f} : K \subset \mathbb{R} \longrightarrow \mathbb{R}^m$ be a function, with K a compact set. Then, the maxima and minima of f are in K.*

2.3.5 Some continuous functions

Example 2.3.1. [1] The identity function is continuous.

Example 2.3.2. [1] Linear transformations, that is, functions that satisfy $f(\vec{x}+\vec{y}) = f(\vec{x}) + f(\vec{y})$ are continuous.

Example 2.3.3. [1] Polynomials of several variables have the form

$$P(\vec{x}) = \sum_{k_1=0}^{p_1} \cdots \sum_{k_n=0}^{p_n} c_{k_1,\ldots,k_n} x_1^{k_1} \cdots , x_n^{k_n} \tag{2.15}$$

and are continuous in all $\mathbb{R}^n$.

Example 2.3.4. [1] Rational functions of polynomials are continuous always as the denominator is not 0.

2.3.6 Properties of continuous scalar fields

Proposition 2.3.5. *[1] Let $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar continuous function and a an interior point of D such that $f(\vec{a}) \neq 0$. Then, there exists a ball such that all its points have the same sign as $f(\vec{a})$.*

Proof. Is an exercise in the onenote. □

If a function is continuous in a point a, we can interchange limits with the function [**?**].

2.4 Uniform continuity

Definition 2.4.1. Let $\vec{f} : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and $A \subseteq F$ be a set. We say $\vec{f}(\vec{x})$ *is uniformly continuous in* A if it satisfies the following condition.

$$\forall \varepsilon \in \mathbb{R}^+ \exists \delta(\varepsilon) \mid \|\vec{x} - \vec{x}'\| < \delta \Rightarrow \left\| \vec{f}(\vec{x}) - \vec{f}(\vec{x}') \right\| < \varepsilon \tag{2.16}$$

The principal difference between continuous and uniformly continuous is that in this new case δ does not depend on the points.
Sum of continuous, product by a scalar, dot product, and components.

Theorem 2.4.1. *Let be $\vec{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a continuous function in a compact set $A \subset \mathbb{R}^n$. Then, $\vec{f}(\vec{x})$ is uniformly continuous in A.*

Bibliography

[1] Tom M Apostol. *Calculus, vol. 2.* Reverté, 1985.

Chapter 3

Functions in $\mathbb{R}^n$. Derivative

3.1 Derivative of scalar functions

We can represent scalar functions with the value $f(\vec{x})$ in an extra axis from the previous ones of $\vec{x}$. In particular, if the functions has two variables x and y, the values $f(x, y)$ can be represented in a three-dimensional space.
As we did with functions of one variable, when we represented a function with a graph we could wonder about what was the slope in a certain point. We can do the same with general scalar functions. However, as we see in the graph, now there is no a "slope" in a certain point because we can move now in more than one directions, so we should talk about "slopes". To specify the followed path, we could say that from a point a we go in a direction $\vec{v}$. If we make the magnitude of $\vec{v}$ to the infinitesimal order, we could see there is a slope in that direction.

Definition 3.1.1. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function, a an interior point of D, and $\vec{v}$ a vector of the associated vector space of $\mathbb{R}^n$. Then ,we define the *derivative of $f(\vec{x})$ in a with respect to $\vec{v}$* as

$$f'(a; \vec{v}) := \lim_{h \to 0} \frac{f(\vec{a} + h\vec{v}) - f(\vec{a})}{h}, \tag{3.1}$$

provided the limit exists.

Note that, different to one variable functions, we need now a to be an interior point, not only accumulation point. We will see why with the following proposition and example.

Proposition 3.1.1. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function, a an interior point of D, and $\vec{v}$ a vector of $\mathbb{R}^n$. If f is derivable at a, then the derivative is unique.*

As we see, if the point was not interior, the derivatives is not necessarily unique [1]. This differs from functions in one variable, whose derivatives only needed the points to be accumulation points.

Example 3.1.1. Let $D = \{(x, y) \in \mathbb{R}^2 \mid x = y\}$, $f : D \longrightarrow \mathbb{R}$ defined by $f(x) = x$, and $a = (a_1, a_2) \in D$, that is, $a_1 = a_2$. Then there are two linear functions L_1, L_2 that work.

Definition 3.1.2. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and $f'(a; \vec{v})$ the derivative of $f(\vec{x})$ in $a \in D$ with respect to a vector $\vec{v}$. If $\|\vec{v}\| = 1$, then we call it the *directional derivative of $f(\vec{x})$ in a along $\vec{v}$.*

Let us take the expression $f(\vec{a} + h\vec{v})$. If we keep $\vec{a}$ and $\vec{v}$ fixed and only change the value of h, we obtain a function $g(h)$ of only one variable. Therefore, in this case we can use all the properties we know about functions of one variable. Let us see an example of that

Theorem 3.1.2. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function, a an interior point of D, $\vec{v}$ a vector of $\mathbb{R}^n$, and $g(t) := f(\vec{a} + t\vec{v})$. If one of the derivatives exist, g' or f', the other exists and satisfies that*

$$g'(t) = f'(\vec{a} + t\vec{v}; \vec{v}). \tag{3.2}$$

In particular, when $t = 0$ we get $g'(0) = f'(\vec{a}; \vec{v})$.

Proof. Let us form the quotient of differences over h. Then, we get

$$\frac{g(h + d) - g(d)}{d} = \frac{f(\vec{a} + h\vec{v} + d\vec{v}) - f(\vec{a} - h\vec{v})}{d} \Rightarrow$$
$$\lim_{d \to 0} \frac{g(h + d) - g(d)}{d} = \lim_{d \to 0} \frac{f(\vec{a} + h\vec{v} + d\vec{v}) - f(\vec{a} - h\vec{v})}{d} \Rightarrow g'(t) = f'(\vec{a} + h\vec{v}, \vec{v}).$$

And we see that one exists if and only if exists the other. □

Theorem 3.1.3 (Intermediate value theorem for scalar functions)**.** *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function, a an interior point of D, and $\vec{v}$ a vector of $\mathbb{R}^n$. Let us suppose $f(\vec{a} + h\vec{v}; \vec{v})$ is derivable in $[\alpha, \beta]$. Then, there exists a real number $\theta \in (\alpha, \beta)$ such that*

$$\frac{f(\vec{a} + \beta\vec{y}) - f(\vec{a} + \alpha\vec{v})}{\beta - \alpha} = f'(\vec{a} + \theta\vec{v}, \vec{v}) \tag{3.3}$$

In particular, when $\alpha = 0$ we have $f(\vec{a} + \beta\vec{y}) - f(\vec{a}) = \beta f'(\vec{a} + \theta\vec{v}, \vec{v})$.

Proof. Let us denote $g(h) = f(\vec{a}+h\vec{v})$. Since $g(t)$ is differentiable in $[\alpha,\beta]$ (and therefore continuous), by the intermediate value theorem for functions of one variable [ ], there exists a real number $\theta \in (\alpha,\beta)$ such that

$$\frac{g(\beta)-g(\alpha)}{\beta-\alpha} = g'(\theta).$$

By the theorem 3.1.2, the derivative of g corresponds to the derivative of f so we obtain that

$$\frac{g(\beta)-g(\alpha)}{\beta-\alpha} = g'(\theta). = g'(\theta) \Rightarrow \frac{f(\vec{a}+\beta\vec{y})-f(\vec{a}+\alpha\vec{v})}{\beta-\alpha} = f'(\vec{a}+\theta\vec{v},\vec{v})$$

$\square$

Definition 3.1.3. Let a be an interior point an open set $\Omega \subseteq \mathbb{R}^n$, $f(\vec{x}_i) : S \longrightarrow \mathbb{R}$ a scalar function and $B = (e_1,\ldots,e_n)$ an orthonormal base the associated vector space $\mathbb{R}^n$ such that $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$. Then, we define the *partial derivative of $f(\vec{x})$ in a point a with respect to $\vec{e}_i$* as the directional derivative (if exists) of $f(\vec{x})$ in a along $\vec{e}_i$.

$$\frac{\partial f(\vec{a})}{\partial x_i} := f'(\vec{a},\vec{e}_i) \tag{3.4}$$

We also denote it by $D_i f(\vec{x})$, $\partial_{x_i} f(\vec{x})$, and $f'_{x_i}(\vec{x})$.

3.1.1 Differentiability

We have seen in functions of one variable that when a function is differentiable, then is continuous [ ]. We could think that this property would extend to multi-variable functions, so let us study if this happens when we deal with several variables.

Example 3.1.2. Let be the function of the example 2.2.1.

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2}, (x,y) \neq (0,0) \\ 0, (x,y) = (0,0) \end{cases}$$

We will study its directional derivative in the origin point with respect to a vector $\vec{v} = (a,b)$. First, we will see what happens when $a \neq 0$.

$$f'(\vec{0},\vec{v}) = \lim_{h\to 0} \frac{f(\vec{0}+h\vec{v})-f(\vec{0})}{h} = \lim_{h\to 0} \frac{f(h\vec{v})}{h} = \lim_{h\to 0} \frac{f((ha,hb))}{h} = \lim_{h\to 0} \frac{1}{h}\frac{ha(hb)^2}{(ha)^2+(hb)^4} =$$
$$\lim_{h\to 0} \frac{1}{h}\frac{h^3ab^2}{h^2(a^2+h^2b^4)} = \lim_{h\to 0} \frac{ab^2}{(a^2+h^2b^4)} = \frac{ab^2}{a^2} = \frac{b^2}{a}$$

Let us see what happens now if $a = 0$ and $b \neq 0$.

$$f'(\vec{0},\vec{v}) = \lim_{h\to 0} \frac{f(\vec{0}+h\vec{v})-f(\vec{0})}{h} = \lim_{h\to 0} \frac{f(h\vec{v})}{h} = \lim_{h\to 0} \frac{f((ha,hb))}{h} = \lim_{h\to 0} \frac{1}{h}\frac{ha(hb)^2}{(ha)^2+(hb)^4} =$$
$$\lim_{h\to 0} \frac{1}{h}\frac{h0(hb)^2}{(h0)^2+(hb)^4} = \lim_{h\to 0} \frac{1}{h}\frac{0}{(hb)^4} = 0$$

Finally, let us study what happens when $a = b = 0$. If the derivative exists, then it would exist in 0 for every possible vector, and as an analogy to one variable functions we could say is differentiable in 0.

$$f'(\vec{0},\vec{v}) = \lim_{h\to 0} \frac{f(\vec{0}+h\vec{v})-f(\vec{0})}{h} = \lim_{h\to 0} \frac{f(h\vec{v})}{h} = \lim_{h\to 0} \frac{f((ha,hb))}{h} = \lim_{h\to 0} \frac{f(\vec{0}}{h} = \lim_{h\to 0} \frac{0}{h} = 0$$

Therefore, f has a derivative in the origin in all directions. Now only remains to see if this function is continuous too in the origin, and to see that we will take the path $(x,y) = (y^2,y)$.

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4} = \lim_{y\to 0} \frac{y^2y^2}{y^4+y^4} = \lim_{y\to 0} \frac{y^4}{3y^4} = \lim_{y\to 0} \frac{1}{2} = \frac{1}{2} \neq f(0,0)$$

Although there exists a directional derivative in the origin for all directions, it is not a continuous function. This means we need another notion of differentiation that can guarantee us this property.

Definition 3.1.4. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a an interior point of D. We say f *is differentiable at* a if and only if there exists a linear function $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that

$$\lim_{\vec{v} \to \vec{0}} \frac{|f(\vec{a}+\vec{v}) - f(\vec{a}) - L(\vec{v})|}{\|\vec{v}\|} = 0. \tag{3.5}$$

In terms of delta epsilon definition, a differentiable function would satisfy that for every $\varepsilon > 0$ exists a $\delta(\varepsilon)$ such that

$$\|\vec{x} - \vec{a}\| < \delta \Rightarrow \frac{|f(\vec{x}) - \vec{a} - L(\vec{x} - \vec{a})|}{\|\vec{x} - \vec{a}\|} < \varepsilon \Leftrightarrow |f(\vec{x}) - \vec{a} - L(\vec{x} - \vec{a})| < \varepsilon \|\vec{x} - \vec{a}\|.$$

While the expression does not tend to zero, it has some real value, and we denote this value by $E(\vec{a};\vec{v})$. This way, the condition of differentiable function is equivalent to satisfy

1. $f(\vec{a}+\vec{v}) = f(\vec{a}) + L(\vec{v}) + \|\vec{v}\| E(\vec{a};\vec{v})$
2. $\lim_{\vec{v}\to\vec{0}} E(\vec{a};\vec{v}) = 0$

Theorem 3.1.4. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, a an interior point of D, and $\vec{v}$ a vector of $\mathbb{R}^n$. If $f(\vec{x})$ is differentiable at a, then there exist all the directional derivatives in a and it satisfies*

$$f'(\vec{a};\vec{v}) = L(\vec{v}) = \sum_{i=1}^{n} v_i \frac{\partial f(\vec{a})}{\partial x_i} \tag{3.6}$$

Proof. If f is differentiable in a, by the definition 3.1.4, we have that $f(\vec{a}+\vec{v}) = f(\vec{a}) + \langle \vec{T}, \vec{v}\rangle_I + \|\vec{v}\| E(\vec{a};\vec{v})$. If we take $\vec{v} = h\vec{y}$, we get

$$f(\vec{a}+h\vec{y}) - f(\vec{a}) = \langle \vec{T}, h\vec{y}_I + \|h\vec{y}\| E(\vec{a};h\vec{y}) = h\langle \vec{T}, \vec{y}\rangle_I + |h| \|\vec{y}\| E(\vec{a};\vec{v}) \Rightarrow$$
$$\frac{f(\vec{a}+h\vec{y}) - f(\vec{a})}{h} = \langle \vec{T}, \vec{y}\rangle_I + \frac{|h|}{h} \|\vec{y}\| E(\vec{a};h\vec{y}).$$

If we now apply the limit to this equality, by the definition 3.1.2 the left side is the directional derivative, and by the definition 3.1.4, the last part of the right side will tend to zero (although the part of $|h|/h$ does not have limit, since it is bounded, the product with E will tend to 0).

$$\lim_{h\to 0} \frac{f(\vec{a}+h\vec{y}) - f(\vec{a})}{h} = \lim_{h\to 0} \langle \vec{T}, \vec{y}\rangle_I + \frac{|h|}{h} \|\vec{y}\| E(\vec{a};h\vec{y}) \Rightarrow f'(\vec{a};\vec{y}) = \langle \vec{T}, \vec{y}\rangle_I + 0$$

Now we know this equality, if we take an orthonormal basis of $\mathbb{R}^n$, we have

$$f'(\vec{a};\vec{v}) = \langle \vec{T}, \vec{v}\rangle_I = \left\langle \vec{T}, \sum_{i=1}^{n} v_i \vec{e}_i \right\rangle_I = \sum_{i=1}^{n} v_i \langle \vec{T}, \vec{e}_i\rangle_I = \sum_{i=!}^{n} v_i f'(\vec{a};\vec{e}_i) = \sum_{i=1}^{n} v_i \frac{\partial f(\vec{a})}{\partial x_i}$$

In this case (where the function is differentiable), $\langle \vec{T}, \vec{v}\rangle_I$ coincides with $\langle \vec{\nabla} f(\vec{a}), \vec{v}\rangle_I$. ☐

In practice, to see if a function is differentiable we calculate the following limit.

$$\lim_{\vec{v}\to\vec{0}} \frac{f(\vec{a}+\vec{v}) - f(\vec{a}) - \langle \vec{\nabla} f(\vec{a}), \vec{v}\rangle_I}{\|\vec{v}\|} \tag{3.7}$$

If and only if the limit tends to zero, the function is differentiable.

Theorem 3.1.5. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a an interior point of D. If f is differentiable at a, then f is continuous at a.*

Proof. Since f is differentiable, by the theorem 3.1.4, $f(\vec{a}+\vec{v}) = f(\vec{a}) + \langle \vec{\nabla} f(\vec{a}), \vec{v}\rangle_I + \|\vec{v}\| E(\vec{a};\vec{v})$. If we make $\vec{v}$ tend to $\vec{0}$, we obtain

$$\lim_{\vec{v}\to\vec{0}} f(\vec{a}+\vec{v}) = \lim_{\vec{v}\to\vec{0}} f(\vec{a}) + \langle \vec{\nabla} f(\vec{a}), \vec{v}\rangle_I + \|\vec{v}\| E(\vec{a};\vec{v}) = f(\vec{a}) + 0 + 0 \Rightarrow \lim_{\vec{v}\to\vec{0}} f(\vec{a}+\vec{v}) = f(\vec{a})$$

Which satisfies the definition 2.3.1 of continuous function. ☐

We have seen that differentiability implies the continuity of a function and that exist directional derivatives of a function. However, we have not seen what property could imply the differentiablity of a function. We will see it now.

Theorem 3.1.6. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a an interior point of D. Let $\partial f(\vec{a})/\partial x_1, \dots, \partial f(\vec{a})/\partial x_n$ be its partial derivatives in a. If they are continuous, then f is differentiable in a.*

Proof. We want to see if the function is differentiable, that is, if $f(\vec{a}+\vec{h}) = f(\vec{a}) + \langle \vec{\nabla} f(\vec{a}) \rangle_I + \left\|\vec{h}\right\| E(\vec{a};\vec{h})$. To do this, we will denote $\vec{h} = \lambda\vec{u}$, with $\|\vec{u}\| = 1$, so $\left\|\vec{h}\right\| = \lambda$. Taking an orthonormal basis, we have too that $\vec{u} = u_1\vec{e}_1 + \cdots + u_n\vec{e}_n$.
Now we will construct the following vectors.

$$\vec{v_0} = \vec{0}, \qquad \vec{v_1} = \vec{0} + u_1\vec{e}_1 = \vec{v}_0 + u_1\vec{e}_1, \qquad \vec{v_2} = \vec{0} + u_1\vec{e}_1 + u_2\vec{e}_2 = \vec{v}_1 + u_2\vec{e}_2,$$
$$\vec{v_n} = \vec{0} + u_1\vec{e}_1 + \cdots + u_n\vec{e}_n = \vec{v}_{n-1} + u_n\vec{e}_n = \vec{u}$$

From these vectors, we can now express $f(\vec{a}+\lambda\vec{u}) - f(\vec{a})$ as a telescopic sum.

$$f(\vec{a}+\lambda\vec{u}) - f(\vec{a}) = f(\vec{a}+\lambda\vec{v}_n) - f(\vec{a}+\lambda\vec{v}_{n-1}) + f(\vec{a}+\lambda\vec{v}_{n-1}) + \cdots + f(\vec{a}+\lambda\vec{v}_0) =$$
$$\sum_{i=1}^{n} f(\vec{a}+\lambda\vec{v}_i) - f(\vec{a}+\lambda\vec{v}_{i-1}) = \sum_{i=1}^{n} f\left(\vec{a}+\lambda\vec{v}_{i-1} + \lambda u_i\vec{e}_i\right) - f(\vec{a}+\lambda\vec{v}_{i-1})$$

In each element there is a difference of functions where the variable is separated by just one axis. Therefore, we can apply the properties for functions in one variable and, in particular, the intermediate value theorem [ ]. With that, we can rephrase the sum as follows.

$$f(\vec{a}+\lambda\vec{u}) - f(\vec{a}) = \sum_{i=1}^{n} f\left(\vec{a}+\lambda\vec{v}_{i-1} + \lambda u_i\vec{e}_i\right) - f(\vec{a}+\lambda\vec{v}_{i-1}) =$$
$$\sum_{i=1}^{n} \lambda u_i \frac{\partial f(\vec{c}_i)}{\partial x_i}, \ \vec{c} = \vec{a} + \lambda\vec{v}_{i-1} + \theta\lambda u_i\vec{e}_i, \ \theta \in (0,1) \Rightarrow$$

Therefore,

$$f(\vec{a}+\lambda\vec{u}) - f(\vec{a}) - \langle \vec{\nabla} f(\vec{a}), \lambda\vec{u} \rangle_I = \sum_{i=1}^{n} \lambda u_i \frac{\partial f(\vec{c}_i)}{\partial x_i} - \langle \vec{\nabla} f(\vec{a}), \lambda\vec{u} \rangle_I \Rightarrow$$
$$\|\lambda\vec{u}\| E(\vec{a};\lambda\vec{u}) = \sum_{i=1}^{n} \lambda u_i \frac{\partial f(\vec{c}_i)}{\partial x_i} - \sum_{i=1}^{n} \lambda u_i \frac{\partial f(\vec{a}_i)}{\partial x_i} \Rightarrow$$
$$\lambda \cdot 1 \cdot E(\vec{a};\lambda\vec{u}) = \lambda \sum_{i=1}^{n} u_i \left(\frac{\partial f(\vec{c}_i)}{\partial x_i} - \frac{\partial f(\vec{a})}{\partial x_i} \right) \Rightarrow E(\vec{a};\lambda\vec{u}) = \sum_{i=1}^{n} u_i \left(\frac{\partial f(\vec{c}_i)}{\partial x_i} - \frac{\partial f(\vec{a})}{\partial x_i} \right)$$

Then, if we prove this last equality, we will have the function is differentiable, completing the theorem. To see this, it needs to satisfy the condition of E, which is that when $\|\vec{v}\| \to 0$, then $E \to 0$. We have seen that $\|\vec{v}\| = \lambda \|\vec{u}\|$, so it is equivalent to make $\lambda \to 0$. Knowing that partial derivatives are continuous ($\lim_{\vec{x}\to\vec{a}} \partial_{x_i} f(\vec{x}) = \partial_{x_i} f(\vec{a})$),

$$\lim_{\lambda\to 0} \sum_{i=1}^{n} u_i \left(\frac{\partial f(\vec{c}_i)}{\partial x_i} - \frac{\partial f(\vec{a})}{\partial x_i} \right) = \lim_{\lambda\to 0} \sum_{i=1}^{n} u_i \left(\frac{\partial f(\vec{a}+\lambda\vec{v}_{i-1}+\theta\lambda u_i\vec{e}_i)}{\partial x_i} - \frac{\partial f(\vec{a})}{\partial x_i} \right) =$$
$$\sum_{i=1}^{n} u_i \left(\frac{\partial f(\vec{a})}{\partial x_i} - \frac{\partial f(\vec{a})}{\partial x_i} \right) = 0$$

Showing it satisfies the condition of E and therefore of differentiability. □

Proposition 3.1.7. *Let be the following functions.*

$$\begin{array}{rl} \vec{r} : \Omega_1 \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n \\ t \longmapsto \vec{r}(t) \end{array} \qquad \begin{array}{rl} f : \Omega_2 \supseteq f(\Omega_1) \longrightarrow \mathbb{R} \\ \vec{r}(t) \longmapsto f(\vec{r}(t)) \end{array}, \tag{3.8}$$

with Ω_1, Ω_2 open sets. If $\vec{r}(t)$ exist and f is differentiable in $\vec{r}(t)$, then the derivative of f with respect to t exist and it is expressed as

$$\frac{df(\vec{r}(dt))}{t} = \left\langle \vec{\nabla} f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle_I \tag{3.9}$$

Proof. Let us denote $f(\vec{r}(t)) = g(t)$ and $\vec{r}(t+h) - \vec{r}(t) = \vec{v}$. Since f is differentiable, we have

$$g(t+h) - g(t) = f(\vec{r}(t+h)) - f(\vec{r}(t)) = f(\vec{r}(t) + \vec{r}(t+h) - \vec{r}(t)) - f(\vec{r}(t)) =$$
$$f(\vec{r}(t) + \vec{v}) - f(\vec{r}(t)) = \langle \vec{\nabla} f(\vec{r}(t)), \vec{v} \rangle_I + \|\vec{v}\|\, E(\vec{r}(t); \vec{v}) \Rightarrow$$
$$\frac{g(t+h) - g(t)}{h} = \left\langle \vec{\nabla} f(\vec{r}(t)), \frac{\vec{r}(t+h) - \vec{r}(t)}{h} \right\rangle_I + \frac{\|\vec{r}(t+h) - \vec{r}(t)\|}{h} E(\vec{r}(t); \vec{r}(t+h) - \vec{r}(t)) \Rightarrow$$
$$\lim_{h\to 0} \frac{g(t+h) - g(t)}{h} = \lim_{h\to 0} \left\langle \vec{\nabla} f(\vec{r}(t)), \frac{\vec{r}(t+h) - \vec{r}(t)}{h} \right\rangle_I +$$
$$\lim_{h\to 0} \frac{\|\vec{r}(t+h) - \vec{r}(t)\|}{h} E(\vec{r}(t); \vec{r}(t+h) - \vec{r}(t)) \Rightarrow g'(t) = \left\langle \vec{\nabla} f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle_I +$$
$$\lim_{h\to 0} \frac{|h|}{h} \left\| \frac{\vec{r}(t+h) - \vec{r}(t)}{h} \right\| E(\vec{r}(t); \vec{r}(t+h) - \vec{r}(t)) \Rightarrow g'(t) =$$
$$\left\langle \vec{\nabla} f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle_I + \lim_{h\to 0} \frac{|h|}{h} \cdot \left\| \frac{d\vec{r}(t)}{dt} \right\| \cdot E(\vec{r}(t); \vec{0}) = \left\langle \vec{\nabla} f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle_I$$

Where we have use that the derivative of $\vec{r}(t)$ exists, so $\vec{r}(t)$ is continuous and $\lim_{h\to} \vec{r}(t+h) = \vec{r}(t)$. □

3.1.2 Differential of scalar functions

We know if a function $f(\vec{x})$ is differentiable, it will satisfy that

$$f(\vec{x} + \vec{v}) - f(\vec{x}) = \langle \vec{\nabla} f(\vec{x}), \vec{v} \rangle_I + \|\vec{v}\|\, E(\vec{x}; \vec{v}). \tag{3.10}$$

In this case, we can denote $\vec{x}' = \vec{x} + \vec{v}$ such that $\vec{v} = \vec{x}' - \vec{x} := \Delta\vec{x}$ and $\Delta f(\vec{x}) = f(\vec{x} + \vec{v}) - f(\vec{x})$. With that the condition of $\|\vec{v}\| \to 0$ is translated to $\|\Delta\vec{x}\| \to 0$, and since the error E tends to 0 when that occurs, we can write the following relation.

$$\lim_{\|\Delta\vec{x}\|\to 0} \Delta f(\vec{x}) = \lim_{\|\Delta\vec{x}\|\to 0} \langle \vec{\nabla} f(\vec{x}), \Delta\vec{x} \rangle_I \tag{3.11}$$

Which can be expressed as

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \tag{3.12}$$

This expression is important and needs a definition.

