Minimum Principle of Pontryagin

A Project

Submitted to the African University of Science and Technology (AUST)

in partial fulfilment of the requirement for

MASTER DEGREE IN PURE AND APPLIED MATHEMATICS

Presented

By

DIADIE SOW

Supervisor:

Doctor Ngalla Djitte

African University of Science and Technology

Abuja

Knowledge is Freedom

Abuja-Nigeria

www.aust.edu.ng
P.M.B 681,Garki,Abuja F.C.T
Nigeria.

CONTENTS

Dedication

This project is specially dedicated to my Mother **Fatou Diouf** who did every thing for me during all my life, unfortunately she has left me at the end of this master, her support and love towards during my training in AUST is my succes today.
I also dedicated this project to my family: my father Madia Sow, my brothers Djily Sow, Cheikh Sow, and to my sisters Khoudia Sow, Betor Sow, Maguete Sow.
Finaly i extend my dedication to my best friends Bay lamine Fall, Cheikh Diop, Abdoul Aziz Mbengue, Mademba Diop, for ever. But also specially to my roomate and friend, Issa Tahir Bachar, his support to me is inexpressible, he was always present for me at the difficult moment of my life i know from him that a friend in need is a friend ended.

Preface

This Project is at the interface between Optimization, Functional analysis and Differential equation. It concerns one of the powerful methods often used to solve optimization problems with constraints; namely Minimum Pontryagin Method. It is more precisely an optimization problem with constrain, an ordinary differential equation. Their applications cover variational calculos as well as applied areas including optimization, economics, control theory and Game theory. But we shall focus on a branch linking minimization and differential equations.
My interest in this subject has been steadily fascinated by the successive lectures delivered at the African University of Sciences and Technology by Prof. C. Chidume (Functional Analysis), Dr. N. Djitte (Sobolev spaces and linear elliptic partial differential equations,Topologie and Variational method), Dr G. Degla (Topics in Differential Analysis) and Prof. Thibault (Measure and Integration theory)

Acknowledgements

First and foremost, I would like to express my sincere gratitude to the Acting President of the African University of Science and Technology, Professor C.E. Chidume, for his advice and leadership at the Institute of Mathematics, and to my supervisor Doctor Ngala Gjitte for motivating and guiding me.
I am also deeply grateful to Professor Doctor Guy Degla for his thorough clarifications about advanced mathematical concepts, and for his frequent remarks and explanations which reinforced my mathematical background
Moreover I am really grateful to all the authorities of the African University of Science and Technology as well as the staff for their hospitality
Eventually, I am heartily thankful to the staff of my home University, also to all de Professors who had given me knowledge since the beginning of my study.

Introduction and Motivations

In order to make the concepts clear, let us recall some keywords. Given a nonempty set X and a function $f : X \longmapsto \mathbb{R}$ which is bounded below, computing the number $\inf_X F := \inf\{F(x) : x \in X\}$ represent a minimization problem posed in X:namely that of finding a minimizing sequence, i.e. $(x_k)_k$ include in X such that $\lim_{k\to\infty} F(x_k) = \inf_X F$. The number $\inf_X F$ is often called the infinimal value of f or more simply the infinimum of f over X. The function f is usually called the objective function or also infinimand. By analogy we have the concepts of supremal value (supremum) and supremand.
An optimal solution of that problem is an element a in X such that
$F(a) \leq F(x)$ for all x in x
such an element a is usually called a minimizer, a minimum point or simply a minimum of f on X. We shall also speak of global minimum. Let us emphasize that the notation

$$\min\{F(x), x \in X\}$$

holds at the same time for a number (when there exists a solution) and a problem to solve.
Likewise one can meet maximization problems but they are all equivalent to minimization problems since for any real valued function g defined on a set X, one has $\sup\{g(x) : x \; inX \;\} = -\inf\{-g(x) : x \in X\}$
When X has a topological structure, another problem related with the above one is to know whether a giving minimizing sequence (x_k) converges to an optimal solution when k tends to ∞.Two conditions are essencial to guarantee a positive answer to the above problem.A topological criterion on the structure of X (e.g., compactness) and a topological criterion on the behavior of the function f (e.g., continuity).
When X is an open set of a real normed linear space (respectively a manifold) and f is Frechet differentiable or just Gateaux differentiable (respectively differentiable in the geometric sense), a necessary condition for a point a in X to be a minimizer (according to Euler) is to be a critical (or stationary) point of f;this means that $F'(x) = 0$ on X(respectively $df(a) = 0$ on T_aX,the tangent space of the manifold X at a).We say that a real number c is a critical value of f if there exists a critical point $a \in X$ such that f(a) = c. In the case of a Hilbert space X

$=$ H endowed with a scalar product $\langle ; \rangle$ and thanks to the Riesz representation theorem, the gradient of a Gateaux differentiable is defined by setting

$$\langle h, \nabla F(x) \rangle = F'(x).h$$

And so in this case, a critical point of f is just a solution of the equation

$$\nabla F(x) = 0$$

Let V be a norm linear space and F be a real value function defined on V. Let $K \subset V$ be a nonempty subset of V. Our aim is to solve the minimization problem

$$\inf_{v \in K} F(v) \tag{1}$$

means to find $\bar{u} \in K$ such that $F(\bar{u}) \leq F(v)$ for all $v \in K$. Such a $\bar{u}$ (if it exists) is then called the solution of (1). The quantity $inf F(v)$(possibly worth $-\infty$) is called the value of (1)
This value is different from $-\infty$ if and only if F is bounded below on K, obviously only this case is of interest. Indeed the value of is (1) is by definition the lower bounded of the subset of $\mathbb{R}, F(K) = \{F(v)\ v \in K\}$.
Finally we write (1) in the form of a minimization problem, which includes also the maximization problems,indeed maximizer a problem is to minimize its opposition. One take $\alpha = inf F(K)$, then there exists a sequence (v_n) of K such that:

$$\lim_{n \to \infty} F(v_n) = \alpha \tag{2}$$

Any sequence of K satisfying (2) is called a minimizing sequence of the problem (2).Note that without any additional hypothesis,there always exists a minimizing sequence of (2)

CHAPTER 1

Preliminaries

1.1 Linear maps

In this part we define linear map and present some basic results concerning them.

Definition :Let X and linear spaces over a scalar field K.A mapping $T : X \longrightarrow Y$ is said to be a linear map if:

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \tag{1.1}$$

for arbitrary $x, y \in X$ and arbitrary scalars $\alpha, \beta \in K$.Some authors use the term linear operator or linear transformation instead of linear map.Condition (1.1) is equivalent to the following two conditions:
$(i)T(x+y) = T(x) + T(y) \forall x, y \in X$
$(ii)T(\alpha x) = \alpha T(x) \forall x \in X$ and for each scalar, α.

1.1.1 A basic result concerning linear maps

We remark first that since linear functionals are special forms of linear maps,any result proved for linear map holds for linear functionals.

Proposition 1.1.1 *Let X and Y be two linear spaces over a scalar field, K, and let $T : X \longrightarrow Y$ be a linear map.Then*

1. $T(0) = 0$
2. *The rang of T, $R(T) = \{y \in Y : T(x) = y$ for some $x \in X\}$ is a linear subspace of Y*
3. *T is one to one if and only if $T(x) = 0$ implies that $x = 0$*
4. *If T is one to one , then T^{-1} exist on $R(T)$ and $T^{-1} : R(T) \longrightarrow X$ is also a linear map.*

Proof. (1) Since T is linear,we have, $T(\alpha x) = \alpha T(x)$ for each $x \in X$ and each scalar α. Take $\alpha = 0$ and (1) follows immediately.
(2)We need to show that for $y_1, y_2 \in R(T)$ and α, β scalars, $\alpha y_1 + \beta y_2 \in R(T)$. Now, $y_1, y_2 \in R(T)$ implies that there exists $x_1, x_2 \in X$ such that $T(x_1) = y_1, T(x_2) = y_2$. Moreover, $\alpha x_1 + \beta x_2 \in X$(since X is a linear space). Furthermore, by the linearity of T,we have

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2)$$

Hence $\alpha y_1 + \beta y_2 \in R(T)$, and so $R(T)$ is a linear subspace of Y.
(3)($\Rightarrow$)Assume that T is one to one. Clearly $T(x) = 0 \Rightarrow T(x) = T(0)$ since T is linear (and so $T(0) = 0$). But T is one-to-one.So,$x = 0$.
($\Leftarrow$)Assume that whenever $T(u) = 0$ then u must be 0. We want to prove that T is one-to-one.So,let $T(x) = T(y)$. Then $T(x) - T(y) = 0$ and by the linearity of $T, T(x - y) = 0$.By hypothesis, $x - y = 0$ which implies $x = y$. Hence T is one to-one.
(4)Assume that T is one-to-one,since the restriction of T on $R(T)$ is always onto, then T is bijective from X into $R(T)$.So T^{-1} exists on $R(T)$. For the linearity, let $y_1, y_2 \in R(T)$ and α a scalar. Then there exists $x_1, x_2 \in X$ such that $y_1 = T(x_1), y_2 = T(x_2)$. so

$$T^{-1}(y_1 + \alpha y_2) = T^{-1}(T(x_1) + \alpha T(x_2)$$

which is equivalent to:

$$T^{-1}(y_1 + \alpha y_2) = T^{-1}(T(x_1 + \alpha x_2)) \text{ by using the linearity of T,}$$

which is also equivalent to

$$T^{-1}(y_1 + \alpha y_2) = x_1 + \alpha x_2 = T^{-1}(y_1) + \alpha T^{-1}(y_2)$$

Therefore T^{-1} is linear

1.1.2 Bounded Linear Maps

Definition : Let X and Y be normed linear spaces over a scalar field K, and let $T : X \longrightarrow Y$ be a linear map. Then T is said to be bounded if there exists some constant $K \geq 0$ such that for each $x \in X$,

$$\|T(x)\| \leq K\|x\|;$$

the constant K is called a bound for T and in this case, T is called a bounded linear map. We denote by $B(X, Y)$, the family of all bounded linear maps from X into Y.
We now turn our attention to linear maps that are continuous. The notion of continuity can be state, for linear maps in several useful equivalent forms. We state these equivalent forms in the following theorem.

Theorem 1.1.1 *Let X and Y be normed linear space and let $T : X \longrightarrow Y$ be a linear map.Then the following are equivalent:*

1. *T is continuous;*

2. *T is continuous at the origin(in the sense that if* $\{x_n\}$ *is a sequence of* X *such that* $x_n \to 0$ *as* $n \to +\infty$,*then* $T(x_n) \to 0$ *as* $n \to +\infty$*)*

3. *T is Lipschitz, ie , there exists a constant* $K \geq 0$ *such that,for each* $x \in X$,

$$\|T(x)\| \leq K\|x\|$$

   ;

4. *If* $D = \{x \in X : \|x\| \leq 1\}$ *is the closed unit disc in* X*, then* $T(D)$ *is bounded(in the sense that there exists a constant* $M \geq 0$ *such that* $\|T(x)\| \leq M$ *for all* $x \in D$*)*

1.2 Banach spaces

Definition Let X be a real linear space, and $\|.\|_X$ a norm on X , and d_X the corresponding metric defined by $d_X(x, y) = \|x-y\|_X \forall x, y \in X$. The norm linear space $(X, \|.\|_X)$ is a real Banach space if the metric space (X, d_X) is complete, ie if any Cauchy sequence of elements of space $(X, \|.\|_X)$ converges in $(X, \|.\|_X.)$ That is every sequence satisfying the the following Cauchy criterion:

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N} : p, q > n_0 \Rightarrow d_X(x_p, x_q) < \epsilon$$

Definition Given any vector space V over a vector field $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), the topological dual space(or simply) dual space of V is the linear space of all bounded linear functionals. We shall denote it by V^*.

