Analytic and geometric properties of steep functions

Martha Eileen Faraggiana

July 13, 2015

Contents

Summary

This thesis deals with *steepness*, a property of continuously differentiable functions defined in open sets in $\mathbb{R}^n$. The steepness property, which is a local property, in a certain way, evaluate the slope of a function at a point. We decided to study this property since it is the key hypothesis of Nekhoroshev's theorem, which deals with Hamiltonian dynamical systems under small perturbations.

In the first chapter we review the general mathematical framework where the steepness property was observed and studied. The chapter does not explain what the steepness property is, but it presents the scenario where it is used. Our aim is to show that such property is particularly important if it holds for an Hamiltonian function.
In the second chapter we give the main definitions. We define both steepness and a weaker property (*arc-steepness*), a property defined through continuous curves rather than analytically.
In the same chapter, we also define quasi-convex and three-jet non degenerate functions. We prove that quasi-convex and three-jet non degenerate functions are steep. In particular, this class of steep functions deserve particular attention since they arise in many physical problems.
The proof for quasi-convex functions is simple, while this is not in the case in the three-jet non degenerate case; indeed, such a proof seem not to be available in literature. We provide an original proof of this last statement. Such proof is divided in two steps: in the first step we prove the property on 1-dimensional subspaces while the second step concerns the k-dimensional spaces for $1 < k < n$. Then, we bring again the problem to the unidimensional case.

The second part of the thesis deals with the steepness property restricted to real analytic functions. The real analytic case is relevant since it is an hypothesis on the Hamiltonian in the Nekhoroshev's theorem.
We prove that a real analytic function, defined on an open set in $\mathbb{R}^n$, is steep if and only if its restriction to proper affine subspaces has only isolated critical points.
The proof of this result is quite geometrical, but one important point is that, in the real analytic context, the definitions of steepness and arc-steepness coincide. Indeed, it is possible to obtain estimates for the classical steep condition, by exploiting the analyticity of the function.

In the last part of the thesis, we give concrete examples and counterexamples of steep functions.
We provide the example of a function which shows the dependence of the steepness property to the dimension of the subspace where the property is verified. We also give a classical example of a nonsteep function for which Nekhoroshev theorem does not hold. Finally, we give a counterexample of a C^k but not real analytic function which does not verify the steepness condition while it verifies the arc-steepness condition.

In the appendices we summarize definitions and theorems not directly related to the steepness property, but used in our proofs.

Acknowledgments:

We thank Professor Massimiliano Guzzo for his generous cooperation, for many suggestions on how to find a proof of Theorem 2.3.2.
We also thank Dott. Domenico Monaco for his contribution in the preparation of the Appendix on the Grassmann manifold.

Introduction

The steepness property, a property of C^1 functions not easy to define (see Definition 2.1.1), is of considerable interest in the study of dynamical systems. Indeed, it is the core hypothesis of a fundamental theorem in Hamiltonian mechanics, i.e. the Nekhoroshev's theorem. Such theorem concerns with exponential stability of "nearly" integrable Hamiltonian systems and is widely used in mathematical physics and especially in celestial mechanics.

The Nekhoroshev's theorem is a part of a classical but still current research topic known as "perturbation theory" which, as the name suggests, is a branch of Hamiltonian mechanics that studies the behavior of "perturbed" dynamical systems. Such systems are close, in a certain way, to simpler systems, with easier to solve equations.

The main goal pursued by many mathematicians of the 20th century was precisely to develop a valid theory which would account for the motion of perturbed systems.
The topic is of considerable interest, even outside the community of mathematicians, to a number of scientists in various fields, especially physics, astronomy and biology.
The classical perturbed Hamiltonian system we are interested in is called "nearly integrable" (See section 1.2 in the first chapter) and the main object of the problem is the Hamiltonian function

$$H(I, \varphi) = h(I) + \varepsilon f(I, \varphi) \tag{1}$$

with its related system of differential equations, given initial data and where $\varphi = (\varphi_1, .., \varphi_n)$ is such that φ_i are angles for every i[1].

[1] $I \in \mathbb{R}^n$ are called action variables. However, everything will be defined in detail in the next chapter.