Definition 3.1.5. Let $f(\vec{x}) : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$. We define the *differential of* f as the limit of $\Delta f(\vec{x})$ when $\|\Delta\vec{x}\|$ tends to 0.

We can interpret the differential geometrically. To see it, let us study the particular case of a function of two variables, that is, $f(x, y) : \mathbb{R}^2 \longrightarrow \mathbb{R}$. We commented before that we could the image of each point in another axis, so we would have the following figure. Let us suppose we know at a point (x_0, y_0) what is the value of the function $f(x_0, y_0)$ and we want to know the value of another point (x, y). If it is an arbitrary function, we can't know it, since it could diverge or have another discontinuity. However, if f is differentiable, we can relate $f(x_0, y_0)$ and $f(x, y)$ by the expression of the definition 3.1.4, knowing there is an small finite error E. This error gets smaller when these two points are closer, so if (x, y) is very closed to (x_0, y_0) we can write

$$f(x, y) \approx f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0). \tag{3.13}$$

Let us denote now the $z = f(x, y) - E$. This point would be approximately equal to $f(x, y)$ at near points of $(x_0, y0)$ and exactly $f(x_0, y_0)$ at the point (x_0, y_0), since the error tends to zero. If we denote $z_0 = z(x_0, y_0) = f(x_0, y_0)$, therefore, we obtain an actual equality of the form

$$0 = (z - z_0) - \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) - \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0). \tag{3.14}$$

This is the equation of a plane, and in particular of a tangent plane at the point (x_0, y_0). As we commented, the values of z are approximately $f(x, y)$ at near points, and we can see this with the plane too.

Proposition 3.1.8. *Let $f : \Omega_1 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and $g : \Omega_2 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be two functions, with Ω_1, Ω_2 open set, and a an interior point of Ω_1 and Ω_2. If f, g are differentiable in a, then $f + g$ is differentiable in a and $d(f + g) = df + dg$.*

Proposition 3.1.9. *Let $f : \Omega_1 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and $g : \Omega_2 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be two functions, with Ω_1, Ω_2 open set, and a an interior point of Ω_1 and Ω_2. If f, g are differentiable in a, then fg is differentiable in a and $d(fg) = gdf + fdg$.*

3.1.3 Geometric interpretation of gradient. Level curves

Let $f(x, y) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a differentiable function. Let us take three set of points: the set of points at which $f(x, y)) = z_1$, the set of points at which $f(x, y) = z_2$, and the set of points at which $f(x, y) = z_3$ (taking $z_1 > z_2 > z_3$). These sets describe curves, and we call it *level curves*, and we will denote them respectively bu C_1, C_2, and C_3.
Now, let us take an infinitesimal step from a point (x, y) of C_2 to another point (x', y') of C_2. The total step is $d\vec{x}$, and since is differentiable, we can write $df = f(x', y') - f(x, y) = \langle \vec{\nabla} f(\vec{x}), \Delta\vec{x} \rangle_I$. However, these two points correspond to the C_2, so they will have the same value and the difference of the left side is 0. This means that, if we move in a path where the function is constant, we will have

$$0 = \langle \vec{\nabla} f(\vec{x}), \Delta\vec{x} \rangle_I \Rightarrow \vec{\nabla} f(\vec{x}) \perp d\vec{x}. \tag{3.15}$$

This means the gradient vector is perpendicular to the path where the function is constant. With that its direction can have two possibilities, to C_1 or C_3. Let us now going from this original point but now going to an arbitrary point (x'', y'') (again infinitely closed). Now, the change of values will be

$$f(\vec{x}'') - f(\vec{x}) = \langle \vec{\nabla} f(\vec{x}), \Delta\vec{x} \rangle_I = \left\| \vec{\nabla} f(\vec{x}) \right\| \|d\vec{x}\| \cos\theta, \tag{3.16}$$

which gets its maximum value when $\theta = 0$. This means the path required to get the maximum positive variation (in this case going to C_1) has the same direction of the gradient vector. Therefore, $\vec{\nabla} f(\vec{x})$ will point to C_1.

3.2 Derivative of vector fields

Definition 3.2.1. Let $\vec{f}(\vec{x}) : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and a point a an interior point of D. We define the *derivative of f at a point a of direction $\vec{v}$*, denoted by $\vec{f}'(\vec{a}; \vec{v})$ as the following expression.

$$\vec{f}'(\vec{a}; \vec{v}) = \lim_{h \to 0} \frac{\vec{f}(\vec{a} + h\vec{v}) - \vec{f}(\vec{a})}{h} \tag{3.17}$$

Proposition 3.2.1. *Let $\vec{f}(\vec{x}) : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function, a a point in D, and an orthonormal base $(\vec{e}_1, \ldots, \vec{e}_m)$ of $\mathbb{R}^m$. Then, the derivative of f at a in the direction of $\vec{v}$ satisfies*

$$\vec{f}'(\vec{a}; \vec{v}) = \sum_{i=1}^{n} f_i'(\vec{a}; \vec{v}) \vec{e}_i \tag{3.18}$$

Proof. Since we have an orthonormal base for $\mathbb{R}^m$ we can decompose the vector $\vec{f}'(\vec{a};\vec{v})$ in its different components. With that, we get

$$\vec{f}'(\vec{a};\vec{v}) = \lim_{h\to 0}\frac{1}{h}\left(\sum_{i=1}^{m} f_i(\vec{a}+h\vec{v})\vec{e}_i - \sum_{i=1}^{m} f_i(\vec{a})\vec{e}_i\right) = \sum_{i=1}^{m}\lim_{h\to 0}\frac{f_i(\vec{a}+h\vec{v})-\vec{f}_i(\vec{a})}{h}\vec{e}_i =$$
$$\sum_{i=1}^{m}\vec{e}_i\lim_{h\to 0}\frac{f_i(\vec{a}+h\vec{v})-\vec{f}_i(\vec{a})}{h} = \sum_{i=1}^{m} f_i'(\vec{a};\vec{v})\vec{e}_i,$$

which is what we wanted to prove. □

3.2.1 Differentiable vector fields

Definition 3.2.2. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a an interior point of D. We say f *is differentiable at* a if and only if there exists a linear function $J : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{\vec{v}\to\vec{0}}\frac{\|f(\vec{a}+\vec{v})-f(\vec{a})-J(\vec{v})\|}{\|\vec{v}\|} = 0. \tag{3.19}$$

In terms of delta epsilon definition, a differentiable function would satisfy that for every $\varepsilon > 0$ exists a $\delta(\varepsilon)$ such that

$$\|\vec{x}-\vec{a}\| < \delta \Rightarrow \frac{\|f(\vec{x})-\vec{a}-J(\vec{x}-\vec{a})\|}{\|\vec{x}-\vec{a}\|} < \varepsilon \Leftrightarrow \|f(\vec{x})-\vec{a}-J(\vec{x}-\vec{a})\| < \varepsilon\,\|\vec{x}-\vec{a}\|\,.$$

While the expression does not tend to zero, it has some real value, and we denote this value by $\vec{E}(\vec{a};\vec{v})$. This way, the condition of differentiable function is equivalent to satisfy

1. $\vec{f}(\vec{a}+\vec{v}) = \vec{f}(\vec{a}) + J(\vec{v}) + \|\vec{v}\|\,\vec{E}(\vec{a};\vec{v})$
2. $\lim_{\vec{v}\to\vec{0}}\vec{E}(\vec{a};\vec{v}) = 0$

Note that since now we want $\hat{T}\vec{v}$ to be a vector, $\hat{T}$ needs to be a $m\times n$ matrix.

Proposition 3.2.2. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a an interior point of D. Then, $\vec{f}$ is differentiable at a if and only if all components of $\vec{f}$ are differentiable at a.*

Theorem 3.2.3. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a differentiable function in an interior point a of D. Then, $\vec{f}'(\vec{a};\vec{v}) = \hat{T}\vec{v}$ and, if $\vec{f} = (f_1,\ldots,f_m)$ and $\vec{v} = (y_1,\ldots,y_n)$ (with orthonormal basis for $\mathbb{R}^m$ and $\mathbb{R}^n$), then*

$$\hat{T}\vec{v} = \vec{f}'(\vec{a};\vec{v}) = \sum_{i=1}^{m}\langle\vec{\nabla}f_i(\vec{a}),\vec{v}\rangle_I\vec{e}_i \tag{3.20}$$

Proof. Since $\vec{f}(\vec{x})$ is differentiable, we know $\vec{f}(\vec{a}+\vec{v})-\vec{f}(\vec{a}) = \hat{T}\vec{v}+\|\vec{v}\|\,\vec{E}(\vec{a};\vec{v})$. Now, let us denote $\vec{v} = h\vec{y}$, such that

$$\vec{f}(\vec{a}+h\vec{y})-\vec{f}(\vec{a}) = \hat{T}(h\vec{y}) + \|h\vec{y}\|\,\vec{E}(\vec{a};h\vec{y})$$

As we did with scalar field, we can manipulate the h and apply the limit.

$$\vec{f}(\vec{a}+h\vec{y})-\vec{f}(\vec{a}) = h\hat{T}\vec{y} + |h|\,\|\vec{y}\|\,\vec{E}(\vec{a};h\vec{y}) \Rightarrow$$
$$\frac{\vec{f}(\vec{a}+h\vec{y})-\vec{f}(\vec{a})}{h} = \hat{T}\vec{y} + \frac{|h|}{h}\vec{E}(\vec{a};h\vec{y}) \Rightarrow$$
$$\lim_{h\to 0}\frac{\vec{f}(\vec{a}+h\vec{y})-\vec{f}(\vec{a})}{h} = \lim_{h\to 0}\hat{T}\vec{y} + \lim_{h\to 0}\frac{|h|}{h}\vec{E}(\vec{a};h\vec{y})$$

The expression of $|h|/h$ has not limit, but it is bounded, and since is multiplied by something than tends do zero, the whole limit will tend to zero [ ]. Therefore, we get

$$\vec{f}'(\vec{a};\vec{y}) = \hat{T}\vec{y} \Leftrightarrow \vec{f}'(\vec{a};\vec{v}) = \hat{T}\vec{v}.$$

And finally, we have seen previously that with an orthonormal base for $\mathbb{R}^m$ we can write the derivative in terms of the derivatives of its components, and with an orthonormal base for $\mathbb{R}^n$ we can write the derivative of each component as the dot product of its gradient vector and the vector $\vec{v}$. Hence, we obtain

$$\hat{T}\vec{v} = \vec{f}'(\vec{a};\vec{v}) = \sum_{i=1}^{m} f_i'(\vec{a};\vec{v})\vec{e}_i = \sum_{i=1}^{m} \langle \vec{\nabla} f_i(\vec{a}), \vec{v} \rangle_I \vec{e}_i,$$

which is what we wanted. □

Now, to determine T, we just need to establish the condition it has to satisfy.

$$\vec{f}'(\vec{a};\vec{v}) = \begin{pmatrix} \langle \vec{\nabla} f_1(\vec{a}), \vec{v} \rangle_I \\ \vdots \\ \langle \vec{\nabla} f_m(\vec{a}), \vec{v} \rangle_I \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1(\vec{a})}{\partial x_1} & \cdots & \dfrac{\partial f_1(\vec{a})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(\vec{a})}{\partial x_1} & \cdots & \dfrac{\partial f_m(\vec{a})}{\partial x_n} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \hat{T}\vec{v} \tag{3.21}$$

Therefore, $\hat{T}$ is the matrix of the left side, and it is called *the Jacobian matrix of the function* $\vec{f}(\vec{x})$ *at the point* a. It can also be denoted by $J\left[\vec{f}'(\vec{a})\right]$.

Theorem 3.2.4. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a differentiable function and a an interior point of D. If $\vec{f}$ is differentiable in a, then $\vec{f}$ is continuous in a.*

Proof. Since $\vec{f}(\vec{x})$ is differentiable, we know that

$$\vec{f}(\vec{a}+\vec{v}) = \vec{f}(\vec{a}) + J\left[\vec{f}'(\vec{a})\right]\vec{v} + \|\vec{v}\|\,\vec{E}(\vec{a};\vec{v}),$$

where the Jacobian matrix is a constant, and $\vec{E}$ is something that tends to zero when $\vec{V}$ tends to zero. Therefore, if we make this equality tend to zero, we finally get

$$\lim_{\vec{v}\to\vec{0}} \vec{f}(\vec{a}+\vec{v}) = \lim_{\vec{v}\to 0} \vec{f}(\vec{a}) + J\left[\vec{f}'(\vec{a})\right]\vec{v} + \|\vec{v}\|\,\vec{E}(\vec{a};\vec{v}) = \vec{f}(\vec{a})$$

which is the definition of continuous function (Def 2.3.1). □

3.2.2 Differential of vector fields

Define it

Proposition 3.2.5. *Let $\vec{f} : \Omega_1 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $\vec{g} : \Omega_2 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be two functions, with Ω_1, Ω_2 open set, and a an interior point of Ω_1 and Ω_2. If $\vec{f}, \vec{g}$ are differentiable in a, then $\vec{f}+\vec{g}$ is differentiable in a and $d(\vec{f}+\vec{g}) = d\vec{f} + d\vec{g}$.*

Proposition 3.2.6. *Let $\vec{f} : \Omega_1 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $\vec{g} : \Omega_2 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be two functions, with Ω_1, Ω_2 open set, and a an interior point of Ω_1 and Ω_2. If $\vec{f}, \vec{g}$ are differentiable in a, then $\vec{f}\vec{g}$ is differentiable in a and $d(\vec{f}+\vec{g}) = \vec{g}d\vec{f} + \vec{f}d\vec{g}$.*

3.2.3 Chain rule for vector fields

Proposition 3.2.7. *Let $\vec{f}(\vec{x}) : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a differentiable function in a point a in D. Then, $\|\vec{f}'(\vec{a};\vec{v})\| / \|\vec{v}\|$ is bounded.*

Proof. Since the function is differentiable, the norm of its derivative can be expressed as the norm of the directional derivatives of its components.

$$\|\vec{f}'(\vec{a};\vec{v})\| = \left\| \sum_{i=1}^{m} \langle \vec{\nabla} f_i(\vec{a}), \vec{v} \rangle_I \vec{e}_i \right\| \leq \sum_{i=1}^{m} \left\| \langle \vec{\nabla} f_i(\vec{a}), \vec{v} \rangle_I \vec{e}_i \right\| = \sum_{i=1}^{m} \left\| \langle \vec{\nabla} f_i(\vec{a}), \vec{v} \rangle_I \right\|$$

Now, by the inequality of Cauchy-Swchartz,

$$\|\vec{f}'(\vec{a};\vec{v})\| \leq \sum_{i=1}^{m} \left\| \langle \vec{\nabla} f_i(\vec{a}), \vec{v} \rangle_I \right\| \leq \sum_{i=1}^{m} \left\| \vec{\nabla} f_i(\vec{a}) \right\| \|\vec{v}\| = \|\vec{v}\| \sum_{i=1}^{m} \left\| \vec{\nabla} f_i(\vec{a}) \right\| \Rightarrow$$
$$\frac{\vec{f}'(\vec{a}+\vec{v})}{\|\vec{v}\|} \leq \sum_{i=1}^{m} \left\| \vec{\nabla} f_i(\vec{a}) \right\| := M_f(\vec{a}),$$

which generates a bound for $\|\vec{f}'(\vec{a};\vec{v})\| / \|\vec{v}\|$. □

Proposition 3.2.8. *Let $\vec{g} : A \subseteq \mathbb{R}^l \longrightarrow \mathbb{R}^n$ and $\vec{f} : B \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be two functions. Let a be an interior point of A, and let $b = \vec{g}(\vec{a})$ be an interior point of B. If $\vec{g}$ is differentiable at a and $\vec{f}$ differentiable at b, then $(\vec{f} \circ \vec{g})$ is differentiable at a and*

$$\left[(\vec{f} \circ \vec{g})(\vec{a})\right] = \left[\vec{f}'(\vec{g}(\vec{a}))\right] [\vec{g}'(\vec{a})]. \tag{3.22}$$

Proof. Let us denote $\vec{g}(\vec{a})$ by $\vec{b}$. Since $\vec{g}(\vec{x})$ is differentiable in a and $\vec{f}(\vec{x})$ is differentiable in $\vec{b}$, they will satisfy (Def 3.2.2)

$$\vec{g}(\vec{a}+\vec{v}) - \vec{g}(\vec{a}) = [\vec{g}'(\vec{a})]\,\vec{v} + \|\vec{v}\|\,\vec{E}_g(\vec{a};\vec{v})$$
$$\vec{f}(\vec{b}+\vec{w}) - \vec{f}(\vec{b}) = \left[\vec{f}'(\vec{b})\right] \vec{w} + \|\vec{w}\|\,\vec{E}_f(\vec{b};\vec{w}).$$

Let us now calculate the difference $(\vec{f} \circ \vec{g})(\vec{a}+\vec{v}) - (\vec{f} \circ \vec{g})(\vec{a})$ to see if it satisfies too the definition of differentiable function.

$$(\vec{f} \circ \vec{g})(\vec{a}+\vec{v}) - (\vec{f} \circ \vec{g})(\vec{a}) = \vec{f}(\vec{g}(\vec{a}) + \vec{g}(\vec{a}+\vec{v}) - \vec{g}(\vec{a})) - \vec{f}(\vec{g}(\vec{a})) =$$
$$\left[\vec{f}'(\vec{g}(a))\right] (\vec{g}(\vec{a}+\vec{v}) - \vec{g}(\vec{a})) + \|\vec{w}\|\,\vec{E}_f(\vec{b};\vec{w}) =$$
$$\left[\vec{f}'(\vec{g}(a))\right] ([\vec{g}'(\vec{a})]\,\vec{v} + \|\vec{v}\|\,\vec{E}_g(\vec{a};\vec{v})) + \|\vec{w}\|\,\vec{E}_f(\vec{b};\vec{w})$$
$$\left[\vec{f}'(\vec{g}(a))\right] [\vec{g}'(\vec{a})]\,\vec{v} + \|\vec{v}\| \left[\vec{f}'(\vec{g}(a))\right] \vec{E}_g(\vec{a};\vec{v}) + \|\vec{w}\|\,\vec{E}_f(\vec{b};\vec{w})$$

As we can see, the matrix product forms a new matrix that multiplies $\vec{v}$, satisfying the first component of the first condition. To satisfy the whole condition, we need to express the second component as a function $\|\vec{v}\|\,\vec{E}_{\vec{f}\circ\vec{g}}(\vec{a};\vec{v})$ that tends to zero.

$$\|\vec{v}\| \left[\vec{f}'(\vec{g}(a))\right] \vec{E}_g(\vec{a};\vec{v}) + \|\vec{w}\|\,\vec{E}_f(\vec{b};\vec{w}) = \|\vec{v}\| \left(\left[\vec{f}'(\vec{g}(a))\right] \vec{E}_g(\vec{a};\vec{v}) + \frac{\|\vec{w}\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) \right)$$

The left side in the parenthesis tends to zero because is the product of something bounded by $\vec{E}_g$, that tends to zero because $\vec{g}(\vec{x})$ is differentiable. Then, it only remains the right side. We are making $\vec{v} \to 0$, and since $\vec{w} = \vec{g}(\vec{a}+\vec{v}) - \vec{g}(\vec{a})$ and the function is continuous (because is differentiable) $\vec{w} \to 0$ too. However, we know which part will win, because this limit is equivalent to

$$\frac{\|\vec{w}\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) = \frac{\|\vec{g}(\vec{a}+\vec{v}) - \vec{g}(\vec{a})\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) = \frac{\left\| [\vec{g}'(\vec{a})]\,\vec{v} + \|\vec{v}\|\,\vec{E}_g(\vec{a};\vec{v}) \right\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}).$$

Now, applying some inequalities and the proposition 3.2.7, we get

$$\frac{\left\| [\vec{g}'(\vec{a})]\,\vec{v} + \|\vec{v}\|\,\vec{E}_g(\vec{a};\vec{v}) \right\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) \leq \frac{\|[\vec{g}(\vec{a})]\,\vec{v}\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) + \left\| \vec{E}_g(\vec{a};\vec{v}) \right\| \vec{E}_f(\vec{b};\vec{w}) \leq$$
$$\frac{M_g(\vec{a})\,\|\vec{v}\|}{\|\vec{v}\|} \vec{E}_f(\vec{b};\vec{w}) + \left\| \vec{E}_g(\vec{a};\vec{v}) \right\| \vec{E}_f(\vec{b};\vec{w}) = M_g(\vec{a}) \vec{E}_f(\vec{b};\vec{w}) + \left\| \vec{E}_g(\vec{a};\vec{v}) \right\| \vec{E}_f(\vec{b};\vec{w}),$$

which tends to zero. Then, $(\vec{f} \circ \vec{g})(\vec{a})$ satisfies the condition of differentiable and with $\hat{T} = \left[\vec{f}'(\vec{g}(\vec{x}))\right] [\vec{g}'(\vec{a})]$. □

This expression is a great generalization of several phenomena we can find with multi-variable functions, some of them already discussed. But before looking and them, we need first to express the derivative of $\vec{f} \circ \vec{g}$ as a sum, which is an simple task. We know the derivative in terms of matrix product is expressed like

$$(\vec{f} \circ \vec{g})'(\vec{a}; \vec{v}) = \left[\vec{f} \circ \vec{g}'(\vec{a})\right] \vec{v} = \left[\vec{f}'(\vec{g}(\vec{a}))\right] [\vec{g}'(\vec{a})]\, v =$$
$$\begin{pmatrix} \dfrac{\partial f_1(\vec{a})}{\partial g_1} & \cdots & \dfrac{\partial f_1(\vec{a})}{\partial g_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(\vec{a})}{\partial g_1} & \cdots & \dfrac{\partial f_m(\vec{a})}{\partial g_n} \end{pmatrix} \begin{pmatrix} \dfrac{\partial g_1(\vec{a})}{\partial x_1} & \cdots & \dfrac{\partial g_1(\vec{a})}{\partial x_l} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_n(\vec{a})}{\partial x_1} & \cdots & \dfrac{\partial g_n(\vec{a})}{\partial x_l} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_l \end{pmatrix},$$

and what we know about the multiplication of matrices [ ], we can express a coefficient of the result as the sum of the coefficients of the previous matrices. Since the result is a matrix $m \times 1$, the coefficient will have the form $(\vec{f} \circ \vec{g})'_{i1}(\vec{a}; \vec{v})$, but we will simplify it by $(\vec{f} \circ \vec{g})'_{i}(\vec{a}; \vec{v})$ (and the same for the vector column of v). With that, we get the following expression.

$$(\vec{f} \circ \vec{g})'(\vec{a}; \vec{v}) = \sum_{i=1}^{m} (\vec{f} \circ \vec{g})'_i(\vec{a}; \vec{v})\vec{e}_i = \sum_{i=1}^{m} \left(\left[\vec{f}'(\vec{g}(\vec{a}))\right] [\vec{g}'(\vec{a})]\, v\right)_i \vec{e}_i = \sum_{i=1}^{m} \sum_{j=1}^{l} \sum_{k=1}^{n} \frac{\partial f_i}{\partial g_k} \frac{\partial g_k}{\partial x_j} v_j \vec{e}_i \tag{3.23}$$

A special case of this is if f becomes a scalar fields of the form $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. In this case, the result does not have several components and the expression is reduced to

$$(f \circ \vec{g})'(\vec{a}; \vec{v}) = \sum_{j=1}^{l} \sum_{k=1}^{n} \frac{\partial f}{\partial g_k} \frac{\partial g_k}{\partial x_j} v_j. \tag{3.24}$$

Besides, if we want to compute a partial derivative with, the vector $\vec{v}$ has only one non-zero component, which will be 1. Taking for example the i-th component (not to be confused with the i of the component of f), the expression results in

$$\frac{\partial (f \circ \vec{g})}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial f}{\partial g_k} \frac{\partial g_k}{\partial x_i}. \tag{3.25}$$

Finally, if $\vec{g}$ is a function of only one variable, the partial derivatives of g are simple derivative of just one variable, let us call t. This way, our final expression is the chain rule we obtain with scalar fields.

$$\frac{d(f \circ \vec{g})}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial g_k} \frac{dg_k}{dt}. \tag{3.26}$$

Example 3.2.1. Let be the following functions

$$\begin{array}{rl} \vec{g} : \mathbb{R}^2 & \longrightarrow \mathbb{R}^2 \\ (\rho, \theta) & \longmapsto (\rho \sin\theta, \rho\cos\theta) \end{array} \qquad \begin{array}{rl} f : \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \sqrt{1 - (x^2 + y^2)} \end{array}$$

The composition is

$$\begin{array}{rl} f \circ \vec{g} : \mathbb{R}^2 & \longrightarrow \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (\rho, \theta) & \longmapsto (\rho \sin\theta, \rho\cos\theta) \longmapsto \sqrt{1 - \rho^2} \end{array}$$

