

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 29, 2020

VERSION HILBERTIAN (COMPLETE) MADNESS

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

L^AT_EX 2_ε INSIDE, VIM POWERED.

JANUARY 29, 2020

VERSION HILBERTIAN (COMPLETE) MADNESS

Contents

1	Complex Numbers and Functions	9
1.1	Complex Numbers	9
1.2	Regions in \mathbb{C}	10
1.2.1	Extended Complex Plane $\hat{\mathbb{C}}$	11
1.3	Elementary Functions	11
1.3.1	Complex Exponentiation	12
1.3.2	Properties of Trigonometric Functions	12
2	Abstract Spaces	15
2.1	Metric Spaces	15
2.2	Convergence and Compactness	19
2.3	Vector Spaces	22
2.3.1	Hölder and Minkowski Inequalities	23
3	Sequences and Series of Functions	25
3.1	Sequences of Functions	25
3.2	Series of Functions	27
3.2.1	Power Series and Convergence Tests	28
4	Infinite Dimensional Spaces	31
4.1	Sequence Spaces	31
4.1.1	Space of Bounded Sequences	32
4.1.2	Space of Sequences Converging to 0	33
4.1.3	$\ell^p(\mathbb{F})$ Spaces	33
4.1.4	Space of Finite Sequences	34
4.2	Function Spaces	34
4.2.1	$C_p(\mathbb{F})$ spaces	35
4.2.2	Function Spaces in \mathbb{R}^n	35
5	Differential Analysis	39
5.1	Digression on the Notation Used	39
5.1.1	Differential Operators	42
5.2	Curves in \mathbb{R}^n	44
5.3	Differentiability in \mathbb{R}^n	46

5.4	Differentiability in \mathbb{C}	49
5.4.1	Holomorphic Functions	50
5.5	Surfaces	52
5.6	Optimization	54
6	Tensors and Differential Forms	61
6.1	Tensors and k -forms	61
6.2	Tangent Space and Differential Forms	65
6.3	Chain Complexes and Manifolds	68
7	Integral Analysis	71
7.1	Measure Theory	71
7.1.1	Jordan Measure	72
7.1.2	Lebesgue Measure	73
7.2	Integration	75
7.2.1	Lebesgue Integration	76
7.3	Calculus of Integrals in \mathbb{R}^2 and \mathbb{R}^3	80
7.3.1	Change of Variables	83
7.3.2	Line Integrals	85
7.3.3	Surface Integrals	87
7.4	Integration in \mathbb{C}	88
7.4.1	Integral Representation of Holomorphic Functions	93
7.5	Integral Theorems in \mathbb{R}^2 and \mathbb{R}^3	96
8	Series Representation and Residues	99
8.1	Series Representation of Functions	99
8.1.1	Taylor Series	99
8.1.2	Laurent Series	102
8.1.3	Multiplication and Division of Power Series	104
8.1.4	Useful Expansions	105
8.2	Residues	106
8.2.1	Singularities and Residues	106
8.2.2	Classification of Singularities, Zeros and Poles	108
8.3	Applications of Residue Calculus	112
8.3.1	Improper Integrals	112
8.3.2	General Rules	115
9	Hilbert Spaces	117
9.1	Hilbert Spaces	117
9.2	Projections	119
9.2.1	Orthogonal Systems and Bases	123
10	Distributions	127
10.1	Linear Functionals	127
10.2	Distributions	130
10.2.1	Local Integrability	130
10.2.2	Regular and Singular Distributions	134

10.2.3	Operations with Distributions	140
10.2.4	A Physicist's Trick, The Dirac δ "Function"	149
10.2.5	Integral Representation of Distributions	150
10.2.6	Distributions in \mathbb{R}^n	150
10.3	Some Applications in Physics	150
10.3.1	Electrostatics	150
10.3.2	Electrodynamics	150
11	Differential Equations	151
11.1	Existence of Solutions	151
11.2	Common Solving Methods	151
11.3	Green Identities and Kelvin Transform	151
12	Fourier Calculus	153
12.1	Bessel Equality and Fourier Coefficients	153
12.1.1	Trigonometric Polynomials in $L^2[-\pi, \pi]$	153
12.2	Fourier Series	153
12.2.1	Fourier Series in $[a, b]$	153
12.2.2	Complex Fourier Series	153
12.2.3	Heat Equation with Fourier Series	153
12.3	Fourier Transform	153
12.3.1	Convolution	153
12.3.2	Heat Equation with Fourier Transforms	153
12.4	Probabilistic Considerations	153

Introduction

The idea of writing this came from a mix of pure laziness and despair in preparing two exams, Analisi Vettoriale (Vector Analysis) and Metodi e Modelli Matematici per la Fisica (Mathematical Models and Methods for Physics, basically just pure Complex Analysis and Functional Analysis), that I have to sustain in my BSc course in Physics here at Sapienza University of Rome.

These two courses are deeply intertwined and it's really difficult to study them apart due to the sheer volume of things that either are done for half in one course and half in the other one, or they simply get generalized in the second, breaking up the logical flow one might get from studying these two seemingly completely different subjects.

There will surely be errors in grammar, typing and probably some mathematical inaccuracies so for any question, inquiry or you just want to say hi to me, don't wait in sending me an email here cheri.1686219@studenti.uniroma1.it.

A huge thanks to anyone of you that'll read this, I hope it will be useful for you as it is to me.

Notation

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

1 Complex Numbers and Functions

§ 1.1 Complex Numbers

Definition 1.1.1 (Complex Numbers). Define with \mathbb{C} the set of *complex numbers*, i.e. the set of numbers $z \in \mathbb{C} : z = (x, y)$ and $x, y \in \mathbb{R}$.

We define the *real and imaginary parts* of z as follows

$$\begin{aligned}\Re(z) &= x \\ \Im(z) &= y\end{aligned}\tag{1.1}$$

Definition 1.1.2 (Operations in \mathbb{C}). Take $z_1, z_2 \in \mathbb{C}$, then we define

$$\begin{aligned}z_1 = z_2 &\iff \Re(z_1) = \Re(z_2), \Im(z_1) = \Im(z_2) \\ z_1 + z_2 &= (\Re(z_1) + \Re(z_2), \Im(z_1) + \Im(z_2)) \\ z_1 z_2 &= (\Re(z_1)\Re(z_2) - \Im(z_1)\Im(z_2), \Re(z_1)\Im(z_2) + \Im(z_1)\Re(z_2))\end{aligned}$$

Theorem 1.1. *With the previous definitions the set \mathbb{C} forms a field.*

Definition 1.1.3 (Imaginary Unit). We define the imaginary unit $i = (0, 1) \in \mathbb{C}$. From this definition and the definition of product of two complex numbers, we have that $i^2 = -1$

With this definition, we have

$$\forall z \in \mathbb{C} \quad z = \Re(z) + i\Im(z)\tag{1.2}$$

Definition 1.1.4 (Complex Conjugate). Taken $z \in \mathbb{C}$, we call the *complex conjugate* of z the number w such that

$$w = \Re(z) - i\Im(z)\tag{1.3}$$

This number is denoted as \bar{z}

Definition 1.1.5 (Complex Module). We define the *module* or *norm* of a complex number, the following operator.

$$\|z\| = \sqrt{z\bar{z}} = \sqrt{\Re^2(z) + \Im^2(z)}\tag{1.4}$$

Definition 1.1.6 (Complex Inverse). The inverse of a complex number $z \in \mathbb{C}$ is defined as z^{-1} and it's calculated as follows

$$z^{-1} = \frac{\bar{z}}{\|z\|^2}\tag{1.5}$$

Definition 1.1.7 (Polar Form). Taken a complex number $z \in \mathbb{C}$ one can define it in polar form with its modulus r and its *argument* θ . We have that, if $z = x + iy$

$$\begin{aligned} r &= \sqrt{x^2 + y^2} = \|z\|^2 \\ \tan(\theta) &= \frac{y}{x} \end{aligned} \quad (1.6)$$

Definition 1.1.8 (Principal Argument). Taken $\arg(z) = \theta$ we can define two different arguments, due to the periodicity of the tan function.

1. $\text{Arg}(z) \in (-\pi, \pi]$ called the *principal argument*
2. $\arg(z) = \text{Arg}(z) + 2k\pi$, $k \in \mathbb{Z}$ called the *argument*

As a rule of thumb, using the previous definition of argument of a complex number $z = x + iy$, we have

$$\text{Arg}(z) = \begin{cases} \arctan(y/x) - \pi & x < 0, y < 0 \\ \arctan(y/x) & x \geq 0, z \neq 0 \\ \arctan(y/x) + \pi & x < 0, y \geq 0 \end{cases} \quad (1.7)$$

Definition 1.1.9 (\arg_+). Given $z \in \mathbb{C}$ we define the $\arg_+(z)$ as the only value of $\arg(z)$ such that $0 \leq \theta < 2\pi$.

In case we have a polydromic function, in order to specify we're using this argument, there will be a $+$ as index.

I.e. $\log_+(z)$, $[z^a]^+$, \sqrt{z}^+ , \dots and so on.

Theorem 1.2 (De Moivre Formula). A complex number $z \in \mathbb{C}$ in polar form can be written with complex exponential and sine and cosine function as follows.

$$z = \|z\|^2 e^{i \arg z} = \|z\|^2 (\cos(\arg z) + i \sin(\arg z)) \quad (1.8)$$

This formula easily generalizes the calculus of exponentials of complex numbers. With this definition, it's obvious that the n -th root of a complex number $\sqrt[n]{z}$ has actually $n - 1$ results, given the 2π -periodicity of the $\arg(z)$ function.

§ 1.2 Regions in \mathbb{C}

Definition 1.2.1 (Line). A line λ in \mathbb{C} , from z_1, z_2 can be written as follows

$$\lambda(t) = z_1 + t(z_2 - z_1) \quad t \in [0, 1] \quad (1.9)$$

If $t \in \mathbb{R}$ this defines the line lying between z_1, z_2 . Its non-parametric representation is the following

$$\{\lambda\} := \left\{ z \in \mathbb{C} \mid \Im\left(\frac{z - z_1}{z_2 - z_1}\right) = 0 \right\} \quad (1.10)$$

Where $z = \lambda(t)$.

Definition 1.2.2 (Circumference). A circumference γ centered in a point $z_0 \in \mathbb{C}$ with radius R is defined as follows

$$\gamma(\theta) = z_0 + R e^{i\theta} \quad \theta \in [0, 2\pi] \quad (1.11)$$

Non parametrically, it can be defined as follows

$$\{\gamma\} := \{z \in \mathbb{C} \mid \|z - z_0\| = R\} \quad (1.12)$$

§§ 1.2.1 Extended Complex Plane $\hat{\mathbb{C}}$

Definition 1.2.3 (Extended Complex Plane). We define the *extended complex plane* $\hat{\mathbb{C}}$ as follows

$$\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\} \quad (1.13)$$

This can be imagined by projecting \mathbb{C} into the Riemann sphere centered in the origin.

Definition 1.2.4 (Points in $\hat{\mathbb{C}}$). Given a point $z \in \mathbb{C}$, $z = x + iy$ we can find its coordinates with the following transformation

$$\hat{z} = (xt, yt, 1 - t) \in \hat{\mathbb{C}} \quad (1.14)$$

Where the condition $\|\hat{z}\| = 1$ must hold, defining the value of $t \in \mathbb{R}$ Inversely, given $\hat{z} = (x_1, x_2, x_3) \in \hat{\mathbb{C}}$ one finds

$$z = \frac{x_1 + ix_2}{1 - x_3} \quad (1.15)$$

§ 1.3 Elementary Functions

Definition 1.3.1 (Exponential). The exponential function $z \mapsto e^z$ with $z \in \mathbb{C}$ is defined as follows

$$e^z = e^{\Re(z) + i\Im(z)} = e^{\Re(z)} (\cos(\Im(z)) + i \sin(\Im(z))) \quad (1.16)$$

This gives

$$\begin{aligned} \|e^z\| &= |e^{\Re(z)}| \\ \arg(e^z) &= \Im(z) + 2\pi k \quad k \in \mathbb{Z} \end{aligned} \quad (1.17)$$

We have therefore, for $z, w \in \mathbb{C}$

$$\begin{aligned} e^z e^w &= e^{z+w} \\ \frac{e^z}{e^w} &= e^{z-w} \end{aligned} \quad (1.18)$$

Definition 1.3.2 (Logarithm). We define the logarithm function $z \mapsto \log z$ as follows

$$\log(z) = \log \|z\| + i \arg(z) \quad (1.19)$$

It's evident how this function has multiple values for the same z value, and therefore is known as a *polydromic function*, like the square root. We also define the principal branch of the logarithm as $\text{Log}(z)$

$$\text{Log}(z) = \log \|z\| + i \text{Arg}(z) \quad (1.20)$$

Lastly we define the $\log_+(z)$ as follows

$$\log_+(z) = \log(\|z\|) + i \arg_+(z) \quad (1.21)$$

Definition 1.3.3 (Branch of the Logarithm). A general branch of the log function is defined as the function $f(z) : D \subset \mathbb{C} \rightarrow \mathbb{C}$ such that

$$e^{f(z)} = z \quad (1.22)$$

§§ 1.3.1 Complex Exponentiation

Definition 1.3.4 (Complex Exponential). Taken $s, z \in \mathbb{C}$, we define the complex exponential as follows, taken z a variable

$$z^s = e^{s \log(z)} \quad z \neq 0 \quad (1.23)$$

Its derivative has the following value

$$\frac{d}{dz} z^s = s e^{(s-1) \log(z)} = s z^{s-1} \quad (1.24)$$

Alternatively, we define

$$s^z = e^{z \log(s)} \quad (1.25)$$

§§ 1.3.2 Properties of Trigonometric Functions

Definition 1.3.5 (Trigonometric Functions). Using De Moivre's formula, we define

$$\begin{aligned} \sin(z) &= \frac{1}{2i} (e^{iz} - e^{-iz}) \\ \cos(z) &= \frac{1}{2} (e^{iz} + e^{-iz}) \end{aligned} \quad (1.26)$$

Definition 1.3.6 (Hyperbolic Functions). We define the hyperbolic functions as follows, given $z = iy$

$$\begin{aligned} \sinh(y) &= -i \sin(iy) \\ \cosh(y) &= \cos(iy) \end{aligned} \quad (1.27)$$

For a general value of z , we define

$$\begin{aligned} \sinh(z) &= \frac{1}{2} (e^z - e^{-z}) \\ \cosh(z) &= \frac{1}{2} (e^z + e^{-z}) \end{aligned} \quad (1.28)$$

Theorem 1.3 (Trigonometric Identities). *Given $z, z_1, z_2 \in \mathbb{C}$ we have*

$$\begin{aligned}
\sin^2(z) + \cos^2(z) &= 1 \\
\sin(z_1 \pm z_2) &= \sin(z_1) \cos(z_2) \pm \cos(z_1) \sin(z_2) \\
\cos(z_1 \pm z_2) &= \cos(z_1) \cos(z_2) \mp \sin(z_1) \sin(z_2) \\
\sin(z) &= \sin(\Re(z)) \cosh(\Im(z)) + i \cos(\Re(z)) \sinh(\Im(z)) \\
\cos(z) &= \cos(\Re(z)) \cosh(\Im(z)) - i \sin(\Re(z)) \sinh(\Im(z)) \\
\|\sin(z)\|^2 &= \sin^2(\Re(x)) + \sinh^2(\Im(y)) \\
\|\cos(z)\|^2 &= \cos^2(\Re(x)) + \sinh^2(\Im(y)) \\
\cosh^2(z) - \sinh^2(z) &= 1 \\
\sinh(z_1 \pm z_2) &= \sinh(z_1) \cosh(z_2) \pm \cosh(z_1) \sinh(z_2) \\
\cos(z_1 \pm z_2) &= \cosh(z_1) \cosh(z_2) \pm \sinh(z_1) \sinh(z_2) \\
\sinh(z) &= \sinh(\Re(z)) \cos(\Im(z)) + i \cosh(\Re(z)) \sin(\Im(z)) \\
\cos(z) &= \cosh(\Re(z)) \cos(\Im(z)) + i \sinh(\Re(z)) \sin(\Im(z)) \\
\|\sinh(z)\|^2 &= \sinh^2(\Re(x)) + \sin^2(\Im(y)) \\
\|\cos(z)\|^2 &= \cosh^2(\Re(x)) + \sin^2(\Im(y))
\end{aligned} \tag{1.29}$$

Definition 1.3.7 (Inverse Trigonometric Functions). Given $z \in \mathbb{C}$ we define

$$\begin{aligned}
\arcsin(z) &= -i \log \left(iz + \sqrt{1 - z^2} \right) \\
\arccos(z) &= -i \log \left(z + i \sqrt{1 - z^2} \right) \\
\arctan(z) &= -\frac{i}{2} \log \left(\frac{i - z}{i + z} \right)
\end{aligned} \tag{1.30}$$

Definition 1.3.8 (Inverse Hyperbolic Functions). Given $z \in \mathbb{C}$ we define

$$\begin{aligned}
\operatorname{asinh}(z) &= \log \left(z + \sqrt{z^2 + 1} \right) \\
\operatorname{arccos}(z) &= \log \left(z + \sqrt{z^2 - 1} \right) \\
\operatorname{atanh}(z) &= \frac{1}{2} \log \left(\frac{1 + z}{1 - z} \right)
\end{aligned} \tag{1.31}$$

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. $d(x, y) = 0 \iff x = y$
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

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

When there won't be doubts 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

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

This set is said to be the *complementary set* of A .

It's also obvious that $A \cap A^c = \{\}$

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

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

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

$$\begin{aligned} x_0 \in A &\implies d(x, x_0) < r \\ \therefore \epsilon = r - d(x, x_0) &> 0 \end{aligned}$$

Then, by definition of open ball we have

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

Then, we can say that

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

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

$$\begin{aligned} x \in \bigcap_{i=1}^n A_i &\implies x \in A_i \\ \therefore \exists \epsilon_i : B_{\epsilon_i}^X(x) &\subset A_i \end{aligned}$$

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

$$B_\epsilon^X(x) \subset B_{\epsilon_i}^X(x) \implies B_\epsilon^X(x) \subset \bigcap_{i=1}^n A_i = A$$

And A is open

If we let $C = A^c$ we have that

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

For the last two we proceed as follows

$$\begin{aligned} x \in A_\alpha &\implies \exists \alpha_0 \in I : x \in A_{\alpha_0} \\ \therefore \exists \epsilon > 0 : B_\epsilon^X(x) &\subset A_{\alpha_0} \subset \bigcup_{\alpha \in I} A_\alpha = B \end{aligned}$$

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

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. $\partial A = \overline{A} \setminus A^\circ = \overline{A} \cup (A^\circ)^c$ 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. $A^\circ = \overline{(A^\circ)^c}^c$
4. $\overline{A} = [(A^\circ)^c]^\circ$
5. $(A \cap B)^\circ = A^\circ \cap B^\circ$
6. $\overline{A \cap B} = \overline{A} \cap \overline{B}$

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 \bar{A} is closed

For the third proposition, we have

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

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. $x \in \bar{A} \iff \forall \epsilon > 0 B_\epsilon(x) \cap A \neq \{\}$
3. $x \in \partial A \iff \forall \epsilon > 0 B_\epsilon(x) \cap A \neq \{\} \wedge B_\epsilon(x) \cap \bar{A} \neq \{\}$

Proof. [1] Let $I(A) := \{x \in X \mid \exists \epsilon > 0 : B_\epsilon(x) \subset A\}$.

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

But

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

[2] For the second proposition, we have

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

[3] For the last one, we have, taking into account the first two proofs

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

□

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

$$f : (X, d) \rightarrow (Y, \rho)$$

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

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

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

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

§ 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 $x_k \rightarrow x \in X$, iff

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

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 $x_k \rightarrow x \wedge x_k \rightarrow y$, 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 $\exists (x_k)_{k \geq 0} \in A : x_k \rightarrow x \in X$. The set of all adherent points of A is called $\text{ad}(A)$

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 $\exists (x_k)_{k \geq 0} : x_k \neq x \wedge x_k \rightarrow x \in \text{ad}(A)$

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

Proof. Let $Y = \text{ad}(A)$, 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 $d(x, x_n) < \frac{1}{n}$, therefore $x \in Y \implies x \in \text{ad}(A)$, and by definition

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

□

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

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 $\forall \epsilon > 0 \exists y \in A : d(x, y) < \epsilon$. 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 $x_k \rightarrow x$, 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 $x_k \rightarrow x \in X$

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

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

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

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

Going the opposite way we have that $\text{ad}(Y) = Y$ 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

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

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

Proof. (X, d) is compact, therefore

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

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

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

□

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

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

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

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

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

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 $T : X \rightarrow X$. T is said to be a *contraction mapping* or *contractor* if

$$\forall x, y \in X \exists q \in [0, 1) : d(T(x), T(y)) \leq qd(x, y) \quad (2.6)$$

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 $T : X \rightarrow X$. Then*

$$\exists! x^* \in X : T(x^*) = x^* \quad (2.7)$$

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

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

It's obvious that

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

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

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

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

$$\forall \epsilon > 0, \exists N \in \mathbb{N} : \forall n > N \ d(x_n, x) < \epsilon$$

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 $x_n \rightarrow x^* \in X$, but, by definition of convergence and limit, we have that by continuity

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

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

$$0 < d(T(x^*), T(y^*)) = d(x^*, y^*) > qd(x^*, y^*) \quad \text{!}$$

Therefore

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

And x^* 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. $\|u\| \geq 0 \quad \forall u \in \mathcal{V}$
2. $\|u\| = 0 \iff u = 0$

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

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

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

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

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

Setting $p = \infty$ we have that

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

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 $c \in \mathbb{F}$

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

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}^* .

§§ 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 $p_q \in \mathbb{N}^*$, where

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

Then

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

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

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

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

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \forall a, b \geq 0$$

Let

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

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

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

Substituting again the definitions of s, t we have

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

□

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

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad (2.12)$$

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

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

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

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

□

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

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

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 $f_n \rightarrow f$, if

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

It can be indicated also as follows

$$\lim_{n \rightarrow \infty} (f_n(x)) = f(x) \quad (3.3)$$

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

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

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

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

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.

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

□

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

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \int_a^b f(x) dx \quad (3.7)$$

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

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

Then, we have that

$$\forall n \geq N_\epsilon \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \|f_n - f\|_\infty (b-a) < \epsilon \quad (3.9)$$

□

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

$$1. \exists x_0 \in I : f_n(x_0) \rightarrow l$$

$$2. f'_n \Rightarrow g \quad \forall x \in I$$

Then

$$f_n(x) \Rightarrow f \implies \forall x \in I, f'(x) = \lim_{n \rightarrow \infty} f'_n(x) = g(x) \quad (3.10)$$

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

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(t) dt$$

Taking the limit we have

$$\begin{aligned} \lim_{n \rightarrow \infty} f_n(x) &= l + \int_{x_0}^x g(t) dt = f(x) \\ \therefore f'(t) &= g(t) \end{aligned}$$

But, we also have that

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

□

§ 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 $f_n : S \rightarrow \mathbb{C}$. We can define the *series of functions* as follows

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

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

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

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

This means that

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

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.

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

□

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

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

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

$$\sum_{k=0}^{\infty} f_k(x) \xrightarrow{A} s(x) \implies \sum_{k=0}^{\infty} |f_k(x)| \rightarrow s(x) \quad (3.15)$$

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

$$\sum_{k=0}^{\infty} f_k(x) \xrightarrow{A} s(x) \implies \sum_{k=0}^{\infty} f_k(x) \rightarrow s(x) \quad (3.16)$$

Proof. Let

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

□

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

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

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

Proposition 10. Let

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

Then

1. $f_n(x) \in C(S) \wedge s_k(x) \rightrightarrows s(x) \implies s(x) \in C(S)$
2. $f_n(x) \in C(S), s_k(x) \rightrightarrows s(x) \implies \int s(x) dx = \lim_{k \rightarrow \infty} \int s_k(x) dx$
3. $s_k(x) \xrightarrow{A} s(x) \implies s_k(x) \rightarrow s(x)$
4. $s_k(x) \rightrightarrows s(x) \implies s_k(x) \xrightarrow{A} s(x)$
5. $s_k(x) \xrightarrow{T} s(x) \implies s_k(x) \rightrightarrows s(x)$

§§ 3.2.1 Power Series and Convergence Tests

Theorem 3.6 (Weierstrass Test). Let $(f_n) : (S, d) \rightarrow \mathbb{C}$ a sequence of functions.

If we have that

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

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

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

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

$$\sum_{k=0}^{\infty} z^k \quad (3.18)$$

We can expand it as follows

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

Taking the limit, we have, therefore

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

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* $R \in \mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$, with the *Cauchy-Hadamard criteria*

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} = \begin{cases} +\infty & \frac{1}{R} = 0 \\ l & 0 < \frac{1}{R} = l < \infty \\ 0 & \frac{1}{R} = +\infty \end{cases} \quad (3.21)$$

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

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

$$\frac{1}{R} = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \implies R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| \quad (3.22)$$

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

$$\int_0^R s(x) dx = \sum_{k=0}^{\infty} \int_0^R a_k x^k dx = \int_0^R \sum_{k=1}^{\infty} a_k x^k dz = \sum_{k=0}^{\infty} a_k \frac{R^{k+1}}{k+1} \quad (3.23)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

This sequence converges pointwise to the null sequence, since

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

§§ 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

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

Therefore, this space is defined as follows

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

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

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

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

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

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

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

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

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

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

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

We have that $\text{diam}(\overline{B_1}) = 2$, therefore this set is bounded and closed by definition.

Take the *canonical sequence of sequences* $((e)_k)_n$, defined as follows:

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

Therefore, $\forall n \neq m$

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

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

$$\ell_0(\mathbb{F}) := \{ (x)_n \in \mathbb{F}^{\mathbb{N}} \mid (x)_n \rightarrow 0 \} \quad (4.5)$$

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.

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

□

§§ 4.1.3 $\ell^p(\mathbb{F})$ Spaces

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

$$\ell^p(\mathbb{F}) := \{ (x)_n \in \mathbb{F}^{\mathbb{N}} \mid \|(x)_n\|_p^p = M \in \mathbb{F} \} \quad (4.6)$$

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) $\forall (x)_n, (y)_n \in \ell^p(\mathbb{F}), (x)_n + (y)_n = (x + y)_n \in \ell^p(\mathbb{F})$

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

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

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

□

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 $(y)_n \rightarrow 0$, therefore $(x)_n \rightarrow 0$ 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

$$\ell_f(\mathbb{F}) := \{ (x)_n \in \mathbb{F}^{\mathbb{N}} \mid (x)_n = 0 \ \forall n > N \in \mathbb{N} \} \quad (4.7)$$

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 $\mathbb{R} \rightarrow \mathbb{R}$, or $\mathbb{R}; \mathbb{C}$, i.e. functions $\mathbb{R} \rightarrow \mathbb{C}$.

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

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

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

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

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

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 $\|\cdot\|_u = \|\cdot\|_\infty$ \square

Remark. Take $f_n \in C_b(\mathbb{F})$ a sequence of functions. The uniform convergence of this sequence means that $f_n \rightarrow f$ 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

$$\|f\|_p := \sqrt[p]{\int_{\mathbb{F}} |f(x)|^p dx} \quad (4.10)$$

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

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

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

§§ 4.2.2 Function Spaces in \mathbb{R}^n

Definition 4.2.3 (Seminorm). A *seminorm* is an application $\|\cdot\|_{\alpha, \beta} : A \longrightarrow \mathbb{F}$ with A a function space and α, β multiindices, where

$$\|f\|_{\alpha, \beta} := \|x^\alpha \partial^\beta f\|_\infty = \sup_{x \in \mathbb{F}} |x^\alpha \partial^\beta f(x)| \quad (4.11)$$

Definition 4.2.4 (Schwartz Space). The space $\mathcal{S}(\mathbb{R}^n)$ is called the *Schwartz space*, and it's defined as follows

$$\mathcal{S}(\mathbb{R}^n) := \left\{ f \in C^\infty(\mathbb{R}^n) \mid \|f\|_{\alpha, \beta} < \infty, \alpha, \beta \text{ multiindices} \right\} \quad (4.12)$$

Example 4.2.1. Taken $p(x) \in \mathbb{R}[x]$ a polynomial, a common example of functions $f(x) \in \mathcal{S}(\mathbb{R})$ is the following.

$$f(x) = p(x)e^{-a|x|^{2n}} \quad (4.13)$$

With $a > 0$, $n \in \mathbb{N}$

Theorem 4.2. A function $f \in C^\infty(\mathbb{R}^n)$ is in $\mathcal{S}(\mathbb{R}^n)$ if $\forall \beta$ multiindex, $\forall a > 0 \exists C_{\alpha,\beta}$ such that

$$\|\partial^\beta f(x)\| \leq \frac{C_{\alpha,\beta}}{(1 + \|x\|^2)^{\frac{\alpha}{2}}} \quad \forall x \in \mathbb{R}^n \quad (4.14)$$

Proof. Taken $n = 1$ and $f \in C^\infty(\mathbb{R})$, then

$$|x^j \partial^k f(x)| = |x|^j |\partial^k f(x)| \leq \frac{C_{j,k} |x|^j}{(1 + x^2)^{\frac{j}{2}}} \leq C_{j,k} \quad \forall x \in \mathbb{R}$$

Therefore

$$\|f\|_{j,k} \leq C_{j,k} < \infty \implies f \in \mathcal{S}(\mathbb{R})$$

Taken $f \in \mathcal{S}(\mathbb{R})$ we have that, if $|x| \geq 1$

$$(1 + x^2)^{\frac{a}{2}} \leq 2^{\frac{a}{2}} |x|^a$$

Taken $j = \lceil a \rceil$

$$|\partial^k f(x)| = \frac{|x^j \partial^k f(x)|}{|x|^j} \leq \frac{\|f\|_{j,k}}{|x|^a} \leq \frac{2^{\frac{a}{2}} \|f\|_{j,k}}{(1 + x^2)^{\frac{a}{2}}} \leq \frac{2^{\frac{a}{2}} \|f\|_{0,k}}{(1 + x^2)^{\frac{a}{2}}} \quad |x| < 1$$

Taken $C_{j,k} = 2^{\frac{a}{2}} \max \|f\|_{\lceil a \rceil, k}, \|f\|_{0,k}$ the assert is proven □

Definition 4.2.5 (Space of Compact Support Function). Given a function with compact support f , we define the space of compact functions $C_c^\infty(\mathbb{R}^n)$ as the space of all such functions. We have obviously that $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$

Theorem 4.3. Both $C_c^\infty(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are dense in $(C_p(\mathbb{R}^n), \|\cdot\|_p)$

Theorem 4.4 (Other Function Spaces). 1. $C(\mathbb{R})$ Space of continuous functions

2. $\mathbb{R}[x]$ Space of real polynomials

3. $C^k(\mathbb{R})$ Space of continuous k -derivable functions

4. $C_c^k(\mathbb{R})$ Space of functions $f \in C^k(\mathbb{R})$ with compact support

5. $C^\infty(\mathbb{R})$ Space of infinitely differentiable (smooth) functions

6. $C_0(\mathbb{R})$ Space of smooth functions with $\lim_{x \rightarrow \pm\infty} f(x) = 0$

7. $C_c^\infty(\mathbb{R})$ Space of smooth functions with compact support
 We have the obvious inclusions

$$\begin{aligned} \ell_f &\subset \ell_p \subset \cdots \subset \ell_0 \subset \mathbb{R}^\infty \\ C_c(\mathbb{R}) &\subset C_0(\mathbb{R}) \subset \cdots \subset C_p(\mathbb{R}) \subset C(\mathbb{R}) \\ \mathbb{R}[x] &\subset C^\infty(\mathbb{R}) \subset \cdots \subset C(\mathbb{R}) \\ C_c^\infty(\mathbb{R}) &\subset C_0^\infty(\mathbb{R}) \subset C^\infty(\mathbb{R}) \end{aligned}$$

Definition 4.2.6 (Banach Space). Given all previous infinite dimensional spaces, we define a *Banach space* as a space which is complete with respect to some norm $\|\cdot\|$

Example 4.2.2. Example of Banach spaces are the following

1. $(\ell_\infty, \|\cdot\|_\infty)$
2. $(\ell_p, \|\cdot\|_p)$, $p \geq 1$
3. $(\ell_0, \|\cdot\|_\infty)$
4. $(C_b(\mathbb{R}), \|\cdot\|_u)$

All the not listed spaces are not complete with respect to p -norms or uniform norms

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}^* be the dual space associated with \mathcal{V} . Then we can write the elements $v \in \mathcal{V}$, $\varphi \in \mathcal{V}^*$ with respect to some basis as follows

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

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 $i = 1, \dots, \dim(\mathcal{V})$.

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

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

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}^* .

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_j^i and b_j^i be two $n \times n$ matrices over the ordered field \mathbb{F} . The multiplication of these two matrices will simply be

$$c_j^i = a_k^i b_j^k \quad (5.5)$$

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

$$\text{tr}(a) = a_i^i \quad (5.6)$$

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

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

In n dimensions, it becomes

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

It’s obvious by definition that this weird entity is completely antisymmetrical 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_j^i

$$\det(a) = \epsilon_{i_1 \dots i_n} a^{1i_1} a^{2i_2} \dots a^{ni_n} = \epsilon_{i_1 \dots i_n} g^{ji_1} a_j^1 g^{ki_2} a_k^2 \dots g^{li_n} a_{i_n}^n \quad (5.9)$$

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

$$\mathbf{c} = \mathbf{v} \times \mathbf{w} \rightarrow c^i = g^{ij} \epsilon_{jkl} v^k w^l \quad (5.10)$$

(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 μ, ν, \dots be our indices and i, j, k, \dots our labels, we can write, for a matrix A_ν^μ

$$\det(A) = A = \epsilon_{\mu_1 \dots \mu_n} g^{\nu\mu_1} g^{\sigma\mu_2} \dots g^{\zeta\mu_n} A_\nu^1 A_\sigma^2 \dots A_\zeta^n \quad (5.11)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

□

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. $a_\nu^\mu b_{\mu\gamma\delta} = c_{\nu\gamma\delta}$. $a_\nu^\mu b_{\mu\gamma\delta} \neq c_{\nu\delta}^\gamma$, $a_\nu^\mu b_{\mu\gamma\delta} \neq c_{\nu\gamma\sigma}$*

2. An index can be repeated only two times per factor and must be contracted diagonally. I.e. $a^\mu b_\mu f_\gamma^\delta = c_\gamma^\delta$ is defined correctly, $a_\mu b_\mu$, $a^\mu b^\mu$ or $a^\mu b_\mu f_\gamma^\mu$ 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

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

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

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

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

$$\mathbf{JF} \longrightarrow \partial_\nu F^\mu = \frac{\partial F^\mu}{\partial x^\nu} = F^{\mu}_{,\nu} \quad (5.20)$$

*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

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

¹It'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

$$\nabla^2 h \longrightarrow g^{\mu\nu} \partial_\nu \partial_\mu h = \partial^\mu \partial_\mu h = \frac{\partial^2 h}{\partial x^\mu \partial x_\mu} = h_{;\mu}{}^\mu \quad (5.22)$$

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 $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ where A is an open set

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

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

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

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

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

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 $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ where, if the components of the two vectors are defined as x^μ, y^μ , is defined as

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

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

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

Definition 5.2.5 (Curves in \mathbb{R}^n). A *curve* is an application $\varphi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$.