According to Poincaré, the solution of a nearly integrable system with Hamiltonian (1) is called the *fundamental problem of dynamics*[2]. Nowadays, concerning (1), there are three types of results that may describe its behavior.
The first important results of the last century concerned a way of studying (1) which is often called "geometric perturbation theory". This approach searches geometric objects invariant under the flow of (1) and these first results were achieved after the fiftys thanks to a group of mainly russian investigators of dynamical systems who set the foundations of modern perturbation theory. In particular, the most significant advance was represented by the *KAM theorem*, stated for the first time by A.N. Kolmogorov (1954) and then extended by V.I. Arnol'd (1962) and J.K. Moser (1961) from which arose the so-called *KAM theory*.
The strength of the KAM theorem is that it ensures the stability of most perturbed motions for infinite times if the Hamiltonian (1) that governs the system satisfies certain assumptions. More precisely, if the hessian of the unperturbed Hamiltonian is non degenerate, i.e. det $h'' \neq 0$ and if $0 < \varepsilon << 1$ is the magnitude of the perturbation, then the evolution of the perturbed system differs for "most" initial data of an order $\sqrt{\varepsilon}$ from the unperturbed one. The KAM theorem is valid only with a specific kind of initial data and this restriction significantly limits its application since it is not valid on any open set of the phase space.

On the other hand, from 1971, N.N. Nekhoroshev, another russian mathematician and student of Arnol'd, introduced a new theorem firstly in [19] and later described in greater depth in [17] and [18] that achieved an old part of the theory known as "classical perturbation theory" concerning with the stability of the motion over long (exponential) time intervals. We are interested in the steepness property which is strictly related to exponential stability.
Before continuing with the discussion on the Nekhoroshev's theorem, we introduce the last main result concerning the study of (1), which is known as "Arnol'd diffusion". Arnol'd diffusion essentially covers behaviors not dealt with by classical and geometric perturbation theory like the future behavior of the trajectories outside the objects discovered by KAM theory or over intervals of time non covered by classical perturbation theory. The name Arnol'd diffusion is dedicated to V.I. Arnol'd who firstly provided in [1] a concrete example of a system with an $O(1)$ order of drift in the action variables I.
However, Nekhoroshev proved that in a perturbed system if the function related to the unperturbed part of the system, specifically named unperturbed

[2]Poincaré H. *Les méthodes nouvelles de la mécanique céleste* Gauthier-Villars, Paris, (1892)

Hamiltonian ($h(I)$ in (1)), verifies the steepness property, then the evolution of all the initial data of the system are controlled for an exponentially long time or in astronomical terms, on a cosmologial time scale.

Therefore, Nekhoroshev's theorem can be applied to systems governed by perturbed Hamiltonians as long as the unperturbed part verifies the steepness property.
A first definition of steepness appeared in [19] so that Nekhoroshev deserves full credit for the discovery of this class of functions. He defined them for the first time and described in a rather difficult way in [18] and [20]. A first purpose of our work is therefore to provide a more undestandable key to identify which ones are the steep functions.

The steepness property, as the name suggests, is a property that "evaluate the slope" of a regular function. The more the function is steep, the more we can find good powers and constants that characterize this property.

The importance of the steepness property in Nekhoroshev's theorem resides in the fact that the estimates provided by the theorem are given in terms of numbers that Nekhoroshev called *steepness indices* and *steepness coefficients* that are essential especially for the time of closeness of the evolutions. Actually, the theorem asserts that the smaller the indices are (so the more the function is steep), the more all the estimates are sharp.
We recall in Appendix C the statement of a theorem that provides estimates on the steepness indices which are in general very difficult to calculate.

In this thesis we calculate the indices of the class of steep function most commonly studied in perturbation theory called quasi-convex and three-jet non-degenerate functions but keep in mind that the total class is much larger. Furthermore, in our work we will show a general result about real analytic functions and steepness.
With a more geometric reading of the steepness property, we prove that a real analytic function is such that on any restriction of its domain it admits only isolated critical points if and only if is steep. This result is local since it holds only on compact sets contained in the domain of the function.
Since in Nekhoroshev's theorem the Hamiltonian is supposed to be real analytic, it happens that many texts that deal with real analytic Hamiltonians, like [2], use this alternative definition of steepness rather than the original one provided by Nekhoroshev.
The proof of this last result uses the concept of *subanalytic functions.* This is a class of functions whose graph is subanalytic in the sense of H. Hironaka, a