If we calculate its partial derivatives, we get

$$\frac{\partial f \circ \vec{g}}{\partial \rho} = \frac{-\rho}{\sqrt{1 - \rho^2}} \qquad = \frac{\partial f \circ \vec{g}}{\partial \theta} = 0.$$

To see this is coherent with the previous theorem, let us calculate the jacobians of $\vec{f}$ and $\vec{g}$ at an arbitrary point (they are differentiable everywhere).

$$[\vec{g}'(\vec{x})] = \begin{pmatrix} \partial g_1/\partial \rho & \partial g_1/\partial \theta \\ \partial g_2/\partial \rho & \partial g_2/\partial \theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\rho \sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$$

$$[f'(\vec{x})] = \begin{pmatrix}\partial f/\partial x & \partial f/\partial y\end{pmatrix} = \begin{pmatrix}-x/\sqrt{1-(x^2+y^2)} & -y/\sqrt{1-(x^2+y^2)}\end{pmatrix}$$

Then, the jacobian of $f \circ \vec{g}$ is

$$[f'(\vec{g}(\vec{x}))]\,[\vec{g}'(\vec{x})] = \begin{pmatrix}-\rho\cos\theta/\sqrt{1-\rho^2} & -\rho\sin\theta/\sqrt{1-\rho^2}\end{pmatrix}\begin{pmatrix}\cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta\end{pmatrix} =$$
$$\begin{pmatrix}-\rho/\sqrt{1-\rho^2} & 0\end{pmatrix} = \begin{pmatrix}\partial(f\circ\vec{g})/\partial\rho & \partial(f\circ\vec{g})/\partial\theta\end{pmatrix},$$

getting the same expression that did before. Another possible way to calculate the partial derivatives of $f \circ \vec{g}$ is doing the matrix product before calculating the actual derivatives. Knowing the image of $\vec{g}$ is the point to which f is applied, we can denote (g_1, g_2) as x, y. This way, the product results in

$$\begin{pmatrix}\partial f/\partial x & \partial f/\partial y\end{pmatrix}\begin{pmatrix}\partial x/\partial\rho & \partial x/\partial\theta \\ \partial y/\partial\rho & \partial y/\partial\theta\end{pmatrix},$$

and the partial derivatives are expressed as

$$\frac{\partial f}{\partial\rho} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\rho} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\rho} \qquad \frac{\partial f}{\partial\theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\theta}.$$

3.2.4 Relation between the jacobian of a function and its inverse

Let $\vec{f} : \mathbb{R}^n \longleftrightarrow \mathbb{R}^n$ be a function. In this case, we can talk about its inverse function $\vec{f}^{-1}(\vec{x})$. In particular, we can wonder about which matrix would be its jacobian. To see it, let us calculate the jacobian of $(\vec{f}^{-1} \circ \vec{f})(\vec{x}) = \vec{\mathrm{id}}(\vec{x})$.

$$\left[\vec{\mathrm{id}}'(\vec{x})\right] = \begin{pmatrix}\dfrac{\partial x_1}{\partial x_1} & \cdots & \dfrac{\partial x_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial x_1} & \cdots & \dfrac{\partial x_n}{\partial x_n}\end{pmatrix} = I_{n\times n}$$

With that, we know the jacobians of $\vec{f}$ and $\vec{f}^{-1}$ are invertible and

$$\left[\vec{f}^{-1}(\vec{x})\right]\left[\vec{f}(\vec{x})\right] = I \Rightarrow \left[\vec{f}^{-1\prime}(\vec{x})\right] = \left[\vec{f}'(\vec{x})\right]^{-1}$$

Proposition 3.2.9. *Let $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be two functions. Let a be an interior point of A and $f(a) = b$ an interior point of B. If $\vec{g}$ is differentiable at b, $\vec{g} \circ \vec{f}$ differentiable at a, and if the jacobian of g is an injective linear function, then $[g'(b)]^{-1}$ and $[\vec{f}'(\vec{a})]$ exist, and*

$$[\vec{f}'(a)] = [g'(\vec{f}(\vec{a}))]^{-1} \circ (\vec{g} \circ \vec{f})'(\vec{a}). \tag{3.27}$$

3.2.5 Mean value theorems

Theorem 3.2.10. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a, b two interior points of D. If $(1-t)a + tb$ are interior point of D for all $(0,1)$, and if f is differentiable at all these points, then there exists a point $c = (1-\tau)a + \tau b$ for some $\tau \in (0,1)$ such that*

$$f(\vec{b}) - f(\vec{a}) = \langle\vec{\nabla} f(\vec{c}), \vec{b} - \vec{a}\rangle_I. \tag{3.28}$$

Corollary 3.2.11. *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function with Ω an open set such that, whenever $a, b \in \Omega$, then $(1-t)a + tb \in \Omega$ for all $t \in (0,1)$ (Ω is convex). If $[\vec{f}'(\vec{x})]$ exists and is the zero function for all $x \in \Omega$, then $\vec{f}(x) = \vec{c}$, where $\vec{c}$ is a constant vector.*

3.3 Higher order derivatives of scalar functions

Sometimes there are functions that we can derive several times. In theses cases we can study several phenomena related to its partial derivatives and the order with which we do it.

Definition 3.3.1. Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar function and $\Lambda \subseteq \Omega$, with Ω an open set. We say f *is a function of class* $C^k(\Lambda)$ with $k \in \mathbb{N}$ if its partial derivatives until k-th degree exist and are continuous in Λ.

Definition 3.3.2. Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar function and $\Lambda \subseteq \Omega$, with Ω an open set. We say f *is a function of class* $C^\infty(\Lambda)$ if it is of class $C^k(\Lambda)$ for all $k \in \mathbb{N}$.

Proposition 3.3.1. *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function with $\Omega \subseteq \mathbb{R}^n$ an open set. If $\partial^2 f(\vec{x})/\partial x_i \partial x_j$ and $\partial f(\vec{x}/\partial x_j \partial x_i)$ exist and are continuous in Ω, then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \tag{3.29}$$

Proof. We will prove only for $n = 2$. Let us take a point $p = (a, b) \in \Omega$. We know there exists a $\delta > 0$ such that $B(p, \delta) \subset \Omega$. Let us take now a number h such that $0 < h < \delta/\sqrt{w}$ and write the "second difference" (we use this expression because we refer to a difference of differences) divided by h^2.

$$\Delta(h) = \frac{1}{h^2}\left[[f(a+u, b+u) - f(a+u, b)] - [f(a, b+u) - f(a, b)]\right]$$

If we define $p(x) = f(x, b+u) - f(x, b)$, we can express $\Delta(h)$ as

$$\Delta(h) = \frac{1}{h^2}\left[g(a+h) - g(a)\right]$$

By the intermediate value theorem, there exists a $\theta_1 \in (0, 1)$ such that

$$\Delta(h) = \frac{1}{h^2}\left[g(a+h) - g(a)\right] = \frac{1}{h^2} g'(a+\theta h)h = \frac{1}{h}\left[\frac{\partial f}{\partial x}(a+\theta_1 h, b+h) - \frac{\partial f}{\partial x}(a+\theta_1 h, b)\right],$$

but we can apply the intermediate value theorem again to this difference, so there exists a $\theta_2 \in (0, 1)$ such that

$$\Delta(h) = \frac{1}{h}\left[\frac{\partial f}{\partial x}(a+\theta_1 h, b+h) - \frac{\partial f}{\partial x}(a+\theta_1 h, b)\right] = \frac{1}{h}\frac{\partial f}{\partial y}\left[\frac{\partial f}{\partial x}(a+\theta_1 h, b+\theta_2 h)h\right] =$$
$$\frac{\partial^2 f}{\partial y \partial x}(a+\theta_1 h, b+\theta_2 h).$$

If we make $h \to 0$, we will have $\theta_1, \theta_2 \to 0$, and since both partial derivatives exist and are continues, we have

$$\lim_{h\to 0} \Delta(h) = \frac{\partial^2 f}{\partial y \partial x}(a, b).$$

Now we can define $q(y) = f(a+h, y) - f(a, y)$. Doing the same process, we get

$$\lim_{h\to 0} \Delta(h) = \lim_{h\to 0} q(b+h) - q(b) = \frac{\partial^2 f}{\partial y \partial x}(a, b).$$

Since both limits tend to the same expression, we conclude they coincide. □

Theorem 3.3.2. *Let $\vec{f} : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a a point in D. If $\partial^2 \vec{f}(\vec{a})/\partial x_i \partial x_j$ and $\partial \vec{f}(\vec{a}/\partial x_j \partial x_i)$ exist and one of them is continuous in a, then the other is also continuous in a and*

$$\frac{\partial f}{\partial x_i \partial x_j}(\vec{a}) = \frac{\partial f}{\partial x_j \partial x_i}(\vec{a}) \tag{3.30}$$

Proof. Apostol pag.340 $\square$

Theorem 3.3.3. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a an interior point of D. If for some δ, $\partial_{x_i} f(\vec{x}), \partial_{x_j} f(\vec{x}), \partial_{x_i}\partial_{x_j} f(\vec{x})$ exist in $B(a,\delta)$ and if $\partial_{x_i}\partial_{x_j} f(\vec{x})$ is continuous in a, then $\partial_{x_j}\partial_{x_i} f(\vec{x})$ exists and*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\vec{a}). \tag{3.31}$$

Example 3.3.1. Let be the following function

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, (x,y) \neq (0,0) \\ 0, (x,y) = (0,0) \end{cases}$$

Now, let us calculate its mixed partial derivative at the point $(0,0)$ and see if they coincide, beginning with $\partial^2/\partial y \partial x$. For that, we need first to calculate the partial derivative with respect to x in general, when $\vec{x} \neq \vec{0}$ and when $\vec{x} = \vec{x}$.

$$\vec{x} \neq \vec{0} : \ \frac{\partial f}{\partial x} = \frac{y(x^4 + 4y^2x^2 - y^4)}{(x^2+y^2)^2}$$

$$\vec{x} = \vec{0} : \ \frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f((h,0)) - f((0,0))}{h} = \lim_{h \to 0} \frac{1}{h}\left(h \cdot 0 \frac{h^2+0^2}{h^2+0^2}\right) = \lim_{h\to 0} 0 = 0$$

Now, we can calculate its derivative with respect to y.

$$\frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = \lim_{k \to 0} \frac{\partial_x f((0,k)) - \partial_x f((0,0))}{k} = \lim_{k \to 0} \frac{1}{k}\frac{k(0^4 + 4k^2 0^2 - k^4)}{(0^2+k^2)^2} = \lim_{k \to 0} -1 = -1$$

Now let us calculate $\partial^2/\partial x \partial y$, beginning by calculating the partial derivative with respect to y.

$$\vec{x} \neq \vec{0} : \ \frac{\partial f}{\partial y} = -\frac{x(y^4 + 4x^2y^2 - x^4)}{(y^2+x^2)^2}$$

$$\vec{x} = \vec{0} : \ \frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{f((0,h)) - f((0,0))}{h} = \lim_{h \to 0} \frac{1}{h} 0 \cdot h \frac{0^2 - h^2}{0^2 + h^2} = \lim_{h \to 0} 0 = 0$$

And calculating now its partial derivative with respect to x at $(0,0)$ we get

$$\frac{\partial f}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{k \to 0} \frac{\partial_y f((k,0)) - \partial_y f((0,0))}{k} = \lim_{k \to 0} \frac{1}{k}(-1)\frac{k(0^4 + 4k^2 0^2 - k^4)}{(0^2+k^2)^2} = \lim_{k \to 0} 1 = 1$$

3.3.1 Notation for partial derivatives and Hessian matrix

Let $f : \Omega \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}$ be a function with Ω an open set. From now we will deal with successive partial derivatives like

$$\frac{\partial^4}{\partial x_3^4}\left(\frac{\partial}{\partial x_2}\left(\frac{\partial^2 f}{\partial x_1^2}\right)\right),$$

so we need to simplify the notation in order to avoid long calculations. In particular, if f is of class $C^n(\Omega)$ (with n arbitrary,)e will use what is called the *multi-index notation* [ ], which consists on using elements of $\mathbb{Z}^m$ with non-negative coordinates, that is,

$$\alpha = (\alpha_1, \ldots, \alpha_m), \quad 0 \leq \alpha_i \in \mathbb{Z}, \quad 1 \leq i \leq m.$$

The number α_i represents the number of times f have been derived with respect to x_i. Since the function is of $C^n(\Omega)$, we can change the order of partial derivatives and group all the times it is derived with respect one variable to represent it with that number. For example,

$$\partial^{(1,2)} f = \frac{\partial}{\partial x_1}\left(\frac{\partial^2 f}{\partial x_2^2}\right), \qquad \partial^{(0,2,1)} f = \frac{\partial^2}{\partial x_2}\left(\frac{\partial f}{\partial x_3}\right).$$

Besides, for some displacement $\vec{h}$ in the using the same meaning of α, we use the following simplifications.

$$|\alpha| = \alpha_1 + \cdots + \alpha_m, \qquad \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m}, \qquad h^\alpha = h_1^{\alpha_1} \cdots h_m^{\alpha_m}, \qquad \alpha! = \alpha_1! \cdots \alpha_m!$$

Now these expressions seem a bit arbitrary but in following parts we will see they are frequently used.

We have seen previously the first derivative of a scalar field can be represented as a matrix product of the form $h^t \nabla f$. For the second derivative there is also a matrix product expression, but is a bit different. Before finding it, let us compute the second derivative by apply the sum expression. Taking a function f, a point a, and a vector $\vec{h}$, and evaluating the second derivative at the same point and with the same direction, we get

$$f'(\vec{a};\vec{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i \Rightarrow$$

$$f''(\vec{a};\vec{h}) = \sum_{j=1}^{m} \frac{\partial}{\partial x_j} \left[f'(\vec{a};\vec{h}) \right] h_j = \sum_{j=1}^{m} \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} \frac{\partial f}{\partial x_i} h_i \right) h_j = \sum_{j=1}^{m} \sum_{i=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_i} h_i h_j$$

If we remember the properties of multiplication by matrices, this formula is equivalent to the next product.

$$\sum_{j=1}^{m} \sum_{i=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_i} h_i h_j = \begin{pmatrix} h_i & \dots & h_m \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x_i^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} = h^T H h$$

This matrix in the center is called the *Hessian matrix*. Let us define it properly.

Definition 3.3.3. Let $f : \Omega \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}$ be a function of class $C^2(\Omega)$, with Ω an open set and a a point of Ω. Then, we define the *Hessian matrix of f at a point a* as the following matrix.

$$\begin{pmatrix} \frac{\partial^2 f(\vec{a})}{\partial x_i^2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{a})}{\partial x_m \partial x_1} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_m^2} \end{pmatrix} \tag{3.32}$$

Note that the case we are studying, where f is of class $C^n(\Omega)$, its partial derivatives are continuous and therefore we can change the order of the mixed partial derivatives. Hence, in the sum we can interchange the partial derivatives and the Hessian matrix is symmetric. In fact, the Hessian matrix can be obtained as a product of matrices, for example, by this one.

$$H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_i^2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_m^2} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_m} \end{pmatrix} = \nabla^t \nabla f$$

This form will serve us to express later the general derivative but again, as with the second derivative, we need first calculate it in as a sum. Applying the same procedure, always at the same point and with the same direction, the n-th derivative can be expressed as

$$f^{(n)}(\vec{a};\vec{h}) = \sum_{i_1,\dots,i_n=1}^{m} \frac{\partial^n f}{\partial x_{i_1} \dots \partial x_{i_n}} h_{i_1} \cdots h_{i_n}. \tag{3.33}$$

If we separate de differential operator from the function as we do with gradient, we can use the properties of sums and express it as follows.

$$\sum_{i_1,\dots,i_n=1}^{m} \frac{\partial^n}{\partial x_{i_1} \dots \partial x_{i_n}} h_{i_1} \cdots h_{i_n} f = \left(\sum_{i_1=1}^{m} \frac{\partial}{\partial x_{i_1}} h_{i_1} \right) \cdots \left(\sum_{i_n}^{m} \frac{\partial}{\partial x_{i_n}} h_{i_n} \right) f =$$

$$(h^t \nabla) \cdots (h^t \nabla) f \Rightarrow (h^t \nabla)^n f$$

Note although the indices were different the sums where the same, se could regroup them as a product of an element by itself which result in an exponentiation. Taking this matrix product and using the dot product notation, we get

$$f^{(n)}(\vec{a};\vec{h}) = \langle \vec{h}, \vec{\nabla} \rangle_I^n f(\vec{a}). \tag{3.34}$$

However, we can still rephrase the n-th derivative in another way. If we express the $h^t\nabla$ again as a sum and use the multinomial theorem [ ], we obtain

$$(h^t\nabla)^n f = \left(\sum_{i=1}^{m} h_i \frac{\partial}{\partial x_i} \right)^n f = \sum_{\alpha_1+\cdots+\alpha_m=n} \frac{n!}{\alpha_1!\cdots\alpha_m!} \left(h_1 \frac{\partial}{\partial x_1} \right)^{\alpha_1} \cdots \left(h_m \frac{\partial}{\partial x_m} \right)^{\alpha_m} f =$$
$$\sum_{\alpha_1+\cdots+\alpha_m=n} \frac{n!}{\alpha_1!\cdots\alpha_m!} h_1^{\alpha_1} \cdots h_m^{\alpha_m} \cdot \left(\frac{\partial}{\partial x_1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_m} \right)^{\alpha_m} f,$$

which has the expressions we mentioned before to be reduced. Following the multi-index notation, the final expression of the n-th derivative results in

$$f^{(n)}(\vec{a};\vec{h}) = \sum_{|\alpha|=n} \frac{n!}{\alpha!} \partial^\alpha f(\vec{a}) h^\alpha \tag{3.35}$$

3.4 Higher order derivatives for vector fields

Theorem 3.4.1. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function and a an interior point of D. If for some δ, $\partial_{x_i}\vec{f}(\vec{x}), \partial_{x_j}\vec{f}(\vec{x}), \partial_{x_i}\partial_{x_j}\vec{f}(\vec{x})$ exist in $B(a,\delta)$ and if $\partial_{x_i}\partial_{x_j}\vec{f}(\vec{x})$ is continuous in a, then $\partial_{x_j}\partial_{x_i}\vec{f}(\vec{x})$ exists and*

$$\frac{\partial^2 \vec{f}}{\partial x_i \partial x_j}(\vec{a}) = \frac{\partial^2 \vec{f}}{\partial x_j \partial x_i}(\vec{a}). \tag{3.36}$$

Bibliography

[1] H. Sagan. *Advanced Calculus: Of Real-valued Functions of a Real Variable and Vector-valued Functions of a Vector Variable.* Houghton Mifflin Company, 1974.

Chapter 4

Maxima and minima

4.1 Taylor expansion for scalar fields

Let us study now if there is a formula analogous to the Taylor polynomial we studied for one variable functions. For that, let us take a scalar field $f : \Omega \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}$ of class C^n, with Ω an open set, and with a a point in Ω. With that, we can construct a one variable function as follows

$$g(t) = f(\vec{a} + t\vec{h}), \tag{4.1}$$

with $\vec{h}$ fixed. Since f is of class C^n, g will be of class C^n too. With that, the Taylor formula around $t = 0$ would be expressed as

$$g(t) = P^{(0)}_{m-1}(t) + R^{(0)}_{n-1}(t) = \sum_{k=0}^{n-1} \frac{g^{(k)}(0)}{k!} t^k + \frac{g^{(n)}(\xi)}{n!} t^n,\ \xi \in (0, t), \tag{4.2}$$

taking the Lagrange rest. This equality shows that the value $g(t)$ can be calculated always by the expression of the right side, but not with complete determination, since ξ has no specific value. Another construction we could generate from this is the Taylor series (if f is of class C^∞), expressed as

$$\sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!} t^k. \tag{4.3}$$

This series is a power series, so it converges with every $x \in (-R, R)$, with R being the radius of convergence. And if we are inside this interval, we can construct an approximation of $g(t)$, the Taylor approximation, specifying the order. Therefore, if we know how to go from the Taylor polynomial of one variable to several variables, we can proceed as we have just done and obtain all Taylor expressions as an immediate consequence.
To obtain the first expression, we need to obtain $g(t)$ and its derivatives. By construction $g(t) = f(\vec{a} + t\vec{h})$, and to see the expression of $g'(t)$, we can use the previous formulas for scalar fields. Denoting $a_i + th_i$ by x_i, we get

$$g'(t) = \frac{df(\vec{a} + t\vec{h})}{dt} = \left\langle \vec{\nabla} f(\vec{a} + t\vec{h}), \frac{d(\vec{a} + t\vec{h})}{dt} \right\rangle_I = \langle \vec{\nabla} f(\vec{a} + t\vec{h}), \vec{h} \rangle_I = \sum_{i=1}^{m} h_i \frac{\partial f}{\partial x_i}$$

Now, to obtain the second derivative we only need to apply the same formula but this time to each term of the sum.

$$g''(t) = \sum_{i=1}^{m} h_i \frac{d}{dt}\left(\frac{\partial f}{\partial x_i}\right) = \sum_{i=1}^{m} h_i \sum_{j=1}^{m} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\frac{dx_j}{dt} = \sum_{i=1}^{m}\sum_{j=1}^{m} h_i h_j \frac{\partial^2 f}{\partial x_j \partial x_i} \tag{4.4}$$

As we see, we obtain the same expressions as if we calculated $f'(\vec{a}; \vec{h})$. This is normal since we proved their equivalence 3.1.2. Therefore, for the n-th derivative we will get

$$g^{(n)}(t) = \sum_{i_1,\ldots,\alpha_n=1}^{m} \frac{\partial^n f}{\partial x_{i_1} \ldots \partial x_{i_n}} h_{i_1} \cdots h_{i_n}. \tag{4.5}$$

Finally, since the Taylor formula is correct for every set where f (or g) is defined, we can select the interval $[0, 1]$ for t and apply the formula for $t = 1$. If we wanted to apply the formula of f in another bigger displacement, we only need to vary $\vec{h}$ maintaining the magnitude of t. This way, all the exponents of t become 1 and disappear in the expression. Said that, we can finally express the Taylor. Since we have seen different versions of the n-th derivative, we will express the Taylor formula in several ways. In the sum notation, the expression is

$$f(\vec{a} + \vec{h}) = \sum_{k=0}^{n-1} \frac{1}{k!} \sum_{i_1,\ldots,\alpha_k=1}^{m} \frac{\partial^k f(\vec{a})}{\partial x_{i_1} \ldots \partial x_{i_n}} h_{i_1} \cdots h_{i_k} + \frac{1}{n!} \sum_{i_1,\ldots,\alpha_n=1}^{m} \frac{\partial^n f(\vec{\xi})}{\partial x_{i_1} \ldots \partial x_{i_n}} h_{i_1} \cdots h_{i_n}. \tag{4.6}$$

If we use the two reduced notation, we have

$$f(\vec{a}+\vec{h}) = \sum_{k=0}^{n-1} \frac{1}{k!}(h^t\nabla)^k f(\vec{a}) + \frac{1}{n!}(h^t\nabla)^n f(\vec{\xi}), \tag{4.7}$$

$$f(\vec{a}+\vec{h}) = \sum_{k=0}^{n-1}\sum_{|\alpha|=k} \frac{\partial^\alpha f(\vec{a})}{\alpha!}h^\alpha + \sum_{|\alpha|=n} \frac{\partial^\alpha f(\vec{a})}{\alpha!}h^\alpha \tag{4.8}$$

4.1.1 Taylor formula of second order

There is a special interest in studying the Taylor formula of the second order. In particular, it will allow us to discuss the topic of maxima and minima that we did in one variable functions but now for multi-variable ones. But before that, we first need to observe some properties about this case of the second order that will serve us in that topic.

