Physicist's

Mathematical Analysis

Handbook

Università degli studi di Roma "La Sapienza"
Physics and Astrophysics BSc

Matteo Cheri

A little handbook on Mathematical Analysis

January 14, 2020

Version 0.1

Physicist's
Mathematical Analysis
Handbook

A little handbook of Mathematical Analysis for Physics Undergraduates

Written by
Matteo Cheri
Università degli Studi di Roma "La Sapienza"
Physics and Astrophysics BSc

LaTeX 2_ε inside, ViM powered.

January 14, 2020

Version Fierce Carbonara

Contents

Notation

- $\mathbb{F}$ Ordered scalar field
- a_n, $(a)_n$ Sequence
- $((a)_k)_n$ Sequence of sequences
- $\mathbf{x}$, x^i Vector
- $\mathbf{x} \wedge \mathbf{y}$, $\epsilon^i_{jk} x^j y^k$ Cross product of $\mathbf{x}, \mathbf{y}$
- ∇f, $\partial_i f$, $\frac{\partial f}{\partial x^i}$ Gradient of $\mathbf{f}$
- $\langle \nabla, \mathbf{f} \rangle$, $\mathrm{div}(\mathbf{f})$, $\partial_i f^i$, $\frac{\partial f^i}{\partial x^i}$ Divergence of $\mathbf{f}$
- $\nabla \wedge \mathbf{f}$, $\mathrm{rot}(\mathbf{f})$, $\epsilon^i_{jk} \partial^j f^k$ Rotor of $\mathbf{f}$
- $\mathbf{Jf}(\mathbf{x})$, $\partial_j f^i$, $\frac{\partial f^i}{\partial x^j}$ Jacobian matrix of $\mathbf{f}$
- $\mathbf{H}_f(\mathbf{x})$, $\partial_i \partial_j f$, $\partial^2_{ij} f$, $\frac{\partial^2 f}{\partial x^i \partial x^j}$ Hessian matrix of the function f
- $\forall^\dagger$ For almost all (for all in a null measure subset)A

1 Complex Numbers and Functions

<++>

2 Abstract Spaces

§ 2.1 Metric Spaces

Definition 2.1.1 (Metric Space)**.** Let X be a non-empty set equipped with an application d, defined as follows

$$\begin{aligned} d\,: X \times X &\longrightarrow \mathbb{F} \\ (x, y) &\rightarrow d(x, y) \end{aligned} \tag{2.1}$$

Where $\mathbb{F}$ is an ordered field.
The couple (X, d) is said to be a *metric space*, if and only if $\forall x, y, z \in X$ the application d satisfies the following properties

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \implies x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, y) \leq d(x, z) + d(z, y)$

Definition 2.1.2 (Ball)**.** Let (X, d) be a metric space. We then define the *open ball of radius* r, centered in x in X (B_r^X), and the *closed ball of radius* r centered in x ($\overline{B_r^X}$) as follows

$$\begin{aligned} B_r^X(x) &:= \{u \in X |\, d(u, x) < r\} \\ \overline{B_r^X}(x) &:= \{u \in X |\, d(u, x) \leq r\} \end{aligned} \tag{2.2}$$

When there won't be doubts on on where the ball is defined, the superscript indicating the set of reference will be omitted.

We're now ready to define the *topology* on a metric space

Definition 2.1.3 (Open Set)**.** Let (X, d) be a metric space, and $A \subseteq X$ a subset. A is said to be an *open set* if and only if

$$\forall x \in X\ \exists B_r^X(x) \subset A \tag{2.3}$$

Definition 2.1.4 (Complementary Set)**.** Let A be a generic set, then the set A^{c} is defined as follows

$$A^{\mathrm{c}} := \{a \notin A\} \tag{2.4}$$

This set is said to be the *complementary set* of A.
It's also obvious that $A \cap A^{\mathrm{c}} = \{\}$

Definition 2.1.5 (Closed Set)**.** Alternatively to the notion of open set, we can say that $E \subseteq X$ is a *closed set*, if and only if

$$\forall x \in E^c \cap X \ \exists B_r^X(x) \subset E^c \cap X \tag{2.5}$$

Remark. A set isn't necessarily open nor closed!

Proposition 1.
1. The set $B_r^X(x)$ is open
2. The set $\overline{B_r^X}(x)$ is closed

Proof. Let $A = B_r^X(x)$. If A is open, we have therefore, applying the definition of open set, that

$$\forall x \in A \ \exists \epsilon > 0 \ : \ B_\epsilon^X(x) \subset A$$

So

$$x_0 \in A \implies d(x, x_0) < r$$
$$\therefore \epsilon = r - d(x, x_0) > 0$$

Then, by definition of open ball we have

$$y \in B_\epsilon^X(x) \implies d(x, y) < \epsilon$$

Then, we can say that

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \epsilon + d(x, x_0) = r$$
$$\therefore y \in B_\epsilon^X(x) \implies y \in B_r^X(x_0) \subset A$$

The demonstration of the second point is exactly the same, whereby we take E as our closed ball and $A = E^c$ $\square$

Proposition 2. Let (X, d)

1. The sets $\{\}, X$ are open
2. The sets $\{\}, X$ are closed
3. If $\{A_i\}_{i=1}^n$ is a collection of open sets, then $A = \bigcap_{i=1}^n A_i$ is open
4. If $\{C_i\}_{i=1}^n$ is a collection of closed sets, then $C = \bigcup_{i=1}^n C_i$ is closed
5. Let $I \subset \mathbb{N}$ be an index set, then
   (a) If $\{A_\alpha\}_{\alpha \in I}$ is a collection of open sets, then $B = \bigcup_{\alpha \in I} A_\alpha$ is open
   (b) If $\{C_\alpha\}_{\alpha \in I}$ is a collection of closed sets, then $D = \bigcap_{\alpha \in I} C_\alpha$ is closed

Proof. The first two statements are of easy proof. Let $B_\epsilon^X \subset \{\}$. This means that B_ϵ^X is empty and therefore $B_\epsilon^X = \{\}$, which makes it open by definition. Therefore we have that $\{\}^c = X$, and X must be closed, but if we reason a bit, we can say that $\forall x \in X \ B_\epsilon^X(x) \subset X$, which means that X is open, thus $X^c = \{\}$ must be closed.
Since we gave a proof for $\{\}$ and X being open, we have that these two sets are both open and

closed. These two sets are said to be *clopen*.
For the other statements we use the De Morgan laws on set calculus, therefore we have

$$x \in \bigcap_{i=1}^{n} A_i \implies x \in A_i$$
$$\therefore \exists \epsilon_i \, : \, B^X_{\epsilon_i}(x) \subset A_i$$

Taking $\epsilon = \min_{i \in I} \epsilon_i$ we have

$$B^X_\epsilon(x) \subset B^X_\epsilon(x) \implies B^X_\epsilon(x) \subset \bigcap_{i=1}^{n} A_i = A$$

And A is open
If we let $C = A^{\mathrm{c}}$ we have that

$$C = A^{\mathrm{c}} = \left(\bigcap_{i=1}^{n} A_i \right)^{\mathrm{c}} = \bigcup_{i=1}^{n} A_i^{\mathrm{c}}$$
$$\therefore C \text{ is closed}$$

For the last two we proceed as follows

$$x \in A_\alpha \implies \exists \alpha_0 \in I \, : \, x \in A_{\alpha_0}$$
$$\therefore \exists \epsilon > 0 \, : \, B^X_\epsilon(x) \subset A_{\alpha_0} \subset \bigcup_{\alpha \in I} = B$$

For the last one, we use the De Morgan laws and the proposition is demonstrated □

Definition 2.1.6 (Internal Points, Closure, Border)**.** Let (X, d) be a metric space and $A \subset X$ a subset.
We define the following sets from A

1. $A^\circ = \bigcup_{\alpha \in I} G_\alpha$ is the set of internal points of A, where I is an index set and $G_\alpha \subset A$ are open
2. $\overline{A} = \bigcap_{\beta \in J} F_\beta$ is the closure of A, where J is another index set and $F_\beta \subset A$ are closed
3. $\partial A = \overline{A} \setminus A^\circ = \overline{A} \cup (A^\circ)^{\mathrm{c}}$ is the border of A

Proposition 3. 1. A is an open set iff $A = A^\circ$

2. A is closed iff $A = \overline{A}$
3. $A^\circ = \overline{(A^\circ)}^{\mathrm{c}}$
4. $\overline{A} = \left[(A^{\mathrm{c}})^\circ \right]^{\mathrm{c}}$
5. $(A \cap B)^\circ = A^\circ \cap B^\circ$
6. $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Proof. Let $\mathcal{O}(A)$ be a collection of open sets, such that $\forall G \in \mathcal{O}(A) \implies G \subset A$, then

$$A = A^\circ \implies A = \bigcup_{G \in \mathcal{O}(A)} G$$

Therefore, being a union of a finite number of open sets, A is open.
For the same reason as before and the previous proposition, we have that $\overline{A}$ is closed
For the third proposition, we have

$$\left(\overline{A^c}\right)^c = \left(\bigcap_{A^c \subset F} F\right)^c = \bigcup_{A^c \subset F} F^c = \bigcup_{G \in \mathcal{O}(A)} G = A^\circ$$

The others are easily demonstrated throw this process, iteratively □

Proposition 4. Let (X, d) be a metric space, and $A \subset X$, $x \in X$

1. $x \in A \iff \exists \epsilon > 0 \, : \, B_\epsilon(x) \subset A$
2. $x \in \overline{A} \iff \forall \epsilon > 0 \; B_\epsilon(x) \cap A \neq \{\}$
3. $x \in \partial A \iff \forall \epsilon > 0 \; B_\epsilon(x) \cap A \neq \{\} \wedge B_\epsilon(x) \cap \overline{A} \neq \{\}$

Proof. $\boxed{1}$ Let $I(A) := \{x \in X | \, \exists \epsilon > 0 \, : \, B_\epsilon(x) \subset A$.

$$x \in I(A) \implies \exists \epsilon > 0 \, : \, B_\epsilon(x) \subset A, \; \therefore x \in \bigcup_{G \subset A} G$$

But

$$\begin{gathered} x \in A^\circ \implies \exists G \subset X \text{ open} \, : \, x \in G \implies \exists \epsilon > 0 \, : \, B_\epsilon(x) \subset G \subset A \\ \therefore A^\circ \subset I(A) \ni x, \; I(A) \subset A \text{ by definition}, \; \therefore I(A) = A^\circ \end{gathered}$$

$\boxed{2}$ For the second proposition, we have

$$\begin{aligned} &\overline{A} = \left[(A^c)^\circ\right]^c \implies x \in A \iff x \in (A^c)^\circ \implies \forall \epsilon > 0 \; B_\epsilon(x) \not\subset A^c \\ &\therefore \forall \epsilon > 0 \; B_\epsilon(x) \cap A \neq \{\} \end{aligned}$$

$\boxed{3}$ For the last one, we have, taking into account the first two proofs

$$\begin{aligned} &x \in \partial A \iff x \in \overline{A} \setminus A^\circ \implies x \in \overline{A} \wedge x \notin A^\circ \\ &\boxed{1} \wedge \boxed{2} \implies x \in \overline{A} \iff \forall \epsilon > 0 \; B_\epsilon(x) \cap A \neq \{\} \\ &\therefore x \notin A^\circ \iff \forall \epsilon > 0 \; B_\epsilon(x) \cap A^c \neq \{\} \end{aligned}$$

□

Definition 2.1.7 (Isometry)**.** Let $(X, d), (Y, \rho)$ be two metric spaces and f an application, defined as follows

$$f : (X, d) \to (Y, d)$$

f is said to be an *isometry* iff

$$\forall x_1, x_2 \in X, \; \rho(f(x_1), f(x_2)) = d(x_1, x_2)$$

Remark. If f is an isometry, then f is injective, but it's not necessarily surjective

Example 2.1.1. Let $X = [0,1]$ and $Y = [0,2]$, therefore

$$\begin{aligned} f :&[0,1] \to [0,2] \\ x &\to f(x) = x \end{aligned}$$

f is obviously an isometry, since, for $x, y \in [0,1]$

$$d(f(x), f(y)) = d(x, y)$$

But it's obviously not surjective.

Definition 2.1.8 (Diameter of a Set)**.** Let A be a set and the couple (A, d) be a metric space. We define the *diameter* of A as follows

$$\operatorname{diam}(A) = \sup_{x,y \in A} (d(x, y))$$

§ 2.2 Convergence and Compactness

Definition 2.2.1 (Convergence)**.** Let (X, d) be a metric space and $x \in X$. A sequence $(x_k)_{k \geq 0}$ in X is said to converge in X and it's indicated as $x_k \to x \in X$, iff

$$\forall \epsilon > 0 \; \exists N > 0 : \forall k \geq N, \; d(x_k, x) < \epsilon \;\; \therefore \lim_{k \to \infty} x_k = x$$

Theorem 2.1 (Unicity of the Limit)**.** *Let (X, d) be a metric space and $(x_k)_{k \geq 0}$ a sequence in X. If $x_k \to x \wedge x_k \to y$, then $x = y$*

Definition 2.2.2 (Adherent point)**.** Let (X, d) be a metric space and $A \subset X$. $x \in X$ is said to be an *adherent point* of A if $\exists (x_k)_{k \geq 0} \in A : x_k \to x \in X$. The set of all adherent points of A is called $\operatorname{ad}(A)$

Definition 2.2.3 (Accumulation point)**.** Let (X, d) be a metric space and $A \subset X$. $x \in X$ is an *accumulation point* of A, or also *limit point* of A if $\exists (x_k)_{k \geq 0} : x_k \neq x \wedge x_k \to x \in \operatorname{ad}(A)$

Proposition 5. Let (X, d) be a metric space and $A \subset X$, then $\overline{A} = \operatorname{ad}(A)$

Proof. Let $Y = \operatorname{ad}(A)$, then

$$\begin{aligned} &x \in \overline{A} \implies \forall \epsilon > 0 \; B_\epsilon(x) \cap A \neq \{\} \\ &\therefore \forall n \in \mathbb{N} \; B_{\frac{1}{n}}(x) \cap A \neq \{\} \implies \forall n \in \mathbb{N} \; \exists x_n \in B_{\frac{1}{n}}(x) \end{aligned}$$

But $d(x, x_n) < n^{-1}$, therefore $x \in Y \implies x \in \operatorname{ad}(A)$, and by definition

$$\begin{aligned} &\exists (x_n)_{n \geq 0} : \forall \epsilon > 0 \; \exists N \in \mathbb{N} : \forall k \geq N \; d(x_k, x) < \epsilon \implies x_N \in B_\epsilon(x) \;\; \therefore x_N \in A \\ &\therefore \forall \epsilon > 0 \; x_N \in B_\epsilon(x) \cap A \neq \{\} \implies x \in \overline{A} \implies Y \subset \overline{A}, \; \therefore Y = \operatorname{ad}(A) = \overline{A} \end{aligned}$$

□

Proposition 6. Let (X,d) be a metric space and $A \subset X$. Then A is closed iff $\exists (x_k)_{k\geq 0} \in A : x_k \to x \in \overline{A} \implies \mathrm{ad}(A) \subset A$

Definition 2.2.4 (Dense Set)**.** Let (X,d) be a metric space and $A, B \subset X$. A is said to be dense in B iff $B \subset \overline{A}$, therefore $\forall \epsilon > 0\ \exists y \in A\ :\ d(x,y) < 0$. One example for this is $\mathbb{Q} \subset \mathbb{R}$, with the usual euclidean distance defined through the modulus.

Definition 2.2.5. Let (X,d) be a metric space and $(x_k)_{k\geq 0} \in X$. The sequence x_k is said to be a *Cauchy sequence* iff

$$\forall \epsilon > 0\ \exists N > 0 : \forall k, n \geq N\ d(x_k, x_n) < \epsilon$$

Proposition 7. Let (X,d) be a metric space and $(x_k)_{k\geq 0} \in X$ a sequence. Then, if $x_k \to x$, x_k is a Cauchy sequence

Definition 2.2.6 (Complete Space)**.** Let (X,d) be a metric space. (X,d) is said to be *complete* iff $\forall (x_k)_{k\geq 0} \in X$ Cauchy sequences, we have $x_k \to x \in X$

Theorem 2.2 (Completeness)**.** *Let (X,d) be a metric space and $Y \subset X$. (Y,d) is complete iff $Y = \overline{Y}$ in X*

Proof. Let (Y,d) be a complete space, then

$$(x_k) \in Y \text{ Cauchy sequence } \implies \exists y \in Y\ :\ x_k \to y$$

Let $z \in \mathrm{ad}(A)$ and η_k a subsequence of x_k, then

$$\exists (\eta_k) \in Y\ :\ \eta_k \to z \implies \exists y \in Y\ :\ \eta_k \to y\ \therefore z = y \implies \mathrm{ad}(Y) \subset Y$$

Going the opposite way we have that $\mathrm{ad}(Y) = Y$ and therefore $Y = \overline{Y}$ □

Definition 2.2.7 (Compact Space)**.** A metric space (X,d) is said to be *compact* or *sequentially compact* if

$$\forall (x_k) \in X\ x_k \to x \in X, \exists (y_k) \text{ Subsequence } :\ y_k \to y \in X$$

Theorem 2.3. *Let (X,d) be a compact space. Then (X,d) is also complete*

Proof. (X,d) is compact, therefore

$$\forall (x_k) \in X \text{ Cauchy sequence } \implies x_k \to x \in X$$

Taken $(x_{n_k})_k \in X$ a subsequence, we have

$$x_k \to x \implies x_{n_k} \to x \in X$$

□

Definition 2.2.8 (Completely Bounded)**.** Let (X,d) be a metric space. X is *totally bounded* iff

$$\exists Y \subset X\ :\ \forall \epsilon > 0, \forall x \in Y\ X = \bigcup_{i=1}^{n} B_\epsilon(x)$$

Definition 2.2.9 (Poligonal Chain)**.** Let $z, w \in \mathbb{C}$. We define a *polygonal* $[z, w]$ as follows

$$[z, w] := \{\, z, w \in \mathbb{C} |\; z + t(w - z),\; t \in [0, 1] \subset \mathbb{R}\}$$

A *polygonal chain* will be indicated as follows $P_{z,w}$ and it's defined as follows

$$P_{z,w} = \bigcup_{k=1}^{n-1} [z_k, z_{k+1}] = [z, z_1, \cdots, z_{n-1}, w]$$

It can also be defined analoguously for every metric space $(X, d) \neq (\mathbb{C}, \|\cdot\|)$, where $\|\cdot\| : \mathbb{C} \to \mathbb{R}$ is the usual complex norm $\|z\| = \sqrt{z\overline{z}} = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$

Definition 2.2.10 (Connected Space)**.** Let (G, d) be a metric space, G is *connected* if

$$\forall z, w \in G\; \exists P_{z,w} \subset G$$

Definition 2.2.11 (Contraction Mapping)**.** Let (X, d) be a complete metric space. Let $T : X \longrightarrow X$. T is said to be a *contraction mapping* or *contractor* if

$$\forall x, y \in X\; \exists q \in [0, 1) \,:\, d(T(x), T(y)) \leq q d(x, y) \tag{2.6}$$

Note that a contractor is necessarily continuous.