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

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

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

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

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

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

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 $\gamma^\mu : [a, b] \rightarrow \mathbb{R}^n$ 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

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

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

$$L_\gamma := \int_a^b \sqrt{\dot{\gamma}_\mu(t) \dot{\gamma}^\mu(t)} dt \quad (5.31)$$

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

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

Its length will be given from the following integral

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

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

$$\varphi^\mu(t) \rightarrow (t, f(t)) \quad (5.34)$$

Its length will be then calculated with the following integral

$$L_\varphi := \int_a^b \sqrt{1 + (\dot{f}(x))^2} dx \quad (5.35)$$

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

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

Then

$$ds = \sqrt{\dot{\varphi}_\mu(t) \dot{\varphi}^\mu(t)} dt \quad (5.37)$$

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

$$L_\varphi = \int_\varphi ds \quad (5.38)$$

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

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

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

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

Definition 5.2.13 (Simple Curve, Closed Curve). A *simple curve* is an injective application $\gamma : [a, b] \rightarrow \mathbb{R}^n$. 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 $\{\gamma\}^c$ is defined as follows

$$\{\gamma\}^c = \{\gamma\}^\circ \cup \text{extr}(\{\gamma\}) \quad (5.41)$$

Note that $\{\gamma\} \subset \mathbb{R}^2$ is the image of the application γ and $\text{extr}(\{\gamma\})$ 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

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

§ 5.3 Differentiability in \mathbb{R}^n

Definition 5.3.1 (Directional Derivative). Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \rightarrow \mathbb{R}$. 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

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

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

$$\frac{\partial f}{\partial x^\mu} = \partial_\mu f = \partial_{x^\mu} f \quad (5.44)$$

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

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

Where we define the function R as follows

$$R(p^\mu + h^\mu) = f(p^\mu + h^\mu) - (f(p^\mu) + a_\mu h^\mu) \quad (5.46)$$

This means that

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

Theorem 5.4 (Consequences of Differentiability). Let $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $f \in C(A)$

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

2. f differentiable in A implies f derivable in A

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

Then

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

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

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

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 $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, 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

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

For Lagrange, we will have

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

Therefore

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

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

□

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

$$\lim_{\sqrt{h_\mu h^\mu} \rightarrow 0} \frac{\|f^\mu(p^\nu + h^\nu) - f^\mu(p^\nu) - J_\nu^\mu h^\nu\|_\mu}{\sqrt{h_\mu h^\mu}} \quad (5.51)$$

Or, equivalently

$$f_\mu(p^\nu + h^\nu) = f^\mu(p^\nu) + J_\nu^\mu h^\nu + \mathcal{O}(\sqrt{h_\mu h^\mu}) \quad (5.52)$$

The then J_ν^μ 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

$$J_\nu^\mu(p^\sigma) = \partial_\nu f^\mu(p^\sigma) \quad (5.53)$$

Theorem 5.7 (Composite Derivation). *Let $f^\mu : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^k$ and $g^\nu : B \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^p$ 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*

$$\partial_\sigma h^\nu = \partial_\mu g^\nu(f^\mu) \partial_\sigma f^\mu \quad (5.54)$$

Since $\mu = 1, \dots, k$, $\nu = 1, \dots, p$, $\sigma = 1, \dots, n$ 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

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

Since f^μ is differentiable, we have

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

Then, we must prove that

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

But

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

And

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

Therefore

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

□

§ 5.4 Differentiability in \mathbb{C}

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

$$\left. \frac{df}{dz} \right|_a = f'(a) = \lim_{z \rightarrow a} \frac{f(z) - f(a)}{z - a} \quad (5.55)$$

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

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

Proof.

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

□

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

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 $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$ and $g : F \subset \mathbb{C} \rightarrow \mathbb{C}$, 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*

$$\left. \frac{d}{dz} (g \circ f) \right|_a = \left. \frac{dg}{dz} \right|_{f(a)} \left. \frac{df}{dz} \right|_a = g'(f(a))f'(a) \quad (5.56)$$

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

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

Thus, rewriting the function inside the limit

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

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*

$$\left. \frac{df^{-1}}{dw} \right|_{f(a)} = \frac{1}{f'(a)} \quad (5.57)$$

Proof. Since f is bijective and continuous we can write

$$\left. \frac{df^{-1}}{dw} \right|_{f(a)} = \lim_{w \rightarrow f(a)} \frac{f^{-1}(w) - f^{-1}(f(a))}{w - f(a)} = \lim_{z \rightarrow a} \frac{z - a}{f(z) - f(a)} = \frac{1}{f'(a)}$$

□

§§ 5.4.1 Holomorphic Functions

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

$$\forall z \in G \exists \frac{df}{dz} = f'(z) \quad (5.58)$$

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 $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$, where G is open and $f \in H(G)$. Then, if we write $z = x + iy$*

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

We have that the function is holomorphic if and only if

$$\begin{cases} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = 0 \\ \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = 0 \end{cases} \quad (5.60)$$

Alternatively, it can be written as follows

$$\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0 \quad (5.61)$$

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 $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$.

$$\begin{cases} \frac{\partial f}{\partial z} = \partial f(z) = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) f(z) \\ \frac{\partial f}{\partial \bar{z}} = \bar{\partial} f(z) = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) f(z) \end{cases} \quad (5.62)$$

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

$$\frac{\partial f}{\partial \bar{z}} = \bar{\partial} f(z) = 0 \quad (5.63)$$

Proof. Let $f(z) = u(x, y) + iv(x, y) : G \subset \mathbb{C} \rightarrow \mathbb{C}$ 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

$$\left. \frac{df}{dz} \right|_{z_0} = \lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

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

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

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

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

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

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

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

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

□

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

Definition 5.4.5 (Singular Point). Let $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$ 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 $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{C}$ 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. $\text{rank}(\partial_\nu r^\mu) = 2$

The image $\text{Im}(r^\mu) = \Sigma \subset \mathbb{R}^3$ is then defined by the following parametric equations

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

The third condition can be rewritten as follows

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

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 $\tilde{u}, \tilde{v} \in I \subset \mathbb{R}$

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

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

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

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

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

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

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

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

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

The line is contained inside the following plane

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

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

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

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

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

For a cartesian surface we have

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

Definition 5.5.5 (Implicit Surface). Let $F : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ be a function such that $F \in C^1(A)$, letting $\Sigma := \{x^\mu \in \mathbb{R}^3 \mid F(x^\mu) = 0\}$. 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

$$\partial_\mu F(x_0^\nu)(x^\mu - x_0^\mu) = 0 \quad (5.74)$$

Definition 5.5.6 (Metric Tensor). Let ds be the curviline coordinate of some curve γ^μ inside a regular surface Σ . Then we have that

$$s^\mu(t) = r^\mu(u(t), v(t)) \quad (5.75)$$

And therefore

$$ds^2 = dr_\mu dr^\mu = \partial_1 r^\mu \partial_1 r_\mu (dx^1)^2 + 2\partial_1 r^\mu \partial_2 r_\mu dx^1 dx^2 + \partial_2 r^\mu \partial_2 r_\mu (dx^2)^2 \quad (5.76)$$

In compact form, we can write

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (5.77)$$

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

$$g_{\mu\nu} = \partial_\mu r^\sigma \partial_\nu r_\sigma \quad (5.78)$$

Or, in matrix notation

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

In usual mathematical notation we have

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

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

$$ds^2 = E (dx^1)^2 + 2F dx^1 dx^2 + G (dx^2)^2 \quad (5.81)$$

§ 5.6 Optimization

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

$$\partial_\mu f(p^\nu) = 0 \quad (5.82)$$

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

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

□

Definition 5.6.1 (Hessian Matrix). Let $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, 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

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

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

$$\partial_{\mu\nu} f = \partial_{\nu\mu} f \quad (5.85)$$

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 critical point of the function $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $A_\nu^\mu \in M_{nn}(\mathbb{R})$, and $(A_k)_\nu^\mu$ be the reduced matrix with order $k \leq n$, then*

1. $\det_{\mu\nu}((A_k)_\nu^\mu) > 0 \implies A_\nu^\mu$ positive definite
2. $(-1)^k \det_{\mu\nu}((A_k)_\nu^\mu) > 0 \implies A_\nu^\mu$ negative definite
3. If $\det_{\mu\nu}((A_{2k})_\nu^\mu) < 0$ or if $\det_{\mu\nu}((A_{2k+1})_\nu^\mu) < 0 \wedge \det_{\mu\nu}((A_{2n+1})_\nu^\mu) > 0$ for $k \neq n$, then A_ν^μ 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*

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

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

$$f(p_{n_k}^\mu) \rightarrow f(p^\mu)$$

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 \square

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.*

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

Then

$$\exists x^\mu \in L : \min_L(f) = f(x^\mu) \quad (5.88)$$

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

$$\inf_L(f) = \lim_{n \rightarrow \infty} f(p_n^\mu) = +\infty \quad \nexists$$

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

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

$$\begin{aligned} \{x^\mu \in \mathbb{R}^n \mid f(x^\mu) < a \in \mathbb{R}\} \\ \{x^\mu \in \mathbb{R}^n \mid f(x^\mu) > b \in \mathbb{R}\} \end{aligned} \quad (5.89)$$

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

$$\begin{aligned} \{x^\mu \in \mathbb{R}^n \mid f(x^\mu) \leq a \in \mathbb{R}\} \\ \{x^\mu \in \mathbb{R}^n \mid f(x^\mu) \geq b \in \mathbb{R}\} \end{aligned} \quad (5.90)$$

Are closed sets

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

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

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

$$f(\lambda x^\mu + (1 - \lambda)y^\mu) \leq \lambda f(x^\mu) + (1 - \lambda)f(y^\mu) \quad \forall x^\mu, y^\mu \in A, \forall \lambda \in [0, 1] \quad (5.92)$$

The function f is also known as a *sublinear* function

Also, the set

$$E_f = \{(x^\mu, \lambda) \in A \times \mathbb{R} \mid f(x^\mu) \leq \lambda\} \quad (5.93)$$

Is convex

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

1. f convex in $A \implies f \in C(A)$
2. f differentiable in $A \implies f$ convex $\iff f(x^\mu) \geq f(p^\mu) + \langle \nabla f(p^\mu), x^\mu - p^\mu \rangle$
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 $A_\nu^\mu(x^\gamma) \in \mathcal{V} \rightarrow M_{mn}(\mathbb{F})$, where $\dim(\mathcal{V}) = n$. We can define a norm for this space as follows

$$\|A_\nu^\mu\|_\infty = \sqrt{m} \sqrt{\max_\mu \sup_{x^\gamma \in \mathcal{V}} A_\nu^{(\mu)} A_{(\mu)}^\nu(x^\gamma)} \quad (5.94)$$

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

$$\|f^\mu(x^\nu) - f^\mu(y^\nu)\|_\mu \leq \|\partial_\nu f^\mu\|_\infty \|x^\nu - y^\nu\|_\nu \quad (5.95)$$

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

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

Therefore

$$\begin{aligned} \|f^\mu(x^\nu) - f^\mu(y^\nu)\|_\mu &\leq \sqrt{m} \sqrt{\max_\mu \sup_\gamma (\partial^\nu f_{(\mu)} \partial_\nu f^{(\mu)})} \|x^\nu - y^\nu\|_\nu = \\ &= \|\partial_\nu f^\mu\|_\infty \|x^\nu - y^\nu\|_\nu \quad \forall x^\nu, y^\nu \in K \end{aligned}$$

□

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

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

Then

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

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

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

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

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

Now take $(X, d) = (C(I, J), \|\cdot\|_\infty)$, with $J = \overline{B}_\epsilon(y_0^\gamma)$, and define an application $H : X \rightarrow X$ 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 $\exists! g^\gamma(x^\nu) : f^\mu(x^\nu, g^\gamma(x^\nu)) = 0 \quad \forall (x^\nu, y^\gamma) \in I \times J$

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

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

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

Therefore

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

I.e. H is a contraction in $C(I, J)$.

Due to the differentiability of f^μ we can write

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

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 $\partial_{y^\gamma} f^\mu = \delta_\gamma^\mu$

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

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

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

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

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

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

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

$$\begin{aligned} \text{Arg max}_A(f) &:= \{x_1^\nu, \dots, x_m^\nu\} \\ \text{Arg min}_A(f) &:= \{y_1^\nu, \dots, y_k^\nu\} \end{aligned} \tag{5.97}$$

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 $\mathcal{M} = \{x^\mu \in A \mid g(x^\mu) = 0\}$ and let $x_0^\mu \in \mathcal{M} : \partial_{\mathcal{M}} u g(x_0^\mu) \neq 0$, then $x_0^\mu \in \text{Arg max}_{\mathcal{M}} f \vee \text{Arg min}_{\mathcal{M}} f$ 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

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

Or, that

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

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

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

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

$$\begin{aligned} \frac{d}{dt} f(\psi^\nu(0)) &= \partial_\mu f \dot{\phi}^\mu(0) + \partial_n f \dot{h}(\phi^\mu(0)) = \partial_\nu f(x_0^\mu) s^\nu \\ \frac{d}{dt} g(\psi^\nu(0)) &= \partial_\nu g(x_0^\mu) s^\nu \end{aligned}$$

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 $\mathcal{M} := \{x^\nu \in A \mid g(x^\nu) = 0\}$. Take $x_0^\nu \in \mathcal{M}$ such that*

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

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

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

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 $T : \mathcal{V}^k \rightarrow \mathbb{R}$ is called *multilinear* if $\forall i = 1, \dots, k, \forall a \in \mathbb{R}, \forall v, w \in \mathcal{V}$

$$T(v_1, \dots, av_i + w_i, \dots, v_k) = aT(v_1, \dots, v_i, \dots, v_k) + T(v_1, \dots, w_i, \dots, v_k) \quad (6.1)$$

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

$$T_{\mu_1 \dots \mu_k} \quad (6.2)$$

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

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

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

This product has the following properties

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

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

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

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}^* . I.e. let $\varphi^\mu \in \mathcal{V}^*$, then $\varphi^\mu : \mathcal{V} \rightarrow \mathbb{R}$.

It's easy to see how $\mathcal{V}^* = \mathcal{T}^1(\mathcal{V})$.

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

$$\mathcal{B}_\mathcal{T} := \{\varphi^{\mu_1} \otimes \dots \otimes \varphi^{\mu_k}, \forall i = 1, \dots, n\} \quad (6.6)$$

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

$$f^*T(v_{\mu_1}, \dots, v_{\mu_k}) = T(f_\nu^\mu v_{\mu_1}, \dots, f_\nu^\mu v_{\mu_k})$$

Theorem 6.3. If g is an inner product on \mathcal{V} (i.e. $g : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$, with the properties of an inner product), there is a basis $v_{\mu_1}, \dots, v_{\mu_n}$ 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 $f_\nu^\mu : \mathbb{R}^n \xrightarrow{\sim} \mathcal{V}$ such that

$$g(f_\nu^\mu x^\nu, f_\nu^\mu y^\nu) = x_\mu y^\mu = g_{\mu\nu} x^\mu y^\nu \quad (6.7)$$

I.e.

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

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

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

Or, compactly

$$\begin{aligned} \omega_{\mu\dots\nu\dots\gamma\dots\sigma} &= -\omega_{\mu\dots\gamma\dots\nu\dots\sigma} \\ \omega_{\mu\dots\nu\dots\nu\dots\gamma} &= 0 \end{aligned} \quad (6.10)$$

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 $\text{Alt} : \mathcal{T}^k(\mathcal{V}) \rightarrow \Lambda^k(\mathcal{V})$ as follows

$$\text{Alt}(T)(v_1^\mu, \dots, v_k^\mu) = \frac{1}{k!} \sum_{\sigma \in \Sigma_k} \text{sgn}(\sigma) T(v_{\sigma(1)}^\mu, \dots, v_{\sigma(k)}^\mu) \quad (6.11)$$

With $\sigma = (i, j)$ a permutation and Σ_k the set of all permutations of natural numbers $1, \dots, k$. 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.

$$\text{Alt}(T_{\mu_1 \dots \mu_k}) = \frac{1}{k!} T_{[\mu_1 \dots \mu_k]} \quad (6.12)$$

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

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

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 $\mathcal{V} \times \cdots \times \mathcal{V} \times \mathcal{V}^* \times \cdots \times \mathcal{V}^*$, we can define the space $\mathcal{T}^k(\mathcal{V} \times \mathcal{V}^*) = \mathcal{W}$. 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

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

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

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

We can define also partially antisymmetric parts as follows

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

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 $\omega \wedge \eta \in \Lambda^{k+l}(\mathcal{V})$

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

With the following properties

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

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

Theorem 6.5. The set

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

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 $v_{\mu_1}, \dots, v_{\mu_n}$ be a basis for \mathcal{V} , and take $\omega \in \Lambda^n(\mathcal{V})$, then, if $w_\mu = a_\mu^\nu v_\nu$

$$\omega(w_{\mu_1} \cdots w_{\mu_n}) = \det_{\mu\nu}(a_\nu^\mu) \omega(v_{\mu_1}, \dots, v_{\mu_n}) \quad (6.18)$$

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

$$\omega_{\mu_1 \dots \mu_n} t^{\mu_1} \cdots t^{\mu_n} = \det_{\mu\nu}(a_\nu^\mu) \omega_{\nu_1 \dots \nu_n} t^{\nu_1} \cdots t^{\nu_n} \quad (6.19)$$

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

$$\eta_{\mu_1 \dots \mu_n} a_{\nu_1}^{\mu_1} a_{\nu_2}^{\mu_2} \cdots a_{\nu_n}^{\mu_n} = \omega_{\mu_1 \dots \mu_n} a_{\nu_1}^{\mu_1} \cdots a_{\nu_n}^{\mu_n}$$

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

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

□

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 $w^\mu = a_\nu^\mu v^\nu$ we have that the two bases belong to the same group iff $\det_{\mu\nu}(a_\nu^\mu) > 0$. 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)

$$\begin{aligned} &[v_\mu] \\ &-[w_\mu] \end{aligned}$$

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 $(v^{\mu_1}, \dots, v^{\mu_n})$, $(w^{\mu_1}, \dots, w^{\mu_n})$ that satisfy the *orthonormality condition* with respect to this scalar product

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

Then

$$\omega_{\mu_1 \dots \mu_n} v^{\mu_1} \cdots v^{\mu_n} = \omega_{\mu_1 \dots \mu_n} w^{\mu_1} \cdots w^{\mu_n} = \det_{\mu\nu}(a_\nu^\mu) = \pm 1$$

Where

$$w^\mu = a_\nu^\mu v^\nu$$

Therefore

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

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

Definition 6.1.8 (Cross Product). Let $v_1^\mu, \dots, v_n^\mu \in \mathbb{R}^{n+1}$ 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

$$z^\mu = v^{\nu_1} \times \dots \times v^{\nu_n} = \epsilon_{\nu_1 \dots \nu_n}^\mu v^{\nu_1} \dots v^{\nu_n}$$

§ 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

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

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

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

A function $f : T_p \mathbb{R}^n \rightarrow \mathbb{R}$ 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

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

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 $f : T_p \mathbb{R}^n \rightarrow \mathbb{R}$ 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

$$df(v^\nu) = \partial_\mu f(v^\nu) \quad (6.25)$$

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

$$\omega_{\mu_1 \dots \mu_k} = \omega_{\mu_1 \dots \mu_k} dx^{\mu_1} \dots dx^{\mu_k} \rightarrow \sum_{i_1 < \dots < i_k} \omega_{i_1 \dots i_k}(p) dx^{i_1} \wedge \dots \wedge dx^{i_k} \quad (6.26)$$

Basically, the vectors dx^μ are the *dual basis* with respect to the canonical basis $(e_\mu)_p$

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

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu = \partial_\mu f dx^\mu \quad (6.27)$$

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

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

Which, in turn, also induces a linear transformation $f^\star : \Lambda^k(T_{f(p)} \mathbb{R}^m) \rightarrow \Lambda^k(T_p \mathbb{R}^n)$, 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

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

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

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

1. $(f^\star)_\nu^\mu(dx^\nu) = df = \partial_\nu f^\mu dx^\nu$
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. $f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta$
5. $f^*(h dx^{[\mu_1 \dots \mu_n]}) = h \circ f \det_{\mu\nu}(\partial_\nu f^\mu) dx^{[\mu_1 \dots \mu_n]}$

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

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

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

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

Theorem 6.9 (Properties of d). 1. $d(\omega + \eta) = d\omega + d\eta$

2. $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$ for $\omega \in \Lambda^k(\mathcal{V})$, $\eta \in \Lambda^l(\mathcal{V})$

3. $dd\omega = d^2\omega = 0$

4. $f^*(d\omega) = d(f^*\omega)$

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

$$d\omega = 0 \quad (6.32)$$

It's called *exact* iff

$$\omega = d\eta \quad (6.33)$$

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 $\omega = d\rho$ for some differential form ρ , therefore

$$d\omega = dd\rho = d^2\rho = 0$$

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

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

$$\omega_\mu = p dx + q dy \quad (6.34)$$

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

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

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

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

Therefore

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

Which, expressed in terms of the basis vectors of $\Lambda^2(\mathbb{R}^2)$, $dx \wedge dy$, we get

$$d\omega = \frac{1}{2}(\partial_x q - \partial_y p) dx \wedge dy + \frac{1}{2}(\partial_y p - \partial_x q) dy \wedge dx = (\partial_x q - \partial_y p) dx \wedge dy \quad (6.35)$$

Therefore

$$d\omega = 0 \iff \partial_x q - \partial_y p = 0 \quad (6.36)$$

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

Definition 6.3.1 (Singular n -cube). A *singular n -cube* is an application $c : [0, 1]^n \rightarrow A \subset \mathbb{R}^n$. In general. A singular 0-cube is a function $f : \{0\} \rightarrow A$ and a singular 1-cube is a curve.

Definition 6.3.2 (Standard n -cube). We define a *standard n -cube* as a function $I^n : [0, 1]^n \rightarrow \mathbb{R}^n$ such that $I^n(x^\mu) = x^\mu$.

Definition 6.3.3 (Face). Given a standard n -cube I^n we define the (i, α) -face of the cube as

$$I^n_{(i, \alpha)} = (x^1, \dots, x^{i-1}, \alpha, x^i, \dots, x^{n-1}) \quad \alpha = 0, 1 \quad (6.37)$$

Definition 6.3.4 (Chain). Given n k -cubes c_i , we define a n -chain s as follows

$$s = \sum_{i=1}^n a_i c_i \quad a_i \in \mathbb{R} \quad (6.38)$$

Definition 6.3.5 (Boundary). Given an n -cube c_i we define the *boundary* as ∂c_i . For a standard n -cube we have

$$\partial I^n = \sum_{i=1}^n \sum_{\alpha=0,1} (-1)^{i+\alpha} I^n_{(i, \alpha)} \quad (6.39)$$

For a k -chain s we define

$$\partial s = \partial \left(\sum_i a_i c_i \right) = \sum_i a_i \partial c_i \quad (6.40)$$

Where ∂s is a $(k-1)$ -chain

Theorem 6.11. For a chain c , we have that $\partial \partial c = \partial^2 c = 0$

Definition 6.3.6 (Manifold). Given a set $M \subset \mathbb{R}^n$, it is said to be a *k-dimensional manifold* if $\forall x^\mu \in M$ we have that

1. $\exists U \subset \mathbb{R}^k$ open set $x^\mu \in U$ and $V \subset \mathbb{R}^n$ and φ a diffeomorphism such that $U \simeq V$ and $\varphi(U \cap M) = V \cap (\mathbb{R}^k \times \{0\})$, i.e. $U \cap M \simeq \mathbb{R}^k \cap \{0\}$
2. $\exists U \subset \mathbb{R}^k$ open and $W \subset \mathbb{R}^k$ open, $x^\mu \in U$ and $f : W \rightarrow \mathbb{R}^n$ a diffeomorphism
 - (a) $f(W) = M \cap U$
 - (b) $\text{rank}(f) = k \forall x^\mu \in W$
 - (c) $f^{-1} \in C(f(W))$

The function f is said to be a *coordinate system in M*

Definition 6.3.7 (Half Space). We define the *k-dimensional half space* $\mathbb{H}^k \subset \mathbb{R}^k$ as

$$\mathbb{H}^k := \{x^\mu \in \mathbb{R}^k \mid x^i \geq 0\} \quad (6.41)$$

Definition 6.3.8 (Manifold with Boundary). A *manifold with boundary* (MWB) is a manifold M such that, given a diffeomorphism h , an open set $U \supset M$ and an open set $V \subset \mathbb{R}^n$

$$h(U \cap V) = V \cap (\mathbb{H}^k \cap \{0\}) \quad (6.42)$$

The set of all points that satisfy this forms the set ∂M called the *boundary of M*

Definition 6.3.9 (Tangent Space). Given a manifold M and a coordinate set f around $x^\mu \in M$, we define the *tangent space of M at $x^\mu \in M$* as follows

$$f : W \subset \mathbb{R}^k \rightarrow \mathbb{R}^n \implies f_*(T_x \mathbb{R}^k) = T_x M \quad (6.43)$$

Definition 6.3.10 (Vector Field on a Manifold). Given a vector field f^μ we identify it as a vector field on a manifold M if $f^\mu(x^\nu) \in T_x M$. Analogously we define a *k-differential form*

7 Integral Analysis

§ 7.1 Measure Theory

Definition 7.1.1 (Lower and Upper Sums). We define the *upper* and *lower Riemann sums* as follows.

Let $f(x)$ be a function, then

$$\begin{cases} \mathcal{U}(f, x) := \sum_{i=1}^n \sup_{t \in [x_k, x_{k+1}]} (f(t)) \\ \mathcal{L}(f, x) := \sum_{i=1}^n \inf_{t \in [x_k, x_{k+1}]} (f(t)) \end{cases} \quad (7.1)$$

A function is said to be Riemann integrable if $\lim_{n \rightarrow \infty} (\mathcal{L}(f, x) - \mathcal{U}(f, x)) = 0$

Definition 7.1.2 (Set Function). Let A be a set. We define the following function $\mathbb{1}_A(x)$ as follows

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \quad (7.2)$$

Theorem 7.1. *The function $\mathbb{1}_{\mathbb{Q}}$ is not integrable over the set $[0, 1]$ with the usual definition of the integral (Riemann sums)*

Proof. Indicating the integral I as usual

$$I = \int_0^1 \mathbb{1}_{\mathbb{Q}}(x) \, dx$$

We see immediately that

$$\begin{aligned} \mathcal{U}(\mathbb{1}_{\mathbb{Q}}, x) &= 1 \\ \mathcal{L}(\mathbb{1}_{\mathbb{Q}}, x) &= 0 \end{aligned}$$

Therefore $\mathbb{1}_{\mathbb{Q}}(x)$ is not integrable in $[0, 1]$ (with the Riemann integral) □

Definition 7.1.3 (Measure). Let $A \subset X$ be a subset of a metric space. We define the measure of the set A , $\mu(A)$ as follows

$$\mu(A) = \int_X \mathbb{1}_A(x) \, dx \quad (7.3)$$

Basically, what we did before, was demonstrating that the set $\mathbb{Q} \cap [0, 1]$ is not measurable in the Riemann integration theory. This is commonly indicated with saying that the set $\mathbb{Q} \cap [0, 1]$ is *not Jordan measurable*.

For clarity, let K be some measure theory. We will say that a set is *K-measurable* if the following calculation exists

$$\mu_K(A) = \int_X \mathbb{1}_A(x) dx \quad (7.4)$$

Definition 7.1.4 (Algebra). Let $X \neq \{\}$ be a set. An *algebra* \mathcal{A} over X is a collection of subsets of X such that

1. $\{\} \in \mathcal{A}$
2. $X \in \mathcal{A}$
3. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
4. $A_1, \dots, A_n \in \mathcal{A} \implies \bigcup_{i=1}^n A_i, \bigcap_{i=1}^n A_i \in \mathcal{A}$

Example 7.1.1 (Simple Set Algebra). Let $X = \mathbb{R}^2$ and call R the set of all rectangles $I_i \subset \mathbb{R}^* \times \mathbb{R}^*$, where $\mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$. It's easy to see that this is not an algebra, since, by taking $[0, 1] \in R$, we have that $[0, 1]^c \notin R$, hence it cannot be an algebra.

But, taken \mathcal{S} as follows

$$\mathcal{S} := \left\{ A \subset \mathbb{R}^2 \mid A = \bigcup_{i=1}^n I_i \quad I_i \in R \right\}$$

We can see easily, using De Morgan law, that \mathcal{S} is an algebra.

§§ 7.1.1 Jordan Measure

Definition 7.1.5 (Disjoint Union). Taken two sets A, B , we define their *disjoint union* the binary operation $A \sqcup B$ as follows

$$A \sqcup B := A \cup B \setminus A \cap B \quad (7.5)$$

Definition 7.1.6 (Simple Set). A set A is a *simple set* iff, for some $R_i \in \mathcal{S}$, we have

$$A = \bigcup_{i=1}^n R_i$$

Definition 7.1.7 (Measure of a Simple Set). Let A be a simple set, the *Jordan measure* of a simple set is given by the sum of the measure of the rectangles, i.e. the “area” of A is given by the sum of the area of each rectangle R_i

$$\mu_J(A) = \sum_{i=1}^n \mu_J(R_i) \quad (7.6)$$

Definition 7.1.8 (External and Internal Measure). We define the external measure $\bar{\mu}_J$ and the internal measure $\underline{\mu}_J$ as follows.

Taken a limited set B and a simple set A we have

$$\begin{aligned} \bar{\mu}_J(B) &= \inf\{\mu_J(A) \mid B \subset A\} \\ \underline{\mu}_J(B) &= \sup\{\mu_J(A) \mid A \subset B\} \end{aligned} \quad (7.7)$$

A set is said to be *Jordan measurable* iff

$$\bar{\mu}_J(B) = \underline{\mu}_J(B) = \mu_J(B)$$

Remark (A Non Measurable Set). A good example for showing that the Jordan measure is the set we were trying to measure, the set $\mathbb{Q} \cap [0, 1]$. We can easily see that

$$\begin{aligned}\bar{\mu}_J(\mathbb{Q} \cap [0, 1]) &= 1 \\ \underline{\mu}_J(\mathbb{Q} \cap [0, 1]) &= 0\end{aligned}$$

Therefore it's not Jordan measurable.

From this we can jump to a new definition of measure, which is the *Lebesgue measure* where instead of covering $\mathbb{Q} \cap [0, 1]$ with a *finite* number of simple sets, we use sets which are formed from the union of *countable infinite* simple sets.

We can define

$$\mathbb{Q} \cap [0, 1] := \{q_1, q_2, \dots\}$$

We then take $\epsilon > 0$ and define the following set

$$A = \bigcup_{n=1}^{\infty} \left[q_n - \frac{\epsilon}{2^n}, q_n + \frac{\epsilon}{2^n} \right]$$

We have that

$$\mu(A) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n-1}} = 2\epsilon$$

But $\bar{\mu}(\mathbb{Q} \cap [0, 1]) \leq \mu(A) \leq 2\epsilon \rightarrow 0$, therefore $\mathbb{Q} \cap [0, 1]$ is measurable with $\mu(\mathbb{Q} \cap [0, 1]) = 0$

§§ 7.1.2 Lebesgue Measure

Definition 7.1.9 (σ -Algebras and Borel Spaces). Given a non empty set X a σ -algebra on X is a collection of subsets \mathcal{F} such that

1. $\forall A \in \mathcal{F}, A \subset X$
2. Let $A_i \in \mathcal{F}, i \in \mathcal{I} : |\mathcal{I}| = \aleph_0$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

The couple (X, \mathcal{F}) is called a *Borel space* or also a *measurable space*

Definition 7.1.10 (Measure). Given a Borel space (X, \mathcal{F}) , we can define an application

$$\mu : \mathcal{F} \longrightarrow [0, \infty] = \mathbb{R}_+^* \quad (7.8)$$

Which satisfies the following properties

1. σ -additivity, given $A_i \in \mathcal{F}$ with $i \in I \subset \mathbb{N}, |I| \leq \aleph_0$, such that $A_n \cap A_k = \{\}$ for $n \neq k$

$$\mu \left(\bigsqcup_{i \in I} A_i \right) = \sum_{i \in I} \mu(A_i)$$

2. If $Y_j \subset X$, with $j \in J \subseteq \mathbb{N}$, $\mu(Y_j) < \infty$ then $X = \bigcup_{j=1}^{\infty} Y_j$

Definition 7.1.11 (Measure Space). A *measure space* is a triplet (X, \mathcal{F}, μ) with \mathcal{F} a σ -algebra and μ a measure.

Remark. The empty set has null measure.

Proof. Due to σ -additivity we have that

$$\mu(\{\}) = \mu(\{\} \cup \{\}) = \mu(\{\}) + \mu(\{\})$$

Therefore, $\mu(\{\}) = 0$ necessarily. \square

Definition 7.1.12 (Lebesgue Measure). Consider again $X = \mathbb{R}^2$ and \mathcal{S} the algebra of simple sets. The *external Lebesgue measure* of a set $B \subset \mathbb{R}^2$ is then defined as follows

$$\bar{\mu}_L(B) := \inf \left\{ \sum_{i=1}^{\infty} \text{Area}(R_i) \mid R_i \in \mathcal{S}, B \subset \bigcup_{i=1}^{\infty} R_i \right\} \quad (7.9)$$

The set B is said to be *Lebesgue measurable* iff, $\forall C \subset \mathbb{R}^2$

$$\bar{\mu}_L(C) = \bar{\mu}_L(C \cap B) + \bar{\mu}_L(C \setminus B) \quad (7.10)$$

If it's measurable, then, $\bar{\mu}_L(B) = \mu_L(B)$ and it's called the *Lebesgue measure* of the set.

In other words $\exists \epsilon > 0 : \exists A, C \subset \mathbb{R}^2$, with $A = A^\circ$, $C = \bar{C}$ such that

$$C \subset B \subset A \vee \bar{\mu}_L(A \setminus C) < \epsilon \quad (7.11)$$

Definition 7.1.13 (Borel Algebra). Let R be the set of all rectangles. The smallest σ -algebra containing R is called the *Borel algebra* and it's indicated as \mathcal{B}

Definition 7.1.14 (Lebesgue Algebra). The set of (Lebesgue) measurable sets is a σ -algebra, which we will indicate as \mathcal{L} . In particular, we have that, if I is a rectangle, $I \in \mathcal{L}$.

If we add the fact that in \mathcal{B} there are null measure sets which have subsets which aren't part of \mathcal{B} , we end up with the conclusion that $\mathcal{B} \subset \mathcal{L}$

Definition 7.1.15 (Null Measure Sets). A set with null measure is a set $X \subset \mathcal{F}$ such that

$$\mu(X) = 0 \quad (7.12)$$

Where μ is a measure function.

It's obvious that sets formed by a single point have null measure.