Proposition 4.1.1. *Let a function of class $C^2(D)$ and a an interior point of D. Then,*

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \frac{h^T\nabla f(\vec{a})}{1!} + \frac{h^T H(\vec{a}+\theta\vec{h})h}{2!},\ \theta \in (0,1) \tag{4.9}$$

is equivalent to

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \frac{h^T\nabla f(\vec{a})}{1!} + \frac{h^T H(\vec{a})h}{2!} + \|\vec{h}\|^2 E_2(\vec{a};\vec{h}). \tag{4.10}$$

Proof. If we compare both expressions, the proposition is equivalent to see that

$$\|\vec{h}\|^2 E_2(\vec{a};\vec{h}) = \frac{1}{2!}h^T\left[H(\vec{a}+\theta\vec{h}) - H(\vec{a})\right]h$$

To simplify the notation, we will denote $\left[H(\vec{a}+\theta\vec{h}) - H(\vec{a})\right]/2!$ by $\hat{H}$, and if we use the sum expressions, we get

$$\|\vec{h}\|^2 E_2(\vec{a};\vec{h}) = \sum_{i,j=1}^m h_i\hat{H}_{ij}h_j.$$

We will now bound this sum using the Cauchy-Schartz inequality, but instead of rewriting the equality all the time, we will just write the simplify the steps.

$$\left(\sum_{i,j=1}^m h_i\hat{H}_{ij}h_j\right)^2 = \left(\sum_{i=1}^m h_i\sum_{j=1}^m \hat{H}_{ij}h_j\right)^2 \le \sum_{i=1}^m h_i^2 \cdot \sum_{k=1}^m\left(\sum_{k=j}^m \hat{H}_{kj}h_j\right)^2 \le$$
$$\sum_{i=1}^m h_i^2\cdot\sum_{k=1}^m\left(\sum_{j=1}^m \hat{H}_{kj}^2\sum_{l=1}^m h_l^2\right) = \sum_{i=1}^m h_i^2\cdot\sum_{i=l}^m h_l^2\cdot\sum_{k=1}^m\sum_{j=1}^m \hat{H}_{kj}^2 = \left(\sum_{i=1}^m h_i^2\right)^2\cdot\sum_{k,j=1}^m \hat{H}_{kj}^2 \le$$
$$\left(\sum_{i=1}^m h_i^2\right)^2\cdot\left(\sum_{i=1}^m |\hat{H}_{kj}|\right)^2 \Rightarrow \left|\sum_{i,j=1}^m h_i\hat{H}_{ij}h_j\right| \le \sum_{i=1}^m h_i^2\cdot\sum_{i=1}^m|\hat{H}_{kj}|$$

With that, the original expression results in

$$\|\vec{h}\|^2|E_2(\vec{a};\vec{h})| \le \sum_{i=1}^m h_i^2\sum_{i=1}^m|\hat{H}_{kj}| \Rightarrow |E_2(\vec{a};\vec{h})| \le \sum_{i,j=1}^m h_i\hat{H}_{ij}h_j = \frac{1}{2!}\sum_{i,j}^m\left[H_{ij}(\vec{a}+\theta\vec{h}) - H_{ij}(\vec{a})\right]$$

Since this relation is true always, it will be too if we make $\vec{h}\to\vec{0}$. And since the Hessian matrix contains second derivatives and the function is of class C^2, all its components are continuous and $H_{ij}(\vec{a}+\theta\vec{h}) - H_{ij}(\vec{a})$ will tend to 0. Therefore, $|E_2(\vec{a};\vec{h})|$ will tend to zero too and consequently, $E_2(\vec{a};\vec{h})$. □

4.2 Maxima and minima of scalar fields

4.2.1 Introduction

Definition 4.2.1. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a a point of D. We say f *has a local maxima in* a if there exists a ball $B(a,r) \subset D$ such that

$$f(\vec{x}) \geq f(\vec{a}), \forall x \in B(a,r) \tag{4.11}$$

Definition 4.2.2. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a a point of D. We say f *has a local minima in* a if there exists a ball $B(a,r) \subset D$ such that

$$f(\vec{x}) \leq f(\vec{a}), \forall x \in B(a,r) \tag{4.12}$$

Definition 4.2.3. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a a point of D. We say f *has a global maxima in* a if

$$f(\vec{x}) \leq f(\vec{a}), \forall x \in D \tag{4.13}$$

Definition 4.2.4. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a a point of D. We say f *has a global minima in* a if

$$f(\vec{x}) \geq f(\vec{a}), \forall x \in D \tag{4.14}$$

We refer to a maxima or minima in general as an extreme. We can also denote a global extreme as a relative extreme and a global extreme as an strict extreme.

Proposition 4.2.1. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and a an interior point of D. If f is differentiable in a and has a local extreme in a, then $\vec{\nabla} f = \vec{0}$ (which is equivalent to say that every component, that is, every partial derivative, is 0).*

Proof. We will prove the proposition for a maxima. Let us suppose $f(\vec{x})$ has a maxima in a. Now, let be $g_i(x_i) = f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ a function that has every component fixed except x_i. Then, g_i is defined in an interval centered in $a_i \in \mathbb{R}$.
Since $f(\vec{x}) \leq f(\vec{a})$ when we are in a ball $B(a,r)$, we can take an interval contained in the ball that will satisfy that $g_i(x_i) \leq g(a_i), \forall x_i \in I_i$. This is the case of one variable function, and we know we can say that []

$$0 = \frac{dg_i}{dx_i}(a_i) = \frac{\partial f}{\partial x_i}(\vec{a})$$

Repeating this for every component, we will have that every partial derivative is zero and therefore, $\vec{\nabla} f = \vec{0}$. ☐

Definition 4.2.5. Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field, with Ω an open set, and a a point in Ω. We say *a is a stationary or extreme point* if $\vec{\nabla}(\vec{a}) = \vec{0}$.

Definition 4.2.6. Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field, with Ω an open set, and a a point in Ω. We say *a is a saddle point* if

$$\forall B(a,r) \exists \vec{x} \mid f(\vec{x}) > f(\vec{a}) \wedge \exists \vec{y} \mid f(\vec{y}) < f(\vec{a}) \tag{4.15}$$

Example 4.2.1. The function $f = x^2 - y^2$ has a saddle point at (0,0).

4.2.2 Properties of Hessian matrix

We have seen previously that the Hessian of a scalar field of class C^2 can be expressed as a matrix product of the form $h^T H h$. This means we could represent it as a bi-linear form, but more particularly, as a quadratic form (since the matrix is symmetric). This association will allow us to use some results from linear algebra to discuss the topic of maxima and minima, as we will see now.

Proposition 4.2.2. *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field of class C^2, with Ω an open set. Let $Q(h)$ quadratic form associated to the Hessian matrix of f in an arbitrary point of Ω. Then, there are two real numbers m, M such that*

$$m\|h\|^2 \leq Q(h) \leq M\|h\|^2 \tag{4.16}$$

Proof. Since H is symmetric, is orthonormally diagonalizable [ ], that is, there exist an orthogonal matrix O such that $OHO^T = D_H$, with D_H a diagonal matrix that contains the eigenvalues of H. Let us take now a vector $x = O^T h$ (that satisfies also that $Ox = h$). With that, we get

$$Q(h) = h^T H h = (Ox)^T H(Ox) = x^T O^T H O x = x^T D_H x \coloneqq \hat{Q}(x)$$

which, since it has the eigenvalues of H, can be expressed as

$$\hat{Q}(x) = \sum_{i=1}^{n} \lambda_i x_i^2$$

In the set of the eigenvalues there is always a maximum and minimum element. With that, we can bound $\hat{Q}(h)$ by

$$\sum_{i=1}^{n} \lambda_{\min} x_i^2 \leq \hat{Q}(x) \leq \sum_{i=1}^{n} \lambda_{\max} x_i^2 \Leftrightarrow \lambda_{\min} \|x\|^2 \leq \hat{Q}(x) \leq \lambda_{\max} \|x\|^2.$$

But $\|x\|^2 = \langle x, x \rangle_I$, which is another bi-linear form. Therefore, the orthogonal matrix won't affect the product and we will have that $\|x\|^2 = x^T I x = h^T O I O^T h = h^T h = \|h\|^2$. Therefore, the final bound of $\hat{Q}(x) = Q(h)$ will be

$$\lambda_{\min} \|h\|^2 \leq Q(h) \leq \lambda_{\max} \|h\|^2.$$

□

This leads to an immediate consequence. If every eigenvalue is positive, then $h^T H h \geq 0$ always that $h \neq 0$ (is *positively defined*), and if every eigenvalue is negative, then $h^T H h \leq 0$ always that $h \neq 0$ (is *negatively defined*). In case we have some eigenvalues positives and some negatives, depending on h, $Q(h)$ will be positive or negative.

Proposition 4.2.3 (Criterion of sufficiency of stationary points)**.** *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field of class C^2, with Ω an open set, and a an stationary point of f in Ω. If the Hessian matrix of f is positively (negatively) defined in a, then a is a minima (maxima) of f.*

Proof. We will prove for a minima. Using the equivalent expression of the Taylor formula, we have

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = \frac{h^T \nabla f(\vec{a})}{1!} + \frac{h^T H(\vec{a}) h}{2!} + \|\vec{h}\|^2 E_2(\vec{a}; \vec{h}),$$

with

$$\lim_{\vec{h} \to \vec{0}} E_2(\vec{a}; \vec{h}) = 0.$$

Since a is a stationary point, by the proposition 4.2.1, $h^T \nabla f(\vec{a}) h$ will be 0, and since $Q(h)$ is positively defined, its value can be bounded by a number $m > 0$. Then, our expression results in

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = \frac{1}{2!} h^t H(\vec{a}) h + \|\vec{h}\|^2 E_2(\vec{a}; \vec{h}) \geq \frac{m}{2} \|\vec{h}\|^2 + \|\vec{h}\|^2 E_2(\vec{a}; \vec{h}).$$

Now, let us take the ball $B(\vec{a}; r)$ such that

$$E_2(\vec{a}; \vec{h}) > -\frac{m}{4}, \qquad \|\vec{h}\| < r.$$

Then, for every point $a + h$ inside the ball we will have

$$f(\vec{a} + \vec{h}) - f(\vec{a}) \geq \frac{m}{2} \|\vec{h}\|^2 + \|\vec{h}\|^2 E_2(\vec{a}; \vec{h}) \geq \frac{m}{2} \|\vec{h}\|^2 - \frac{m}{4} \|\vec{h}\|^2 = \frac{m}{4} \|\vec{h}\|^2 > 0,$$

obtaining that $f(\vec{a} + \vec{h}) > f(\vec{a})$, which is the definition of a minima. □

Note that, if the quadratic form is not positively defined or negatively defined, the stationary point a will be a saddle point.

Theorem 4.2.4. *Let $f : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a scalar field of class C^2, with Ω an open set, and a an stationary point of f in Ω. Then,*

- *If* $\det\{H(\vec{a})\} > 0$ *and* $\partial^2 f(\vec{a})/\partial x^2 > 0$*, then f has a local minima in a.*
- *If* $\det\{H(\vec{a})\} > 0$ *and* $\partial^2 f(\vec{a})/\partial x^2 < 0$*, then f has a local maxima in a.*
- *If* $\det\{H(\vec{a})\} < 0$*, then f has a saddle point in a.*
- *If* $\det\{H(\vec{a})\} = 0$*, we can't determine the point.*

Proof. Let us diagonalize the Hessian matrix of f by calculating the characteristic polynomial.

$$\det\{(H(\vec{a}) - \lambda I)\} = 0 \Leftrightarrow \lambda^2 - \mathrm{Tr}\{H(\vec{a})\}\lambda - \det\{H(\vec{a})\} = 0 \tag{4.17}$$

The determinant and trace of H are independent of the selected basis, so we will calculate them when the matrix is diagonalized. If we denote by λ_1, λ_2 the eigenvalues of H, the determinant would be $\lambda_1\lambda_2$ and the trace $\lambda_1 + \lambda_2$. Therefore, we have that

$$\lambda_1\lambda_2 = \det\{H(\vec{a})\} = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$
$$\lambda_1 + \lambda_2 = \mathrm{Tr}\{H(\vec{a})\} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Suppose now that the determinant is positive. Then, λ_1 and λ_2 will have the same sign. Besides the left side also has to be positive, and therefore $\partial_x^2 f \partial_y^2 f$ too. Then, we can identify two cases: where $\partial_x^2 f$ and $\partial_y^2 f$ are both positive and where are both negative. If the second partial derivative with respect to x is positive (and therefore with respect to y), the trace will be positive. Hence, $\lambda_1 + \lambda_2$ is positive and with the same sign, which shows us that $\lambda_1, \lambda_2 > 0$ and consequently that $H(\vec{a})$ is positively defined. By the proposition 4.2.3, this is equivalent to say that a is a minima. If $\partial_x^2 f$ is negative (and therefore $\partial_y^2 f$ too), the trace is negative, and following the same reasoning, we get that λ_1 and λ_2 are negative, which proves that $H(\vec{a})$ is negatively defined and consequently that a is a maxima.
Let us discuss now the case where $\det\{H(\vec{a})\}$ is negative. By the first equation, we get that $\lambda_1\lambda_2 < 0$, and therefore, they have different signs. This means that the quadratic form is not positively defined or negatively defined, so it will be a saddle point. If the determinant is 0, by the first equation, at least one of the eigenvalues will be 0. Therefore, the expression of the diagonalized quadratic form ($\sum \lambda_i x_i^2$) will have some direction where it will be 0. Taking $\lambda_1 = 0$, for example, one direction possible would be one of the form $\vec{x} = (0, x_2)$ (it would be translated to the direction $\vec{h}$ by multiplying it by the orthogonal matrix). Then, in these directions $h^T H(\vec{a}) h$ will be 0 and the Taylor formula will be presented as

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = \tau(\vec{a}),$$

where $\tau(\vec{a})$ is a continuation of the Taylor formula. It is not necessarily E_2, and we can't figure it out because it depends on if we could compute its higher-order partial derivatives. Therefore, there is not enough information to determine the stationary point. □

In fact, this is just a particular case from a result we already know in linear algebra []. As we know, if we take H_k as the $k \times k$ matrix from the upper left corner, computing their determinants can tell us the nature of the matrix. If every determinant is positive, the H is positively defined, and if the determinant is positive for k even and negative for k odd, then H is negatively defined. If it has no a sign defined, it is a saddle point, and if $\det\{H(\vec{a})\} = 0$, for the same reason as before, we can't determine the stationary point.

Chapter 5

Implicit function

5.1 Inverse function Theorem

Theorem 5.1.1 (Inverse function Theorem). *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a function with Ω an open set. If $\vec{f} \in C^1(\Omega)$, and if, for some point $a \in \Omega$, $\operatorname{rank}[\vec{f}'(\vec{a})] = n$, the there exist open sets $U, V \subseteq \mathbb{R}^n$ such that $a \in U, \vec{f}(\vec{a}) \in V$, and*

1. *$\vec{f} : U \longrightarrow V$ is injective,*
2. *$\vec{f}^{-1} : V \longrightarrow U$ exists,*
3. *$\vec{f}^{-1} \in C^1(V)$,*
4. *$\det\Big\{[\vec{f}(\vec{x})]\Big\} \neq 0$ for all $x \in U$, and*
5. *$\det\Big\{[\vec{f}^{-1}(\vec{x})]\Big\} \neq 0$ for all $y \in V$.*

5.2 Conditions for existence and uniqueness

Let us suppose now a functional equation expressed in the implicit form, that is,

$$F(x, y) = 0, \qquad x, y \in \mathbb{R}. \tag{5.1}$$

We hope that, by certain conditions, the set of solutions $\big\{(x, y) \in \mathbb{R}^2 \,\big|\, F(x, y) = 0\big\}$ to be a curve from which we could isolate y and get an explicit expression of the form

$$y = f(x) \tag{5.2}$$

or of the form

$$x = g(y). \tag{5.3}$$

If we want this to happen, by the definition of function, we need that every x has exactly one y associated. It could happen in all the domain or just one region, so our task is to determine these regions. Another detail we have to consider is that there could be the case where $F(x, y) = 0$ does not determine any expression of the form $y = f(x)$. However, it does not mean that the expression does not have solution (see example 5.2.1).

Example 5.2.1. Let us consider the following implicit relations

$$F(x, y) = x^2 + y^2 + 1 = 0 \qquad G(x, y) = x^2 + y^2 = 0.$$

The first relation does not determine any explicit relation because there is no solution. However, although the second one does not determine any relation neither, it has a solution, the point $(0, 0)$.

Theorem 5.2.1 (Theorem of the implicit function). *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field of class C^1, with Ω an open set and a a point in Ω such that $f(\vec{a}) = 0$ and $\partial f / \partial x_n(\vec{a}) \neq 0$. Then, there exists a ball $B(\vec{a}; r) \subset \Omega$, an open set $\Gamma \subseteq \mathbb{R}^{n-1}$ that contains a, and a function $g : \Gamma \longrightarrow \mathbb{R}$ of class $C^1(\Gamma)$ such that*

$$x \in B(\vec{a}; r) \qquad \text{and} \qquad f(\vec{x}) = 0 \tag{5.4}$$

if and only if

$$(x_1, \ldots, x_{n-1}) \in \Gamma \qquad \text{and} \qquad x_n = g(x_1, \ldots, x_{n-1}). \tag{5.5}$$

Proof. content... □

Besides, we have that

$$\frac{\partial g}{\partial x_i} = -\frac{\partial_{x_i} f}{\partial_{x_n} f} \tag{5.6}$$

5.2.1 2-D case

Proposition 5.2.2. *Let $f : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a scalar field of class C^2, with Ω an open set, and (x_0, y_0) a point in Ω such that $f((x_0, y_0)) = 0$ and $\partial f/\partial xy((x_0, y_0)) \neq 0$. Then, there exist a rectangle $R = [x_0 - \alpha, x_0 + \alpha] \times [y_0 - \beta, y_0 + \beta]$ such that for every $x \in R$ the equation $F(x, y) = 0$ there is a unique solution $y = g(x)$ where $y \in [y_0 - \beta, y_0 + \beta]$.*

Proposition 5.2.3. *If $y = g(x)$, then g is continuous, its derivative g' is continuous and*

$$g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} \tag{5.7}$$

Proof. PROVE THE FIRST PART.
Proof of the second part.
If we change x an amount h, y will change in a way that $f = 0$ remains true. We can denote this change in y by $k(h)$, since it depends on the variation of x. Therefore, proving that g is continuous is equivalent to see that

$$\lim_{h\to 0} g(x+h) - f(x) = 0 = \lim_{h\to 0} k(h).$$

To do that, let us discuss the following relation

$$0 = f(x+h, y+k(h)) - f(x, y),$$

which is true because $f = 0$ in the rectangle R. Here we can apply the mean value theorem [ ], and since the function is of class C^1 its partial derivatives are continuous and therefore f is differentiable, which allows us to express the theorem as (denoting (x, y) by $\vec{a}$ and $(h), k)$ by $\vec{v}$)

$$0 = f(x+h, y+k(h)) - f(x, y) = f'(\vec{a} + \theta\vec{v}; \vec{v}) = \langle \vec{\nabla} f(\vec{a} + \theta\vec{v}), \vec{v}\rangle_I =$$
$$h\partial_x f(x + \theta h, y + \theta k) + k\partial_y f(x + \theta h, y + \theta k),\ \theta \in (0, 1).$$

Hence,

$$\frac{k(h)}{h} = -\frac{\partial_x f(x+\theta h, y+\theta k(h))}{\partial_y f(x+h, y+k(h))} \Rightarrow \frac{|k(h)|}{|h|} = \frac{|\partial_x f(x+\theta h, y+\theta k(h))|}{|\partial_y f(x+h, y+k(h))|} \leq \frac{M}{m/2} = \frac{2M}{m}.$$

This allows us to bound $k(h)$ with h multiplied by some number, and since we are making h tend to zero, we will get that $k(h)$ too.

$$|k(h)| \leq \frac{2M}{m} h \Rightarrow \lim_{h\to 0} |k(h)| \leq \lim_{h\to 0} \frac{2M}{m}|h| = 0 \Rightarrow \lim_{h\to 0} k(h) = 0$$

Proof of the third part.
To compute the derivative, we have that

$$g'(x) = \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = \lim_{h\to 0} \frac{y + k(h) - y}{h} = \lim_{h\to 0} \frac{k(h)}{h} = \lim_{h\to 0} -\frac{\partial_x(f+\theta h, y+\theta k)}{\partial_y(x+\theta h, y+\theta l)} =$$
$$-\frac{\partial_x f(x, y)}{\partial_y f(x, y)}$$

Where we have resolved the limit considering that partial derivatives are continuous.
Alternative way
Since $y = g(x)$ is a solution, we have that $f(x, g(x)) = 0$ always (in the ball). Therefore, this last expression is not an equation but un identity, that is, a property that satisfies $y = g(x)$. If we compute the derivative with respect to x, we get

$$0 = f(x, g(x)) \Rightarrow 0 = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dg}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} g'(x)$$

Note that in the second partial derivative we compute it with respect to y and not $g(x)$ because what we do is first calculate the derivative with respect to the variable y and then substitute y by the value $g(x)$. After that, we obtain

$$\frac{\partial f}{\partial y} g'(x) = -\frac{\partial f}{\partial x} \Rightarrow g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))}$$

$\square$

Example 5.2.2. Let be the following implicit relation

$$f(x,y) = (x^2+y^2)^2 - 2a^2(x^2-y^2) = 0$$

The following graph is
Let us suppose we want to determine the maxima and minima of the curve. One way is by isolating y, but the operations would be too long. Fortunately, we can use the consequences of the implicit function theorem, in particular, the expression of $g'(x)$ in terms of the partial derivatives of $f(x,y)$. The only task we have is compute $\partial_x f$ and make it equal to 0 (as long as $\partial_y f \neq 0$).

$$f(x,y) = (x^2+y^2)^2 - 2a^2(x^2-y^2) \Rightarrow$$
$$\frac{\partial f}{\partial x} = 4x(x^2+y^2-a^2), \qquad \frac{\partial f}{\partial y} = 4y(x^2+y^2+a^2)$$

The partial with respect to y only is 0 when $y=0$. The partial with respect to x is 0 when $x=0$ and when $x^2+y^2=a^2$. However, at $x=0$ the unique value of y that satisfies $f=0$ is $y=0$, and at this case the partial with respect to y is null. Therefore, this value is not a solution.
To see the other case, we have to consider that $f(x,y)=0$ as we did before, so the solutions are determined by the following system.

$$\begin{cases} x^2+y^2 = a^2 \\ a^4 - 2a^2(x^2-y^2) = 0 \end{cases}$$

The second equation can be expressed as $x^2-y^2 = a^2/2$, the we can add and subtract and see the solutions are $x = \pm a\sqrt{3}/2$ and $y = \pm a/2$. Note that, since both partial derivatives of f are 0, we can't isolate locally x or y and therefore can't define any ball to express $y=g(x)$ or $x=h(y)$.

5.2.2 Extra: difference between identity and equation

The following expression is an equation

$$x-2=0$$

Since it is an equation, we can't compute the derivative. If we did it, we would get

$$x-2=0 \Rightarrow 1-0=0,$$

which is a contradiction. However, if we have an identity as the following one

$$\arcsin \sin x = x,$$

we do can compute the derivative, and will be true.

5.2.3 Curves

Let $f : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a scalar field of class C^1, with Ω an open set and a a point in Ω such that $f(\vec{a}) = 0$ and $\partial f/\partial y(\vec{a}) \neq 0$. By the previous theorem, we know we can isolate y and make it determined by a function $g(x)$. Generally finding the expression $y = g(x)$ is complicated, but we can start with other task. Instead of calculating the points of the curve, we can find the tangent line of the curve at the point (x_0, y_0).
In fact, there are three ways of doing it, and we will start with the gradient. Note that the curve is defined as $f = 0 = \text{ctt}$, and if we differentiate it we get $df = 0$. As we saw previously, this means the gradient is perpendicular to this path. Therefore, the tangent line can be obtained by

$$\langle (\vec{r}-\vec{r}_0), \vec{\nabla} f(\vec{r}_0)\rangle_I = 0 \Rightarrow (x-x_0)\frac{\partial f}{\partial x} + (y-y_0)\frac{\partial f}{\partial y} = 0 \Rightarrow y = -\frac{\partial_x f}{\partial_y f}(x-x_0)+y_0. \tag{5.8}$$

The second method is using directly the Implicit Function Theorem. If we have the expression $y = h(x)$, the slope of the line is $h'(x)$, and by the theorem, this can be calculated by the partial derivatives of f.

$$y = h'(x_0)(x-x_0)+y_0 \Rightarrow y = -\frac{\partial_x f}{\partial_y f}(x-x_0)+y_0$$

Finally, the tangent line can be understood too as an approximation of f of first order. Using the Taylor formula of the section [ ], we get

$$0 = f(x_0, y_0) + \partial_x f(x - x_0) + \partial_y f(y - y_0) = \partial_x f(x - x_0) + \partial_y f(y - y_0) \Rightarrow y = -\frac{\partial_x f}{\partial_y f}(x - x_0) + y_0.$$

Now we can wonder about the perpendicular line, and this can be done too in several ways, but we will present only the gradient manner because the others follow the same reasoning. Since the gradient is already perpendicular, the displacement in this line will be proportional to the gradient vector, so

$$\vec{r} - \vec{r}_0 = \lambda \vec{\nabla} f(x_0, y_0) \Rightarrow x - x_0 = \lambda \partial_x f, y - y_0 = \lambda \partial_y f \Rightarrow \frac{y - y_0}{x - x_0} = \frac{\partial_x f}{\partial_y f} \Rightarrow y = \frac{\partial_x f}{\partial_y f}(x - x_0) + y_0.$$

Note this is coherent with the tangent line since they slopes multiplied result in -1, which is what a pair of perpendicular lines satisfies.

5.2.4 Surfaces

Let $f : \Omega \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}$ be a scalar field of class C^1, with Ω an open set and a a point in Ω such that $f(\vec{a}) = 0$ and $\partial_z(\vec{a}) \neq 0$. Again, we can take the tangent lineal variety to the surface, which will be a plane instead of a line like before. This expression can be obtained by the three methods presented before, but we will omit the last two fo the same reasons. Since the gradient is perpendicular to this surface, our expression will be determined by

$$\langle (\vec{r} - \vec{r}_0), \vec{\nabla} f(\vec{r}_0) \rangle_I = 0 \Rightarrow (x - x_0)\partial_x f + (y - y_0)\partial_y f + (z - z_0)\partial_z f = 0$$

5.3 System of equations

Generalization from the previous theorem

Theorem 5.3.1 (Implicit function Theorem). *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a function with Ω an open set and $\vec{f}$ of class $C^1(\Omega)$. If there is a point $(a, b) \in \Omega$ such that $\vec{f}(\vec{a}, \vec{b}) = 0$ and*

$$\det\left\{ J_{\vec{f},\vec{x}}(\vec{a}, \vec{b}) \right\} = \begin{vmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}, \vec{b}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{a}, \vec{b}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{a}, \vec{b}) & \dots & \frac{\partial f_n}{\partial x_n}(\vec{a}, \vec{b}) \end{vmatrix} \neq 0, \tag{5.9}$$

then there exists an open set $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing (a, b), an open set $V \subseteq \mathbb{R}^m$ containing b, and a function $\vec{g} : V \longrightarrow \mathbb{R}^n$ such that $\vec{h}(\vec{b}) = \vec{a}$ and $\vec{f}(\vec{h}(\vec{y}), \vec{y}) = 0$ for all $y \in V$. Furthermore, $\vec{g} \in C^1(V)$ and $\vec{g}$ is uniquely determined by $(g(y), y) \in U$ for all $y \in V$.