Remark 1.2.1 1. The dual space V^* has a canonical norm defined by

$$\|f\| = \sup_{x \in V, \|x\| \neq 0} \frac{|f(x)|}{\|x\|}, \forall f \in V^*. \tag{1.2}$$

2. The dual of every real normed linear space, endowed with its canonical norm is a Banach space.

In order to define other useful topologies on dual spaces, we recall the following:

Definition (Initial topology) Let X be an nonempty set, $\{Y_i\}_{i \in I}$ be a family of topological spaces(where I is an arbitrary index set)and $\phi_i : X \longrightarrow Y; i \in I$ a family of maps.
The smallest topology on X such that the map $\phi_i, i \in I$ are continuous is called the initial topology.

Next, we define the weak topology of a normed vector space X and the weak star topology of its dual space X^* which are special initial topologies.

Definition (weak topology) Let X be a real normed linear space, and let us associate to each $f \in X^*$ the map $\phi_f : X \longrightarrow \mathbb{R}$ given by $\phi_f(x) = f(x) \forall x \in X$
The weak topology on X is the smallest topology on X for which all the ϕ_f are continuous.
We write ω-topology for the weak topology.

Definition (weak star topology) Let X be a real normed linear space and X^* its dual. Let us associate to each $x \in X$ the map $\phi_x : X^* \longrightarrow \mathbb{R}$ given by $\phi_x(f) = f(x) \forall f \in X^*$.
The weak star topology on X^* is the smallest topology on X^* for which all the ϕ_x are continuous. We write ω^*-topology for the weak star topology

Proposition 1.2.1 *Let X be a real normed linear space and X^* its dual space. Then, there exists on X^* three standard topologies, the strong topology given by the canonical norm $\|.\|_{x^*}$, the weak topology(ω-topology) and the weak star topology ((ω^*-topology) such that:*
$(X^*, \omega^*) \hookrightarrow (X^*, \omega) \hookrightarrow (X^*, \|.\|_{X^*})$.

The following part of this section is devoted to reflexive spaces. For any normed real linear space X; the space X^* of all bounded linear functionals on X is a Banach space and as a linear space, it has its own corresponding . Let X be a Banach space, there exists a natural mapping $J : X \longrightarrow X^{**}$ of X into X^{**} defined, for each $x \in X$ by $J(x) = \phi_x$ where $\phi_x : X \longrightarrow \mathbb{R}$ is given by $\phi_x(f) = \langle f, x \rangle$, for each $f \in X^*$. Thus,

$$\langle J(x), f \rangle = \langle f, x \rangle \text{ for each }, \quad f \in X^*$$

We verify the following properties of J:
(i)J is linear(which is trivial);
(ii)$\|J(x)\| = \|x\|$ for all $x \in X$;ie J is an isometry.In fact, for each $f \in X^*, \|J(x)\| = \sup_{\|f\|=1, f \in X^*} |\langle f, x \rangle| = \|x\|$.
In general, the map J need not to be onto. Consequently, we always identify X as a subspace of X^{**}. Since an isometry is always injective, it follows that J is an isomorphism onto $J(X) \subset X^{**}$. The mapping J defined above is called the canonical map(or canonical embedding) of X into X^{**}, and the space X is said to be embedded in X^{**}. This leads to the following definition.

Definition Let X be a norm linear space and let J be the canonical embedding of X into X^{**}. If J is onto, then X is called reflexive. Thus, a reflexive Banach space is one in which the canonical embedding is onto.

We now state the following important theorem.

Theorem 1.2.1 *(Eberlein-Smul'yan theorem)*
A real Banach space X is reflexive if and only if every (norm) bounded sequence in X has a subsequence which converges weakly to an element of X.

1.3 Hilbert Spaces

Definition 1. A map $\phi : E \times E \longrightarrow \mathbb{C}$ is **Sesquilinear** if
(i)$\phi(x + y, z + w) = \phi(x, z) + \phi(x, w) + \phi(y, z) + \phi(y, w)$
(ii)$\phi(ax, by) = \bar{a}b\phi(x, y)$, where the "bar" indicates the complex conjugation

for all $x, y, z, w \in E$ and for all $a, b \in \mathbb{C}$.

2. A**Hermitian form** is a sesquilinear form $\phi : E \times E \longrightarrow \mathbb{C}$ such that $\phi(x, y) = \bar{\phi}(y, x)$
3. A **positive Hermitian form** is a Hermitian form such that $\phi(x, x) \geq 0$ for all $x \in E$.
4. A **definite Hermitian form** is a Hermitian form such that $\phi(x, x) = 0 \Rightarrow x = 0$
5. An **inner product** on E is a **positive definite Hermitian form** and will be denoted $\langle ., . \rangle := \phi(., .)$. The pair$(E, \langle , \rangle)$ is called an **inner product space**.
   We shall simply write E for the inner product space (E, $\langle , \rangle$) when the inner product $\langle , \rangle$) is known.
   In the case where we are using more than one inner product spaces, specification will be made by writing h; $\langle , \rangle_E$ when talking about the inner product space (E; $\langle , \rangle$)

Definition Two vectors x and y in an inner product space E are said to be orthogonal if $\langle x, y \rangle = 0$. For a subset F of E, we have x is orthogonal to F if x is orthogonal to y for all y in F.

Proposition 1.3.1 *Let E be an inner product space and$x; y \in E$: Then*

$$\|\langle x, y \rangle\| \leq (\langle x, x \rangle)^{\frac{1}{2}} (\langle y, y \rangle)^{\frac{1}{2}}$$

For an inner product space (E; $\langle , \rangle$),the function $\|.\| : E \longrightarrow \mathbb{R}$ defined by

$$\|x\|_E = \sqrt{\langle x, x \rangle_E}$$

is a norm on E
Thus, $(E; \|.\|_E$) is a normed vector space, hence a metric space endowed with the distance
$d_E : E \times E \longrightarrow \mathbb{R}$ defined by $d_E(x, y) = \|x - y\|_E$.

Definition (Hilbert Space). An inner product space E is called a **Hilbert space** if it is complete.

Remark 1.3.1
1. Hilbert spaces are thus a special class of Banach spaces.
2. Every finite dimension inner product space is complete and simply called **Euclidian Space**.

Proposition 1.3.2 *: Let H be a Hilbert space. Then,*
for all $u \in H, T_u(v) := \langle u, v \rangle$ defines a bounded linear functional,ie $T_u \in H^*$. Furthermore $\|u\|_H = \|T_u\|_{H^*}$

Theorem 1.3.1 *(Riesz Representation theorem) Let H be a Hilbert space and let f be a bounded linear functional on H: Then,*

1. *There exists a unique vector* $y_0 \in H$ *such that*

$$f(x) = \langle x, y_0 \rangle \text{ for each }, \quad x \in H$$

2. *Moreover,* $\|f\| = \|y_0\|$.

Remark 1.3.2 The map $T : H \longrightarrow H^*$ defined by $T(u) = T_u$ is linear,(anti-linear in the complex case) and isometric. Therefore the canonical embedding is an isometry showing that "**any Hilbert space is reflexive**".

At the end of this part, we state this important proposition which is just a corollary of Eberlein-Smul'yan theorem.

Proposition 1.3.3 *Let H be a Hilbert space, then any bounded sequence in H has a subsequence which converges weakly to an element of H*

1.4 Differential Calculus in Banach spaces

In this section we define the derivative of a map defined between real Banach spaces.

Definition (Directional Differentiability) Let f be a function defined from a real linear space X into a real normed linear space Y and let x_0 in X and v in X/$\{0\}$. The function f is said to be differentiable at x_0 in the direction v if the function $t \longmapsto f(x_0 + tv)$ is differentiable at $t = 0$. i.e.

$$t \longmapsto \frac{f(x_0 + tv) - f(x_0)}{t}; t \neq 0$$

has a limit in Y when t tends to 0.This,when it exists is denoted $f'(x_o, v)$ or $\frac{\partial f}{\partial v}(x_0)$

Definition (Gateau Differentiability) A function f defined from a real linear space X into a real normed linear space Y is Gateaux Differentiable at a point x_0 in X if :

1. f is differentiable at x_0 in every direction v
2. there exists a bounded linear map $A : X \longrightarrow Y$ such that $f'(x_0, v) = A(v)$;in others words,the map$v \longmapsto f'(x_0, v)$ is a bounded linear map from X into Y
   In this case the map $f'(x_0; :)$ is called the Gateaux differential of f at x_0 and is denoted by $D_G f(x_0; .)$ or $f'_G(x_0)$.

Definition (Frechet Differentiability) A map $f : U \subset X \longrightarrow Y$ whose domain U is an open set of a real Banach space X and whose range is a real Banach space Y is (Frechet) differentiable at $x \in U$ if there exists a bounded linear map $A : X \longrightarrow Y$ such that

$$\lim_{\|u\| \to 0} \left\| \frac{f(x + u) - f(x) - Au}{\|u\|} \right\| = 0$$

or equivalently to

$$f(x + u) - f(x) - Au = o(\|u\|)$$

Proposition 1.4.1 *: If $f : U \subset X \longrightarrow Y$ is Frechet Differentiable, then f is Gateaux Differentiable.*

Proof. *Indeed by taking $u = tv$; in the definition of Frechet Differentiability we have:$f(x + tv) - f(x) = A(tv) + o(\|tv\|)$. Simplifying by t and using the fact that A is linear,we have:*

$$\frac{f(x+tv) - F(x)}{t} = (A(v) + \frac{(o\|tv\|)}{t})$$

by the Frechet Differentiability of f And since as $t \to 0, u \to 0$ so we have

$$\lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} = Av$$

Proposition 1.4.2 *Let X be a real Banach space and Y be a real normed linear space. Then*

1. *The set of Gateaux differentiable mappings from X into Y is a linear subspace of the linear space of all the mappings defined from X into Y space is contained in $B(X;Y)$;*

2. *The set of Frechet Differentiable mappings from X into Y is also a subspace of $B(X;Y)$*

1.5 Convex sets and convex functions

Let us recall first the following basic notions of optimization.

Definition Let V be a norm linear space, and let $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$.

1. A point $\bar{u} \in K$ is a minimizer of F on K if

$$F(\bar{u}) \leq F(v) \forall v \in K$$

2. A point $\bar{u} \in K$ is a local minimizer of F on K if there exists $r > 0$ such that

$$F(\bar{u}) \leq F(v) \forall v \in K \cap B(\bar{u}, r)$$

3. A point $\bar{u} \in K$ is a strict minimizer of F on K if

$$F(\bar{u}) < F(v) \forall v \in K, v \neq \bar{u}$$

4. A point $\bar{u} \in K$ is a strict local minimizer of F on K if there exists $r > 0$ such that

$$F(\bar{u}) < F(v) \forall v \in K \cap B(\bar{u}, r), v \neq \bar{u}$$

Definition Let X be a norm linear space.A set $C \subseteq X$ is convex if and only if: $\forall x, y \in C, \forall \lambda \in (0,1)$, we have $\lambda x + (1 - \lambda) y \in C$
ie $[x, y] \subseteq C \quad , \quad \forall x, y \in C$

Definition Let C be a convex subset of a norm linear space X. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is convex on C if $\forall x, y \in C, \forall \lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \tag{1.3}$$

If (1.3) is strict for $x, y \in C$, with $x \neq y$ and $f(x), f(y)$ finite then f is strictly convex.
f linear functional implies that f is convex and concave

Theorem 1.5.1 *(Slope Inequality) Let I be an interval of $\mathbb{R}$ and let $f : I \longrightarrow \mathbb{R} \cup \{+\infty\}$ convex. Let $r_1, r_2, r_3 \in I : r_1 < r_2 < r_3$;with $f(r_1), f(r_2)$ finite then:*

$$\frac{f(r_1) - f(r_2)}{r_2 - r_1} \leq \frac{f(r_3) - f(r_1)}{r_3 - r_1} \leq \frac{f(r_3) - f(r_2)}{r_3 - r_2}$$

Proof. Set $\lambda = \dfrac{r_2 - r_1}{r_3 - r_1} \in (0,1)$. We have $1 - \lambda = \dfrac{r_3 - r_2}{r_3 - r_1} \in (0,1)$ and $r_2 = (1-\lambda)r_1 + \lambda r_3$. Using the convexity of f, it follows that $f(r_2) \leq \lambda f(r_3) + (1-\lambda)f(r_1) = \dfrac{r_3 - r_2}{r_3 - r_1} f(r_1) + \dfrac{r_2 - r_1}{r_3 - r_1} f(r_3)$. Adding $-f(r_1)$ in both sides,we get $f(r_2) - f(r_1) \leq \dfrac{r_1 - r_2}{r_3 - r_1} f(r_1) + \dfrac{r_2 - r_1}{r_3 - r_1} f(r_3)$. Which implies that $\dfrac{f(r_2) - f(r_1)}{r_2 - r_1} \leq \dfrac{f(r_3) - f(r_1)}{r_3 - r_1}$. This prove first inequality. Replacing $-f(r_1)$ by $-f(r_3)$ in this first inequality and using the same argument, we get the second inequality.