Theorem 2.4 (Banach Fixed Point)**.** *Let (X, d) be a complete metric space, with $X \neq \{\}$ and equipped with a contractor $T : X \longrightarrow X$. Then*

$$\exists! x^\star \in X \,:\, T(x^\star) = x^\star \tag{2.7}$$

Proof. Take $x_0 \in X$ and a sequence $x_n : \mathbb{N} \longrightarrow X$, where

$$x_n = T(x_{n-1}), \quad \forall n \in \mathbb{N}$$

It's obvious that

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq q d(x_n, x_{n-1}) \leq q^n d(x_1, x_0)$$

We need to prove that x_n is a Cauchy sequence. Let $m, n \in \mathbb{N} \,:\, m > n$, then

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + \cdots + d(x_{n+1}, x_n) \leq q^{m-1} d(x_1, x_0) + \cdots + q^n d(x_1, x_0)$$

Regrouping, we have

$$d(x_m, x_n) \leq q^n d(x_1, x_0) \sum_{k=0}^{m-n-1} q^k \leq q^n d(x_1, x_0) \sum_{k=0}^{\infty} q^k = q^n d(x_1, x_0) \left(\frac{1}{1-q}\right)$$

By definition of convergence, we have then

$$\forall \epsilon > 0,\; \exists N \in \mathbb{N} \,:\, \forall n > N\; d(s_n, s) < \epsilon$$

Then

$$\frac{q^n d(x_1, x_0)}{1-q} < \epsilon \implies q^n < \frac{\epsilon(1-q)}{d(x_1, x_0)}, \quad \forall n > N$$

Therefore, after taking $m > n > N$, we have

$$d(x_m, x_n) < \epsilon$$

Therefore x_n is a Cauchy sequence. Since (X, d) is a complete metric space, this sequence must have a limit $x_n \to x^\star \in X$, but, by definition of convergence and limit, we have that by continuity

$$x^\star = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T\left(\lim_{n\to\infty} x_{n-1}\right) = T(x^\star)$$

This point is unique. Take $y^\star \in X$ such that $T(y^\star) = y^\star \neq x^\star$, then

$$0 < d(T(x^\star), T(y^\star)) = d(x^\star, y^\star) > qd(x^\star, y^\star) \quad \lightning$$

Therefore

$$\exists! x^\star \in X \,:\, T(x^\star) = x^\star$$

And $x^\star$ is the fixed point of the contractor T □

§ 2.3 Vector Spaces

Definition 2.3.1 (Vector Space)**.** A vector space $\mathcal{V}$ over a field $\mathbb{F}$ is a set, where $\mathcal{V} \neq \{\}$ and it satisfies the following properties, $\forall u, v, w \in \mathcal{V}$ and $a, b \in \mathbb{F}$

1. $u + v \in \mathcal{V}$ sum closure
2. $av \in \mathcal{V}$ scalar closure
3. $u + v = v + u$
4. $(u + v) + w = u + (v + w)$
5. $\exists! 0 \in \mathcal{V} \,:\, u + 0 = 0 + u = u$
6. $\exists! v \in \mathcal{V} \,:\, u + v = 0 \implies v = -u$
7. $\exists! 1 \in \mathcal{V} \,:\, 1 \cdot u = u$
8. $(ab)u = a(bu) = b(au) = abu$
9. $(a + b)u = au + bu$
10. $a(u + v) = au + av$

Definition 2.3.2 (Norm)**.** Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$, then the *norm* is an application defined as follows

$$\|\cdot\| : \mathcal{V} \longrightarrow \mathbb{F}$$

Where it satisfies the following properties

1. $\|u\| \geq 0 \; \forall u \in \mathcal{V}$
2. $\|u\| = 0 \iff u = 0$

3. $\|cu\| = |c|\|u\| \; \forall u \in \mathcal{V} \; c \in \mathbb{F}$
4. $\|u + v\| \leq \|u\| + \|v\| \; \forall u, v \in \mathcal{V}$

Definition 2.3.3 (Normed Vector Space)**.** A *normed vector space* is defined as a couple $(\mathcal{V}, \|\cdot\|)$, where $\mathcal{V}$ is a vector space over a field $\mathbb{F}$.

Proposition 8. A normed vector space (NVS), is also a metric vector space (MVS) if we define our distance as follows

$$d(u, v) = \|u - v\| \; \forall u, v \in \mathcal{V}$$

Definition 2.3.4 (Vector Subspace)**.** Let $\mathcal{V}$ be a vector space and $\mathcal{U} \subset \mathcal{V}$. $\mathcal{U}$ is a *vector subspace* of $\mathcal{V}$ iff

1. $u, v \in \mathcal{U} \implies u + v \in \mathcal{U}$
2. $u \in \mathcal{U}, \; a \in \mathbb{F} \implies au \in \mathcal{U}$

Proposition 9. If $(\mathcal{V}, \|\cdot\|)$ is an normed vector space and $\mathcal{W} \subset \mathcal{V}$ is a subspace of $\mathcal{V}$, then $(\mathcal{W}, \|\cdot\|)$ is a normed vector space

Definition 2.3.5 (p-norm)**.** Let $(\mathcal{V}, \|\cdot\|_p)$ be a normed vector space. The norm $\|\cdot\|_p$ is said to be a *p-norm* if it's defined as follows

$$\|v\|_p := \left(\sum_{i=1}^{\dim(\mathcal{V})} (v_i)^p \right)^{\frac{1}{p}}, \; \forall v \in \mathcal{V}, \; \forall p \in \mathbb{N}^\star := \mathbb{N} \cup \{\pm\infty\} \tag{2.8}$$

Setting $p = \infty$ we have that

$$\|v\|_\infty = \max_{i \leq \dim(\mathcal{V})} |v_i| \tag{2.9}$$

Definition 2.3.6 (Dual Space)**.** Let $\mathcal{V}$ be a vector space over the field $\mathbb{F}$, we define a *linear functional* as an application $\varphi : \mathcal{V} \longrightarrow \mathbb{F}$ such that $\forall u, v \in \mathcal{V}$ and $c \in \mathbb{F}$

$$\begin{aligned} \varphi(u + v) &= \gamma(u) + \varphi(v) \\ \varphi(\lambda u) &= \lambda \varphi(u) \end{aligned} \tag{2.10}$$

Defining the sum of two linear functionals as $(\varphi_1 + \varphi_2)(v) = \varphi_1(v) + \varphi_2(v)$ we immediately see that the set of all linear functionals forms a vector space over $\mathcal{V}$, which will be called the *dual space* $\mathcal{V}^\star$.

§§ 2.3.1 Hölder and Minkowski Inequalities

Having defined p-norms, we can prove two inequalities that work with these norms, the *Minkowski inequality* and the *Hölder Inequality*

Theorem 2.5 (Hölder Inequality)**.** *Let $p_q \in \mathbb{N}^\star$, where*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then

$$\forall x, y \in \mathbb{R}^n \; \|x\|_p \|y\|_q \geq \sum_{k=1}^{n} |x_k y_k| \tag{2.11}$$

Proof. Taking $p = 1$, we have $q = \infty$, and the demonstration is obvious

$$\|x\|_p \|y\|_q = \|x\|_1 \|p\|_\infty = \max_{k \le n} |y_k| \sum_{k=1}^{n} |x_k| \ge \sum_{k=1}^{n} |x_k y_k|$$

Else, if $p > 1$, we ave that

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \ \forall a, b \ge 0$$

Let

$$s = \frac{x}{\|x\|_p}, \ t = \frac{y}{\|y\|_q}$$

We have

$$\sum_{k=1}^{n} \|s\|^p = \frac{1}{\|x\|_p^p} \sum_{k=1}^{n} |x_k|^p = 1 = \sum_{k=1}^{n} |t|^q = \frac{1}{\|y\|_q^q} \sum_{k=1}^{n} |y|^p$$

Therefore

$$\sum_{k=1}^{n} |s_k t_k| \le \frac{1}{p} \sum_{k=1}^{n} |s_k|^p + \frac{1}{q} \sum_{k=1}^{n} |t_k|^q$$

Substituting again the definitions of s, t we have

$$\sum_{i=1}^{n} |y_k x_k| = \|x\|_p \|y\|_q \sum_{k=1}^{n} |s_k t_k| \le \|x\|_p \|y\|_q$$

$\square$

Theorem 2.6 (Minkowski Inequality)**.** *Let $p \ge 1$, therefore $\forall x, y \in \mathbb{R}^n$ we have*

$$\|x + y\|_p \le \|x\|_p + \|y\|_p \tag{2.12}$$

Proof. We begin by writing explicitly the p-norm

$$\|x + y\|_p^p = \sum_{k=1}^{n} (|x_k| + |y_k|)^p = \sum_{k=1}^{n} (|x_k| + |y|_k) (|x_k| + |y_k|)^{p-1}$$

Letting $u_k = (|x_k| + |y_k|)^{p-1}$ we have, after imposing the condition on q of the p-norm as $q(p+1) = p$ and using that the sum is Abelian, we have

$$\begin{cases} \sum_{k=1}^{n} |x_k| u_k \le \|x\|_p \|u\|_q = \|x\|_p \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{1}{q}} \\ \sum_{k=1}^{n} |y_k| u_k \le \|y\|_p \|u\|_q = \|y\|_p \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p \right)^{\frac{1}{q}} \end{cases}$$

Therefore, summing and imposing that $1 - q^{-1} = p$ we have that

$$\|x + y\|_p \le \|x\|_p + \|y\|_q$$

$\square$

3 Sequences and Series of Functions

§ 3.1 Sequences of Functions

Definition 3.1.1 (Sequence of Functions)**.** Let S be a set and (X,d) a metric space, a *sequence of functions* is defined as follows

$$\begin{aligned} f_n :& S \longrightarrow (X,d) \\ & s \to f_n(s) \end{aligned} \tag{3.1}$$

Where, $\forall n \in \mathbb{N}$ a function $f_{(n)} : S \longrightarrow (X,d)$ is defined

Definition 3.1.2 (Pointwise Convergence)**.** A sequence of functions $(f_n)_{n\geq 0}$ is said to converge pointwise to a function $f : S \longrightarrow (X,d)$, and it's indicated as $f_n \to f$, if

$$\forall \epsilon > 0,\ \forall x \in S\ \exists N_\epsilon(x) \in \mathbb{N} : d(f_n(x), f(x)) < \epsilon\ \forall n \geq N_\epsilon(x) \tag{3.2}$$

It can be indicated also as follows

$$\lim_{n\to\infty} (f_n(x)) = f(x) \tag{3.3}$$

Definition 3.1.3 (Uniform Convergence)**.** Defining an $\|\cdot\|_\infty = \sup_{i\leq n} |\cdot|$ we have that the convergence of a sequence of functions is uniform, and it's indicated as $f_n \rightrightarrows f$, iff

$$\forall \epsilon > 0\ \exists N_\epsilon \in \mathbb{N} : d(f_n(x), f(x)) < \epsilon\ \forall n \geq N_\epsilon\ \forall x \in S \tag{3.4}$$

Or, using the norm $\|\cdot\|_\infty$

$$\forall \epsilon > 0\ \exists N_\epsilon \in \mathbb{N} : \|f_n - f\|_\infty < \epsilon \tag{3.5}$$

Theorem 3.1 (Continuity of Uniformly Convergent Sequences)**.** *Let $(f_n)_{n\geq 0} : (S, d_S) \longrightarrow (X,d)$ be a sequence of continuous functions. Then if $f_n \rightrightarrows f$, we have that $f \in C(S)$, where $C(S)$ is the space of continuous functions*

Proof.

$$\begin{aligned} &\forall x \in S,\ \exists \epsilon > 0 : f_n \rightrightarrows f,\ \therefore \forall n \geq N_\epsilon \in \mathbb{N} : d(f_n(x), f(x)) < \frac{\epsilon}{3} \\ &f_n \in C(S) \implies \exists \delta_\epsilon > 0 : d(f_n(x), f_n(y)) < \frac{\epsilon}{3},\ \forall x, y \in S : d_S(x,y) < \delta \\ &\therefore d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) < \epsilon \iff d_S(x,y) < \delta_\epsilon \end{aligned} \tag{3.6}$$

□

Theorem 3.2 (Integration of Sequences of Functions)**.** *Let* $(f_n)_{n\geq 0}$ *be a sequence of functions such that* $f_n \rightrightarrows f$ *Then we can define the following equality*

$$\lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x \tag{3.7}$$

Proof. We already know that in the closed set $[a,b]$ we can say, since $f_n \rightrightarrows f$, that

$$\forall\epsilon > 0\ \exists N_\epsilon \in \mathbb{N} : \forall n \geq N_\epsilon\ \|f_n - f\|_\infty < \frac{\epsilon}{b-a} \tag{3.8}$$

Then, we have that

$$\forall n \geq N_\epsilon\ \left|\int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x\right| \leq \|f_n - f\|_\infty (b-a) < \epsilon \tag{3.9}$$

□

Theorem 3.3 (Differentiation of a Sequence of Functions)**.** *Define a sequence of functions as* $f_n : I \longrightarrow \mathbb{R}$*, with* $f_n(x) \in C^1(I)$*. If*

1. $\exists x_0 \in I : f_n(x_0) \to l$
2. $f_n' \rightrightarrows g\ \forall x \in I$

Then

$$f_n(x) \rightrightarrows f \implies \forall x \in I,\ f'(x) = \lim_{n\to\infty} f_n'(x) = g(x) \tag{3.10}$$

Proof. For the fundamental theorem of integral calculus, we can write, using the regularity of the $f_n(x)$ that

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f_n(t)\,\mathrm{d}t$$

Taking the limit we have

$$\begin{gathered}\lim_{n\to\infty} f_n(x) = l + \int_{x_0}^x g(t)\,\mathrm{d}t = f(x)\\ \therefore f'(t) = g(t)\end{gathered}$$

But, we also have that

$$\begin{gathered}\forall\epsilon > 0\ \|f_n' - f'\|_\infty \leq |f_n(x_0) - l| + \|f_n' - g\|_\infty (b-a) < \epsilon\\ \therefore f_n \rightrightarrows f,\ f_n' \rightrightarrows f'\end{gathered}$$

□

§ 3.2 Series of Functions

Let now, for the rest of the section, $(X, d) = \mathbb{C}$.

Definition 3.2.1 (Series of Functions)**.** Let $(f_n)_{n\geq 0} \in \mathbb{C}$ be a sequence of functions, such that $f_n : S \to \mathbb{C}$. We can define the *series of functions* as follows

$$s_n(x) = \sum_{k=1}^{n} f_k(x) \tag{3.11}$$

Definition 3.2.2 (Convergent Series)**.** A series of functions $s_n(x) : S \to \mathbb{C}$ is said to be *convergent* or *pointwise convergent* if

$$s_n(x) = \sum_{k=0}^{n} f_k(x) \longrightarrow s(x) \tag{3.12}$$

Where $s(x) : S \to \mathbb{C}$ is the *sum* of the series.
This means that

$$\forall x \in S, \ \lim_{k\to\infty} s_k(x) = \sum_{k=0}^{\infty} f_k(x) = s(x) \tag{3.13}$$

Theorem 3.4. *Necessary Condition for the convergence of a series of functions:*
Let $(f_n) \in \mathbb{C}$ be a succession, then the series $s_n(x)$ defined as follows, converges to the function $s(x)$

$$s_n(x) = \sum_{k=0}^{n} f_k(x) = s(x) = \sum_{k=0}^{\infty} f_k(x)$$

Proof.

$$\forall x \in S \ \lim_{k\to\infty} f_k(x) = \lim_{n\to\infty} (s_n(x) - s_{n+1}(x)) = 0$$

□

Definition 3.2.3 (Uniform Convergence)**.** A series of functions is said to be *uniformly convergent* if and only if

$$\sum_{k=0}^{\infty} f_k(x) \rightrightarrows s(x) \iff s_n(x) = \sum_{k=0}^{n} f_k(x) \rightrightarrows s(x) \tag{3.14}$$

Definition 3.2.4 (Absolute Convergence)**.** A series of functions is said to be *absolutely convergent* if and only if

$$\sum_{k=0}^{\infty} f_k(x) \overset{\mathrm{A}}{\to} s(x) \implies \sum_{k=0}^{\infty} |f_k(x)| \to s(x) \tag{3.15}$$

Theorem 3.5. *Let $\sum_{k=0}^{\infty} f_k(x) \overset{\mathrm{A}}{\to} s(x)$, then*

$$\sum_{k=0}^{\infty} f_k(x) \overset{\mathrm{A}}{\to} s(x) \implies \sum_{k=0}^{\infty} f_k(x) \to s(x) \tag{3.16}$$

Proof. Let

$$\begin{aligned}
s_n(x) = \sum_{k=0}^{n} f_k(x) \ \therefore \exists g(x) : (S,d) \longrightarrow \mathbb{C},\ \exists N_\epsilon(x) \in \mathbb{N} : \left| g(x) - \sum_{k=0}^{\infty} f_k(x) \right| = \\
= \sum_{k=n+1}^{\infty} |f_k(x)| < \epsilon \ \forall n \geq N_\epsilon(x) \\
\therefore \forall n, m \in \mathbb{N}, m > n \\
|s_m(x) - s_n(x)| = \left| \sum_{k=n+1}^{m} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| < \epsilon \ \forall x \in S \\
\therefore (s_n(x)) \text{ is a Cauchy series in } \mathbb{C} \implies s_k(x) \to s(x)
\end{aligned}$$

□

Definition 3.2.5 (Total Convergence)**.** A series of functions $s_k(x)$ is said to be *totally convergent* if

1. $\exists M_k \ : \sup_S |f_k(x)| \leq M_k \ \forall k \geq 1$
2. $\sum_{k=0}^{\infty} M_k \to M$

The total convergence is then indicated as $s_k(x) \xrightarrow{\text{T}} s(x)$

Proposition 10. Let

$$s_n(x) = \sum_{k=0}^{n} f_n(x)$$

Then

1. $f_n(x) \in C(S) \land s_k(x) \rightrightarrows s(x) \implies s(x) \in C(S)$
2. $f_n(x) \in C(S),\ s_k(x) \rightrightarrows s(x) \implies \int s(x)\,\mathrm{d}x = \lim_{k\to\infty} \int s_k(x)\,\mathrm{d}x$
3. $s_k(x) \xrightarrow{\text{A}} s(x) \implies s_k(x) \to s(x)$
4. $s_k(x) \rightrightarrows s(x) \implies s_k(x) \xrightarrow{\text{A}} s(x)$
5. $s_k(x) \xrightarrow{\text{T}} s(x) \implies s_k(x) \rightrightarrows s(x)$

§§ 3.2.1 Power Series and Convergence Tests

Theorem 3.6 (Weierstrass Test)**.** *Let* $(f_n) : (S,d) \to \mathbb{C}$ *a sequence of functions.*
If we have that

$$\begin{gathered}
\forall n > N_\epsilon \in \mathbb{N} \ \exists M_n > 0 \ : \ |f_n(x)| \leq M_n \\
\therefore \forall x \in S \ \sum_{k=0}^{n} f_k(x) \leq \sum_{k=1}^{\infty} M_k \to M \therefore \sum_{k=0} f_k(x)^n \rightrightarrows s(x)
\end{gathered}$$

Definition 3.2.6 (Power Series)**.** Let $z, z_0, (a_n) \in \mathbb{C}$. A *power series centered in* z_0 is defined as follows

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k \tag{3.17}$$

Example 3.2.1. Take the *geometric series*. This is the best example of a power series centered in $z_0 = 0$, and it has the following form

$$\sum_{k=0}^{\infty} z^k \tag{3.18}$$

We can expand it as follows

$$\sum_{k=0}^{m} z^k = (1 - z)\left(1 + z + z^2 + \cdots + z^m\right) = 1 - z^{n+1} = \frac{1 + z^{n+1}}{1 - z} \ \forall |z| \neq 1 \tag{3.19}$$

Taking the limit, we have, therefore

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1 - z} \ \forall |z| < 1 \tag{3.20}$$

Theorem 3.7 (Cauchy-Hadamard Criteria)**.** *Let* $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ *be a power series, with* $a_n, z, z_0 \in \mathbb{C}$*. We define the Radius of convergence* $R \in \mathbb{R}^\star = \mathbb{R} \cup \{\pm\infty\}$*, with the Cauchy-Hadamard criteria*