Example 5.3.1. Let $\phi : \Omega \subseteq \mathbb{R}^4 \longrightarrow \mathbb{R}$ and $\psi : \Omega \subseteq \mathbb{R}^4 \longrightarrow \mathbb{R}$ be two functions of class $C^1(\Omega)$ such that $\phi = \psi = 0$ and

$$\begin{vmatrix} \partial_x \phi & \partial_y \phi \\ \partial_x \psi & \partial_y \psi \end{vmatrix} \neq 0$$

around (x_0, y_0, u_0, v_0). Then, the Implicit Function Theorem guaranties that we can isolate y and x in the forms

$$y = g(u, v) \in C^1(\Omega), \qquad x = h(u, v) \in C^1(\Omega). \tag{5.10}$$

Besides, we can find a linear approximation of these functions by the following expressions.

$$y = y_0 + \partial_u g(u - u_0) + \partial_v g(v - v_0), \qquad x = x_0 + \partial_u h(u - u_0) + \partial_v h(v - v_0) \tag{5.11}$$

Now our task is to find these partial derivatives. We know that $\phi(x, y, u, v) = 0$, and since $y = g(u, v)$ and $x = h(u, v)$, we can substitute them in ϕ such that ϕ will satisfy that condition always. And if we derive with respect to u the equation, we get

$$\phi(h(u, v), g(u, v), u, v) = 0 \Rightarrow \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial u} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial u} = 0 \Rightarrow \frac{\partial \phi}{\partial x}\frac{\partial h}{\partial u} + \frac{\partial \phi}{\partial y}\frac{\partial g}{\partial u} + \frac{\partial \phi}{\partial u} = 0,$$

where we have used the chain rule for partial derivatives presented in the equation 3.25. Doing the same with ψ results in an identical expression. With these we can form a system of two equations of the form

$$\begin{cases} \partial_x\phi\partial_u h + \partial_y\phi\partial_y g + \partial_u\phi = 0 \\ \partial_x\psi\partial_u h + \partial_y\phi\partial_u g + \partial_u\psi = 0 \end{cases} \Leftrightarrow \begin{pmatrix} \partial_x\phi & \partial_y\phi \\ \partial_x\psi & \partial_y\psi \end{pmatrix} \begin{pmatrix} \partial_u h \\ \partial_u g \end{pmatrix} = -\begin{pmatrix} \partial_u\phi \\ \partial_u\psi \end{pmatrix}.$$

We have obtained the matrix presented at the beginning, and since we imposed its determinant is not null, this system has unique solutions and we can determine $\partial_y h$ and $\partial_u g$. To determine now the partial derivatives with respect to v, the procedure to follow is the same. We calculate de derivative with respect to v and transform the obtained system in a matrix product.

$$\begin{cases} \partial_x\phi\partial_v h + \partial_y\phi\partial_v g + \partial_v\phi = 0 \\ \partial_x\psi\partial_v h + \partial_y\phi\partial_v g + \partial_v\psi = 0 \end{cases} \Leftrightarrow \begin{pmatrix} \partial_x\phi & \partial_y\phi \\ \partial_x\psi & \partial_y\psi \end{pmatrix} \begin{pmatrix} \partial_v h \\ \partial_v g \end{pmatrix} = -\begin{pmatrix} \partial_v\phi \\ \partial_v\psi \end{pmatrix}$$

5.4 Case where partial derivatives are zero

We have been imposing some conditions in the implicit equations in order to isolate one of the variables in an explicit way. We will discuss now what happens when these conditions are not satisfied. In particular, the conditions were that partial derivatives are not null or a determinant of them, so the topic consists in what happens whey the do are null. By the previous theorems we cannot say anything about this case, and the only experience about this case is with the lemniscata [ ]. In that case partial derivatives at $(0,0)$ were null and y couldn't be isolated, but in fact this does not happens always. Let us see now some examples where y can be expressed as a function of x although the partial derivatives are 0.

Example 5.4.1. Let be the implicit equation $f = (y-x)^2 = 0$, defined in $\mathbb{R}^2$. If we compute their partial derivatives at $(0,0)$, we get

$$\frac{\partial f}{\partial x} = -2(y-x) = 0, \qquad \frac{\partial f}{\partial y} = 2(y-x) = 0.$$

However, this equation has an explicit solution, $y = x$, that is besides of class $C^\infty(\mathbb{R})$.

Example 5.4.2. Let be the implicit equation $f = y^3 - x^2 = 0$, defined in $\mathbb{R}^2$. If we compute their partial derivatives at $(0,0)$, we get

$$\frac{\partial f}{\partial x} = -2x = 0, \qquad \frac{\partial f}{\partial y} = 3y^2 = 0.$$

However, this equation has an explicit solution, $y = x^{3/2}$ (whose derivative does not exist at $(0,0)$).

Therefore, in general, if partial derivatives are zero, we cannot say what happens.
This is because we only knew that f was a function of class $C^1(\Omega)$. If it is from another superior class, it could admit a further Taylor expansion and we would be able to know more. For example, if we take a function of class $C^2(\Omega)$ where the first partial derivatives are zero, we have

$$f(x,y) =$$
$$f(x_0,y_0) + \frac{\partial f}{\partial x}(x-x_0) + \frac{\partial f}{\partial y}(y-y_0) + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}(x-x_0)^2 + \frac{1}{2!}\frac{\partial^2 f}{\partial y^2}(y-y_0)^2 + \frac{\partial^2 f}{\partial x\partial y}(x-x_0)(y-y_0)$$
$$\Rightarrow 0 = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x-x_0)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(y-y_0)^2 + \frac{\partial^2 f}{\partial x\partial y}(x-x_0)(y-y_0) \Rightarrow$$
$$\partial_y^2 f(y-y_0) = \left[\partial_{xy} f \pm \sqrt{\partial_{xy}^2 f - \partial_x^2 f \partial_y^2 f}\right](x-x_0)$$

And if we impose that $\partial_y^2 f \neq 0$, we get

$$y = \frac{\partial_{xy} f \pm \sqrt{\partial_{xy}^2 f - \partial_x^2 f \partial_y^2 f}}{\partial_y^2 f}(x-x_0) + y_0. \tag{5.12}$$

Depending on the discriminant, we will be able to find out the number of solutions.

Example 5.4.3. Let be the equation of the lemniscata, that is, $(x^2+y^2)^2-2a^2(x^2-y^2)=0$. If we apply the previous formula with $(x_0,y_0)=(0,0)$, we get

$$y=\pm x.$$

Therefore, at that point there is not a unique solution $y=g(x)$, which is what we saw the last time.

5.5 Conditioned extremes

Let us suppose we whant to know the point that has the least distance with respect to the origin. In $\mathbb{R}^2$, this id determine by the function $f(x,y)=\sqrt{x^2+y^2}$. We can solve it calculating the minima of the function with the gradient (although it is easy to see the solution is $(0,0)$), so problems like this are already solved. However, let us suppose now we ask again what is the point at least distance but also that satisfies that $\phi(x,y)=0$, for a given ϕ. This requires a new procedure, and it is what we will discuss now.
One possible way is, having $\phi(x,y)=0$, isolate y (in case it is possible), and the calculate the maxima and minima of $f(x,y(x))$. However, there is a better method.

5.5.1 Lagrange multipliers' method

Geometrical justification

Our problem is finding the extreme of a function f under the condition that $\phi=0$. If we first think only about finding the extreme of f, we can see it in a graph of level curves (taking $\mathbb{R}^2$). After having the different level curves, let us draw the curve $\phi=0$. The curve could not lie in the extreme, so it would change its path to another direction. However, we know at some point it will be at the minimum distance of this extreme. At this point, the curve will change the direction and therefore will be perpendicular to its gradient. Besides, at this point we can draw too a level curve of f, and its gradient will also be perpendicular. Therefore, at this point f and ϕ are tangent and

$$\vec{\nabla}f=\lambda\vec{\nabla}\phi.$$

With that, we obtain the following system of equations

$$\begin{cases}\partial_x f=\lambda\partial_x\phi\\ \partial_y f=\lambda\partial_y\phi\\ \phi=0,\end{cases},$$

where our unknowns are x, y, and λ. One way to do this procedure in a schematic way would be constructing the function $L(x,y,\lambda)=f(x,y)-\lambda\phi(x,y)$ and then impose that $\partial_x F=\partial_y F=\partial_\lambda=F$, which results in the same system of equations. If we have several conditions, our function would be $L=f-\lambda_1\phi_1-\cdots-\lambda_n\phi_n$ (here we have more equations than unknowns and it could happen that there is no solution).

Example 5.5.1. Suppose we want to know the least distance to the origin that is in the curve $(x-1/2)^2+y^2=9/4$. The function distance is $f=\sqrt{x^2+y^2}$, so our function L is

$$L=\sqrt{x^2+y^2}-\lambda\left[\left(x-\frac{1}{2}\right)^2+y^2-\frac{9}{4}\right].$$

Although the method consists in constructing this function, we can do a small variation. The distance is a function that is always positive, so finding the extreme of f is equivalent to finding the extreme of f^2. Therefore, we will put f^2 in L and the function will take the following form.

$$L=x^2+y^2-\lambda\left[\left(x-\frac{1}{2}\right)^2+y^2-\frac{9}{4}\right]$$

Now we can imposed the conditions we discussed before and get

$$\begin{cases} \partial_x L = 2x - 2\lambda\,(x - 1/2) = 0 \\ \partial_y L = 2y - 2\lambda y = 0 \\ \partial_\lambda L = (x - 1/2)^2 + y^2 - 9/4 = 0 \end{cases}$$

From the second equation we obtain that $y = 0$ or $\lambda = 1$. If the second one is true, the first equation will result in

$$0 = 2x - 2\left(x - \frac{1}{2}\right) = 2x - 2x + 1 = 1,$$

which is a contradiction and hence, $y = 0$. Substituting it in the last equation, we have

$$0 = \left(x - \frac{1}{2}\right)^2 - \frac{9}{4} \Rightarrow x = \frac{1}{2} \pm \frac{3}{2},$$

where $x = -1$ is the minima and $x = 2$ the maxima.

If we want to isolate one variable from the condition $\phi = 0$, by the Implicit Function Theorem [ ], it must happen that $\phi_x(x_0, y_0) \neq 0$ or $\phi_y(x_0, y_0) \neq 0$, or which is equivalent, that $\phi_x^2(x_0, y_0) + \phi_y^2(x_0, y_0) \neq 0$. Let us suppose we can isolate y, that is, $\phi_y(x_0, y_0) \neq 0$ and that the rest of conditions are satisfied. Then, with $y = g(x)$ we have that $\phi(x, g(x)) = 0$ always, and therefore, $g'(x) = -\phi_x/\phi_y$. Now we want to find the extreme of $f(x, y)$, but since y is determined because of ϕ, the function we need to extreme is $f(x, g(x))$, which can be done by making its derivative equal to zero.

$$f' = 0 = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dg}{dx}$$

We can substitute the expression of g' in the last equation, but the Lagrange multiplier's method consists in another procedure. Since $\phi_y \neq 0$, we have that $f_y(x_0, y_0) = \lambda\phi_y(x_0, y_0)$ for some value of λ. With that, now susbtituting g' in the last equation, we get

$$0 = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}(-1)\frac{\partial \phi}{\partial x}\left(\frac{\partial \phi}{\partial y}\right)^{-1} = \frac{\partial f}{\partial x} - \lambda\frac{\partial \phi}{\partial x} \Rightarrow \frac{\partial f}{\partial x} = \lambda\frac{\partial \phi}{\partial x},$$

which is the second part of Lagrange multipliers' conditions.

Theorem 5.5.1. *Let $u : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function of the form $u = f(x_1, \ldots, x_n)$ where the variables are conditioned by $m < n$ equations $\phi_1 = \cdots = \phi_m = 0$, with $\phi_i : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ a function as f defined in an open set Ω. If $f, \phi_1, \ldots, \phi_m$ are functions of class $C^2(\Omega)$ and not every $\det\{[\phi_i]\}$ with respect m of the n variables is zero at the extreme, then by introducing $\lambda_1, \ldots, \lambda_m$ Lagrange multipliers and making the $n + m$ derivatives of $F(\lambda_1, \ldots, \lambda_m, x_1, \ldots, x_n) = f + \lambda_1\phi_1 + \cdots \lambda_m\phi_m$ equal to zero, we obtain the equations*

$$\frac{\partial F}{\partial \lambda_1} = \cdots = \frac{\partial F}{\partial \lambda_m} = \frac{\partial F}{\partial x_1,} = \cdots = \frac{\partial F}{\partial x_n} = 0. \tag{5.13}$$

These equations form a system of $n + m$ equations for the $n + m$ variables $\lambda_1, \ldots, \lambda_m, x_1, \ldots, x_n$ that must satisfy the extreme of f.

Example 5.5.2. Let be the function $f(x, y, z) = x^2 + y^2$ and suppose we want to know its minima with the restriction of being in the intersection of the following surfaces.

$$\begin{cases} z = 0 \\ z^2 = (y - 1)^3 \end{cases}$$

From the figure, we can se the minima is at the point $(0, 1, 0)$. If we calculate the gradient of f, z, and the surface at that point, we get

$$\vec{\nabla} f((0,1,0)) = 2x\vec{e}_x + 2y\vec{e}_y = 2\vec{e}_y,$$
$$\vec{\nabla}((0,1,0)) = \vec{e}_z,$$
$$\vec{\nabla}(z^2 - (y-1)^3)((0,1,0)) = -3(y-1)^2\vec{e}_y + 2z\vec{e}_z = \vec{0}.$$

Therefore, the equations are

$$\vec{\nabla} f = \lambda_1 \vec{\nabla} z + \vec{\nabla}(z^2 - (y-1)^3)((0,1,0)) \Rightarrow \begin{cases} 0 = \lambda_1 0 + \lambda_2 0 \\ 2 = \lambda_1 0 + \lambda_2 0 \\ 0 = \lambda_1 + \lambda_2 0 \end{cases},$$

but this system has no solution, so the Lagrange multipliers' method does not work. Note that in this example, the jacobians are zero, so we cannot assure the method could work.

Chapter 6

Integrals

6.1 Construction of general Riemann Integral

6.1.1 Partitions

Definition 6.1.1. Let $\Pi_j = \{x_0^{(j)}, \dots, x_{k_j}^{(j)}\}$ with $x_0^{(j)} = a_j$, $x_{k_j}^{(j)} = b_j$, and $j = 1, \dots, n$ be sets that form partitions of $[a_j, b_j]$. Then, we define the *partition of the n-dimensional interval I* as

$$\Pi = \Pi_1 \times \cdots \times \Pi_n. \tag{6.1}$$

Note that a degenerate closed interval cannot be partitioned. We can see the partition divides the interval in $k_1 k_2 \cdots k_n$ n-dimensional intervals. We will denote this number by μ.

Definition 6.1.2. Let I be a closed n-dimensional interval, $f : I \longrightarrow \mathbb{R}$ be a bounded function, and Π a partition of I which divides I into μ n-dimensional closed intervals $I_1, \dots, I_\mu$. Then, we define, we define the following notations

$$m_k(f) \coloneqq \inf_{x \in I_k} f(x), \qquad M_k(f) \coloneqq \sup_{x \in I_k} f(x). \tag{6.2}$$

Note that, since f is bounded and $\mathbb{R}$ is complete [ ], the infimum and supremum exist.

Definition 6.1.3. Let I be a closed n-dimensional interval, $f : I \longrightarrow \mathbb{R}$ be a bounded function, and Π a partition of I which divides I into μ n-dimensional closed intervals $I_1, \dots, I_\mu$. Then, we define the *lower sum of f on I* as

$$\underline{S}(f; \Pi) \coloneqq \sum_{k=1}^{\mu} m_k(f) |I_k|. \tag{6.3}$$

Definition 6.1.4. Let I be a closed n-dimensional interval, $f : I \longrightarrow \mathbb{R}$ be a bounded function, and Π a partition of I which divides I into μ n-dimensional closed intervals $I_1, \dots, I_\mu$. Then, we define the *upper sum of f on I* as

$$\bar{S}(f; \Pi) \coloneqq \sum_{k=1}^{\mu} M_k(f) |I_k|. \tag{6.4}$$

Definition 6.1.5. Let Π, Π' be two partitions of an n-dimensional interval I. We say Π' *is a refinement of* Π, $\Pi' \supseteq \Pi$, if and only if every point in Π is also a point in Π'.

Now we have these generalized definitions, we can prove the same properties of one variable integrals [ ] to the general case.

Proposition 6.1.1. *Let I be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a bounded function. Then,*

1. $\underline{S}(f; \Pi) \leq \bar{S}(f; \Pi)$*, for every partition* Π *of* I*,*
2. $\underline{S}(f; \Pi) \leq \underline{S}(f; \Pi'), \bar{S}(f; \Pi) \geq \bar{S}(f; \Pi')$*, if* $\Pi' \supseteq \Pi$*,*
3. $\underline{S}(f; \Pi) \leq \bar{S}(f; \Pi')$ *for any partitions* Π, Π' *of* I*.*

Definition 6.1.6. Let I be an n-dimensional closed interval, $f : I \longrightarrow \mathbb{R}$ a bounded function, and Π a partition of I. Then, we define the *lower integral* and *upper integral*, respectively, as

$$\underline{\int_I} f \coloneqq \sup_{(\Pi)} \underline{S}(f; \Pi), \qquad \overline{\int_I} f \coloneqq \inf_{(\Pi)} \bar{S}(f; \Pi). \tag{6.5}$$

Proposition 6.1.2. *Let I be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a bounded function. Then, the lower and upper integral exist and*

$$\underline{\int_I} f \leq \overline{\int_I} f. \tag{6.6}$$

Definition 6.1.7. Let $f : I \longrightarrow \mathbb{R}$ be a function. We say f *is Riemann-integrable on* I if and only if

$$\underline{\int_I} f = \overline{\int_I} f. \tag{6.7}$$

In this case, call it the *Riemann integral of* f *over* I and denote it by

$$\int_I f. \tag{6.8}$$

6.1.2 Existence of the integral

Theorem 6.1.3. *[1] Let I be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a bounded function. Then, f is Riemann-integrable on I if and only if it satisfies the following condition*

$$\forall \varepsilon > 0 \exists\ \Pi(\varepsilon) \mid \bar{S}(f; P) - \underline{S}(f; P) < \varepsilon, \forall \Pi \supset \Pi(\varepsilon). \tag{6.9}$$

Proposition 6.1.4. *Let I be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a continuous function. Then, f is Riemann-integrable.*

Definition 6.1.8. Let $S \subseteq \mathbb{R}^n$ be a set. We say S *has an n-dimensional Lebesgue measure zero,* $\lambda(S) = 0$, if and only if one can find, for every $\varepsilon > 0$, a sequence of n-dimensional open intervals $\{I_k\}$ such that

1. $S \subseteq \bigcup_{k \in \mathbb{N}} I_k$,
2. $\sum_{k=1}^{\infty} |I_k| < \varepsilon$.

Theorem 6.1.5. *Let I be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a function that is continuous on I except on a set of n-dimensional Lebesgue measure zero. Then, f is Riemann-integrable.*

Proposition 6.1.6. *Let I be an n-dimensional closed interval, $f, g : I \longrightarrow \mathbb{R}$ two Riemann-integrable functions and λ, μ two real numbers. Then, $\lambda f + \mu g$ is Riemann-integrable and*

$$\int_I \lambda f + \mu g = \lambda \int_I f + \mu \int_I g. \tag{6.10}$$

6.1.3 General sets

Definition 6.1.9. Let $D \subseteq \mathbb{R}^n$ be a bounded set and $f : D \longrightarrow \mathbb{R}$ a bounded function. Then, we define

$$f_D(x) = \begin{cases} f(x), & \text{if } x \in D, \\ 0, & \text{if } x \notin D. \end{cases} \tag{6.11}$$

Proposition 6.1.7. *[1] Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a bounded function with D a bounded set. Let $f_D : I \longrightarrow \mathbb{R}$ be the extension of F in an n-dimensional closed interval I that contains D. If the Riemann integral of f_D on I exists, then*

$$\int_I f_D = \int_{I'} f_D \tag{6.12}$$

for all n-dimensional intervals I' that contain D.

Definition 6.1.10. Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a bounded function with D a bounded set. Then, we say f *is Riemann-integrable over* D if and only if f_D is Riemann-integrable over (at least) one n-dimensional interval $I \supseteq D$. In that case,

$$\int_D f := \int_I f_D. \tag{6.13}$$

Theorem 6.1.8. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a bounded function with D a bounded set. If $\lambda(\partial D) = 0$ and f is continuous almost everywhere on D, then $\int_D f$ exists.*

6.1.4 Jordan content

Definition 6.1.11. Let $D \subseteq \mathbb{R}^n$ be a bounded set and I an n-dimensional closed interval such that $I \supseteq D$. We define the *characteristic function* $\chi_D : \mathbb{R}^n \to \mathbb{R}$ *of* D as the following function

$$\chi_D(x) = \begin{cases} 1 & \text{if } x \in D, \\ 0 & \text{if } x \notin D. \end{cases} \tag{6.14}$$

Definition 6.1.12. Let $D \subseteq \mathbb{R}^n$ be a bounded set, I an n-dimensional closed interval such that I, and Π a partition on I. We define the *inner n-dimensional Jordan content of D* as

$$\underline{J}(D) \coloneqq \sup_{(\Pi)} \underline{S}(\chi_D, \Pi) \tag{6.15}$$

Definition 6.1.13. Let $D \subseteq \mathbb{R}^n$ be a bounded set, I an n-dimensional closed interval such that I, and Π a partition on I. We define the *outer n-dimensional Jordan content of D* as

$$\underline{J}(D) \coloneqq \inf_{(\Pi)} \bar{S}(\chi_D, \Pi) \tag{6.16}$$

Definition 6.1.14. [1] Let $D \subseteq \mathbb{R}^n$ be a bounded set, I an n-dimensional closed interval such that I, and Π a partition on I. We say D *has n-dimensional Jordan content* $J(D)$ if and only if

$$\underline{J}(D) = \bar{J}(D). \tag{6.17}$$

Then, $J(D) = \underline{J}(D) = \bar{J}(D)$. A bounded set which has Jordan content is called *Jordan-measurable*.

Theorem 6.1.9. *Let $D \subseteq \mathbb{R}^n$ be a bounded set and I an n-dimensional closed interval such that I. Then, D is Jordan-measurable if and only if the characteristic function χ_D of D is integrable on I. In that case,*

$$J(D) = \int_I \chi_D. \tag{6.18}$$

Theorem 6.1.10. *Let $D \subseteq \mathbb{R}^n$ be a bounded set and I an n-dimensional closed interval such that I. Then, D is Jordan-measurable if and only if ∂D has Jordan content zero.*

Some more things...

6.1.5 Integration by iteration

Definition of Darboux integrals

Proposition 6.1.11. *Let $I, J \subseteq \mathbb{R}^n$ be two closed intervals and let $f : I \times J \to \mathbb{R}$ be a bounded function. If we define $\phi_x : J \to \mathbb{R}$ by $\phi_x(y) = f(x, y)$, then*

$$\underline{\Phi}(x) = \underline{\int_J} \phi_x, \qquad \bar{\Phi}(x) = \overline{\int_J} \phi_x \tag{6.19}$$

exist for all $x \in I$. If Π_I, Π'_I are two partitions of I and Π_B, Π'_B two partitions of B, then

$$\underline{S}(f; \Pi) \leq \underline{S}(\underline{\Phi}; \Pi_I) \leq \underline{S}(\bar{\Phi}; \Pi_I) \leq \bar{S}(\bar{\Phi}; \Pi'_I) \leq \bar{S}(f; \Pi'), \tag{6.20}$$

where $\Pi = \Pi_A \times \Pi_b$ and $\Pi' = \Pi'_A \times \Pi'_B$. If we define $\psi_y; I \to \mathbb{R}$ by $\psi_y = f(x, y)$, then

$$\underline{\Psi}(y) = \underline{\int_I} \psi_y, \qquad \bar{\Psi}(y) = \overline{\int_I} \psi_y \tag{6.21}$$

exist for all $y \in J$ and

$$\underline{S}(f; \Pi) \leq \underline{S}(\underline{\Psi}; \Pi_J) \leq \underline{S}(\bar{\Psi}; \Pi_J) \leq \bar{S}(\bar{\Psi}; \Pi'_J) \leq \bar{S}(f; \Pi'). \tag{6.22}$$

Theorem 6.1.12 (Fubini's Theorem)**.** *Let $I, J \in \mathbb{R}^n$ be two closed intervals. If $f : I \times J \to \mathbb{R}$ is integrable on $I \times J$, then*

$$\int\limits_{I\times J} f = \int\limits_{I} \underline{\Phi} = \int\limits_{I} \bar{\Phi} = \int\limits_{J} \underline{\Psi} = \int\limits_{J} \bar{\Psi}. \tag{6.23}$$

Theorem 6.1.13. *Let $I, J \subseteq \mathbb{R}^n$ be two closed intervals. If $f : I \times J \to \mathbb{R}$ is integrable on $I \times J$, then*

1. *if ϕ_x is integrable on J, then*

$$\int\limits_{I\times J} f = \int\limits_{I}\int\limits_{J} \phi_x, \tag{6.24}$$

2. *if ψ_y is integrable on I, then*

$$\int\limits_{I\times J} f = \int\limits_{J}\int\limits_{I} \psi_y, \tag{6.25}$$

3. *and if both ϕ_x, ψ_y are integrable on I, J (that is,f is continuous on $I \times J$), then*

$$\int\limits_{I\times J} f = \int\limits_{I}\int\limits_{J} \phi_x = \int\limits_{J}\int\limits_{I} \psi_y, \tag{6.26}$$

or more explicitly,

$$\int\limits_{I\times J} f(\vec{x}, \vec{y}) d\hat{x} d\hat{y} = \int\limits_{I}\int\limits_{J} f(\vec{x}, \vec{y}) d\hat{y} d\hat{x} = \int\limits_{J}\int\limits_{I} f(\vec{x}, \vec{y}) d\hat{x} d\hat{y}. \tag{6.27}$$

Corollary 6.1.14. *Let $I = [a_1, b_1] \times \cdots [a_n, b_n] \subseteq \mathbb{R}^n$ be an n-dimensional closed interval and $f : I \longrightarrow \mathbb{R}$ a continuous function. Then,*

$$\int\limits_{I} f = \int\limits_{a_1}^{b_1} \cdots \int\limits_{a_n}^{b_n} f(\vec{x}) dx_1 \ldots dx_n. \tag{6.28}$$

Theorem 6.1.15. *Let $\alpha, \beta : [a, b] \longrightarrow \mathbb{R}$ be two continuous functions such that $\alpha(x) \leq \beta(x)$ for all $x \in [a, b]$ and let*

$$D = \left\{ (x, y) \in \mathbb{R}^2 \,\middle|\, a \leq x \leq b, \alpha(x) \leq y \leq \beta(x) \right\}. \tag{6.29}$$

If $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, then the Riemann integral of f on D exists and

$$\int\limits_{D} f = \int\limits_{a}^{b} \int\limits_{\alpha(x)}^{\beta(x)} f(x, y) dy dx. \tag{6.30}$$

Theorem 6.1.16. *Let $\alpha, \beta, \gamma, \delta$ be four continuous functions, and let*

$$D = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, a \leq x \leq b, \alpha(x) \leq y \leq \beta(x), \gamma(x, y) \leq z \leq \delta(x, y) \right\}. \tag{6.31}$$

If $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, the the Riemann integral of f on D exists and

$$\int\limits_{D} f = \int\limits_{a}^{b} \int\limits_{\alpha(x)}^{\beta(x)} \int\limits_{\gamma(x,y)}^{\delta(x,y)} f(x, y, z) dz dy dx. \tag{6.32}$$

Bibliography

[1] H. Sagan. *Advanced Calculus: Of Real-valued Functions of a Real Variable and Vector-valued Functions of a Vector Variable*. Houghton Mifflin Company, 1974.