Corollary 1.5.1 *Let $f : I \longrightarrow \mathbb{R}$ derivable on I, then The following are equivalent:*

1. *f is convex*
2. *f' is increasing*
3. *$f(y) \geq f(x) + f'(x)(y - x) \forall x, y \in I$*
   If f is twice derivable on I the f is convex if and only if $f'(x) \geq 0 \forall x \in I$

Proof.
(1) $\Rightarrow$ (2). Let $r < t$.By the first inequality of the **slope inequality** we have $\dfrac{f(s) - f(r)}{s - r} \leq \dfrac{f(t) - f(r)}{t - r}$, for $r < s < t$. So we have, $f'(r) = \lim\limits_{s \searrow r} \dfrac{f(s) - f(r)}{s - r} \leq \dfrac{f(t) - f(r)}{t - r}$. By taking $s \nearrow t$, and by using the second inequality of the **slope inequality**, we have $\dfrac{f(t) - f(r)}{t - r} \leq \lim\limits_{s \nearrow t} \dfrac{f(s) - f(t)}{s - t} = f'(t)$. Therefore we have $f'(r) \leq f'(t) \forall r < t$. Thus f' is increasing.
(2)$\Rightarrow$ (3)
Let $x \in I$, let us define $g(y) = f(y) - f(x) - f'(x)(y - x)$. We have that g is differentiable and $g'(y) = f'(y) - f'(x)$. this give us:$g'(y) \geq 0$ if and only if $f'(y) \geq f'(x)$, if and only if $y \geq x$ since f' is increasing. And also we have,$g'(y) \leq 0$if and only if $y \leq x$. So by studding the function g, we have that x is a minimizer of this function. So we have $g(y) \geq g(x) \forall y$. Since $g(x) = 0$,this implies that $g(y) \geq 0$ which implies $f(y) - f(x) - f'(x)(y - x) \geq 0$,

ie $f(y) \geq f(x) + f'(x)(y-x)$, ie (3)
(3)$\Rightarrow$ (1)
We have by assumption that $f(y) \geq f(x) + f'(x)(y-x) \forall x \in I$, which implies that $f(y) \geq \sup\limits_{x\in I}[f(x) + f'(x)(y-x)]$,which is equivalent to $f(y) \geq \sup\limits_{x\in I} h_x(y)$, where $h_x(y) = f(x)+f'(x)(y-x) = f'(x).y+f(x)-f'(x).x$. Since for $x = y$ we have equality on the previous inequality,then we have: $f(y) = \sup\limits_{x\in I} h_x(y)$, which give us that f is convex.

Theorem 1.5.2 *Let X a norm linear space, $U \subseteq X$, is open, convex, nonempty. Let $f : U \longrightarrow \mathbb{R}$ differentiable, then the following are equivalent:*

1. *f is convex*
2. *$f' : U \longrightarrow X^*$ is monotone increasing, ie $\langle f'(y)-f'(x), y-x\rangle \geq 0 \forall x, y \in U$*
3. *$f(y) \geq f(x) + \langle f'(x), y-x\rangle \forall x, y \in U$*
   If f is twice differentiable then f is convex if and only if f"(x) is positive semi-definite ie:
   $\langle f'(x).y, y\rangle \geq 0 \forall y \in X, \forall x \in U$

Proof.
Let $x, y \in U$, define $I = \{s \in \mathbb{R} : x + s(y-x) \in U\}$
Claim:I is an interval of $\mathbb{R}$ such that $0, 1 \in I$
Let $s_1, s_2 \in I, t \in (0,1)$, then we have: $x + s_1(y-x) \in U$ and $x + s_2(y-x) \in U$. Our aim is to show that $x + (ts_1 + (1-t)s_2)(y-x) \in U$ to get that $ts_1 + (1-t)s_2(y-x) \in I$, which is equivalent to say that I is an interval. For that, we have $x + (ts_1 + (1-t)s_2)(y-x) = x + ts_1(y-x) + (1-t)s_2(y-x)$

$$= tx + (1-t)x + ts_1(y-x) + (1-t)s_2(y-x)$$

$$= t(x + s_1(y-x)) + (1-t)(x + s_2(y-x))$$

Since $x + s_1(y-x) \in U$ and $x + s_2(y-x) \in U$ and U is convex, $t \in (0,1)$, we have $= t(x+s_1(y-x)) + (1-t)(x+s_2(y-x)) \in U$. Which is equivalent to say $x + (ts_1 + (1-t)s_2)(y-x) \in U$, ie I is an interval.
Assume that (1) is true, we want to show (2), ie f' is monotone. Let us define $h : I \longrightarrow \mathbb{R}$ by $h(s) = f(x + s(y-x))$
We have that:
i)h is derivable on I
ii)h is convex if and only if f is convex on U
f convex implies that h is convex. so according to the previous theorem, we have h' is increasing, which give us that $h'(1) \geq h'(0)$
Buy differentiating h, we have $h'(s) = \langle f'(x + s(y-x)), y-x\rangle$. So $h'(0) = \langle f'(x), y-x\rangle$ and $h'(1) = \langle f'(y), y-x\rangle$. So using our inequality $h'(1) \geq h'(0)$, we have $\langle f'(y), y-x\rangle \geq \langle f'(x), y-x\rangle$, which is equivalent to f' is monotone.
we assume (2), ie a dire f is monotone, we have to prove first that f' monotone implies that h' is increasing. For that, let us evaluate the difference $(h'(s_2) - h'(s_1))(s_2 - s_1)$. We have that
$(h'(s_2) - h'(s_1))(s_2 - s_1) = (\langle f'(x+s_2(y-x) - f'(y+s_1(y-x), y-x\rangle)(s_2 - s_1)$

$$= \langle f'(x + s_2(y-x) - f'(y + s_1(y-x), y-x, (s_2 - s_1)(y-x)\rangle$$

$$= \langle f'(z_2) - f'(z_1), z_2 - z_1 \rangle$$

By using the monotony of f', we have that $\langle f'(z_2) - f'(z_1), z_2 - z_1 \rangle \geq 0$,which implies that $(h'(s_2) - h'(s_1))(s_2 - s_1) \geq 0$, which means that h' is increasing. So buy using our previous theorem, we have:$h(s) \geq h(r) + h'(r)(s-r) \forall s, r \in I$. So $h(1) \geq h(0) + h'(0)$, which implies that $f(y) \geq f(x) + \langle f'(x), y-x \rangle \forall x, y \in U$, ie (3)
we assume (3) and we have to (1), ie f is convex. We have by hypothesis that $\Rightarrow f(y) \geq f(x) + \langle f'(x), y - x \rangle$, which implies that

$$f(y) = \sup_{x \in U}(f(x) + \langle f'(x), y - x \rangle) = \sup_{x \in U} f_x(y)$$

where $f_x(y) = f(x) + \langle f'(x), y - x \rangle$.
Therefore f is convex.

Definition (Domain of a Function) Let $F : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a map. The domain of F is the set defined by

$$\text{dom}(F) := \{x \in X : F(x) < +\infty\}$$

Domain of F is sometimes called the effective domain of F. The map F is called Proper if $D(F) \neq \emptyset$.Recall that this means there exists at least one $x_0 \in D(F)$ such that $F(x_0) \in \mathbb{R}$ or F is not identically $+\infty$

1.5.1 Notation and Further definitions

Let X be a real normed space and $D \subset X$ is a convex subset of X. Let $f : D \longrightarrow \mathbb{R}$ be a convex function. We define the convex extension of f, $F : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ by

$$F(x) = \begin{cases} f(x), & \text{if } x \in D \\ +\infty, & \text{if } x \in X \backslash D \end{cases} \tag{1.4}$$

We observe that f is convex on D if and only if F is convex on X. Moreover $\text{dom}(f) = \text{dom}(F)$.

Definition The epigraph of F is the subset of $X] \times \mathbb{R}$ denoted by $\text{epi}(F)$ and defined by

$$\text{epi}(F) = \{(x, \alpha) \in X \times \mathbb{R} : x \in \text{dom}(F) \text{ and, } \quad F(x) \leq \alpha\}.$$

Definition Let $\alpha \in \mathbb{R}$. We have the following definitions:

$$S_{F,\alpha} = \{x \in X : F(x) \leq \alpha\} = \{x \in D(F) : F(x) \leq \alpha\}$$

Proposition 1.5.1 *F is convex if and only if epi(F) is convex*

Proof.
$\Rightarrow$)Assume that F is convex. Let $(x_1, \alpha_1) \in epi(F), (x_2, \alpha_2) \in epi(F), t \in (0, 1)$. Now $F(x_1) \leq \alpha_1, F(x_2) \leq \alpha_2$. Hence $F(tx_1 + (1-t)x_2) \leq tF(x_1) + (1-t)F(x_2)$

$$\leq t\alpha_1 + (1 - t)\alpha_2.$$

Thus $(tx_1 + (1-t)x_2, t\alpha_1 + (1-t)\alpha_2) \in epi(F)$. But $(tx_1 + (1-t)x_2, t\alpha_1 + (1-t)\alpha_2) = t(x_1, \alpha_1) + (1-t)(x_2, \alpha_2)$. Hence, $epi(F)$ is convex.
$(\Leftarrow)$ Assume that $epi(F)$ is convex. We show that F is convex. Let $x_1, x_2 \in D(F), t \in (0,1)$. Then $x_1, x_2 \in D(F)$ implies that $F(x_1) \in \mathbb{R}, F(x_2) \in \mathbb{R}$. Thus $(x_1, F(x_1)) \in epi(F)$ and $(x_2, F(x_2) \in epi(F)$. But $epi(F)$ is convex,thus $t(x_1, F(x_1) + (1-t)(x_2, F(x_2)) \in epi(F)$ if and only if $F(tx_1 + (1-t)x_2) \leq tF(x_1) + (1-t)F(x_2)$ if and only if f is convex.

1.6 Lower Semi-Continuous Functions

Let V be a real norm linear space, let $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a map.