japanese mathematician who introduced, in 1973 ([12]), the term "subanalytic" to identify a class of sets whose study was previously developed by Lojasiewicz (1964).
We will not go further into this topic, but we will state the main results that will be used in our present work. For a complete theory on subanalytic geometry, we refer to [12], [5] and [26].
The proof of this theorem, that leads to this equivalence, is based on the subanalyticity and on the continuity of the projection of the Hamiltonian's gradient, restricted to a compact analytic subset of the domain.
The projection of the gradient is a continuous, real analytic and subanalytic function. The geometry of the domain leads to prove the existence of a real analytic curve[3] contained in the compact restriction. Such curve contains a point that verifies an inequality called, in subanalytic geometry, *Lojasiewicz's Inequality*, which is essentially equivalent to steepness.
Compactness of the restriction is a relevant hypothesis. Indeed, we take advantage of the finite subcover since we use the existence of such analytic curve on each component of such subcover.
The final part concerns with various and elementary examples and counterexamples.

[3] *Curve Selection Lemma* , [5] and [16].

Chapter 1

A General Framework

In this first chapter we would like to present the mathematical context where the steepness property was considered for the first time and to explain its importance. In order to be able to proceed, we will assume that are known the foundamental notions on dynamical systems or Hamiltonian mechanics. The reason is that the steepness property for a function has been concieved as a foundamental hypothesis for a very strong theorem used in what is known as "perturbation theory" which is, as we show in a while, a very important branch of the theory of hamiltonian integrable systems.
We will recall the basic notions of an hamiltonian integrable system in order to let the reader understand the framework where the steepness property is most commonly used. When not necessary and for sake of brevity, we will omit detailed statements of theorems that would require too many definitions, beyond the purpose of this thesis.
For further information, details or reminders on theorems and main definitions in this chapter we refer to [3] and [2].

1.1 Hamiltonian Systems on $\mathbb{R}^n \times \mathbb{T}^n$

Given a domain $D \subseteq \mathbb{R}^n$, we define an action-angle Hamiltonian system as a dynamical system of ordinary differential equations of the first order defined on $\mathbb{R}^n \times \mathbb{T}^n$ where $\mathbb{T}^n$ is an n-dimensional torus. To avoid confusions, in the context of Hamiltonian systems the n-dimensional torus is defined as the quotient space

$$\mathbb{T}^n := \frac{\mathbb{R}^n}{2\pi\mathbb{Z}^n}$$

defined through the equivalent relation that identifies $x, y \in \mathbb{R}^n$ in the following way: $x \sim y$ if and only if $x_i - y_i = 2\pi k_i$ for every $i = 1, ..., n$ and for some $k \in \mathbb{Z}$.
An action-angle Hamiltonian system has the form

$$\begin{cases} \dot{I}(t) = -\partial_\varphi h\,(I,\varphi) \\ \dot{\varphi}(t) = \partial_I h\,(I,\varphi) \end{cases} \tag{1.1}$$

with initial conditions

$$(I(0), \varphi(0)) = (I_0, \varphi_0)$$

where $(I, \varphi) \in D \times \mathbb{T}^n$ are called action-angle variables; the dot denotes the differentiation with respect to the time variable while ∂_I is the I-gradient, that is

$$\partial_I h(I,\varphi) = (\partial_{I_1} h(I,\varphi), ..., \partial_{I_n} h(I,\varphi))$$

and, in the same way, ∂_φ is the φ-gradient

$$\partial_\varphi h(I,\varphi) = (\partial_{\varphi_1} h(I,\varphi), ..., \partial_{\varphi_n} h(I,\varphi))$$

and the function

$$h : D \times \mathbb{T}^n \to \mathbb{R} \tag{1.2}$$

is a smooth and scalar Hamiltonian[1] function defined on the phase space $D \times \mathbb{T}^n$ and 2π-periodic in every component φ_i of φ.
We consider action-angle systems since in this coordinates the solution of (1.1) can be found analytically and this property is provided by a theorem often known as *Liouville-Arnol'd's theorem*.
Liouville-Arnol'd's theorem states that if a Hamiltonian system with n degrees of freedom has also n first integrals[2] that are independent from each other in a certain way, then the system is integrable by quadratures, that is, the solution is an explicit integral of a known function.
Action-angle variables are essential in perturbation theory, where by the term perturbation theory we mean the collection of methods for studying solutions of perturbed problems which are closed to completely solvable non-perturbed problems.
With completely solvable non-perturbed problems we mean an integrable Hamiltonian system such that the Hamiltonian can be written in the form

$$h(I,\varphi) = h(I)$$

[1]We will focus our attention on real analytic Hamiltonians.