Chapter 7

Line integrals

7.1 Functions of one variable

Let us take a function $f(x,y)$ and define $g(y) = f(x,y)$ with x fixed. Then, if $g(y)$ is integrable, we have

$$F(x) = \int_a^b g(y)dy = \int_a^b f(x,y)dy. \tag{7.1}$$

If the new function $F(x)$ is defined in an interval $[\alpha, \beta]$, the original function should be defined in a rectangle $[\alpha, \beta] \times [a, b]$ in order to be integrable.
Another case would be taking the limits of integration not as constant numbers but some expressions of x, that is,

$$F(x) = \int_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy.$$

Then, the function f is not defined in a rectangle but in a region where $\alpha \leq x \leq \beta$ and $\psi_1(x) \leq y \leq \psi_2(x)$. About this new function F we can ask some questions like how could we compute its derivative.

Example 7.1.1. Let be a sphere of radius r with a density $\varrho(\rho)$. Then, the total mass will be

$$m = \int_0^r \varrho(\rho) 4\pi\rho^2 d\rho.$$

This is a case of the first case of integral. Now, let us suppose ρ is not constant but is expanding at a constant velocity v and ϱ also depends on time. Then, the total mass would be

$$m(t) = \int_0^{vt} \varrho(\rho, t) 4\pi\rho^2 d\rho.$$

This is an example of the second case of the integral.

Theorem 7.1.1. *Let $f(x,y)$ be a continuous function in a rectangle $[\alpha, \beta] \times [a, b]$. Then, the integral $F(x) = \int_a^b f(x,y)dy$ is a continuous function in $[\alpha, \beta]$.*

Proof.

$$|F(x+h) - F(x)| = \left| \int_a^b f(x+h,y)dy - \int_a^b f(x,y)dy \right| = \left| \int_a^b f(x+h,y) - f(x,y)dy \right| \leq$$
$$\int_a^b |f(x+h,y) - f(x,y)|\, dy$$

Since f is continuous and it is defined in a compact, then f is uniformly continuous. Hence, there is a $\delta(\varepsilon)$ such that $|f(x+h,y) - f(x,y)| < \varepsilon/(b-a)$. Then,

$$|F(x+h) - F(x)| \leq \int_a^b |f(x+h,y) - f(x,y)|\, dy \leq \int_a^b \frac{\varepsilon}{b-a} dy = \frac{b-a}{b-a}\varepsilon = \varepsilon$$

□

Theorem 7.1.2. *Let $f(x,y)$ be a continuous function in region delimited by $\psi_1(x)$ and $\psi_2(x)$, with $x \in [\alpha, \beta]$ and ψ_1, ψ_2 continuous functions. Then, the integral $F(x) = \int_a^b f(x,y)dy$ is a continuous function in $[\alpha, \beta]$.*

Proof.

$$|F(x+h)-F(x)| = \left| \int_{\psi_1(x+h)}^{\psi_2(x+h)} f(x+h,y)dy - \int_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy \right| =$$

$$\left| \int_{\psi_1(x+h)}^{\psi_1(x)} f(x+h,y)dy + \int_{\psi_1(x)}^{\psi_2(x)} f(x+h,y) - f(x,y)dy + \int_{\psi_2(x)}^{\psi_2(x+h)} f(x+h,y)dy \right| \leq$$

$$\left| \int_{\psi_1(x+h)}^{\psi_1(x)} f(x+h,y)dy \right| + \left| \int_{\psi_1(x)}^{\psi_2(x)} f(x+h,y) - f(x,y)dy \right| + \left| \int_{\psi_2(x)}^{\psi_2(x+h)} f(x+h,y)dy \right| \leq$$

$$\int_{\psi_1(x+h)}^{\psi_1(x)} |f(x+h,y)|dy + \int_{\psi_1(x)}^{\psi_2(x)} |f(x+h,y) - f(x,y)|dy + \int_{\psi_2(x)}^{\psi_2(x+h)} |f(x+h,y)|dy$$

Making $h \to 0$, all tend to zero. $\square$

Theorem 7.1.3. *Let $R = [\alpha, \beta] \times [a,b]$ be a rectangle and $f : R \longrightarrow \mathbb{R}$ a Riemann-integrable function on $[a,b]$ for all $y \in [c,d]$. If $\partial_y f$ is continuous on R, then the function $F : [c,d] \longrightarrow \mathbb{R}$ is of class $C^1[c,d]$ and*

$$\frac{dF}{dy} = \frac{d}{dy} \int_a^b f(x,y)dx = \int_a^b \frac{\partial f}{\partial y}(x,y)dx. \tag{7.2}$$

Theorem 7.1.4. *Let $f(x,y)$ be a continuous function in $[\alpha, \beta] \times [a,b]$. If the function has $\partial_x f$ continuous, then*

$$\frac{dF}{dx} = \frac{d}{dx} \int_a^b f(x,y)dy = \int_a^b \frac{\partial f}{\partial x}(x,y)dy \tag{7.3}$$

and F' is continuous.

Proof. Let us derive F

$$\lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\to 0} \frac{1}{h} \left[\int_a^b f(x+h,y)dy - \int_a^b f(x,y)dy \right] =$$

$$\lim_{h\to 0} \frac{1}{h} \left[\int_a^b f(x+h,y) - f(x,y)dy \right]$$

By the mean value theorem,

$$\lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\to 0} \frac{1}{h} \left[\int_a^b \frac{\partial f}{\partial x}(x+\theta h, y)hdy \right] = \lim_{h\to 0} \int_a^b \frac{\partial f}{\partial x}(x+\theta h, y)dy =$$

$$\int_a^b \lim_{h\to 0} \frac{\partial f}{\partial x}(x+\theta h, y)dy$$

and since $\partial_x f$ is continuous, $\partial_x f(x+h,y) \to \partial_x f(x,y)$,

$$\frac{dF}{dx} = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = \int_a^b \frac{\partial f}{\partial x}(x,y)dy$$

$\square$

Theorem 7.1.5. *Let $F(x) = \int_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy$ and ψ_1, ψ_2 are of class $C^1(I)$, with $I = [\alpha, \beta]$. If f and $\partial_x f$ are continuous in $[\alpha, \beta] \times [a, b]$, then*

$$\frac{dF}{dx} = \frac{d}{dy} \int\limits_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy = \int\limits_{\psi_1(x)}^{\psi_2(x)} \frac{\partial f}{\partial x}(x,y)dy + f(x, \psi_2(x))\psi_2'(x) - f(x, \psi_1(x))\psi_1'(x). \tag{7.4}$$

Proof. Let us define the following function

$$\phi(u, v, x) = \int\limits_u^v f(x,y)dy$$

By construction, we have $F(x) = \phi(\psi_1(x), \psi_2(x), x)$. Then, the derivative of F is just the derivative of ϕ using the chain rule [ ].

$$\frac{dF}{dx} = \frac{d\phi}{dx}(\psi_1(x), \psi_2(x), x) = \frac{\partial \phi}{\partial u}(\psi_1(x), \psi_2(x), x)\frac{d\psi_1'}{dx} + \frac{\partial \phi}{\partial v}(\psi_1(x), \psi_2(x), x)\frac{d\psi_2'}{dx} + \frac{\partial \phi}{\partial x} =$$

$$-f(x,u)\psi_1'(x) + f(x,v)\psi_2'(x) + \frac{\partial}{\partial x}\int\limits_u^v f(x,y)dy =$$

$$-f(x, \psi_1(x))\psi_1'(x) + f(x, \psi_2(x))\psi_2'(x) + \int\limits_{\psi_1(x)}^{\psi_2(x)} \frac{\partial f}{\partial x}(x,y)dy$$

□

Theorem 7.1.6. *Let $f(x,y)$ be a continuous function defined in the rectangle $[\alpha, \beta] \times [a, b]$. Then,*

$$\int\limits_a^b \int\limits_\alpha^\beta f(x,y)dxdy = \int\limits_\alpha^\beta \int\limits_a^b f(x,y)dydx \tag{7.5}$$

Proof. Let us define the following function

$$y(x,y) = \int\limits_\alpha^x \int\limits_a^y f(x', y')dy'dx'.$$

With this construction, $u(\beta, b)$ is the integral we want to calculate, let us denote it by I. Besides, $u(x, a) = 0$. Let us calculate the partial of u with respect to y

$$\frac{\partial u}{\partial y} = \frac{\partial}{\partial y}\int\limits_\alpha^x \int\limits_a^y f(x', y')dy'dx' = \int\limits_\alpha^x \frac{\partial}{\partial y}\int\limits_a^y f(x', y')dy'dx' = \int\limits_\alpha^x f(x', y)dx'$$

Now, if we integrate u with respect to y,

$$\int\limits_a$$

CONTINUE THE PROOF. □

7.2 Line integrals

Definition 7.2.1. Let $\Gamma \subseteq \mathbb{R}^n$ be a set. We say Γ *is a curve in* $\mathbb{R}^n$ if and only if there exists an interval $I \subseteq \mathbb{R}$ and a continuous function $\gamma : I \longrightarrow R^n$ such that $\mathrm{Im}\,(\gamma) = \Gamma$. In that case, we call

$\vec{r}: I \longrightarrow \Gamma$ a *parametrization of* Γ. We call

$$\begin{aligned} x_1 &= \phi_1(t) \\ &\vdots \\ x_n &= \phi_n(t) \end{aligned} \tag{7.6}$$

the *parameter representation of* Γ and t the *parameter.*

Definition 7.2.2. Let $\Gamma \subseteq \mathbb{R}^n$ be a set. We say Γ *is a simple curve in* $\mathbb{R}^n$ if and only if there exists an interval $I \subseteq \mathbb{R}$ and a continuous injective function $\gamma : I \longrightarrow \mathbb{R}^n$ such that $\operatorname{Im}(\gamma) = \Gamma$. In that case, we call $\gamma : I \longrightarrow \Gamma$ an *injective parametrization of* Γ.

Definition 7.2.3. Let $\Gamma \subseteq \mathbb{R}$ be a curve with a parametrization $\gamma : I \longrightarrow \Gamma$ and $\varphi : J \longrightarrow I$ a diffeomorphism. Then, $\gamma \circ \varphi : J \longrightarrow \Gamma$ is a *reparametrization of* Γ.

Definition 7.2.4. Let $\Gamma \subseteq \mathbb{R}^n$ be a curve with a parametrization $\gamma : I \longrightarrow \Gamma$. We say Γ *is regular* if it is differentiable and its parametrization γ never vanishes.

Definition 7.2.5. Let $\Gamma \subseteq \mathbb{R}^n$ be a curve with a parametrization $\gamma : I \longrightarrow \Gamma$. We say Γ *is piecewise-regular* if

$$\gamma(t) = \begin{cases} \gamma_1(t), & t \in [t_1, t_2] \\ \vdots & \\ \gamma_m(t), & t \in [t_n, t_{n+1}] \end{cases}, \tag{7.7}$$

with $\gamma_1, \ldots, \gamma_m$ being regular curves.

Definition 7.2.6. Let $\mathbb{C}$ be a curve. Then, we define an arc of a curve as a function $\alpha : [a, b] \longrightarrow \mathscr{C}$.

Definition 7.2.7. Let α be an arc of a curve $\mathscr{C}$. Then, we say $\vec{\alpha}$ is a simple curve if $\vec{\alpha}$ is bijective.

This is what we call a parametrization of a curve $\mathscr{C}$.

Definition 7.2.8. Let α be an arc of a curve $\mathscr{C}$. Then, we say $\vec{\alpha}$ is a regular curve if $\vec{\alpha}$ is of class C^1 in pieces, that is, the discontinuity is only present in a finite number of points.

Example 7.2.1. If the variable of $\vec{\alpha}$ is t, then the output could be the position vector.

Definition 7.2.9. Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a vector field. Let be a simple and regular arc of a curve $\mathscr{C}$. Then, we define the *line integral over the path* $\mathscr{C}$ as

$$\int_{\mathscr{C}} \langle \vec{f}, d\vec{r} \rangle_I \coloneqq \lim_{\Pi} \sum_{i=1} \langle \vec{f}(\vec{x}_i'), \vec{x}_{i+1} - \vec{x}_i \rangle_I, \qquad \vec{x}_i' \text{ between } \vec{x}_{i+1}, \vec{x}_i. \tag{7.8}$$

Given the parametrization $\vec{\alpha}$, $\vec{x}_i = \vec{\alpha}(t_i)$, so there is a partition too for the interval $[a, b]$. We can take $t_{i+1} - t_i = \Delta t = (b - a)/n$. Then,

$$\vec{x}_0 = \vec{\alpha}(t_0) = \vec{\alpha}(a), \qquad \vec{x}_1 = \vec{\alpha}(t_1) = \vec{\alpha}(a + \Delta), \ldots, \vec{x}_n = \vec{\alpha}(t_n) = \vec{\alpha}(b)$$

Since $\vec{\alpha} \in C^1$, we can apply the mean value theorem for vector fields [ ], that is,

$$\vec{\alpha}(t_{i+1}) - \vec{\alpha}(t_i) = \Delta t \frac{d\vec{\alpha}}{dt}(t_i')$$

With that,

$$\lim_{\Pi} \sum_{i=1} \langle \vec{f}(\vec{x}_i'), \vec{x}_{i+1} - \vec{x}_i \rangle_I = \lim_{\Pi} \sum_{i=1} \langle \vec{f}(\vec{\alpha}(t_i')), \vec{\alpha}(t_{i+1}) - \vec{\alpha}(t_i') \rangle_I =$$

$$\lim_{\Pi} \sum_{i=1} \left\langle \vec{f}(\vec{\alpha}(t_i')), \frac{d\vec{\alpha}}{dt}(t_i') \right\rangle_I \Delta t = \int_a^b \left\langle \vec{f}(\vec{\alpha}(t)), \frac{d\vec{\alpha}}{dt}(t) \right\rangle_I dt$$

Now we know how does the line integral work, that is, we can express the integral as a Riemann integral, already defined. However, if the path has not a parametrization, we still do not know how does the integral work.
The first notation can lead to some confusions because $d\vec{r}$ is not the variable of integration. For that, another common notation is (more for mathematics)

$$\int_{\mathscr{C}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I . \tag{7.9}$$

Example 7.2.2. Work is an example of a line integral of a vector field $\vec{F}$. Also the circulation of electric and magnetic fields.

Some of the possible questions is what happens when we parametrize a curve in two different ways. Another better definition:

Definition 7.2.10. Let $\Gamma \subseteq \mathbb{R}^n$ be a curve of class $C^1(I)$ and let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a continuous function. Then, we define the *line integral of $\vec{f}$ along* Γ as

$$\int_{\vec{f}} \langle :=, \mathrm{d}\mathbf{r} \rangle_I \int_a^b \langle \vec{f}(\gamma(t)), \gamma'(t) \rangle_I dt. \tag{7.10}$$

7.2.1 Properties of line integrals for vector fields

Proposition 7.2.1. *Let ... Then,*

$$\int_{\mathscr{C}} \langle \lambda \vec{f}(\vec{\alpha}) + \mu \vec{g}(\vec{\alpha}), d\vec{\alpha} \rangle_I = \lambda \int_{\mathscr{C}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I + \mu \int_{\mathscr{C}} \langle \vec{g}(\vec{\alpha}), d\vec{\alpha} \rangle_I . \tag{7.11}$$

Proof.

$$\int_{\mathscr{C}} \langle \lambda \vec{f}(\vec{\alpha}) + \mu \vec{g}(\vec{\alpha}), d\vec{\alpha} \rangle_I = \int_a^b \left\langle \lambda \vec{f}(\vec{\alpha}(t)) + \vec{g}(\vec{\alpha}(t)), \frac{d\vec{\alpha}}{dt}(t) \right\rangle_I dt =$$

$$\lambda \int_a^b \left\langle \vec{f}(\vec{\alpha}(t)), \frac{d\vec{\alpha}}{dt}(t) \right\rangle_I dt + \mu \int_a^b \left\langle \vec{g}(\vec{\alpha}(t)), \frac{d\vec{\alpha}}{dt}(t) \right\rangle_I dt = \lambda \int_{\mathscr{C}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I + \mu \int_{\mathscr{C}} \langle \vec{g}(\vec{\alpha}), d\vec{\alpha} \rangle_I .$$

□

Proposition 7.2.2. *Let... Then,*

$$\left| \int_{\mathscr{C}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I \right| \leq \int_a^b \left| \left\langle \vec{f}(\vec{\alpha}(t)), \frac{d\vec{\alpha}}{dt}(t) \right\rangle_I \right| dt \tag{7.12}$$

Proposition 7.2.3 (Additivity)**.** *Let $\mathscr{C} = \mathscr{C}_1 + \mathscr{C}_2$ ... Then,*

$$\int_{\mathscr{C}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I = \int_{\mathscr{C}_1} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I + \int_{\mathscr{C}_2} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I . \tag{7.13}$$

Proposition 7.2.4. *The line integral depend on the orientation*

$$\int_{\mathscr{C}_{A \to B}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I = - \int_{\mathscr{C}_{B \to A}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha} \rangle_I \tag{7.14}$$

Let $\vec{\alpha}$ be a parametrization as a function of t. Then, the integral is

$$\int_{\mathscr{C}_{A\to B}} \langle \vec{f}(\vec{\alpha}), d\vec{\alpha}\rangle_I = \int_{\mathscr{C}_{A\to B}} \langle \vec{f}(\vec{\alpha}), \frac{d\alpha}{dt}\rangle_I dt$$

Let us reparametrize this with $t = \varphi(\varrho)$, with $\varphi \in C^1$ (so we can derivate). Then, changing the variable in the integral, we have

$$\int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} \langle \vec{f}(\vec{\alpha}), \frac{d\alpha}{dt}\rangle_I \frac{d\varphi}{d\varrho} d\varrho$$

This leads to two possible case. We want that t and φ are one-to-one related, so it could happen that ϱ_0 corresponds to $t = a$ and ϱ_1 to $t = b$ or viceversa. It cannot happen that two ρ correspond to one value of t because the function would not be bijective.
In the first scenario, we have

$$\int_{\varrho_0}^{\varrho_1} \langle \vec{f}(\vec{\alpha}), \frac{d\alpha}{dt}\rangle_I \frac{d\varphi}{d\varrho} d\varrho = \int_{\varrho_0}^{\varrho_1} \langle \vec{f}(\vec{\alpha}), \frac{d\vec{\beta}}{d\varrho}\rangle_I d\varrho = \int_{\varrho_0}^{\varrho_1} \langle \vec{f}(\vec{\alpha}), d\vec{\beta}\rangle_I, \qquad \vec{\beta} = \vec{\alpha}(\varphi(\varrho))$$

If we are in the second case, we will have that $\varrho_1 \le \varrho \le \varrho_0$. Then, we are not calculating the integral in an increasing integral. Therefore, to satisfy the definition of line integral, we need to change the sign of the whole integral.

Proposition 7.2.5. *If the parametrization preserve the orientation of the curve, the line integral does not depend on the parametrization*

7.2.2 Scalar fields

Definition 7.2.11. For scalar fields

$$\int_{\mathscr{C}} \varrho(\vec{x})dl, \tag{7.15}$$

where dl is the length differential.

We make another parametrization $\vec{\alpha} : [a, b] \to \mathbb{R}^n$. We make a partition in t such that $t_i < t_{i+1}$. With that, the integral is

$$\int_{\mathscr{C}} \varphi(\vec{x})dl = \lim_{\Pi} \sum_{i=1} \varphi(\vec{x}_i')\|\vec{x}_{i+1} - \vec{x}_i\|$$

Again, we use the main value theorem and take $\Delta t = (b - a)/n$,

$$\int_{\mathscr{C}} \varphi(\vec{x})dl = \lim_{\Pi} \sum_{i=1} \varphi(\vec{\alpha}(t_i')) \left\| \frac{d\vec{\alpha}}{dt}(t_i') \right\| |\Delta t|$$

If we use $\Delta t > 0$,

$$\int_{\mathscr{C}} \varphi(\vec{x})dl = \lim_{\Pi} \sum_{i=1} \varphi(\vec{\alpha}(t_i')) \left\| \frac{d\vec{\alpha}}{dt}(t_i') \right\| \Delta t = \int_a^b \varphi(\vec{\alpha}(t)) \left\| \frac{d\vec{\alpha}}{dt} \right\| dt$$

Note that now the orientation does not affect the integral because we take the absolute value of $d\vec{x}$ (always than $dt > 0$ and does not change the sign).
This is the definition.

Definition 7.2.12. Let $\Gamma \subseteq \mathbb{R}^n$ be a curve with a parametrization $\gamma \in C^1(I)$ and $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ a continuous function. Then, we define the *line integral of f along* Γ as

$$\int_f := \mathrm{d}r \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt. \tag{7.16}$$

Example 7.2.3. The length of a curve $\mathscr{C}$ is

$$L = \int_a^b \left\| \frac{d\vec{\alpha}}{dt} \right\| dt$$

Example 7.2.4. Let us suppose we have a string with a density $\varphi(\vec{x})$. Then, the mass of the spring is

$$\int_{\text{spring}} \varphi dl = \int_a^b \varphi(\vec{\alpha}(t)) \left\| \frac{d\vec{\alpha}}{dt} \right\| dt$$

doing a parametrization that follows the shape of the string

Example 7.2.5. For scalar fields: calculating the mass of a cylinder with density $\varrho(\vec{x})$ or the charge.

Example 7.2.6. The center of mass of the spring of the previous example is

$$\vec{r}_{CM} = \int_{\text{spring}} \varphi \vec{r} dl = \int_{\text{spring}} (\alpha_x \vec{e}_x + \alpha_y \vec{e}_y + \alpha_z \vec{e}_z) \varphi(\vec{\alpha}(t)) \|\vec{\alpha}'(t)\| dt$$

doing a parametrization over a variable t. We do one integral for each component.

Example 7.2.7. Let $\vec{f} = xy^2\vec{e}_1 + y^2\vec{e}_2$ and the path $\mathscr{C}^* : x^2 + y^2 = 1$ (the asterisc means it is an increasing curve). To parametrize it, we will use $\vec{x} = \vec{\alpha}(\theta) = (\cos\theta, \sin\theta)$, with $\theta \in [0, \pi/2]$. We can see this parametrization follows the direction of the path. Then, the integral is

$$\int_0^{\pi/2} \left\langle \varphi(\vec{\alpha}(\theta)), \frac{d\vec{\alpha}}{d\theta} \right\rangle_I d\theta = \int_0^{\pi/2} \left\langle \varphi((\cos\theta, \sin\theta)), \frac{d\vec{\alpha}}{d\theta} \right\rangle_I d\theta =$$

$$\int_0^{\pi/2} \left\langle \cos\theta \sin^2\theta \vec{e}_1 + \sin^2\theta \vec{e}_2, -\sin\theta \vec{e}_1 + \cos\theta \vec{e}_2 \right\rangle_I d\theta =$$

$$\int_0^{\pi/2} -\cos\theta \sin^3\theta + \sin^2\theta \cos\theta d\theta = \frac{1}{12}$$

Let us do the same with another parametrization, which is $\vec{x} = \vec{\alpha}(x) = (x, \sqrt{1-x^2})$, where we take the positive square root because we are in the first quadrant. To follow the direction of the path, we put the limits from 1 to 0.

$$\int_1^0 \left\langle \varphi(\vec{\alpha}(x)), \frac{d\vec{\alpha}}{dx} \right\rangle_I dx = \int_1^0 \left\langle \varphi((x, \sqrt{1-x^2})), \frac{d\vec{\alpha}}{dx} \right\rangle_I dx =$$

$$\int_1^0 \left\langle x(1-x^2)\vec{e}_x + (1-x^2)\vec{e}_y, x\vec{e}_1 + \sqrt{1-x^2}\vec{e}_y \right\rangle_I dx =$$

$$\int_1^0 x(1-x^2)\left(1 - \frac{1}{\sqrt{1-x^2}}\right) dx = \frac{1}{12}$$

Example 7.2.8. Let be $\varphi = xy$ in the same path $\mathscr{C}^*$ as before. Again, we do the parametrization with θ.

$$\int_0^{\pi/2} \cos\theta \sin\theta d\theta = \frac{1}{2}$$

With the second parametrization,

$$\int_0^1 x\sqrt{1-x^2}\frac{1}{\sqrt{1-x^2}}dx = \frac{1}{2}$$

Arc length

Definition 7.2.13. Let $\Gamma \subseteq \mathbb{R}$ be a curve with a parametrization $\vec{r} : I \longrightarrow \Gamma$ and Π a partition of I. Then, we define the *length of the polygonal* as

$$L(\vec{r};\Pi) := \sum_{i=1}^{n} \|\vec{r}(t_i) - \vec{r}(t_{i-1})\|. \tag{7.17}$$

Definition 7.2.14. Let $\Gamma \subseteq \mathbb{R}$ be a curve with a parametrization $\vec{r} : I \longrightarrow \Gamma$ and Π a partition of I. Then, we define the length of Γ as

$$L(\Gamma) = \sup_{\Pi} L(\vec{r};\Pi). \tag{7.18}$$

Definition 7.2.15. Let $\Gamma \subseteq \mathbb{R}$ be a curve. We say *the curve* Γ *is rectifiable* if and only if its length is finite.