Definition : We says that F is lower semi-continuous(lsc)at $x_0 \in Dom(F)$ if:

$$\text{Given,} \quad \epsilon > 0, \exists \eta > 0 : \|x - x_0\| < \eta \Rightarrow F(x_0) - \epsilon < F(x).$$

Definition One says that F is lsc if and only if $epi(F)$ is closed in $X \times \mathbb{R}$

Proposition 1.6.1 *: Let $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$. F is lower semi-continuous if and only if for all sequence (x_n) of V converging to $x \in V$, we have $F(x) \leq$* $\liminf F(x_n.)$

Proof. Assume that F is lsc and $x_n \to x$ in V. We take a subsequence (x_{nk}) of (x_n) such that $\lim_k F(x_{nk}) = \liminf_n F(x_n)$. It follows that$(x_{nk}, F(x_{nk})$ is a sequence of $epi(F)$ converging to $(x, \liminf_n F(x_n))$. Since $epi(F)$ is closed, then $(x, \liminf_n F(x_n)) \in epi(F)$. Thus
$F(x) \leq \liminf_n F(x_n)$
Reversely suppose that$(x_n \to x\text{in } V) \Rightarrow F(x) \leq \liminf_n F(x_n)$. Let (a_n, x_n) be a sequence of $epi(F)$ converging to (x, a) in $V \times \mathbb{R}$. Then $x_n \to x$ and $a_n \to a$. So by hypothesis we have $F(x) \leq \liminf_n F(x_n) \leq \liminf a_n = \lim_n a_n = a$
therefore $(x, a) \in epi(F)$

Proposition 1.6.2 *: A function $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ is lsc if and only if for all $a \in \mathbb{R}, S_{a,F} = \{x \in V : F(x) \leq a\}$ is closed.*

Proof. Assume that F is lsc. Let $a \in \mathbb{R}$ and let (x_n) be a sequence of $S_{a,F}$ converges to $x \in V$.
Since $F(x_n) \leq a \forall n$,then $\liminf_n F(x_n) \leq a$. So according to proposition 1, $F(x) \leq \liminf_n F(x_n) \leq a$.
Thus $x \in C_a$, which means that C_a is closed.
Reversely assume that $S_{a,F}$ is closed for each $a \in \mathbb{R}$. Let x_n be a sequence of V converging to $x \in V$. There exists a subsequence x_{nk} of x_n such that

$$\lim_k F(x_{nk}) = \liminf_n F(x_n)$$

Assume that $F(x) > \liminf_n F(x_n)$. Then there exists $a \in \mathbb{R}$ such that

$$\liminf_n F(x_n) < a < F(x). \tag{1.5}$$

So there exists $N \in IN$ such that $F(x_{nk} < a$ for all $k \geq N$. So $x_{nk} \in S_{a,F} \forall k \geq N$. Since $x_{nk} \to x$ and $S_{a,F}$ is closed then $x \in S_{a,F}$ or equivalent to say that $F(x) \leq a$ which contradict (2.1) Thus $F(x) \leq \liminf_{n} F(x_n)$

Proposition 1.6.3 *Let $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function on a compact set K. Then there exists $\bar{x} \in K$ such that:*

$$F(\bar{x}) = \inf_{x \in K} F(x)$$

.

Proof. Let $\{x_n\}$ be a minimizing sequence of F on K. Since K is compact,there exists a subsequence (x_{nk}) of (x_n) such that $x_{nk} \to \bar{x} \in K$. From Proposition 2 we have:

$$F(\bar{x}) \leq \liminf_{k} F(x_{nk}) = \lim_{k} F(x_{nk}) = \inf_{x \in K} F(x).$$

Therefore $F(\bar{x}) = \inf_{x \in K} F(x)$.

1.7 Existence Result

A concept that will be needed in the proof of one of our fundamental theorems of optimization is that of coercivity. We introduce this concept now

Definition : Let X be a topological space. Let K be a subset of X. The set K is said to be sequentially compact if every sequence in K has a subsequence which converges to an element of K. The subset K is called Weakly sequentially compact if every sequence in K has a subsequence which converges weakly to an element of K.

Definition : Let X be a topological space. A function $F : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is called (weakly, respectively) sequentially coercive if the closure of every section $S_{\lambda,F}$ is (weakly, respectively) sequentially compact in X

Definition :let X be a real reflexive.A function $F : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is called coercive if:

$$\lim_{\|x\| \to +\infty} F(x) = +\infty$$

In fact, we have the following Proposition:

Proposition 1.7.1 *: Let X be a real reflexive space. Then $F : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is weakly sequentially coercive if and only if $\lim_{\|x\| \to +\infty} F(x) = +\infty$*

Proof. ($\Rightarrow$) Let F be weakly sequentially coercive. We prove that $lim_{\|x\| \to +\infty} F(x) = +\infty$ i.e. , $\forall M \in \mathbb{R}, \exists N_0 \in \mathbb{R}$:if $\|x\| > N_0$, then $F(x) > M$.

Assume by contradiction that this is not the case. Then there exists a sequence $\{x_n\}$ such that $\lim_{n \to +\infty} \|x_n\| = +\infty$and $\{F(x_n)\}$ is bounded. Let $\lambda \in \mathbb{R}$ be such that $|F(x_n)| \leq \lambda \forall n \geq 1$.Since F is weakly sequentially coercive, the section

$S_{\lambda,F} \subset X$ is weakly sequentially compact. Hence the sequence $\{x_n\}$ which is in the section $S_{\lambda,F} = \{x \in X : F(x) \leq \lambda\}$, has a subsequence $\{x_{nj}\}$ which converges weakly to an element of $S_{\lambda,F}; F(x_{nj}) \leq \lambda$ whereas $\lim \|x_{nj}\| = +\infty$. This subsequence is therefore bounded(Since every weakly convergent sequence is bounded), which is a contradiction.
($\Leftarrow$)We assume that $\lim\limits_{\|x\|\to+\infty} F(x) = +\infty$
we want to prove that F is weakly sequentially coercive, ie for arbitrary $\lambda \in \mathbb{R}, S_{\lambda,F}$ is weakly sequentially compact, or that any sequence in $S_{\lambda,F}$ has a weakly convergent subsequence. So let $\{x_n\}$ be an arbitrary sequence in $S_{\lambda,F}$ for arbitrary $\lambda \in \mathbb{R}$.
Then $\{x_n\}$ is bounded. Since X is reflexive, $\{x_n\}$ has a subsequence $\{x_{nk}\}$ converges weakly to an element of $\bar{S}_{\lambda,F}$. Thus F is weakly sequentially coercive.

Remark 1.7.1 If X is reflexive and if $\lim\limits_{\|x\|\to+\infty} F(x) = +\infty$, we simply say that F is coercive. This condition actually implies that

$$\forall A > 0, \exists B > 0 : \forall x \in X, \|x\| > B \Rightarrow F(x) > A$$

Equivalently,this means that $F(x) \leq A \Rightarrow \|x\| \leq B$, i.e. ,if the range of F is bounded, then the domain of F is also bounded.

Theorem 1.7.1 *(Existence of minimum in finite dimension)*
Let K be a nonempty closed subset of $\mathbb{R}^n$ and F a continuous real value application on K satisfying the following property:

$$\forall (u^n) \subset K, \lim_{n\to\infty} \|u^n\| = +\infty \Rightarrow \lim_{n\to\infty} F(u^n) = +\infty \tag{1.6}$$

Then there exists a least one minimum point of F over K. Moreover for any minimizing sequence of F over K,one can take a subsequence converging to a minimum point.

Proof.
Let (u^n) be a minimizing sequence of F over K. The condition (1.3) implies that (u^n) is bounded since $(F(u^n))$ is a bounded real sequence. So by Bonzano-weistrass there exists a subsequence (u^{nk}) of(u^n) converging to a point u of $\mathbb{R}^n$. But u is in K since K is closed,and $F(u^{nk})$ converges to $F(u)$ by the continuity of F, thus $F(U) = \inf\limits_{v\in K} F(v)$

Theorem 1.7.2 *(Existence in infinite dimension)*
Let K be a closed,convex and nonempty subset of a reflexive real Banach space V. Let $F : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ convex and lsc.
If K is bounded or if F is coercive, then there exists $\bar{u} \in K$ such that $F(\bar{u}) = \min\limits_{v\in K} F(v)$
Moreover if F is strictly convex, then $\bar{u}$ is unique

Proof. Assume that $F \neq +\infty$, ie a there exists at least $x_0 \in V : F(x_0) \neq +\infty$. In the case where $F = +\infty$, we obviously have $\inf\limits_{v\in K} F(v) = +\infty$
K is nonempty,let (v_n) be a minimizing sequence of F on K. K bounded or

F coercive implies that (v_n) is bounded. Indeed, if K is bounded,then v_n is bounded as a sequence of K, if F is coercive,assume by contradiction that v_n is not bounded,then v_n has a subsequence $v_{nk} : v_{nk} \to +\infty$ as $k \to +\infty$. Since F is coercive,then $\lim\limits_{k\to+\infty} F(v_{nk}) = +\infty$. But since (v_n) is a minimizing sequence of F, we have that $\lim\limits_{k\to+\infty} F(v_{nk}) = \inf\limits_{v\in K} F(v)$ which implies that $\inf\limits_{v\in K} F(v) = +\infty$. Therefore we have $F(x) = +\infty$ which is a contradiction,thus v_n is bounded.
V reflexive implies that there exists (v_{nk}) subsequence of v_n such that v_{nk} converges weakly to $\bar{u}$ in V. K convex and closed implies that K is weakly closed,which implies that $\bar{u} \in K$ So in one hand we have $\lim\limits_{k\to+\infty} F(v_{nk}) = \inf\limits_{v\in K} F(v)$,and in a other hand, we have F convex and lsc implies that $epi(f)$ is convex and closed, which implies that F is weakly lsc.Or v_{nk} converges weakly to $\bar{u}$,so we have $F(\bar{u}) \leq \liminf F(v_{nk}) = \inf\limits_{v\in K} F(v)$. Thus $F(\bar{u}) = \inf\limits_{v\in K} F(v)$. but we have $\bar{u} \in K$,which implies that $F(\bar{u}) = \min\limits_{v\in K} F(v)$ For the uniqueness, assume that F is strictly convex. Let $\bar{u}$ and $\bar{w}$ two minimum of F over K Assume by contradiction that $\bar{u} \neq \bar{w}$. K convex implies that $\frac{1}{2}\bar{u} + \frac{1}{2}\bar{w} \in K$. So we have $\min\limits_{v\in K} F(v) \leq F(\frac{1}{2}\bar{u} + \frac{1}{2}\bar{w})$. Using the strict convexity of F,we have $\min\limits_{v\in K} F(v) < \frac{1}{2}F(\bar{u}) + \frac{1}{2}F(\bar{w})$, which is equivalent that

$$\min_{v\in K} F(v) < \min_{v\in K} F(v)$$

which is a contradiction, therefore $\bar{u} = \bar{w}$, ie the minimum is unique

Corollary 1.7.1 *Let A be an $(n \times n)$ positif defined matrix. Then there exists $\delta > 0$ such that*

$$\langle Ax, x\rangle \geq \delta \|x\|^2 \forall x \in \mathbb{R}^n$$

Proof. Consider the following problem

$$\begin{cases} \min J(x) := \langle Ax, x\rangle \\ \\ \|x\| = 1 \end{cases} \tag{1.7}$$

-J is continuous
-$\dim(\mathbb{R}^n) < \infty$
-Define $K = \{x \in \mathbb{R}^n : \|x\| = 1\}$ is compact
Then by Weistrass theorem there exists $\bar{x} \in K : \bar{x} = arg \min J(x)$.
ie $\exists \bar{x} \in K : J(\bar{x}) \leq J(\frac{x}{\|x\|}) \forall x \neq 0$. Taking $\delta = J(\bar{x}) > 0$,we have
$\frac{1}{\|x\|^2}\langle Ax, x\rangle \geq \delta$,
which implies that $\langle Ax, x\rangle \geq \delta\|x\|^2$

1.8 Optimality condition:

-Convex case
In this section, we will try to get the necessary condition,and sometimes sufficient, for minimizer. Our aim is now in one hand more practical than the last section, since these condition will more often be useful for computing a minimizer.

These conditions will be achieved by using the first derivative(necessary conditions of first order) or second(necessary conditions of second order)

Theorem 1.8.1 *: Let K be a nonempty convex set of norm linear space V. Let $F : V \longrightarrow \mathbb{R}$ derivable (in the sens of Gateaux).*
If $\bar{u}$ is a local minimum of F over K then

$$F'(\bar{u})(v - \bar{u}) \geq 0 \forall v \in K \tag{1.8}$$

Moreover if $\bar{u} \in Int(K)$ then (1.8) become

$$F'(\bar{u}) = 0 \quad EULER\ EQUATION\ ,$$

ie $\bar{u}$ is a critical point.