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \begin{cases} +\infty & \dfrac{1}{R} = 0 \\ l & 0 < \dfrac{1}{R} = l < \infty \\ 0 & \dfrac{1}{R} = +\infty \end{cases} \tag{3.21}$$

Then $s_k(z) \rightrightarrows s(z) \ \forall |z| \in (-R, R)$

Theorem 3.8 (D'Alambert Criteria)**.** *From the power series we have defined before, we can write the D'Alambert criteria for convergence as follows*

$$\frac{1}{R} = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| \implies R = \lim_{k\to\infty} \left| \frac{a_k}{a_{k+1}} \right| \tag{3.22}$$

Where R *is the previously defined radius of convergence*

Theorem 3.9 (Abel)**.** *Let* $R > 0$*, then if a power series converges for* $|z| = R$*, it converges uniformly* $\forall |z| \in [r, R] \subset (-R, R]$*. It is valid analogously for* $x = -R$

Remark (Power Series Integration). If the series has $R > 0$ and it converges in $|z| = R$, calling $s(x)$ the sum of the series, with $x = |z|$ we can say that

$$\int_0^R s(x)\,\mathrm{d}x = \sum_{k=0}^{\infty} \int_0^R a_k x^k \,\mathrm{d}x = \int_0^R \sum_{k=1}^{\infty} a_k x^k \,\mathrm{d}z = \sum_{k=0}^{\infty} a_k \frac{R^{k+1}}{k+1} \tag{3.23}$$

Remark (Power Series Derivation). If Abel's theorem holds, we have also that, if we have $s(x)$ our power series sum, we can define the $n-$th derivative of this series as follows

$$\frac{\mathrm{d}^n s}{\mathrm{d}x^n} = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1) a_k x^{k-n} \tag{3.24}$$

4 Infinite Dimensional Spaces

§ 4.1 Sequence Spaces

Notation (The Field $\mathbb{F}$)**.** Here in this section, the field $\mathbb{F}$ should be intended as either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$

Definition 4.1.1 (Sequence Space)**.** As a n-tuple in the field $\mathbb{F}^n$ can be seen as a sequence, as follows

$$x \in \mathbb{F}^n,\ x = (x_1, x_2, \cdots, x_n) = (x_k)_{k=1}^n$$

We can imagine a sequence as a point in a space. We will call this space $\mathbb{F}^{\mathbb{N}}$, and an element of this space will be indicated as follows

$$x \in \mathbb{F}^{\mathbb{N}},\ x = (x)_n = (x_1, x_2, \cdots, x_n, \cdots) = (x_k)_{k=1}^{\infty}$$

Therefore, every point in $\mathbb{F}^{\mathbb{N}}$ is a sequence. Note that the infinite sequence of 0s and 1s will be indicated as $0 = (0)_n,\ 1 = (1)_n$

Definition 4.1.2 (Sequence of Sequences)**.** We can see a sequence of sequences as a mapping from $\mathbb{N}$ to the space $\mathbb{F}^{\mathbb{N}}$, as follows

$$\begin{aligned} x :& \mathbb{N} \longrightarrow \mathbb{F}^{\mathbb{N}} \\ & n \to ((x)_k)_n \end{aligned}$$

It's important to note how there are two indexes, since every element of the sequence is a sequence in itself (i.e. $((x)_k)_n \in \mathbb{F}^{\mathbb{N}}$ for any fixed $n \in \mathbb{N}$)

Definition 4.1.3 (Convergence of a Sequence of Sequences)**.** A sequence of sequences is said to converge to a sequence in $\mathbb{F}^{\mathbb{N}}$ if and only if

$$\lim_{n\to\infty} \|(x)_k - ((x)_k)_n\| = 0 \tag{4.1}$$

For some norm $\|\cdot\|$

Definition 4.1.4 (Pointwise Convergence)**.** A sequence of sequence is said to converge *pointwise* to a sequence in $\mathbb{F}^{\mathbb{N}}$ if and only if

$$\forall k \in \mathbb{N}\ \lim_{n\to\infty} ((x)_k)_n = (x)_k \tag{4.2}$$

And it's indicated as $((x)_k)_n \to (x)_k$

Example 4.1.1. Take the following sequence of sequences in $\mathbb{F}^{\mathbb{N}}$

$$\left((x)_k\right)_n = \frac{k}{n}$$

This sequence converges pointwise to the null sequence, since

$$\lim_{n\to\infty} \left((x)_k\right)_n = \lim_{n\to\infty} \frac{k}{n} = (0)_k$$

§§ 4.1.1 Space of Bounded Sequences

Definition 4.1.5 (Limited Sequence Space)**.** Let $(x)_k \in \mathbb{F}^{\mathbb{N}}$. Calling the space of bounded sequences as $\ell^\infty(\mathbb{F})$, we have that $(x)_k \in \ell^\infty(\mathbb{F})$ if and only if

$$\sup_{n\in\mathbb{N}} |(x)_n| = M \in \mathbb{F} \tag{4.3}$$

Therefore, this space is defined as follows

$$\ell^\infty(\mathbb{F}) := \left\{ (x)_n \in \mathbb{F}^{\mathbb{N}} \middle| \sup_{n\in\mathbb{N}} |(x)_n| < M,\ M \in \mathbb{F} \right\} \tag{4.4}$$

Theorem 4.1. *The application* $\|\cdot\|_\infty = \sup_{n\in\mathbb{N}} |\cdot|$ *is a norm in* $\ell^\infty(\mathbb{F})$

Proof. 1) $\|(x)_n\| \geq 0\ \forall (x)_n \in \mathbb{F}^{\mathbb{N}},\ \|(x)_n\|_\infty = 0 \iff (x)_n = (0)_n$, by definition of sup the first statement is obvious, meanwhile for the second

$$0 \leq |(x)_n| \leq \sup_{n\in\mathbb{N}} |(x)_n| = 0 \implies |(x)_n| = 0 \ \therefore (x)_n = (0)_n$$

2) $\|c(x)_n\|_\infty = |c| \|(x)_n\|_\infty$

$$\|c(x)_n\|_\infty = \sup_{n\in\mathbb{N}} |c(x)_n| = \sup_{n\in\mathbb{N}} |c||(x)_n| = |c| \sup_{n\in\mathbb{N}} |(x)_n| = |c| \|(x)_n\|_\infty$$

3) $\|(x)_n + (y)_n\|_\infty \leq \|(x)_n\|_\infty + \|(y)_n\|_\infty$

$$\sup_{n\in\mathbb{N}} |(x)_n + (y)_n| \leq \sup_{n\in\mathbb{N}} \left(|(x)_n| + |(y)_n|\right) = \sup_{n\in\mathbb{N}} |(x)_n| + \sup_{n\in\mathbb{N}} |(y)_n| = \|(x)_n\|_\infty + \|(y)_n\|_\infty$$

Since $\ell^\infty(\mathbb{F})$ is a vector space, the couple $(\ell^\infty(\mathbb{F}), \|\cdot\|_\infty)$ is a normed vector space ☐

Remark. Let $\mathcal{V}$ be a vector space over some field $\mathbb{F}$. If $\dim(\mathcal{V}) = \infty$, a closed and bounded subset $\mathcal{W} \subset \mathcal{V}$ isn't necessarily compact, whereas, a compact subset $\mathcal{Z} \subset \mathcal{V}$ is necessarily closed and bounded.

Example 4.1.2. Take $\mathcal{V} = \ell^\infty(\mathbb{F})$ and $\mathcal{W} = \overline{B_1((0)_n)}$, where

$$\overline{B_1((0)_n)} := \left\{ (x)_n \in \mathbb{F}^\infty \middle| \|(x)_n\|_\infty \leq 1 \right\}$$

We have that $\operatorname{diam}(\overline{B_1}) = 2$, therefore this set is bounded and closed by definition.
Take the *canonical sequence of sequences* $\left((e)_k\right)_n$, defined as follows:

$$\left((e)_k\right)_n = ((0)_k, (0)_k, \cdots, (0)_k, (1)_k, (0)_k, \cdots), \text{ for some } k \in \mathbb{N}$$

Therefore, $\forall n \neq m$

$$\left\| \left((e)_k\right)_n - \left((e)_k\right)_m \right\|_\infty = \|(1)_k\|_\infty = 1$$

Therefore there aren't converging subsequences, and therefore $\overline{B_1}$ can't be compact.

§§ 4.1.2 Space of Sequences Converging to 0

Definition 4.1.6 (Space of Sequences Converging to 0)**.** The space of sequences converging to 0 is indicated as $\ell_0(\mathbb{F})$ and is defined as follows

$$\ell_0(\mathbb{F}) := \left\{ (x)_n \in \mathbb{F}^{\mathbb{N}} \middle| \, (x)_n \to 0 \right\} \tag{4.5}$$

Proposition 11. $\ell_0(\mathbb{F}) \subset \ell^\infty(\mathbb{F})$, and the couple $(\ell_0(\mathbb{F}), \|\cdot\|_\infty)$ is a normed vector space, where the norm $\|\cdot\|_\infty$ gets induced from the space $\ell^\infty(\mathbb{F})$

Proof.

$$\lim_{k\to\infty} (x)_k = 0 \implies \forall \epsilon > 0 \; \exists N \in \mathbb{N} \, : \, |(x)_n| < \epsilon \; \forall n \geq N$$
$$\therefore \sup_{n\in\mathbb{N}} |(x)_n| = \epsilon \leq M \in \mathbb{F} \implies (x)_n \in \ell^\infty(\mathbb{F}), \; \therefore \ell_0(\mathbb{F}) \subset \ell^\infty(\mathbb{F})$$

□

§§ 4.1.3 $\ell^p(\mathbb{F})$ Spaces

Definition 4.1.7. The sequence space $\ell^p(\mathbb{F})$ is defined as follows

$$\ell^p(\mathbb{F}) := \left\{ (x)_n \in \mathbb{F}^{\mathbb{N}} \middle| \, \|(x)_n\|_p^p = M \in \mathbb{F} \right\} \tag{4.6}$$

Where $\|\cdot\|_p$ is the usual $p-$norm

Proposition 12. The application $\|\cdot\|_p : \ell^p(\mathbb{F}) \longrightarrow \mathbb{F}$ is a norm in $\ell^p(\mathbb{F})$, and the couple $(\ell^p(\mathbb{F}), \|\cdot\|_p)$ is a normed vector space

Proof. We begin by proving that $\ell^p(\mathbb{F})$ is actually a vector space, therefore 1) $\forall (x)_n, (y)_n \in \ell^p(\mathbb{F}), \; (x)_n + (y)_n = (x+y)_n \in \ell^p(\mathbb{F})$

$$(x+y)_n \in \ell^p(\mathbb{F}) \implies \sum_{n=0}^{\infty} |(x)_n + (y)_n|^p = \|(x)_n + (y)_n\|_p^p < M \in \mathbb{F}$$
$$\|(x)_n + (y)_n\|_p^p \leq \|(x)_n\|_p^p + \|(y)_n\|_p^p < M \in \mathbb{F}$$

2) $\forall (x)_n \in \ell^p(\mathbb{F}), \; c \in \mathbb{F}, \; c(x)_n \in \ell^p(\mathbb{F})$

$$c(x)_n \in \ell^p(\mathbb{F}) \implies \|c(x)_n\|_p^p < M \in \mathbb{F}$$
$$\|c(x)_n\|_p^p = \sum_{n=0}^{\infty} |c(x)_n|^p = |c|^p \sum_{n=0}^{\infty} |(x)_n|^p = |c|^p \|(x)_n\|_p^p < M \in \mathbb{F}$$

□

Remark. $(x)_n \in \ell^p(\mathbb{F}) \implies (x)_n \in \ell_0(\mathbb{F})$.

Proof. The proof is simple, taking $(y)_n = |(x)_n|^p$, we can see that $(y)_n \to 0$, therefore $(x)_n \to 0$ and $(x)_n \in \ell_0(\mathbb{F})$ □

§§ 4.1.4 Space of Finite Sequences

Definition 4.1.8 (Space of Finite Sequences)**.** The space of finite sequences is indicated as $\ell_f(\mathbb{F})$ and it's defined as follows

$$\ell_f(\mathbb{F}) := \{ (x)_n \in \mathbb{F}^{\mathbb{N}} \big| \, (x)_n = 0 \; \forall n > N \in \mathbb{N} \} \tag{4.7}$$

It's already obvious that $\ell_f(\mathbb{F}) \subset \ell^p(\mathbb{F}) \subset \ell^q(\mathbb{F}) \subset \ell_0(\mathbb{F}) \subset \ell^\infty(\mathbb{F})$, where $p < q \in \mathbb{R}^+ \setminus \{0\}$ where $p < q \in \mathbb{R}^+ \setminus \{0\}$

§ 4.2 Function Spaces

Notation. In this case, when there will be written the field $\mathbb{F}$, we might either mean $\mathbb{R}$ only, i.e. functions $\mathbb{R} \longrightarrow \mathbb{R}$, or $\mathbb{R}; \mathbb{C}$, i.e. functions $\mathbb{R} \longrightarrow \mathbb{C}$.

Definition 4.2.1 (Some Function Spaces)**.** We are already familiar from the basic courses in one dimensional real analysis, about the space of continuous functions $C(A)$, where $A \subset \mathbb{R}$. We can define three other spaces directly, adding some restrictions.

1. $C_b(\mathbb{F}) := \{ f \in C(\mathbb{F}) | \sup_{x \in \mathbb{F}} (f(x)) \leq M \in \mathbb{F} \}$
2. $C_0(\mathbb{F}) := \{ f \in C(\mathbb{F}) | \lim_{x \to \infty} (f(x)) = 0 \}$
3. $C_c(\mathbb{F}) := \{ f \in C(\mathbb{F}) | \, f(x) = 0 \quad \forall x \in A^c \subset \mathbb{F} \}$ i.e. $C_c(\mathbb{F}) := \{ f \in C(\mathbb{F}) | \operatorname{supp}(f) \text{ is compact } \}$, where with supp we indicate the following set $\operatorname{supp}_{\mathbb{F}}(f) := \overline{\{ x \in \mathbb{F} | \, f(x) \neq 0 \}}$

Due to the properties of continuous functions, these spaces are obviously vector spaces.

Proposition 13. We have $C_c(\mathbb{F}) \subset C_0(\mathbb{F}) \subset C_b(\mathbb{F}) \subset C(\mathbb{F})$, the application

$$\|f\|_u = \|f\|_\infty = \sup_{x \in A} |f(x)| \tag{4.8}$$

Is a norm in $C(A)$, whereas

$$\|f\|_u = \|f\|_\infty = \sup_{x \in \mathbb{F}} |f(x)| \tag{4.9}$$

Is a norm in the other three spaces

Proof. The inclusion of these spaces is obvious, due to the definition of these. For the proof that the application $\|\cdot\|_u$ is a norm, it's immediately given from the proof that the application $\|\cdot\|_\infty$ is a norm in $\ell^\infty(\mathbb{F})$, and that $\|\cdot\|_u = \|\cdot\|_\infty$ □

Remark. Take $f_n \in C_b(\mathbb{F})$ a sequence of functions. The uniform convergence of this sequence means that $f_n \to f$ in the norm $\|\cdot\|_u = \|\cdot\|_\infty$

Proposition 14. If $f \in C_0(\mathbb{F})$, then f is uniformly continuous

Proof. Let $f \in C_0(\mathbb{F})$, then

$$\forall \epsilon > 0 \; \exists l \, : \, |x| \geq l \implies |f(x)| < \frac{\epsilon}{2}$$

Since every continuous function is uniformly continuous in a closed set, then

$$\forall\epsilon > 0\ \exists\delta\ :\ \forall x, y \in [-L-1, L+1] \wedge |x-y| < \delta \implies |f(x) - f(y)| < \epsilon$$

Hence we can have two cases. We either have $|x-y| < \delta$ or $x, y \in [-L-1, L+1]$. Hence we have, in the first case

$$|f(x) - f(y)| \leq |f(x)| + |f(y)| < \epsilon$$

Or, in the second case

$$\forall\epsilon > 0\ \exists\delta > 0\ :\ |x-y| < \delta \implies |f(x) - f(y)| < \epsilon$$

Demonstrating our assumption □

§§ 4.2.1 $C_p(\mathbb{F})$ spaces

Definition 4.2.2. We can define a set of function spaces analogous to the $\ell^p(\mathbb{F})$ spaces. These spaces are the $C_p(\mathbb{F})$ spaces. We define analogously the $p-$norm for functions as follows

$$\|f\|_p := \sqrt[p]{\int_{\mathbb{F}} |f(x)|^p \, \mathrm{d}x} \tag{4.10}$$

Thanks to what said about $\ell^p(\mathbb{F})$ spaces and $p-$norms, it's already obvious that these spaces are normed vector spaces

Remark. Watch out! $C_p(\mathbb{F}) \not\subset C_0(\mathbb{F})$, and $C_p(\mathbb{F}) \not\subset C_q(\mathbb{F})$ for $1 \leq p \leq q$. It's easy to find counterexamples

Proposition 15. If $1 \leq p \leq q$, then

$$C_p(\mathbb{F}) \cap C_b(\mathbb{F}) \subset C_q(\mathbb{F})$$

Proof. Let $f \in C_p(\mathbb{F}) \cap C_b(\mathbb{F})$. Therefore $\sup_{x\in\mathbb{F}} |f(x)| < M \in \mathbb{F}$, then

$$\int_{\mathbb{F}} |f(x)|^q \, \mathrm{d}x = \int_{\mathbb{F}} |f(x)|^p |f(x)|^{q-p} \, \mathrm{d}x \leq M^{q-p} \int_{\mathbb{F}} |f(x)|^p \, \mathrm{d}x \leq \infty$$

Therefore $f \in C_p(\mathbb{F}) \cap C_b(\mathbb{F}) \implies f \in C_q(\mathbb{F})$ □

5 Differential Analysis

§ 5.1 Digression on the Notation Used

In this chapter (and from now on, mostly), we will use a notation which is called *abstract index notation* with the *Einstein summation convention.* This is usually abbreviated in common literature as the *Einstein index notation.* We will give here a brief explanation of how this notation actually works, and why it's so useful in shortening mathematical expressions. Let $\mathcal{V}$ be a vector space and $\mathcal{V}^\star$ be the dual space associated with $\mathcal{V}$. Then we can write the elements $v \in \mathcal{V}$, $\varphi \in \mathcal{V}^\star$ with respect to some basis as follows

$$\begin{aligned} v &= (v_1, v_2, \cdots, v_n) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ \varphi &= (\varphi_1, \varphi_2, \cdots, \varphi_n) = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_n \end{pmatrix} \end{aligned} \tag{5.1}$$

The first notation is the *ordered tuple* notation, meanwhile the second notation is the usual column/row notation for vectors utilized in linear algebra. In Einstein notation we will have that

$$\begin{aligned} v &\longrightarrow v^i \\ \varphi &\longrightarrow \varphi_i \end{aligned} \tag{5.2}$$

Where the vector in the space will be indicated with a raised index (*index, not power!*) and the covector with a lower index, where the index will span all the values $i = 1, \cdots, \dim(\mathcal{V})$.
Let's represent the scalar product in Einstein notation. Let's say that we want to write the scalar product $\langle v, v\rangle$

$$\langle v, v\rangle = \sum_{i=1}^{\dim(\mathcal{V})} v_i v_i \longrightarrow v_i v^i \tag{5.3}$$