Proposition 7.2.6. *Let* $\Gamma \subseteq \mathbb{R}$ *be a curve with a parametrization* $\vec{r} : I \longrightarrow \Gamma$*. If* $\vec{r} \in C^1(I)$*, then* Γ *is rectifiable and*

$$L(\Gamma) = \int_a^b \|\vec{r}'(t)\|dt. \tag{7.19}$$

7.3 Conditions of gradient

7.3.1 Definitions

Definition 7.3.1. Let $(\mathbb{M}, d)$ be a metric space and $S \subseteq \mathbb{M}$ a set. We say *S is connected* if and only if it cannot be represented as the union of two or more disjoint non-empty open subsets. If not, we say *S is disconnected*

Definition 7.3.2. Let $(\mathbb{M}, d)$ be a metric space and $S \subseteq \mathbb{M}$ a set. We say *S is path-connected* if every pair of points can be connected by a continuous path that belongs to the set.

Definition 7.3.3. A *path* is a continuous function $f : [0,1] \longrightarrow \mathbb{R}^n$. We call $f(0)$ the *initial point* and $f(a)$ the *terminal point.*

Definition 7.3.4. Let S be a connected set. We say *it is simply connected* if it is path-connected and every path between two points can be continuously transformed into any other such path while preserving the two endpoints in question. Equivalently, S is simply connected if it is path connected and any loop in S can be contracted to a point. Otherwise, we say it is multiply connected.

Definition 7.3.5. Let S be a simply connected set. We say *S is convex* if for all pair of points $a, b \in S$, the segment defined by

$$[a,b] = \{x \mid x = (1-t)a + tb, 0 \le t \le 1\} \tag{7.20}$$

is contained in S, that is, if every pair of points can be connected by a straight line that belongs to the set.

Theorem 7.3.1 (Gradient Theorem)**.** *Let* $\Gamma \subseteq \mathbb{R}^n$ *be curve with a parametrization* $\gamma : [a,b] \longrightarrow \Gamma$ *and* $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ *a differentiable function with* D *a connected set. Then,*

$$\int_{\vec{\nabla} f} \langle =, \mathrm{d}\mathbf{r}\rangle_I f(\gamma(b)) - f(\gamma(a)). \tag{7.21}$$

Corollary 7.3.2. *Let $\Gamma \subseteq \mathbb{R}^n$ be closed curve with a parametrization $\gamma : [a,b] \longrightarrow \Gamma$ and $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ a differentiable function with D a connected set. Then,*

$$\oint_\Gamma \langle \vec{\nabla} f, d\vec{r} \rangle_I = 0. \tag{7.22}$$

Theorem 7.3.3. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a continuous function, with Ω a connected and open set. If the line integral of $\vec{f}$ between two pints is independent on the curve and, given a point $a \in \Omega$ we define the function*

$$\varphi(\vec{x}) = \int_a^x \langle \vec{f}, d\vec{r} \rangle_I = \int_\gamma \langle \vec{f}, \mathrm{d}\mathbf{r} \rangle_I \tag{7.23}$$

with Γ an arbitrary piece-wise regular. Then,

$$\vec{f} = \vec{\nabla} \varphi, \qquad \forall x \in \Omega. \tag{7.24}$$

Proof. Calculate partial derivataives by definition, using the definition of φ. Then, see $\partial_k \varphi = f_k$. ☐

Theorem 7.3.4. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function of class $C^0(\Omega)$ (continuous), with Ω a connected an open set. Then, the following conditions are equivalent.*

1. *$\vec{f} = \vec{\nabla} \varphi$ for some $\varphi(\vec{x})$.*
2. *The line integral of $\vec{f}$ does not depend on the path*
3. *The line integral of $\vec{f}$ over every piece-wise regular closed path contained in Ω is zero.*

Proof. We have seen 2 implies 1 and 1 implies 3, so it only remains to see 3 implies 2. We make a closed path that intersects two points a and b and divide it in two segments delimited by these two points. Then,

$$\mathscr{C} = \mathscr{C}_{1,a\to b} + \mathscr{C}_{2,b\to a} = 0 \Rightarrow \mathscr{C}_{1,a\to b} = -\mathscr{C}_{2,b\to a} \Rightarrow \mathscr{C}_{1,a\to b} = \mathscr{C}_{2,a\to b}$$

Before each $\mathscr{C}$ should be an integral. We conclude the integral does not depend on the path. ☐

Theorem 7.3.5. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ a function of class $C^1(D)$, with D not necessarily connected. If $\vec{f} = \vec{\nabla}\varphi$, then*

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}. \tag{7.25}$$

In particular, in $\mathbb{R}^3$, if $\vec{f} = \vec{\nabla}\varphi$ then $\vec{\nabla} \times \vec{f} = \vec{0}$.

We will see the other implication is not true (only if the set is connected, we will see it later).

Example 7.3.1. Let be

$$\vec{f} = -\frac{y}{x^2+y^2}\vec{e}_x + \frac{x}{x^2+y^2}\vec{e}_y.$$

It is defined in $\mathbb{R}^2$ $(0,0)$ (not connected set). We can see $\partial_y f_1 = (y^2 - x^2)/(x^2+y^2)^2 = \partial_x f_2$ (and the rotational is zero). However, if we make the circulation integral, we do not get zero. For that, let us take the curve $\mathscr{C} : x^2 + y^2 = R^2$. We do the parameterization of θ, and the integral is

$$\oint_{\mathscr{C}} \langle \vec{f}, d\vec{\alpha} \rangle_I = \int_0^{2\pi} \langle -\sin\theta/R\vec{e}_x + \cos\theta/R\vec{e}_y, -R\sin\theta\vec{e}_x + R\cos\theta\vec{e}_y \rangle_I = \int_0^{2\pi} d\theta = 2\pi \neq 0$$

Therefore, $\vec{f}$ is not a gradient. In fact, every closed path that contains the pint $(0,0)$ inside will leads to the same value of the integral (we will see later why), and every closed path that does not contain that point will lead to a null integral. (This is equivalent, in complex variable, to the theorem of residue of Cauchy.)

Theorem 7.3.6. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function of class $C^1(\Omega)$ with Ω a convex and open set. Then, $\vec{f} = \vec{\nabla}\varphi$ if and only if*

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}, \qquad \forall \vec{x}_i \in \Omega. \tag{7.26}$$

Proof. The right direction of implication it is already proved. Let us see the left direction.
Second implication. Since it is convex, two points can be connected by a straigth line. We define

$$\varphi(\vec{x}) = \int_{\vec{a}}^{\vec{x}} \vec{f}, d\vec{\alpha}$$

We follow straith line from a to x. From this, we can see $\partial_{x_k}\varphi = f_k$.

□

As we see, the double implication depends deeply in the kind of set of the domain.
In particular, in $\mathbb{R}^3$, the fact that $\partial_{x_j} f_i = \partial_{x_i} f_j$ is equivalent to $\vec{\nabla} \times \vec{f} = \vec{0}$

Example 7.3.2. Let be again

$$\vec{f} = -\frac{y}{x^2+y^2}\vec{e}_x + \frac{x}{x^2+y^2}\vec{e}_y.$$

In this case the second implication is not true (as we saw in [ ]) because the set is not convex.

7.4 Plane integrals

Definition 7.4.1. Let $[a, b] \times [c, d]$ be a closed rectangle. If we make a partition Π_1 in $[a, b]$ and Π_2 in $[c, d]$, then we define the partition of the rectangle as a collection of sets $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$.

Definition 7.4.2. Let f be a fuction. We say f *is scalonated* if there is a partition of a rectangle R such that f is constant in every point of the open rectangles that define the partition.

Proposition 7.4.1. *If f and g are scalonated functions, then $c_1 f + c_2 g$ is a scalonated function.*

Definition 7.4.3. Let f be a scalonaed function and $R = [a, b] \times [c, d]$ a rectangle of domain of f. Then, given a partition Π over R, we define the subrectangle r as $[x_{i-1}, x_i] \times y_{j-1}, y_j$. In each subrectangle, the function behaves as a constant c_{ij}. Then, we define

$$\iint_R f = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}(x_i - x_{i-1})(y_j - y_{j-1}). \tag{7.27}$$

Note that, if we refine the parition, the result won't change. We will see some properties.

Proposition 7.4.2. *Let f be a function defined in a rectangle R. Then,*

$$\iint_R f = \int_a^b \int_c^d f dydx = \int_c^d \int_a^b f dxdy \tag{7.28}$$

Proof. From the definition of sum and knowing that

$$\sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f dx = \int_a^b f dx$$

□

Properties

Proposition 7.4.3.

$$\int_a^b \int_c^d \lambda f + \mu g dydx = \lambda \int_a^b \int_c^d f dydx + \mu \int_a^b \int_c^d g dydx \tag{7.29}$$

Proposition 7.4.4. *If we have a rectangle R that is $R_1 \cup R_2$. Then,*

$$\iint_R f = \iint_{R_1} f + \iint_{R_2} f \tag{7.30}$$

Theorem 7.4.5 (Theorem of comparation)**.** *If $R \subseteq T$, then*

$$\iint_R f \leq \iint_T f \tag{7.31}$$

Integral of a defined and bounded function in the rectangle

We don't need it to be scalonated

Definition 7.4.4. Let f be a bounded fnction, that is, $|f| \leq M$. Let s and t two scalonated functions such that $s \leq f \leq t$ (that exist because f is bounded). If there is a unique number I such that

$$\iint_R s \leq \iint_R f \leq \iint_R t$$

for every pair of functions that satisfy the condition we presented before, then I is called double integral of f over the rectangle R and it is denoted by

$$\iint_R f(x,y)dxdy. \tag{7.32}$$

Is is interesting note that this notation is more used in physics than mathematics. In mathematics, is more used the notation

$$\iint_R f$$

because the value of the integral does not depend on the variables. In some other cases (not very common), they use the notation

$$\iint_R f(x,y)d(\hat{xy})$$

As we did with Riemann integral, we can do with two variables.

Definition 7.4.5. Let S

$$S := \left\{ \iint_R s \,\middle|\, s \leq f, \forall (x,y) \in R \right\} \tag{7.33}$$

Definition 7.4.6. Let S

$$T := \left\{ \iint_R t \,\middle|\, t \geq f, \forall (x,y) \in R \right\} \tag{7.34}$$

Observation: these sets are non-empty because f is bounded.
Observation: we can see $s \leq t$, so by the theorem of comparation,

$$\iint_R s \leq \iint_R t.$$

We will study now the supreme of S and the infime of T. We will see the infime is always greater than the supreme.

Proof. Suppose the supreme of S is greater than the infime of T. By the inequality we saw before, every integral in t is an upper bound of the integral of s, and in particular, the infime of T too. Then, if the infime is lower than the supp of S, it would mean there is an upper bound smaller than the supprem. This is a contradiction because the supreme is the smallest. Then, the supreme of S is always lower or equal than the infime of T. $\square$

Theorem 7.4.6. *We say f is integrable if the supreme of S and the infime of T are equal.*

Proof. Since the integral in f is between the integral in s and in t, when the limit coincide, the integral of f has a unique value. $\square$

Consequences

Proposition 7.4.7. *Additivity, ... LOOK IN THE PDF MORE*

Theorem 7.4.8. *Let $f : R \longrightarrow \mathbb{R}$ be an integrable function in the rectangle $R = [a, b] \times [c, d]$. Let us suppose there exist*

$$A(y) = \int_a^b f(x, y)dx, \forall y \in [c, d], \qquad \int_c^d A(y)dy.$$

Then,

$$\iint_R f = \int_c^d \int_a^b f(x, y)dxdy \tag{7.35}$$

Proof. Using the theorem of comparation, the fact that s and t are scalonated (we know how to compute integrals for scalonated functions), ant that f is integrable. $\square$

Corollary 7.4.9. *Integrating first by y and then by x is the same.*

Now we will see if continuous functions are integrable.

Theorem 7.4.10. *Let $f : R \longrightarrow \mathbb{R}$ be a continuous function defined in a rectangle $R = [a, b] \times [c, d]$. Then, f is integrable in R and the integral can be calculated by iterated integrals.*

Proof. $\square$

7.4.1 Bounded functions with discontinuities

Definition 7.4.7. Let $A \subseteq \mathbb{R}^2$ be a bounded set. We say A *has content/measurement/area null* if $\forall \varepsilon > 0$ there is a finite covering of A with rectangles of area $< \varepsilon$.

Example 7.4.1. A point in $\mathbb{R}^2$ and a segment (finite piece of a line) satisfy the previous definition.

Let us make some observations about these sets

1. A finite set of points have measure zero.
2. The finite union of bounded sets of measure zero is a set of measure zero.
3. Every subset of a set of measure zero has a measure zero.
4. Every segment of a line has a measure zero.

We will see now that the value of the integral won't be afected by what happens to the function (if it is bounded) in the reagion of the set of measure zero.

Theorem 7.4.11. *Let f be a bounded function defined in $R = [a, b] \times [c, d]$. If the set of discontinuities of f has measure zero, then f is integrable in R.*

7.4.2 Integration over general regions

Let S be a bounded region of $\mathbb{R}^2$ and R a rectangle such that $S \subseteq R$. Then, we define in R the function

$$\hat{f} = \begin{cases} f(x,y), & (x,y) \in S \\ 0, & (x,y) \notin S \end{cases}.$$

Now we have a function defined in a rectangle, where we know how to integrate. However, the function is not more continuous, because in the boundary of S the function is not continuous. We need to see how do these discontinuities behave to know if we will be able to integrate $\hat{f}$ or not. We it is possible, we define the integral in S as

$$\iint_S f := \iint_R \hat{f}.$$

We can find two possible cases of regions.
Region of first kind:

$$S_1 = \{(x,y) | a \leq x \leq b, \varphi_2(x) \leq y \leq \varphi_1(x)\}, \qquad \varphi_1, \varphi_2 \text{ continuous}$$

And the second kind:

$$S_1 = \{(x,y) | c \leq y \leq d, \psi_2(x) \leq y \leq \psi_1(x)\}, \qquad \psi_1, \psi_2 \text{ continuous}$$

We need a previous result

Theorem 7.4.12. *Let $\varphi(x)$ be a continuous function defined in $[a,b]$. Then, the graph of $\varphi(x)$ is a set of measure zero.*

Theorem 7.4.13. *Let S be a region of first kind. Let f be a continuous function define in S. Then, the double integral of f in S exists and*

$$\iint_S f = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dydx. \tag{7.36}$$

Theorem 7.4.14. *If S is of second kind, then the integral exists and*

$$\iint_S f = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dxdy. \tag{7.37}$$

There are sets of both kinds, like the rectangle or the disk. In general, if we have now integrals of more general shape, we can divide the region in other ones of first and second kind, and then add them.

7.4.3 Areas of plane regions

Let be a region S_1, and evaluate the integral

$$\iint_{S_1} 1,$$

or, using the properties of S_1,

$$\int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} 1dydx = \int_a^b \varphi_2(x) - \varphi_1(x)dx$$

This can be interpreted as the area in the region S_1.

Double integrals of volume

If the function f is not 1, then the double integral can be interpreted as a volume.

$$V = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} z(x,y)dydx \tag{7.38}$$

This volume has three plane sides: the top side and the two sides at $x = a$ and $x = b$. This is because the integral from $\varphi_1(x)$ to $\varphi_2(x)$ of f with respect to y can be interpreted as the area $A(X)$, so we are taking and infinite sum of areas, which leads to a volume. If we want now to compute the area compressed between two functions z, it will be

$$V = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} z_2(x,y) - z_1(x,y)dydx$$

Differential of volume

We can think the volume as a sum of infinite differential volumes, expressed as

$$dV = f(x,y)dxdy$$

Example 7.4.2. Let us suppose we want to calculate the volume of an ellipsoid.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$

The volume will be

$$\iint_S 2c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$$

We need to define now S. This region, in general, is the projection over the plane xy. In this case, this region is determined by $x^2/a^2 + y^2/b^2 \leq 1$. We can represent it as a region of first kind (and second kind).

$$S = \left\{(x,y) \in \mathbb{R}^2 \middle| -a \leq x \leq a, -b\sqrt{1 - x^2/a^2} \leq y \leq b\sqrt{1 - x^2/a^2}\right\}$$

Then, the integral is

$$\int_{-a}^{a} \int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} 2c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}dydx = 8c\int_0^a \int_b^{b\sqrt{1-x^2/a^2}} \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}dydx,$$

where we have used the function is even (with respect to x and with respect to y) and the integral is done in intervals of the form $[-c, c]$. If we denote now $1 - x^2/a^2$ by A^2, we have

$$8c\int_0^a \int_0^{bA} \sqrt{A^2 - \frac{y^2}{b^2}}dydx = 8c\int_0^a A\int_0^{bA} \sqrt{1 - \frac{y^2}{(Ab)^2}}dydx.$$

Now, with the substitution $y/Ab = \sin\theta$, we have

$$8c\int_0^a A\int_0^{\pi/2} b\cos^2\theta d\theta dx.$$

To calculate the first integral, we will use the following procedure (it is intuitive with some diagrams)

$$\int_0^{\pi/2} \cos^2\theta d\theta = \int_0^{\pi/2} \sin^2\theta d\theta \Rightarrow \int_0^{\pi/2} d\theta = \int_0^{\pi/2} \sin^2\theta + \cos^2\theta d\theta = 2\int_0^{\pi/2} \cos^2\theta d\theta \Rightarrow$$

$$\int_0^{\pi/2} \cos^2\theta d\theta = \frac{1}{2}\int_0^{\pi/2} d\theta = \frac{\pi}{4}$$

Then,

$$8c\int_0^a A^2 \int_0^{\pi/2} b\cos^2\theta d\theta dx = 8c\int_0^a A^2 b\frac{\pi}{4}dx = 2bc\pi\int_0^a 1 - \frac{x^2}{a^2}dx = 2\pi bc\left(a - \frac{a}{3}\right) = \frac{4}{3}\pi abc$$

We get the correct result, but there is a faster way to obtain this formula. By symmetry, we know the role played by a, b, and c are equal (if we exchange the variables the volume must be the same), so the volume must have the form

$$V = kabc$$

which is coherent too with the dimensions of volume. To find the constant, we can take $a = b = c = r$ and obtain the formula of a sphere, which it is known, so

$$V_e = kr^3 = \frac{4}{3}\pi r^3 \Rightarrow k = \frac{4}{3}\pi \Rightarrow V = \frac{4}{3}\pi abc$$

7.4.4 Mass

If we have a surface density $\sigma(x, y)$, the total quantity will be

$$\iint_S \sigma(x, y) \tag{7.39}$$

Center of mass

Going from discrete to continuous distribution of mass,

$$x_{CM} = \iint_S x\sigma(x, y), \qquad y_{CM} = \iint_S y\sigma(x, y) \tag{7.40}$$

7.5 Green's theorem for plane integrals

Definition 7.5.1. A *Jordan curve* $\Gamma \subseteq \mathbb{R}^n$ is a closed simple piece-wise regular curve.

Note that is bijective always and that, since it i closed, if the extremes are at a and b, then $\vec{\alpha}(a) = \vec{\alpha}(b)$.

Theorem 7.5.1. *Let $P, Q : D \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ two functions of class $C^1(D)$- Let $\Gamma \subseteq D$ be a Jordan curve and $R = \Gamma \cup \operatorname{int}\Gamma$ (which is simply connected). If R is a region of first and second kind, then*

$$\iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dxdy = \oint_{\partial R} Pdx + Qdy \tag{7.41}$$

If the region is not simultaneously of first and second kind, we can decompose it in several regions of both kinds, and then add it together. This is true because, we integrate two times over the boundaries, but these sets are sets of measure zero, and do not affect to the actual value (for the plane integral). For the line integral, it does not affect either because the shared path between two curves is canceled always (in one path we integrate it in one direction and in the other in the opposite).

Corollary 7.5.2. *The area can be calculated as*

$$\oint_{\partial R} xdy = \oint_{\partial R} -ydx \tag{7.42}$$

and infinite other always that $\partial_x Q - \partial_y P = 1$.

7.5.1 Green's theorem consequences

Theorem 7.5.3. *Let $\vec{f} : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function defined by $\vec{f} = P\vec{e}_x + Q\vec{e}_y$ and of class $C^1(\Omega)$, with Ω a simply connected and open set. Then, $\vec{f} = \vec{\nabla}\varphi$ if and only if*

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}. \tag{7.43}$$

Proof. The fact that the gradient implies the equality was already proven. We need to see the second implication.
From the Green theorem, since Ω is simply connected,

$$0 = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dxdy = \oint_{\partial R} Pdx + Qdy = \int_{\mathscr{C}} \langle \vec{f}, d\vec{\alpha} \rangle_I$$

So $\vec{f}$ is a gradient. □

For that reason, in the previous example, if we did an integral in a region that does not contain a discontinuity, the integral resulted in zero.

Theorem 7.5.4 (Green's Theorem for multiply connected regions)**.** *Let $\Gamma_1, \dots, \Gamma_k$ be Jordan curves such that*

1. $\Gamma_i \cap \Gamma_j = \emptyset$,
2. $\forall i, \Gamma_i \subseteq \operatorname{int}\Gamma_1$,
3. $\forall i \neq j \geq 2 \Gamma_i \subseteq \operatorname{ext}\Gamma_j \Leftrightarrow \operatorname{int}\Gamma_i \cap \operatorname{int}\Gamma_j = \emptyset$.

Let $R = \Gamma_1 \cup \operatorname{int}\Gamma_1 - \bigcup_{i=2}^k \operatorname{int}\Gamma_i$ (which is multiply connected if $k \geq 2$). Let $\vec{f} : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function defined by $\vec{f} = P\vec{e}_x + Q\vec{e}_y$ and of class $C^1(\Omega)$, with $\Omega \supseteq R$ a connected and open set. Then,

$$\iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dxdy = \oint_{\Gamma_1} Pdx + Qdy - \sum_{i=2}^n \oint_{\Gamma_i} Pdx + Qdy \tag{7.44}$$

Theorem 7.5.5 (Invariance over path deformation)**.** *Let $P, Q : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ two functions of class $C^1(\Omega)$ with Ω a connected and open set and such that $\partial_y P = \partial_x Q$ for all point of Ω. Let $\Gamma_1, \Gamma_2 \in \Omega$ be two Jordan curves such that*

1. $\Gamma_2 \subseteq \operatorname{int}\Gamma_2$,
2. $\operatorname{int}\Gamma_1 \cap \operatorname{ext}\Gamma_2 \subseteq \Omega$.

Then,

$$\oint_{\Gamma_1} Pdx + Qdy = \oint_{\Gamma_2} Pdx + Qdy$$

Theorem 7.5.6. *Let f be a function in Ω. Let $u = \phi(x, y)$ and $v = \psi(x, y)$ be two bijective functions of class $C^1(\Omega)$ and such that $\partial(\phi, \psi)/\partial(x, y) \neq 0$. Then.*

$$\iint_\Omega f(x, y)dxdy = \iint_\Upsilon f(\phi^{-1}(u, v), \psi^{-1}(u, v)) \left\| \frac{\partial(x, y)}{\partial(u, v)} \right\| dudv. \tag{7.45}$$

Proof. We have $dA = dxdy$ and $\vec{r}(u,v) = x(u,v)\vec{e}_x + \vec{e}_y$. If we move in u direction,

$$d\vec{r} = \frac{\partial \vec{r}}{\partial u} du$$

And if we move in v direction,

$$d\vec{r} = \frac{\partial \vec{r}}{\partial v} dv$$

In small regions, the area of two displacements will be a parallelogram (this can be more rigorously justified). This area is calculated by norm of cross product.

$$dA = \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| dudv$$

Let us calculate the norm.

$$\left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} = ... = \left| \frac{\partial(x,y)}{\partial(u,v)} \right|$$

□

One could thinks this absolute value is a strange element in the change of variable because it has not been used in one variable integrals. However, we can see the absolute value is in fact part of the integration in one variable with this example.

Example 7.5.1. Suppose we have

$$\int_1^2 f(ax)dx$$

and make the change of variable $ax = t$
We integrate always in an increasing way. If $a > 0$,

$$\int_a^{2a} f(t)\frac{1}{a}dt = \int_a^{2a} f(t)\frac{1}{|a|}dt$$

IF $a < 0$,

$$\int_a^{2a} f(t)\frac{1}{a}dt = \int_{2a}^{a} f(t)\frac{1}{-a}dt = \int_{2a}^{a} f(t)\frac{1}{|a|}dt$$

As we see, since the interval of integration is always increasing, it is necessary to write an absolute vale so the limits of integration are ordered.