Proof.
$\bar{u}$ minimum local implies that there exists $r \geq 0$ such that $F(\bar{u}) \leq F(v), \forall v \in K \cap B(\bar{u}, r)$. Let $v \in K, v \neq \bar{u}$, let $t \in]0, 1[$, let $w_t = \bar{u} + t(v - \bar{u}) = (1 - t)\bar{u} + tv \in K$. For w_t to be in $B(\bar{u}, r)$, it suffices to have $\|w_t - \bar{u}\| < r$, ie $0 < t < \dfrac{r}{\|v - \bar{u}\|}$. let $0 < t < min\{\dfrac{r}{\|v - \bar{u}\|}, 1\} = \delta$, for all $t \in]0, \delta[$, we have $w_t \in K \cap B(\bar{u}, r)$, which implies that $F(\bar{u}) \leq F(\bar{u} + t(v - \bar{u}))$; which gives that $\dfrac{F(\bar{u}) - F(\bar{u} + t(v - \bar{u}))}{t} \geq 0 \forall t \in]0, \delta[$. Letting t goes to 0, we get $F'(\bar{u})(v - \bar{u}) \geq 0$
Assume $\bar{u} \in int(K)$, so $\exists r > 0, B(\bar{u}, r) \subset K$. we already know that $\forall u \in K, F'(u)(v - \bar{u}) \geq 0$, let $h \neq 0$. Consider for all $t \in \mathbb{R}, \bar{u} + th$. So for all $t$such that $|t| < \dfrac{r}{\|r\|}$,we have $\bar{u} + th \in K$,which implies that for all t such that $|t| < \dfrac{r}{\|r\|}$, we have $tF'(\bar{u}).h \geq 0$. this give us $tF'(u).(-h) = 0$, which implies that $F'(u).h = 0 \forall h \neq 0$, which means that $F'(u) = 0$
-General case

Definition : Let K be a nonempty subset of a norm linear space V. Let $u \in K$, let $d \in V, d \neq 0$.
One says that d is an **admissible direction** of K on U if it exists $\delta > 0$ such that $u + td \in K$, for all $t \in]o, \delta[$. We denote by $D_{ad}(u)$ the set of admissible direction on the point u.

Theorem 1.8.2 *Consider*

$$\begin{cases} MinF(v) \\ v \in K \neq \emptyset \end{cases} \tag{1.9}$$

If F is differentiable and if F is a local minimum of F over K, then $F'(u).d \geq 0 \forall d \in D_{ad}(u)$
In particular if $u \in int(K)$, then $F'(u) = 0$

Proof.
u minimum local of F implies that there exists$r \geq 0 : F(\bar{u}) \leq F(v)$, for all $v \in K \cap B(u, r)$. Let $d \in D_{ad}(u_)$, it follows that $\delta : u + td \in K, \forall t \in]0, \delta[$.So for

$u+td$ to be on $B(u,r)$,it suffices to have $0 < t < \dfrac{r}{\|d\|}$. Now let $displaystyle\beta = min\{\delta, \frac{r}{\|d\|}\}$, so for all $t \in]0,\beta[$, we have $u+td \in K \cap B(u,r)$, therefore $F(u+td) \geq F(u), \forall t \in]0,\beta[$ implies that $\dfrac{F(u+td)-F(u)}{t} \geq 0, \forall t \in]0,\beta[$, letting t goes to 0, we have $F'(u).d \geq 0$
If $u \in int(K)$,then $D_{ad} = V/\{0\}$ implies that $F'(u).d \geq 0, \forall d \in V/\{0\}$, which implies that $F'(u).d \leq 0, \forall d \in V/\{0\}$, which implies that $F'(u) = 0$

Theorem 1.8.3 *(necessary and sufficient condition of second order)*
One assume that K is nonempty, F is of class C^2.
If $u \in K$ is a local minimum of F on K then:
1) $F'(u).d \geq 0 \forall d \in D_{ad}(U)$
2) *If $d \in D_{ad}(u)$ and $F'(u).d = 0$ then $F''(u).d.d \geq 0$*
3) *If $F'(u).d = 0$ and $F''(u).d.d \geq \alpha\|d\|^2, \forall d \neq 0, \alpha > 0$, then u is strict local minimum of F over K*

Proof.
(1)Let $d \in D_{ad}(u) : F'(u).d = 0$, let $\delta > 0 : \forall t \in]0,\delta[, u+td \in K \cap B(u,r)$. We have by Taylor expansion that:

$$F(u+td) = F(u) + tF'(u).d + \frac{t^2}{2}F''(u).d.d + t^2\|d\|^2\varepsilon(t)$$

where $\varepsilon(t) \to 0$ as $t \to 0$. This implies that $t^2[\frac{1}{2}F''(u).d.d + \|d\|^2\varepsilon(t) \geq 0$. By simplifying the last inequality by t^2 and bu letting t goes to 0 we have:$F'(u).d \geq 0$
(2) Let $v \in V, v \neq u$,we have also by the Taylor expansion of F that

$$F(v) - F(u) = F'(u)(v-u) + \frac{1}{2}.(v-u)(v-u) + \|v-u\|^2\varepsilon(v)$$

where $\varepsilon(v) \to 0$ as $v \to 0$. This implies that $F(v) - F(u) \geq \|v-u\|^2[\dfrac{\alpha}{2} + \varepsilon(v)]$
But $\frac{\alpha}{2} > 0$ and $varepsilon(v) \to 0$ as $v \to 0$ implies that there exists $r > 0$: $\|v-u\| < r$ implies that $\varepsilon(v)| < \frac{\alpha}{2}$. This implies that for $\|v-u\| < r$,we have $-\dfrac{\alpha}{2} < \varepsilon(v) < \dfrac{\alpha}{2}$,which implies that for all $v \in B(u,r)$,we have $\dfrac{\alpha}{2} + \varepsilon(v) > 0$, which give us: for all $v \in B(u,r), v \neq u$, we have $F(v) > F(u)$
Therefore u is a strict local minimum If F is convex, then $F'(u).(v-u) \geq 0$ is a (**Necessary and sufficient condition**)

1.9 Optimization with equality constrains

In this part we are interested with optimization problem with equality constrains in $\mathbb{R}^n$. Given Ω an open set of $\mathbb{R}^n$, F and $F_1, cdots, F_p$ functions defined on Ω taking value on $\mathbb{R}$. One consider the following problem:

$$\inf_{v \in K} F(v) \tag{1.10}$$

with

$$\{v \in \Omega, g_i(v) = 0, 1 \leq i \leq m\} \tag{1.11}$$

The function F is called objective or cost.The function g_i define the equality constrains,the elements of K are called the admissible elements. Obviously the first question we have to ask is the one on the existence of solution of (1.10), for that we use our last results. Indeed they have been obtained in spaces which are more general, they are applicable in particularly in $\mathbb{R}^n$. It is preferable in the following to denote in a more synthetic form the constrains. For that, one pose.

$$G(x) = (g_1(x), \cdots, g_m(x)) \forall x \in \Omega$$

Hence the set of the constrains is written $K = G^{-1}\{0\}$.
Consider the following optimization problem

$$\begin{cases} \min F(x) \\ g_i(x) = 0, i = 1, ...p \\ x \in \mathbb{R}^N (p \leq N) \end{cases} \tag{1.12}$$

where $F, g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$
Set $C = \{x \in \mathbb{R}^N : g_i(x) = 0, \forall i = 1, \cdots, p\}$
So our problem (1.4) become:

$$\begin{cases} \min F(x) \\ x \in C \end{cases} \tag{1.13}$$

Definition A point $\bar{x}$ satisfies the qualification condition (QC) or we can say $\bar{x}$ is regular if $\nabla g_i(\bar{x}), i = 1, ...p$ are linearly independant
If $p = 1, \bar{x}$ is regular if $\nabla g_1(\bar{x}) \neq 0$
$G : \mathbb{R}^N \longrightarrow \mathbb{R}^p$
$x \mapsto G(x) = (\nabla g_1(\bar{x}), \cdots, \nabla g_p(\bar{x}))$
$C = \{x \in \mathbb{R}^N / G(x) = 0\}$
$\bar{x}$ is regular if and only if $JG(\bar{x})$ has rank p

Theorem 1.9.1 *(Lagrange)*
Assume that $\bar{x} \in C$ is regular. If $\bar{x}$ is an extremum then:

$$\exists \lambda_1, \cdots, \lambda_p \in \mathbb{R} :$$

$$\begin{cases} \nabla F(x) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = 0 \\ g_i(\bar{x}) = 0, i = 1, \cdots, p \end{cases} \tag{1.14}$$

The real $\lambda_1, \cdots, \lambda_p$ are called the **Lagrange multipliers** associated with the constrains of the problem and the minimum local point. The uniqueness of these reals is assured by the qualification condition(QC)

Proposition 1.9.1 *Let Ω be an open convex subset of $\mathbb{R}^n$. Let $F : \Omega \to \mathbb{R}$ be a convex differentiable function on Ω and $g_i : \Omega \to \mathbb{R}$ be affine. Let $\bar{x} \in K$ such that there exists $\lambda_1, \lambda_2, \cdots, \lambda_p \in \mathbb{R}$ such that*

$$\nabla F(x) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = 0. \tag{1.15}$$

Then, $\bar{x}$ is a minimizer of F on K.

Proof. Let $x \in K$,by convexity of F we have:

$$F(x) - F(\bar{x}) \geq \nabla F(\bar{x})(x - \bar{x}) \tag{1.16}$$

Using (1.15) it follows that:

$$F(x) - F(\bar{x}) \geq - \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x})(x - \bar{x}) \tag{1.17}$$

Since g_i are affine and $g_i(x) = g_i(\bar{x}) = 0$,we have also the identities

$$g_i(x) - g_i(\bar{x}) = \nabla g_i(\bar{x}).(x - \bar{x}) = 0$$

By replacing in (1.17),we obtain $F(\bar{x}) \leq F(x)$

CHAPTER 2

Pontryagin minimum method principle

Our aim in this work is to present the goal of optimal control, more precisely the protryagin method. It focus in the study of dynamical system of optimization. Roughly speaking this is best to control a system(mechanic, chemestre , physic, economic...) to minimize a cost function. To formulate this problem mathematically, it is obviously necessary first to define the variable that define the stat of the system, the model that describes the evolution of these variables and the actions that can be exercise on the system(ie the control or command). Before able to specify the jargon,it is unnecessary to emphasize that we are all doing optimal control without paying any attention(and more or less): for example : go from a place to an other as quickly as possible,or by the shortest way,maximizing the return on investment or placing or minimizing debt are problem of this type. More precisely,we are interested in a system(a rocket,an airplane,a robot,a portfolio of stocks,bonds or obtions etc) whose state x(t) (at the time t) is Given the solution of the ordinary differential equation:

$$\frac{dx}{dt} = f(t, x, u) \text{ on }]t_0, t_1[, \;\; x(t_0) = x_0,$$

where x_0 is in $\mathbb{R}^n$ and t_0,t_1 in $\mathbb{R}$ are fixed with $t_0 < t_1$;f is a function given from $\mathbb{R} \times \mathbb{R}^n \times U$ into $\mathbb{R}^n$. As we often see it,one of our arguments of f is a function u defined on the interval $]t_0$,$t_1[$ taking value in the given set U.We always assume,to simplify,that U is a closed,nonempty set of $\mathbb{R}^m$. This function u translates mathematically the actions (or decisions) that we have on the evolution of system,the set U corresponds to any restrictions or constrains that must respect the controls. (for example limited resources,limits on the acceleration or speed for the driving). Formulate an optimal control problem is to define the state of the system and the differential equation governing its evolution, the class admissible controls u(t) and finally a criterion of evolution or a cost function f, most often it is a cumulative in time added to a final cost, whose typical form is:

$$\int_{t_0}^{t_1} g(t, x(t), u(t))dt + h(x(t_1)) \tag{2.1}$$

where g and h are given function define respectively on $\mathbb{R} \times \mathbb{R}^n \times U and \mathbb{R}^n$ in to $\mathbb{R}$.