Note how we have omitted the sum over the repeated index. Now one might ask why it's not written as $v_i v_i$ (or $v^i v^i$, since $v \in \mathcal{V}$), and this is easily explained introducing the matrix g_{ij}, which is the matrix of the scalar product.
Applying this matrix to v^i we have $g_{ij} v^i$. Note how the low index j is free and i is being summed over, hence is a dummy index, this means that the result must have a lower index j for consistency. So we can write $v_j = g_{ij} v^i$, and due to the lower index we already know that this is a covector,

i.e. a linear functional $\mathcal{V} \longrightarrow \mathbb{F}$, hence it will "eat" a vector and "spew" a scalar (with no indices!). Feeding to this covector the vector v^j we have finally

$$\langle v, v \rangle = v_j v^j = g_{ij} v^i v^j \tag{5.4}$$

Where, algebraically we have "omitted" the definition of $\iota(\cdot) = \langle v, \cdot \rangle$, which is the canonical isomorphism between $\mathcal{V}$ and $\mathcal{V}^\star$.
With this definition we have defined what mathematically are called *musical isomorphisms*, applications which raise and lower indexes. Ironically, this operation is called *index gymnastics*, since we're raising and lowering indices. Thanks to these conventions operations with matrices (and tensors) become much much easier. Let a^i_j and b^i_j be two $n \times n$ matrices over the ordered field $\mathbb{F}$. The multiplication of these two matrices will simply be

$$c^i_j = a^i_k b^k_j \tag{5.5}$$

Note how the k index gets "eaten". This mathematical cannibalism is called *contraction of the index* k. So, the trace for a matrix a^i_j will be

$$\mathrm{tr}(a) = a^i_i \tag{5.6}$$

And now comes the tricky part. In order to write determinants we need to define a symbol, the so called *Levi-Civita symbol*, $\epsilon_{i_1 \dots i_n}$. In three dimensions it's ϵ_{ijk}, and follows the following rules

$$\epsilon_{ijk} = \begin{cases} 1 & \text{even permutation of the indices} \\ -1 & \text{uneven permutation of the indices} \\ 0 & i = j \vee j = k \vee k = i \end{cases} \tag{5.7}$$

In n dimensions, it becomes

$$\epsilon_{i_1 \dots i_k} = \begin{cases} 1 & \text{even permutation of indices} \\ -1 & \text{uneven permutation of indices} \\ 0 & i_i = i_j \text{ for some } i, j \end{cases} \tag{5.8}$$

It's obvious by definition that this weird entity is completely antysimmetrical and unitary, and therefore it's perfect for representing permutations (it's also known as permutation symbol for a reason). Therefore, remembering the definition of the determinant we can write, for an $n \times n$ matrix a^i_j

$$\det(a) = \epsilon_{i_1 \dots i_n} a^{1 i_1} a^{2 i_2} \cdots a^{n i_n} = \epsilon_{i_1 \dots i_n} g^{j i_1} a^1_j g^{k i_2} a^2_k \cdots g^{l i_n} a^n_{i_n} \tag{5.9}$$

If $\dim(\mathcal{V}) = 3$, we can therefore immediately define the cross product of two vectors as follows

$$\mathbf{c} = \mathbf{v} \times \mathbf{w} \longrightarrow c^i = g^{ij} \epsilon_{jkl} v^k w^l \tag{5.10}$$

(Note how we had to raise the index i).
From now on, we will start to use Greek letters for indices and Latin letters for labels, in order to avoid confusions, simply look again at the formula for the determinant, it's much clearer this way. In fact, letting $\mu, \nu, \cdots$ be our indices and $i, j, k, \cdots$ our labels, we can write, for a matrix A^μ_ν

$$\det(A) = A = \epsilon_{\mu_1 \dots \mu_n} g^{\nu \mu_1} g^{\sigma \mu_2} \cdots g^{\zeta \mu_n} A^1_\nu A^2_\sigma \cdots A^n_\zeta \tag{5.11}$$

See? Much clearer, at least in my opinion.
Now we might want to understand how to write norms with this notation. For the usual Euclidean norm it's quite easy. So we can easily write

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \longrightarrow \sqrt{v_\mu v^\mu} \tag{5.12}$$

In case we have a vector function $f^\mu(x^\nu)$, the following notation will be used

$$\|\mathbf{f}(\mathbf{x})\| \longrightarrow \sqrt{f_\mu f^\mu(x^\nu)} \tag{5.13}$$

Or, for the sum of two functions $g^\mu(x^\nu) \pm f^\mu(y^\nu)$

$$\|\mathbf{g}(\mathbf{x}) \pm \mathbf{f}(\mathbf{y})\| \longrightarrow \sqrt{g_\mu g^\mu(x^\nu) + f_\mu f^\mu(y^\nu) \pm 2 g_\mu(x^\nu) f^\mu(y^\nu)} \tag{5.14}$$

Or

$$\|\mathbf{g}(\mathbf{x}) \pm \mathbf{f}(\mathbf{y})\| \longrightarrow= \sqrt{g_{\mu\nu}(g^\mu(x^\gamma) \pm f^\mu(y^\gamma))(g^\nu(x^\gamma) \pm f^\nu(y^\gamma))}$$

A shorthand notation can be created by directly using the norm symbol, but with the contracted index in the upper or lower position as follows

$$\|\mathbf{g}(\mathbf{x}) \pm \mathbf{f}(\mathbf{y})\| \longrightarrow \|g^\mu(x^\nu) \pm f^\mu(y^\nu)\|_\mu \tag{5.15}$$

For p-norms we have to watch out for a little detail. We have to add a square root in order to "fix" the squaring of every element. So we get

$$\|\mathbf{v}\|_p \longrightarrow \sqrt[p]{(v_\mu)^{\frac{p}{2}} (v^\mu)^{\frac{p}{2}}} = \left((v_\mu)^{\frac{p}{2}} (v^\mu)^{\frac{p}{2}} \right)^{\frac{1}{p}} \tag{5.16}$$

Theorem 5.1. $\left((v_\mu v^\mu)^{p/2}\right)^{1/p}$ *is wrong*

Proof. It's easy to see why it doesn't work by expanding the sum on μ

$$\begin{aligned} (v_\mu v^\mu)^p &= \left(v_1 v^1 + v_2 v^2 + \cdots + v_n v^n\right)^p \\ (v_\mu)^{\frac{p}{2}} (v^\mu)^{\frac{p}{2}} &= \left((v_1 v^1)^p + (v_2 v^2)^p + \cdots + (v_n v^n)^p\right) \end{aligned} \tag{5.17}$$

□

Moreover, it's time to bring down some formal rules for the usage of this notation

Theorem 5.2 (Rules for Index Calculus in Einstein Notation)**.**

1. *Free indices must be consistent in both sides of the equation. I.e.* $a^\mu_\nu b_{\mu\gamma\delta} = c_{\nu\gamma\delta}$*.* $a^\mu_\nu b_{\mu\gamma\delta} \neq c^\gamma_{\nu\delta}$*,* $a^\mu_\nu b_{\mu\gamma\delta} \neq c_{\nu\gamma\sigma}$

2. *An index can be repeated* only two times per factor and must be contracted diagonally*. I.e.* $a^\mu b_\mu f^\delta_\gamma = c^\delta_\gamma$ *is defined correctly,* $a_\mu b_\mu$*,* $a^\mu b^\mu$ *or* $a^\mu b_\mu f^\mu_\gamma$ *are ill defined*

3. *Dummy indices can be replaced at will, since they don't contribute to the "index equation"*

§§ 5.1.1 Differential Operators

Differential operators will be defined formally in the next sections, but for now we will simply explain how they actually work with this notation (and what are the advantages of such), alongside the usual boldface notation.
We will begin by defining the derivative along the coordinate vectors (usually indicated with x^μ).
We will use the differential operator *del* (∂).
This operator will be used as follows

1. If there is no ambiguity for the coordinate system, the derivative alongside the coordinates x^μ will be indicated as ∂_μ
2. In case of ambiguity, something will be added in order to distinguish the operators. I.e. let (x^μ, y^ν) be our coordinate system, then we will have ∂_{x^μ} or ∂_{y^ν}
3. In every single case, even the differential operator *must follow the index calculus rules*

Now let $f(x^\mu)$ be some (scalar, there are no free indices) function of the coordinates. The derivative (or gradient, it will soon be defined properly) can be written in various ways. In boldface notation it's usual to indicate this as ∇f, which can be translated as follows

$$\nabla f \longrightarrow \partial_\mu f = \frac{\partial f}{\partial x^\mu} = f_{,\mu} \tag{5.18}$$

Note how in the RHS it's obvious that this quantity must be a vector due to the free index. The last one is the *comma notation* for derivation, used for compacting (even more) the notation (Also check how in the second notation, even if the index is raised, it behaves as a lower index, we will check deeply this part in the section on differential forms).
Now comes the fun part. Higher order derivatives.
For the same function, we can define the *Hessian matrix* (the matrix of second derivatives) **Hf**, simply applying two times the ∂ operator

$$\mathbf{Hf} \longrightarrow \partial_\nu \partial_\mu f = \partial^2_{\mu\nu} f = \partial_{\mu\nu} f = \frac{\partial^2 f}{\partial x^\mu \partial x^\nu} = f_{,\mu\nu} \tag{5.19}$$

Derivatives of order > 2 can then be defined recursively.
Now we might ask, what if we have a vector field F^μ? Nothing changes. We simply have to remember to not repeat indices in order not to represent scalar products.
We have **JF** as the *Jacobian matrix* of F^μ, basically the derivative matrix which in Einstein notation, as before, has a quite obvious nature

$$\mathbf{JF} \longrightarrow \partial_\nu F^\mu = \frac{\partial F^\mu}{\partial x^\nu} = F^\mu_{,\nu} \tag{5.20}$$

And so on, and so on...[1]
Let's now define the divergence and curl operators. Take now a vector field g^μ. We then have

$$\begin{aligned} \nabla \cdot \mathbf{g} &\longrightarrow g_{\mu\nu} g^{\mu\delta} \partial_\delta g^\nu = \partial_\mu g^\mu = \frac{\partial g^\mu}{\partial x^\mu} = g^\mu_{,\mu} \\ \nabla \times \mathbf{g} &\longrightarrow \epsilon_{\mu\nu\sigma} g^{\nu\delta} \partial_\delta g^\sigma = \epsilon_{\mu\nu\sigma} \partial^\nu g^\sigma = \epsilon_{\mu\nu\sigma} \frac{\partial g^\sigma}{\partial x_\mu} = \epsilon_{\mu\nu\sigma} g^{\sigma,\nu} \end{aligned} \tag{5.21}$$

[1] It's quite fun to dive into the dumpster of Einstein notation, isn't it?

And therefore, defining the Laplacian as $\nabla^2 = \nabla \cdot \nabla$, we will simply have, for whatever function h

$$\nabla^2 h \longrightarrow g^{\mu\nu}\partial_\nu\partial_\mu h = \partial^\mu\partial_\mu h = \frac{\partial^2 h}{\partial x^\mu \partial x_\mu} = h^{,\mu}{}_{,\mu} \tag{5.22}$$

Note how the operator ∂^μ appears. This can be seen as a derivation along the covector basis ($x_\mu = g_{\mu\nu}x^\nu$).
Now, we can go back to our mathematical rigor.

§ 5.2 Curves in $\mathbb{R}^n$

Definition 5.2.1 (Scalar Field). A *scalar field* is a function $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ where A is an open set

Definition 5.2.2 (Vector Field). A *vector field* is a function $f^\mu : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ where A is an open set

Definition 5.2.3 (Continuity). A scalar field $f : A \longrightarrow \mathbb{R}$ is said to be continuous in a point $p^\mu \in A$ if

$$\forall \epsilon > 0\ \exists \delta_p\ :\ \|x^\mu - p^\mu\| < \delta \implies |f(x^\mu) - f(p^\mu)| < \epsilon \tag{5.23}$$

A vector field $f^\mu : A \to \mathbb{R}^m$ is said to be continuous instead if

$$\forall \epsilon > 0\ \exists \delta_p\ :\ \|x^\nu - p^\nu\|_\nu < \delta \implies \|f^\mu(x^\nu) - f^\mu(p^\nu)\|_\mu < \epsilon \tag{5.24}$$

If this function is continuous $\forall p^\mu \in A$, then the vector field is said to be part of the space $C(A)$, with $A \subseteq \mathbb{R}^n$

Definition 5.2.4 (Canonical Scalar Product). Let $x^\mu, y^\mu \in \mathbb{R}^n$, the *canonical scalar product* is a bilinear application $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ where, if the components of the two vectors are defined as x^μ, y^μ, is defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x^i y^i \longrightarrow x_\mu y^\mu \tag{5.25}$$

It's easy to see that the canonical scalar product induces the euclidean norm as follows

$$\|\mathbf{v}\| = \|\mathbf{v}\|_2 = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{v_\mu v^\mu} \tag{5.26}$$

Definition 5.2.5 (Curves in $\mathbb{R}^n$). A *curve* is an application $\varphi : [a, b] \subset \mathbb{R} \to \mathbb{R}^n$.
The function $\varphi^\mu(t) = p^\mu$, with $t \in [a, b]$ is called the *parametric representation* of the curve.
Remembering how indexes work in this notation, we already know that this application can be represented with an ordered $n-$tuple or a vector in $\mathbb{R}^n$

Definition 5.2.6 (Regular Curves). A curve $\varphi^\mu(t)$ is said to be continuous if all its components are continuous. A curve is said to be regular iff

$$\begin{gathered} \varphi^\mu(t) \in C^1([a, b]) \\ \sqrt{\varphi_\mu(t)\varphi^\mu(t)} \neq 0 \quad t \in (a, b) \end{gathered} \tag{5.27}$$

A curve is said to be *piecewise regular* if it's not regular in $[a, b]$ but it's regular in a finite number of subsets $[a_n, b_n] \subset [a, b]$

Definition 5.2.7 (Homotopy of Curves). Let γ^μ, η^μ be two curves from a set $[a, b]$, $[c, d]$ respectively. These two curves are said to be *homotopic* to one another, and it's indicated as $\gamma^\mu \sim \eta^\mu$ iff

$$\exists h : [c, d] \xrightarrow{\sim} [a, b],\ h \in C([c, d]), h^{-1} \in C([a, b]),\ h(s) > h(t) \text{ for } s > t\ :\ \eta^\mu = \gamma^\mu \circ h \tag{5.28}$$

Definition 5.2.8 (Tangent Vector)**.** The tangent vector of a regular curve is defined as the following vector.

$$T^\mu(t) = \frac{\dot{\gamma}^\mu}{\sqrt{(\dot{\gamma}_\mu \dot{\gamma}^\mu)}} \tag{5.29}$$

Where with $\dot{\gamma}^\mu(t)$ we indicate the derivative of γ^μ with respect to the only variable t.

Definition 5.2.9 (Tangent Line)**.** A curve $\gamma^\mu : [a, b] \longrightarrow \mathbb{R}^n$ is said to have *tangent line* at a point $t_0 \in [a, b]$ if it's regular, therefore the line will have parametric equations

$$p^\mu(t) = \gamma^\mu(t_0) + \dot{\gamma}^\mu(t_0)(t - t_0) \tag{5.30}$$

Definition 5.2.10 (Length of a Curve)**.** The *length* of a curve $\gamma : [a, b] \longrightarrow \mathbb{R}^n$ is defined as follows

$$L_\gamma := \int_a^b \sqrt{\dot{\gamma}_\mu(t)\dot{\gamma}^\mu(t)}\, \mathrm{d}t \tag{5.31}$$

Remark. In $\mathbb{R}^2$, if a curve is defined in polar coordinates, it will appear as follows

$$\rho = \rho(\theta), \quad \theta \in [\theta_0, \theta_1] \tag{5.32}$$

Its length will be given from the following integral

$$L_\rho := \int_{\theta_0}^{\theta_1} \sqrt{(\rho'(\theta))^2 + (\rho(\theta))^2}\, \mathrm{d}\theta \tag{5.33}$$

The graph of a function $f : [a, b] \longrightarrow \mathbb{R},\ f \in C^1(a, b)$ can also be parametrized from a curve $\varphi^\mu(t)$, where

$$\varphi^\mu(t) \to (t, f(t)) \tag{5.34}$$

Its length will be then calculated with the following integral

$$L_\varphi := \int_a^b \sqrt{1 + (\dot{f}(x))^2}\, \mathrm{d}x \tag{5.35}$$

Definition 5.2.11 (Curviline Coordinate)**.** Let $\varphi : [a, b] \longrightarrow \mathbb{R}^n$, we can define a function $s(t)$ as follows

$$s(t) = \int_a^t \sqrt{\dot{\varphi}_\mu(\tau)\dot{\varphi}_\mu(\tau)}\, \mathrm{d}\tau \tag{5.36}$$

Then

$$\mathrm{d}s = \sqrt{\dot{\varphi}_\mu(t)\dot{\varphi}^\mu(t)}\, \mathrm{d}t \tag{5.37}$$

And the length of a curve can also be indicated as follows

$$L_\varphi = \int_\varphi \mathrm{d}s \tag{5.38}$$

Definition 5.2.12 (Curvature, Normal Vector)**.** The *curvature* of a curve is defined as follows

$$\kappa(s) = \sqrt{T_\mu(s)T^\mu(s)} = \sqrt{\ddot{\varphi}_\mu(s)\ddot{\varphi}^\mu(s)} \tag{5.39}$$

(Note that $\|\varphi'(s)\| = 1$) The *normal vector* is similarly defined as

$$N^\mu(s) = \frac{\dot{T}^\mu(s)}{\kappa(s)} \tag{5.40}$$

Definition 5.2.13 (Simple Curve, Closed Curve)**.** A *simple curve* is an injective application $\gamma : [a, b] \longrightarrow \mathbb{R}^n$. A curve is said to be closed iff $\gamma^\mu(a) = \gamma^\mu(b)$

Theorem 5.3 (Jordan Curve)**.** *Let γ^μ be a simple and closed curve in $\mathbb{R}^2$ or $\mathbb{C}$ (note that $\mathbb{C} \simeq \mathbb{R}^2$), then the set $\{\gamma\}^{\mathrm{c}}$ is defined as follows*

$$\{\gamma\}^{\mathrm{c}} = \{\gamma\}^\circ \cup \operatorname{extr}(\{\gamma\}) \tag{5.41}$$

Note that $\{\gamma\} \subset \mathbb{R}^2$ is the image of the application γ and $\operatorname{extr}(\{\gamma\})$ is the set of points that lay outside of the closed curve.