[2]A first integral for a system of ordinary differential equations of first order is a non-constant continuously differentiable function F such that is is locally constant on any solution of the system.

that is, it depens only on the action variables and then Hamilton's equations of motion (1.1) assume the simplified form

$$\begin{cases} \dot{I}(t) = 0 \\ \dot{\varphi}(t) = \partial_I h(I) \end{cases} \tag{1.3}$$

with initial conditions

$$(I(0), \varphi(0)) = (I_0, \varphi_0)$$

It is clear that all the motions are linear and the actions do not move for all the times. Then the motion lies permanently on an n-dimensional manifold known as *invariant torus*.
To be precise, the solution of the system (1.3) is

$$\begin{cases} I(t) = I_0 \\ \varphi(t) = \varphi_0 + \omega t \end{cases} \tag{1.4}$$

where

$$\omega := \omega(I_0) = \frac{\partial h}{\partial I}(I_0)$$

is called the *frequency vector* of the solution and the solution curve is a straight line which, due to the identification of the angular coordinates modulus 2π, is called a *winding* on the invariant torus $\mathcal{T} = \{I_0\} \times \mathbb{T}^n$.

However, the collection of integrable problems in the form (1.1) is not very large. We are interested in the study of those known as *nearly integrable hamiltonian systems* or *hamiltonian systems close to integrable.*

1.2 Nearly-Integrable Hamiltonian Systems

With the term nearly integrable hamiltonian systems we mean a system governed by a hamiltonian in the form

$$H(I, \varphi, \varepsilon) = h(I) + \varepsilon f(I, \varphi, \varepsilon) \tag{1.5}$$

where $(I, \varphi) \in D \times \mathbb{T}^n$ are the usual action-angle variables. $\varepsilon << 1$ is a real positive number and it is called the magnitude of the perturbation. We will assume that the hamiltonian H is real analytic in all the components, that is, both the unperturbed Halimiltonian h and the perturbation f are real analytic functions. We also assume that f is 2π-periodic in every component φ_i of φ.

For $\varepsilon = 0$ the system (1.5) returns to be integrable and the space is foliated into invariant tori $\mathcal{T}$.
We recall that the frequency vector ω is said to be *rationally indipendent* if for every $k \neq 0 \in \mathbb{Z}^n$ it is $\omega \cdot k \neq 0$. The tori where the frequencies are rationally indipendent are called *non resonant* and this means that the trajectory fills the non-resonant torus everywhere densely.

The integrable system is said to be *non degenerate* if

$$\det\left(\frac{\partial \omega}{\partial I}\right) = \det\left(\frac{\partial^2 h}{\partial I^2}\right) \neq 0$$

Now with this notations, we state an important theorem first introduced in 1954 by Andrey Kolmogorov and then extended by V. Arnol'd and J. Moser in 1962 and in 1961 respectively. The theorem is known as the KAM theorem[3].
The original theorem was stated for real analytic unperturbed Hamiltonians but it is valid also for C^k functions, for $k > 2n$.

Theorem 1.2.1 (KAM)**.** *[2] If the unperturbed Hamiltonian system is non-degenerate or Iso-Energetically Non-degenerate[4] then under sufficiently small perturbations most of the non-resonant invariant tori do not disappear but are only slightly deformed, so that in the phase space of the perturbed system there also exist invariant tori filled everywhere densely with phase curves winding around them quasi periodically with the number of frequencies equal to the number of degrees of freedom. These invariant tori form a majority in the sense that the measure of the complement of their union is small together with the perturbation. In the case of isoenergetic non-degeneracy the invariant tori form a majority on each energy level manifold.*

The theorem states that the motion of a perturbed system with non-degenerate Hamiltonian is preserved almost always. The word "almost" means that the set outside invariant tori has small measure and KAM theorem asserts that this happens with a small probability.
In particular, the invariant tori constructed in this theorem are called Kolmogorov's tori and then their union form the *Kolmogorov's set.* The measure of the complementof the Kolmogorov's set does not exceed a quantity of order $\sqrt{\varepsilon}$ and the deviation of the perturbed torus depends on the arithmetic properties of the frequencies.
The scenary changes when the motion outside the invariant tori is studied. One has to distinguish the cases with two or higher degrees of freedom.