The problem to solve is now to determine the optimal cost and optimal control, ie to solve the following optimization problem:

$$\begin{cases} \inf_{u\in L^2(o,T,U)} \left(\int_{t_0}^{t_1} g(t,x(t),u(t))dt + h(x(t_1)) \right) \\ \frac{dx}{dt} = f(t,x,u) \text{ on }]t_0,t_1[, \; x(t_0) = x_0, \end{cases} \tag{2.2}$$

During the following, a control will be a measurable function u defined on a interval $]t_0,t_1[$ taking value in U, with $u \in L^2(]t_0,t_1[,U)$
More over one always assume,to make sens to (2.1),(2.2),that the functions f, g and h satisfy the following hypothesis:

$$\begin{cases} f \text{ and } g \text{ are continuous with respect to } x \text{ in } \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \\ \text{and the derivatives } \frac{\partial f}{\partial x}, \frac{\partial g}{\partial x} \text{ are continuous on } \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \end{cases} \tag{2.3}$$

$h \in C^1(\mathbb{R}^n)$,

$$g, \text{ and } h \text{ are respectively bonded below on } \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \text{ and } \mathbb{R}^n; \tag{2.4}$$

$$\|f(t,x,u)\| \le C(1+u) + \|x\| + \|u\|), \forall (t,x,u) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \tag{2.5}$$

$$\|\frac{\partial f}{\partial x}(t,x,u)\| + \|\frac{\partial g}{\partial x}(t,x,u)\| \le C_R(1+\|u\|), \text{ for } \|x\| \le R, |t| \le R, u \in U \tag{2.6}$$

$$|g(t,x,u) \le C_R(1+\|u\|^2), for\|x\| \le R, |t| \le R, u \in U \tag{2.7}$$

In these hypothesis, C is a positive constant independant from x, t and u when C_R is a positive real, which depend only on the real R(arbitrary on $\mathbb{R}_+^*$)
The hypothesis make on f will alow us to be sure that there exists a unique solution x of the differential system (2.1). To be sure, if $u(t)$ is fixed,the function $\bar{f}(t,x) = f(t,x,u)$ satisfy:

$$\|\frac{\partial \bar{f}}{\partial x}(t,x)\| \le C_R(1+\|u(t)\|$$

,
for $\|x\| \le R$ and for all $t \in]-\mathbb{R},\mathbb{R}[$,hence that
$\|\bar{f}(t,x)\| \le C(1+\|x\|+\|u(t)\|)$
for $x \in \mathbb{R}^n$ and for almost every $t \in]t_0,t_1[$. According to the general results for the ordinary differential equations, these boundednes is enough when ever
$u \in L^2(]t_o,t_1[;\mathbb{R}^m)$ to assure the existence of a unique solution $x \in C([t_0,t_1],\mathbb{R}^n)$ of (2.1), means satisfying:

$$x(t) = x_0 \int_{t_0}^{t} f(s,x(s),u(s))\,ds$$

, for all $t \in]t_0,t_1[$.More over $x(t)$ depend continual on x_0. In particular,the control u has been fixed, we have: $\|x(t)\| \le R$ for all $t \in [t_0,t_1]$ for a constant $R > 0$ (depending on u). Because of the hypothesis (2.7) made on g, this allow the criterion J given by (2.1) to be well defined. Finaly the hypothesis (2.4) made on g and h assure that the infinimum (2.1) is finite ($\inf J > -\infty$)

2.0.1 Towards the principle of pontryagin

Definition We define the Hamiltonien for the problem (2.2),the function denoted by H and defined from $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times U$ into $\mathbb{R}$ by:
$H(t, x, p, u) := g(t, x, u) + < p, f(t, x, u) >$ for all (t, x, p, u) on $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times U$

Definition Let $u(t)$ be a control and $x(t)$ be the state associated ie

$$\frac{dx}{dt} = f(t, x, u), x(t_0) = x_0,$$

We define the adjoint state associated to $u(t), x(t)$, the unique solution denoted $p(t)$ of the following system

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x}(t, x(t), p(t), u(t)) \text{ on }]t_0, t_1[\quad , p(t_1) = h'(x(t_1)) \tag{2.8}$$

Theorem 2.0.2 *One assume that the hypothesis* (2.3) *and* (2.5) *are satisfied. Let* $\bar{u}$*:*$]t_0,t_1[\longrightarrow \mathbb{R}^n$ *solution of* (2.1). *Then for all t in* $]t_0,t_1[$

$$\bar{u}(t) \text{ realize the minimum over } U \text{ of the function } u \longmapsto H(t, \bar{x}(t), \bar{p}(t), u) \tag{2.9}$$

where $\bar{x}$ *and* $\bar{p}$ *are the state and the adjoint state associated to* $\bar{u}(t)$.

To fixe the idea, let us consider a concrete example.

One consider a car running on a road.Let us denote by $x_1(t)$:the position of the car at time t; by $x_2(t)$: the speed of the car at time t; and by $u(t)$:the acceleration of the car at time t. So we have

$$\begin{cases} \dfrac{dx_1}{dt} = x_2(t) \\ \\ \dfrac{dx_2}{dt} = u(t) \end{cases} \tag{2.10}$$

which give us:$\frac{dx}{dt} = f(t, x(t), u(t))$
where $f(t, x, u) = (x_2, u)$ and $x = (x_1, x_2)$.
We know that the consumption of the fuel by the car depend on the acceleration, so our cost function J is given by the consumption of the fuel, and our control is the acceleration; $x_1(0) = x_0^1$ and $x_2(0) = x_0^2$. Our pontryagin problem is so

given by:

$$\begin{cases} \inf_{u} J(u) := \text{ consumption of fuel} \\ \\ \dfrac{dx}{dt} = f(t, x(t), u(t)) \\ \\ x(o) = (x_0^1, x_0^2) \end{cases} \tag{2.11}$$

Therefore our problem (2.11) minimize the consumption of the fuel of the car by using the acceleration

CHAPTER 3

Minimum Principle of Pontryagin: Linear Quadratic Case

Consider the following quadratic problem given by the cost function:

$$J(u) = \frac{1}{2}\int_0^T < Q(x(t) - x^0(t)), x(t) - x^0(t) > dt + \frac{1}{2}\int_0^T (Ru(t), u(t))dt$$
$$+\frac{1}{2} < D(x(T) - x^0(T)), x(T) - x^0(T) > \tag{3.1}$$

and with the constrains given by the following ODE:

$$\dot{x}(t) = Ax(t) + Bu(t) + f(t) \text{ for } 0 < t < T,\ x(0) = x_0 \text{ in } \mathbb{R}^n \tag{3.2}$$

where A is a $n\times n$matrix, B a $n\times m$ matrix, Q and D two $n\times n$ symmetric matric semi positive define, R a $m \times m$ matrix symmetric positive define; T>0,x_0 in $\mathbb{R}^n$,f in $L^1([0,T];\mathbb{R}^n)$ and x^0 in $C([0,T];\mathbb{R}^n)$ are fixed.Finaly, the set of controls u is in $L^2(]0,T[;U)$ and one assume that U is a nonempty subset convex and closed of $\mathbb{R}^m$
So we have the typical problem:

$$\begin{cases} \inf_{u\in L^2(0,T,U)} J(u) := \int_{t_o}^{t_1} g(t,x(t),u(t))dt + h(x(t_1)) \\ \dfrac{dx}{dt} = F(t,x(t),u(t)) \\ x(t_0) = x_0 \end{cases} \tag{3.3}$$

Where $g : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \longrightarrow \mathbb{R}$ is given by

$$g(t,x,u) = \frac{1}{2}(\langle Q(x(t) - x^0(t)), (x(t) - x^0(t)\rangle + \langle Ru(t), u(t)\rangle)$$

and $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ is given by $h(x) = \frac{1}{2}\langle D(x - x^0), x - x^0\rangle$ and
$F : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \longrightarrow \mathbb{R}^n$ is given by $F(t,x,u) = Ax + Bu + f(t)$. We easily get

the Hamiltonian as the function $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ defined by

$$\begin{aligned} H(t,x,p,u) &= \frac{1}{2}(\langle Q(x(t)-x^0(t)),(x(t)-x^0(t)\rangle + \langle Ru(t),u(t)\rangle) \\ &\quad + \langle p, Ax+Bu+f(t)\rangle. \end{aligned}$$

3.1 Existence and uniqueness of the optimal control

Lemma 3.1.1 *There exists a bilinear form symmetric,continuous and coercive on* $L^2([0,T];\mathbb{R}^m)$,*a, a linear and continuous form* l *on* $L^2([0,T];\mathbb{R}^m)$ *and a constant* C *such that*

$J(u)=a(u,u)+l(u)+C$

To prove this, we need the following lemma

Lemma 3.1.2 *: One consider the system*

$$\begin{cases} \dot{x} = Ax + Bu + f, t \ \ in \ \ [0,1] \\ \\ x(0) = X_0 \end{cases} \tag{3.4}$$

Then the map $u \longmapsto x(u)$,*where x(u) is the unique solution of* (3.4), *is affine,continuous on* $L^2([0,T];\mathbb{R}^m)$ *into* $c([0,T];\mathbb{R}^n)$

Proof.
Let $\bar{y}$ be the solution of (3.4) corresponding to u=0 and $x_h(u)$ be the solution of the homogeneous ODE:

$$\begin{cases} \dot{x} = Ax + Bu, t \ \text{in} \ \ [0,T] \\ \\ x(0) = 0 \end{cases} \tag{3.5}$$

Then $x(u) = x_h(u) + \bar{y}$.
Indeed,we have

$$\frac{d\bar{y}}{dt} = A\bar{y} + f, \bar{y}(0) = X_0 \ \text{ and } \ \frac{dx_h(u)}{dt} = Ax_h(u) + BU, x_h(u) = 0$$

This implies that $\frac{dx_h(u)}{dt} + \frac{d\bar{y}}{dt} = Ax_h(u) + A\bar{y} + Bu + f, \bar{y}(0) = 0, x_h(u)(0) = 0$

which implies that $\frac{d(x_h(u)+\bar{y})}{dt} = A(x_h(u) + \bar{y}) + Bu + f;\ (x_h(u) + \bar{y})(0) = X_0$

thus $x_h(u) + \bar{y}$ solve (3.4)

therefore by the uniqueness of the solution of (3.4), we have $x(u) = x_h(u) + \bar{y}$.