In $\mathbb{C}$ everything that was said about curves holds, however one must watch out for the definition of modulus, for a curve $\gamma^\mu \in \mathbb{C}$ we will have

$$|\dot{\gamma}(t)| = \sqrt{(\Re'(\gamma))^2 + (\Im'(\gamma))^2} = \sqrt{\overline{\gamma}(t)\gamma(t)} \tag{5.42}$$

§ 5.3 Differentiability in $\mathbb{R}^n$

Definition 5.3.1 (Directional Derivative)**.** Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \longrightarrow \mathbb{R}$. The function is said to be *derivable* with respect to the direction $v^\mu \in A$ at a point $p^\mu \in A$, if the following limit is finite

$$\partial_{v^\mu} f(p^\nu) = \lim_{h \to 0} \frac{f(p^\mu + hv^\mu) - f(p^\mu)}{h} \tag{5.43}$$

If $v^\mu = x^\mu$ then this is called a *partial derivative*, and it will be indicated in the following ways

$$\frac{\partial f}{\partial x^\mu} = \partial_\mu f = \partial_{x^\mu} f \tag{5.44}$$

Definition 5.3.2 (Differentiability)**.** A scalar field $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, with A open, is said to be *differentiable* in a point $p^\mu \in A$ if and only if there exists a linear application $a_\mu(p^\nu) = a_\mu$, such that the following limit is finite

$$\lim_{\sqrt{h_\mu h^\mu} \to 0} \frac{R(p^\mu + h^\mu)}{\sqrt{h_\mu h^\mu}} = 0 \tag{5.45}$$

Where we define the function R as follows

$$R(p^\mu + h^\mu) = f(p^\mu + h^\mu) - (f(p^\mu) + a_\mu h^\mu) \tag{5.46}$$

This means that

$$f(p^\mu + h^\mu) = f(p^\mu) + a_\mu h^\mu + \mathcal{O}\left(\sqrt{h_\mu h^\mu}\right) \tag{5.47}$$

Theorem 5.4 (Consequences of Differentiability)**.** *Let $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable scalar field in every point of A, then*

1. $f \in C(A)$
2. *f is differentiable in A, and $a_\mu = \partial_\mu$, where is the vector differential operator, composed by the partial derivatives*

3. *f has directional derivatives in A and the following equation holds*

$$\partial_{v^\mu} f(p^\nu) = \partial_\mu f(p^\nu) v^\mu$$

4. *$\partial_\mu f$ indicates the maximum and minimum growth of the function f*

5. *There exist a tangent hyperplane to the graphic of the function at the point $(p^\mu, f(p^\mu)) \in \mathbb{R}^{n+1}$ and has the following equation*

$$x^{n+1} = f(p^\mu) + \partial_\mu f(p^\nu)(x^\mu - p^\mu)$$

Proof. 1. f differentiable in A implies $f \in C(A)$

$$\lim_{\sqrt{h_\mu h^\mu}\to 0} f(p^\mu + h^\mu) = \lim_{\sqrt{h_\mu h^\mu}\to 0} (f(p^\mu) + a_\mu h^\mu + \mathcal{O}\left(\sqrt{h_\mu h^\mu}\right)) = f(p^\mu)$$

2. f differentiable in A implies f derivable in A

$$\partial_i f(p^\mu) = \lim_{h\to 0} \frac{f(p^\mu + he^i) - f(p^\mu)}{h} = \lim_{h\to 0} \frac{a^i h + \mathcal{O}(h)}{h} = a^i \in \mathbb{R}$$

Then

$$R(p^\mu + h^\mu) = f(p^\mu + h^\mu) - f(p^\mu) + \partial_\mu f(p^\nu) h^\mu = \mathcal{O}\left(\sqrt{h_\mu h^\mu}\right)$$

3. $\partial_{v^\mu} f = \partial_\mu f v^\mu$

$$\partial_{v^\mu} f(p^\nu) = \lim_{h\to 0} \frac{f(p^\mu + hw^\mu) - f(p^\mu)}{h} = \lim_{h\to 0} \frac{h\partial_\mu f(p^\nu) v^\mu + \mathcal{O}(h)}{h} = \partial_\mu f(p^\nu) v^\mu$$

4. $\partial_\mu f$ indicates the direction of maximum growth.
For Cauchy-Schwartz, we have

$$\sqrt{\partial_{v^\mu} f \partial^{v_\mu} f} = \sqrt{\partial_\mu f v^\mu \partial_\nu f v^\nu} \leq \sqrt{\partial_\mu f \partial^\mu f}\sqrt{v_\nu v^\nu}$$

□

Theorem 5.5 (Continuous Differentiation)**.** *Let $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, with A open. If $f \in C^1(A)$ (i.e. the derivatives of f are continuous), then f is differentiable in A, the vice versa is also true*

Proof. We can write the following equation

$$f(p^\mu + h^\mu) = f(p^1 + h^1, \cdots, p^n) - f(p^1, \cdots, p^n) + \cdots + f(p^1, \cdots, p^n + h^n) - f(p^1, \cdots, p^n) \quad (5.48)$$

For Lagrange, we will have

$$f(p^i + h^i) = h^i \partial_i f(p^1, \cdots, q^i, \cdots, p^n) = h^i \partial_i f(c_i) \quad (5.49)$$

Therefore

$$\lim_{\sqrt{h_\mu h^\mu}\to 0} \frac{|f(p^\mu + h^\mu) - f(a^\mu) - \partial_\mu f h^\mu|}{\sqrt{h_\mu h^\mu}} \leq \lim_{h\to 0} \sum_{i=1}^{n} |\partial_i f(c_i) - \partial_i f(p^\mu)| \frac{|h^i|}{\sqrt{h_\mu h^\mu}} = 0 \quad (5.50)$$

Therefore $\partial_i f(p^\mu)$ is continuous and the function is differentiable. □

Theorem 5.6 (Differentiability of Vector Fields, Jacobian Matrix)**.** *Let $f^\mu : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a vector field and A an open set, then the function f^μ is differentiable iff exists a matrix $J^\mu_\nu \in M_{nm}(\mathbb{R})$ such that*

$$\lim_{\sqrt{h_\mu h^\mu}\to 0} \frac{\left\| f^\mu(p^\nu + h^\nu) - f^\mu(p^\nu) - J^\mu_\nu h^\nu \right\|_\mu}{\sqrt{h_\mu h^\mu}} \tag{5.51}$$

Or, equivalently

$$f_\mu(p^\nu + h^\nu) = f^\mu(p^\nu) + J^\mu_\nu h^\nu + \mathcal{O}\left(\sqrt{h_\mu h^\mu}\right) \tag{5.52}$$

The then J^μ_ν is the matrix of partial derivatives of the vector field, called the Jacobian matrix of the vector field f^μ, and can be calculated as follows

$$J^\mu_\nu(p^\sigma) = \partial_\nu f^\mu(p^\sigma) \tag{5.53}$$

Theorem 5.7 (Composite Derivation)**.** *Let $f^\mu : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and $g^\nu : B \subseteq \mathbb{R}^k \longrightarrow \mathbb{R}^p$ be two differentiable functions in $p^\sigma \in A$, $f^\mu(p^\sigma) \in B$ and A, B open sets, then $h^\nu = g^\nu \circ f^\mu$ is differentiable, and*

$$\partial_\sigma h^\nu = \partial_\mu g^\nu(f^{\tilde{\mu}})\partial_\sigma f^\mu \tag{5.54}$$

Since $\mu = 1, \cdots, k$, $\nu = 1, \cdots, p$, $\sigma = 1, \cdots, n$ it's obvious that $\partial_\sigma h^\nu \in M_{p,n}(\mathbb{R})$, $\partial_\mu g^\nu \in M_{p,k}(\mathbb{R})$, $\partial_\sigma f^\mu \in M_{k,n}(\mathbb{R})$.

Proof. We have that $(g^\nu \circ f^\mu) = g^\nu(f^\mu(p^\sigma))$. Then g^ν is differentiable at $f^\mu(p^\sigma)$ if

$$g^\nu(f^\mu(p^\sigma + s^\sigma)) = g^\nu(f^\mu(p^\sigma)) + \partial_\sigma\left(g^\nu \circ f^{\tilde{\mu}}\right)(f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)) + \mathcal{O}\Big(\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu\Big)$$

Since f^μ is differentiable, we have

$$\begin{aligned} g^\nu(f^\mu(p^\sigma + s^\sigma)) = {} & g^\nu(f^\mu(p^\sigma)) + \partial_\sigma\left(g^\nu \circ f^{\tilde{\mu}}\right)\partial_\sigma f^\mu + \\ & + \partial_\sigma\left(g^\nu \circ f^\mu\right)\mathcal{O}\left(\sqrt{s_\mu s^\mu}\right) + \mathcal{O}\Big(\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu\Big) \end{aligned}$$

Then, we must prove that

$$\lim_{\sqrt{s_\mu s^\mu}\to 0} \frac{\partial_\sigma\left(g^\nu \circ f^\mu\right)\mathcal{O}\Big(\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu\Big)}{\sqrt{s_\mu s^\mu}} = 0$$

But

$$\frac{\partial_\sigma\left(g^\nu \circ f^\mu\right)\mathcal{O}(\sqrt{s_\mu s^\mu})}{\sqrt{s_\mu s^\mu}} \to 0$$

And

$$\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu \le \sqrt{\partial^\sigma f_\mu \partial_\sigma f^\mu}\sqrt{s_\mu s^\mu} + \mathcal{O}\left(\sqrt{s_\mu s^\mu}\right) \le C\sqrt{s_\mu s^\mu}$$

Therefore

$$\frac{\mathcal{O}\Big(\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu\Big)}{\sqrt{s_\mu s^\mu}} = \frac{\mathcal{O}\Big(\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu\Big)}{\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu}\frac{\left\| f^\mu(p^\sigma + s^\sigma) - f^\mu(p^\sigma)\right\|_\mu}{\sqrt{s_\mu s^\mu}} \to 0$$

$\square$

§ 5.4 Differentiability in $\mathbb{C}$

Definition 5.4.1 (Differentiability)**.** A function $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ with G open, is said to be *differentiable* or *derivable* at a point $a \in G$ if exists finite the following limit

$$\left.\frac{\mathrm{d}f}{\mathrm{d}z}\right|_a = f'(a) = \lim_{z\to a} \frac{f(z) - f(a)}{z - a} \tag{5.55}$$

As usual, if this holds $\forall a \in G$, the function is derivable in G

Theorem 5.8. *If $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ is derivable in $a \in G$, then f is continuous in a*

Proof.

$$\lim_{z\to a}(f(z) - f(a)) = \lim_{z\to a}(z - a) \lim_{z\to a} \frac{f(z) - f(a)}{z - a} = 0$$

□

Theorem 5.9 (Some Simple Rules)**.** *Let $f, g : G \subset \mathbb{C} \longrightarrow \mathbb{C}$*

1. $(f \pm g)'(z) = f'(z) \pm g'(z)$
2. $(fg)'(z) = f'(z)g(z) + f(z)g'(z)$
3. $(f/g)'(z) = f'(z)/g(z) - f(z)g'(z)/g^2(z) \quad \forall z \in G \,:\, g(z) \neq 0$
4. $f(z) = c \implies f'(z) = 0$
5. $f(z) = z^n \implies f'(z) = nz^{n-1}$

Theorem 5.10 (Composite Function Derivation)**.** *Let $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ and $g : F \subset \mathbb{C} \longrightarrow \mathbb{C}$, where $f(G) \subset F$. If f is derivable in $a \in G$ and g is derivable in $f(a) \in F$, then $g \circ f$ is derivable, and its derivative is calculated as follows*

$$\left.\frac{\mathrm{d}}{\mathrm{d}z}(g \circ f)\right|_a = \left.\frac{\mathrm{d}g}{\mathrm{d}z}\right|_{f(a)} \left.\frac{\mathrm{d}f}{\mathrm{d}z}\right|_a = g'(f(a))f'(a) \tag{5.56}$$

Proof. Since G is open, $\exists B_r(a) \subset G$. Therefore, taking a sequence $(z)_n \in B_r(a) \,:\, \lim_{n\to\infty}(z)_n = a$. Letting $f(z_n) \neq a$, we can directly write in the definition of derivative

$$\lim_{n\to\infty} \frac{(g \circ f)(z_n) - (g \circ f)(a)}{z_n - a} = (g \circ f)'(a) = g'(f(a))f'(a)$$

Thus, rewriting the function inside the limit

$$\frac{(g \circ f)(z_n) - (g \circ f)(a)}{z_n - a} = \frac{(g \circ f)(z_n) - (g \circ f)(a)}{f(z_n) - f(a)} \frac{f(z_n) - f(a)}{z_n - a} \to 0$$

Since f is continuous in $a \in G$ □

Theorem 5.11 (Inverse Function Derivation)**.** *Let $f : G \subset \mathbb{C} \xrightarrow{\sim} \mathbb{C}$ be a bijective continuous map, with $f^{-1}(w) = z$. If $f(a) \neq 0$ and it's derivable at that same point, we have*

$$\left.\frac{\mathrm{d}f^{-1}}{\mathrm{d}w}\right|_{f(a)} = \frac{1}{f'(a)} \tag{5.57}$$

Proof. Since f is bijective and continuous we can write

$$\left.\frac{\mathrm{d}f^{-1}}{\mathrm{d}w}\right|_{f(a)} = \lim_{w\to f(a)} \frac{f^{-1}(w) - f^{-1}(f(a))}{w - f(a)} = \lim_{z\to a} \frac{z-a}{f(z)-f(a)} = \frac{1}{f'(a)}$$

□

§§ 5.4.1 Holomorphic Functions

Definition 5.4.2 (Holomorphic Function)**.** A function $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ is said to be *holomorphic* in its domain G if G is open, and

$$\forall z \in G \; \exists \frac{\mathrm{d}f}{\mathrm{d}z} = f'(z) \tag{5.58}$$

It is indicated as $f \in H(G)$. It's easy to demonstrate that this set is a vector space.

Theorem 5.12 (Cauchy-Riemann Equation)**.** *Let $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$, where G is open and $f \in H(G)$. Then, if we write $z = x + iy$*

$$\begin{cases} \Re(f(z)) = u(x,y) \\ \Im(f(z)) = v(x,y) \end{cases} \tag{5.59}$$

We have that the function is holomorphic if and only if

$$\begin{cases} \dfrac{\partial u}{\partial x} - \dfrac{\partial v}{\partial y} = 0 \\ \dfrac{\partial v}{\partial x} + \dfrac{\partial u}{\partial y} = 0 \end{cases} \tag{5.60}$$

Alternatively, it can be written as follows

$$\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} = 0 \tag{5.61}$$

Definition 5.4.3 (Wirtinger Derivatives)**.** Before demonstrating the previous theorem, we define the *Wirtinger derivatives* as follows.
Let $z \in \mathbb{C}$, $z = x + iy$ and $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$.

$$\begin{cases} \dfrac{\partial f}{\partial z} = \partial f(z) = \dfrac{1}{2}\left(\dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y}\right) f(z) \\ \dfrac{\partial f}{\partial \overline{z}} = \overline{\partial} f(z) = \dfrac{1}{2}\left(\dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y}\right) f(z) \end{cases} \tag{5.62}$$

Then, the Cauchy-Riemann equations will be equivalent to the following equation

$$\frac{\partial f}{\partial \overline{z}} = \overline{\partial} f(z) = 0 \tag{5.63}$$

Proof. Let $f(z) = u(x,y) + iv(x,y) : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ be a differentiable function in a point z_0, then as we defined, we have that $f \in H(B_\epsilon(z_0))$, and therefore

$$\left.\frac{\mathrm{d}f}{\mathrm{d}z}\right|_{z_0} = \lim_{h\to 0} \frac{f(z_0+h) - f(z_0)}{h}$$

And, therefore, along the imaginary axis and the real axis, we have

$$\begin{aligned}
\lim_{\Re(h)\to 0} \frac{f(z_0 + \Re(h)) - f(z_0)}{\Re(h)} &= \left.\frac{\partial f}{\partial x}\right|_{z_0} \\
\lim_{\Im(h)\to 0} \frac{f(z_0 + i\Im(h)) - f(z_0)}{i\Im(h)} &= \frac{1}{i}\left.\frac{\partial f}{\partial y}\right|_{z_0} = -i\left.\frac{\partial f}{\partial y}\right|_{z_0}
\end{aligned}$$

Due to the continuity of the derivative ($f \in H(B_\epsilon(z_0))$) we must have an equality between these limits

$$\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} = 2\frac{\partial f}{\partial \overline{z}} = 0, \; \therefore f \in H(B_\epsilon(z_0)) \implies \frac{\partial f}{\partial \overline{z}} = 0$$

But, since $f(z) = u(x,y) + iv(x,y)$, we will have that

$$\begin{aligned}
\frac{\partial f}{\partial \overline{z}} &= \frac{\partial}{\partial \overline{z}}(u(x,y) + iv(x,y)) = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(u(x,y) + iv(x,y)) \\
&= \frac{1}{2}\left(\frac{\partial u}{\partial x} + i\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial x} - \frac{\partial v}{\partial y}\right) = 0 \\
&\therefore \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = i\frac{\partial v}{\partial x} + i\frac{\partial u}{\partial y}
\end{aligned}$$

Rewriting the previous equation in a system, we immediately get back the Cauchy-Riemann equations

$$\begin{aligned}
\frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} &= 0 \\
\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} &= 0
\end{aligned}$$

$\square$

Definition 5.4.4 (Whole Function)**.** A function $f : \mathbb{C} \longrightarrow \mathbb{C}$ is said to be *whole* iff $f \in H(\mathbb{C})$

Definition 5.4.5 (Singular Point)**.** Let $f : G \subset \mathbb{C} \longrightarrow \mathbb{C}$ be function such that if $D = B_\epsilon(z_0) \setminus \{z_0\}$ and $f \in H(D)$, then z_0 is said to be a *singular point* of f

For functions $f : [a,b] \subset \mathbb{R} \longrightarrow \mathbb{C}$ every theorem already stated for curves in $\mathbb{R}^n$ with $n = 2$ holds, since $\mathbb{C} \simeq \mathbb{R}^2$. The only thing that should be checked thoroughly is that

$$f(t) = \begin{pmatrix} \Re(f(t)) \\ \Im(f(t)) \end{pmatrix} \in \mathbb{C}|_{\mathbb{R}}$$

Is written as

$$f(t) = \Re(f(t)) + i\Im(f(t)) \in \mathbb{C}$$

§ 5.5 Surfaces

Definition 5.5.1 (Regular Surface)**.** Let $K \subset \mathbb{R}^2$, $K = \overline{E}$ where E is an open and connected subset. A *regular surface* in $\mathbb{R}^3$ is an application

$$r^\mu : K \longrightarrow \mathbb{R}^3$$

Such that

1. $r^\mu \in C^1(K)$, i.e. $\exists \partial_\nu r^\mu \in C(K)$
2. r^μ is injective in K
3. $\operatorname{rank}(\partial_\nu r^\mu) = 2$

The image $\operatorname{Im}(r^\mu) = \Sigma \subset \mathbb{R}^2$ is then defined by the following parametric equations

$$r^\mu(u, v) = \begin{cases} x(u, v) = r^1(u, v) \\ y(u, v) = r^2(u, v) \\ z(u, v) = r^3(u, v) \end{cases} \tag{5.64}$$

The third condition can be rewritten as follows

$$\epsilon^\mu{}_{\nu\sigma} \partial_1 r^\nu \partial_2 r^\sigma = \epsilon^\mu{}_{\nu\sigma} \partial_1 r^\nu \partial_2 r^\sigma \neq 0 \quad \forall (u, v) \in K^\circ \tag{5.65}$$

Remark. A function $f \in C^1(K)$ defines automatically a surface with parametric equations $r^\mu(u, v) = (u, v, f(u, v))$. This surface is always regular since $\epsilon_{\mu\nu\sigma} \partial_1 r^\nu \partial_2 r^\sigma = (-2u, -2v, 1) \neq 0 \quad \forall (u, v) \in K$

Definition 5.5.2 (Coordinate Lines)**.** The curves obtained fixing one of the two variables are called *coordinate lines* in the surface Σ. We have therefore, for a parametric surface $r^\mu(u, v)$ and two fixed values $\tilde{u}, \tilde{v} \in I \subset \mathbb{R}$

$$\begin{aligned} x_1^\mu(t) &= r^\mu(t, \tilde{v}) \\ x_2^\mu(t) &= r^\mu(\tilde{u}, t) \end{aligned} \tag{5.66}$$

Example 5.5.1. The sphere centered in a point $p_0^\mu \in \mathbb{R}^3$, $p_0^\mu = (x_0, y_0, z_0)$ with radius $R \geq 0$ has the following parametric equations

$$\begin{aligned} x &= x_0 + R \sin(u) \cos(v) \\ y &= y_0 + R \sin(u) \sin(v) \\ z &= z_0 + R \cos(v) \end{aligned} \tag{5.67}$$