[3]For the complete statement we refer to [22].
[4]This property is defined in 2.1.8 in the next chapter.

Remark 1.2.1. We remind that in a isoenergetically nondegenerate systems with two degrees of freedom ($n = 2$) the existence of a large number of invariant tori implies the absence of actions for *all* the initial conditions and not just for most. That is, in such systems, for all initial conditions the action variables remain forever near their initial values (See [2] or [19]). This happens because if $n = 2$, the tori separate the three-dimensional $(2n - 1)$ energy surface into two components, and thus no trajectory can pass from one side of the torus to the other. Since the torus is defined by setting each action to a fixed constant (the initial condition I_0), then it is clear that the variation of each action stays bounded for all time.

But when the degrees of freedom are greater or equal to three, then the n-dimensional tori do not divide the $(2n-1)$-dimensional energy level manifold: there is then no a priori reason for stability, that is, there might be trajectories along which the action variables go away from the initial conditions.

1.2.1 Nekhoroshev's Theorem

We now introduce the Nekhoroshev's theorem as a complement of the KAM's theorem. We said that KAM theorem does not hold with a small probability while, on the other hand, Nekhoroshev's theorem asserts that if the perturbed action drifts away from the initial condition, then it happens extremely slowly. The hypotesis of this theorem is the topic of our thesis: if the unperturbed Hamiltonian is a real analytic steep function, then the action dritfs away slowly. As we saw, the theorem is intresting when the system has at least three degrees of freedom.
A first definition of steepness can be given as follows

> *Steepness* A continously differentiable function $h : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be steep in U with positive steepness coefficients $\widetilde{\xi}, C_1, ..., C_{n-1}$ and steepness indices $\delta_1, ..., \delta_{n-1}$ real numbers greater or equal to 1 if
>
> $$\inf_{I \in U} \|\partial_I h(I)\| > 0$$
>
> and if for any $I \in U$ and any k-dimensional linear subspace $V \subseteq \mathbb{R}^n$ orthogonal to $\partial_I h(I)$ with $k \in \{1, ..., n-1\}$, it is
>
> $$\max_{0 \le \eta \le \xi} \min_{\substack{\|u\| = \eta \\ u \in V}} \|P_V \partial_I h(I + u)\| \geq C_k \xi^{\delta_k}$$

We will return with more details on this definition in the next chapter.
We set $U - r := \Big\{ I \in U \text{ such that } \overline{B_r(I)} \subseteq U \Big\}$.

However, we introduced all the elements that lead Nekhoroshev to state the main result in classical perturbation theory and then, following [9] and the definition of steepness provided above, we state

Theorem 1.2.2 (Nekhoroshev). *Let H be a real analytic perturbed Hamiltonian as in (1.5) and let the unperturbed Hamiltonian h be a steep function with steepness indices $\delta_1, ..., \delta_{n-1}$ and let*

$$p_1 := \prod_{k=1}^{n-2} \delta_k \ , \qquad a := \frac{1}{2np_1} \ , \qquad b := \frac{a}{\delta_{n-1}} \tag{1.6}$$

Then, there exist positive constants $\varepsilon_0, R_0, T, c > 0$ such that for any $0 \leq \varepsilon < \varepsilon_0$ the solution $(I(t), \varphi(t))$ of the Hamilton equations for $H(I, \varphi)$ with initial data (I_0, φ_0) with

$$I_0 \in U - 2R_0\varepsilon^b$$

that have the form (1.3), satisfies

$$\|I(t) - I_0\| \leq R_0\varepsilon^b \tag{1.7}$$

for any time t such that

$$|t| \leq \frac{T}{\sqrt{\varepsilon}} \exp\left(\frac{c}{\varepsilon^a}\right) \tag{1.8}$$

It can be seen from the statement that, since $\delta_k \geq 1$, then the maximum value for the exponent a and b, that leads to the maximal estimate on the time of stability, is $p_1 = 1$, that is, all the steepness indices are equal to 1.
This is a particular case where the unperturbed Hamiltonian is *quasi convex*. Quasi-convexity is the class of "steepest" functions, as we will prove in the next chapter. However, what we wish to highlight is that the quasi-convex unperturbed Hamiltonian is the best situation which can happen in terms of prediction of stability.