Claim:$u \longmapsto x_h(u)$ is linear and continuous
Ended,
Let u_1, u_2 in $L^2(o,T;\mathbb{R}^m)$

We have that

$$\begin{cases} \dfrac{dx_h(u_1)}{dt} = Ax_h(u_1) + Bu_1 \\ x_h(u_1)(0) = 0 \end{cases} \tag{3.6}$$

and

$$\begin{cases} \dfrac{dx_h(u_2)}{dt} = Ax_h(u_2) + Bu_2 \\ x_h(u_2)(0) = 0 \end{cases} \tag{3.7}$$

So (3.6)+(3.7) implies that

$$\begin{cases} \dfrac{d(x_h(u_1) + x_h(u_2))}{dt} = A(x_h(u_1) + x_h(u_2)) + B(u_1 + u_2) \\ (x_h(u_1) + x_h(u_2))(0) = 0 \end{cases} \tag{3.8}$$

which is equivalent to say that $x_h(u_1) + x_h(u_2)$ solve the ODE:

$$\begin{cases} \dfrac{dx}{dt} = Ax + B(u_1 + u_2) \\ x(o) = 0 \end{cases} \tag{3.9}$$

since $x_h(u_1 + u_2)$ is the unique solution corresponding to the ODE (3.9), then we have:
$x_h(u_1 + u_2) = x_h(u_1) + x_h(u_2)$
Let u in $L^2(0, T; \mathbb{R}^m)$ and α in $\mathbb{R}$
then

$$\begin{cases} \dfrac{dx_h(u)}{dt} = Ax_h(u) + Bu \\ x_h(u)(0) = 0 \end{cases} \tag{3.10}$$

which implies

$$\begin{cases} \alpha\dfrac{dx_h(u)}{dt} = \alpha Ax_h(u) + \alpha Bu \\ \alpha x_h(u)(0) = 0 \end{cases} \tag{3.11}$$

which is equivalent

$$\begin{cases} \dfrac{d(\alpha x_h(u))}{dt} = A\alpha x_h(u) + B\alpha u \\ \alpha x_h(u)(0) = 0 \end{cases} \tag{3.12}$$

which means that $\alpha x_h(u)$ solve the ODE

$$\begin{cases} \dfrac{dx}{dt} = Ax + B(\alpha u) \\ \\ x(0) = 0 \end{cases} \tag{3.13}$$

Since $x_h(\alpha u)$ is the unique solution corresponding to the Ode (3.13) then
$x_h(\alpha u) = \alpha x_h(u)$
therefore $u \longmapsto x_h(u)$ is linear. Since $x(u) = x_h(u) + \bar{y}$, then it follows that $u \longmapsto x(u)$ is affine

To show the continuity of $u \longmapsto x(u)$,it is enough to show the one of the map $u \longmapsto x_h(u)$. Since $x_h(u)$ is the solution corresponding to the homogeneous ODE (3.5), then it is defined from $L^2(0,T;\mathbb{R}^m)$ into $C([0,T];\mathbb{R}^m)$ by:

$$x_h(u)(t) = \int_0^t e^{A(t-s)} Bu(s)\, ds \tag{3.14}$$

which implies that

$$\|x_h(u)(t)\|_{\mathbb{R}^n} \leq \int_0^t e^{\|A(t-s)\|} \|Bu(s)\|_{\mathbb{R}^m}\, ds \tag{3.15}$$

which is equivalent that $\|x_h(u)(t)\|_{\mathbb{R}^n} \leq e^{\|AT\|}\|B\| \int_0^T \|u(s)\|_{\mathbb{R}^m}\, ds$
and by using the Cauchy-Swatch inequality, we have

$$\|x_h(u)(t)\|_{\mathbb{R}^n} \leq e^{\|AT\|}\|B\|(\int_0^T \|u(s)\|^2_{\mathbb{R}^m}\, ds)^1|2\sqrt(T)$$

Which is equivalent that $\|x_h(u)(t)\|_{\mathbb{R}^n} \leq C\|u\|_{L^2(0,T;\mathbb{R}^m)}$;
where $C = e^{\|AT\|}\|B\|\sqrt(T)$.
By taking suppremum over [0,T], we get
$\|x_h(u)(t)\|_C([0,T] : \mathbb{R}^n) \leq C\|u\|_{L^2(0,T;\mathbb{R}^m)}$
Since x_h is linear, therefore it is continuous
Proof. (Lemma3.1.1)

By applying **Lemma3.1.2**,we get

$$\begin{aligned} J(u) \quad = \quad & 1/2\int_0^T < Q((x_h(u)(t) + \bar{y} - x^0(t))), x_h(u)(t) + \bar{y} - x^0(t)\, dt \\ & +1/2\int_0^T < Ru(t), u(t) > \, dt + 1/2 < D(x_h(T) \\ & +\bar{y}(T) - x^(0), x_h(T) + \bar{y}(T) - x^0(T) > . \end{aligned}$$

So we obtain

$$\begin{aligned}
J(u) &= 1/2\int_0^T < Qx_h(u), x_h(u) > \, dt + 1/2\int_0^T < Ru(t), u(t) > \, dt \\
&\quad +1/2 < Dx_(u)(T), x_h(u)(T) > + \int_0^T < Q(\bar{y} - x^0), x_h(u)) > \, dt \\
&\quad + < D(\bar{y}(T) - x^0(T)), x_h(u)(T) > \\
&\quad +1/2\int_0^T < Q(\bar{y} - x^0), \bar{y} - x^0 > \, dt \\
&\quad +1/2 < D(\bar{y}(T) - x^0(T), \bar{y}(T) - x^0(T) > \\
&= a(u,u) + l(u) + c
\end{aligned}$$

where

$$\begin{aligned}
a(u,v) &= 1/2\int_0^T < Qx_h(u), x_h(v) > \, dt + 1/2\int_0^T < Ru(t), v(t) > \, dt \\
&\quad +1/2 < Dx_(u)(T), x_h(v)(T) >,
\end{aligned}$$

$$l(v) = \int_0^T < Q(\bar{y} - x^0), x_h(v) > \, dt + < D(\bar{y}(T) - x^0(T)), x_h(v)(T) >$$

and

$$C = \frac{1}{2}\int_0^T < Q(\bar{y} - x^0), \bar{y} - x^0 > \, dt + \frac{1}{2} < D(\bar{y}(T) - x^0(T)), \bar{y}(T) - x^0(T) >$$

Claim: a is bilinear symmetric and continuous and l is linear and continuous.

In fact let u, v_1, v_2 in $L^2(0,T;\mathbb{R}^m)$ $\alpha \in \mathbb{R}$, we have

$$\begin{aligned}
a(u, v_1 + \alpha v_2) &= 1/2\int_0^T < Qx_h(u), x_h(v_1 + \alpha v_2) > \, dt \\
&\quad +1/2\int_0^T < Ru(t), v_1 + \alpha v_2 > \, dt \\
&\quad +1/2 < Dx_(u)(T), x_h(v_1 + \alpha v_2)(T) > \\
&= 1/2\int_0^T < Qx_h(u), x_h(v_1) > \, dt \\
&\quad +1/2\alpha\int_0^T < Qx_h(u), x_h(v_2) > \, dt \\
&\quad +1/2\int_0^T < Ru(t), v_1) \, dt + \alpha 1/2\int_0^T < Ru(t), v_2) \, dt \\
&\quad +1/2 < Dx_(u)(T), x_h(v_1) \, dt + \alpha 1/2 < Dx_(u)(T), x_h(v_2) \, dt \\
&= a(u, v_1) + \alpha a(u, v_2)
\end{aligned}$$

By using the same argument, $a(v_1 + \alpha v_2, u) = a(v_1, u) + \alpha a(v_2, u)$. Using the continuity of x_h, it follows that a is continuous.
The same arguments give the linearity and continuity of l.
j is continuous since $a(,)$ and l are continuous

Q and D are symmetric semi positive defined then $< Qx_h(u), x_h(u) >\geq 0$ and $\langle Dx_h(u)(T), x_h(u)(T)\rangle \geq 0$, which implies that $\int_0^T \langle Qx_h(u), x_h(u)\rangle \geq 0$ and $\int_0^T \langle Dx_h(u)(T), x_h(u)(T)\rangle \geq 0$. since R is positive defined then $\langle Ru(t), u(t)\rangle \geq \alpha\|u(t)\|^2$, so this implies that $a(u,u) \geq \frac{\alpha}{2}\int_0^T \|u(t)\|^2\, dt$, which is equivalent to say that $a(u,u) \geq \beta\|u\|^2_{L^2(0,T;\mathbb{R}^m)}$.
l is linear and continuous so $|l(u)| \leq k\|u\|_{L^2(0,T;\mathbb{R}^m)}$.But $l(u) \leq |l(u)| \leq k\|u\|_{L^2(0,T;\mathbb{R}^m)}$, which implies that $-l(u) \geq -k\|u\|_{L^2(0,T;\mathbb{R}^m)}$.
Thus $J(u) \geq \beta\|u\|^2_{L^2(0,T;\mathbb{R}^m)} - k\|u\|_{L^2(0,T;\mathbb{R}^m)}$, so $\lim_{\|u\|\to\infty} J(u) = +\infty$. Therefore J is coercive.
Since R is positive definite and symmetric,then a(,) is strictly convex. Ended, $u \longmapsto (Ru,u)$ is strictly convex, to prove this,let define $S(u) = (Ru,u)$, then for all u, v in $L^2(0,T;\mathbb{R}^m)$, for all t in]0,1[, we have:
$S(tu+(1-t)v)-tS(u)-(1-t)S(v) = (R(tu+(1-t)v, tu+(1-t)v)-t(Ru,u)-(1-t)(Rv,v)$
$= (Rtu, tu+(1-t)v) + (R(1-t)v, tu+(1-t)v) - t(Ru,u) - (1-t)(Rv,v)$
$= t^2(Ru,u) + t(1-t)(Ru,v) + (1-t)t(Rv,u) + (1-t)^2(Rv,v)$
$- t(Ru,u) - (1-t)(Rv,v)$
$= (t^2-t)(Ru,u) + t(1-t)[(Ru,v) + (Rv,u) - (R(u-v), u-v)]$
$+ ((1-t)^2 - (1-t))(Rv,v)$
$= -t(1-t)(R(u-v), u-v) < 0$
we got this last inequality by using the fact that R is positive define therefore by using **Corollary 1.4.2** there exist α strictly positive such that $(Rv,v) \geq \alpha\|v\|^2$ for all v in $L^2(0,T;\mathbb{R}^m)$
So $S(tu+(1-t)v) < tS(u) + (1-t)S(v)$. Therefore S is strictly convex,ie $u \longmapsto (Ru,u)$ is strictly convex.
For the same argument,$u \longmapsto < Qx_h(u), x_h(u) >$ and $u \longmapsto < Dx_h(u)(T), x_h(u)(T) >$ are convex by the fact that Q and D are positive define. Since Integral is linear, therefore J is strictly convex
U is convex,closed so $L^2(0,T;U)$ is also convex and closed. Then, by using the existence theorem we have that J has a unique minimum on $L^2(0,T;U)$ called $\bar{u}$

3.2 characterization of the optimal control

Lemma 3.2.1 *J is Gateau differentiable on $L^2(0,T;\mathbb{R}^m)$, and more over we have:*

$$\begin{aligned}(J'(u), w) = &\int_0^T \langle Q(x(u)-x^0), x_h(w)\rangle\, dt + \int_0^T (Ru(t), w(t))\, dt \\ &+ < D(x(u)(T) - x^0(T)), x_h(w)(T) >\end{aligned} \tag{3.16}$$

Proof.
a is a continuous bilinear form,so it is Frechet differentiable on $L^2(0,T;\mathbb{R}^m)$, and therefore Gateau differentiable on $L^2(0,T;\mathbb{R}^m)$.
l is a linear and continuous map on $L^2(0,T;\mathbb{R}^m)$ so it is Gateau differentiable

on $L^2(0,T;\mathbb{R}^m)$.
Thus $u \longmapsto J(u) = a(u,u) + l(u) + C$ is Gateau differentiable on $L^2(0,T;\mathbb{R}^m)$ and for each w in $L^2(0,T;\mathbb{R}^m)$, we have:

$$(J'(u),w) = da(u,u).(1,1).w + dl(u).w$$
$$= da(u,u).(w,w) + dl(u).w$$

Since a is bilinear symmetric, then $da(u,u).(w,w) = 2a(u,w)$. And l is linear so $dl(u).w = l(w)$,
which implies that $(J'(u),w) = 2a(u,w) + l(w)$

$$\begin{aligned} &= \int_0^T \langle Qx_h(u), x_h(w)\rangle\, dt + \int_o^T \langle Ru(t), u(t)\rangle\, dt \\ &\quad + \langle Dx_h(u)(T), x_h(w)(T)\rangle + \int_0^T \langle Q(\bar{y} - x^0), x_h(w)\rangle\, dt \\ &\quad + \langle D(\bar{y}(T) - x^0(T)), x_h(w)(T)\rangle \\ &= \int_0^T \langle Q(x_h(u) + \bar{y} - x^0), x_h(w)\rangle\, dt + \int_0^T \langle Ru(t), u(t)\rangle\, dt \\ &\quad + \langle D(x_h(U)(T) + \bar{y}(T) - x^0(T)), x_h(w)(T)\rangle. \end{aligned}$$

But since $x(u) = x_h(u) + \bar{y}$, then we have

$$\begin{aligned} (J'(u),w) &= \int_0^T \langle Q(x(u) - x^0), x_h(w)\rangle\, dt + \int_0^T \langle Ru(t), u(t)\rangle\, dt \\ &\quad + \langle D(x(u)(T) - x^0(T)), x_h(w)(T)\rangle. \end{aligned}$$