With $(u, v) \in [0, \pi] \times [0, 2\pi]$. It's a regular surface, since

$$\left\| \epsilon^\mu{}_{\nu\sigma} \partial_1 r^\nu \partial_2 r^\sigma \right\|_\mu = R^2 \sin(u) > 0 \quad \forall (u, v) \in [0, \pi] \times [0, 2\pi]$$

Definition 5.5.3 (Curve on a Surface)**.** Let $\gamma : [a, b] \subset \mathbb{R} \longrightarrow K \subset \mathbb{R}^3$ be a regular curve, and $\mathbf{r} : K \longrightarrow \Sigma$, with the following parametric equations

$$\gamma^\mu(t) = \begin{cases} u = u(t) \\ v = v(t) \end{cases} \tag{5.68}$$

The regular curve $p^\mu(t) = r^\mu(u(t), v(t))$ has $\operatorname{Im} p^\mu \subset \Sigma$. If it passes for a point $p_0^\mu = (u_0, v_0)$ it has tangent line

$$p^\mu(t) = p_0^\mu + \dot{r}^\mu(t)(t - t_0) = p_0^\mu + \partial_u r^\mu(u(t), v(t))\dot{u}(t) + \partial_v r^\mu(u(t), v(t))\dot{v}(t) \tag{5.69}$$

The line is contained inside the following plane

$$\det \begin{pmatrix} (x - x_0) & (y - y_0) & (z - z_0) \\ \partial_1 r^1 & \partial_1 r^2 & \partial_3 r^1 \\ \partial_2 r^1 & \partial_2 r^2 & \partial_3 r^3 \end{pmatrix} \tag{5.70}$$

For a *cartesian surface*, i.e. the surface generated from the graph of a function $f(x, y)$, the tangent plane will be

$$z = f(x_0^\mu) + \partial_\mu f(x_0^\nu)(x^\mu - x_0^\mu) \tag{5.71}$$

Definition 5.5.4 (Normal Vector)**.** The *normal vector* to a surface Σ, $n^\mu(u, v)$ is the vector

$$n^\mu(u, v) = \frac{1}{\sqrt{\epsilon^\mu{}_{\nu\sigma}\epsilon_\mu{}^{\delta\gamma}\partial_1 r^\nu \partial_2 r^\sigma \partial_1 r_\delta \partial_v r_\gamma}} \epsilon^\mu{}_{\nu\sigma}\partial_u r^\nu \partial_v r^\sigma \tag{5.72}$$

For a cartesian surface we have

$$n^\mu(x, y) = \frac{1}{\sqrt{1 + \partial_\mu f \partial^\mu f}} \begin{pmatrix} -\partial_1 f \\ -\partial_2 f \\ 1 \end{pmatrix}^\mu \tag{5.73}$$

Definition 5.5.5 (Implicit Surface)**.** Let $F : A \subset \mathbb{R}^3 \longrightarrow \mathbb{R}$ be a function such that $F \in C^1(A)$, letting $\Sigma := \{ x^\mu \in \mathbb{R}^3 \big| \, F(x^\mu) = 0 \}$. If $x_0^\nu \in \Sigma$ and $\partial_\mu F(x_0^\nu) \neq 0$, Σ coincides locally to a cartesian surface, and the equation of the tangent plane at the point x_0^ν is the following

$$\partial_\mu F(x_0^\nu)(x^\mu - x_0^\mu) = 0 \tag{5.74}$$

Definition 5.5.6 (Metric Tensor)**.** Let $\mathrm{d}s$ be the curviline coordinate of some curve γ^μ inside a regular surface Σ. Then we have that

$$s^\mu(t) = r^\mu(u(t), v(t)) \tag{5.75}$$

And therefore

$$\mathrm{d}s^2 = \mathrm{d}r_\mu \, \mathrm{d}r^\mu = \partial_1 r^\mu \partial_1 r_\mu \left(\mathrm{d}x^1\right)^2 + 2\partial_1 r^\mu \partial_2 r_\mu \, \mathrm{d}x^1 \, \mathrm{d}x^2 + \partial_2 r^\mu \partial_2 r_\mu \left(\mathrm{d}x^2\right)^2 \tag{5.76}$$

In compact form, we can write

$$\mathrm{d}s^2 = g_{\mu\nu} \, \mathrm{d}x^\mu \, \mathrm{d}x^\nu \tag{5.77}$$

And, the *metric tensor* $g_{\mu\nu}$ can defined as follows

$$g_{\mu\nu} = \partial_\mu r^\sigma \partial_\nu r_\sigma \tag{5.78}$$

Or, in matrix notation

$$g_{\mu\nu} = \begin{pmatrix} \partial_1 r^\mu \partial_1 r_\mu & \partial_1 r^\mu \partial_2 r_\mu \\ \partial_2 r^\mu \partial_1 r_\mu & \partial_2 r^\mu \partial_2 r_\mu \end{pmatrix}_{\mu\nu} \tag{5.79}$$

In usual mathematical notation we have

$$g_{\mu\nu} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}_{\mu\nu} \tag{5.80}$$

And it's called the *first fundamental quadratic form* in the language of differential geometry. Then, we can write

$$\mathrm{d}s^2 = E\left(\mathrm{d}x^1\right)^2 + 2F\,\mathrm{d}x^1\,\mathrm{d}x^2 + G\left(\mathrm{d}x^2\right)^2 \tag{5.81}$$

§ 5.6 Optimization

Theorem 5.13 (Fermat)**.** *Let $p^\nu \in A$ be a point of local minimal or maximal for the function $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ with $f \in C^1(A)$. If f is differentiable in p^ν we have*

$$\partial_\mu f(p^\nu) = 0 \tag{5.82}$$

The point p^ν satisfying this condition is then called a stationary point or a critical point for the function f

Proof. Let $v^\mu \in A$ be a direction. The function $g(t) = f(p^\mu + tv^\mu)$ has a point of local maximal or minimal for $t = 0$. Then

$$F'(0) = \partial_{v^\mu} f(p^\nu) = \partial_\mu f(p^\nu) v^\mu = 0 \implies \partial_\mu f(p^\nu) = 0 \tag{5.83}$$

□

Definition 5.6.1 (Hessian Matrix)**.** Let $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, and let $f \in C^1(A)$, then we define the *Hessian matrix* as the matrix of the second partial derivatives of the function f

$$\partial_\mu \partial_\nu f(x^\gamma) = \partial_{\mu\nu} f(x^\gamma) \begin{pmatrix} \partial_{11} f & \cdots & \partial_{1n} f \\ \vdots & \ddots & \vdots \\ \partial_{n1} f & \cdots & \partial_{nn} f \end{pmatrix}_{\mu\nu} (x^\gamma) \tag{5.84}$$

Theorem 5.14 (Schwarz)**.** *Let $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $f \in C^2(A)$, then*

$$\partial_{\mu\nu} f = \partial_{\nu\mu} f \tag{5.85}$$

Definition 5.6.2 (Nature of Critical Points)**.** Let p^γ be a critical point for a function $f \in C^1(A)$. Then

1. $\partial_{\mu\nu} f(p^\gamma)$ is definite positive, then p^γ is a local minimum
2. $\partial_{\mu\nu} f(p^\gamma)$ is definite negative, then p^γ is a local maximum
3. $\partial_{\mu\nu} f(p^\gamma)$ is indefinite, then p^γ is a saddle point

Theorem 5.15. *Here is a list of some rules in order to determine the definition of the matrix $\partial_{\mu\nu} f$. Let $v^\mu \in A$ be a direction, and $p^\gamma \in A$ a critial point of the function $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ then*

1. *If $\partial_{\mu\nu} f(p^\gamma) v^\mu v^\nu > 0 \quad \forall v^\mu \in A \implies \partial_{\mu\nu} f(p^\gamma)$ positive definite*

2. *If* $\partial_{\mu\nu} f(p^\gamma) v^\mu v^\nu < 0 \quad \forall v^\mu \in A \implies \partial_{\mu\nu} f(p^\gamma)$ *negative definite*

3. *If* $\partial_{\mu\nu} f(p^\gamma) v^\mu v^\nu \geq 0 \quad \forall v^\mu \in A \implies \partial_{\mu\nu} f(p^\gamma)$ *semi-positive definite*

4. *If* $\partial_{\mu\nu} f(p^\gamma) v^\mu v^\nu \leq 0 \quad \forall v^\mu \in A \implies \partial_{\mu\nu} f(p^\gamma)$ *semi-negative definite*

5. *If* $v^\mu \neq w^\mu$ *are two directions, and* $\partial_{\mu\nu} f(p^\gamma) v^\mu v^\nu > 0 \wedge \partial_{\mu\nu} f(p^\gamma) w^\mu w^\nu < 0 \implies \partial_{\mu\nu} f(p^\gamma)$ *indefinite*

Theorem 5.16 (Sylvester's Criteria)**.** *Let* $A^\mu_\nu \in M_{nn}(\mathbb{R})$, *and* $(A_k)^\mu_\nu$ *be the reduced matrix with order* $k \leq n$, *then*

1. $\det_{\mu\nu}\left((A_k)^\mu_\nu\right) > 0 \implies A^\mu_\nu$ *positive definite*

2. $(-1)^k \det_{\mu\nu}\left((A_k)^\mu_\nu\right) > 0 \implies A^\mu_\nu$ *negative definite*

3. *If* $\det_{\mu\nu}\left((A_{2k})^\mu_\nu\right) < 0$ *or if* $\det_{\mu\nu}\left((A_{2k+1})^\mu_\nu\right) < 0 \wedge \det_{\mu\nu}\left((A_{2n+1})^\mu_\nu\right) > 0$ *for* $k \neq n$, *then* A^μ_ν *is indefinite*

Theorem 5.17 (Compact Weierstrass)**.** *Let* $f : K \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $f \in C(K)$, *with* K *a compact set, then*

$$\exists p^\mu, q^\mu \in K \,:\, \min_K(f) = f(p^\mu) \leq f(x^\mu) \leq \max_K(f) = f(q^\mu) \ \forall x^\mu \in K \tag{5.86}$$

Proof. Being K a compact set, we have that every sequence $(p^\mu)_n$ converges inside the set, therefore, letting $(p^\mu)_n$ being a minimizing sequence for f. Then there exist a converging subsequence $(p^\mu)_{n_k}$ such that

$$f(p^\mu_{n_k}) \to f(p^\mu)$$

But, since $(p^\mu)_n$ is a minimizing sequence, we have

$$f(p^\mu) = \min_K(f)$$

By definition of minimizing sequence. Analogously, one can define a maximizing sequence and obtain the same result for the maximum of the function in K □

Theorem 5.18 (Closed Weierstrass)**.** *Let* $f : L \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$. *If* $L = \overline{L}$ *and* $f \in C(L)$ *is a coercitive function, i.e.*

$$\lim_{\sqrt{x_\mu x^\mu} \to \infty} f(x^\mu) = +\infty \tag{5.87}$$

Then

$$\exists x^\mu \in L \,:\, \min_L(f) = f(x^\mu) \tag{5.88}$$

Proof. Let $(p^\mu)_n$ be a minimizing sequence for f in L. If this sequence wasn't limited, we would have that $\sqrt{(p_\mu)_n (p^\mu)_n} \to \infty$, and therefore

$$\inf_L(f) = \lim_{n \to \infty} f(p^\mu_n) = +\infty \quad \lightning$$

Therefore $(p^\mu)_n$ must be limited, and the proof is the same as in the case of a compact set. □

Theorem 5.19 (Topology and Functions)**.** *Let* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$*,* $f \in C(\mathbb{R}^n)$*. Then*

$$\begin{aligned} &\{ x^\mu \in \mathbb{R}^n |\, f(x^\mu) < a \in \mathbb{R} \} \\ &\{ x^\mu \in \mathbb{R}^n |\, f(x^\mu) > b \in \mathbb{R} \} \end{aligned} \tag{5.89}$$

Are open sets in $\mathbb{R}^n$ *with the standard topology, and*

$$\begin{aligned} &\{ x^\mu \in \mathbb{R}^n |\, f(x^\mu) \le a \in \mathbb{R} \} \\ &\{ x^\mu \in \mathbb{R}^n |\, f(x^\mu) \ge b \in \mathbb{R} \} \end{aligned} \tag{5.90}$$

Are closed sets

Definition 5.6.3 (Convex Set)**.** A set $A \subset \mathbb{R}^n$ is said to be *convex* if

$$\lambda x^\mu + (1-\lambda) y^\mu \in A \quad \forall x^\mu, y^\mu \in A,\ \forall \lambda \in [0,1] \tag{5.91}$$

Analogously, a function $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be convex, if

$$f\left(\lambda x^\mu + (1-\lambda) y^\mu\right) \le \lambda f(x^\mu) + (1-\lambda) f(y^\mu) \quad \forall x^\mu, y^\mu \in A,\ \forall \lambda \in [0,1] \tag{5.92}$$

The function f is also known as a *sublinear* function
Also, the set

$$E_f = \{ (x^\mu, \lambda) \in A \times \mathbb{R} |\, f(x^\mu) \le y \} \tag{5.93}$$

Is convex

Theorem 5.20 (Convexity)**.** *Let* $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$*.*

1. f *convex in* $A \implies f \in C(A)$
2. f *differentiable in* $A \implies f$ *convex* $\iff f(x^\mu) \ge f(p^\mu) + \langle \nabla f(p^\mu), x^\mu - p^\mu \rangle$
3. $f \in C^2(A) \implies f$ *convex* $\iff \partial_{\mu\nu} f(x^\gamma)$ *is positive semidefinite*

Definition 5.6.4 (Matrix Infinite Norm)**.** Let $A^\mu_\nu(x^\gamma) \in \mathcal{V} \longrightarrow M_{mn}(\mathbb{F})$, where $\dim(\mathcal{V}) = n$. We can define a norm for this space as follows

$$\|A^\mu_\nu\|_\infty = \sqrt{m} \sqrt{\max_\mu \sup_{x^\gamma \in \mathcal{V}} A^{(\mu)}_\nu A^\nu_{(\mu)}(x^\gamma)} \tag{5.94}$$

Theorem 5.21 (Average Value)**.** *Let* $f^\mu : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$*, with* $f \in C^1(A)$*,* A *an open set and* $K \subset A$ *a compact convex subset, then*

$$\|f^\mu(x^\nu) - f^\mu(y^\nu)\|_\mu \le \|\partial_\nu f^\mu\|_\infty \|x^\nu - y^\nu\|_\nu \tag{5.95}$$

Proof. Let $r^\nu(t) = (1-t) y^\nu + t x^\nu$ be a smooth parametrization of a segment connecting the two points x^ν, y^ν, then

$$\begin{aligned} \|f^\mu(r^\nu(1) - f^\mu(r^\nu(0)\|^2_\mu \le \partial^\nu f_\mu \partial_\nu f^\mu(r^\nu(t)) &\le \sup_{r^\gamma} (\partial^\nu f_\mu \partial_\nu f^\mu(r^\gamma)) \|x^\nu - y^\nu\|^2_\nu \\ &\le m \max_\mu \sup_\gamma \left(\partial^\nu f_{(\mu)} \partial_\nu f^{(\mu)}(x^\gamma) \right) \|x^\nu - y^\nu\|^2_\nu \end{aligned}$$

Therefore

$$\|f^\mu(x^\nu) - f^\mu(y^\nu)\|_\mu \le \sqrt{m}\sqrt{\max_\mu \sup_\gamma \left(\partial^\nu f_{(\mu)} \partial_\nu f^{(\mu)}\right)} \|x^\nu - y^\nu\|_\nu = $$
$$= \|\partial_\nu f^\mu\|_\infty \|x^\nu - y^\nu\|_\nu \quad \forall x^\nu, y^\nu \in K$$

□

Theorem 5.22 (Implicit Functions, Dini). *Let $f^\mu : A \subseteq \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$, where $f^\mu \in C^1(A)$. Also let $(x_0^\nu, y_0^\gamma) \in A$ such that*

$$f^\mu(x_0^\nu, y_0^\gamma) = 0$$
$$\det_{\mu\nu}\left(\frac{\partial f^\mu}{\partial y^\gamma}\right) \neq 0$$

Then

$$\exists B_\epsilon(x_0^\nu) = I \subset \mathbb{R}^m,\ B_\epsilon(y_0^\gamma) = J \subset \mathbb{R}^n \ :\ f^\mu(x^\nu, y^\gamma) = 0\ \forall (x^\nu, y^\gamma) \in I \times J$$

Has a unique *solution $y^\gamma = g^\gamma(x^\nu) \in J$, with $g^\gamma \in C^1(I)$, and*

$$\frac{\partial g^\gamma}{\partial x^\nu} = -\left(\frac{\partial f^\mu}{\partial y^\gamma}\right)^{-1} \frac{\partial f^\mu}{\partial x^\nu} \tag{5.96}$$

Proof. Let $B_\mu^\gamma = \left(\partial_{y_0^\gamma} f^\mu\right)^{-1}$, then we know that

$$f^\mu(x^\nu, y^\gamma) = 0 \iff B_\mu^\gamma f^\mu(x^\nu, y^\sigma) = 0 \iff G^\gamma(x^\nu, y^\sigma) = y^\gamma - B_\mu^\gamma f^\mu(x^\nu, y^\sigma) = 0$$

We have therefore

$$G^\gamma(x^\nu, g^\sigma(x^\nu)) = g^\gamma(x^\nu) - B_\mu^\gamma f^\mu(x^\nu, g^\sigma(x^\nu)) = g^\gamma(x^\nu)\ \forall x^\nu \in \overline{B}_r(x_0^\nu) = I$$
$$\frac{\partial G^\gamma}{\partial y^\sigma} = \delta_\sigma^\gamma - B_\mu^\gamma \frac{\partial f^\mu}{\partial y^\sigma}$$
$$\frac{\partial G^\gamma}{\partial y_0^\sigma} = \delta_\sigma^\gamma - B_\mu^\gamma \frac{\partial f^\mu}{\partial y^\sigma} = \delta_\sigma^\gamma - \delta_\sigma^\gamma = 0$$

Now take $(X, d) = (C(I, J), \|\cdot\|_\infty)$, with $J = \overline{B}_\epsilon(y_0^\gamma)$, and define an application $H : X \longrightarrow X$ such that

$$H^\gamma(w^\sigma(x^\nu)) = G^\gamma(x^\nu, w^\sigma(x^\nu))$$

We need to demonstrate that this application is a contraction, i.e. that $\exists! g^\gamma(x^\nu)\ :\ f^\mu(x^\nu, g^\gamma(x^\nu)) = 0\ \forall (x^\nu, y^\gamma) \in I \times J$

$$\|H^\gamma(w^\sigma(x^\nu)) - y_0^\gamma\|_\gamma = \|G^\gamma(x^\nu, w^\sigma(x^\nu)) - y_0^\gamma\|_\gamma \le$$
$$\le \|G^\gamma(x^\nu, w^\sigma(x^\nu)) - G^\gamma(x^\nu, y_0^\sigma)\|_\gamma + \|G^\gamma(x^\nu, y_0^\sigma) - G^\gamma(x_0^\nu, y_0^\nu)\|_\gamma \le$$
$$\le \left\|\frac{\partial G^\gamma}{\partial y^\sigma}\right\|_\infty \|w^\sigma(x^\nu) + y_0^\sigma\|_\sigma + \|G^\gamma(x^\nu, y_0^\sigma) - G^\gamma(x_0^\nu, y_0^\sigma)\|_\gamma \le \epsilon$$

Since $\|G^\gamma(x^\nu, y_0^\gamma) - G^\gamma(x_0^\nu, y_0^\gamma)\|_\gamma \le \epsilon/2$ and $\|w^\sigma(x^\nu) - y_0^\sigma\|_\sigma \le \epsilon$, $\forall (x^\nu, y^\gamma) \in I \times J$, we have

$$\left\|\frac{\partial G^\gamma}{\partial y^\sigma}\right\|_\infty \le \frac{1}{2}$$