I.e take a set $A = \{a\}$, then it can be seen as a rectangle with 0 area, and therefore

$$\mu(\{a\}) = 0 \quad (7.13)$$

Theorem 7.2. Every set such that $|A| = \aleph_0$ has null measure

Corollary 7.1.1. Every line in \mathbb{R}^2 has null measure

Proof. Take the set $A = \{a_1, a_2, a_3, \dots\}$. Then, due to σ -additivity, we have

$$\mu(\{a_1, a_2, a_3, \dots\}) = \mu\left(\bigsqcup_{k=1}^{\infty} \{a_k\}\right) = \sum_{k=1}^{\infty} \mu(\{a_k\}) = 0 \quad (7.14)$$

For the corollary, it's obvious if the line is thought as a rectangle in \mathbb{R}^2 with null area \square

§ 7.2 Integration

Definition 7.2.1 (Measurable Function). Given a Borel space (X, \mathcal{F}) a *measurable function* is a function $f : X \rightarrow \mathbb{F}$ such that, $\forall k \in \mathbb{F}$ the following set is measurable

$$I_f := \{k \in \mathbb{F} \mid f(x) < k\} \quad (7.15)$$

Or, in other words $I_f \in \mathcal{F}$, with \mathcal{F} the given σ -algebra of the Borel space. The space of all measurable functions on X will be identified as $\mathcal{M}(X)$

Theorem 7.3. Given a set $A \in \mathcal{F}$ with \mathcal{F} a σ -algebra, the function $\mathbb{1}_A(x)$ is measurable

Proof. We have that

$$I_{\mathbb{1}_A} = \begin{cases} A & k > 1 \\ \{\} & t \leq 1 \end{cases}$$

Therefore $I_{\mathbb{1}_A} \in \mathcal{F}$ and $\mathbb{1}_A(x)$ is measurable □

Definition 7.2.2 (Simple Measurable Function). Given a Borel space (X, \mathcal{F}) , a *simple measurable function* is a function $f : X \rightarrow \mathbb{F}$ which can be written as follows

$$f(x) = \sum_{k=1}^n c_k \mathbb{1}_{A_k}(x) \quad (7.16)$$

Where $A_k \in \mathcal{F}$, $c_k \in \mathbb{F}$ $0 \leq k \leq n$

Definition 7.2.3 (Integral). Given a measure space (X, \mathcal{F}, μ) and a simple function $f(x)$, we can define the *integral* of the function f with respect to the measure μ over the set X as follows

$$\int_X f(x) \mu(dx) = \sum_{k=1}^n c_k \mu(A_k) \quad (7.17)$$

For non negative functions we define the integral as follows

$$\int_X f(x) \mu(dx) = \sup \left\{ \int_X g(x) \mu(dx) \right\} \quad (7.18)$$

Where $g(x)$ is a simple measurable function such that $0 \leq g \leq f$.

If f assumes both negative and positive values we can write

$$f = f^+ - f^- \quad (7.19)$$

Where

$$\begin{cases} f^+ = \max f, 0 \\ f^- = \max -f, 0 \end{cases} \quad (7.20)$$

The integral, due to linearity, then will be

$$\int_X f(x) \mu(dx) = \int_X f^+(x) \mu(dx) - \int_X f^-(x) \mu(dx) \quad (7.21)$$

With the only constraint that the function $f(x)$ must be measurable in the σ -algebra \mathcal{F}

Definition 7.2.4 (\mathcal{L}^p spaces). With the previous definitions, we can define an *infinite dimensional function space* with the following properties

Given a measure space (X, \mathcal{F}, μ) we have the following definition

$$\mathcal{L}^p(X, \mathcal{F}, \mu) = \mathcal{L}^p(\mu) := \left\{ f : X \rightarrow \mathbb{F} \mid I_f \in \mathcal{F} \wedge \int_X |f|^p \mu(dx) < \infty \right\} \quad (7.22)$$

Defining the integral as an *operator* $\hat{K}_\mu[f]$ we can see easily that this is a vector spaces due to the properties of \hat{K}_μ .

It's easy to note that if the chosen σ -algebra and measure are the Lebesgue ones, then this integral is simply an extension of the usual Riemann integral.

It's important to note that a norm in $\mathcal{L}^p(\mu)$ can't be defined as an usual integral p -norm, since there are nonzero functions which have actually measure zero.

Definition 7.2.5 (Almost Everywhere Equality). Taken two functions $f, g \in \mathcal{L}^p(\mu)$ we say that these two function are *almost everywhere equal* if, given a set $A := \{x \in X \mid f(x) \neq g(x)\}$ has null measure. Therefore

$$f \sim g \iff \mu(A) = 0 \quad (7.23)$$

This equivalence relation creates equivalence classes of functions compatible with the vector space properties of $\mathcal{L}^p(\mu)$.

Definition 7.2.6 (L^p -Spaces). With the definition of the almost everywhere equality we can then define a quotient space as follows

$$L^p(\mu) = \mathcal{L}^p(\mu) \setminus \sim \quad (7.24)$$

This is a vector space, obviously, where the elements are the equivalence classes of functions $f \in \mathcal{L}^p(\mu)$, indicated as $[f]$.

If we define our σ -algebra and measure as the Lebesgue ones, this space is called the *Lebesgue space* $L^p(X)$, where an integral p -norm can be defined.

§§ 7.2.1 Lebesgue Integration

Note:

In this section the differential dx will actually indicate the Lebesgue measure $\mu(dx)$ used previously, unless stated otherwise.

Theorem 7.4. Let $f : E \rightarrow \mathbb{F}$ be a measurable function over E .

Given

$$F_{+\infty} = \{x \in E \mid f(x) = +\infty\} \wedge F_{-\infty} = \{x \in E \mid f(x) = -\infty\}$$

Assuming $E \subset X$, with (X, \mathcal{L}, μ) a Lebesgue measure space, we have that

$$\mu(F_{+\infty}) = \mu(F_{-\infty}) = 0$$

Proof. We can immediately say that

$$F_{+\infty} = \bigcap_{k \geq 0} F_k \in \mathcal{L}$$

Letting $r > 0$ we will indicate with $\mathbb{1}_r(x)$ the set function of the set $F_{+\infty} \cap B_r(0)$, therefore we have that

$$f^+(x) \geq k\mathbb{1}_r(x) \quad \forall k \in \mathbb{N}$$

Therefore

$$\mu(F_{+\infty} \cap B_r(0)) = \int \mathbb{1}_r(x) dx \leq \frac{1}{k} \int_E f^+(x) dx \longrightarrow 0$$

The proof is analogous for $F_{-\infty}$ □

Theorem 7.5. *Let (X, \mathcal{L}, μ) be a measure space, where \mathcal{L} is the Lebesgue σ -algebra and μ is the Lebesgue measure. Given a function $f \in L^1(X)$ we have that*

$$\int_X f(x) dx = 0 \iff f \sim 0 \quad (7.25)$$

Proof. Let $F_0 = \{x \in X \mid f(x) > 0\} = \bigcap_{k \geq 0} F_{1/k}$.
Since $f(x) > 1/k$, $\forall x \in F_{1/k}$, we have that, $\forall k \in \mathbb{N}$

$$\mu(F_{1/k}) \leq \int_X f(x) dx = 0$$

Through induction, we obtain that $\mu(F_0) = 0$ □

Theorem 7.6 (Monotone Convergence (B. Levi)). *Let $(f)_k$ be a sequence of measurable functions over a Borel space E , such that*

$$0 \leq f_1(x) \leq \dots \leq f_k(x) \leq \dots \quad \forall x \in F \subset E, \quad \mu(F) = 0$$

If $f_k(x) \rightarrow f(x)$, we have that

$$\int_E f(x) dx = \lim_{k \rightarrow \infty} \int_E f_k(x) dx \quad (7.26)$$

Or, in another notation

$$\int_E f_k(x) dx \rightarrow \int_E f(x) dx \quad (7.27)$$

Proof. Let $F_{0k} = \{0 < y < f_k(x)\}$ and $F_0 = \{0 < y < f(x)\}$ be two sets defined as seen. They are all measurable since $f_k(x), f(x)$ are measurable, and due to the monotony of $f_k(x)$ we have that

$$F_{01} \subset F_{02} \subset \dots \subset F_{0k} \subset \dots \quad \wedge \quad F_0 = \bigcup_{k=1}^{\infty} F_{0k}$$

Due to σ -additivity of the measure function, we have that F_0 is measurable, and that

$$\mu(F_0) = \sum_{k=1}^{\infty} \mu(F_{0k}) \quad \therefore \mu(F_0) = \lim_{k \rightarrow \infty} \mu(F_{0k})$$

□

Notation (For Almost All). We now introduce a new (unconventional) symbol in order to avoid writing too much, which would complicate the already difficult to understand theorems.

In order to indicate that we're picking *almost all* elements of a set we will use a new quantifier, which means that we're picking all elements of a null measure subset of the set in question. The quantifier in question will be the following

$$\forall^\dagger \quad (7.28)$$

Corollary 7.2.1. Let $f_k(x)$ be a sequence of non-negative measurable functions over a measurable set E , then $\forall^\dagger x \in E$

$$\int_E \sum_{k \geq 0} f_k(x) dx = \sum_{k \geq 0} \int_E f_k(x) dx \quad (7.29)$$

Theorem 7.7 (Fatou). Let $f_k(x)$ be a sequence of measurable functions over a measurable set E , such that $\forall^\dagger x \in E \exists \Phi(x)$ measurable: $f_k(x) > \Phi(x)$, then

$$\int_E \liminf_{k \rightarrow \infty} f_k(x) dx \leq \liminf_{k \rightarrow \infty} \int_E f_k(x) dx$$

Analogously happens with the \limsup of the sequence

Proof. Let $h_k(x) = f_k(x) - \Phi(x) \geq 0 \quad \forall^\dagger x \in E$ and $g_j(x) = \inf_{k \geq j} h_k(x)$, then $\forall k \geq j$ we have

$$\int_E g_j(x) dx \leq \int_E h_k(x) dx$$

It's also (obviously) true taking the \limsup of the RHS, and for the theorem on the monotone convergence, we have that

$$\begin{aligned} \int_E \lim_{j \rightarrow \infty} g_j(x) dx &= \lim_{j \rightarrow \infty} \int_E g_j(x) dx \leq \int_E h_k(x) dx \\ \therefore \lim_{j \rightarrow \infty} g_j(x) &= \sup_j g_j(x) = \sup_j \inf_{k \geq j} h_k(x) = \liminf_{k \rightarrow \infty} h_k(x) \end{aligned}$$

□

Theorem 7.8 (Dominated Convergence (Lebesgue)). Let $h(x) \geq 0$ be a measurable function on the measurable set E such that for a sequence of measurable functions $f_k(x)$ we have that

$$|f_k(x)| \leq h(x) \quad \forall^\dagger x \in E$$

And

$$f(x) = \lim_{k \rightarrow \infty} f_k(x) \quad \forall^\dagger x \in E$$

Then

$$\int_E f(x) dx = \lim_{k \rightarrow \infty} \int_E f_k(x) dx$$

Proof. By definition we have that $-h(x) \leq f_k(x) \leq h(x) \quad \forall^\dagger x \in E$, and we can apply Fatou's theorem

$$\int_E f(x) dx \leq \liminf_{k \rightarrow \infty} \int_E f_k(x) dx \leq \limsup_{k \rightarrow \infty} \int_E f_k(x) dx \leq \int_E f(x) dx$$

□

Corollary 7.2.2. Let E be a measurable set such that $\mu(E) < \infty$ and let $f_k(x)$ be a sequence of functions in E such that $|f_k(x)| \leq M \quad \forall^\dagger x \in E$ and $f_k(x) \rightarrow f(x), \quad \forall^\dagger x \in E$. Then the theorem (7.8) is valid.

Example 7.2.1. Take the sequence of functions $f_k(x) = kxe^{-kx}$ over $E = [0, 1]$. We already know that $f_k(x) \rightarrow f(x) = 0$ for $x \in E$, but $f_k(x) \not\rightarrow f(x)$ in E . We have that

$$\sup_E f_k(x) = e^{-1} = h(x) \neq f(x)$$

We have that $h(x)$ is measurable in E and we can apply the theorem (7.8)

Definition 7.2.7 (Carathéodory Function). Let (X, \mathcal{L}, μ) be a measure space and $A \subset \mathbb{R}^n$. $f : X \times A \rightarrow \mathbb{R}$ is a *Carathéodory function* iff $f(x^\mu, a^\nu) \in C(A) \quad \forall a^\nu \in A$ and $f(x^\mu, a^\nu) \in \mathcal{M}(X) \quad \forall^\dagger x^\mu \in X$

Definition 7.2.8 (Locally Uniformly Integrably Bounded). Let $f : X \times A \rightarrow \mathbb{R}$ be a Carathéodory function. It's said to be *locally uniformly integrably bounded* if $\forall a^\nu \in A \exists h_{a^\nu} : X \rightarrow \mathbb{R}$ measurable, and $\exists B_\epsilon(a^\nu) \subset A$, such that

$$\forall y^\nu \in B_\epsilon(x^\mu) \quad |f(x^\mu, y^\nu)| \leq h_{a^\nu}(x^\mu)$$

Note that if μ is a finite measure, then f bounded $\implies f$ locally uniformly integrably bounded or LUIB.

Theorem 7.9 (Leibniz's Derivation Rule). Let (X, \mathcal{F}, μ) be a measure space and $A \subset \mathbb{R}^n$ an open set. If $f : X \times A \rightarrow \mathbb{R}$ is a LUIB Carathéodory function we can define

$$g(a^\mu) = \int_X f(x^\nu, a^\mu) d\mu(x^\sigma) \in C(A)$$

Then

$$\partial_{x^\mu} f(x^\nu, a^\sigma) \in C(A)$$

Is LUIB, and therefore

$$g(a^\mu) \in C^1(A)$$

And

$$\partial_\mu g = \int_X \partial_{a^\mu} f(a^\nu, x^\sigma) d\mu(x^\gamma)$$

In other terms

$$\partial_{a^\mu} \int_X f(a^\nu, x^\sigma) d\mu(x^\gamma) = \int_X \partial_{a^\mu} f(a^\nu, x^\sigma) d\mu(x^\gamma) \quad (7.30)$$

Proof. Since f is a LUIB Carathéodory function we have that $\exists h_{a^\mu}(x^\nu) : X \rightarrow \mathbb{R}$ and $B_\epsilon(a^\mu) \subset A : \forall y^\mu \in B_\epsilon(a^\nu)$

$$|f(y^\mu, x^\nu)| \leq h_{a^\mu}(x^\nu)$$

Therefore

$$|g(a^\mu)| \leq \int_X h_{a^\mu}(x^\nu) d\mu(x^\sigma) < \infty$$

Now take a sequence $(a^\mu)_n : (a^\mu)_n \rightarrow a^\mu$, then $f \in C(A) \implies f(a_n^\mu, x^\nu) \rightarrow f(a^\mu, x^\nu) \quad \forall^\dagger x^\mu \in X, \forall a_n^\mu \in B_\epsilon(a^\mu)$

$$\therefore \exists N \in \mathbb{N} : \forall n \geq N \quad |f(a_n^\mu, x^\nu)| \leq h_{a^\mu}(x^\nu)$$

Then

$$g(a_n^\mu) = \int_X f(a_n^\mu, x^\nu) d\mu(x^\sigma) \rightarrow \int_X f(a^\mu, x^\nu) d\mu(x^\sigma) = g(a^\mu)$$

Since f is differentiable and its derivative is measurable, we have for the mean value theorem

$$f(a^\mu + te^\mu, x^\nu) - f(a^\mu, x^\nu) = t \partial_\mu f(\xi^\nu(t, x^\sigma), x^\gamma)$$

If $\xi^\mu(t, x^\nu) \in B_\epsilon(a^\mu)$ we have that

$$|t \partial_\mu f(\xi^\nu(t, x^\sigma), x^\gamma)| \leq h_{a^\mu}(x^\nu)$$

And therefore

$$\frac{g(a^\mu + te^\mu) - g(a^\mu)}{t} = \frac{1}{t} \int_X t \partial_\mu f(\xi^\nu(t, x^\sigma), x^\gamma) d\mu(x^\delta)$$

For $t \rightarrow 0$ $\partial_\mu f(\xi^\nu, x^\sigma) \rightarrow \partial_\mu f(a^\nu, x^\sigma)$, and the LHS is simply the gradient of g . Therefore for theorem (7.8)

$$\partial_\mu g(a^\nu) = \frac{\partial}{\partial a^\mu} \int_X f(a^\nu, x^\sigma) d\mu(x^\gamma) = \int_X \partial_\mu f(a^\nu, x^\sigma) d\mu(x^\gamma)$$

□

§ 7.3 Calculus of Integrals in \mathbb{R}^2 and \mathbb{R}^3

Theorem 7.10. Let $E \subset \mathbb{R}^2$ and $F \subset \mathbb{R}^3$. Define $E_x := \{y \in \mathbb{R} \mid (x, y) \in E\}$ the sections of E parallel to the y axis, then

$$\mu(E) = \int_{\mathbb{R}} \mu_1(E_x) dy \quad (7.31)$$

Where with μ_i we indicate the i -dimensional measure on \mathbb{R}^n .

Analogously, we define $F_z := \{(x, y) \in \mathbb{R}^2 \mid (x, y, z) \in F\}$ then

$$\mu(F) = \int_{\mathbb{R}} \mu_2(F_z) dz \quad (7.32)$$

If we define $F_{xy} := \{z \in \mathbb{R} \mid (x, y, z) \in F\}$ we have

$$\mu(F) = \iint_{\mathbb{R}^2} \mu_1(F_{xy}) dx dy \quad (7.33)$$

Proof. Let $A \subset \mathbb{R}^2$ open, and let $Y_k \subset \mathbb{R}^2$ be rectangles such that

$$Y_1 \subset Y_2 \subset Y_3 \subset \dots$$

$$A = \bigcup_{k=1}^{\infty} Y_k$$

Then, due to σ -additivity, we have

$$\mu_2(A) = \lim_{k \rightarrow \infty} \mu_2(Y_k) = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} \mu_1(Y_{kx}) dx$$

But

$$Y_{1x} \subset Y_{2x} \subset \dots$$

$$A_x = \bigcup_{k=1}^{\infty} Y_{kx}$$

Due to σ -additivity and the Beppo-Levi theorem we have that

$$\int_{\mathbb{R}} \mu_1(A_x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} \mu_1(Y_{kx}) dx$$

Let $E \subset \mathbb{R}^2$ be a measurable set. Define a sequence of compact sets K_i and a sequence of open sets A_j such that

$$K_1 \subset \dots \subset K_j \subset E \subset A_j \subset \dots \subset A_1$$

We have that $\lim_{j \rightarrow \infty} \mu_2(A_j) = \lim_{j \rightarrow \infty} \mu_2(K_j) = \mu_2(E)$ and that $K_{jx} \subset E \subset A_{jx}$. From the previous derivation we can write that

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}} (\mu_1(A_{jx}) - \mu_1(K_{jx})) dx = 0$$

Building a sequence of non-negative functions $f_j(x) = \mu_1(A_{jx}) - \mu_1(K_{jx})$ we have that $f_j(x) \leq f_{j-1}(x)$ and due to Beppo-Levi we have that

$$\lim_{j \rightarrow \infty} \int_{\mathbb{R}} f_j(x) dx = \int_{\mathbb{R}} \lim_{j \rightarrow \infty} f_j(x) dx$$

And therefore $\mu_1(K_{jx}) = \mu_1(A_{jx})$, and

$$\forall^\dagger x \in \mathbb{R} \quad \mu_2(K_j) = \int_{\mathbb{R}} \mu_1(K_{jx}) dx \leq \int_{\mathbb{R}} \overline{\mu_1}(E_x) dx \leq \int_{\mathbb{R}} \mu_1(A_{jx}) dx = \mu_2(A_j)$$

□

Theorem 7.11 (Fubini). *Let $f(x, y)$ be a measurable function in \mathbb{R}^2 , then*

1. $\forall^\dagger x \in \mathbb{R} \quad y \mapsto f(x, y)$ is measurable in \mathbb{R}
2. $g(x) = \int_{\mathbb{R}} f(x, y) dy$ is measurable in \mathbb{R}
3. $\iint_{\mathbb{R}^2} f(x, y) dx dy = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy$

Proof. Let $f(x, y) \geq 0$. Defining $F_0 := \{(x, y) \in E \times \mathbb{R} \mid 0 < z < f(x, y)\} \subset \mathbb{R}^3$, we have that F_0 is measurable, and

$$\mu_3(F_0) = \iint_{\mathbb{R}^2} f(x, y) dx dy$$

But F_{0x} is also measurable $\forall^\dagger x \in \mathbb{R}$ and therefore

$$\mu_3(F_0) = \int_{\mathbb{R}} \mu_2(F_{0x}) dx = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy$$

□

Theorem 7.12 (Tonelli). *Let $f(x, y)$ be a measurable function and $E \subset \mathbb{R}^2$ be a measurable set. If one of these integrals exists, the others also exist and have the same value*

$$\iint_{\mathbb{R}^2} f(x, y) \, dx \, dy \quad \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) \, dx \right) dy \quad \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) \, dy \right) dx$$

Theorem 7.13 (Integration Over Rectangles). *Let $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ be a rectangle, and $f(x, y)$ a measurable function over R . Then*

1. *If $\forall^\dagger x \in [a, b] \exists G(x) = \int_c^d f(x, y) \, dy$, the function $G(x)$ is measurable in $[a, b]$ and*

$$\iint_R f(x, y) \, dx \, dy = \int_a^b G(x) \, dx = \int_a^b \int_c^d f(x, y) \, dy \, dx$$

2. *If $\forall^\dagger y \in [c, d] \exists F(y) = \int_a^b f(x, y) \, dx$, the function $F(y)$ is measurable in $[c, d]$ and*

$$\int_{\mathbb{R}^2} f(x, y) \, dx \, dy = \int_c^d F(y) \, dy = \int_c^d \int_a^b f(x, y) \, dx \, dy$$

If both are true, then

$$\int_R f(x, y) \, dx \, dy = \int_a^b dx \int_c^d f(x, y) \, dy = \int_c^d dy \int_a^b f(x, y) \, dx \quad (7.34)$$

Definition 7.3.1 (Normal Set). A set $E \subset \mathbb{R}^2$ is said to be *normal* with respect to the x axis if

$$E = \{ (x, y) \in \mathbb{R}^2 \mid a \leq x \leq b \, \alpha(x) \leq y \leq \beta(x) \}$$

The definition is analogous for the other axes.

Theorem 7.14 (Integration over Normal Sets). *Let $E \subset \mathbb{R}^2$ be a normal set with respect to the x axis, and $f(x, y)$ is a measurable function over E . Then*

$$\int_E f(x, y) \, dx \, dy = \int_a^b dx \int_{\alpha(x)}^{\beta(x)} f(x, y) \, dy \quad (7.35)$$

Theorem 7.15 (Dirichlet Inversion Formula). *Take the triangle $T := \{ (x, y) \in \mathbb{R}^2 \mid a \leq y \leq x \leq b \}$. It can be considered normal with respect to both axes, and we can use the inversion formula*

$$\iint_T f(x, y) \, dx \, dy = \int_a^b dx \int_a^x f(x, y) \, dy = \int_a^b dy \int_y^b f(x, y) \, dx \quad (7.36)$$

Theorem 7.16 (Wire Integration). *Let $E \subset \mathbb{R}^3$ be a normal set with respect to the z axis. If $f(x, y, z)$ is measurable in E we have*

$$\iiint_E f(x, y, z) \, dx \, dy \, dz = \iint_D dx \, dy \int_{h(x, y)}^{g(x, y)} f(x, y, z) \, dz \quad (7.37)$$

This is called the wire integration formula

Theorem 7.17 (Section Integration). *Let $F \subset \mathbb{R}^3$ be a measurable set bounded by the planes $z = a$ and $z = b$ with $a < b$. Taken $z \in [a, b]$ we can define F_z and we have*

$$\iiint_F f(x, y, z) \, dx \, dy \, dz = \int_a^b dz \iint_{F_z} f(x, y, z) \, dx \, dy \quad (7.38)$$

This is called the section integration formula

Theorem 7.18 (Center of Mass). *Take a plane $E \subseteq \mathbb{R}^2$ with surface density $\rho(x, y) > 0$. We define the total mass M as follows*

$$M = \iint_E \rho(x, y) \, dx \, dy \quad (7.39)$$

The coordinates of the center of mass will be the following

$$\begin{aligned} x_G &= \frac{1}{M} \iint_E \rho(x, y) x \, dx \, dy \\ y_G &= \frac{1}{M} \iint_E \rho(x, y) y \, dx \, dy \end{aligned} \quad (7.40)$$

Theorem 7.19 (Moment of Inertia). *Taken the same plane E , we define the moment of inertia with respect to a line r as the following integral*

$$I_r = \iint_E \rho(x, y) (d(p^\mu, r))^2 \, dx \, dy \quad (7.41)$$

Where $d(p^\mu, r)$ is the distance function between the point (x, y) and the rotation axis r . Both formulas are easily generalizable in \mathbb{R}^3

§§ 7.3.1 Change of Variables

Definition 7.3.2 (Diffeomorphism). *Let $M, N \subset X$ be two subsets of a metric space X . The two sets are said to be *diffeomorphic* if $\exists f : M \xrightarrow{\sim} N$ an isomorphism such that $f \in C^1(M)$ and $f^{-1} \in C^1(N)$. The application f is called a *diffeomorphism*.*

Two diffeomorphic sets are indicated as follows

$$M \simeq N$$

Theorem 7.20. *Let $A, B \subset \mathbb{R}^n$ be two open sets and $\varphi^\mu : A \xrightarrow{\sim} B$ a diffeomorphism, such that*

$$\varphi^\mu(E) = F$$

If $f : E \subset B \rightarrow \mathbb{R}$ is measurable, we have that

$$\int_E f(y^\mu) \, dy^\mu = \int_{\varphi^{-1}(E)} f(\varphi^\mu(x^\nu)) \left| \det \partial_\mu \varphi^\nu \right| dx^\mu = \int_F f(\varphi^\mu) \left| \det \partial_\mu \varphi^\nu \right| dx^\mu$$

Theorem 7.21 (Change of Variables). *Let $\varphi^\mu : \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n$ be a diffeomorphism such that*

$$\varphi^\mu(x^\nu) = x^\mu \quad \forall \|x^\mu\|_\mu > 1$$

And $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a function such that $\text{supp } f = K \subset \mathbb{R}^n$ is a compact set. If f is measurable, we have that

$$\int_{\mathbb{R}^n} f(y^\mu) \, dy^\mu = \int_{\mathbb{R}^n} f(\varphi^\mu(x^\nu)) \left| \det \partial_\mu \varphi^\nu \right| dx^\mu \quad (7.42)$$

Proof. Take $n = 2$ without loss of generality. We can immediately write that

$$g(y^1, y^2) = \int_{-\infty}^{y^1} f(\eta, y^2) d\eta$$

Then, for the fundamental theorem of integral calculus

$$\partial_1 g(y^1, y^2) = f(y^1, y^2)$$

Taken $c \in \mathbb{R}$, $c > 1$: $K \subset Q = [-c, c] \times [-c, c]$, we have that $\varphi^\mu(x^\nu) = \delta_\nu^\mu \forall \|x^\mu\|_\mu > 1 \wedge f(x^\mu) = 0 \forall x^\mu \notin Q$.

Therefore $f(\varphi^\mu) = 0$ also and we have

$$\int_{\mathbb{R}^n} f(\varphi^\mu) \det_{\mu\nu} \partial_\mu \varphi^\nu dx^\gamma = \int_Q f(\varphi^\mu) \det_{\mu\nu} \partial_\mu \varphi^\nu dx^\gamma = \int_Q \partial_1 g(\varphi^\mu) \det_{\mu\nu} \partial_\mu \varphi^\nu dx^\gamma$$

But we have that

$$g(y^\mu) = 0 \quad \forall |y^1| \geq c \vee |y^1| < -c$$

Define the following matrix $H_{\mu\nu}$

$$H_{\mu\nu} = \begin{pmatrix} \partial_\mu g(\phi^\gamma) \\ \partial_\mu \varphi^2 \end{pmatrix}$$

Then we have that

$$\det_{\mu\nu} H_{\mu\nu} = \partial_1 g(\varphi^\mu) \det_{\mu\nu} \partial_\mu \varphi^\nu$$

Writing $g(\varphi^\mu) = G(x^\mu)$ we have

$$\det_{\mu\nu} H_{\mu\nu} = \partial_1 G \partial_2 \varphi^2 - \partial_2 G \partial_1 \varphi^2$$

Thanks to the integration formula (7.34) we can then write

$$\int_Q \det_{\mu\nu} H_{\mu\nu} dx^\gamma = \int_{-c}^c dx^2 \int_{-c}^c \partial_1 G \partial_2 \varphi^2 dx^\nu$$

Integrating by parts we get

$$\int_Q \det_{\mu\nu} H_{\mu\nu} dx^\gamma = G \partial_2 \varphi^2 \Big|_{-c}^c - \int_{-c}^c G \partial_{21}^2 \varphi^2 dx^1 - G \partial_1 \varphi^2 \Big|_{-c}^c - \int_{-c}^c G \partial_{12}^2 \varphi^2 dx^2$$

But $\forall x^\mu \in \partial Q \quad \varphi^\mu(x^\nu) = x^\mu \implies G(-c, x^2) = g(-c, x^2) = 0 \wedge G(c, x^2) = g(c, x^2)$

$$\therefore \int_Q \det_{\mu\nu} H_{\mu\nu} dx^\gamma = \int_Q f(x^\mu) dx^\gamma$$

□

Theorem 7.22 (Common Coordinate Transformation in \mathbb{R}^2 and \mathbb{R}^3). *1. Polar Coordinates*

$$\varphi^\mu(x^\nu) = \begin{cases} x(\rho, \theta) = \rho \cos \theta & \rho \in \mathbb{R}^+ \\ y(\rho, \theta) = \rho \sin \theta & \theta \in [0, 2\pi) \end{cases} \quad (7.43a)$$

$$\begin{aligned}\partial_\mu \varphi^\nu &= \begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{pmatrix} \\ \det_{\mu\nu} \partial_\mu \varphi^\nu &= \rho\end{aligned}\tag{7.43b}$$

2. Spherical Coordinates

$$\varphi^\mu(x^\nu) = \begin{cases} x(\rho, \theta, \phi) = \rho \sin \phi \cos \theta & \rho \in \mathbb{R}^+ \\ y(\rho, \theta, \phi) = \rho \sin \phi \sin \theta & \theta \in [0, 2\pi) \\ z(\rho, \theta, \phi) = \rho \cos \phi & \phi \in [0, \pi] \end{cases}\tag{7.44a}$$

$$\begin{aligned}\partial_\mu \varphi^\nu &= \begin{pmatrix} \sin \phi \cos \theta & -\rho \sin \phi \sin \theta & \rho \cos \phi \cos \theta \\ \sin \phi \sin \theta & \rho \sin \phi \cos \theta & \rho \cos \phi \sin \theta \\ \cos \phi & 0 & -\rho \sin \phi \end{pmatrix} \\ \det_{\mu\nu} \partial_\mu \varphi^\nu &= \rho^2 \sin \phi\end{aligned}\tag{7.44b}$$

3. Cylindrical Coordinates

$$\varphi^\mu(x^\nu) = \begin{cases} x(\rho, \theta, z) = \rho \cos \theta & \rho \in \mathbb{R}^+ \\ y(\rho, \theta, z) = \rho \sin \theta & \theta \in [0, 2\pi) \\ z(\rho, \theta, z) = z & z \in \mathbb{R} \end{cases}\tag{7.45a}$$

$$\det_{\mu\nu} \partial_\mu \varphi^\nu = \rho\tag{7.45b}$$

Definition 7.3.3 (Rotation Solids). Let $D \subset \mathbb{R}^2$ be a bounded measurable set contained in the half-plane $y = 0, x > 0$. Suppose we let D “pop up” into \mathbb{R}^3 through a rotation by an angle θ_0 around the z axis. What has been obtained is a *rotation solid* $E \subset \mathbb{R}^3$. We have that

$$\mu(E) = \iiint_E dx dy dz = \iint_D \int_0^{\theta_0} \rho d\rho d\theta dz = \theta_0 \iint_D \rho d\rho dz = \theta_0 \iint_D x dx dy\tag{7.46}$$

Or

$$\mu(E) = \theta_0 x_G \mu_2(D)$$

Theorem 7.23 (Guldino). *The measure of a rotation solid is given by the measure of the rotated figure times the circumference described by the center of mass of the solid. This is exactly the previous formula.*

§§ 7.3.2 Line Integrals

Definition 7.3.4 (Line Integral of the First Kind). Given a scalar field $f : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ and a smooth curve $\{\gamma\} \subset \mathbb{R}^3$, we define the *line integral of the first kind* as follows

$$\int_\gamma f ds = \int_a^b f(\gamma^\mu) \left\| \frac{d\gamma^\mu}{dt} \right\|_\mu dt\tag{7.47}$$

Theorem 7.24 (Center of Mass of a Curve). *Given a curve $\gamma^\mu : [a, b] \rightarrow \mathbb{R}^3$ with linear mass density $m : \{\gamma\} \rightarrow \mathbb{R}$, we define the total mass of γ as follows*

$$M = \int_{\gamma} m \, ds = \int_a^b m(\gamma^\mu) \left\| \frac{d\gamma^\mu}{dt} \right\|_{\mu} dt \quad (7.48)$$

The center of mass is then defined as follows

$$x_G^\mu = \frac{1}{M} \int_{\gamma} x^\mu m(x^\nu) \, ds \quad (7.49)$$

Definition 7.3.5 (Line Integral of the Second Kind). *Given a vector field $f^\mu : A \rightarrow \mathbb{R}^3$ and a smooth curve $\gamma^\mu : [a, b] \rightarrow A \subset \mathbb{R}^3$ we define the *line integral of the second kind* as follows*

$$\int_{\gamma} f^\mu T_\mu \, ds = \int_a^b f^\mu(\gamma^\nu) \frac{d\gamma_\mu}{dt} dt \quad (7.50)$$

Defining a differential form $\omega = f^\mu dx_\mu$ we can also see this integral as follows

$$\int_{\gamma} \omega = \int_{\gamma} f^\mu T_\mu \, ds \quad (7.51)$$

Where T^μ is the tangent vector of the curve

Definition 7.3.6 (Conservative Field). *Let $f^\mu : A \rightarrow \mathbb{R}^3$ be a vector field such that $f^\mu \in C^1(A)$ and A is open and connected. This field is said to be *conservative*, if $\forall x^\mu \in A$*

$$\exists U(x^\mu) \in C^2(A) : f^\mu = -\partial^\mu U \quad (7.52)$$

The function $U(x^\mu)$ is called the *potential* of the field.

Theorem 7.25 (Line Integral of a Conservative Field). *Given a conservative field $f^\mu : A \rightarrow \mathbb{R}^3$ and a smooth curve $\{\gamma\} \subset A$, $\gamma^\mu : [a, b] \rightarrow \mathbb{R}^3$ with A open and connected, we have that*

$$\int_{\gamma} f^\mu T_\mu \, ds = U(\gamma(a)) - U(\gamma(b)) \quad (7.53)$$

Where $U(x^\mu)$ is the potential of the vector field.