Example 7.5.2. We will calculate the volume of a sphere integrating in a region

$$V = 8\iint_S \sqrt{R^2 - x^2 - y^2}dxdy =$$

We will make the change $x = \rho\cos\theta, y = \rho\sin\theta$

$$V = 8\iint_T \sqrt{R^2 - \rho^2} \left\| \frac{\partial(x,y)}{\partial(\rho,\theta)} \right\| d\rho d\theta = 8\int_0^R \int_0^{\pi/2} \sqrt{R^2 - \rho^2}\rho d\theta d\rho = \frac{4}{3}\pi R^3$$

Note that this jacobian gets null at $\rho = 0$. However, this is not a problem because it is a set of measure zero, so does not affect the value.

7.5.2 About the differential elements on the integral

In a double integral, we see $dxdy$. Is this a product? No, and we can see it with an example. If we make $x = \rho\cos\theta, y = \rho\sin\theta$, it results in

$$dxdy = (d\rho\cos\theta - \rho\sin\theta d\theta)(d\rho\sin\theta + \rho\cos\theta d\theta)$$

In this product there will be $d\rho d\rho$ and $d\theta d\theta$, so it does not correspond to the jacobian. Besides, For that reason some books do not write $dxdy$ but $d(xy)$ or $d(\hat{xy})$.
In fact, it has been studied and this "product" is a exterior product or wedge product, which has the following properties

1. $a \wedge b = -b \wedge a$
2. $a \wedge a = 0$

From this definition, we can see

$$dx \wedge dy = \rho d\rho \wedge d\theta$$

Example 7.5.3. We want to evaluate in the region $x = 0, y = 0, x + y = 2$.

$$I = \iint\limits_S e^{\frac{y-x}{y+x}} dxdy$$

If we make the change $u = y - x$, $v = y + x$, we get

$$I = \iint\limits_T e^{\frac{u}{v}} \left\| \frac{\partial(x,y)}{\partial(u,v)} \right\| dudv = \iint\limits_T e^{\frac{u}{v}} \frac{1}{2} dudv$$

Note in this case it is easy to calculate x and y as functions of u, v. In some case where it will more difficult, we can calculate $\partial(u,v)/\partial(x,y)$ and then take the inverse. Now, we need to calculate the limits of integration.

$$\begin{gathered} x = 0, 0 \leq y \leq 2 \Rightarrow u = y = v \in [0,2] \\ y = 0, x \in [0,2] \Rightarrow -u = x = v \in [0,2] \\ x + y = 2 \Rightarrow v = 2 \end{gathered}$$

With this we can see an area bounded by the new limits. Now one could wonder if the interior of the new region corresponds to the interior of the first one or the exterior. We can see the interior corresponds to the interior because we need the second area to be finite, since the first is finite. Another option is to make small changes in the values of the boundary and see if moving to the interior in one corresponds to moving to the interior in the other.
Now, we can integrate

$$I = \frac{1}{2} \int\limits_0^2 \int\limits_{-v}^{v} e^{u/v} dudv = e - \frac{1}{e}$$

Important: if we have a region with a gap, by transformation the new set will also have a gap (because it must be bijective).

7.5.3 Generalization of change of variables for multiple integrals

Theorem 7.5.7 (Jacobi's Theorem)**.** *[?] Let $\Omega \subseteq \mathbb{R}^n$ be an open set and D a Jordan-measurable set such that $D \cup \partial D \subseteq \Omega$. If $g : \Omega \longrightarrow \mathbb{R}^n$, $g \in C^1(\Omega)$, $g : \operatorname{int}(D) \longrightarrow \mathbb{R}^n$, $J[g(x)] \neq 0$ for all $x \in \operatorname{int}(D)$, then $g(D)$ is Jordan-measurable and its Jordan content is given by*

$$J(g(D)) = \int\limits_D \|J[\vec{g}(\vec{x})]\| \, d\hat{x}. \tag{7.46}$$

If, in addition, $f : g(D) \longrightarrow \mathbb{R}$ is bounded and continuous, then

$$\int\limits_{\vec{g}(D)} f(\vec{y})d\hat{y} = \int\limits_{D} (f \circ \vec{g})(\vec{y}) \left\| J[g(\vec{y})] \right\| d\hat{y}. \tag{7.47}$$

Or more explicitly, with $g(\vec{y}) = (\phi_1, \dots, \phi_n)$ and $\phi_i = \phi_i(\vec{y})$,

$$\int\!\!\cdots\!\!\int\limits_{\vec{g}(D)} f(x_1, \dots, x_n)dx_1 \dots dx_n = \int\!\!\cdots\!\!\int\limits_{D} f(\phi_1, \dots, \phi_n) \left\| \begin{matrix} \partial_{y_1}\phi_1 & \dots & \partial_{y_n}\phi_1 \\ \vdots & \ddots & \vdots \\ \partial_{y_1}\phi_n & \dots & \partial_{y_n}\phi_n \end{matrix} \right\| dy_1 \dots dy_n \tag{7.48}$$

Triple integrals

We must not say volume integrals because double integrals can represent volume too
In double integrals we had two options of regions. Now we have three.

Definition 7.5.2. let $V \subseteq \mathbb{R}^3$. We say V *is projectable in xy* if there exists the following set

$$V_{xy} = \{(x, y) \in S | \phi_1(x, y) \le z \le \phi_2(x, y)\},$$

with ϕ_1, ϕ_2 continuous functions.

In this case, we can do

$$\iint\limits_{R_{xy}} \int\limits_{\phi_1}^{\phi_2} F dzdxdy$$

Definition 7.5.3.

$$R_{xz} = \{(x, z) \in S | \psi_1(x, z) \le y \le \psi_2(x, z)\} \tag{7.49}$$

Definition 7.5.4.

$$R_{yz} = \{(y, z) \in S | \psi_1(y, z) \le x \le \psi_2(y, z)\} \tag{7.50}$$

Example 7.5.4. Cylindrical coordinates

$$\frac{\partial(x, y, z)}{\partial(\rho, \theta, z)} = \begin{vmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho \Rightarrow dxdydz = \rho d\rho d\theta dz$$

Again, it has a problem in $\rho = 0$ but has measure zero.

Important, if we are in three dimensions, a gap that is an area (only two dimensions), it is a set of measure zero.

Example 7.5.5. Spherical coordinates

$$\frac{\partial(x, y, z)}{\partial(\rho, \theta, \varphi)} = \begin{vmatrix} \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & \rho\sin\theta\cos\varphi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} = \rho^2\sin\theta \Rightarrow dxdydz = \rho^2\sin\theta d\rho d\theta d\varphi$$

Example 7.5.6. We want to calculate the area of an ellipse. The integral is (we did it before)

$$V = 2\iint\limits_{S} c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} dxdy$$

We can make the changes $x = a\rho\cos\varphi, y = b\rho\sin\varphi$, with $\varphi \in [0, 2\pi], \rho \in [0, 1]$. Then,

$$V = 2c\int\limits_0^{2\pi}\int\limits_0^1 \sqrt{1 - \rho^2} \left\| \frac{\partial(x, y)}{\partial(\rho, \varphi)} \right\| d\rho d\varphi = 2abc\int\limits_0^{2\pi}\int\limits_0^1 \rho\sqrt{1 - \rho^2} d\rho d\varphi = \frac{4}{3}\pi abc$$

Exercise: calculate the volume of the ellipsoid by using spherical coordinates (adapted to a, b, c as before) with a triple integral.

Chapter 8

Differential Geometry

8.1 Representation of parametric surfaces

Definition 8.1.1. [1] Let $\Sigma \subseteq \mathbb{R}^3$ be a set. We say Σ *is a surface* if and only if there exists a connected set $\Omega \subseteq \mathbb{R}^3$ with nonempty interior and a continuous function $\vec{r} : \Omega \longrightarrow \mathbb{R}^3$ such that $\text{Im}(f) = \Sigma$.
In this case, we call $\sigma : \Omega \longrightarrow \Sigma$ *the parametrization of* Σ. We call

$$\begin{aligned} x_1 &= \phi_1(u, v), \\ x_2 &= \phi_2(u, v), \\ x_3 &= \phi_3(u, v) \end{aligned} \tag{8.1}$$

the parameter representation of Σ and (u, v) the *parameter.*

Definition 8.1.2. [1] Let $\Sigma \subseteq \mathbb{R}^3$ be a set. We say Σ *is a simple surface* if and only if there exists a connected set $\Omega \subseteq \mathbb{R}^3$ with nonempty interior and a injective continuous function $\sigma : \Omega \longrightarrow \mathbb{R}^3$ such that $\text{Im}(f) = \Sigma$.
In this case, we call $\sigma : \Omega \longrightarrow \Sigma$ *the injective parametrization of* Σ.

From class: Let $\Omega \subseteq \mathbb{R}^3$. We say Ω *is a simple surface* if there exists a bijective function $\vec{r} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ of class C^1 such that $\text{Im}(f) = \Omega$.
Besides, if we have a function $\vec{r}(u, v)$, we can move in $\vec{r}(u, v + dv)$ or $\vec{r}(u + du, v)$ direction. In order to be a surface, these two directions must be independent, otherwise we would have only a curve (one dimension of the image). This is will be accomplished if it satisfies

$$\left\| \frac{\partial \vec{r}}{\partial v} \times \frac{\partial \vec{r}}{\partial u} \right\| \neq 0 \tag{8.2}$$

And, if this happens, that only occurs in a region of measure zero.
Now we can wonder about the area of a surface. Each small region of the surface will have a correspondence to the area of the domain. As we have seen before, these differential of areas can be expressed as

$$\left\| \frac{\partial \vec{r}}{\partial v} \times \frac{\partial \vec{r}}{\partial u} \right\| dudv$$

Then, the area will be define as the sum of these infinitesimal areas.

Definition 8.1.3. Let Σ be a surface. Then, we define the *area of* Σ as

$$\iint_{\Omega} \left\| \frac{\partial \vec{r}}{\partial v} \times \frac{\partial \vec{r}}{\partial u} \right\| dudv \tag{8.3}$$

We can also define the length of a curve contained in a surface. The point in the curve are points of the surface that are constrained to be in the points of the curve. This can be done with an extra parametrization. If the image was dependent on u, v, this will follow a curve if we make $u = u(t), v = v(t)$. Now, to calculate the length, we can obtain the differential of length by $dl = \|d\vec{r}\|$. Let us calculate $d\vec{r}$ now.

$$\vec{r} = \vec{r}(u(t), v(t)) \Rightarrow d\vec{r} = \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv, \qquad du = \frac{du}{dt} dt, dv = \frac{dv}{dt} dt \tag{8.4}$$

Then,

$$dl^2 = \langle d\vec{r}, d\vec{r} \rangle_I = \left\langle \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv, \frac{\partial \vec{r}}{\partial u} du + \frac{\partial \vec{r}}{\partial v} dv \right\rangle_I = \left\| \frac{\partial \vec{r}}{\partial u} \right\|^2 du^2 + 2 \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I dudv + \left\| \frac{\partial \vec{r}}{\partial v} \right\|^2 dv^2$$

Then, the length will be

$$L = \int dl = \int_a^b \sqrt{\left\| \frac{\partial \vec{r}}{\partial u} \right\|^2 \left(\frac{du}{dt} \right)^2 + 2 \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I \frac{du}{dt} \frac{dv}{dt} + \left\| \frac{\partial \vec{r}}{\partial v} \right\|^2 \left(\frac{dv}{dt} \right)^2} dt$$

Notice the differential of length can also be written as

$$dl^2 = \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} \left\| \frac{\partial \vec{r}}{\partial u} \right\|^2 & \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I \\ \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I & \left\| \frac{\partial \vec{r}}{\partial v} \right\|^2 \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \tag{8.5}$$

This is called the metric of the surface, and is like a transformation of the euclidean space to the properties of the surface. If we took a plane surface, this matrix would be I_2. In fact, knowing the expression of the length of a curve in a surface, we know everything about the surface. To see that, let us use the following notation

$$\left\| \frac{\partial \vec{r}}{\partial u} \right\|^2 \equiv E, \qquad \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I \equiv F, \qquad \left\| \frac{\partial \vec{r}}{\partial v} \right\|^2 \equiv G$$

These quantities are called the *Gauss' fundamental quantities*. We have seen the surface is caracterized by the cross product of partial derivatives. And as we know, this cross product can be rephrased as

$$\left\| \frac{\partial \vec{r}}{\partial v} \times \frac{\partial \vec{r}}{\partial u} \right\|^2 = \left\| \frac{\partial \vec{r}}{\partial u} \right\|^2 \left\| \frac{\partial \vec{r}}{\partial v} \right\|^2 - \left\langle \frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v} \right\rangle_I^2 = EG - F^2 \tag{8.6}$$

Example 8.1.1. Let Σ be a surface. The points can be expressed using spherical coordinates with $\rho = R$

$$\begin{cases} x = R \sin\theta \cos\varphi \\ y = R \sin\theta \sin\varphi \\ z = R\cos\theta \end{cases} \tag{8.7}$$

Now, we need to calculate the position vector with respect to the parametritzation variables.

$$\partial_\theta \vec{a} = R\cos\theta\cos\varphi\vec{e}_x + R\cos\theta\sin\varphi\vec{e}_y - R\sin\theta\vec{e}_z$$
$$\partial_\varphi \vec{r} = -R\sin\theta\sin\varphi\vec{e}_x + R\sin\theta\cos\varphi\vec{e}_y$$

Then,

$$= R^2\cos\theta\sin\theta\vec{e}_z + R^2\sin^2\theta\sin\varphi\vec{e}_y + R^2\sin^2\theta\cos\varphi\vec{e}_X \Rightarrow norm = R^2\sin\theta$$

The area will be

$$\int_0^\pi \int_0^{2\pi} R^2 \sin\theta d\varphi d\theta = 4\pi R^2.$$

Example 8.1.2. Length of a curve over a sphere.

$$\left\| \frac{\partial \vec{r}}{\partial \theta} \right\|^2 = R^2$$
$$\left\| \frac{\partial \vec{r}}{\partial \theta} \right\|^2 = R^2\sin^2\theta$$
$$= 0$$

so

$$\begin{pmatrix} d\theta & d\varphi \end{pmatrix} \begin{pmatrix} R^2 & 0 \\ 0 & R^2\sin^2\theta \end{pmatrix} \begin{pmatrix} d\theta \\ d\varphi \end{pmatrix} = R^2 d\theta^2 + R^2\sin^2\theta d\varphi^2$$

Let us calculate the length of the curve $\theta = \theta_0$. This means $d\theta = 0$, and we can parametrize only φ. We can do it directly with the variable φ. Then,

$$L = \int_0^{2\pi} R\sin\theta_0 d\varphi = 2\pi R\sin\theta_0$$

Now with $\varphi = \varphi_0$ ($d\varphi = 0$), we get

$$L = \int_0^\pi R d\theta = \pi R$$

8.2 Surface integrals

8.2.1 Scalar fields

Definition 8.2.1. Let $f : D \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}$ be a function and a surface $\Sigma \subseteq D$ with a parametrization $\sigma : \Omega \longrightarrow \Sigma$ of class $C^1(\Omega)$. If f is bounded in Σ, then we define the *surface integral of f* as

$$\iint_\Sigma f ds := \iint_\Omega f \left\| \frac{\partial \sigma}{\partial u} \times \frac{\partial \sigma}{\partial v} \right\| dudv. \tag{8.8}$$

Let $u, v \in T$ two paremeters of a surface and $t, s \in M$ two other parameters of the same surface Σ such that the relation beteween these four are bijective, of class C^1 and such that $J \neq 0$ (except in sets of measure zero).

$$\begin{aligned} u &= u(s,t), & v &= v(s,t) \\ s &= s(u,v), & t &= t(u,v) \end{aligned}$$

Then, $\vec{r}(s,t) = \vec{r}(u(s,t), v(s,t))$. We could wonder if the integral with respect one parametrization and the other have the same value, that is, if

$$\iint_S f \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| dudv = \iint_T f \left\| \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right\| dsdt \tag{8.9}$$

We can see that (prove it)

$$\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \frac{\partial(u,v)}{\partial(s,t)}$$

Then, if we have another parametrization such that the jacobian of the right is positive, the orientation is presserved. If is negative, the orientation is changed and the vector surface is inverted. In particular, making $s = v, t = u$, the orientation is changed.
However, since the integral contains these terms in absolute value, we have

$$\iint_T f \left\| \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right\| dsdt = \iint_T f \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| \left\| \frac{\partial(u,v)}{\partial(s,t)} \right\| dsdt = \iint_S f \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| dudv \tag{8.10}$$

When we have used the formula of change of variable for plane regions (the regions of the parameters s, t, u, v are planes).

Theorem 8.2.1. *Under these conditions, the integral does not depend on the parametrization.*

8.2.2 Vector fields

Definition 8.2.2. Let $\vec{f} : D \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be a function and a surface $\Sigma \subseteq D$ with a parametrization $\sigma : \Omega \longrightarrow \Sigma$ of class $C^1(\Omega)$. If $\vec{f}$ is bounded in Σ, then we define the *surface integral of f* as

$$\int_\Sigma \langle \vec{f}, \mathrm{d}\mathbf{s} \rangle_I := \iint_\Omega \left\langle \vec{f}, \frac{\partial \sigma}{\partial u} \times \frac{\partial \sigma}{\partial v} \right\rangle_I dudv \tag{8.11}$$

Notice is not the same the order of the product. Now, since there is not absolute value, the change of parametrization change the value. If this parametrization does not change the orientation (positive jacobian), the value is the same, and if it does change the orientation (negative jacobian), the value only will change the sign.

8.2.3 Notation for differential forms

We have seen

$$\int\limits_{\Sigma} \langle \vec{f}, d\vec{s} \rangle_I = \iint\limits_{S} \left\langle \vec{f}, \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\rangle_I dudv$$

Let us calculate the cross product

$$\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} = \frac{\partial(y,z)}{\partial(u,v)}\vec{e}_x + \frac{\partial(z,x)}{\partial(u,v)}\vec{e}_y + \frac{\partial(x,y)}{\partial(u,v)}\vec{e}_z$$

where we have changed the sign of the second term by inverting the order of the variables. We can see the order of the variables and vectors are cyclic, that is (putting the vector at beginning), $(1,2,3),(2,3,1),(3,1,2)$. The integral can be calculated then by

$$\iint\limits_{S} f \cdot d\vec{s} = \iint\limits_{S} f_x \frac{\partial(y,z)}{\partial(u,v)} dudv + f_y \frac{\partial(z,x)}{\partial(u,v)} dudv + f_z \frac{\partial(x,y)}{\partial(u,v)} dudv \tag{8.12}$$

With the notation of exterior product of differential forms,

$$\int\limits_{\Sigma} \langle \vec{f}, \mathrm{d}\mathbf{s} \rangle_I = \iint\limits_{\Sigma} f_x dy \wedge dz + f_y dz \wedge dx + f_z dx \wedge dy \tag{8.13}$$

We can see this is true by the properties of differential forms

$$dy \wedge dz = \left(\frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv \right) \wedge \left(\frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv \right) =$$

8.3 Stokes' Theorem

Theorem 8.3.1 (Stokes' Theorem). *Let $\vec{f} : D \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be a function of class C^1 in a simple surface $\Sigma \subseteq D$ (with a parametrization $\sigma : \Omega \longrightarrow \Sigma$ of class $C^2(\Omega)$ in all points except in sets of measure zero), Ω of first and second kind, and the Jordan curve $\Gamma = \partial\Sigma$ piece-wise regular. Then,*

$$\int\limits_{\Sigma} \langle \vec{\nabla} \times \vec{f}, \mathrm{d}\mathbf{s} \rangle_I = \oint\limits_{\partial\Sigma} \langle \vec{f}, d\vec{r} \rangle_I, \tag{8.14}$$

where the direction of the surface vector is defined by the right-hand rule with respect the direction of rotation of the curve.

Theorem 8.3.2 (Stokes' Theorem for non-convex sets). *Let $S = \vec{r}(T)$ a simple surface with T a plane region of first and second kind, with $\vec{r}$ of class C^2 except of sets of measure zero and the S being bounded by a Jordan curve Γ regular in pieces. Let $\vec{F} = (P, Q, R)$ defined in S of class C^1. Then,*

$$\int\limits_{\Sigma} \langle \vec{\nabla} \times \vec{F}, d\vec{s} \rangle_I = \oint\limits_{\partial\Sigma} \langle \vec{F} \cdot d\vec{\alpha} \rangle_I - \oint\limits_{C} \langle \vec{F} \cdot d\vec{\alpha} \rangle_I, \tag{8.15}$$

where the direction of the surface vector is defined by the right-hand rule with respect the direction of rotation of the curve and C is the boundary of the hole.

Theorem 8.3.3. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function of class $C^1(D)$ and let B be an orthonormal basis. Then,*

- *If D is convex, then $\partial_i f_j = \partial_j f_i \Leftrightarrow \vec{f} = \vec{\nabla}\phi$ [ ],*
- *If D is simply connected and $m = 2$, then $\partial_i f_j = \partial_j f_i \Leftrightarrow \vec{f} = \vec{\nabla}\phi$ [ ],*
- *If D is simply connected and $m = 3$, then $\partial_i f_j = \partial_j f_i \Leftrightarrow \vec{f} = \vec{\nabla}\phi$.*

Example 8.3.1. Let be the following function

$$\vec{f} = -\frac{y}{x^2+y^2}\vec{e}_x + \frac{x}{x^2+y^2}\vec{e}_y \tag{8.16}$$

We saw previously this function was not a gradient and the circulation along a curve that contained the origin resulted in 2π. Let us make now the function $\vec{g}(x, y, z)$ such that $\vec{g} = \vec{f}$ for all z. This new function has a singularity in the z axis, and the rotational is not null there. Even trying to generate a surface that does not passes through this region, it is not possible. Then, in this case the Stoke's theorem is not applicable.

Example 8.3.2. Let Σ be a surface with a parametrization σ that has the form of a cylinder. In that case, the surface is not more simple, since two point will have the same image (in order to close the cylinder).
Let us now cut this surface. In this situation, the Stoke's theorem it is applicable (before cutting is was not possible). The surface integral of the equaility is the same as the integral we would do in the original surface because the cut only affects in a set of measure zero, so the value does not change. With respect the line integral, the extra path traveled is canceled because in one part is followed in one direction and in the other in the opposite one. Then, we conclude the Stoke's theorem works in this surface.

Example 8.3.3. Let now be the Möbius strip. We could think this is equivalent to the cylinder and then we could use the Stokes' theorem. However, in this situation the theorem does not work because the surface is not orientable. The Stokes' theorem only works if the surface is orientable.

8.4 Divergence Theorem

Theorem 8.4.1 (Gauss' Theorem)**.** *Let $V \subseteq \mathbb{R}^3$ be a symmetric projectable solid that is limited by an orientable surface and $\vec{f} : D \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ a function of class $C^1(D)$. If $d\vec{s}_{ext}$ is the exterior differential of surface, then*

$$\iiint_V \langle \vec{\nabla}, \vec{f} \rangle_I dv = \oiint_{\partial V} \langle \vec{f}, d\vec{s}_{ext} \rangle_I. \tag{8.17}$$

8.4.1 Consequences

Theorem 8.4.2. *Let $\vec{f} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be a function of class C^1 in a closed ball $B(a, t)$. Then,*

$$\langle \vec{\nabla}, \vec{f}(a) \rangle_I = \lim_{t \to 0} \frac{1}{|V(t)|} \oiint_{\partial V(t)} \langle \vec{f}, d\vec{s}_{ext} \rangle_I. \tag{8.18}$$

Similarly, using the stokes' theorem, we can see

Theorem 8.4.3. *Let $\vec{f} : D \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be a function of class C^1 in a closed disc $\Sigma(t) \subseteq D$. Then*

$$\langle \vec{n}, \vec{\nabla} \times \vec{f}(a) \rangle_I = \lim_{t \to 0} \frac{1}{|\Sigma(t)|} \oint_{\partial \Sigma(t)} \langle \vec{f}, d\vec{r} \rangle_I. \tag{8.19}$$

where

- *The surface is arbitrary*
- *The curve is the boundary of S*
- *$\vec{n}$ is a unitary vector perpendicular to S*
- *the integral is done with the right-hand rule and the vector $\vec{n}$*

Both theorems do not require a sphere and a disk, they can be more general sets.

Theorem 8.4.4. *With the same conditions as the Gauss' theorem, then*

$$\iiint_V \vec{\nabla} \times \vec{f} dv = -\oiint_{\partial V} \vec{f} \times d\vec{s}_{ext}. \tag{8.20}$$

8.5 Properties of nabla operator

Proposition 8.5.1. *We have*

- $\langle\vec{\nabla}, \vec{f} \times \vec{n}\rangle_I = \langle\vec{n}, \vec{\nabla} \times \vec{f}\rangle_I - \langle\vec{f}, \vec{\nabla} \times \vec{n}\rangle_I$
- $\langle\vec{n}, \vec{f} \times d\vec{s}\rangle_I = \langle d\vec{s}, \vec{n} \times \vec{f}\rangle_I = \langle\vec{f}, d\vec{s} \times \vec{n}\rangle_I$

8.6 Others

Theorem 8.6.1. *Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and $g : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ two functions and a an interior point of D. Then,*

$$\vec{\nabla} \times \vec{\nabla} f = \vec{0}, \qquad \langle\vec{\nabla}, \vec{\nabla} \times \vec{g}\rangle_I = 0. \tag{8.21}$$

Important: el profe ha dit que si sobra temps farem càlcul amb altres bases de vectors. Demanar-li

Bibliography

[1] Hans Sagan. Advanced Calculus, 2014.