Therefore

$$\|H^\gamma(w^\gamma(x^\nu)) - H^\gamma(v^\sigma(x^\nu))\|_\gamma \leq \frac{1}{2}\|w^\gamma(x^\nu) - v^\gamma(x^\nu)\|_\gamma$$

I.e. H is a contraction in $C(I, J)$.
Due to the differentiability of f^μ we can write

$$\forall \epsilon > 0\ \exists \eta_\epsilon\ :\ \|h^\nu\|_\nu, \|k^\gamma\|_\gamma \leq \eta_\epsilon \implies \|f^\mu(x^\nu, g^\gamma(x^\nu) + k^\gamma) - f^\mu(x^\nu, g^\gamma(x^\nu)) - \partial_{x^\nu} f^\mu h^\nu - \partial_{y^\gamma} f^\mu k^\gamma\|_\mu \leq \\ \leq \epsilon(\|h^\nu\|_\nu + \|k^\gamma\|_\gamma)$$

Letting $k^\gamma = g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)$ we have by definition

$$f^\mu(x^\nu + h^\nu, g^\gamma(x^\nu) + k^\gamma) - f^\mu(x^\nu, g^\gamma(x^\nu)) = 0$$

And therefore, putting $\partial_{y^\gamma} f^\mu = \delta^\mu_\gamma$

$$\|g^\gamma(x^\nu + h^\nu) + \partial_{x^\nu} f^\mu h^\nu\|_\mu \leq \epsilon\left(\|h^\nu\|_\nu + \|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)\|_\gamma\right)$$

Letting $\epsilon = 1/2$ we have $\|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)\| \leq \eta_{1/2}$, and we have

$$\begin{aligned} \|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)\|_\gamma &\leq \|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu) - \partial_{x^\nu} f^\gamma h^\nu\|_\gamma \|\partial_{x^\nu} f^\mu h^\nu\|_\gamma \\ &\leq \frac{1}{2}\left(\|h^\nu\|_\nu + \|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)\|_\gamma\right) + \|\partial_{x^\nu} f^\mu\|_\infty \|h^\nu\| \end{aligned}$$

Which implies

$$\|g^\gamma(x^\nu + h^\nu) - g^\gamma(x^\nu)\|_\gamma \leq \|h^\nu\|_\nu \left(1 + 2\|\partial_{x^\nu} f^\mu\|_\infty\right)$$

Which implies that $g^\gamma(x^\nu)$ is continuously differentiable in I. Whenever $\partial_{y^\gamma} f^\mu \neq \delta^\mu_\gamma$ we can find a transformed function $\tilde{f}^\mu$ such that $\partial_{y^\gamma} \tilde{f}^\mu = \delta^\mu_\gamma$ □

Definition 5.6.5 (Vinculated Critical Points). Let $f : K \subset A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ with A open and $f \in C^1(A)$. Let $\partial K = \bigcup_{k \leq n} \gamma_k$ with $\gamma_k : [a_k, b_k] \longrightarrow \mathbb{R}^2$. The critical points of $f|_{\partial K}$ are found in $P_{ik} = \gamma_k(t_i) \in \partial K$, for which

$$\frac{\mathrm{d}}{\mathrm{d}t} f(P_{ik}) = 0$$

Definition 5.6.6 (Argmax, Argmin). Let $f^\mu : A \longrightarrow \mathbb{R}^n$ a function which reach its maximum in $x^\nu_i \in A\ i = 1, \cdots, m$ and its minimum at $y^\nu_j \in A\ j = 1, \cdots, k$ Then we can define

$$\begin{aligned} \operatorname*{Arg\,max}_A(f) &:= \{x^\nu_1, \cdots, x^\nu_m\} \\ \operatorname*{Arg\,min}_A(f) &:= \{y^\nu_1, \cdots, y^\nu_k\} \end{aligned} \tag{5.97}$$

Theorem 5.23 (Lagrange Multipliers). *Let $f, g : A \longrightarrow \mathbb{R}$, $f, g \in C^1(A)$, $A \subseteq \mathbb{R}^n$ open, and $\mathcal{M} = \{\, x^\mu \in A |\, g(x^\mu) = 0\}$ and let $x^\mu_0 \in \mathcal{M}\ :\ \partial_m u g(x^\mu_0) \neq 0$, then $x^\mu_0 \in \operatorname{Arg\,max}_{\mathcal{M}} f \vee \operatorname{Arg\,min}_{\mathcal{M}} f$ if it's a free critical point of the Lagrangian*

$$\mathcal{L}(x^\mu, \lambda) = f(x^\mu) - \lambda g(x^\mu) \quad (x^\mu, \lambda) \in A \times \mathbb{R} \tag{5.98}$$

I.e. $\exists \lambda_0 \in \mathbb{R} : (x_0^\mu, \lambda_0)$ *solves*

$$\begin{cases} \partial_\mu f(x^\nu) = \lambda \partial_\mu g(x^\nu) \\ \quad g(x^\mu) = 0 \end{cases} \tag{5.99}$$

Or, that

$$\operatorname{rank} \begin{pmatrix} \partial_\mu f(x_0^\nu) \\ \partial_\mu g(x_0^\nu) \end{pmatrix} = 1$$

Proof. Let $\partial_n g \neq 0$, then we can see $\mathcal{M}$ as a graph of a regular implicit function of g, $h : \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$, where

$$g(x^\mu, h(x^\mu)) = 0 \quad \forall x^\mu \in B_r(\tilde{x^\mu}_0) \subset \mathbb{R}^{n-1}$$

Letting $\varphi : (-\epsilon, \epsilon) \longrightarrow B_r(x_0^\mu)$ a smooth curve, such that $\varphi^\mu(0) = x_0^\mu$, we have that $\psi^\nu(t) = (\varphi^\mu(t), h(t)) \in \mathcal{M}$ is the parameterization of a smooth curve that passes through $x_0^\mu \in \mathcal{M}$. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\psi^\nu(0)) = \partial_\mu f \dot{\phi}^\mu(0) + \partial_n f \dot{h}(\phi^\mu(0)) = \partial_\nu f(x_0^\mu) s^\nu$$
$$\frac{\mathrm{d}}{\mathrm{d}t} g(\psi^\nu(0)) = \partial_\nu g(x_0^\mu) s^\nu$$

With $s^\nu = \dot{\psi}^\nu(0)$, therefore $\partial_\nu f \|_{\psi^\nu(0)} \partial_\nu g$ □

Theorem 5.24 (Generalized Lagrange Multiplier Method)**.** *Let* $f, g_i : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$*,* $0 < i < n$*,* $f, g_i \in C^1(A)$ *with* A *an open set, let* $\mathcal{M} := \{x^\nu \in A |\, g(x^\nu) = 0\}$*. Take* $x_0^\nu \in \mathcal{M}$ *such that*

$$\operatorname{rank} \partial_\nu g^\mu(x_0^\gamma) = k$$

Then x_0^ν *is a critical point for* $f|_\mathcal{M}$*, and it's a free critical point for the Lagrangian* $\mathcal{L}$

$$\mathcal{L}(x^\gamma, \lambda^\mu) = f(x^\gamma) - \lambda_\nu g^\nu(x^\gamma)$$

I.e. $\exists (x_0^\gamma, \lambda_0^\nu) \in A \times \mathbb{R}$ *solution of the system*

$$\begin{cases} \partial_\nu f(x^\gamma) = \partial_\nu g^\mu(x^\gamma) \lambda^\nu \\ \ g^\nu(x^\gamma) = 0 \end{cases}$$

Alternatively, one can check that

$$\operatorname{rank}(A) = \begin{pmatrix} \partial_\mu f(x_0^\gamma) \\ \partial_\nu g^\mu(x^\gamma) \end{pmatrix} = k$$

6 Tensors and Differential Forms

§ 6.1 Tensors and k-forms

Definition 6.1.1 (Multilinear Functions, Tensors)**.** Let $\mathcal{V}$ be a real vector space, and take $\mathcal{V}^k = \mathcal{V} \times \cdots \times \mathcal{V}$ $k-$times. A function $T : \mathcal{V}^k \longrightarrow \mathbb{R}$ is called *multilinear* if $\forall i = 1, \cdots, k$, $\forall a \in \mathbb{R}$, $\forall v, w \in \mathcal{V}$

$$T(v_1, \cdots, av_i + w_i, \cdots, v_k) = aT(v_1, \cdots, v_i, \cdots, v_k) + T(v_1, \cdots, w_i, \cdots, v_k) \tag{6.1}$$

A multilinear function of this kind is called *k-tensor* on $\mathcal{V}$. The set of all $k-$tensors is denoted as $\mathcal{T}^k(\mathcal{V})$ and is a real vector space.
The tensor T is usually denoted as follows

$$T_{\mu_1 \dots \mu_k} \tag{6.2}$$

Where each index indicates a slot of the multilinear application $T(-, \cdots, -)$

Definition 6.1.2 (Tensor Product)**.** Let $S \in \mathcal{T}^k(V), T \in \mathcal{T}^l(\mathcal{V})$, we define the *tensor product* $S \otimes T \in \mathcal{T}^{k+l}(\mathcal{V})$ as follows

$$(S \otimes T)(v_1, \cdots, v_k, v_{k+1}, \cdots, v_{k+l}) = S(v_1, \cdots, v_k)T(v_{k+1}, \cdots, v_{k+l}) \tag{6.3}$$

This product has the following properties

$$\begin{aligned} (S_1 + S_2) \otimes T &= S_1 \otimes T + S_2 \otimes T \\ S \otimes (T_1 + T_2) &= S \otimes T_1 + S \otimes T_2 \\ (aS) \otimes T &= S \otimes (aT) = a(S \otimes T) \\ (S \otimes T) \otimes U &= S \otimes (T \otimes U) = S \otimes T \otimes U \end{aligned} \tag{6.4}$$

If $S = S_{\mu_1 \dots \mu_k}$ and $T = T_{\mu_{k+1} \dots \mu_{k+l}}$ we have

$$(S \otimes T)_{\mu_1 \dots \mu_k \mu_{k+1} \dots \mu_{k+l}} = S_{\mu_1 \dots \mu_k} T_{\mu_{k+1} \dots \nu_{k+l}} \tag{6.5}$$

Definition 6.1.3 (Dual Space)**.** We define the *dual space* of a real vector space $\mathcal{V}$ as the space of all *linear functionals* from the space to the field over it's defined, and it's indicated with $\mathcal{V}^\star$. I.e. let $\varphi^\mu \in \mathcal{V}^\star$, then $\varphi^\mu : \mathcal{V} \longrightarrow \mathbb{R}$.
It's easy to see how $\mathcal{V}^\star = \mathcal{T}^1(\mathcal{V})$.

Theorem 6.1. *Let $\mathcal{B} = \{v_{\mu_1}, \cdots, v_{\mu_n}\}$ be a basis for the space $\mathcal{V}$, and let $\mathcal{B}^\star := \{\varphi^{\mu_1}, \cdots, \varphi^{\mu_n}\}$ be the basis of the dual space, i.e. $\varphi^\mu v_\nu = \delta^\mu_\nu\ \forall \varphi^\mu \in \mathcal{B}^\star,\ v_\mu \in \mathcal{B}$, then the set of all k-fold tensor products has basis $\mathcal{B}_\mathcal{T}$, where*

$$\mathcal{B}_\mathcal{T} := \{\varphi^{\mu_1} \otimes \cdots \otimes \varphi^{\mu_k},\ \forall i = 1, \cdots, n\} \tag{6.6}$$

Theorem 6.2 (Linear Transformations on Tensor Spaces)**.** *If $f^\mu_\nu : \mathcal{V} \longrightarrow \mathcal{W}$ is a linear transformation, $f^\nu_\mu \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, one can define a linear transformation $f^\star : \mathcal{T}^k(W) \longrightarrow \mathcal{T}^k(V)$ as follows*

$$f^\star T(v_{\mu_1}, \cdots, v_{\mu_k}) = T(f^\mu_\nu v_{\mu_1}, \cdots, f^\mu_\nu v_{\mu_k})$$

Theorem 6.3. *If g is an inner product on $\mathcal{V}$ (i.e. $g : \mathcal{V} \times \mathcal{V} \longrightarrow \mathbb{R}$, with the properties of an inner product), there is a basis $v_{\mu_1}, \cdots, v_{\mu_n}$ of $\mathcal{V}$ such that $g(v_\mu, v_\nu) = g_{\mu\nu} = g_{\nu\mu} = g(v_\nu, v_\mu) = \delta_{\mu\nu}$. This basis is called orthonormal with respect to T. Consequently there exists an isomorphism $f^\mu_\nu : \mathbb{R}^n \xrightarrow{\sim} \mathcal{V}$ such that*

$$g(f^\mu_\nu x^\nu, f^\mu_\nu y^\nu) = x_\mu y^\mu = g_{\mu\nu} x^\mu y^\nu \tag{6.7}$$

I.e.

$$f^\star g(\cdot, \cdot) = g_{\mu\nu} \tag{6.8}$$

Definition 6.1.4 (Alternating Tensor)**.** Let $\mathcal{V}$ be a real vector space, and $\omega \in \mathcal{T}^k(\mathcal{V})$. ω is said to be *alternating* if

$$\begin{aligned} &\omega(v_{\mu_1}, \cdots, v_{\mu_i}, \cdots, v_{\mu_j}, \cdots, v_{\mu_k}) = -\omega(v_{\mu_1}, \cdots, v_{\mu_j}, \cdots, v_{\mu_i}, \cdots, v_{\mu_k}) \\ &\omega(v_{\mu_1}, \cdots, v_{\mu_i}, \cdots, v_{\mu_i}, \cdots, v_{\mu_k}) = 0 \end{aligned} \tag{6.9}$$

Or, compactly

$$\begin{aligned} \omega_{\mu\ldots\nu\ldots\gamma\ldots\sigma} &= -\omega_{\mu\ldots\gamma\ldots\nu\ldots\sigma} \\ \omega_{\mu\ldots\nu\ldots\nu\ldots\gamma} &= 0 \end{aligned} \tag{6.10}$$

The space of all alternating $k-$tensors on $\mathcal{V}$ is indicated as $\Lambda^k(\mathcal{V})$, and we obviously have that $\Lambda^k(\mathcal{V}) \subset \mathcal{T}^k(\mathcal{V})$.
We can define an application $\mathrm{Alt} : \mathcal{T}^k(\mathcal{V}) \longrightarrow \Lambda^k(\mathcal{V})$ as follows

$$\mathrm{Alt}(T)(v^\mu_1, \cdots, v^\mu_k) = \frac{1}{k!} \sum_{\sigma \in \Sigma_k} \mathrm{sgn}(\sigma) T(v^\mu_{\sigma(1)}, \cdots, v^\mu_{\sigma(k)}) \tag{6.11}$$

With $\sigma = (i, j)$ a permutation and Σ_k the set of all permutations of natural numbers $1, \cdots, k$ Compactly, we define an operation on the indices, indicated in square brackets, called the *antisymmetrization* of the indices inside the brackets.
This definition is much more general, since it lets us define a partially antisymmetric tensor, i.e. antisymmetric on only some indices.

$$\mathrm{Alt}(T_{\mu_1\ldots\mu_k}) = \frac{1}{k!} T_{[\mu_1\ldots\mu_k]} \tag{6.12}$$

As an example, for a $2-$tensor $a_{\mu\nu}$ we can write

$$a_{[\mu\nu]} = \frac{1}{2}\left(a_{\mu\nu} - a_{\nu\mu}\right) = \tilde{a}_{\mu\nu} \in \Lambda^2(\mathcal{V}) \tag{6.13}$$

This is valid for general tensors. If we define a $k-$tensor over the product repeated k times for $\mathcal{V}$ and k for its dual space $\mathcal{V} \times \cdots \mathcal{V} \times \mathcal{V}^\star \times \cdots \times \mathcal{V}^\star$, we can define the space $\mathcal{T}^k(\mathcal{V} \times \mathcal{V}^\star) = \mathcal{W}$. Let the basis for this space be the following

$$\mathcal{B}_\mathcal{W} := \{v_{\mu_1} \otimes \cdots \otimes v_{\mu_k} \otimes \varphi^{\nu_1} \otimes \cdots \otimes \varphi^{\nu_k}\}$$

Then an element $\mathcal{Y}$ of the space $\mathcal{W}$ can be written as follows

$$\mathcal{Y}(v_{\mu_1}, \cdots, v_{\mu_k}, \varphi^{\nu_1}, \cdots, \varphi^{\nu_k}) = \mathcal{Y}^{\nu_1 \dots \nu_k}_{\mu_1 \dots \mu_k}$$

We can define a new element $Y \in \Lambda^k(\mathcal{V} \times \mathcal{V}^\star)$ using the antisymmetrization brackets

$$Y^{\nu_1 \dots \nu_k}_{\mu_1 \dots \mu_k} = \mathcal{Y}^{[\nu_1 \dots \nu_k]}_{[\mu_1 \dots \mu_k]}$$

We can define also partially antisymmetric parts as follows

$$R^{\nu_1 \dots \nu_k}_{\mu_1 \dots \mu_k} = \mathcal{Y}^{\nu_1 \dots [\nu_i \nu_{i+1}] \dots \nu_k}_{\mu_1 \dots [\mu_l \mu_{l+1}] \dots \mu_k} = \frac{1}{4!}\left(\mathcal{Y}^{\nu_1 \dots \nu_i \nu_{i+1} \dots \nu_k}_{\mu_1 \dots \mu_l \mu_{l+1} \dots \mu_k} - \mathcal{Y}^{\nu_1 \dots \nu_{i+1} \nu_i \dots \nu_k}_{\mu_1 \dots \mu_l \mu_{l+1} \dots \mu_k} + \mathcal{Y}^{\nu_1 \dots \nu_i \nu_{i+1} \dots \nu_k}_{\mu_1 \dots \mu_l \mu_{l+1} \dots \mu_k} - \mathcal{Y}^{\nu_1 \dots \nu_i \nu_{i+1} \dots \nu_k}_{\mu_1 \dots \mu_{l+1} \mu_l \dots \mu_k}\right)$$

Note how the indexes in the expressions with the label i and l simply got switched, and in the new definition, the tensor R is antisymmetric in both the *covariant* (lower) indexes μ_l, μ_{l+1} and in the *contravariant* (upper) indexes ν_i, ν_{i+1}, where obviously $i, l \leq k$

Theorem 6.4. *Let $T \in \mathcal{T}^k(\mathcal{V})$ and $\omega \in \Lambda^k(\mathcal{V})$. Then*

$$\begin{aligned} T_{[\mu_1 \dots \mu_k]} &\in \Lambda^k(\mathcal{V}) \\ \omega_{[\mu_1 \dots \mu_k]} &= \omega_{\mu_1 \dots \mu_k} \\ T_{[[\mu_1 \dots \mu_k]]} &= T_{[\mu_1 \dots \mu_k]} \end{aligned} \tag{6.14}$$

Definition 6.1.5 (Wedge Product)**.** Let $\omega \in \Lambda^k(\mathcal{V}),\ \eta \in \Lambda^l(\mathcal{V})$. In general $\omega \otimes \eta \notin \Lambda^{k+l}(\mathcal{V})$, hence we define a new product, called the *wedge product*, such that $\omega \wedge \eta \in \Lambda^{k+l}(\mathcal{V})$

$$\omega_{\mu_1 \dots \mu_k} \wedge \eta_{\nu_1 \dots \nu_k} = \frac{(k+l)!}{k!l!}\omega_{[\mu_1 \dots \mu_k}\eta_{\nu_1 \dots \nu_l]} \tag{6.15}$$