Definition 7.3.7 (Rotor). *Given a vector field $f^\mu : A \rightarrow \mathbb{R}^3$ with $f^\mu \in C^1(A)$, we define the *rotor* of the vector field as follows*

$$\text{rot}(f^\mu) = \epsilon_{\nu\gamma}^\mu \partial^\nu f^\gamma \quad (7.54)$$

Theorem 7.26. *Given f^μ a conservative vector field on an open connected set A , we have that*

$$\epsilon_{\nu\gamma}^\mu \partial^\nu f^\gamma = 0 \quad (7.55)$$

Alternatively, if $\gamma^\mu : [a, b] \rightarrow \mathbb{R}^3$ is the parameterization of a smooth closed curve, we have that

$$\oint_{\gamma} f^\mu T_\mu \, ds = 0 \quad (7.56)$$

§§ 7.3.3 Surface Integrals

Definition 7.3.8 (Area of a Surface). Given $r^\mu : K \subset \mathbb{R}^2 \rightarrow \Sigma \subset \mathbb{R}^3$ a smooth surface, we have that given its metric tensor $g_{\mu\nu}(u, v)$ we have that

$$\mu(\Sigma) = \int_{\Sigma} d\sigma = \iint_K \sqrt{\det g_{\mu\nu}} du dv = \iint_K \sqrt{EG - F^2} du dv \quad (7.57)$$

For a cartesian surface S we have that

$$\mu(S) = \int_S ds = \iint_K \sqrt{1 + (\|\partial_\mu f\|_\mu)^2} dx dy \quad (7.58)$$

Definition 7.3.9 (Rotation Surface). Given a smooth curve $\gamma^\mu : [a, b] \rightarrow \mathbb{R}^3$, the rotation of this curve around the z -axis generates a smooth surface Σ with the following parameterization

$$r^\mu(t, \theta) = \begin{cases} \gamma^1(t) \cos \theta \\ \gamma^2(t) \sin \theta \\ \gamma^3(t) \end{cases} \quad (t, \theta) \in [a, b] \times [0, \theta_0] \quad (7.59)$$

The area of a rotation surface is calculated as follows

$$\mu(\Sigma) = \theta_0 \int_a^b \gamma^1(t) \sqrt{\left(\frac{d\gamma^1}{dt}\right)^2 + \left(\frac{d\gamma^2}{dt}\right)^2} dt \quad (7.60)$$

Theorem 7.27 (Guldino II). Given Σ a smooth rotation surface defined as before, we have that its area will be

$$\mu(\Sigma) = \theta_0 \int_{\gamma} x^1 ds = \theta_0 x_G^1 L_{\gamma} \quad (7.61)$$

Where x_G^1 is the first coordinate of the center of mass of the curve, calculated as follows

$$x_G^1 = \frac{1}{L_{\gamma}} \int_{\gamma} x^1 ds$$

Definition 7.3.10 (Surface Integral). Given a smooth surface $\Sigma \subset \mathbb{R}^3$ with parameterization $r^\mu : K \rightarrow \Sigma$ and a scalar field $h : \mathbb{R}^3 \rightarrow \mathbb{R}$, we define the *surface integral* of h as follows

$$\int_{\Sigma} h(x^\mu) d\sigma = \iint_K h(r^\mu) \sqrt{\det g_{\mu\nu}} du dv \quad (7.62)$$

If Σ is a cartesian surface, we have

$$\int_{\Sigma} h(x^\mu) d\sigma = \iint_K h(x^1, x^2, f) \sqrt{1 + (\|\partial_\mu f\|_\mu)^2} dx dy \quad (7.63)$$

Definition 7.3.11 (Center of Mass of a Surface). Given a smooth surface Σ with parameterization $r^\mu(u, v)$ and mass density δ , we define its total mass as follows

$$M = \int_{\Sigma} \delta d\sigma \quad (7.64)$$

Its center of mass x_G^μ will be calculated as follows

$$x_G^\mu = \frac{1}{M} \int_{\Sigma} x^\mu \delta(x^\nu) d\sigma \quad (7.65)$$

Definition 7.3.12 (Moment of Inertia of a Surface). Given a smooth surface Σ with parameterization $r^\mu(u, v)$ and mass density δ we define its moment of inertia around an axis r , I , as the following integral

$$I = \int_{\Sigma} \delta(x^\mu) (d(p^\mu, r))^2 d\sigma \quad p^\mu \in \Sigma \quad (7.66)$$

Definition 7.3.13 (Orientable Surface). A smooth surface with parameterization $r^\mu : K \subset \mathbb{R}^2 \rightarrow \Sigma \subset \mathbb{R}^3$ is said to be *orientable* if $\forall \gamma : [a, b] \rightarrow \Sigma$ smooth closed curve, we have, given n^μ the normal vector of the surface

$$n^\mu(\gamma^\nu(a)) = n^\mu(\gamma^\nu(b)) \quad (7.67)$$

Another way of formulating it is

$$n^\mu(x^\nu) \in C(K) \quad (7.68)$$

Definition 7.3.14 (Boundary of a Surface). Given a smooth surface as before, we define the *boundary* $\partial\Sigma$ as follows

$$\partial\Sigma = \bar{\Sigma} \setminus \Sigma \quad (7.69)$$

Note how, given the parameterization r^μ , we have $r^\mu(\partial K) = \partial\Sigma$

Definition 7.3.15 (Closed Surface). A surface $\Sigma \subset \mathbb{R}^3$ is said to be *closed* iff $\partial\Sigma = \{\}$

Definition 7.3.16 (Flux). Given a vector field $f^\mu : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and a smooth orientable surface $\Sigma \subset A$, we define the *flux* of the vector field f^μ on the surface as follows

$$\Phi_\Sigma(f^\mu) = \int_{\Sigma} f^\mu n_\mu d\sigma = \iint_K f^\mu(r^\nu) \epsilon_{\mu\gamma\sigma} \partial_1 r^\gamma \partial_2 r^\sigma du dv \quad (7.70)$$

§ 7.4 Integration in \mathbb{C}

Definition 7.4.1 (Piecewise Continuous Function). Let $\gamma : [a, b] \rightarrow \mathbb{C}$ be a piecewise continuous curve such that $\{\gamma\} \subset D \subset \mathbb{C}$, and $f : D \rightarrow \mathbb{C}$, $f \in C(D)$. Then the function $(f \circ \gamma) \gamma'(t) : [a, b] \rightarrow \mathbb{C}$ is a *piecewise continuous function*

Definition 7.4.2 (Line Integral in \mathbb{C}). Let $\gamma : [a, b] \rightarrow D \subset \mathbb{C}$ be a piecewise continuous curve and $f : D \rightarrow \mathbb{C}$ a measurable function $f \in C(D)$.

We define the *line integral over γ* the result of the application of the integral operator $\hat{K}_\gamma[f]$, where

$$\hat{K}_\gamma[f] = \int_\gamma f(z) dz = \int_a^b (f \circ \gamma) \gamma'(t) dt \quad (7.71)$$

Where $\forall^\dagger z \in \{\gamma\}$ $f(z)$ is defined

Theorem 7.28 (Properties of the Line Integral). Let $z, w, t \in \mathbb{C}$, $f, g \in \mathcal{M}(\mathbb{C})$ and $\{\gamma\}, \{\eta\}, \{\kappa\}$ three smooth curves, then

1. $\hat{K}_\gamma[zf + wg] = z\hat{K}_\gamma[f] + w\hat{K}_\gamma[g]$
2. $\gamma \sim \eta \implies \hat{K}_\gamma[f] = \hat{K}_\eta[f]$

$$3. \gamma = \eta + \kappa \implies \hat{K}_\gamma[f] = \hat{K}_{\eta+\kappa}[f] = \hat{K}_\eta[f] + \hat{K}_\kappa[f]$$

$$4. \hat{K}_{\gamma+w}[f(z)] = \hat{K}_\gamma[f(z+w)]$$

Notation. If a measurable function $f(z)$ has the same value of the integral for different curves between two points $z_1, z_2 \in \mathbb{C}$, we will write directly

$$\int_\gamma f(z) dz = \int_{z_1}^{z_2} f(z) dz$$

Theorem 7.29 (Darboux Inequality). *Let $f : D \rightarrow \mathbb{C}$ be a measurable function and $\gamma : [a, b] \rightarrow \{\gamma\} \subset D \subseteq \mathbb{C}$ piecewise smooth. Then*

$$\left\| \int_\gamma f(z) dz \right\| \leq L_\gamma \sup_{z \in \{\gamma\}} \|f(z)\|$$

Proof. The proof is quite straightforward using the definition given for the line integral

$$\begin{aligned} \left\| \int_\gamma f(z) dz \right\| &= \left\| \int_a^b (f \circ \gamma) \gamma'(t) dt \right\| \leq \int_a^b \|(f \circ \gamma) \gamma'(t)\| dt \leq \\ &\leq \sup_{z \in \{\gamma\}} \|f(z)\| \int_a^b \|\gamma'(t)\| dt = L_\gamma \sup_{z \in \{\gamma\}} \|f(z)\| \end{aligned}$$

□

Definition 7.4.3 (Primitive). Let $f : D \rightarrow \mathbb{C}$ and $F : D \rightarrow \mathbb{C}$ be two functions and $D \subset \mathbb{C}$ an open and connected set. $F(z)$ is said to be the *primitive function* or *antiderivative* of f in D if

$$\frac{dF}{dz} = f(z) \quad \forall z \in D \quad (7.72)$$

Notation. Given a closed curve γ and a measurable function $f(z)$ we define the following notation

$$\int_\gamma f(z) dz = \oint_\gamma f(z) dz$$

Theorem 7.30 (Existence of the Primitive Function). *Let $f : D \rightarrow \mathbb{C}$ $f \in C(D)$ with $D \subset \mathbb{C}$ open and connected. Then these statements are equivalent*

1. $\exists F : D \rightarrow \mathbb{C} : F'(z) = f(z)$
2. $\forall z_1, z_2 \in D, \forall \{\gamma\} \subset D$ piecewise smooth $\int_\gamma f(z) dz = \int_{z_1}^{z_2} f(z) dz$
3. $\forall \gamma : [a, b] \rightarrow \{\gamma\} \subset D$ closed piecewise smooth $\oint_\gamma f(z) dz = 0$

Proof. 1 \implies 2. As with the hypothesis we have that $\exists F : D \rightarrow \mathbb{C} : F'(z) = f(z) \forall z \in D$. Given two points $z_1, z_2 \in D$ and taken a smooth curve $\gamma : [a, b] \rightarrow D : \gamma(a) = z_1 \wedge \gamma(b) = z_2$. Therefore

$$\int_\gamma f(z) dz = \int_a^b (f \circ \gamma) \gamma'(t) dt = \int_a^b (F' \circ \gamma) \gamma'(t) dt$$

The result of the integral is obviously $F(z_2) - F(z_1)$, therefore we can immediately write that, if

$$\exists F : D \longrightarrow \mathbb{C} : F'(z) = f(z) \implies \int_{\gamma} f(z) dz = \int_{z_1}^{z_2} f(z) dz$$

2 \implies 1 Taken a point $z_0 \in D$, any point $z \in D$ can be connected with a polygonal to z_0 since D is connected. The integral of f over this polygonal is obviously path-independent, hence we can define the following function

$$F(z) = \int_{z_0}^z f(w) dw$$

Since D is open we can define $\delta_z \in \mathbb{R}$, $\delta_z > 0 \wedge \exists B_{\delta_1}(z) \subset D$. Taken $\Delta z \in \mathbb{C} : \|\Delta z\| < \delta_1$ we have that

$$F(z + \Delta z) - F(z) = \int_z^{z+\Delta z} f(w) dw$$

Dividing by Δz and taking the limit as $\Delta z \rightarrow 0$ we have that using the Darboux inequality we get that

$$\left\| \frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) \right\| = \frac{1}{\|\Delta z\|} \left\| \int_z^{z+\Delta z} f(w) dw \right\| \leq \epsilon$$

2 \implies 3. Taken an arbitrary piecewise smooth curve γ and $z_1 \neq z_2 \in \{\gamma\}$. We can now find two curves such that $\gamma(t) = \gamma_1(t) - \gamma_2(t)$. Since the integral of f is path independent, we get

$$\int_{\gamma} f(z) dz = \int_{\gamma_1} f(z) dz - \int_{\gamma_2} f(z) dz = 0$$

3 \implies 2 is exactly as before but with the opposite reasoning. □

Example 7.4.1. Let's calculate the integral of functions $f_n(x) = z^{-n}$ $n \in \mathbb{N}$ for a closed simple piecewise smooth curve γ such that $0 \notin \{\gamma\}$.

For $n > 1$ we have that $f \in C(D)$ where $D = \mathbb{C} \setminus \{0\}$, and we have that

$$\int \frac{1}{z^n} dz = -\frac{z^{-(n-1)}}{n-1} + w \quad w \in \mathbb{C}$$

Therefore, for every closed simple piecewise smooth curve $\gamma : 0 \notin \{\gamma\}$ we have

$$\oint_{\gamma} \frac{1}{z^n} dz = 0$$

For $n = 1$ we still have that $f \in C(D)$ but $\nexists F(z) : D \longrightarrow \mathbb{C}$ primitive of $f_1(z)$, but there exists one in the domain G of holomorphy of the logarithm.

Although we have that $G \subset D$, and we can take a curve $\gamma : 0 \in \text{extr } \gamma$, and therefore $\{\gamma\} \subset G$ and we have that

$$\oint_{\gamma} \frac{1}{z} dz = 0$$

If we otherwise have $0 \in \gamma^\circ$ the integral is non-zero.

Take a branch of the logarithm σ and a curve η has only one point of intersection with such

branch at $z_i = u_0 e^{i\alpha}$. Taken $\eta(a) = \eta(b) = u_0 e^{i\alpha}$, we define $\eta_\epsilon : [a + \epsilon, b + \epsilon] \rightarrow \mathbb{C}$ with $\epsilon > 0 : \eta_\epsilon(t) = \eta(t) \forall t \in [a + \epsilon, b + \epsilon]$, then

$$\oint_{\eta} \frac{1}{z} dz = \lim_{\epsilon \rightarrow 0} \oint_{\eta_\epsilon} \frac{1}{z} dz$$

Therefore, $\forall z \in \mathbb{C} \setminus \{\sigma\}$ we have that

$$\frac{d \log z}{dz} = \frac{1}{z}$$

And therefore

$$\oint_{\eta_\epsilon} \frac{1}{z} dz = \log(\eta(b - \epsilon)) - \log(\eta(a + \epsilon))$$

For $\epsilon \rightarrow 0$ we have

$$\int_{\eta} \frac{1}{z} dz = (\log(u_0) + i(\alpha + 2\pi)) - (\log(u_0) + i\alpha) = 2\pi i$$

Example 7.4.2. Let's calculate the integral of $f(z) = \sqrt{z}$ along a closed simple piecewise smooth curve $\gamma : [a, b] \rightarrow \mathbb{C} : 0 \in \gamma^\circ$ and it intersects the line $\sigma_\alpha = u_0 e^{i\alpha}$, where

$$\sqrt{z} = \sqrt{r} e^{i\frac{\theta}{2}} \quad r \in \mathbb{R}^+, \theta \in (\alpha, \alpha + 2\pi], \alpha \in \mathbb{R}$$

Taken a parametrization $\gamma(t) : \gamma(a) = \gamma(b) = u_0 e^{i\alpha}$ we have that $f(z) \in H(D)$ where $D = \mathbb{C} \setminus \{\sigma_\alpha\}$. Proceeding as before, we have

$$\oint_{\gamma} \sqrt{z} dz = \lim_{\epsilon \rightarrow 0} \oint_{\gamma_\epsilon} \sqrt{z} dz$$

Since it has a primitive in D we can write

$$\lim_{\epsilon \rightarrow 0} \oint_{\gamma_\epsilon} \sqrt{z} dz = \frac{2}{3} \lim_{\epsilon \rightarrow 0} z \sqrt{z} \Big|_{\gamma_\epsilon(a+\epsilon)}^{\gamma_\epsilon(b-\epsilon)} = \frac{2}{3} u_0 \sqrt{u_0} e^{\frac{3}{2}i(\alpha+2\pi)} - \frac{2}{3} u_0 \sqrt{u_0} e^{\frac{3}{2}i\alpha} = -\frac{4}{3} u_0 \sqrt{u_0} e^{\frac{3}{2}i\alpha}$$

Lemma 7.4.1. Taken a closed simple pointwise smooth curve $\gamma : [a, b] \rightarrow \mathbb{C}$ and taken $D = \{\gamma\}^\circ \cup \gamma = \overline{\{\gamma\}^\circ}$ and a function $f \in H(D)$, for a finite cover of D , \mathcal{Q} composed by squares $Q_j \in \mathcal{Q} \forall j \in [1, N] \subset \mathbb{N}$, we have that

$$\exists z_j \in Q_j \cap \overline{\{\gamma\}^\circ} : \left\| \frac{f(z) - f(z_j)}{z - z_j} - \frac{df}{dz} \Big|_{z_j} \right\| < \epsilon \quad \forall z \in Q_j \cap \overline{\{\gamma\}^\circ} \setminus \{z_j\}$$

Proof. Going by contradiction, let's say that

$$\exists \epsilon > 0 : \nexists z_j \in Q_j \cap \overline{\{\gamma\}^\circ}$$

Taken a finite subcover \mathcal{Q}_n where $\text{diam}(Q_j^n) = \frac{d}{2^n} \forall Q_j \in \mathcal{Q}$ we can define for some $k \in K \subset \mathbb{N}$

$$A_n = \bigcup_{k \in K} Q_k^n \cap \overline{\{\gamma\}^\circ} \quad \forall n \in \mathbb{N}$$

We have that $A_{n+1} \subset A_n$, and taking a sequence $(w)_n \in \overline{\{\gamma\}^\circ}$ we have due to the compactness of $\overline{\{\gamma\}^\circ}$ that $\exists (w)_{n_j} \rightarrow w \in \overline{\{\gamma\}^\circ}$. Since $f \in H(\{\gamma\}^\circ)$ we have that f is holomorphic in w , therefore

$$\forall \epsilon > 0 \exists \delta_\epsilon > 0 : \left\| \frac{f(z) - f(w)}{z - w} - \frac{df}{dz} \Big|_w \right\| < \epsilon \quad \forall z \in B_{\delta_\epsilon}(w) \setminus \{w\}$$

Taken an \tilde{n} such that $\text{diam}(Q_j^{\tilde{n}}) = \frac{\sqrt{2}}{2^{\tilde{n}}} d < \delta$ we have that still $w \in A_n \quad \forall n \in \mathbb{N}$, and due to its closedness we can also say

$$\exists N_j \in \mathbb{N} : \forall n_j > N_j \quad (w)_{n_j} \in A_n$$

Therefore

$$\exists k_0 \in \mathbb{N} : w \in Q_{k_0}^{\tilde{n}} \cap \overline{\{\gamma\}^\circ} \subset A_{\tilde{n}} \quad \nexists$$

□

Theorem 7.31 (Cauchy-Goursat). *Taken $\gamma : [a, b] \rightarrow \mathbb{C}$ a closed simple piecewise smooth curve and $D = \{\gamma\} \cup \{\gamma\}^\circ$ and a function $f \in H(D)$, we have*

$$\oint_{\gamma} f(z) dz = 0 \quad (7.73)$$

Proof. Using the previous lemma we can say that for a finite cover $\{\gamma\}, Q_j \in \mathcal{Q} \exists z_j \in Q_j \cap \overline{\{\gamma\}^\circ}$ and a function

$$\delta_j(z) = \begin{cases} \frac{f(z) - f(z_j)}{z - z_j} - f'(z_j) & z \neq z_j \\ 0 & z = z_j \end{cases}$$

Which is continuous and $\delta_j(z) < \epsilon \quad \forall z \in Q_j \cap \overline{\{\gamma\}^\circ}$.

Taken a curve $\{\eta_j\} = \partial(Q_j \cap \overline{\{\gamma\}^\circ})$, and the expansion of $f(z)$ in the region, we have that

$$\begin{aligned} f(z) &= f(z_j) + f'(z_j)(z - z_j) + \delta_j(z)(z - z_j) \\ \oint_{\eta_j} f(z) dz &= (f(z_j) - z_j f'(z_j)) \oint_{\eta_j} dz + f'(z_j) \oint_{\eta_j} z dz + \oint_{\eta_j} \delta_j(z)(z - z_j) dz \end{aligned}$$

The first two integrals on the second line are null, and we have therefore

$$\oint_{\eta_j} f(z) dz = \oint_{\eta_j} \delta_j(z)(z - z_j) dz$$

By definition $\{\gamma\} = \bigcup_{j=1}^N \{\eta_j\}$ and therefore

$$\oint_{\gamma} f(z) dz = \sum_{j=1}^N \oint_{\eta_j} \delta_j(z)(z - z_j) dz$$

Using the Darboux inequality we have immediately that

$$\left\| \oint_{\gamma} f(z) dz \right\| \leq \sum_{j=1}^N \left\| \oint_{\eta_j} \delta_j(z)(z - z_j) dz \right\| \leq \sum_{j=1}^N \epsilon \sqrt{2} d (4d + L_j)$$

Using the theorem on the Jordan curve, we have that $\exists Q_n \in \mathcal{Q}$ such that $\{\gamma\} \subset Q_n$. Taken $\text{diam}(Q_n) = D$

$$\left\| \oint_{\gamma} f(z) dz \right\| \leq \sum_{j=1}^N \epsilon \sqrt{2} D (4D + L) \rightarrow 0$$

□

Definition 7.4.4 (Simple Connected Set). An open set $G \subset X$ with X some metric space, is said to be *simply connected* iff $\forall \{\gamma_j\} \subset G$ simple curves we have that $\gamma_j \sim 0$. $\gamma \sim 0$ implies that the curve is homotopic to a point

Theorem 7.32 (Cauchy-Goursat II). *Let $G \subset \mathbb{C}$ open and simply connected. Then, $\forall f \in H(G)$, $\{\gamma\} \subset G$ with γ simple closed and smooth*

$$\oint_{\gamma} f(z) dz = 0$$

Proof. 1. The curve γ doesn't intersect itself.

$$\oint_{\gamma} f(z) dz = \oint_0 f(z) dz = 0$$

2. The curve γ intersects itself $n - 1$ times.

Then $\{\gamma\} = \bigcup_{k=1}^n \{\gamma_k\}$ with γ_k simple smooth non intersecting curves. Since $\{\gamma_k\} \subset G \forall k = 1, \dots, n$, $\{\gamma_k\} \sim 0$, we have

$$\oint_{\gamma} f(z) dz = \sum_{k=1}^n \oint_{\gamma_k} f(z) dz = 0$$

□

Theorem 7.33. *Let $G \subset \mathbb{C}$ be a simply connected open set. If $f \in H(G)$, then there exists a primitive for $f(z)$*

§§ 7.4.1 Integral Representation of Holomorphic Functions

Definition 7.4.5 (Positively Oriented Curve). The parametrization of a curve in \mathbb{C} is said to be *positively oriented* if its parametrization is taken such the path taken results counterclockwise.

Notation. The integral over a closed positively oriented parametrization of a curve γ is indicated as follows

$$\oint_{\gamma}$$

Theorem 7.34 (Cauchy Integral Representation). *Taken a positively oriented closed simple piecewise smooth curve $\gamma : [a, b] \rightarrow \mathbb{C}$ and a function $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$ such that if $D = \{\gamma\} \cup \{\gamma\}^\circ \subset G$, $f \in H(D)$, we have that*

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{w - z} dw \quad \forall w \in \{\gamma\}^\circ \quad (7.74)$$

Proof. Taken $\gamma_\rho(\theta) = z + \rho e^{i\theta}$ such that $\gamma_\rho \sim \gamma$, $\{\gamma_\rho\} \subset \{\gamma\}^\circ$ is a simple curve, we have

$$\oint_{\gamma} \frac{f(w)}{w-z} dw = \oint_{\gamma_\rho} \frac{f(w)}{w-z} dw$$

Then, using that

$$\oint_{\gamma} \frac{1}{w-z} dw = 2\pi i$$

We get

$$\oint_{\gamma} \frac{f(z)}{w-z} dw - 2\pi i f(z) = \oint_{\gamma_\rho} \frac{f(w) - f(z)}{w-z} dw$$

Since $f \in H(\{\gamma\}^\circ)$ we have that

$$\forall \epsilon > 0 \exists \delta_\epsilon > 0 : \|z - w\| < \delta_\epsilon \implies \|f(z) - f(w)\| < \epsilon$$

Taken $\rho < \delta_\epsilon$ we get, using the Darboux inequality

$$\left\| \oint_{\gamma_\rho} \frac{f(w) - f(z)}{w-z} dw \right\| \leq 2\pi\epsilon \implies \oint_{\gamma} \frac{f(w) - f(z)}{w-z} dw = 0$$

□

Theorem 7.35 (Derivatives of a Holomorphic Function). *Let $D \subset \mathbb{C}$ be an open set and $f : D \rightarrow \mathbb{C}$ a function $f \in H(D)$, then $f \in C^\infty(D)$ and*

$$\frac{d^n f}{dz^n} = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w-z)^{n+1}} dw \quad (7.75)$$

Where γ is a closed simple piecewise smooth curve such that $z \in \{\gamma\}^\circ$ and $\overline{\{\gamma\}} \subset D$

Corollary 7.4.1. Let $f \in H(D)$, then

$$\forall n \in \mathbb{N} \quad \frac{d^n f}{dz^n} \in H(D)$$

Theorem 7.36 (Morera). *Let $D \subset \mathbb{C}$ be an open and connected set. Take $f : D \rightarrow \mathbb{C} : f \in C(D)$. Then, if $\forall \{\gamma\} \subset D$ closed piecewise smooth*

$$\oint_{\gamma} f(z) dz = 0 \implies f \in H(D) \quad (7.76)$$

Proof. Since $f \in C(D) \exists F(z) \in C^1(D) : f(z) = F'(z)$. Since $C^1(\mathbb{C}) \simeq H(\mathbb{C})$ we have that, due to the previous corollary

$$\frac{dF}{dz} = f(z) \in H(D)$$

□

Theorem 7.37 (Cauchy Inequality). *Let $f \in H(B_R(z_0))$ with $z_0 \in \mathbb{C}$. If $\|f(z)\| \leq M \forall z \in B_R(z_0)$*

$$\left\| \frac{df}{dz} \Big|_{z_0} \right\| \leq \frac{n!M}{R^n} \quad (7.77)$$

Proof. Take $\gamma_r(\theta) = z_0 + re^{i\theta}$ with $\theta \in [0, 2\pi]$, $r > R$, then the derivative $\frac{d^n f}{dz^n} \Big|_{z_0}$ can be written using the Cauchy integral representation, since $f \in H(B_r(z_0))$

$$\frac{d^n f}{dz^n} \Big|_{z_0} = \frac{n!}{2\pi i} \oint_{\gamma_r} \frac{f(w)}{(w - z_0)^{n+1}} dw$$

Using the Darboux inequality we have then

$$\left\| \frac{d^n f}{dz^n} \Big|_{z_0} \right\| \leq \frac{n!}{r^n} \sup_{z \in \{\gamma_r\}} \|f(z)\| \leq \frac{n!M}{r^n}$$

Since $r < R$ we therefore have

$$\left\| \frac{d^n f}{dz^n} \Big|_{z_0} \right\| \leq \frac{n!M}{R^n}$$

□

Theorem 7.38 (Liouville). *Let $f : \mathbb{C} \rightarrow \mathbb{C}$ a function such that $f \in H(\mathbb{C})$, i.e. whole. If $\exists M > 0 : \|f(z)\| \leq M \forall z \in \mathbb{C}$ the function $f(z)$ is constant*

Proof. $f \in H(\mathbb{C})$, $\|f(z)\| \leq M$ and we can write, taken $\gamma_R(\theta) = z + Re^{i\theta}$ with $\theta \in [0, 2\pi]$

$$f'(z) = \frac{1}{2\pi i} \oint_{\gamma_R} \frac{f(w)}{(w - z)^2} dz$$

For Darboux

$$\|f'(z)\| \leq \frac{1}{2\pi} \left\| \oint_{\gamma_R} \frac{f(w)}{(w - z)^2} dz \right\| \leq \frac{\sup_{z \in \{\gamma_R\}} \|f(z)\|}{R} \leq \frac{M}{R}$$

Since $R > 0$ is arbitrary, we can say directly that $\|f'(z)\| = 0$ and therefore $f(z)$ is constant $\forall z \in \mathbb{C}$. □

Theorem 7.39 (Fundamental Theorem of Algebra). *Take a polynomial $P_n(z) \in \mathbb{C}_n[z]$, where $\mathbb{C}_n[z]$ is the space of complex polynomials with variable z and degree n . If we have*

$$P_n(z) = \sum_{k=0}^n a_k z^k, \quad z, a_k \in \mathbb{C}, \quad a_n \neq 0$$

We can say that $\exists z_0 \in \mathbb{C} : P_n(z_0) = 0$

Proof. As an absurd, say that $\forall z \in \mathbb{C}, P_n(z) \neq 0$. Then $f(z) = 1/P_n(z) \in H(\mathbb{C})$.

Since $\lim_{z \rightarrow \infty} P_n(z) = \infty$, we have that $\|f(z)\| \leq M \forall z \in \mathbb{C}$, and $\lim_{z \rightarrow \infty} f(z) = 0$.

Therefore $\exists R > 0 : \forall \|z\| > R, \|f(z)\| < 1$. Since $f \in H(\mathbb{C})$, we have that $f \in C(\overline{B_R}(z))$. Due to the Liouville theorem we have that $f(z)$ is constant \nexists □

§ 7.5 Integral Theorems in \mathbb{R}^2 and \mathbb{R}^3

Theorem 7.40 (Gauss-Green). *Given $D \subset \mathbb{R}^2$ a set with a piecewise smooth parameterization of ∂D and two functions $\alpha, \beta : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ and $\overline{D} \subset A$*

$$\iint_D \partial_x \beta \, dx \, dy = \int_{\partial^+ D} \beta(x, y) \, dy, \quad \iint_D \partial_y \alpha \, dx \, dy = - \int_{\partial^+ D} \alpha(x, y) \, dx \, dy \quad (7.78)$$

Theorem 7.41 (Stokes). *Given $D \subset \mathbb{R}^2$ an open set with ∂D piecewise smooth and a vector field $f^\mu : A \rightarrow \mathbb{R}^2$ with $D \subset A$*

$$\int_D \epsilon_{3\mu\nu} \partial^\mu f^\nu \, dx \, dy = \int_{\partial^+ D} f^\mu t_\mu \, ds \quad (7.79)$$

Where t^μ is the vector tangent to $\partial^+ D$

Theorem 7.42 (Gauss 1). *Given $D \subset \mathbb{R}^n$ open set with ∂D piecewise smooth and a vector field $f^\mu : A \rightarrow \mathbb{R}^n$ with $D \subset A$*

$$\iint_D \partial_\mu f^\mu \, dx \, dy = \int_{\partial^+ D} f^\mu n_\mu \, ds \quad (7.80)$$

Where n^μ is the normal vector to $\partial^+ D$

Theorem 7.43 (Stokes for Surfaces). *Given a smooth surface $\Sigma \subset \mathbb{R}^3$ with parameterization r^μ and a vector field $f^\mu : A \rightarrow \mathbb{R}^3$ with $\Sigma \subseteq A$*

$$\int_\Sigma n^\mu \epsilon_{\mu\nu\gamma} \partial^\nu f^\gamma \, d\sigma = \int_{\partial^+ \Sigma} f^\mu t_\mu \, ds \quad (7.81)$$

Where t^μ is the tangent vector to the border of the surface

Theorem 7.44 (Useful Identities). *Given $u, v \in C^2(\Omega)$ and a vector field $f^\mu \in C^2(\Omega, \mathbb{R}^3)$*

$$\begin{aligned} \int_\Omega \partial_\mu \partial^\mu v \, dx \, dy \, dz &= \int_{\partial\Omega} n^\mu \partial_\mu v \, d\sigma \\ \int_\Omega u \partial_\mu f^\mu \, dx \, dy \, dz &= - \int_\Omega f^\mu \partial_\mu u \, dx \, dy \, dz + \int_{\partial\Omega} u f^\mu n_\mu \, d\sigma \\ \int_\Omega u \partial_\mu \partial^\mu v \, dx \, dy \, dz &= - \int_\Omega \partial_\mu u \partial^\mu v \, dx \, dy \, dz + \int_{\partial\Omega} u n^\mu \partial_\mu v \, d\sigma \\ \int_\Omega (u \partial_\mu \partial^\mu v - v \partial_\mu \partial^\mu u) \, dx \, dy \, dz &= \int_{\partial\Omega} (u n^\mu \partial_\mu v - v n^\mu \partial_\mu u) \, d\sigma \end{aligned} \quad (7.82)$$

We can analogously write these theorems in the language of differential forms and manifolds, after giving a couple of definitions

Definition 7.5.1 (Volume Element). *Given a k -dimensional compact oriented manifold M with boundary and $\omega \in \Lambda^k(M)$ a k -differential form on M , we define the *volume* of M as follows*

$$V(M) = \int_M dV = \int_M \omega \quad (7.83)$$

Where dV is the *volume element* of the manifold, given by the unique $\omega \in \Lambda^k(M)$, defined as follows

$$\omega = f dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k} \quad (7.84)$$

With f an unique function.

For $M \subset \mathbb{R}^3$ with n^μ as outer normal and $\omega \in \Lambda^2(M)$ we can write immediately, by definition

$$\omega_{\mu\nu} v^\mu w^\nu = n^\mu \epsilon_{\mu\nu\gamma} v^\nu w^\gamma = dA$$

Therefore

$$dA = \|\epsilon_{\mu\nu\gamma} v^\nu w^\gamma\|^\mu \quad (7.85)$$

Which is the already known formula.