Which is (3.19).
Let us simplify $(J'(u),w)$ by using the adjoint state

Lemma 3.2.2 *:*

$$(J'(\bar{u}),w) = \int_0^T (R\bar{u}(t) + B^t\bar{p}(t), w(t)) \tag{3.17}$$

Proof.
We know that:

$$\begin{aligned} (J'(u),w) = &\int_0^T < Q(x(u) - x^0), x_h(w) > \, dt + \int_0^T (Ru(t), w(t))\, dt \\ + &< D(x(u)(T) - x^0(T)), x_h(w)(T) > \end{aligned} \tag{3.18}$$

One denote by $\bar{x}$, the state corresponding to the optimal control $\bar{u}$, and by $\bar{p}$ the adjoint state associated.
So we have:

$$\begin{aligned} (J'(\bar{u}),w) = &\int_0^T < Q(\bar{x} - x^0), x_h(w) > \, dt + \int_0^T (R\bar{u}(t), w(t))\, dt \\ + &< D\bar{x}(T) - x^0(T)), x_h(w)(T) > \end{aligned} \tag{3.19}$$

But $\bar{p}$ is solution of the ODE:

$$\begin{cases} \dfrac{dp}{dt} = -\dfrac{\partial H}{\partial x}(t, x, p, u) \\ p(T) = h'(x(T)) \end{cases} \tag{3.20}$$

ie

$$\dot{\bar{p}} = -A^t\bar{p} - Q(\bar{x} - x^0), \quad 0 < t < T, \quad \bar{p}(T) = D(\bar{x}(T) - x^0(T)) \tag{3.21}$$

So by using (3.21) we have

$(J'(\bar{u}), w) - \int_0^T \langle R\bar{u}, w\rangle \, dt = \int_0^T \langle -A^t\bar{p} - \dot{\bar{p}}, x_h(w)\rangle \, dt + \langle p(T), x_h(w)(T)\rangle$

$= \int_0^T \langle -A^t\bar{p}, x_h(w)\rangle \, dt - \int_0^T \langle \dot{\bar{p}}, x_h(w)\rangle \, dt + \langle p(T), x_h(w)(T)\rangle$

Integrating $\int_0^T \langle \dot{\bar{p}}, x_h(w)\rangle \, dt$ by part, we get

$(J'(\bar{u}), w) - \int_0^T \langle R\bar{u}, w\rangle \, dt = -\int_0^T \langle A^t\bar{p}, x_h(w)\rangle \, dt + \int_0^T \langle \bar{p}(T), \frac{dx_h(w)}{dt}\rangle \, dt - [\langle \bar{p}, x_h(w)\rangle]_0^T + \langle \bar{p}(T), x_h(w)\rangle$

Since $x_h(w)$ is solution of the ODE : $\frac{dx}{dt} = Ax + Bw; x(0) = 0$, then we have $\frac{dx_h(w)}{dt} = Ax_h(w) + Bw$

which implies that: $(J'(\bar{u}), w) - \int_0^T \langle R\bar{u}, w\rangle \, dt = -\int_0^T \langle A^t\bar{p}, x_h(w)\rangle \, dt$

$+ \int_0^T \langle \bar{p}(T), Ax_h(w) + Bw\rangle \, dt - \langle \bar{p}(T), x_h(w)(T)\rangle + \langle \bar{p}(T), x_h(w)(T)\rangle$

$= -\int_0^T \langle A^t\bar{p}, x_h(w)\rangle \, dt + \int_0^T \langle A^t\bar{p}, x_h(w)\langle \, dt + \int_0^T \langle B^t\bar{p}, w\rangle dt$

which implies that $(J'(\bar{u}), w) = \int_0^T \langle B^t\bar{p}, w\rangle dt + \int_0^T \angle R\bar{u}, w\rangle \, dt$

$$= \int_0^T \langle Ru + B^t\bar{p}, w\rangle \, dt$$

Therefore $(J'(u), w) = \int_0^T \langle Ru + B^t\bar{p}, w\rangle \, dt$

So by the optimality condition,we have that:the unique optimal control $\bar{u}$ is characterized by the condition: $\bar{u}$ in $L^2(0, T; U)$ and $(J'(u), u - \bar{u}) \geq 0$ for all u in $L^2(0, T; U)$, i.e.

$$\bar{u} \in L^2(0, T; U) \text{ and } \int_0^T \langle Ru + B^t\bar{p}, u(t) - \bar{u}(t)\rangle \, dt \geq 0 \forall u \in L^2(0, T; U) \tag{3.22}$$

Lemma 3.2.3 *The variational inequality (3.22) is equivalent to say that: $\bar{u}$ is the unique point minimum on $L^2(0, T; U)$ of*

$\bar{J}(u) = 1/2 \int_0^T \langle Ru(t), u(t)\rangle \, dt + \int_0^T \langle B^t\bar{p}(t), u(t)\rangle \, dt$

Proof. Indeed,
$\Leftarrow$)assume that $\bar{u}$ satisfies (3.22), we know that $\bar{J}$ is strictly convex,since $u \mapsto (Ru, u)$ is strictly convex, by the fact that R is positive define(proved on the Lemma3.1.1), and $u \mapsto \int_0^T \langle B^t\bar{p}(t), u(t)\rangle\, dt$ is linear so convex.
By the convex inequality we have:
$\bar{J}(u) \geq \bar{J}(\bar{u}) + \bar{J}'(\bar{u})(u - \bar{u})$
but $\bar{j}'(\bar{u}).(u - \bar{u}) = 1/2db(\bar{u}, \bar{u}).(u - \bar{u}) + l(\bar{u}).(u - \bar{u})$
where $b(u, u) = \int_0^T \langle Ru, u\rangle\, dt$ is bilinear and $l(u) = \int_0^T \langle B^t\bar{p}(t), u(t)\rangle\, dt$ is linear so $db(\bar{u}, \bar{u}) = 2b(\bar{u}, u - \bar{u})$ and
$dl(\bar{u}).(u - \bar{u}) = l(u - \bar{u})$
which implies $\bar{J}'(u)(u - \bar{u}) = \int_0^T \langle Ru + B^t\bar{p}, u - \bar{u}\rangle\, dt$, and so by assumption $\bar{J}'(u)(u - \bar{u}) \geq 0$
which implies that $\bar{J}(u) \geq \bar{J}(\bar{u})$ for all u in $L^2(0, T; U)$. Therefore $\bar{u}$ is a minimizer of $\bar{j}$. Since $\bar{J}$ is strictly convex, so it has a unique minimizer. Therefore $\bar{u}$ is the unique minimizer.
$\Rightarrow$) conversely
assume that $\bar{u}$ is a unique minimizer of $\bar{j}$. Since $\bar{J}$ is convex on $L^2(0, T; U)$ and $L^2(0, T; U)$ is also convex and closed,then by **Theorem 1.4.5** $\bar{J}'(u)(u - \bar{u}) \geq 0$, which is equivalent that $\int_0^T < Ru + B^t\bar{p}, u - \bar{u} > dt \geq 0$

Lemma 3.2.4 *$\bar{u}$ is the minimizer over U of the function $v \mapsto 1/2(Rv, v) + (B^t\bar{p}(t), v)$*

Proof. Let $\bar{v}(t)$, for each t in $[0, T]$, be the minimum over U of the function $v \mapsto 1/2(Rv, v) + (B^t\bar{p}(t), v)$, which is convex, continuous and coercive on the convex closed and nonempty set U. Ended, for the continuity, since $\bar{v}(t)$ is the minimizer over U of the function $v \mapsto 1/2(Rv, v) + (B^t\bar{p}(t), v)$ for each $t \in [0, T]$, we have the derivative of this function apply at $\bar{v}(t)$ is equal to zero, ie

$$(1/2(Rv,) + (B^t\bar{p}(t),))'(\bar{v}(t)) = o,$$

which is equivalent to

$$R\bar{v}(t) + B^t\bar{p}(t) = 0$$

which is equivalent to $\bar{v}(t) = R^{-1}B^tp(t)$
ie $\bar{v} = R^{-1}B^t\bar{p}$.Since $\bar{p}$ is is the adjoint state associated to the optimal control $\bar{u}$, so solution of the ODE

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x}(t, x, p, u)$$

therefore the function $\bar{v}$ is continuous. Moreover as we saw it above using the fact that R is symmetric positive define we have that $\bar{v}$ is convex and coercive, which prove that $\bar{v}$ is convex, continuous and coercive.
Since $\bar{v}$ is the minimizer of $v \mapsto 1/2(Rv, v) + (B^t\bar{p}(t), v)$, we have:
$1/2(R\bar{v}, \bar{v}) + (B^t\bar{p}(t), \bar{v}) \leq 1/2(R\bar{u}, \bar{u}) + (B^t\bar{p}(t), \bar{u})$
Using the fact that $\bar{v}$ is continuous, we get that our above function is integrable, and by integrating the last inequality we get: $\bar{J}(\bar{v}) \leq \bar{J}(\bar{u})$. So since $\bar{u}$ realize the minimum over U of $\bar{J}$, then we have $\bar{u} = \bar{v}$. i.e $\bar{u}$ is the minimizer over U of the function $v \mapsto 1/2(Rv, v) + (B^t\bar{p}(t), v)$

3.2.1 Riccati equation

The goal of this section is to show that in the case without constrains, i.e $U = \mathbb{R}^m$, the optimal control and the adjoint state depend linearly on the optimal trajectory.

Theorem 3.2.1 *One assume that the hypothesis made on the theorem* **Theorem 2.0.2** *are satisfied. One assume also that* $f = 0, , x_0 = 0$ *and that* $U = \mathbb{R}^m$*. Then there exists a unique matrix* $P(t)$ *of order* n *and of class* C^1 *on [0,T] such that the optimal control* $\bar{u}$*, the adjoint state* $\bar{p}$*, and the optimal trajectory* $\bar{x}$ *the following relation:*

$$\bar{u}(t) = -R^{-1}B^t\bar{p}(t), \ \ \bar{p}(t) = P(t)\bar{x}(t), \ \ \forall t \in [0, T]. \tag{3.23}$$

The matrice $P(t)$ *is the solution of the following ordinary differential equation:*

$$\dot{P} = -A^tP - PA + PBR^{-1}B^tP - Q, \ \ t \in [0, T], \ \ P(T) = 0 \tag{3.24}$$

Moreover, for all $t \in [0, T]$*,the matrix* $P(t)$ *is symmetric semi positive define, and is positive define is* D *is.*

Conclusion 1 *Given a dynamical optimization problem,we have proved the existence and uniqueness of solution to that problem.Moreover we summarize the problem to a simple optimization problem in finite dimension given by* $\inf_{u \in U \subseteq \mathbb{R}^m} F(v) := 1/2(Rv, v) + (B^t\bar{p}(t), v)$*,and the solution to this problem represent our control* $\bar{u}$ *which is the solution to our dynamical system.*

3.3 Bibliography

BIBLIOGRAPHY

[1] H. BREZIS; *Analyse Fonctionnelle:Theorie et Applications*, Masson (1983) et Dunod (1999).

[2] C.E. Chidume; *Applicable functional analysis*, International Center for Theoretical Physics Trieste, Italy, July 2006.

[3] G. Degla, *Lecture notes on Ordinary Differential Equation*, African University of Science and Technology, 2010.

[4] N. Djitte ; *Cours Optimisation en dimension infinie: Calcul des Variations, Controle Optimal: Principe de Pontryagin*, Universite Gaston Berger,Saint Louis, Senegal 2005-2006.

[5] N. Djitte, *Lecture notes on Convex analysis and Variational method*, African university of science and Technology, 2011

[6] L. ThIbault; *Lecture Notes on Convex Analysis*, African University of Science and Technology, 2010.

[7] , L. Thilbault, *lecture notes on Measure and Integration Theory*, African university of science and Technology, 2011