With the following properties
$\forall \omega, \omega_1, \omega_2 \in \Lambda^k(\mathcal{V}),\ \forall \eta, \eta_1, \eta_2 \in \Lambda^l(\mathcal{V}),\ \forall a \in \mathbb{R}, \forall f^\star \in \mathcal{L} : \mathcal{T}^k(\mathcal{V}) \longrightarrow \mathcal{T}^l(\mathcal{V})\ \forall \theta \in \Lambda^m(\mathcal{V})$

$$\begin{aligned} (\omega_1 + \omega_2) \wedge \eta &= \omega_1 \wedge \eta + \omega_2 \wedge \eta \\ \omega \wedge (\eta_1 + \eta_2) &= \omega \wedge \eta_1 + \omega \wedge \eta_2 \\ (\omega \wedge \eta) \wedge \theta &= \omega \wedge (\eta \wedge \theta) \\ a\omega \wedge \eta &= \omega \wedge a\eta = a(\omega \wedge \eta) \\ \omega \wedge \eta &= (-1)^{kl}\eta \wedge \omega \\ f^\star(\omega \wedge \eta) &= f^\star(\omega) \wedge f^\star(\eta) \end{aligned} \tag{6.16}$$

Theorem 6.5. *The set*

$$\{\varphi^{\mu_1} \wedge \cdots \wedge \varphi^{\mu_k},\ k < n\} \subset \Lambda^k(\mathcal{V}) \tag{6.17}$$

Is a basis for the space $\Lambda^k(\mathcal{V})$*, and therefore*

$$\dim(\Lambda^k(\mathcal{V})) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Where $\dim(\mathcal{V}) = n$.
Therefore, $\dim(\Lambda^n(\mathcal{V})) = 1$

Theorem 6.6. *Let* $v_{\mu_1}, \cdots, v_{\mu_n}$ *be a basis for* $\mathcal{V}$*, and take* $\omega \in \Lambda^n(\mathcal{V})$*, then, if* $w_\mu = a^\nu_\mu v_\nu$

$$\omega(w_{\mu_1} \cdots w_{\mu_n}) = \det_{\mu\nu}(a^\mu_\nu)\omega(v_{\mu_1}, \ldots, v_{\mu_n}) \tag{6.18}$$

Or using the basis representation of a vector $t^\mu = t^\mu w_\mu = t^\mu a^\nu_\mu v_\nu$ *we have*

$$\omega_{\mu_1 \ldots \mu_n} t^{\mu_1} \cdots t^{\mu_n} = \det_{\mu\nu}(a^\mu_\nu)\omega_{\nu_1 \ldots \nu_n} t^{\nu_1} \cdots t^{\nu_n} \tag{6.19}$$

Proof. Define $\eta_{\mu_1 \ldots \mu_n} \in \mathcal{T}^n(\mathbb{R}^n)$ as

$$\eta_{\mu_1 \ldots \mu_n} a^{\mu_1}_{\nu_1} a^{\mu_2}_{\nu_2} \cdots a^{\mu_n}_{v_n} = \omega_{\mu_1 \ldots \mu_n} a^{\mu_1}_{\nu_1} \cdots a^{\mu_n}_{\nu_n}$$

Hence $\eta \in \Lambda^n(\mathbb{R}^n)$ so $\eta = \lambda \det(\cdot)$ for some λ, and

$$\lambda = \eta_{\mu_1 \ldots \mu_n} e^{\mu_1} \cdots e^{\mu_n} = \omega_{\mu_1 \ldots \mu_n} v^{\mu_1} \cdots v^{\mu_n}$$

□

Definition 6.1.6 (Orientation)**.** The previous theorem shows that a $\omega \in \Lambda^n(\mathcal{V}),\ \omega \neq 0$ splits the bases of $\mathcal{V}$ in two disjoint sets.
Bases for which $\omega(\mathcal{B}_v) > 0$ and for which $\omega(\mathcal{B}_w) < 0$. Defining $w^\mu = a^\mu_\nu v^\nu$ we have that the two bases belong to the same group iff $\det_{\mu\nu}(a^\mu_\nu) > 0$. We call this the *orientation* of the basis of the space. The *usual orientation* of $\mathbb{R}^n$ is

$$[e_\mu]$$

Given another two basis of $\mathbb{R}^n$ we can define (taking the first two examples)

$$[v_\mu]$$
$$-[w_\mu]$$

Definition 6.1.7 (Volume Element)**.** Take a vector space $\mathcal{V}$ such that $\dim(\mathcal{V}) = n$ and it's equipped with an inner product g, such that there are two bases $(v^{\mu_1}, \cdots, v^{\mu_n}),\ (w^{\mu_1}, \cdots, w^{\mu_n})$ that satisfy the *orthonormality condition* with respect to this scalar product

$$g_{\mu\nu} v^{\mu_i} v^{\nu_j} = g_{\sigma\gamma} w^{\sigma_i} w^{\gamma_j} = \delta_{ij} \tag{6.20}$$

Then

$$\omega_{\mu_1 \ldots \mu_n} v^{\mu_1} \cdots v^{\mu_n} = \omega_{\mu_1 \ldots \mu_n} w^{\mu_1} \cdots w^{\mu_n} = \det_{\mu\nu}(a^\mu_\nu) = \pm 1$$

Where

$$w^\mu = a^\mu_\nu v^\nu$$

Therefore

$$\exists! \omega \in \Lambda^n(\mathcal{V}) \,:\, \exists! [w^{\mu_1}, \cdots, w^{\mu_n}] = O$$

Where O is the *orientation* of the vector space.

Definition 6.1.8 (Cross Product)**.** Let $v_1^\mu, \cdots, v_n^\mu \in \mathbb{R}^{n+1}$ and define $\varphi_\nu w^\nu$ as follows

$$\varphi_\nu w^\nu = \det \begin{pmatrix} v^{\mu_1} \\ \vdots \\ v^{\mu_n} \\ w^\nu \end{pmatrix}$$

Then $\varphi \in \Lambda^1(\mathbb{R}^{n+1})$, and

$$\exists! z^\mu \in \mathbb{R}^{n+1} \; : \; z^\mu w_\mu = \varphi_\nu w^\nu$$

z^μ is called the *cross product*, and it's indicated as

$$z^\mu = v^{\nu_1} \times \cdots \times v^{\nu_n} = \epsilon^\mu{}_{\nu_1 \ldots \nu_n} v^{\nu_1} \cdots v^{\nu_n}$$

§ 6.2 Tangent Space and Differential Forms

Definition 6.2.1 (Tangent Space)**.** Let $p \in \mathbb{R}^n$, then the set of all pairs $\{(p, v^\mu) | \, v^\mu \in \mathbb{R}^n\}$ is denoted as $T_p\mathbb{R}^n$ and it's called the *tangent space* of $\mathbb{R}^n$ (at the point. This is a vector space defining the following operations

$$(p, av^\mu) + (p, aw^\mu) = (p, a(v^\mu + w^\mu)) = a(p, v^\mu + w^\mu) \quad \forall v^\mu, w^\mu \in \mathbb{R}^n, \; a \in \mathbb{R}$$

Remark. If a vector $v^\mu \in \mathbb{R}^n$ can be seen as an arrow from 0 to the point v, a vector $(p, v^\mu) \in T_p\mathbb{R}^n$ can be seen as an arrow from the point p to the point $p + v$. In concordance with the usual notation for vectors in physics, we will write $(p, v^\mu) = v^\mu$ directly, or v_p^μ when necessary to specify that we're referring to the vector $v^\mu \in T_p\mathbb{R}^n$. The point $p + v$ is called the *end point* of the vector v_p^μ.

Definition 6.2.2 (Inner Product in $T_p\mathbb{R}^n$)**.** The *usual inner product* of two vectors $v_p^\mu, w_p^\mu \in T_p\mathbb{R}^n$ is defined as follows

$$\begin{aligned} \langle \cdot, \cdot \rangle_p : & T_p\mathbb{R}^n \times T_p\mathbb{R}^n \longrightarrow \mathbb{R} \\ & v_p^\mu w_\mu^p = v^\mu w_\mu = k \end{aligned} \tag{6.21}$$

Analogously, one can define the usual orientation of $T_p\mathbb{R}^n$ as follows

$$[(e^{\mu_1})_p, \cdots, (e^{\mu_n})_p]$$

Definition 6.2.3 (Vector Fields, Again)**.** Although we already stated a definition for a vector field, we're gonna now state the actual precise definition of vector field
Let $p \in \mathbb{R}^n$ be a point, then a function $f^\mu(p) : \mathbb{R}^n \longrightarrow T_p\mathbb{R}^n$ is called a vector field, if $\forall p \in A \subseteq \mathbb{R}^n$ we can define

$$f^\mu(p) = f^\mu(p)(e_\mu)_p \tag{6.22}$$

Where $(e_\mu)_p$ is the canonical basis of $T_p\mathbb{R}^n$
All the previous *(and already stated)* considerations on vector fields hold with this definition.

Definition 6.2.4 (Differential Form)**.** Analogously to vector fields, one can define $k-$forms on the tangent space. These are called *differential (k-)forms* and "live" on the space $\Lambda^k(T_p\mathbb{R}^n)$.

Let $\varphi_p^{\mu_1}, \cdots, \varphi_p^{\mu_k} \in (T_p\mathbb{R}^n)^\star$ be a basis on such space, then the differential form $\omega \in \Lambda^k(T_p\mathbb{R}^n)$ is defined as follows

$$\omega_{\mu_1\dots\mu_k}(p) = \omega_{\mu_1\dots\mu_k}\varphi_p^{[\mu_1}\cdots\varphi_p^{\mu_k]} \to \sum_{i_1<\dots<i_k} \omega_{i_1\dots i_k}(p)\varphi_{i_1}(p)\wedge\cdots\wedge\varphi_{i_k}(p) \tag{6.23}$$

A function $f : T_p\mathbb{R}^n \longrightarrow \mathbb{R}$ is defined as $f \in \Lambda^0(T_p\mathbb{R}^n)$, or a $0-$form. In general, so, we can write without incurring in errors

$$f(p)\omega = f(p)\wedge\omega = f(p)\omega_{\mu_1\dots\mu_k} \tag{6.24}$$

Definition 6.2.5 (Differential)**.** Now we will omit that we're working on a point $p \in \mathbb{R}^n$ and we'll use the usual notation.
Let $f : T_p\mathbb{R}^n \longrightarrow \mathbb{R}$ be a smooth (i.e. continuously differentiable) function, where $f \in C^\infty$, then, using operatorial notation we have that $\partial_\mu f(v) \in \Lambda^1(\mathbb{R}^n)$, therefore, with a small modification, we can define

$$\mathrm{d}f(v_p^\nu) = \partial_\mu f(v^\nu) \tag{6.25}$$

It's obvious how $\mathrm{d}x^\mu(v_p^\nu) = \partial_\nu x^\mu(v^\nu) = v^\mu$, therefore $\mathrm{d}x^\mu$ is a basis for $\Lambda^1(T_p\mathbb{R}^n)$, which we will indicate as $\mathrm{d}x^\mu$, therefore $\forall\omega \in \Lambda^k(T_p\mathbb{R}^n)$

$$\omega_{\mu_1\dots\mu_k} = \omega_{\mu_1\dots\mu_k}\,\mathrm{d}x^{[\mu_1}\cdots\mathrm{d}x^{\mu_k]} \to \sum_{i_1<\dots<i_k} \omega_{i_1\dots i_k}(p)\,\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k} \tag{6.26}$$

Basically, the vectors $\mathrm{d}x^\mu$ are the *dual basis* with respect to the canonical basis $(e_\mu)_p$

Theorem 6.7. *Since* $\mathrm{d}f(v_p^\nu) = \partial_\nu f(v^\nu)$ *we have, expressing the differential of a function with the basis vectors,*

$$\mathrm{d}f = \frac{\partial f}{\partial x^\mu}\,\mathrm{d}x^\mu = \partial_\mu f\,\mathrm{d}x^\mu \tag{6.27}$$

Definition 6.2.6. Having defined a smooth linear transformation $f_\nu^\mu : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, it induces another linear transformation $\partial_\gamma f_\nu^\mu : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, which with some modifications becomes the application $(f_\star)_\nu^\mu : T_p\mathbb{R}^n \longrightarrow T_{f(p)}\mathbb{R}^m$ defined such that

$$(f_\star)_\nu^\mu(v^\nu) = \left(\mathrm{d}f|_{f(p)}\right)_\nu^\mu(v^\nu) \tag{6.28}$$

Which, in turn, also induces a linear transformation $f^\star : \Lambda^k(T_{f(p)}\mathbb{R}^m) \longrightarrow \Lambda^k(T_p\mathbb{R}^n)$, defined as follows. Let $\omega_p \in \Lambda^k(\mathbb{R}^m)$, then we can define $f^\star\omega \in \Lambda^k(T_{f(p)}\mathbb{R}^n)$ as follows

$$(f^\star\omega_p)(v_{\mu_1},\dots,v_{\mu_k}) = \omega_{f(p)}\left((f_\star)_{\nu_1}^{\mu_1}v_{\mu_1},\cdots,(f_\star)_{\nu_k}^{\mu_k}v_{\mu_k}\right) \tag{6.29}$$

(Just remember that in this way we are writing explicitly the chosen base, watch out for the indexes!)

Theorem 6.8. *Let* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a smooth function, then*

1. $(f^\star)_\nu^\mu(\mathrm{d}x^\nu) = \mathrm{d}f = \partial_\nu f^\mu\,\mathrm{d}x^\nu$
2. $f^\star(\omega_1+\omega_2) = f^\star\omega_1 + f^\star\omega_2$
3. $f^\star(g\omega) = (g\circ f)f^\star\omega$

4. $f^\star(\omega \wedge \eta) = f^\star\omega \wedge f^\star\eta$
5. $f^\star\left(h\,\mathrm{d}x^{[\mu_1}\cdots\mathrm{d}x^{\mu_n]}\right) = h\circ f\,\det_{\mu\nu}(\partial_\nu f^\mu)\,\mathrm{d}x^{[\mu_1}\cdots\mathrm{d}x^{\mu_n]}$

Definition 6.2.7 (Exterior Derivative)**.** We define the operator d as an operator $\Lambda^k(T_p\mathcal{V}) \xrightarrow{\mathrm{d}} \Lambda^{k+1}(T_p\mathcal{V})$ for some vector space $\mathcal{V}$. For a differential form ω it's defined as follows

$$(\mathrm{d}\omega)_{\nu\mu_1\dots\mu_k} = \partial_{[\nu}\omega_{\mu_1\dots\mu_k]} \tag{6.30}$$

This, using the classical mathematical notation can be written as follows

$$\begin{aligned}\mathrm{d}\omega &= \sum_{i_1<\dots<i_k}\mathrm{d}\omega_{i_1,\dots,i_k}\wedge\mathrm{d}x^{i_1}\wedge\dots\wedge\mathrm{d}x^{i_k}\\ \mathrm{d}\omega &= \sum_{i_1<\dots<i_k}\sum_{j=1}^{n}\frac{\partial}{\partial x^j}\omega_{i_1,\dots,i_k}\,\mathrm{d}x^j\wedge\mathrm{d}x^{i_1}\wedge\dots\wedge\mathrm{d}x^{i_k}\end{aligned} \tag{6.31}$$

Theorem 6.9 (Properties of d)**.** *1.* $\mathrm{d}(\omega+\eta) = \mathrm{d}\omega + \mathrm{d}\eta$

2. $\mathrm{d}(\omega\wedge\eta) = \mathrm{d}\omega\wedge\eta + (-1)^k\omega\wedge\mathrm{d}\eta$ *for* $\omega\in\Lambda^k(\mathcal{V}),\ \eta\in\Lambda^l(\mathcal{V})$
3. $\mathrm{d}\mathrm{d}\omega = \mathrm{d}^2\,\omega = 0$
4. $f^\star(\mathrm{d}\omega) = \mathrm{d}(f^\star\omega)$

Definition 6.2.8 (Closed and Exact Forms)**.** A form ω is called *closed* iff

$$\mathrm{d}\omega = 0 \tag{6.32}$$

It's called exact iff

$$\omega = \mathrm{d}\eta \tag{6.33}$$

Theorem 6.10. *Let ω be an exact differential form. Then it's closed*

Proof. The proof is quite straightforward. Since ω is exact we can write $\omega = \mathrm{d}\rho$ for some differential form ρ, therefore

$$\mathrm{d}\omega = \mathrm{d}\mathrm{d}\rho = \mathrm{d}^2\,\rho = 0$$

Hence $\mathrm{d}\omega = 0$ and ω is closed. □

Example 6.2.1. Take $\omega\in\Lambda^1(\mathbb{R}^2)$, where it's defined as follows

$$\omega_\mu = p\,\mathrm{d}x + q\,\mathrm{d}y \tag{6.34}$$

The external derivative will be of easy calculus by remembering the mnemonic rule $\mathrm{d}\to\partial_\mu\wedge\mathrm{d}x^\mu$, or also as $\partial_{[\nu}$ then we have

$$\mathrm{d}\omega_{\mu\nu} = \partial_{[\nu}\omega_{\mu]}$$

But

$$\partial_\nu\omega_\mu = \begin{pmatrix}\partial_1\omega_1 & \partial_1\omega_2\\ \partial_2\omega_1 & \partial_2\omega_2\end{pmatrix}_{\mu\nu}$$

And

$$\partial_{[\nu}\omega_{\mu]} = \frac{1}{2}(\partial_\nu\omega_\mu - \partial_\mu\omega_\nu) = \frac{1}{2}(\partial\omega - \partial\omega^{\mathrm{T}})$$

Therefore

$$\mathrm{d}\omega_{\mu\nu} = \frac{1}{2}\begin{pmatrix} 0 & \partial_x q - \partial_y p \\ \partial_y p - \partial_x q & 0 \end{pmatrix}_{\mu\nu}$$

Which, expressed in terms of the basis vectors of $\Lambda^2(\mathbb{R}^2)$, $\mathrm{d}x \wedge \mathrm{d}y$, we get

$$\mathrm{d}\omega = \frac{1}{2}(\partial_x q - \partial_y p)\,\mathrm{d}x \wedge \mathrm{d}y + \frac{1}{2}(\partial_y p - \partial_x q)\,\mathrm{d}y \wedge \mathrm{d}x = (\partial_x q - \partial_y p)\,\mathrm{d}x \wedge \mathrm{d}y \tag{6.35}$$

Therefore

$$\mathrm{d}\omega = 0 \iff \partial_x q - \partial_y p = 0 \tag{6.36}$$

Definition 6.2.9 (Star Shaped Set)**.** A set A is said to be *star shaped with respect to a point* a iff $\forall x \in A$ the line segment $[a, x] \subset A$

Lemma 6.2.1 (Poincaré's)**.** Let $A \subset \mathbb{R}^n$ be an open star shaped set, with respect to 0. Then every closed form on A is exact

§ 6.3 Chain Complexes and Manifolds

<++> *E M P T Y* <++>

Bibliography

[Ces18] F. Cesi. *Rudimenti di Analisi Infinito Dimensionale, Versione 0.3 (Armadillo)*. 2018.

[Fla18] Eugenio Montefusco Flavia Lanzara. *Esercizi Svolti di Analisi Vettoriale con Elementi di Teoria*. Edizioni LaDotta, 2018. ISBN: 9788898648429.

[Pre13] C. Presilla. *Elementi di Analisi Complessa, Funzioni di una Variabile*. Springer, 2013. ISBN: 9788847055001.

[Spi65] Michael Spivak. *Calculus on Manifolds, a Modern Approach to Classical Theorems of Advanced Calculus*. Addison-Wesley Publishing Company, 1965. ISBN: 0805390219.