For a 2-manifold we can write immediately the following formulas

$$dA = n^1 dy \wedge dz + n^2 dz \wedge dx + n^3 dx \wedge dy \quad (7.86)$$

And, on M

$$\begin{cases} n^1 dA = dy \wedge dz \\ n^2 dA = dz \wedge dx \\ n^3 dA = dx \wedge dy \end{cases} \quad (7.87)$$

Theorem 7.45 (Gauss-Green-Stokes-Ostogradskij). *Given M a smooth manifold with boundary, c a p -cube in M and $\omega \in \Lambda(M)$ we have*

$$\int_c d\omega = \int_{[0,1]^p} c^* d\omega = \int_{\partial c} \omega \quad (7.88)$$

In general, we can write

$$\int_M d\omega = \int_{\partial M} \omega \quad (7.89)$$

Definition 7.5.2 (Gauss-Green, Differential Forms). Given $M \subset \mathbb{R}^2$ a compact 2-manifold with boundary and two functions $\alpha, \beta : M \rightarrow \mathbb{R}$ with $\alpha, \beta \in C^1(M)$ defining

$$\omega = \alpha dx + \beta dy \quad (7.90)$$

We have

$$\int_{\partial M} \alpha dx + \beta dy = \int_{\partial M} \omega = \int_M d\omega = \iint_M \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dx \wedge dy \quad (7.91)$$

Proof. Take $\omega = \alpha dx + \beta dy$, then

$$d\omega = d(\alpha dx + \beta dy) = \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dx \wedge dy$$

□

Theorem 7.46 (Gauss, Differential Forms). *Given M a 3-manifold smooth with boundary and compact with outer normal n^μ and a vector field $f^\mu \in C^1(M)$, we have*

$$\int_M \partial_\mu f^\mu dV = \int_{\partial M} f^\mu n_\mu dA \quad (7.92)$$

Proof. Taken the following differential form

$$\omega = f^1 dy \wedge dz + f^2 dz \wedge dx + f^3 dx \wedge dy$$

We have, using the formulas (7.87)

$$\omega = f^\mu n_\mu dA$$

And

$$d\omega = \partial_\mu f^\mu dV$$

Therefore

$$\int_M \partial_\mu f^\mu dV = \int_M d\omega = \int_{\partial M} \omega = \int_{\partial M} f^\mu n_\mu dA$$

□

Theorem 7.47 (Stokes, Differential Forms). *Given $M \subset \mathbb{R}^3$ a compact oriented smooth 2-manifold with boundary with n^μ as outer normal and t^μ as tangent vector in ∂M , given a vector field $f^\mu \in C^1(A)$ where $M \subset A$, we have*

$$\int_M n^\mu \epsilon_{\mu\nu\gamma} \partial^\nu f^\gamma dA = \int_{\partial M} f^\mu t_\mu ds \quad (7.93)$$

Proof. Taking the following differential form

$$\omega = f^\mu dx_\mu$$

We have that

$$d\omega = (\partial_2 f^3 - \partial_3 f^2) dy \wedge dz + (\partial_3 f^1 - \partial_1 f^3) dz \wedge dx + (\partial_1 f^2 - \partial_2 f^1) dx \wedge dy$$

Using the formulas (7.87) we have

$$d\omega = n^\mu \epsilon_{\mu\nu\gamma} \partial^\nu f^\gamma dA$$

Since in \mathbb{R}^2 we have $t^\mu ds = dx^\mu$ we therefore have

$$f^\mu t_\mu ds = f^\mu dx_\mu = \omega$$

And therefore

$$\int_M n^\mu \epsilon_{\mu\nu\gamma} \partial^\nu f^\gamma dA = \int_M d\omega = \int_{\partial M} \omega = \int_{\partial M} f^\mu t_\mu ds$$

□

These last formulas are a good example on how they can be generalized through the use of differential forms, bringing an easy way of calculus in \mathbb{R}^n of the various integral theorems, all condensed in one formula, the *Gauss-Green-Stokes-Ostogradskij theorem*

8 Series Representation and Residues

§ 8.1 Series Representation of Functions

§§ 8.1.1 Taylor Series

Theorem 8.1 (Taylor Series Expansion). *Let $f : D \rightarrow \mathbb{C}$ be a function such that $f \in H(B_R(z_0))$, with $B_r(z_0) \subseteq D$. Then*

$$f(z) = \sum_{n=0}^n \frac{1}{n!} \left. \frac{d^n f}{dz^n} \right|_{z_0} (z - z_0)^n \quad \|z - z_0\| < r \quad (8.1)$$

Proof. Taken $z \in B_r(z_0)$ and $\gamma(t) = z_0 + re^{it}$ $t \in [0, 2\pi]$ and $\|z - z_0\| < r < R$ we can write, using the integral representation of f

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z)} dw = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z_0) - (z - z_0)} dw$$

From basic calculus we know already that if $z \neq w$

$$\begin{aligned} \frac{1}{w - z} &= \frac{1}{w} \left(\frac{1 - (z/w)^n}{1 - z/w} + \frac{1}{1 - z/w} \left(\frac{z}{w} \right)^n \right) = \\ &= \frac{1}{w - z} \left(\frac{z}{w} \right)^n + \sum_{k=0}^{n-1} \frac{1}{w} \left(\frac{z}{w} \right)^k \end{aligned}$$

Therefore, inserting it back into the integral representation, we have

$$f(z) = \sum_{k=0}^n \frac{(z - z_0)^k}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z_0)^{k+1}} dw + \frac{(z - z_0)^n}{2\pi i} \oint_{\gamma} \frac{f(w)}{[(w - z_0) - (z - z_0)](w - z_0)^n} dw$$

On the RHS as first term we have the k -th derivative of f and on the right there is the so called *remainder* $R_n(z)$. Therefore

$$f(z) = \sum_{k=0}^n \frac{1}{k!} \left. \frac{d^k f}{dz^k} \right|_{z_0} (z - z_0)^k + R_n(z)$$

It's easy to demonstrate that $R_n(z) \xrightarrow{n \rightarrow \infty} 0$, and therefore

$$f(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \left. \frac{d^k f}{dz^k} \right|_{z_0} (z - z_0)^k$$

□

Definition 8.1.1 (MacLaurin Series). Taken a Taylor series, such that $z_0 = 0$, we obtain a MacLaurin series.

$$f(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \left. \frac{d^k f}{dz^k} \right|_{z=0} z^k \quad (8.2)$$

Definition 8.1.2 (Remainders). We can have two kinds of remainder functions while calculating series:

1. Peano Reminders, $R_n(z) = \mathcal{O}(\|z - z_0\|^n)$
2. Lagrange Reminders, $R_n(x) = (n+1)!^{-1} f^{(n+1)}(\xi)(x - x_0)^{n+1}$, $x, x_0 \in \mathbb{R}$ $\xi \in (x, x_0)$

What we saw before as $R_n(z)$ is the remainder function for functions $f : D \subset \mathbb{C} \rightarrow \mathbb{C}$. A particularity of remainder function is that $R_n(z) \rightarrow 0$ always, if f is holomorphic

Theorem 8.2 (Integration of Power Series II). Let $f, g : B_R(z_0) \rightarrow \mathbb{C}$ and $\{\gamma\} \subset B_R(z_0)$ a piecewise smooth path. Taken

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad g \in C(\{\gamma\})$$

We have that

$$\sum_{n=0}^{\infty} a_n \int_{\gamma} g(z) (z - z_0)^n dz = \int_{\gamma} g(z) f(z) dz \quad (8.3)$$

Proof. Since $f, g \in C(\{\gamma\})$ by definition, and $f \in H(\overline{B_r(z_0)})$ with $r < R$, we have that $\exists \hat{K}_{\gamma}[fg]$. Firstly we can write that $\forall z \in B_R(z_0)$

$$g(z)f(z) = \sum_{k=0}^{n-1} a_k g(z) (z - z_0)^k + g(z)R_n(z) = \sum_{k=0}^{n-1} a_k g(z) (z - z_0)^k + g(z) \sum_{k=n}^{\infty} a_k (z - z_0)^k$$

Then we can write

$$\int_{\gamma} g(z)f(z) dz = \sum_{k=0}^{n-1} a_k \oint_{\gamma} g(z) (z - z_0)^k dz + \int_{\gamma} g(z)R_n(z) dz$$

Letting $M = \sup_{z \in \{\gamma\}} \|g(z)\|$, and noting that $\|R_n(z)\| < \epsilon$ for $\forall \epsilon > 0$ and for some $n \geq N_{\epsilon} \in \mathbb{N}$, $z \in \{\gamma\}$ we have, using the Darboux inequality

$$\left\| \int_{\gamma} g(z)R_n(z) dz \right\| \leq ML_{\gamma} \epsilon \rightarrow 0$$

□

Theorem 8.3 (Holomorphy of Power Series). *If a function $f(z)$ is expressible as a power series $f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k$, $\|z - z_0\| < R$ we have that $f \in H(B_R(z_0))$*

Proof. Take the previous theorem on the integration of power series, and choose $g(z) = 1$. Since $g(z) \in H(\mathbb{C})$ we also have that it'll be continuous on all paths $\{\gamma\} \subset \mathbb{C}$ piecewise smooth. Take now a closed piecewise smooth path $\{\gamma\}$, then we can write

$$\oint_{\gamma} f(z)g(z) dz = \oint_{\gamma} \sum_{k=0}^{\infty} a_k(z - z_0)^k = \sum_{k=0}^{\infty} a_k \oint_{\gamma} (z - z_0)^k dz$$

Since the function $h(z) = (z - z_0)^k \in H(\mathbb{C}) \forall k \neq 1$, we have, thanks to the Morera and Cauchy theorems

$$\oint_{\gamma} f(z) dz = 0 \implies f(z) \in H(B_R(\overline{\{\gamma\}}))$$

□

Corollary 8.1.1 (Derivative of a Power Series II). Take $f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k$ $\|z - z_0\| < R$. Then, $\forall z \in B_R(z_0)$ we have that

$$\frac{df}{dz} = \sum_{k=1}^{\infty} a_k k (z - z_0)^{k-1} \quad (8.4)$$

Proof. Taken $z \in B_R(z_0)$ and a continuous function $g(w) \in C(\{\gamma\})$, with $\{\gamma\} \subset B_R(z_0)$ a closed simple piecewise smooth path. If $z \in \{\gamma\}^\circ$ and

$$g(w) = \frac{1}{2\pi i} \left(\frac{1}{(w - z)^2} \right)$$

We have, using the integral representation for holomorphic functions

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z)^2} dw = \sum_{k=0}^{\infty} \frac{a_k}{2\pi i} \oint_{\gamma} \frac{(w - z_0)^k}{(w - z)^2} dw$$

Since $h(w) = (w - z_0)^k \in H(\mathbb{C}) \forall k \neq 1$ we have, using again the integral representation for holomorphic functions

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z)^2} dw &= \frac{df}{dz} \\ \frac{1}{2\pi i} \oint_{\gamma} \frac{(w - z_0)^k}{(w - z)^2} dw &= k(z - z_0)^{k-1} \end{aligned}$$

Therefore

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f(w)}{(w - z)^2} dw = \sum_{k=0}^{\infty} a_k k (z - z_0)^{k-1} = \frac{df}{dz}$$

□

Corollary 8.1.2 (Uniqueness of the Taylor Series). Taken an holomorphic function $f \in H(D)$ with $D \subset \mathbb{C}$ some connected open set, we have that

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k \quad a_k = \frac{1}{k!} \left. \frac{d^k f}{dz^k} \right|_{z_0} \quad \forall \|z - z_0\| < R$$

Proof. Taken $g(z)$ a continuous function over a closed piecewise simple path $\{\gamma\} \subset \mathbb{C}$, where

$$g(z) = \frac{1}{2\pi i} \left(\frac{1}{(z - z_0)^{k+1}} \right)$$

We have that

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz = \sum_{k=1}^{\infty} \frac{a_k}{2\pi i} \oint_{\gamma} (z - z_0)^{k-n-1} dz$$

The integral on the RHS evaluates to δ_n^k , and thanks to the integral representation of $f(z)$ we can write

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz = \frac{1}{n!} \left. \frac{d^n f}{dz^n} \right|_{z_0} = n! a_n$$

□

§§ 8.1.2 Laurent Series

Definition 8.1.3 (Annulus Domain). Let $0 \leq r < R \leq \infty$ and $z_0 \in \mathbb{C}$, we define the *annulus set* as follows

$$A_{r,R}(z_0) := \{z \in \mathbb{C} \mid r < \|z - z_0\| < R\} \quad (8.5)$$

Special cases of this are the ones where $r = 0$, $R = \infty$ and $r = 0$, $R = \infty$

$$\begin{aligned} A_{0,R}(z_0) &= B_R(z_0) \setminus \{z_0\} \\ A_{r,\infty}(z_0) &= \mathbb{C} \setminus \overline{B_r}(z_0) \\ A_{0,\infty}(z_0) &= \mathbb{C} \setminus \{z_0\} \end{aligned}$$

Theorem 8.4 (Laurent Series Expansion). Let $f : A_{R_1 R_2}(z_0) \rightarrow \mathbb{C}$ be a function such that $f \in H(A_{R_1 R_2}(z_0))$, and $\{\gamma\} \subset A_{R_1 R_2}(z_0)$ a closed simple piecewise smooth curve. Then f is expandable in a generalized power series or a Laurent series as follows

$$f(z) = \sum_{n=0}^{\infty} c_n^+ (z - z_0)^n + \sum_{n=1}^{\infty} \frac{c_n^-}{(z - z_0)^n} = \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k \quad (8.6)$$

Where the coefficients are the following

$$\begin{aligned} c_n^- &= \frac{1}{2\pi i} \oint_{\gamma} f(z) (z - z_0)^{n-1} dz \quad n \geq 0 \\ c_n^+ &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad n > 0 \\ c_k &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{k+1}} dz \quad k \in \mathbb{Z} \end{aligned} \quad (8.7)$$

Proof. Taken a random point $z \in A_{R_1 R_2}(z_0)$, a closed simple piecewise smooth curve $\{\gamma\} \subset A_{R_1 R_2}(z_0)$ and two circular smooth paths $\{\gamma_2\}, \{\gamma_3\} : \{\gamma_2\} \cup \{\gamma_3\} = \partial A_{r_1 r_2}(z_0) \subset A_{R_1 R_2}(z_0) \wedge \{\gamma\} \subset A_{r_1 r_2}(z_0)$ and a third circular path $\{\gamma_3\} \subset A_{r_1 r_2}(z_0)$, we can write immediately, using the omotopy between all the paths

$$\oint_{\gamma_2} \frac{f(w)}{w-z} dw = \oint_{\gamma_1} \frac{f(w)}{w-z} dw + \oint_{\gamma_3} \frac{f(w)}{w-z} dw$$

Using the Cauchy integral representation we have that the integral on γ_3 yields immediately $2\pi i f(z)$, hence we can write

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma_2} \frac{f(w)}{(w-z_0)-(z-z_0)} dw + \frac{1}{2\pi i} \oint_{\gamma_1} \frac{f(w)}{(z_0-z)-(w-z_0)} dw$$

Using the two following identities for $z \neq w$

$$\begin{aligned} \frac{1}{w-z} &= \frac{1}{w-z} \left(\frac{z}{w}\right)^n + \sum_{k=0}^{n-1} \frac{1}{w} \left(\frac{z}{w}\right)^k \\ \frac{1}{z-w} &= \frac{1}{z-w} \left(\frac{w}{z}\right)^n + \sum_{k=1}^n \frac{1}{w} \left(\frac{w}{z}\right)^k \end{aligned}$$

We obtain that

$$f(z) = \sum_{k=0}^{n-1} \frac{(z-z_0)^k}{2\pi i} \oint_{\gamma_2} \frac{f(w)}{(w-z_0)^{k+1}} dw + \rho_n(z) + \sum_{k=1}^n \frac{1}{2\pi i (z-z_0)^k} \oint_{\gamma_1} f(w)(w-z_0)^{k-1} dw + \sigma_n(z)$$

Where, after choosing appropriate substitutions with some coefficients c_k^+, c_k^- we have

$$f(z) = \sum_{k=0}^{n-1} c_k^+ (z-z_0)^k + \rho_n(z) + \sum_{k=1}^n \frac{c_k^-}{(z-z_0)^k} + \sigma_n(z)$$

Where ρ_n, σ_n are the two remainders of the series expansion, and are

$$\begin{aligned} \rho_n(z) &= \frac{(z-z_0)^n}{2\pi i} \oint_{\gamma_2} \frac{f(w)}{[(w-z_0)-(z-z_0)](w-z_0)^n} dw \\ \sigma_n(z) &= \frac{1}{2\pi i (z-z_0)^n} \oint_{\gamma_1} \frac{f(w)}{(w-z_0)-(z-z_0)} dw \end{aligned}$$

In order to prove the theorem we now need to demonstrate that $\rho_n, \sigma_n \xrightarrow{n \rightarrow \infty} 0$. Taken $M_1 = \sup_{z \in \{\gamma_1\}} \|f(z)\|$, $M_2 = \sup_{z \in \{\gamma_2\}} \|f(z)\|$, we have, using the fact that both γ_1, γ_2 are circular

$$\begin{aligned} \|\rho_n(z)\| &\leq \frac{M_2}{1 - \frac{1}{r_2} \|z-z_0\|} \left(\frac{\|z-z_0\|}{r_2}\right)^n \xrightarrow{n \rightarrow \infty} 0 \quad \|z-z_0\| < r_2 \\ \|\sigma_n(z)\| &\leq \frac{M_1}{\frac{1}{r_1} \|z-z_0\| - 1} \left(\frac{r_1}{\|z-z_0\|}\right)^n \xrightarrow{n \rightarrow \infty} 0 \quad r_1 < \|z-z_0\| \end{aligned}$$

And the theorem is proved. \square

Theorem 8.5 (Convergence of a Laurent Series). *Being defined on an annulus set, the Laurent series of a function must have two radii of convergence. Given a function f holomorphic on a set $A_{R_1 R_2}(z_0)$ we have*

$$\begin{aligned} \frac{1}{R_2} &= \limsup_{n \rightarrow \infty} \sqrt[n]{\|c_n\|} \\ R_1 &= \limsup_{n \rightarrow \infty} \sqrt[n]{\|c_{-n}\|} \end{aligned} \quad (8.8)$$

It's equivalent of showing the convergence of the two series

$$f(z) = \sum_{k=0}^{\infty} c_k^+(z - z_0)^k + \sum_{k=1}^{\infty} \frac{c_k^-}{(z - z_0)^k}$$

Theorem 8.6 (Integral of a Laurent Series). *Let $f(z) \in H(A_{R_1 R_2}(z_0))$ and take $\{\gamma\} \subset A_{R_1 R_2}(z_0)$ a piecewise smooth curve, and $g \in C(\{\gamma\})$, then we have*

$$\sum_{n=-\infty}^{\infty} c_n \oint_{\gamma} g(z)(z - z_0)^n dz = \oint_{\gamma} g(z)f(z) dz$$

Proof. We begin by separating the sum in two parts, ending up with the following

$$\begin{aligned} \sum_{n=0}^{\infty} c_n^+ \oint_{\gamma} g(z)(z - z_0)^n dz &= \oint_{\gamma} g(z)f_+(z) dz \\ \sum_{n=1}^{\infty} c_n^- \oint_{\gamma} \frac{g(z)}{(z - z_0)^n} dz &= \oint_{\gamma} g(z)f_-(z) dz \end{aligned}$$

Which is analogous to the integration of Taylor series. The same could be obtained keeping the bounds of the sum in all \mathbb{Z} \square

As for Taylor series, in a completely analogous fashion, a Laurent series is holomorphic and unique.

The derivative of a Laurent series, is then obviously the following

$$\frac{df}{dz} = \sum_{n=-\infty}^{\infty} c_n n (z - z_0)^{n-1}$$

§§ 8.1.3 Multiplication and Division of Power Series

Theorem 8.7 (Product of Power Series). *Take $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$, $z \in B_{R_1}(z_0)$ and $g(z) = \sum_{n=0}^{\infty} b_n(z - z_0)^n$, $z \in B_{R_2}(z_0)$. Then*

$$f(z)g(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n, \quad c_n = \sum_{k=0}^n a_k b_{n-k} \quad \|z - z_0\| < \min(R_1, R_2) = R$$

Proof. Due to the holomorphy of both f and g , we have that the function fg has a Taylor series expansion

$$f(z)g(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n \quad \|z - z_0\| < R$$

We have then, using Leibniz's derivation rule

$$\begin{aligned}
 c_n &= \frac{1}{n!} \frac{d^n}{dz^n} f(z)g(z) = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \left. \frac{d^k f}{dz^k} \right|_{z_0} \left. \frac{d^{n-k} g}{dz^{n-k}} \right|_{z_0} = \\
 &= \sum_{k=0}^n \frac{1}{k!} \left. \frac{d^k f}{dz^k} \right|_{z_0} \frac{1}{(n-k)!} \left. \frac{d^{n-k} g}{dz^{n-k}} \right|_{z_0} = \\
 &= \sum_{k=0}^n a_k b_{n-k}
 \end{aligned}$$

□

Theorem 8.8 (Division of Power Series). *Taken the two functions as before, with the added necessity that $g(z) \neq 0$, we have that*

$$\frac{f(z)}{g(z)} = \sum_{n=0}^{\infty} d_n (z - z_0)^n \quad d_n = \frac{1}{b_0} \left(a_n - \sum_{k=0}^{n-1} d_k b_{n-k} \right)$$

Proof. Everything hold as in the previous proof. Remembering that $(f/g)g = f$ and using the previous theorem, we obtain

$$a_n = \sum_{k=0}^n d_k b_{n-k}$$

And therefore, inverting

$$d_n = \frac{a_n}{b_0} - \frac{1}{b_0} \sum_{k=0}^{n-1} d_k b_{n-k}$$

□

§§ 8.1.4 Useful Expansions

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad \|z\| < \infty \quad (8.9)$$

$$\sin(z) = \frac{1}{2i} (e^{iz} - e^{-iz}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \quad \|z\| < \infty \quad (8.10)$$

$$\cos(z) = \frac{d}{dz} \sin(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} \quad \|z\| < \infty \quad (8.11)$$

$$\cosh(z) = \cos(iz) = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} \quad \|z\| < \infty \quad (8.12)$$

$$\sinh(z) = \frac{d}{dz} \cosh(z) = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} \quad \|z\| < \infty \quad (8.13)$$

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n \quad \|z\| < 1 \quad (8.14)$$

$$\frac{1}{1+z} = \sum_{n=0}^{\infty} (-1)^n z^n \quad \|z\| < 1 \quad (8.15)$$

$$\frac{1}{z} = \sum_{n=0}^{\infty} (-1)^n (z-1)^n \quad \|z-1\| < 1 \quad (8.16)$$

$$(1+z)^s = \sum_{n=0}^{\infty} \binom{s}{n} z^n \quad s \in \mathbb{C}, \|z\| < 1 \quad (8.17)$$

$$e^{\frac{1}{z}} = \sum_{n=0}^{\infty} \frac{1}{n! z^n} \quad 0 < \|z\| < \infty \quad (8.18)$$

§ 8.2 Residues

§§ 8.2.1 Singularities and Residues

Definition 8.2.1 (Singularity). Given a function $f : G \rightarrow \mathbb{C}$ we define a *singularity* a point $z_0 \in G$ such that

$$\forall \epsilon > 0 \exists z \in B_\epsilon(z_0) : f(z) \text{ is holomorphic} \quad (8.19)$$

Definition 8.2.2 (Isolated Singularity). Given a function $f : G \rightarrow \mathbb{C}$ we define an *isolated singularity* a point $z_0 \in G$ such that

$$\exists r > 0 : f \in H(A_{0r}(z_0)) \quad (8.20)$$

Definition 8.2.3 (Residue). Let $z_0 \in G$ be an isolated singularity of $f : G \rightarrow \mathbb{C}$, then $\exists r > 0 : \forall z \in A_{0r}(z_0)$ the following Laurent series expansion holds

$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z-z_0)^n} = \sum_{n=-\infty}^{\infty} c_n (z-z_0)^n$$

The *residue* of the function f in z_0 is defined as follows

$$\operatorname{Res}_{z=z_0} f(z) = b_1 = c_{-1} \quad (8.21)$$

A second definition is given by the following contour integral

$$\operatorname{Res}_{z=z_0} f(z) = \frac{1}{2\pi i} \oint_{\gamma} f(z) dz$$

Where γ is a simple closed path around z_0

Definition 8.2.4 (Winding Number). Given a closed curve $\{\gamma\}$ we define the *winding number* or *index* of the curve around a point z_0 the following integral

$$n(\gamma, z_0) = \frac{1}{2\pi i} \oint_{\gamma} \frac{dz}{z-z_0} \quad (8.22)$$

Theorem 8.9 (Residue Theorem). *Given a function $f : G \rightarrow \mathbb{C}$ such that $f \in H(D)$ where $D = G \setminus \{z_1, \dots, z_n\}$ and z_k are isolated singularities, we have, taken a closed piecewise smooth curve $\{\gamma\}$, such that $\{z_1, \dots, z_n\} \subset \{\gamma\}^\circ$*

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{k=0}^{\infty} n(\gamma, z_k) \operatorname{Res}_{z=z_k} f(z) \quad (8.23)$$

Proof. Firstly we can say that $\gamma \sim \sum_k \gamma_k$ where γ_k are simple curves around each z_k , then since the function is holomorphic in the regions $A_{0r}(z_k)$ with $k = 1, \dots, n$ we can write

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_k)^n$$

Therefore, we have

$$\oint_{\gamma} f(z) dz = \sum_{k=0}^n \oint_{\gamma_k} f(z) dz = \sum_{k=0}^n \sum_{j=-\infty}^{\infty} c_j \oint_{\gamma_k} (z - z_k)^j dz$$

We can then use the linearity of the integral operator and write

$$\oint_{\gamma} f(z) dz = \sum_{k=0}^n \sum_{j=-\infty}^{-2} c_j \oint_{\gamma_k} (z - z_k)^j dz + c_{-1} \oint_{\gamma_k} \frac{dz}{z - z_k} + \sum_{j=0}^{\infty} c_j \oint_{\gamma_k} (z - z_k)^j dz$$

Thanks to the Cauchy theorem we already know that the first and last integrals on the RHS must be null, therefore

$$\oint_{\gamma} f(z) dz = \sum_{k=0}^n c_{-1} \oint_{\gamma_k} \frac{dz}{z - z_k}$$

Recognizing the definition of residue and the winding number of the curve, we have the assert

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{k=0}^n n(\gamma, z_k) \operatorname{Res}_{z=z_k} f(z)$$

□

Definition 8.2.5 (Residue at Infinity). Given a function $f : G \rightarrow \mathbb{C}$ and a piecewise smooth closed curve γ . If $f \in H(\{\gamma\} \cup \operatorname{extr}\{\gamma\})$ we have

$$\oint_{\gamma} f(z) dz = -2\pi i \operatorname{Res}_{z=\infty} f(z) = 2\pi i \operatorname{Res}_{z=0} \frac{1}{z^2} f\left(\frac{1}{z}\right) \quad (8.24)$$

Theorem 8.10. *Given a function $f : G \rightarrow \mathbb{C}$ as before, if the function has z_k singularities with $k = 1, \dots, n$*

$$\operatorname{Res}_{z=\infty} f(z) = \sum_{k=1}^n \operatorname{Res}_{z=z_k} f(z) \quad (8.25)$$

§§ 8.2.2 Classification of Singularities, Zeros and Poles

Definition 8.2.6 (Pole). Given a function $f(z)$ with an isolated singular point in $z_0 \in \mathbb{C}$, we have that in $A_{0r}(z_0)$ the function can be expanded with a Laurent series

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k + \sum_{k=1}^{\infty} \frac{b_k}{(z - z_0)^k}$$

The point z_0 is called a *pole of order m* if $b_k = 0 \ \forall k > m$

Definition 8.2.7 (Removable Singularity). Given $f(z), z_0$ as before, we have that z_0 is a *removable singularity* if $b_k = 0 \ \forall k \geq 1$

Definition 8.2.8 (Essential Singularity). Given $f(z), z_0$ as before, we have that z_0 is an *essential singularity* if $b_k \neq 0$ for infinite values of k

Definition 8.2.9 (Meromorphic Function). Let $f : G \subset \mathbb{C} \rightarrow \mathbb{C}$ be a function. f is said to be *meromorphic* if $f \in H(\tilde{G})$ where $\tilde{G} = G \setminus \{z_1, \dots, z_n\}$ where $z_k \in G$ are poles of the function

Theorem 8.11. Let z_0 be an isolated singularity of a function $f(z)$. z_0 is a pole of order m if and only if

$$f(z) = \frac{1}{(m-1)!} \left. \frac{d^{m-1}g}{dz^{m-1}} \right|_{z_0} \quad g \in H(B_\epsilon(z_0)) \quad \epsilon > 0 \quad (8.26)$$

Proof. Let $f : G \rightarrow \mathbb{C}$ be a meromorphic function and $g : G \rightarrow \mathbb{C}$, $g \in H(G)$ where $f(z)$ has a pole in $z_0 \in G$ and $g(z_0) \neq 0$

$$f(z) = \frac{g(z)}{(z - z_0)^m}$$

Since $g(z)$ is holomorphic in z_0 we have that, for some r

$$g(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k g}{dz_0^k} (z - z_0)^k \quad z \in B_r(z_0)$$

And therefore, $\forall z \in A_{0r}(z_0)$

$$f(z) = \frac{1}{(z - z_0)^m} g(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k g}{dz_0^k} (z - z_0)^{k-m}$$

Since $g(z_0) \neq 0$ we have the assert.

Alternatively we start by hypothesizing that z_0 is already a pole of order m for f , and therefore we can write the following Laurent expansion for some $r > 0$

$$f(z) = \sum_{k=-m}^{\infty} c_k (z - z_0)^k \quad \forall z \in A_{0r}(z_0)$$

Where $c_{-m} \neq 0$. Therefore, we write

$$g(z) = \begin{cases} (z - z_0)^m f(z) & z \in A_{0r}(z_0) \\ c_{-m} & z = z_0 \end{cases}$$

And, expanding $g(z)$ for $z \in B_r(z_0)$ we obtain

$$g(z) = c_{-m} + c_{-m+1}(z - z_0) + \cdots + c_{-1}(z - z_0)^{m-1} + \sum_{k=0}^{\infty} c_k(z - z_0)^{k+m}$$

$g(z)$ is holomorphic in the previous domain of expansion, and therefore we have, since the Taylor expansion is unique

$$c_{-1} = \frac{1}{(m-1)!} \frac{d^{m-1}g}{dz_0^{m-1}} = \operatorname{Res}_{z=z_0} f(z)$$

□

Definition 8.2.10 (Zero). Let $f : G \rightarrow \mathbb{C}$ be a holomorphic function. Taken $z_0 \in G$, it's said to be a *zero of order m* if

$$\begin{cases} \frac{d^k f}{dz_0^k} = 0 & k = 1, \dots, m-1 \\ \frac{d^m f}{dz_0^m} \neq 0 \end{cases}$$

Theorem 8.12. The point $z_0 \in G$ is a zero of order m for f if and only if

$$f(z) = \frac{g(z)}{(z - z_0)^m} \quad g(z_0) \neq 0, \quad g \in H(G)$$

Proof. Taken $f(z) = (z - z_0)^m g(z)$ such that $g(z_0) \neq 0$ we can expand $g(z)$ with Taylor and at the end obtain

$$f(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k g}{dz_0^k} (z - z_0)^{k+m}$$

Since this is a Taylor expansion also for $f(z)$ we have that, for $j = 1, \dots, m-1$

$$\frac{d^j f}{dz_0^j} = 0 \quad \frac{d^m f}{dz_0^m} = m!g(z_0) \neq 0$$

The same is obtainable with the vice versa demonstrating the theorem

□

Notation. Let f be a meromorphic function. We will define the following sets of points accordingly

1. Z_f^m as the set of zeros of order m
2. S_f as the set of isolated singularities of f
3. P_f^m as the set of poles of order m

We immediately see some special cases

1. P_f^∞ is the set of essential singularities of f
2. P_f^1 is the set of removable singularities of f

Theorem 8.13. Let $f : D \rightarrow \mathbb{C}$ be a function such that $f \in H(D)$, with D an open set, then

1. $f(z) = 0 \quad \forall z \in D$
2. $\exists z_0 : f^{(k)}(z_0) = 0 \quad \forall k \geq 0$
3. $Z_f \subset D$ has a limit point

Proof. 3) \implies 2)

Take $z_0 \in D$ as the limit point of Z_f . Since $f \in C(D)$ we have that $z_0 \in Z_f^m$. therefore

$$f(z) = (z - z_0)^m g(z) \quad g(z_0) \neq 0, \quad g \in H(D) \implies \exists \delta > 0 : g(z) \neq 0 \quad \forall z \in B_\delta(z_0)$$

Therefore

$$f(z) \neq 0 \quad \forall z \in A_{0\delta}(z_0) \not\equiv$$

2) \implies 1)

Suppose that $Z_{f^{(k)}} := \{z \in D \mid f^{(k)}(z) = 0\} \neq \{\}$. We have to demonstrate that this set is clopen in D .

Take $z \in \overline{Z_{f^{(k)}}}$ and a sequence $(z)_k \in Z_{f^{(k)}}$ such that $z_k \rightarrow z$. We have then

$$f^{(k)}(z) = \lim_{k \rightarrow \infty} f^{(k)}(z_k) = 0$$

Therefore $Z_{f^{(k)}} = \overline{Z_{f^{(k)}}}$ and the set is closed.

Take then $z \in Z_{f^{(k)}} \subset D$, since D is open we have that $\exists r > 0 : B_r(z) \subset D$, therefore

$$\forall w \in B_r(z), \quad z \neq w \quad f(w) = \sum_{k=0}^{\infty} a_k (w - z)^k = 0 \implies \begin{cases} z = w \\ a_k = 0 \quad \forall k \geq 0 \end{cases}$$

Since $w \neq z$ we have that $B_r(z) \subset Z_{f^{(k)}}$ and the set is open. Taking both results we have that the set is clopen and $D = Z_{f^{(k)}}$ \square

Corollary 8.2.1. Let $f, g : D \rightarrow \mathbb{C}$ and $f, g \in H(D)$. We have that $f = g$ iff the set $\{f(z) = g(z)\}$ has a limit point in D

Corollary 8.2.2 (Zeros of Holomorphic Functions). Let $f : D \rightarrow \mathbb{C}$ be a non-constant function $f \in H(D)$ with D an open connected set. Then

$$\forall z \in Z_f^m \quad m < \infty$$

Proof. Take $z_0 \in Z_f$, then since f is non-constant we have that Z_f has no limit points in D , therefore

$$\exists \delta > 0 : f(z) \neq 0 \quad \forall z \in A_{0\delta}(z_0) \quad \wedge \quad \exists m \geq 1 : \frac{d^k f}{dz_0^k} = 0 \quad k \in [0, m), \quad \frac{d^m f}{dz_0^m} \neq 0$$

Therefore $z_0 \in Z_f^m$ \square

Theorem 8.14. Let $f : D \rightarrow \mathbb{C}$ be a meromorphic function, such that

$$f(z) = \frac{p(z)}{q(z)} \quad p, q \in H(D)$$

If $z_0 \in Z_q^m$ such that $p(z_0) \neq 0$, then $z_0 \in P_f^m$

Proof. $z_0 \in Z_q^m$ is an isolated singularity of q , therefore

$$\exists \delta > 0 : q(z) \neq 0 \quad \forall z \in A_{0\delta}(z_0) \quad \therefore z_0 \in S_{p/q}$$

We therefore can take $q(z) = (z - z_0)^m g(z)$ and we have

$$f(z) = \frac{p(z)}{g(z)(z - z_0)^m} = \frac{h(z)}{(z - z_0)^m}$$

Where $h(z)$ is a holomorphic function such that $h(z_0) \neq 0$. By definition of pole we have $z_0 \in P_f^m$ \square

Theorem 8.15 (Quick Calculus of Residues for Rational Functions). *If $f(z) = p(z)/q(z)$ as before, there is a quick rule of thumb for calculating the residue in z_0 . We can write*

$$\text{Res}_{z=z_0} f(z) = \frac{1}{(m-1)!} \frac{d^{m-1} h}{dz_0^{m-1}}$$

If the pole is a removable singularity, we have $z_0 \in P_f^1$ and

$$\text{Res}_{z=z_0} f(z) = \frac{p(z_0)}{q'(z_0)}$$

Theorem 8.16. *Let f be a meromorphic function. If $z_0 \in P_f^m$ we have*

$$\lim_{z \rightarrow z_0} f(z) = \infty$$

Proof.

$$z_0 \in P_f^m \implies f(z) = \frac{g(z)}{(z - z_0)^m}, \quad z_0 \notin Z_g$$

Then

$$\lim_{z \rightarrow z_0} \frac{1}{f(z)} = \lim_{z \rightarrow z_0} \frac{(z - z_0)}{g(z)} = 0$$

\square

Theorem 8.17. *If $z_0 \in P_f^1$, $\exists \epsilon > 0$ such that $f \in A_{0\epsilon}(z_0)$ and $\|f(z)\| \leq M$, $\forall z \in A_{0\epsilon}(z_0)$*

Proof. By definition we have that

$$\exists r > 0 : f \in H(A_{0r}(z_0))$$

And therefore the function is Laurent representable in this set as follows

$$f(z) = \sum_{k=0}^{\infty} c_k (z - z_0)^k \quad 0 < \|z - z_0\| < \epsilon$$

Taken the following holomorphic function

$$g(z) = \begin{cases} f(z) & z \in A_{0\epsilon}(z_0) \\ \sum_{k=0}^{\infty} c_k (z - z_0)^k & z = z_0 \end{cases}$$

We have that $g \in H(\overline{B_\epsilon}(z_0))$ and therefore $\|f(z)\| \leq M \quad \forall z \in A_{0\epsilon}(z_0)$ \square

Lemma 8.2.1 (Riemann). Take a function $f \in H(A_{0\epsilon}(z_0))$ for some $\epsilon > 0$, then if $\|f(z)\| \leq M \forall z \in A_{0\epsilon}(z_0)$

The point z_0 is a removable singularity for f

Proof. In the set of holomorphy the function is representable with Laurent, therefore

$$f(z) = \sum_{k=0}^{\infty} c_k^+(z - z_0)^k + \sum_{k=1}^{\infty} \frac{c_k^-}{(z - z_0)^k}$$

We have that the coefficients c_k^- are the following, where we integrate over a curve $\{\gamma\} := \{z \in \mathbb{C} \mid \|z - z_0\| = \rho < \epsilon\}$

$$c_k^- = \frac{1}{2\pi i} \oint_{\gamma} f(z)(z - z_0)^{k-1} dz$$

The function is limited, and therefore for Darboux

$$c_k^- \leq \rho^k M \rightarrow 0 \quad \forall k \geq 1$$

Therefore $z_0 \in P_f^1$ □

Theorem 8.18 (Quick Calculus Methods for Residues). *Let f be a meromorphic function, then*

1. $z_0 \in P_f^n$ then

$$\text{Res}_{z=z_0} f(z) = \frac{1}{(n-1)!} \lim_{z \rightarrow z_0} \frac{d^{n-1}}{dz^{n-1}} (z - z_0)^n f(z) \quad (8.27)$$

2. $z_0 \in P_f^m$ and $f(z) = p(z)/(z - z_0)^m$, where $p \in \mathbb{C}_k[z]$ with $k \leq m - 2$ and $p(z_0) \neq 0$, then

$$\text{Res}_{z=z_0} f(z) = \text{Res}_{z=z_0} \frac{p(z)}{(z - z_0)^m} = 0$$

§ 8.3 Applications of Residue Calculus

§§ 8.3.1 Improper Integrals

Definition 8.3.1 (Improper Integral). An *improper integral* is defined as the integral of a function in a domain where such function has a divergence, or where the interval is infinite. Some examples of such integrals, given a function $f(x)$ with divergences at $a, b \in \mathbb{R}$ are the following

$$\begin{aligned} \int_c^\infty f(x) dx &= \lim_{R \rightarrow \infty} \int_c^R f(x) dx \\ \int_{-\infty}^d f(x) dx &= \lim_{R \rightarrow \infty} \int_{-R}^d f(x) dx \\ \int_{-\infty}^\infty f(x) dx &= \int_{\mathbb{R}} f(x) dx = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx \\ \int_a^b f(x) dx &= \lim_{\epsilon \rightarrow 0^+} \int_{a+\epsilon}^{b-\epsilon} f(x) dx \\ \int_e^h f(x) dx &= \lim_{\epsilon \rightarrow 0^+} \left(\int_e^{a-\epsilon} f(x) dx + \int_{a+\epsilon}^h f(x) dx \right) \quad a \in (e, h) \end{aligned}$$

Definition 8.3.2 (Cauchy Principal Value). The previous definitions give rise to the following definition, the *Cauchy principal value*. Given an improper integral we define the Cauchy principal value as follows

Let $f(x)$ be a function with a singularity $c \in (a, b)$, and $g(x)$ another function then

$$\begin{aligned} \text{PV} \int_{-\infty}^{\infty} g(x) dx &= \text{PV} \int_{\mathbb{R}} g(x) dx = \lim_{R \rightarrow \infty} \int_{-R}^R g(x) dx \\ \text{PV} \int_a^b f(x) dx &= \lim_{\epsilon \rightarrow 0^+} \left(\int_a^{c-\epsilon} f(x) dx + \int_{c+\epsilon}^b f(x) dx \right) \end{aligned}$$

In the first case. PV is usually omitted.

For a complex integral, if $\gamma_R(t) = Re^{it}$ is a circumference, we have

$$\text{PV} \int_{\gamma_R} f(z) dz = \lim_{R \rightarrow \infty} \int_{\gamma_R} f(z) dz$$

Notation (Circumferences and Parts of Circumference). For a quick writing of the integrals in this section, we will use this notation for the following circumferences

$$\begin{aligned} C_R(t) &= Re^{it} \quad t \in [0, 2\pi] \\ C_{R\alpha\beta} &= Re^{it} \quad t \in [\alpha, \beta] \\ C_R^+(t) &= Re^{it} \quad t \in [0, \pi] \\ C_R^-(t) &= Re^{-it} \quad t \in [0, \pi] \\ \tilde{C}_R^\pm &= C_R^\pm \times [-R, R] \end{aligned}$$

Hypothesis 1. Let $R_0 > 0$ and $f \in C(D)$, where $D := \{z \in \mathbb{C} \mid \|z\| \geq R_0\} \cup \mathbb{R}$ and

$$\lim_{z \rightarrow \infty} zf(z) = 0$$

Hypothesis 2. Let $R_0 > 0$ and $f \in C(D)$, where $D := \{z \in \mathbb{C} \mid \|z\| \geq R_0\} \cup \mathbb{R}$ and

$$\lim_{z \rightarrow \infty} f(z) = 0$$

Theorem 8.19. If (1) holds true, then

$$\text{PV} \int_{\gamma_R} f(z) dz = 0 \quad \gamma_R = C_R, C_R^+, C_R^- \quad (8.28)$$

Also, if $f(x)$ is a real function

$$\int_{\mathbb{R}} f(x) dx = \text{PV} \int_{\tilde{C}_R^+} f(z) dz = \text{PV} \int_{\tilde{C}_R^-} f(z) dz \quad (8.29)$$

Theorem 8.20. Let $f(z)$ be an even function, if (1) holds we have

$$\int_0^\infty f(x) dx = \frac{1}{2} \text{PV} \int_{\tilde{C}_R^+} f(z) dz = \frac{1}{2} \text{PV} \int_{\tilde{C}_R^-} f(z) dz \quad (8.30)$$

Theorem 8.21. Let $f(z) = g(z^k)$, $k \geq 2$. If (1) holds

$$\int_0^\infty f(x) dx = \frac{1}{1 - e^{\frac{2i\pi}{k}}} \text{PV} \int_{\tilde{C}_{R0, 2\pi/k}} f(z) dz \quad (8.31)$$

Theorem 8.22. If (2) holds

$$\begin{aligned} \int_{\mathbb{R}} f(x) e^{i\lambda x} dx &= \text{PV} \int_{\tilde{C}_R^+} f(z) e^{i\lambda z} dz \quad \lambda > 0 \\ \int_{\mathbb{R}} f(x) e^{i\lambda x} dx &= \text{PV} \int_{\tilde{C}_R^-} f(z) e^{i\lambda z} dz \quad \lambda > 0 \end{aligned} \quad (8.32)$$

From this, we can write then, for $\lambda > 0$

$$\begin{aligned} \int_{\mathbb{R}} f(x) \cos(i\lambda x) dx &= \Re \left(\text{PV} \int_{\tilde{C}_R^+} f(z) e^{i\lambda z} dz \right) \quad \lambda > 0 \\ \int_{\mathbb{R}} f(x) \sin(i\lambda x) dx &= \Im \left(\text{PV} \int_{\tilde{C}_R^+} f(z) e^{i\lambda z} dz \right) \quad \lambda > 0 \end{aligned} \quad (8.33)$$

Hypothesis 3. Let $f(z) = g(z)h(z)$ with $g(z)$ a meromorphic function such that $S_g \not\subset \mathbb{R}^+$ and

1. $h \in H(\mathbb{C} \setminus \mathbb{R}^+)$
2. $\lim_{z \rightarrow \infty} z f(z) = 0$
3. $\lim_{z \rightarrow 0} z f(z) = 0$

Definition 8.3.3 (Pacman Path). Let $\Gamma_{Rr\epsilon}$ be what we will call as the *pacman path*, this path is formed by 4 different paths

$$\begin{aligned} \gamma_1(t) &= r e^{it} \quad t \in [\epsilon, 2\pi - \epsilon] \\ \gamma_2 &= [-R, R] \\ \gamma_3(t) &= R e^{it} \quad t \in [\epsilon, 2\pi - \epsilon] \\ \gamma_4 &= [-R, R] \end{aligned} \quad (8.34)$$

We will abbreviate this as Γ

Theorem 8.23. Given $f(x)$ a function such that (3) holds, we have that

$$\int_0^\infty g(x) \Delta h(x) dx = \text{PV} \int_{\Gamma} g(z) h(z) dz \quad (8.35)$$

Where

$$\Delta h(x) = \lim_{\epsilon \rightarrow 0^+} (h(x + i\epsilon) - h(x - i\epsilon)) \quad (8.36)$$

In general, we have the following conversion table

$h(z)$	$\Delta h(x)$
$-\frac{1}{2\pi i} \log_+(z)$	1
$\log_+(z)$	$-2\pi i$
$\log_+^2(z)$	$-2\pi i \log(x) + 4\pi^2$
$\log_+(z) - 2\pi i \log_+(z)$	$-4\pi i \log(x)$
$\frac{i}{4\pi} \log_+^2(z) + \frac{1}{2} \log_+(z)$	$\log(x)$
$[z^\alpha]^+$	$x^\alpha (1 - e^{2\pi i \alpha})$

All the previous integrals are solved through a direct application of the residue theorem.

§§ 8.3.2 General Rules

Theorem 8.24 (Integrals of Trigonometric Functions). *Let $f(\cos \theta, \sin \theta)$ be some rational function of cosines and sines. Then we have that*

$$\int_0^{2\pi} f(\cos \theta, \sin \theta) d\theta = \int_{\|z\|=1} f\left(\frac{z+z^{-1}}{2}, \frac{z-z^{-1}}{2i}\right) \frac{dz}{iz} \quad (8.38)$$

Theorem 8.25 (Integrals of Rational Functions). *Let $f(x) = p_n(x)/q_m(x)$ with $m \geq n+2$ and $q_m(x) \neq 0 \quad \forall x \in \mathbb{R}$, then*

$$\int_{\mathbb{R}} \frac{p_n(x)}{q_m(x)} dx = 2\pi i \sum_k \operatorname{Res}_{z=z_k} \frac{p_n(z)}{q_m(z)} \quad (8.39)$$

Lemma 8.3.1 (Jordan's Lemma). *Let $f(z)$ be a holomorphic function in $A := \{z \in \mathbb{C} \mid \|z\| > R_0, \Im(z) \geq 0\}$.*

Taken $\gamma(t) = Re^{it}$ $0 \leq t \leq \pi$ with $R > R_0$.

If $\exists M_R > 0 : \|f(z)\| \leq M_R \quad \forall z \in \{\gamma\}$ and $M_R \rightarrow 0$, we have that

$$\operatorname{PV} \int_{\gamma} f(z) e^{iaz} dz = 0 \quad a > 0 \quad (8.40)$$

Theorem 8.26. *Let $f(x) = p_n(x)/q_m(x)$ and $m \geq n+1$ with $q_m(x) \neq 0 \quad \forall x \in \mathbb{R}$, then $\forall a > 0$ we have that*

$$\int_{\mathbb{R}} \frac{p_n(x)}{q_m(x)} e^{iax} dx = 2\pi i \sum_k \operatorname{Res}_{z=z_k} \frac{p_n(z)}{q_m(z)} e^{iaz} \quad (8.41)$$

Lemma 8.3.2. Let $f(z)$ be a meromorphic function such that $z_0 \in P_f^1$ and γ_r^\pm are semi circumferences parametrized as follows

$$\gamma_r^\pm(t) = z_0 + re^{\pm i\theta} \quad \theta \in [-\pi, 0]$$

Then

$$\text{PV} \int_{\gamma_r^\pm} f(z) dz = \pm \pi i \text{Res}_{z=z_0} f(z) \quad (8.42)$$

Theorem 8.27. Let $f(x) = p_n(x)/q_m(x)$ with $m \geq n + 2$ and $q_m(x)$ has $x_j \in Z_g^1|_{\mathbb{R}}$ then

$$\int_{\mathbb{R}} \frac{p_n(x)}{q_m(x)} dx = 2\pi i \sum_k \text{Res}_{z=z_k} \frac{p_n(z)}{q_m(z)} + \pi i \sum_j \text{Res}_{z=x_j} \frac{p_n(z)}{q_m(z)} \quad (8.43)$$

If $g(x) = r_\alpha(x)/s_\beta(x)e^{iax}$ and $\beta \geq \alpha + 1$ with $x_j \in Z_g^1|_{\mathbb{R}}$, then $\forall a > 0$

$$\int_{\mathbb{R}} \frac{r_\alpha(x)}{s_\beta(x)} e^{iax} dx = 2\pi i \sum_k \text{Res}_{z=z_k} \frac{r_\alpha(z)}{s_\beta(z)} e^{iaz} + \pi i \sum_j \text{Res}_{z=x_j} \frac{r_\alpha(z)}{s_\beta(z)} e^{iaz} \quad (8.44)$$

z_k are all the zeros of q, s contained in the plane $\{\Im(z) > 0\}$

9 Hilbert Spaces

§ 9.1 Hilbert Spaces

Definition 9.1.1 (Hermitian Product). Given \mathcal{V} a complex vector space, and an application $\langle \cdot, \cdot \rangle : \mathcal{V} \longrightarrow \mathbb{C}$ such that $\forall u, v, z \in \mathcal{V}, c, d \in \mathbb{C}$

1. $\langle v, v \rangle \geq 0$
2. $\langle v, v \rangle = 0 \iff v = 0$
3. $\langle u, v \rangle = \overline{\langle v, u \rangle}$
4. $\langle u + v, z \rangle = \langle u, z \rangle + \langle v, z \rangle$
5. $\langle cu, dv \rangle = c\bar{d}\langle u, v \rangle$

The application $\langle \cdot, \cdot \rangle$ is called an *Hermitian product* in \mathcal{V} , and the couple $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ is called an *Euclidean space*

Remark. It's usual in physics that for a Hermitian product, we have that

$$\langle cu, v \rangle = \bar{c}\langle u, v \rangle \quad (9.1)$$

Definition 9.1.2 (Hilbert Space). Given $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ an euclidean space. It's said to be a *Hilbert space* if it's complete

Theorem 9.1 (Cauchy-Schwartz Inequality). *Given $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ a complex euclidean space, then $\forall u, v \in \mathcal{V}$*

$$\|\langle u, v \rangle\|^2 \leq \langle u, u \rangle \langle v, v \rangle \quad (9.2)$$

Proof. Taken $t \in \mathbb{C}$, we define $p(t) = \langle tu + v, tu + v \rangle$. Then by definition of the Hermitian product, we have

$$p(t) = \|t\|^2 \langle u, u \rangle + t \langle u, v \rangle + \bar{t} \langle v, u \rangle + \langle v, v \rangle$$

Writing $\langle u, v \rangle = \rho e^{i\theta}$, $t = se^{-i\theta}$ we have

$$p(se^{-i\theta}) = s^2 \langle u, u \rangle + 2s\rho + \langle v, v \rangle \geq 0 \quad \forall s \in \mathbb{R}$$

Then, by definition, we have

$$\rho^2 = \|\langle u, v \rangle\|^2 \leq \langle u, u \rangle \langle v, v \rangle$$

□

Theorem 9.2 (Induced Norm). *Given a Hermitian product $\langle \cdot, \cdot \rangle$ we can define an induced norm $\|\cdot\|$ by the definition*

$$\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle} \quad (9.3)$$

Theorem 9.3. *Addition, multiplication by a scalar and the scalar product are continuous in an euclidean space \mathcal{V} , then, given two sequences $u_n \rightarrow u \in \mathcal{V}$, $v_n \rightarrow v \in \mathcal{V}$ and*

$$\begin{aligned} u_n + v_n &\rightarrow u + v \\ c \in \mathbb{C} &\implies cu_n \rightarrow u \\ \langle u_n, v_n \rangle &\rightarrow \langle u, v \rangle \end{aligned}$$

Proof. Thanks to Cauchy-Schwartz we have

$$\begin{aligned} |\langle u, v \rangle - \langle u_n, v_n \rangle| &= |\langle u, v \rangle - \langle u, v_n \rangle + \langle u, v_n \rangle - \langle u_n, v_n \rangle| \leq |\langle u, v \rangle - \langle u, v_n \rangle| + |\langle u, v_n \rangle - \langle u_n, v_n \rangle| = \\ &= \|u\| \|v - v_n\| + \|v_n\| \|u - u_n\| \end{aligned}$$

Since the successions are convergent, we have that $\exists M > 0 : \|v_n\| \leq M \forall n \in \mathbb{N}$, therefore

$$|\langle u, v \rangle - \langle u_n, v_n \rangle| \leq \max\{\|u\|, M\} (\|v - v_n\| + \|u - u_n\|) \rightarrow 0$$

□

Example 9.1.1 (Some Euclidean Spaces). 1) $\ell^2(\mathbb{C})$

Given $x, y \in \ell^2(\mathbb{C})$ we define the scalar product as

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

2) $\ell^2(\mu)$, a weighted sequence space, where

$$\ell^2(\mu) := \left\{ x \in \mathbb{C}^{\mathbb{N}} \mid \sum_{i=1}^{\infty} \mu_i |x_i|^2 < \infty, \mu_i \in \mathbb{R}, \mu_i > 0 \forall i \right\}$$

Given $x, y \in \ell^2(\mu)$ we define

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \mu_i x_i \overline{y_i}$$

3) $C_2(\mathbb{C})$ Given $f, g \in C_2(\mathbb{C})$ we define

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x) \overline{g(x)} dx$$

4) $C_2(\mathbb{C}, p(x) dx)$, weighted function spaces, where

$$C_2(\mathbb{C}, p(x) dx) := \left\{ f \in C(\mathbb{C}) \mid \int_{\mathbb{R}} f(x) \overline{f(x)} p(x) dx < \infty, p(x) \in C(\mathbb{R}; \mathbb{R}^+) \right\}$$

Given $f, g \in C_2(\mathbb{C}, p(x) dx)$ we define

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x) \overline{g(x)} p(x) dx$$

The spaces C_2 aren't complete therefore they're not Hilbert spaces. The spaces $L^2(\mathbb{C})$ and the weighted alternative are the completion of such spaces and are therefore Hilbert spaces

Theorem 9.4 (Polarization Identity). *Given a complex euclidean space $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ we have, $\forall u, v \in \mathcal{V}$*

$$\langle u, v \rangle = \frac{1}{4} \left(\|u + v\|^2 - \|u - v\|^2 + i \left(\|u + iv\|^2 - \|u - iv\|^2 \right) \right) \quad (9.4)$$

Theorem 9.5 (Parallelogram Rule). *Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space. A necessary and sufficient condition that the norm is induced by a scalar product is that*

$$\|u + v\|^2 + \|u - v\|^2 = 2 \left(\|u\|^2 + \|v\|^2 \right) \quad \forall u, v \in \mathcal{V} \quad (9.5)$$

§ 9.2 Projections

Definition 9.2.1 (Angle). *Given a real euclidean space $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ we define the angle $\theta = \angle u, v$ as follows*

$$\theta = \arccos \left(\frac{\langle u, v \rangle}{\|u\| \|v\|} \right) \quad (9.6)$$

Definition 9.2.2 (Orthogonal Complement). *Given an euclidean vector space \mathcal{V} and two vectors u, v , we say that the two vectors are orthogonal $u \perp v$ if*

$$\langle u, v \rangle = 0 \quad (9.7)$$

If $X \subset \mathcal{V}$ and $\forall x \in X$ we have that

$$\langle u, x \rangle = 0$$

We say that $u \in X^\perp$ where this space is called the *Orthogonal Complement* of X , i.e.

$$X^\perp := \{ v \in \mathcal{V} \mid \langle v, w \rangle = 0 \ \forall w \in X \} \quad (9.8)$$

Theorem 9.6. *Given $X \subset \mathcal{V}$ with \mathcal{V} an euclidean space, the set X^\perp is a closed subspace of \mathcal{V}*

Proof. X^\perp is a subspace, hence $\forall v_1, v_2 \in X^\perp$ and $c_1, c_2 \in \mathbb{C}$ we have

$$\langle c_1 v_1 + c_2 v_2, w \rangle = c_1 \langle v_1, w \rangle + c_2 \langle v_2, w \rangle \quad \forall w \in X$$

Hence $c_1 v_1 + c_2 v_2 \in X^\perp$.

Given a sequence $(v)_n \in X^\perp : (v)_n \rightarrow v \in \mathcal{V}$ we have, given $w \in X$

$$\langle v_n, w \rangle = 0 \quad \forall n \in \mathbb{N}$$

Due to the continuity of the scalar product we have that

$$\lim_{n \rightarrow \infty} \langle v_n, w \rangle = 0 = \langle v, w \rangle$$

Therefore $v \in X^\perp$ and the subspace X^\perp is closed in \mathcal{V} □

Theorem 9.7. *Given $X, Y \subset \mathcal{V}$ with \mathcal{V} an euclidean space, we have*

$$\begin{aligned} X \subset Y &\implies Y^\perp \subset X^\perp \\ X^\perp &= (\overline{X})^\perp = \left(\overline{\text{span}(X)} \right)^\perp \end{aligned}$$

Proof. Taken $v \in Y^\perp$ we have by definition

$$\langle v, y \rangle = 0 \quad \forall y \in Y$$

Since $X \subset Y$ we have then

$$\langle v, x \rangle = 0 \quad \forall x \in X$$

Therefore $Y^\perp \subset X^\perp$.

By definition we have that $X \subset \overline{X} \subset \overline{\text{span}(X)}$, and thanks to the previous proof

$$X^\perp \supset (\overline{X})^\perp \supset (\overline{\text{span}(X)})^\perp$$

Taken $w \in \text{span}(X)$ we have

$$w = \sum_i c_i w_i \quad w_i \in X$$

And given $v \in X^\perp$, we get

$$\langle v, w \rangle = \sum_i c_i \langle v, w_i \rangle = 0$$

Now take $w \in \overline{\text{span}(X)}$. Take a sequence $(w)_n \in \overline{\text{span}(X)}$ such that $(w)_n \rightarrow w$. Thanks to the continuity of the scalar product we have

$$\langle v, w \rangle = \langle v, \lim_{n \rightarrow \infty} w_n \rangle = \lim_{n \rightarrow \infty} \langle v, w_n \rangle = 0$$

Demonstrating that $X^\perp = (\overline{\text{span}(X)})^\perp$ □

Lemma 9.2.1. Let \mathcal{V} be a Hilbert space. Given $\mathcal{W} \subset \mathcal{V}$ a closed subspace. Given $v \in \mathcal{V}$

$$\exists! w_0 \in \mathcal{W} : \forall w \in \mathcal{W} \quad d = \|v - w_0\| \leq \|v - w\|$$

Proof. Take $d = \inf_{w \in \mathcal{W}} \|v - w\|$. By definition of infimum we have that $\exists (w)_n \in \mathcal{W}$ such that

$$\lim_{n \rightarrow \infty} \|v - w_n\| = d$$

Using the parallelogram rule, we have that

$$\|w_n - w_k\|^2 = \|(w_n - v) + (v - w_k)\|^2 = 2\|v - w_n\|^2 + 2\|v - w_k\|^2 - 4\left\|\frac{1}{2}(w_n + w_k) - v\right\|^2$$

Since $\frac{1}{2}(w_n + w_k) \in \mathcal{W}$ we have by definition of d

$$\left\|\frac{1}{2}(w_n + w_k) - v\right\| \geq d$$

Therefore, we can rewrite

$$\|w_n - w_k\|^2 \leq 2\|v - w_n\|^2 + 2\|v - w_k\|^2 - 4d^2$$

Therefore, we have

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} : \forall n, k \geq N \quad \|w_n - w_k\|^2 \leq 4(d + \epsilon)^2 - 4d^2$$

Hence $(w)_n$ is a Cauchy sequence. Since by definition \mathcal{V} is complete and $\mathcal{W} \subset \mathcal{V}$ is closed, we have that \mathcal{W} is also complete, therefore $(w)_n \rightarrow w_0 \in \mathcal{W}$ and we have

$$\|v - w_0\| = d$$

Now suppose that $\exists w_1, w_2 \in \mathcal{W}$ such that the previous is true, i.e.

$$\|v - w_1\| = \|v - w_2\| \leq \|v - w\| \quad \forall w \in \mathcal{W}$$

Taken $w_3 = 1/2(w_1 + w_2)$ we have

$$\|v - w_3\|^2 = \|v - w_1\|^2 - \frac{1}{4}\|w_2 - w_1\|^2$$

Taken $d = \|v - w_1\| = \|v - w_2\|$, $z_1 = v - w_3$ and $z_2 = 1/2(w_1 - w_2)$ we get

$$\|z_1 + z_2\|^2 + \|z_1 - z_2\|^2 = 2(\|z_1\|^2 + \|z_2\|^2)$$

And therefore

$$d^2 = \frac{1}{2}(\|v - w_1\|^2 + \|v - w_2\|^2) = \|v - w_3\|^2 + \frac{1}{4}\|w_1 - w_2\|^2$$

I.e. if $w_1 \neq w_2$, w_3 is the infimum between $v \in \mathcal{V}$ and $\mathcal{W} \nsubseteq$

□

Theorem 9.8 (Projection). *Given $\mathcal{W} \subset \mathcal{V}$ closed subspace of a Hilbert space, we have*

$$v = w + z \quad \forall v \in \mathcal{V}, w \in \mathcal{W}, z \in \mathcal{W}^\perp$$

Proof. Given $v \in \mathcal{V}$, due to the previous lemma we have that $\exists! w \in \mathcal{W}$ such that

$$d = \|v - w\| \leq \|v - w'\| \quad \forall w' \in \mathcal{W}$$

Taken $z = v - w$, and an element $x \in \mathcal{W}$, define the vector $w + tx$ with $t \in \mathbb{C}$. Since \mathcal{W} is a subspace $w + tx \in \mathcal{W}$ and $\forall t \in \mathbb{C}$ we have

$$d^2 \leq \|v - (w + tx)\|^2 = \|v - w\|^2 - \bar{t}\langle v - w, x \rangle - t\langle x, v - w \rangle + \|t\|^2\|x\|^2$$

Writing $\langle x, v - w \rangle = \|\langle x, v - w \rangle\|e^{i\theta}$ and $t = se^{-i\theta}$ with $s \in \mathbb{R}$ we have

$$-2s\|\langle v - w, x \rangle\| + s^2\|x\|^2 \geq 0 \quad \forall s \in \mathbb{R}$$

Which implies

$$\langle v - w, x \rangle = 0 \implies z = v - w \in \mathcal{W}^\perp$$

Therefore there exists a representation $v = w + z$ with $w \in \mathcal{W}$, $z \in \mathcal{W}^\perp$. Now, we suppose that $v = w' + z'$, then

$$0 = (w - w') + (z - z')$$

Therefore

$$0 = \|(w - w') + (z - z')\|^2 = \|w - w'\|^2 + \|z - z'\|^2$$

Therefore the representation is unique.

□

Theorem 9.9. *If $\mathcal{W} \subset \mathcal{V}$ with \mathcal{V} a Hilbert space, we have*

$$(\mathcal{W}^\perp)^\perp = \overline{\mathcal{W}}$$

If \mathcal{W} is closed

$$(\mathcal{W}^\perp)^\perp = \mathcal{W}$$

Proof. Taken $w \in \mathcal{W}$ we have that $w \perp v$ with $v \in \mathcal{W}^\perp$, therefore $w \in (\mathcal{W}^\perp)^\perp$. Therefore

$$\mathcal{W} \subset (\mathcal{W}^\perp)^\perp$$

Since the space on the right is closed, we have

$$\overline{\mathcal{W}} = \overline{(\mathcal{W}^\perp)^\perp} = (\mathcal{W}^\perp)^\perp$$

Now taken $w \in (\mathcal{W}^\perp)^\perp$, since $\overline{\mathcal{W}}$ is a closed subspace by definition, we can write

$$w = v + z \quad v \in \overline{\mathcal{W}}, \quad z \in \overline{\mathcal{W}}^\perp = \mathcal{W}^\perp$$

We have $w \perp z$, and therefore

$$\|z\|^2 = \langle z, w - v \rangle = 0 \implies w = v \in \overline{\mathcal{W}}$$

□

Definition 9.2.3 (Orthogonal Projection). Given a closed subspace $\mathcal{W} \subset \mathcal{V}$ we can define an operator $\hat{\pi}_{\mathcal{W}} : \mathcal{V} \longrightarrow \mathcal{W}$ such that

$$\hat{\pi}_{\mathcal{W}}v = w \iff w \in \mathcal{W}, \quad v - w \in \mathcal{W}^\perp \quad (9.9)$$

$\hat{\pi}_{\mathcal{W}}$ is linear and called a *orthogonal projection*

Theorem 9.10. *Given $\mathcal{W} \subset \mathcal{V}$ a closed subspace of the Hilbert space \mathcal{V} , then given an orthogonal projection $\hat{\pi}_{\mathcal{W}} : \mathcal{V} \longrightarrow \mathcal{W}$ we have, $\forall v, z \in \mathcal{V}$ and another closed subspace $\mathcal{Z} \subset \mathcal{V}$*

1. $\hat{\pi}_{\mathcal{W}}^2 = \hat{\pi}_{\mathcal{W}}$
2. $\langle \hat{\pi}_{\mathcal{W}}v, z \rangle = \langle v, \hat{\pi}_{\mathcal{W}}z \rangle \geq 0$
3. If $\mathcal{Z} \subseteq \mathcal{W}^\perp$ $\hat{\pi}_{\mathcal{W}} \circ \hat{\pi}_{\mathcal{Z}} = \hat{\pi}_{\mathcal{Z}} \circ \hat{\pi}_{\mathcal{W}} = 0$
4. If $\mathcal{Z} \subset \mathcal{W}$ $\hat{\pi}_{\mathcal{W}} \circ \hat{\pi}_{\mathcal{Z}} = \hat{\pi}_{\mathcal{Z}} \circ \hat{\pi}_{\mathcal{W}} = \hat{\pi}_{\mathcal{W}}$

Definition 9.2.4 (Direct Sum). An euclidean space \mathcal{V} is called the *direct sum* of closed subspaces $\mathcal{V}_i \subset \mathcal{V}$ and it's indicated as follows

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots = \bigoplus_{k=1}^{\infty} \mathcal{V}_k \quad (9.10)$$

If

1. The spaces \mathcal{V}_k are orthogonal in couples
2. $\forall v \in \mathcal{V} \quad v = \sum_{k=1}^{\infty} v_k$ with $v_k \in \mathcal{V}_k$

Corollary 9.2.1. Given a Hilbert space \mathcal{V} and a closed subspace \mathcal{W} , then

$$\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp \quad (9.11)$$

§§ 9.2.1 Orthogonal Systems and Bases

Definition 9.2.5 (Orthogonal System). A set of vectors $X \subset \mathcal{V}$, $X \neq \{\}$ is said to be an *orthogonal system* if $\forall u, v \in X$, $u \neq v \implies u \perp v$.

Definition 9.2.6 (Orthonormal System). Given an orthogonal system $X \subset \mathcal{V}$ such that $\forall u \in X$ we have $\|u\| = 1$, the system X is called an *orthonormal system*.

Theorem 9.11. Given an orthogonal system $X \subset \mathcal{V}$ with $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ an euclidean space, then we have that X is a system of linearly independent vectors.

Definition 9.2.7 (Basis). Given an orthogonal and complete set of vectors $(v)_\alpha$ in an euclidean space \mathcal{V} it's said to be an *orthogonal basis* of \mathcal{V} . If it's an orthonormal and complete set of vectors it's said to be an *orthonormal basis* of \mathcal{V} .

Lemma 9.2.2. Given an orthogonal system $(v)_{k=1}^n \in \mathcal{V}$ and let $u \in \mathcal{V}$ an arbitrary vector. Then given $z \in \mathcal{V}$ as follows

$$z := u - \sum_{k=1}^n \frac{\langle u, v_k \rangle}{\|v_k\|^2} v_k$$

We have that $z \perp v_i \forall 1 \leq i \leq n$ and therefore $z \perp \text{span}\{v_1, \dots, v_n\}$

Proof. $\forall i = 1, \dots, n$ it's obvious that $\langle z, v_i \rangle = 0$, therefore

$$z \in \{v_1, \dots, v_n\}^\perp$$

Therefore

$$z \in \{v_1, \dots, v_n\}^\perp = \overline{\text{span}\{v_1, \dots, v_n\}}^\perp = \text{span}\{v_1, \dots, v_n\}^\perp$$

□

Theorem 9.12 (Gram-Schmidt Orthonormalization). Given \mathcal{V} an euclidean space and $(v)_{n \in \mathbb{N}} \in \mathcal{V}$ a set of linearly independent vectors. Then $\exists (u)_{n \in \mathbb{N}} \in \mathcal{V}$ orthonormal system such that

1. u_n is a linear combination of $v_i \forall 0 \leq i \leq n$, i.e.

$$u_n = \sum_{k=1}^n a_{nk} v_k \quad a_{nn} \neq 0$$

2. v_n can be written as follows

$$v_n = \sum_{k=1}^n b_{nk} u_k \quad b_{nn} \neq 0$$

Therefore, $\forall n \in \mathbb{N}$ we have that

$$\text{span}\{v_1, \dots, v_n\} = \text{span}\{u_1, \dots, u_n\}$$

Proof. Defining

$$w_n = v_n - \sum_{k=1}^{n-1} \frac{\langle v_n, v_k \rangle}{\|w_k\|^2} w_k \quad u_n = \frac{w_n}{\|w_n\|}$$

We can say immediately that $\forall n \geq 1 \ w_n \in \{w_1, \dots, w_{n-1}\}^\perp$. By induction we can say that it holds $\forall (w)_{n \in \mathbb{N}}$, therefore $(w)_n$ is an orthogonal system and $(u)_{n \in \mathbb{N}}$ is an orthonormal system. We can also say that

$$v_n = w_n + \sum_{j=1}^{n-1} \beta_{nj} w_j$$

I.e. $\forall n \geq 1 \ w_n$ is a linear combination of $\{v_1, \dots, v_n\}$, therefore, by definition

$$\text{span}\{v_1, \dots, v_n\} = \text{span}\{w_1, \dots, w_n\} = \text{span}\{u_1, \dots, u_n\}$$

□

Example 9.2.1. 1) Legendre Polynomials

Using the Gram-Schmidt orthonormalization procedure, we can find an orthonormal system $\{p_0, \dots, p_n\} \subset C_2[-1, 1]$ starting from the following system

$$(v)_n := \{1, x, x^2, x^3, x^4, \dots, x^n\}$$

The final result will be called the *Legendre Polynomials*. We begin by taking the canonical scalar product in $C_2[-1, 1]$ as follows

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) \, dx$$

And using that

$$\int_{-1}^1 x^n \, dx = \begin{cases} \frac{2}{n+1} & n = 2k \in \mathbb{N} \\ 0 & n = 2k+1 \in \mathbb{N} \end{cases}$$

Therefore, we have that

$$\begin{aligned} w_0 &= 1 \quad \|w_0\|^2 = \int_{-1}^1 dx = 2 \\ w_1 &= x - \frac{1}{2} \int_{-1}^1 x \, dx = x \quad \|w_1\|^2 = \int_{-1}^1 x^2 \, dx = \frac{2}{3} \\ w_2 &= x^2 - \frac{1}{3} \quad \|w_2\|^2 = \int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 \, dx = \frac{8}{45} \\ &\vdots \end{aligned}$$

Normalizing, we find

$$\begin{aligned} p_0 &= \frac{1}{\sqrt{2}} \\ p_1(x) &= \sqrt{\frac{3}{2}}x \\ p_2(x) &= \sqrt{\frac{45}{8}} \left(x^2 - \frac{1}{3} \right) = \frac{1}{2} \sqrt{\frac{5}{2}} (3x^2 - 1) \\ &\vdots \end{aligned}$$

And so on. The set $(p)_n$ is called the set of *Legendre polynomials* 2) Hermite Polynomials

Using the same procedure, we can find the *Hermite polynomials* $H_n(x)$ in the space $C_2(\mathbb{R}, e^{-x^2} dx)$. The first 5 are the following

$$\begin{aligned} H_1 &= 1 \\ H_2(x) &= x \\ H_3(x) &= x^2 - \frac{1}{2} \\ H_4(x) &= x^3 - \frac{3}{2}x \\ H_5(x) &= x^4 - 3x^2 + \frac{3}{4} \end{aligned}$$

Theorem 9.13 (Existence of an Orthonormal Basis). *Given a separable or complete euclidean space \mathcal{V} , there always exists an orthonormal basis*

Proof. Taken \mathcal{V} a separable euclidean space and $(v)_{n \in \mathbb{N}}$ a dense subset of \mathcal{V} .

Removing the linearly independent elements of this subset, we can call the new linearly independent set $(w)_{n \in \mathbb{N}}$. We have obviously

$$\begin{aligned} \text{span} \{(w)_{n=1}^\infty\} &= \text{span} \{(v)_{n=1}^\infty\} \\ \overline{\text{span} \{(w)_{n=1}^\infty\}} &= \overline{\text{span} \{(v)_{n=1}^\infty\}} = \mathcal{V} \end{aligned}$$

Orthonormalizing the system $(w)_{n \in \mathbb{N}}$ with the Gram-Schmidt procedure I obtain then a new set $(u)_{n \in \mathbb{N}}$ such that

$$\overline{\text{span} \{(u)_{n \in \mathbb{N}}\}} = \overline{\text{span} \{(w)_{n \in \mathbb{N}}\}} = \mathcal{V}$$

$(u)_{n \in \mathbb{N}}$ is complete and therefore a basis. □

Remark. If \mathcal{V} is a complete euclidean space but not separable, we can find thanks to Zorn's lemma a maximal orthonormal basis, but

1. There isn't a standard procedure for finding this basis
2. Taken the basis $(u)_{\alpha \in I}$ it can't be numerable. If it was then $\mathcal{V} = \overline{\text{span} \{(u)_\alpha\}}$. Taken $X = \text{span}_{\mathbb{Q}} \{(u)_k\}$ as the set of finite linear combinations with rational coefficients, we have that $\bar{X} = \text{span} \{(u)_k\}$ and therefore $\bar{X} = \mathcal{V}$ contradicting the fact that \mathcal{V} is not separable

10 Distributions

§ 10.1 Linear Functionals

Definition 10.1.1 (Dual Space II). Given $(\mathcal{V}, \|\cdot\|)$ a normed vector space over a field \mathbb{F} , we define the following set

$$\mathcal{V}^* := \{f : \mathcal{V} \longrightarrow \mathbb{F} \mid v \mapsto f(v)\}$$

An element $f \in \mathcal{V}^*$ is called a *continuous linear functional*, with the following properties, $\forall f, g \in \mathcal{V}^*$, $u, v \in \mathcal{V}$, $\lambda, \lambda_1, \lambda_2 \in \mathbb{F}$

$$\begin{aligned} f(\lambda_1 u + \lambda_2 v) &= \lambda_1 f(u) + \lambda_2 f(v) \\ (f + g)(u) &= f(u) + g(u) \\ (\lambda f)(u) &= f(\lambda u) = \lambda f(u) \end{aligned} \tag{10.1}$$

With these properties the set \mathcal{V}^* has a vector space structure, and it's called the *dual space*

Definition 10.1.2 (Bounded Linear Functional). Given $f \in \mathcal{V}^*$, we call f a *bounded linear functional* if, $\forall x \in \mathcal{V}$

$$\sup_{\|x\| \leq 1} |f(x)| < \infty \tag{10.2}$$

Definition 10.1.3 (Dual Norm). Given a normed vector space $(\mathcal{V}, \|\cdot\|)$ we can give a structure of normed vector space to its dual with the couple $(\mathcal{V}^*, \|\cdot\|_*)$, where the application $\|\cdot\|_*$ is called the *dual norm*

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

The dual norm is defined in two ways. $\forall f \in \mathcal{V}^*$, $v \in \mathcal{V}$

$$\|f\|_{*1} = \sup_{v \neq 0} \frac{|f(v)|}{\|v\|} \tag{10.3a}$$

$$\|f\|_{*2} = \sup_{\|v\| \leq 1} |f(v)| \tag{10.3b}$$

Where $f \in \mathcal{V}^*$ is a bounded linear functional.

Theorem 10.1. *For a bounded linear functional $f \in \mathcal{V}^*$ the two definitions of the dual norm coincide, i.e.*

$$\|f\|_{\star_1} = \|f\|_{\star_2} = \|f\|_{\star}$$

And we can define the following inequality

$$|f(v)| \leq \|f\|_{\star} \|v\| \quad \forall v \in \mathcal{V} \quad (10.4)$$

Proof. Since $f \in \mathcal{V}^*$ is bounded we have that both norms exist and must be finite. We can therefore write, thanks to the homogeneity of f , taken $v \in \mathcal{V}$

$$\|f\|_{\star_1} = \sup_{v \neq 0} \left| f \left(\frac{v}{\|v\|} \right) \right| = \sup_{\|v\|=1} |f(v)| \leq \sup_{\|v\| \leq 1} |f(v)| = \|f\|_{\star_2}$$

I.e. $\|f\|_{\star_1} \leq \|f\|_{\star_2}$, analogously

$$\|f\|_{\star_2} = \sup_{\|v\| \leq 1} |f(v)| \leq \sup_{0 < \|v\| \leq 1} \frac{|f(v)|}{\|v\|} \leq \sup_{v \neq 0} \frac{|f(v)|}{\|v\|} = \|f\|_{\star_1}$$

I.e. $\|f\|_{\star_1} \geq \|f\|_{\star_2}$, therefore we have

$$\|f\|_{\star_1} = \|f\|_{\star_2} = \|f\|_{\star}$$

The inequality is obvious taken the definition of supremum □

Theorem 10.2. *Given $f \in \mathcal{V}^*$ with \mathcal{V} normed vector space, we have that the following assumptions are equivalent*

1. f is continuous
2. f is continuous at the origin
3. f is bounded

Proof. 1) \implies 2)

Since $f \in \mathcal{V}^*$ is linear by definition, we have that it's also continuous and injective, therefore

$$\lim_{v \rightarrow 0} f(v) = f(0) = 0$$

2) \implies 3)

Since f is continuous at the origin, we have by definition of continuity and limit

$$\forall \epsilon > 0 \exists \delta > 0 : \|x\| < \delta \implies |f(x)| < \epsilon \quad x \in \mathcal{V}$$

Taken $u = \delta x \in \mathcal{V}$, we have $\|u\| = |\delta| \|x\|$ and

$$|f(x)| = \left| f \left(\frac{u}{\delta} \right) \right| = \frac{1}{|\delta|} |f(u)| \leq \epsilon$$

Therefore, if $\|x\| \leq 1$ we have that $|f(u)| \leq \delta \epsilon$, therefore

$$|f(u)| \leq \|f\|_{\star} \|u\| \implies \|f\|_{\star} \|x\| \leq \|f\|_{\star} \leq \epsilon$$

3) \implies 1) By definition of continuity and boundedness we have

$$\forall u, w \in \mathcal{V}, \forall \epsilon > 0 \exists \delta > 0 : \|u - w\| < \delta \implies |f(u) - f(w)| < \epsilon$$

Through the linearity of f we have

$$|f(u) - f(w)| = |f(u - w)| \leq \|f\|_* \|u - w\| \leq \delta \|f\|_*$$

Taken $\delta = \epsilon \|f\|_*^{-1}$, we have

$$|f(u) - f(w)| \leq \delta \|f\|_* = \epsilon$$

□

Corollary 10.1.1. Given $f \in \mathcal{V}^*$, we have that, given $C \in \mathbb{F}$ a constant

$$\|f\|_* = C \iff \begin{cases} \forall x \in \mathcal{V} & |f(x)| \leq C\|x\| \\ \forall \epsilon > 0 \exists x \in \mathcal{V} : & |f(x)| \geq (C - \epsilon)\|x\| \end{cases}$$

Proof. We have, by definition of $\|f\|_*$

$$\|f\|_* = \sup_{\|x\| \leq 1} |f(x)| \leq \sup_{\|x\| \leq 1} |c\|x\|| = C$$

In the second case, supposing $\|f\|_* < C$ and taken $\epsilon = 1/2(C - \|f\|_*)$

$$\exists x \in \mathcal{V} : |f(x)| \geq (C - \|f\|_*)\|x\| \quad \nexists$$

□

Definition 10.1.4 (Kernel). Given a linear functional $f \in \mathcal{V}^*$ we define the *kernel* of f as the set of zeros of f , i.e.

$$\ker f = \{v \in \mathcal{V} \mid f(v) = 0\} \quad (10.5)$$

Theorem 10.3 (Riesz Representation Theorem). *Given $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ a Hilbert space, we can uniquely define its dual \mathcal{V}^* through the isomorphism $\phi_v : \mathcal{V} \xrightarrow{\sim} \mathbb{F}$ defined as follows*

$$\forall u, v \in \mathcal{V} \quad \phi_v(u) = \langle u, v \rangle$$

I.e. $v \mapsto \langle \cdot, v \rangle$.

This isomorphism has the following properties

$$\begin{aligned} 1) & \forall v \in \mathcal{V} \quad \phi_v \in \mathcal{V}^*, \quad \|\phi_v\|_* = \|v\| \\ 2) & \forall c_1, c_2 \in \mathbb{F}, \quad v_1, v_2 \in \mathcal{V} \quad \phi_{c_1 v_1 + c_2 v_2} = \overline{c_1} \phi_{v_1} + \overline{c_2} \phi_{v_2} \\ 3) & f \in \mathcal{V}^* \implies \exists v \in \mathcal{V} : f = \phi_v \end{aligned} \quad (10.6)$$

Proof. Taken $w, v, z \in \mathcal{V}$ and $c, s \in \mathbb{F}$ we have by definition

$$\begin{aligned} \phi_v(w + z) &= \langle w + z, v \rangle = \langle w, v \rangle + \langle z, v \rangle = \phi_v(w) + \phi_v(z) \\ \phi_v(cw) &= \langle cw, v \rangle = c\langle w, v \rangle = c\phi_v(w) \\ \phi_{cw + sz}(v) &= \langle v, cw + sz \rangle = \overline{c}\langle v, w \rangle + \overline{s}\langle v, z \rangle = \overline{c}\phi_w(v) + \overline{s}\phi_z(v) \end{aligned}$$

We also have, thanks to Cauchy-Schwartz

$$\begin{aligned} |\phi_v(z)| &= |\langle z, v \rangle| \leq \|z\| \|v\| \\ |\phi_v(v)| &= |\langle v, v \rangle| \leq \|v\|^2 \quad \therefore \|\phi_v\|_* = \|v\| \end{aligned}$$

For the last one we take $f \in \mathcal{V}^*$. If $\ker f = \mathcal{V}$ we have by definition that

$$f(v) = 0 \quad \forall v \in \mathcal{V} \quad \therefore f = 0$$

Taken $v = 0$ we then have

$$\phi_0(v) = \langle 0, v \rangle = 0 = f(v) \quad \forall v \in \mathcal{V}$$

Supposing now $\ker f \neq \mathcal{V}$ there must be $w \in \ker f^\perp : w \neq 0$. Taken $u \in \mathcal{V}$ we can write

$$u = u_1 + \frac{f(u)}{f(w)}w = u - \frac{f(u)}{f(w)}w + \frac{f(u)}{f(w)}w$$

Therefore, through the linearity of f

$$f(u_1) = f\left(u - \frac{f(u)}{f(w)}w\right) = f(u) - f(w) = f(u - w) = 0 \quad \therefore u_1 \in \ker f$$

Since $w \in \ker f^\perp$, we have that

$$\langle u_1, w \rangle = 0$$

By definition of ϕ_v , then

$$\phi_v(u) = \langle v, u_1 \rangle + \langle v, \frac{f(u)}{f(w)}w \rangle = \langle \frac{f(w)}{\|w\|^2}w, u_1 \rangle + \langle \frac{f(w)}{\|w\|^2}w, \frac{f(u)}{f(w)}w \rangle$$

Then

$$\phi_v(u) = \frac{f(w)}{\|w\|^2} \langle w, u_1 \rangle + \frac{f(w)}{\|w\|^2} \frac{f(u)}{f(w)} \langle w, w \rangle = f(u)$$

Therefore

$$\phi_v(u) = f(u) \quad \forall f \in \mathcal{V}^*, \forall u \in \mathcal{V}$$

□

§ 10.2 Distributions

§§ 10.2.1 Local Integrability

Definition 10.2.1 (Weak Convergence). Given a normed vector space $(\mathcal{V}, \|\cdot\|)$, a sequence $(v)_{k \in \mathbb{N}} \in \mathcal{V}$ is said to be *weakly convergent* to $v \in \mathcal{V}$ and it's indicated as $v_k \rightharpoonup_w v$ if

$$\forall f \in \mathcal{V}^* \quad \lim_{k \rightarrow \infty} f(v_k) = f(v)$$

Theorem 10.4. Given a normed vector space \mathcal{V} and a sequence $(v)_{k \in \mathbb{N}} \in \mathcal{V}$ such that $v_k \rightarrow v \in \mathcal{V}$ (i.e. converges strongly), we have that

$$v_k \rightarrow v \implies v_k \rightharpoonup_w v$$

Proof. Given $v_k \rightarrow v$, we have by definition that

$$v_k \rightarrow v \iff \lim_{k \rightarrow \infty} \|v_k - v\| = 0$$

Now, writing the definition of weak convergence and applying the linearity of the limit

$$v_k \rightharpoonup_w v \iff \forall f \in \mathcal{V}^* \quad \lim_{k \rightarrow \infty} (f(v_k) - f(v)) = 0$$

Therefore

$$\forall f \in \mathcal{V}^* \quad \lim_{k \rightarrow \infty} (f(v_k - v)) \leq \|f\|_* \lim_{k \rightarrow \infty} \|v_k - v\|$$

Therefore we have that, since f is bounded

$$\exists C \in \mathbb{R} : \lim_{k \rightarrow \infty} (f(v_k - v)) \leq C \lim_{k \rightarrow \infty} \|v_k - v\| \quad \forall f \in \mathcal{V}^*$$

And therefore

$$v_k \rightarrow v \implies v_k \rightharpoonup_w v$$

□

Remark. The opposite isn't necessarily true.

Proof. Take $(\mathcal{V}, \|\cdot\|) = (\ell^2, \|\cdot\|_2)$ and take the standard sequence $(e_i)_j$ where

$$(e_i)_j = (0, 0, \dots, 0, \underbrace{(1)_j}_{i\text{-th}}, 0, \dots)$$

And $(1)_j$ is the identity sequence.

We have that $(e_i)_j \not\rightharpoonup (a)_i$ since

$$\|(e_i)_j - (e_i)_k\|_2 = \sqrt{2} \quad \forall j \neq k$$

Now, considering that $(\ell^2, \|\cdot\|_2)$ is a Hilbert space, we have that, due to Riesz representation theorem

$$\forall f \in \ell^{2*} \quad f = \langle \cdot, a \rangle = \phi_a \quad a \in \ell^2$$

Therefore

$$f(e_{ij}) = \langle e_{ij}, a_i \rangle = \sum_{i=1}^{\infty} a_i e_{ij} = a_j \rightarrow 0$$

Therefore

$$\forall (a)_j \in \ell^2 \quad \forall f \in \ell^{2*} \quad \lim_{j \rightarrow \infty} f(e_{ij}) = 0$$

Hence, by definition $e_{ij} \rightharpoonup_w 0$

□

Definition 10.2.2 (Isolated Singularities II). Given a function $f : \mathbb{R} \rightarrow \mathbb{C}$, it's said to have *isolated singularities* if

$$S_f := \left\{ x \in \mathbb{R} \mid \lim_{y \rightarrow x} f(y) \neq f(x) \right\}$$

Doesn't have accumulation points, i.e. if $|S_f| < \infty$

Definition 10.2.3 (Piecewise Continuity). A function $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{C}$ is said to be *piecewise continuous* if

$$1. S_f \neq \{\} \text{ and } |S_f| < \infty$$

$$2. \forall u \in S_f$$

$$\begin{aligned} \exists \lim_{x \rightarrow u^+} f(x) &= f(u^+) \\ \exists \lim_{x \rightarrow u^-} f(x) &= f(u^-) \end{aligned} \tag{10.7}$$

$$3.$$

$$\begin{aligned} \exists \lim_{x \rightarrow a^+} f(x) &= f(a^+) \\ \exists \lim_{x \rightarrow b^-} f(x) &= f(b^-) \end{aligned} \tag{10.8}$$

Definition 10.2.4 (Jump). Given a piecewise continuous function f we define the *jump* of the function at a discontinuity $x \in S_f$ as follows

$$\Delta f(x) = f(x^+) - f(x^-)$$

Definition 10.2.5 (Piecewise Differentiability). A piecewise continuous function $f : [a, b] \rightarrow \mathbb{C}$ is said to be *piecewise differentiable* if and only if

$$1. f \text{ is piecewise continuous in } [a, b]$$

$$2. f' \text{ is piecewise continuous in } [a, b]$$

Definition 10.2.6 (Local Integrability). A function $f : \mathbb{R} \rightarrow \mathbb{C}$ is said to be *locally integrable* if

$$1. S_f \neq \{\} \text{ and } |S_f| < \infty$$

$$2. \forall a, b \in \mathbb{R}, a < b \text{ we have that}$$

$$\int_a^b |f(x)| dx < \infty$$

The set of locally integrable functions forms a subspace of $L^1(\mathbb{R})$ and it's indicated as $L^1_{loc}(\mathbb{R})$

Theorem 10.5. Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a piecewise continuous function, then $f \in L^1_{loc}(\mathbb{R})$

Proof. Take $f(x)$ a piecewise continuous function, then $\forall a, b \in \mathbb{R} \exists (u)_{0 \leq k \leq n} \in \mathbb{R} : u_0 = a, u_n = b$ for which we have $f \in C((u_{k-1}, u_k)) \forall 0 \leq k \leq n$. I can therefore define $\tilde{f}_k : [u_{k-1}, u_k] \rightarrow \mathbb{C}$, where

$$\tilde{f}_k(x) = \begin{cases} f(x) & x \in (u_{k-1}, u_k) \\ f(u_{k-1}^+) & x = u_{k-1} \\ f(u_k^-) & x = u_k \end{cases}$$

We have that

$$f(x) = \sum_{k=0}^{\infty} \tilde{f}_k(x)$$

And therefore

$$\int_a^b |f(x)| dx = \sum_{k=0}^n \int_{u_{k-1}}^{u_k} |f(x)| dx = \sum_{k=0}^n \int_{u_{k-1}}^{u_k} |\tilde{f}_k(x)| dx$$

Therefore $f(x) \in L_{loc}^1(\mathbb{R})$ and the theorem is proved \square

Theorem 10.6 (Integration by Parts in $L_{loc}^1(\mathbb{R})$). *Let $f, g \in L_{loc}^1([a, b])$, then indicating the evaluation of a function at two points as $[f(x)]_z^w$ we have*

$$\begin{aligned} [fg]_{a+}^{b-} &= \int_a^b (f'(x)g(x) + f(x)g'(x)) dx + \\ &+ \sum_{x \in S_f \cup S_g} (f(x^-)\Delta g(x) + g(x^-)\Delta f(x) + \Delta f(x)\Delta g(x)) \end{aligned} \quad (10.9)$$

Proof. Take $H : [a, b] \rightarrow \mathbb{C}$, then for the fundamental theorem of calculus

$$H(b^-) - H(a^+) = \int_a^b \frac{dH}{dx} dx + \sum_{x \in S_H} \Delta H(x)$$

Taken $H(x) = f(x)g(x)$ we have

$$\Delta(fg)(x) = f(x^+)g(x^+) - f(x^-)g(x^-)$$

With some manipulation we have

$$\begin{aligned} \Delta(fg)(x) &= f(x^+)g(x^+) - f(x^+)g(x^-) + f(x^+)g(x^-) - f(x^-)g(x^-) = f(x^+)\Delta g(x) + g(x^-)\Delta f(x) = \\ &= (f(x^-) + \Delta f(x))\Delta g(x) + g(x^-)\Delta f(x) = f(x^-)\Delta g(x) + g(x^-)\Delta f(x) + \Delta f(x)\Delta g(x) \end{aligned}$$

Therefore

$$(fg)(b^-) - (fg)(a^+) = \int_a^b (fg)'(x) dx + \sum_{x \in S_f \cup S_g} (f(x^-)\Delta g(x) + g(x^-)\Delta f(x) + \Delta f(x)\Delta g(x))$$

\square

Definition 10.2.7 (Test Functions). A function $f : \mathbb{R} \rightarrow \mathbb{C}$ is said to be a *test function* if $\overline{\text{supp}\{f\}}$ is compact and $f \in C^\infty(\mathbb{R})$, i.e. $f \in C_c^\infty(\mathbb{R})$

This space is usually denoted as follows $C_c^\infty(\mathbb{R}) = \mathcal{K}$

Definition 10.2.8 (\mathcal{K} -convergence). Given a sequence $(f)_{n \in \mathbb{N}} \in \mathcal{K}$, it's said to be \mathcal{K} -convergent if

$$1. \exists I \subset \mathbb{R} : \forall x \in I^c \ f_n(x) = 0$$

$$2. \forall k \in \mathbb{N} \ f_n^{(k)}(x) \Rightarrow f^{(k)}(x)$$

Then, it's indicated as $f_n \xrightarrow{\mathcal{K}} f$

§§ 10.2.2 Regular and Singular Distributions

Definition 10.2.9 (Distribution). A *distribution* is a continuous linear functional $\varphi : \mathcal{K} \rightarrow \mathbb{C}$, i.e.

$$\forall (f)_n \in \mathcal{K}, f_n \xrightarrow{\mathcal{K}} f \implies \varphi(f_n) = \varphi(f)$$

By definition of dual space, we have that $\varphi \in \mathcal{K}^*$

Theorem 10.7. Given $g \in L^1_{loc}(\mathbb{R})$ and $\varphi_g : \mathcal{K} \rightarrow \mathbb{C}$, if we have

$$\varphi_g(f) = \int_{\mathbb{R}} g(x)f(x) dx \quad \forall f \in \mathcal{K}$$

Then $\varphi_g \in \mathcal{K}^*$

Proof. Using the fact that $f \in \mathcal{K}$ we can immediately say that $|f(x)| \leq M \quad \forall x \in [-a, a] \subset \mathbb{R}$, therefore

$$\int_{\mathbb{R}} |g(x)f(x)| dx = \int_{-a}^a |g(x)f(x)| dx \leq M \int_{-a}^a |g(x)| dx < \infty$$

Alternatively, using the definition of integral we can say that

$$\varphi_g(\alpha f + \beta h) = \alpha \varphi_g(f) + \beta \varphi_g(h) \quad \forall \alpha, \beta \in \mathbb{C}, \forall f, h \in \mathcal{K}$$

We only need to show that this application is \mathcal{K} -continuous. I take $f_n \xrightarrow{\mathcal{K}} f$, with $f_n \in \mathcal{K}$.

It's obvious that $f_n \rightrightarrows f$, and using the linearity and that $g \in L^1_{loc}(\mathbb{R})$, calling $A = \|g\|$, we have

$$|\varphi_g(f) - \varphi_g(f_n)| \leq A \|f - f_n\|_u \rightarrow 0$$

Therefore $\varphi_g(f_n) \rightarrow \varphi_g(f) \quad \forall f_n \in \mathcal{K}, f_n \xrightarrow{\mathcal{K}} f$ □

Definition 10.2.10 (Regular Distribution). A distribution $f \in \mathcal{K}^*$ is said to be *regular* if $\exists g \in L^1_{loc}(\mathbb{R})$ such that

$$f = \langle \cdot, g \rangle = \varphi_g$$

I.e.

$$f(h) = \varphi_g(h) = \langle h, g \rangle = \int_{\mathbb{R}} h(x)g(x) dx$$

Definition 10.2.11 (Ceiling and Floor Functions). We define the *ceiling function* $\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$ and the *floor functions* $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ as follows

$$\begin{aligned} \lfloor x \rfloor &= \max \{ m \in \mathbb{Z} \mid m \leq x \} \\ \lceil x \rceil &= \min \{ n \in \mathbb{Z} \mid n \geq x \} \end{aligned} \tag{10.10}$$

Definition 10.2.12 (Heaviside Function). We define the *Heaviside function* as a function $H : \mathbb{R} \rightarrow \mathbb{R}$ as follows

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

It's obviously piecewise continuous and therefore $H \in L^1_{loc}(\mathbb{R})$

A secondary definition is the one that follows $\tilde{H} : \mathbb{R} \rightarrow \mathbb{R}$

$$\tilde{H}(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Example 10.2.1 (Floor and Ceiling Distributions). Since both the floor and ceiling distributions are locally integrable, we can build a regular distribution $\varphi \in \mathcal{K}^*$ as follows

$$\begin{aligned}\varphi_{[x]}(f) &= \int_{\mathbb{R}} f(x)[x] \, dx = \sum_{k \in \mathbb{Z}} \int_k^{k+1} k f(x) \, dx \\ \varphi_{[x]}(f) &= \int_{\mathbb{R}} f(x)[x] \, dx = \sum_{k \in \mathbb{Z}} \int_k^{k+1} (k+1) f(x) \, dx\end{aligned}$$

Example 10.2.2 (Theta Distribution). Given $H \in L^1_{loc}(\mathbb{R})$ we can define the associated *theta distribution* $\varphi_H = \vartheta \in \mathcal{K}$, as follows

$$\vartheta(f) = \int_{\mathbb{R}} f(x)H(x) \, dx = \int_{\mathbb{R}^+} f(x) \, dx$$

It's already obvious that $\varphi_{\tilde{H}} = \varphi_H$, but it's better to formalize it in the following theorem

Theorem 10.8. *Let $f, g \in L^1_{loc}(\mathbb{R})$ such that $\{x \in \mathbb{R} \mid f(x) \neq g(x)\} = \{u_1, \dots, u_n\}$. Then we have $\varphi_g = \varphi_f \in \mathcal{K}^*$*

Proof. We have by Riesz theorem that

$$\varphi_g(f) = \int_{\mathbb{R}} g(x)f(x) \, dx = \int_{-\infty}^{u_1} f(x)g(x) \, dx + \sum_{k=2}^{n-1} \int_{u_k}^{u_{k+1}} f(x)g(x) \, dx + \int_{u_n}^{\infty} f(x)g(x) \, dx$$

But, since $f(x) = g(x)$, $\forall x \in (u_k, u_{k+1})$ we have

$$\begin{aligned}\int_{\mathbb{R}} f(x)g(x) \, dx &= \int_{-\infty}^{u_1} f(x)g(x) \, dx + \sum_{k=2}^{n-1} \int_{u_k}^{u_{k+1}} f(x)g(x) \, dx + \int_{u_n}^{\infty} f(x)g(x) \, dx = \\ &= \int_{-\infty}^{u_1} f(x)h(x) \, dx + \sum_{k=2}^{n-1} \int_{u_k}^{u_{k+1}} f(x)h(x) \, dx + \int_{u_n}^{\infty} f(x)h(x) \, dx = \int_{\mathbb{R}} f(x)h(x) \, dx\end{aligned}$$

Therefore

$$\varphi_g = \varphi_h$$

□

Definition 10.2.13 (Singular Distribution). A *singular distribution* is a distribution $f \in \mathcal{K}^*$ for which, given $g \in L^1_{loc}(\mathbb{R})$, $f \neq \varphi_g$, where

$$\varphi_g(h) = \int_{\mathbb{R}} h(x)g(x) \, dx$$

Definition 10.2.14 (Dirac Delta Distribution). An example of singular distribution is the *Dirac delta distribution* $\delta_a \in \mathcal{K}^*$. This distribution is defined as follows

$$\delta_a(f) = f(a) \quad \forall f \in \mathcal{K}, \quad a \in \mathbb{R}$$

Theorem 10.9. *Given the Dirac delta distribution δ_a , $\exists \delta(x) \in L^1_{loc}(\mathbb{R})$ such that (taken $a = 0$ without loss of generality)*

$$\delta_0(f) = \int_{\mathbb{R}} f(x) \delta(x) dx$$

Proof. Let's say that $\exists \delta(x) \in L^1_{loc}(\mathbb{R})$, therefore, we could define $\delta_0 = \langle \cdot, \delta(x) \rangle$. This function therefore must have these properties

$$\delta(x) \neq 0 \iff x = 0, x \in S_\delta$$

But, since for $b \in \mathbb{R}$, $b \neq 0$, we have that $\delta(x) \rightarrow \delta(b)$ continuously, i.e. $\delta(b) = A > 0$, therefore

$$\exists \epsilon > 0 : \delta(x) \geq \frac{A}{2} \quad \forall x \in [b - \epsilon, b + \epsilon]$$

But, then $\exists f \in \mathcal{K}$ such that $f(x) \rightarrow f(b)$ continuously, $f > 0$, $f(b) = 1$. Taken $\epsilon < b$ we can say

$$\exists \epsilon' \in (0, \epsilon) : f(x) \geq \frac{1}{2} \quad \forall x \in [b - \epsilon', b + \epsilon']$$

Therefore, we have that

$$\int_{\mathbb{R}} f(x) \delta(x) dx = \int_{-\epsilon}^{\epsilon} f(x) \delta(x) dx \geq \int_{-\epsilon}^{\epsilon} \frac{A}{2} \frac{1}{2} dx = \frac{\epsilon' A}{2}$$

But, by definition $f(0) = 0$, therefore

$$\int_{\mathbb{R}} f(x) \delta(x) dx = 0 \implies \varphi_\delta(f) = \varphi_0(f) \quad \nexists$$

Therefore, the distribution $\delta_0 \neq \langle \cdot, \delta(x) \rangle$ □

Notation (Common Abuse of Notation). A common abuse of notation in the usage of the Dirac delta distribution, is supposing that it's writable like a regular distribution, i.e. supposing that $\exists \delta(x) \in L^1_{loc}(\mathbb{R})$, the “Dirac delta function” such that

$$\delta_a(f) = \int_{\mathbb{R}} f(x) \delta(x - a) dx = f(a) \quad \forall f \in \mathcal{K}$$

This is obtainable only if the “delta function” is defined as follows

$$\delta(x) = \begin{cases} +\infty & x = 0 \\ 0 & x \neq 0 \end{cases} \quad (10.11)$$

Note that

$$\delta_0(1) = \int_{\mathbb{R}} \delta(x) dx = 1$$

Theorem 10.10. *Given $g_n \in L^1_{loc}(\mathbb{R})$ such that $g_n \in C[-a_n, a_n]$ with $a_n \rightarrow 0$, $a_n > 0 \quad \forall n \in \mathbb{N}$, we have that, if*

$$\int_{\mathbb{R}} g_n(x) dx = 1$$

Then

$$\lim_{n \rightarrow \infty} \varphi_{g_n}(f) = \delta_0(f)$$

Proof. We use the definition of limit, therefore we have

$$\left| \int_{\mathbb{R}} f(x)g_n(x) dx - \delta_0(f) \right| \leq \int_{-a_n}^{a_n} |f(x) - f(0)|g_n(x) dx$$

Using that the integral over the real axis of g_n is unitary

$$\left| \int_{\mathbb{R}} f(x)g_n(x) dx - \delta_0(f) \right| \leq \sup_{|x| \leq a_n} |f(x) - f(0)| \rightarrow 0$$

Therefore

$$\lim_{n \rightarrow \infty} \varphi_{g_n}(f) = \delta_0(f)$$

□

Corollary 10.2.1. Given $g \in C_1(\mathbb{R})$ a non-negative function $g \geq 0$ such that

$$\int_{\mathbb{R}} g(x) dx = 1$$

Then, if we put $g_n(x) = ng(nx)$ we have that

$$\lim_{n \rightarrow \infty} \varphi_{g_n}(f) = \delta_0(f)$$

Proof. As before, we use the definition, and therefore we have

$$\left| \int_{\mathbb{R}} f(x)g_n(x) dx - \delta_0(f) \right| \leq \int_{\mathbb{R}} |f(x) - f(0)|g_n(x) dx = n \int_{\mathbb{R}} |f(x) - f(0)|g(nx) dx$$

Using the substitution $u = nx$ we therefore get

$$\int_{\mathbb{R}} |f(x) - f(0)|g(nx)n dx = \int_{\mathbb{R}} \left| f\left(\frac{u}{n}\right) - f(0) \right| g(u) du$$

But, by definition of $g(x)$, we have that $\exists L > 0 : \exists \epsilon > 0$ for which

$$1 - \epsilon \leq \int_{-L}^L g(u) du \leq 1 \quad \wedge \quad \int_{|u| \geq L} g(u) du \leq \epsilon$$

And therefore

$$\int_{\mathbb{R}} \left| f\left(\frac{u}{n}\right) - f(0) \right| g(u) du = \int_{-L}^L \left| f\left(\frac{u}{n}\right) - f(0) \right| g(u) du + \int_{|u| \geq L} \left| f\left(\frac{u}{n}\right) - f(0) \right| g(u) du$$

By using the properties of the integral, we know that

$$\int_{\mathbb{R}} \left| f\left(\frac{u}{n}\right) - f(0) \right| g(u) du \leq \sup_{x \leq \frac{L}{n}} |f(x) - f(0)| + \epsilon \left(\sup_{u \in \mathbb{R}} \left| f\left(\frac{u}{n}\right) \right| - |f(0)| \right)$$

Since $f(x) \rightarrow f(0)$ continuously, due to the fact that $f \in \mathcal{K}$, we have that

$$\int_{\mathbb{R}} |f(x) - f(0)|g(nx)n dx \leq 2\epsilon \|f\|_u$$

Therefore

$$\lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}} f(x) g_n(x) dx - \delta_0(f) \right| = 0$$

□

Definition 10.2.15 (PV (x^{-n}) –Distribution). Taken $n > 0$, $n \in \mathbb{N}$, we have that $x^{-n} \notin L^1_{loc}(x)$ and therefore there is no associated distribution $\varphi_{x^{-n}} \in \mathcal{K}^*$. A useful thing we could do is utilizing the definition of the Cauchy principal value of the function. Therefore, we define the following singular distribution

$$\text{PV} \left(\frac{1}{x^n} \right) [f] = \lim_{a \rightarrow \infty} \int_{-a}^a \frac{1}{x^n} \left(f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k \right) dx \quad (10.12)$$

Theorem 10.11. The application $\text{PV}(x^{-n}) : \mathcal{K} \rightarrow \mathbb{R}$ defined as before is a distribution, hence $\text{PV}(x^{-n}) \in \mathcal{K}^*$

Proof. We already know that this distribution is linear, since $\forall f, g \in \mathcal{K}, \forall c, d \in \mathbb{R}$

$$\begin{aligned} \text{PV} \left(\frac{1}{x^n} \right) [cf + dg] &= \lim_{a \rightarrow \infty} \int_{-a}^a \frac{1}{x^n} \left(cf(x) + dg(x) + \sum_{k=0}^{n-1} \frac{1}{k!} (cf^{(k)}(0) + dg^{(k)}(0)) x^k \right) dx = \\ &= c \text{PV} \left(\frac{1}{x^n} \right) [f] + d \text{PV} \left(\frac{1}{x^n} \right) [g] \end{aligned}$$

Secondly we must show that the integral is well defined.

Let $f \in \mathcal{K}$, then

$$I_R^n(f) = \int_{-R}^R \frac{1}{x^n} \left(f(x) - \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k)}(0) x^k \right) dx, \quad n \in \mathbb{N}, \quad R > 0$$

Using the Lagrange formulation for the remainder of the McLaurin expansion, we have

$$R_{n-1}(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k = x^n r_{n-1}(x)$$

And therefore we have

$$I_R^n(x) = \int_{-R}^R r_{n-1}(x) dx$$

Since $r_{n-1}(x) \in C^\infty(\mathbb{R})$ (it's a polynomial) and $\text{supp } f \subset [-b, b]$ for some $b \in \mathbb{R}$, $b > 0$, we have that for $R > b$

$$I_R^n(f) = I_b^n(f) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} \int_{b \leq |x| \leq R} \frac{1}{x^{n-k}} dx$$

The last integral evaluates to

$$\int_{b \leq |x| \leq R} \frac{1}{x^{n-k}} dx = \begin{cases} \frac{2}{n-k-1} \left[\frac{1}{b^{n-k-1}} - \frac{1}{R^{n-k-1}} \right] & n-k \bmod 2 = 0, \quad n-k > 2 \\ 0 & \text{else} \end{cases}$$

Therefore, introducing the following dummy index $j = n - k$ where $j \bmod 2 = 0$, we have

$$\text{PV} \left(\frac{1}{x^n} \right) (f) = I_b^n(f) - \sum_{j=0}^{n-1} \frac{f^{(j)}(0)}{j!} \frac{2}{(j-1)b^{j-1}}$$

Therefore the integral defining the distribution is well defined.

Lastly we need to prove that $\text{PV}(x^{-n})(f_n) \xrightarrow{\mathcal{K}} \text{PV}(x^{-n})(f)$, $\forall (f)_n \in \mathcal{K} : f_n \rightarrow f$, i.e. that it's a continuous functional.

Going back to the previous definition we have that

$$r_{n-1}(x) = \frac{R_{n-1}(x)}{x^n} = \frac{1}{x^n} \frac{d^{n-1}f}{dx^{n-1}}(\xi(x)) \quad \xi(x) \in (0, x)$$

Therefore

$$|I_R^n(f)| \leq 2 \sup_{x \in [-R, R]} |r_{n-1}(x)| \leq \frac{2R}{n!} \|f\|_u$$

Which means

$$\left| \text{PV} \left(\frac{1}{x^n} \right) (f) \right| \leq \frac{2b}{n!} \|f^{(n)}\|_u + \sum_{j=0}^{n-1} \frac{|f^{(j)}(0)|}{j!} \frac{2}{(j-1)b^{j-1}}$$

Taken $b_1 = \max\{b, 1\}$ we have that $\text{supp } f \subset [-b_1, b_1]$ and whenever $j-1 > 1$, $b_1 \geq 1$

$$\frac{2}{(j-1)b_1^{j-1}} \leq 2$$

Therefore, finally

$$\left| \text{PV} \left(\frac{1}{x^n} \right) (f) \right| \leq \frac{2}{n!} \max\{b, 1\} \|f^{(n)}\|_u + 2 \sum_{j=0}^{n-1} \frac{|f^{(j)}(0)|}{j!}$$

Now, defined $f_j \in \mathcal{K}$ and $h_j(x) = f(x) - f_j(x)$ with $f_j \rightarrow f$ we have that $\text{supp } h_j \subset [-b, b]$, $\forall j \in \mathbb{N}$.

By definition we have that $h_j \rightrightarrows 0$ and $h_j^{(n)} \rightrightarrows 0 \forall n \in \mathbb{N}$

$$\text{PV} \left(\frac{1}{x^n} \right) (h_j) \leq \frac{2}{n!} \max\{b, 1\} \|h_j^{(n)}\|_u + 2 \sum_{j=0}^{n-1} \frac{|h_j^{(j)}(0)|}{j!}$$

This means

$$\forall k \in \mathbb{N}, \forall \epsilon > 0, \exists N_R \in \mathbb{N} : \forall k \geq N_j \left| h_j^{(k)}(x) \right| \leq \frac{\epsilon}{2(b+1+\epsilon)}$$

Chosen $N = \max_{0 \leq k \leq n-1} N_k$ we have that $\forall j \geq N$

$$\left| \text{PV} \left(\frac{1}{x^n} \right) (h_j) \right| \leq \frac{2\epsilon}{2(b+1+\epsilon)} \left[\max\{b, 1\} + \sum_{j=0}^{n-1} \frac{1}{j!} \right] \leq \epsilon \rightarrow 0$$

Therefore, $\text{PV}(x^{-n}) \in \mathcal{K}^*$

□

§§ 10.2.3 Operations with Distributions

Definition 10.2.16 (Weak Derivative). Given $u \in L^1([a, b])$, we define the weak derivative $v \in L^1([a, b])$ if, $\forall h \in C_c^\infty([a, b])$ we have

$$\int_a^b u(x)h'(x) \, dx = - \int_a^b v(x)h(x) \, dx$$

The function $v(x)$ will then be identified as follows

$$v(x) = D u(x)$$

Theorem 10.12 (Operations with Distributions). Given $\varphi, \gamma \in \mathcal{K}^*$, $f \in \mathcal{K}$, $h \in C^\infty(\mathbb{R})$ and $c \in \mathbb{C}$ we define the following operations in \mathcal{K}^*

$$\begin{aligned} + : \mathcal{K}^* \times \mathcal{K}^* &\longrightarrow \mathcal{K}^* \\ \cdot : \mathbb{C} \times \mathcal{K}^* &\longrightarrow \mathcal{K}^* \\ \circ : C^\infty \times \mathcal{K}^* &\longrightarrow \mathcal{K}^* \\ D : \mathcal{K}^* &\longrightarrow \mathcal{K}^* \end{aligned} \tag{10.13}$$

Where, they act as follows

$$\begin{aligned} +(\varphi(f), \gamma(f)) &= (\varphi + \gamma)(f) = \varphi(f) + \gamma(f) \\ \cdot(c, \varphi(f)) &= (c\varphi)(f) = c\varphi(f) \\ \circ(h, \varphi(f)) &= (h\varphi)(f) = \varphi(hf) \\ D \varphi(f) &= -\varphi(f') \end{aligned}$$

The last operation is the distributional derivative.

Theorem 10.13. Given $g(x) \in L_{loc}^1(\mathbb{R})$ and $\varphi_g \in \mathcal{K}^*$ its associated distribution. Then, $\forall h \in C^\infty$, $f \in \mathcal{K}$

$$(h\varphi_g)(f) = \varphi_g(hf) = \varphi_{hg}(f) \tag{10.14}$$

Proof. The proof is quite straightforward

$$(h\varphi_g)(f) = \varphi_g(hf) = \int_{\mathbb{R}} g(x) (h(x)f(x)) \, dx = \int_{\mathbb{R}} f(x) (g(x)h(x)) \, dx = \varphi_{hg}(f)$$

□

Theorem 10.14. Taken $g \in L_{loc}^1(\mathbb{R})$ and $g \mapsto \varphi_g \in \mathcal{K}^*$ its associated distribution, we have that

$$D \varphi_g = \varphi_{Dg}$$

Where Dg is the weak derivative of g

Proof. $\forall f \in \mathcal{K}$ we have that

$$D \varphi_g(f) = -\varphi_g(f') = - \int_{\mathbb{R}} f'(x)g(x) \, dx = \int_{\mathbb{R}} f(x) Dg(x) \, dx = \varphi_{Dg}(f)$$

□

Theorem 10.15. $\forall g, f \in L^1_{loc}(\mathbb{R})$ and given $\varphi_g, \gamma_f \in \mathcal{K}^*$, $c \in \mathbb{C}$ their associated distributions, we have that

$$\begin{aligned}\varphi_g + \varphi_f &= \varphi_{g+f} \\ \varphi_{cf} &= c\varphi_f\end{aligned}$$

Proof. The proof is obvious using the linearity of the integral operator □

Theorem 10.16 (Useful Identities). *Here is a list of some useful identities*

$$\begin{aligned}x\delta_0 &= 0 \\ x\text{PV}\left(\frac{1}{x}\right) &= \varphi_1 \\ xD\delta_0 &= -\delta_0\end{aligned}\tag{10.15}$$

Proof. 1) $x\delta_0 = 0$.

Taken $f \in \mathcal{K}$

$$x\delta_0(f) = \delta_0(xf) = 0f(0) = 0$$

2) $x\text{PV}\left(\frac{1}{x^n}\right) = \varphi_1$

Again, taken $g \in \mathcal{K}$

$$x\text{PV}\left(\frac{1}{x}\right)(g) = \text{PV}\left(\frac{1}{x}\right)(xg) = \lim_{\epsilon \rightarrow 0} \int_{|x| \geq \epsilon} f(x) dx = \varphi_1$$

For the last one, taken $h \in \mathcal{K}$

$$xD\delta_0(f) = D\delta_0(xf) = -\delta_0(f + xf') = -f(0) - 0f'(0) = -f(0) = -\delta_0(f)$$

□

Notation (Abuse of Notation). Given $g \in L^1_{loc}(\mathbb{R})$ and its associated distribution $g \mapsto \varphi_g$, is quite common to use the original function to indicate actually the distribution. Together with this, the distributional derivative D gets indicated as an usual derivative, therefore

$$\begin{aligned}\varphi_g &\rightarrow g(x) \\ D\varphi_g &\rightarrow g'(x) = \frac{dg}{dx}\end{aligned}$$

Therefore it isn't uncommon to see identities like this

$$x\text{PV}\left(\frac{1}{x}\right) = 1$$

Where actually $1 \rightarrow \varphi_1$ is the identity distribution.

Or

$$\delta'_0 = -\delta_0$$

This makes an easy notation for calculating distributional derivatives and have some calculations, but one should watch out to this common abuse of notation

Definition 10.2.17 (\mathcal{K}^* Convergence). Given $(\sigma)_{n \in \mathbb{N}} \in \mathcal{K}^*$ a sequence of distributions, it's say to \mathcal{K}^* -converge if

$$(\sigma)_n(f) \rightarrow \sigma(f) \in \mathcal{K}^*$$

It's indicated as follows

$$(\sigma)_n \rightarrow_* \sigma$$

Theorem 10.17. Given $\kappa \in \mathcal{K}^*$ a distribution, we have that $f \in C^\infty(\mathcal{K})$ in the sense of distributional derivatives, where the n -th derivative is defined as follows

$$D^n \kappa(f) = (-1)^n \kappa(f^{(n)}) \quad \forall f \in \mathcal{K}$$

Proof. Taken $f \in \mathcal{K}$ we have that $f \in C_c^\infty(\mathbb{R}) \subset C^\infty(\mathbb{R})$ hence f is smooth, therefore we have, for $\kappa \in \mathcal{K}^*$

$$D \kappa(f) = -\kappa(f')$$

Iterating, we get for $n = 2$

$$D^2 \kappa(f) = -D \kappa(f') = \kappa(f'')$$

Iterating till n we get finally

$$D^n \kappa(f) = (-1)^n \kappa(f^{(n)})$$

Hence the derivability depends only on the test function f , and since it's smooth, the distributional derivatives can be defined $\forall n \in \mathbb{N}$ □

Example 10.2.3. Take for example the delta distribution $\delta_0 \in \mathcal{K}^*$, given $f \in \mathcal{K}$ we want to calculate the following value

$$D^n \delta_0(f)$$

Using the previous identity, we have that

$$D^n \delta_0(f) = (-1)^n \delta_0[f^{(n)}] = (-1)^n f^{(n)}(0)$$

Theorem 10.18 (Chain Rule). Given $h \in C^\infty(\mathbb{R})$ and $\eta \in \mathcal{K}^*$, we have that

$$D(h\eta) = h'\eta + h D \eta$$

Where the distributional derivative is identical to the usual derivative for h and it's the usual distributional derivative for η

Proof. Taken $h \in C^\infty$, $\eta \in \mathcal{K}^*$ and $f \in \mathcal{K}$ we have

$$(h\eta)(f) = \eta(fh)$$

Therefore, using the identity for derivating a distribution we have that

$$\begin{aligned} D(h\eta)[f] &= -(h\eta)[f'] = \\ &= -\eta[hf'] = -\eta[(hf)' - h'f] = -\eta[(hf)'] + \eta[h'f] = \\ &= (h D \eta)[f] + (h'\eta)[f] \end{aligned}$$

□

Theorem 10.19. *Given $g_n(x) : \mathbb{R} \rightarrow \mathbb{C}$ a continuous sequence of functions such that $g_n \rightrightarrows g \in C(\mathbb{R})$. Then, taken $g_n \mapsto \varphi_{g_n} \in \mathcal{K}^*$ the associated distribution, we have that*

$$\varphi_{g_n} \rightarrow_\star \varphi_g \in \mathcal{K}^*$$

Proof. By definition of convergence, we have that, given $f \in \mathcal{K}$, where $\text{supp } f \subset [-a, a]$

$$|\varphi_{g_n}(f) - \varphi_g(f)| \leq \varphi_{|g_n - g|}(|f|) \leq 2a \|g_n - g\|_u \|f\|_u \rightarrow 0$$

Therefore, $\varphi_{g_n} \rightarrow_\star \varphi_g$ □

Theorem 10.20. *Given $(\eta)_n \in \mathcal{K}^*$ a sequence of distribution, then if $(\eta)_n \rightarrow_\star \eta$ we have that*

$$D \eta_n \rightarrow_\star D \eta$$

Proof. By definition, we have

$$\lim_{n \rightarrow \infty} D \eta_n(f) = - \lim_{n \rightarrow \infty} \eta_n(f') = -\eta(f') = D \eta(f)$$

□

Example 10.2.4 (The Absolute Value Distribution). Taken $g(x) = |x|$ we have that (obviously) $|x| \in L^1_{loc}(\mathbb{R})$, therefore there exists a distribution $\varphi_{|x|} \in \mathcal{K}^*$ defined as follows

$$\varphi_{|x|}(f) = \int_{\mathbb{R}} |x| f(x) dx$$

We have that the distributional derivative it's actually the function $\text{sgn}(x)$, (in this case, since it's a locally integrable function it coincides with its weak derivative), therefore it's not unusual to see expressions like this

$$\frac{d}{dx} |x| = \text{sgn}(x) \quad \text{weakly/distributionally}$$

The proof of this is quite easy. Taken $f \in \mathcal{K}$

$$\begin{aligned} D \varphi_{|x|}(f) &= -\varphi_{|x|}(f') = - \int_{\mathbb{R}} |x| f'(x) dx = \\ &= [xf(x)]_{\mathbb{R}^-} - [xf(x)]_{\mathbb{R}^+} + \int_{\mathbb{R}^+} f(x) dx - \int_{\mathbb{R}^-} f(x) dx = \\ &= \int_{\mathbb{R}} \text{sgn}(x) f(x) dx = \varphi_{\text{sgn}(x)}(f) \end{aligned}$$

Since we have that

$$D \varphi_g = \varphi_{Dg}$$

Where Dg is intended as the weak derivative of g , we have that

$$D |x| = \text{sgn}(x)$$

Where in order to emphasize that this is a weak derivative, one could use the notation

$$D_w |x| = \text{sgn}(x)$$

Note that in literature it's common to use the abuse of notation written beforehand in general and in this particular case.

Also note that if the function we had used would have been an ordinarily derivable function, the weak derivative would have coincided with the ordinary derivative.

Theorem 10.21. *In general, given $g \in C^1(\mathbb{R})$, we can say that $g(|x|)$ has the following weak derivative*

$$Dg(|x|) = Dg(|x|) \operatorname{sgn}(x)$$

Proof. Since $g(|x|) \in L^1_{loc}(\mathbb{R})$ we can say that exists $g \mapsto \varphi_g \in \mathcal{K}^*$ such that

$$D\varphi_{g(|x|)} = \varphi_{Dg(|x|)}$$

Where Dg is the weak derivative of $g(|x|)$. Since it's a derivative (and it's also demonstrable) we can use the rules for composite derivation, and knowing that $D|x| = \operatorname{sgn}(x)$ we have

$$D\varphi_{g(|x|)} = \varphi_{Dg(|x|)} = \varphi_{Dg D|x|} = \varphi_{\operatorname{sgn} Dg}$$

□

Theorem 10.22 (Derivative of the ϑ Distribution). *Given $\vartheta \in \mathcal{K}^*$ the theta distribution, we define*

$$\vartheta(f) = \int_{\mathbb{R}} f(x)H(x) dx = \int_{\mathbb{R}^+} f(x) dx$$

The derivative of this distribution is

$$D\vartheta = \delta_0$$

Proof. We have $\forall f \in \mathcal{K}$

$$D\vartheta(f) = -\vartheta(f') = -\int_{\mathbb{R}} f'(x)H(x) dx = -\int_{\mathbb{R}^+} f'(x) dx = f(0) = \delta_0(f)$$

Therefore

$$D\vartheta = \delta_0$$

Or, using a common abuse of notation

$$\vartheta = \vartheta(x) = H(x), \quad DH(x) = \delta(x)$$

□

Notation (Piecewise Derivative). Given $g : \mathbb{R} \rightarrow \mathbb{C}$ a piecewise differentiable function, we define the differential operator D_\circ as follows

$$D_\circ g(x) = \begin{cases} g'(x) & \exists g'(x) \\ 0 & \nexists g'(x) \end{cases} \quad (10.16)$$

Theorem 10.23. *Given $f : \mathbb{R} \rightarrow \mathbb{C}$ a piecewise differentiable function and $S_f := \{u_1, \dots, u_k\}$ isolated singularities, then*

$$Df(x) = D_\circ f(x) + \sum_{i=1}^{\infty} \Delta f(u_i) \delta(u_i)$$

Proof. Taken $S_f := \{u\}$ without loss of generality, we have that

$$\begin{aligned}\lim_{x \rightarrow u^+} f(x) &= f(u^+) \\ \lim_{x \rightarrow u^-} f(x) &= f(u^-) \\ \Delta f(u) &= f(u^+) - f(u^-)\end{aligned}$$

And, $\forall g \in \mathcal{K}$, taken $f \mapsto \varphi_f \in \mathcal{K}^*$

$$D \varphi_f(g) = -\varphi_f(g') = -\int_{-\infty}^u g'(x)f(x) dx - \int_u^{\infty} g(x)f(x) dx$$

Integrating by parts and rebuilding the definition of $\Delta f(u)$ we have

$$D \varphi_f(g) = g(u)\Delta f(u) + \int_{\mathbb{R}} g(x) D_{\circ} f(x) dx = \varphi_{D_{\circ} g}(f) + \Delta f(u)\delta_u(f)$$

And therefore

$$D f(x) = D_{\circ} f + \Delta f(u)\delta(u)$$

□

Example 10.2.5 (Deriving the Sign Function). Take $a \in \mathbb{R}$ and the function $\text{sgn}(x - a)$, using the previous formula we have that $S_{\text{sgn}(x-a)} = \{a\}$ and $\Delta \text{sgn}(a) = 2$, therefore

$$D \text{sgn}(x - a) = D_{\circ} \text{sgn}(x - a) + 2\delta(x - a)$$

Since $D_{\circ} \text{sgn}(x - a) = 0$ we have finally

$$D \text{sgn}(x - a) = 2\delta(x - a)$$

And equivalently

$$D \text{sgn}(a - x) = -2\delta(x - a)$$

Example 10.2.6 (A General Piecewise Differentiable Function). Take the function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined as follows

$$g(x) = \begin{cases} 1 & x < 0 \\ x - 2 & x \geq 0 \end{cases}$$

We have that

$$D_{\circ} g(x) = H(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Since the discontinuity is in the origin and we have $\Delta g(0) = -3$ we have

$$D g(x) = H(x) - 3\delta(x)$$

Example 10.2.7 (Derivative of the Floor Function). Given the floor function $[x] : \mathbb{R} \rightarrow \mathbb{Z}$ we have that $S_{[x]} = \mathbb{Z}$, $\Delta[x] = 1$ and $D_{\circ}[x] = 0$ therefore

$$D[x] = \sum_{k \in \mathbb{Z}} \delta(x - k) = \sum_{k=-\infty}^{\infty} \delta(x - k)$$

Given instead a different function, $[x^2]$ we have that $D_\circ[x^2] = 0$ and $S_{[x^2]} = \{k \in \mathbb{Z} \mid \pm\sqrt{k}\} \subset \mathbb{R}$ with $\Delta[\pm\sqrt{k}] = \pm 1$, therefore

$$D[x^2] = \sum_{k \geq 1} \delta(x - \sqrt{k}) - \delta(x - \sqrt{-k}) = \sum_{k=1}^{\infty} \delta(x - \sqrt{k}) - \delta(x - \sqrt{-k})$$

Theorem 10.24. *Given $h \in C^\infty(\mathbb{R})$ a smooth function and $a \in \mathbb{R}$, we have that*

$$h(x) D^n \delta(x - a) = \sum_{k=0}^n (-1)^k \binom{n}{k} h^{(k)}(a) D^{n-k} \delta(x - a)$$

Proof. Taken $f \in \mathcal{K}$ we have that

$$(h(x) D^n \delta_a)(f) = D^n \delta_a(hf) = \delta_a((hf)^{(n)})$$

Using now Leibnitz's composite derivation rule, we have

$$\frac{d^n}{dx^n} h(x) f(x) = \sum_{k=0}^n \binom{n}{k} \frac{d^k h}{dx^k} \frac{d^{n-k} f}{dx^{n-k}}$$

We have that

$$\begin{aligned} (h(x) D^n \delta_a)(f) &= \delta_a \left[(-1)^n \sum_{k=0}^n \binom{n}{k} h^{(k)}(x) f^{(n-k)}(x) \right] = \\ &= \sum_{k=0}^n \binom{n}{k} (-1)^k h^{(k)}(x) D^{n-k} \delta_a(f) \end{aligned}$$

□

Example 10.2.8. Take as an example the following distributional derivative

$$D^7 (xe^x D^2 \delta_0)$$

We have from that formula that

$$xe^x D^2 \delta_0 = \frac{d^2}{dx^2} (xe^x) \delta_0 = \frac{d}{dx} (e^x + xe^x) \delta_0 = (2 + x) e^x \delta_0$$

And therefore $h'(0) = 1$, $h''(0) = 2$

Which means that

$$xe^x D^2 \delta_0 = \binom{2}{2} h(0) D^2 \delta_0 - \binom{2}{1} h'(0) D \delta_0 + \binom{2}{0} h''(0) \delta_0$$

Inserting the values, we get

$$xe^x D^2 \delta_0 = 2 D \delta_0 - 2 \delta_0$$

The derivative is now trivial, and we get

$$D^7 (xe^x D^2 \delta_0) = 2 D^7 (\delta_0 - D \delta_0) = 2 (D^7 \delta_0 - D^8 \delta_0)$$

Theorem 10.25. Given $k, m \in \mathbb{N}$ and a function $f \in C^k(\mathbb{R})$ we have that

$$D^k(x^m f)(0) = \begin{cases} 0 & k < m \\ \binom{k}{m} m! f^{(k-m)}(0) & k \geq m \end{cases}$$

And

$$D^i(x^m)(0) = m! \delta_m^i$$

Proof. We apply immediately the Leibnitz chain rule and we have

$$D^k(x^m f)(0) = \sum_{j=0}^m \binom{k}{j} f^{(k-j)}(0) m! \delta_{jm} = m! \binom{k}{m} f^{(k-m)}(0) \quad \forall k \geq m$$

□

Theorem 10.26 (Properties of the $PV(x^{-n})$ Distribution). Here there will be listed some properties of the $PV(x^{-n})$ distribution

1. $D \log |x| = PV(x^{-n})$
2. Given $n, m \in \mathbb{N}$, then

$$x^m PV\left(\frac{1}{x^n}\right) = \begin{cases} x^{m-n} & m \geq n \\ PV\left(\frac{1}{x^{n-m}}\right) & m < n \end{cases}$$

3. $D^m PV(x^{-n}) = (-m)! PV(x^{-n-m})$

Proof. 1) Taken $\log |x| \in L_{loc}^1(\mathbb{R})$ we know that $\exists \varphi_{\log |x|} \in \mathcal{K}^*$ such that $\log |x| \mapsto \varphi_{\log |x|}$ and therefore we can write as follows the previous derivative, that $\forall f \in \mathcal{K}$

$$D \varphi_{\log |x|}(f) = -\varphi_{\log |x|}(f') = -\lim_{\epsilon \rightarrow 0^-} \int_{-\infty}^{\epsilon} f'(x) \log(-x) dx - \lim_{\epsilon \rightarrow 0^+} \int_{\epsilon}^{\infty} f'(x) \log(x) dx$$

Integrating by parts we get

$$D \varphi_{\log |x|}(f) = \lim_{\epsilon \rightarrow 0^+} \log \epsilon (f(\epsilon) - f(-\epsilon)) + \lim_{\epsilon \rightarrow 0^+} \int_{|x| \geq \epsilon} \frac{f(x)}{x} dx = PV\left(\frac{1}{x^n}\right)(f)$$

Or

$$D \log |x| = PV\left(\frac{1}{x^n}\right)$$

- 2) Taken $f \in \mathcal{K}$

$$\left(x^m PV\left(\frac{1}{x^n}\right)\right)(f) = PV\left(\frac{1}{x^n}\right)(x^m f) = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{1}{x^n} \left(x^m f(x) - \sum_{k=0}^{n-1} \frac{D^k(x^m f)(0)}{k!} x^k\right) dx$$

Taken $m \geq n$ we have that the sum is null $\forall k \leq m$, therefore since $k \leq n-1 < m$ it sums to 0
Therefore

$$\lim_{R \rightarrow \infty} I_R^n(x^m f) = \text{PV} \left(\frac{1}{x^n} \right) (x^m f) = \int_{\mathbb{R}} x^{m-n} f(x) dx = \varphi_{x^{m-n}}(f)$$

Taken $m < n$ we have that

$$\sum_{k=0}^{n-1} \frac{x^k}{k!} D^k(x^m f)(0) = \sum_{k=m}^{n-1} \binom{k}{m} \frac{m!}{k!} f^{(k-m)}(0) = \sum_{k=0}^{n-m-1} \frac{f^{(k)}(0)}{k!} x^k$$

Therefore

$$\lim_{R \rightarrow \infty} I_R^n(x^m f) = \lim_{R \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \int_{-R}^R \frac{1}{x^{n-m}} \left(f(x) - \sum_{k=0}^{n-m-1} \frac{f^{(k)}(0)}{k!} x^k \right) dx = \text{PV} \left(\frac{1}{x^{n-m}} \right) (f)$$

3) Taken $n \in \mathbb{N}$, $n > 0$, $m = 1$ we have

$$D \text{PV} \left(\frac{1}{x^n} \right) (f) = -\text{PV} \left(\frac{1}{x^n} \right) (f') = -\lim_{R \rightarrow \infty} I_R^n(f')$$

Integrating by parts we have

$$\begin{aligned} I_R^n(f') &= \left[\frac{1}{x^n} \left(f(x) - f(0) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(0)}{(k+1)!} x^{k+1} \right) \right]_{-R}^R + \\ &+ \int_{-R}^R \frac{n}{x^{n+1}} \left(f(x) - f(0) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(0)}{(k+1)!} x^{k+1} \right) dx = \\ &= \frac{1}{R^n} \left(f(R) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} R^k \right) - \frac{1}{(-R)^n} \left(f(-R) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} (-R)^k \right) + n I_R^{n+1}(f) \end{aligned}$$

Since $\text{supp } f \subset [-b, b]$ we have two cases, $R > b$ and $R < b$.

Supposing $R > b$ we have

$$I_R^n(f') = n I_R^{n+1}(f) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} \left(\frac{1}{(-R)^{n-k}} - \frac{1}{R^{n-k}} \right)$$

Taken a new index $j = n - k$ such that $j \bmod 2 = 0$ we have

$$I_R^n(f') = n I_R^{n+1}(f) - 2 \sum_{j=0}^n \frac{f^{(j)}(0)}{j!} \frac{1}{R^{n-k}} \rightarrow -n \text{PV} \left(\frac{1}{x^{n+1}} \right) (f)$$

In case that $m > 1$, by iteration we get

$$D^m \text{PV} \left(\frac{1}{x^n} \right) = (-m)! \text{PV} \left(\frac{1}{x^{n+m}} \right)$$

□

Theorem 10.27 (Quello con distribuzione complessa). $\langle ++ \rangle$

$\langle ++ \rangle$

§§ 10.2.4 A Physicist's Trick, The Dirac δ "Function"

Definition 10.2.18 (Composite Delta). Supposing that a Dirac δ exists also as a function, we can imagine describing it as a composite function $(\delta \circ f)(x) = \delta(f(x))$.

Watch out, since $\delta(f(x)) \neq \delta_0(f)$

Definition 10.2.19 (Giving it Some Meaning). Taken $g_n \in L_{loc}^1(\mathbb{R})$ a sequence of functions such that $g_n \in C[-a_n, a_n]$ with $a_n > 0$, $a_n \rightarrow 0$, we have that $\exists! g_n \mapsto \varphi_{g_n}$ such that $\varphi_{g_n} \rightarrow_\star \delta_0$. A good choice for this sequence would be the Gaussian function γ_n defined as follows

$$\gamma_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2} \quad (10.17)$$

Then, we have that $\varphi_{\gamma_n} \rightarrow_\star \delta_0$, or written in a clear abuse of notation with non-existing functions

$$\lim_{n \rightarrow \infty} \gamma_n(x) = \delta(x)$$

Which, actually means that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(x) f(x) dx = f(0) = \delta_0(f) \quad \forall f \in \mathcal{K}$$

Then, letting $b(x) \in C^1(\mathbb{R})$ we can define the following quantity

$$\int_{\mathbb{R}} f(x) \delta(b(x)) dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(b(x)) f(x) dx \quad \forall f \in \mathcal{K}$$

Theorem 10.28. Let $g \in C^1(\mathbb{R})$ be a function with isolated and simple zeros $Z_g := \{x_1, x_2, x_3, \dots\}$, then $\forall f \in \mathcal{K}$ we have

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(g(x)) f(x) dx = \int_{\mathbb{R}} f(x) \delta(g(x)) dx = \sum_{x_k \in Z_g} \frac{\delta_{x_k}(f)}{|g'(x_k)|} \quad (10.18)$$

Proof. Taken $g(x) \in C^1(\mathbb{R})$ such that $g(x) \geq g(y) \quad \forall x \geq y$, supposing that $Z_g = \{x_0\}$ we have, since it's a simple zero, that $g'(x_0) \neq 0$

Taking in the previous integral the substitution $y = g(x)$, we have that $x = g^{-1}(y)$ and $dx = Dg^{-1}(y) dy$, therefore

$$I = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \gamma_n(g(x)) f(x) dx = \lim_{n \rightarrow \infty} \int_{\inf(g)}^{\sup(g)} \gamma_n(y) \frac{f(g^{-1}(y))}{g'(g(y))} dy$$

Evaluating the integral we have

$$I = \frac{f(g^{-1}(0))}{g'(g^{-1}(0))} = \frac{f(x_0)}{g'(x_0)} = \frac{1}{|g'(x_0)|} \delta_{x_0}$$

This is easily generalizable if there is more than one zero □

Example 10.2.9 (Composition with a Constant). Take $\delta(ax)$ with $g(x) = ax$ and $a \neq 0$. Since $Z_g = \{0\}$ using the previous formula we have $g'(0) = a$ and therefore

$$\delta(ax) = \frac{1}{|a|} \delta(x)$$

Example 10.2.10. Taken $g(x) = \text{atan}(x)$ we have that $Z_g := \{0\}$ and

$$g'(x) = \frac{1}{1+x^2} \quad g'(0) = 1$$

Therefore

$$\delta(\text{atan}(x)) = \delta(x)$$

Example 10.2.11 (Composition with a Polynomial). Take $g(x) = (x^2 - a^2)$, then $Z_g := \{-a, a\}$ and $g'(\pm a) = \pm 2a$, therefore, if $a \neq 0$

$$\delta(x^2 - a^2) = \frac{1}{2|a|} (\delta(x - a) + \delta(x + a))$$

Example 10.2.12 (Composition with an Exponential). Taken $g(x) = e^x$ we have that $Z_g = \{\}$ therefore $\delta(e^x) = 0$

Example 10.2.13 (Composition with a Trigonometric Function). Taken now $g(x) = \cos(x)$ we have that $Z_g = \{k \in \mathbb{Z} | k\pi/2\}$ and $g'(x) = -\sin(x)$, therefore

$$g'(x_k) = -\sin\left(\frac{k\pi}{2}\right) = (-1)^{k+1}$$

Therefore

$$\delta(\cos(x)) = \sum_{k \in \mathbb{Z}} \delta\left(x - \frac{k\pi}{2}\right) = \sum_{k=-\infty}^{\infty} \delta\left(x - \frac{k\pi}{2}\right)$$

Example 10.2.14 (Calculating an Integral with the Composite Delta). Take the following integral

$$\int_{\mathbb{R}} e^{-|x|} \delta(\sin(2x)) dx$$

In order for solving it we expand the composite delta. We have that $Z_g = \{k \in \mathbb{Z} | k\pi/2\}$ and $g'(x_k) = (-1)^k 2$, therefore

$$\int_{\mathbb{R}} e^{-|x|} \delta(\sin(2x)) dx = \frac{1}{2} \sum_{k=-\infty}^{\infty} e^{-|x|} \delta\left(x - \frac{k\pi}{2}\right) = \frac{1}{2} \sum_{k=-\infty}^{\infty} e^{-|\frac{k\pi}{2}|}$$

Fixing the sum and summing, we have

$$\int_{\mathbb{R}} e^{-|x|} \delta(\sin(2x)) dx = \sum_{k \in \mathbb{N}} e^{-\frac{k\pi}{2}} - \frac{1}{2} = \frac{1}{1 - e^{\frac{\pi}{2}}} - \frac{1}{2}$$

§§ 10.2.5 Integral Representation of Distributions

<++>

§§ 10.2.6 Distributions in \mathbb{R}^n

§ 10.3 Some Applications in Physics

§§ 10.3.1 Electrostatics

§§ 10.3.2 Electrodynamics

11 Differential Equations

§ 11.1 Existence of Solutions

<++>

§ 11.2 Common Solving Methods

§ 11.3 Green Identities and Kelvin Transform

12 Fourier Calculus

§ 12.1 Bessel Equality and Fourier Coefficients

<++>

§§ 12.1.1 Trigonometric Polynomials in $L^2[-\pi, \pi]$

§ 12.2 Fourier Series

§§ 12.2.1 Fourier Series in $[a, b]$

§§ 12.2.2 Complex Fourier Series

§§ 12.2.3 Heat Equation with Fourier Series

§ 12.3 Fourier Transform

§§ 12.3.1 Convolution

§§ 12.3.2 Heat Equation with Fourier Transforms

§ 12.4 Probabilistic Considerations

Appendices

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.