Chapter 2

General properties of steep functions

2.1 Definitions

We start our work by giving a list of notations and definitions of the elements that we will be using.

Let h be a C^k real valued function (with $k \geq 1$ or $k = +\infty$) defined in an open domain G contained in $\mathbb{R}^n$ and assume that h is regular on G, that is, $\nabla h(I) \neq 0$ for every $I \in G$.
We denote by $\|\cdot\|$ the usual Euclidean norm and by $\nabla h(I)^\perp$ the $(n-1)-$dimensional subspace of $\mathbb{R}^n$ orthogonal to the gradient of h, that is

$$\nabla h(I)^\perp = \{x \in \mathbb{R}^n : \nabla h(I) \cdot x = 0\}$$

where the dot denotes the scalar product. By $h''(I) := H_h(I)$ we denote the hessian matrix of h at the point I and, for every $x \in \mathbb{R}^n$, we will denote by $h^{(k)}[\cdot, x, ..., x](I)$ the vector with j_1th component defined by[1]

$$\sum_{j_2,..j_k} \frac{\partial^k h}{\partial I_{j_1} ... \partial I_{j_k}}(I)\, x_{j_2} ... x_{j_k} \tag{2.1}$$

For every $k \in \{1, ..., n-1\}$ we denote by $V^k := V^k(I)$ both the affine subspace

[1]See Appendix A.

contained in $\nabla h(I)^{\perp}$ such that $\dim V^k = k$ and passing through the point I and the corrisponding k-dimensional vector subspace.
Even when not specified, we will assume that the function h has these features.
We also denote with $\bar{u}$ a versor in V^k, *i.e.*, a vector with unit norm.

The main definitions of this thesis are the followings

Definition 2.1.1. *Fix $k \in \{1, ..., n-1\}$. A C^1 function $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be* **steep** *at a point $I \in G$ on a k-dimensional vector space V^k if exist positive constants $\widetilde{\xi}_k, C_k, \delta_k$ such that the inequality*

$$\max_{0<\eta\leq\xi} \min_{\substack{\bar{u}\in V^k\cap G \\ \|\bar{u}\|=1}} \|P_{V^k}\nabla h(I+\eta\bar{u})\| \geq C_k \xi^{\delta_k} \tag{2.2}$$

holds for every $0 < \xi \leq \widetilde{\xi}_k$ and where P_{V^k} is the orthogonal projection on V^k.

The numbers $\widetilde{\xi}_k$ and C_k are called ***steepness coefficients*** *while δ_k are called* ***steepness indices****.*

Definition 2.1.2. *A C^1 function $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be* **steep** *at a point I if the following conditions hold:*

1. *$\nabla h(I) \neq 0$, that is, I is regular for h.*
2. *$\forall\ k \in \{1, ..., n-1\}$ there exist positive constants ξ_k, C_k, δ_k such that h verifies condition (2.2) at I on every affine subspace $V^k \subseteq \nabla h(I)^{\perp}$.*

Definition 2.1.3. *A C^1 function h is said to be* **steep** *over an open domain $G \subseteq \mathbb{R}^n$ with positive coefficients $C_1, ..., C_{n-1}, \xi_1, .., \xi_{n-1}$ and indices $\delta_1, .., \delta_{n-1}$ if there are no critical points in G and if h is steep at every point $I \in G$ with uniform coefficients and indices.*

In the essence we defined a steep function as a regular function such that the norm of the projection of the gradient $\|P_{V^k}\nabla h(\cdot)\|$ vanishes in I, but then, moving from I inside V^k, $\|P_{V^k}\nabla h(\cdot)\|$ grows at least as a power of the distance from I.
Later on, when we talk about steep functions, we will refer to Definition 2.1.3, unless otherwise specified.

Definition 2.1.4. *Consider an open set $G \subseteq \mathbb{R}^n$. A C^1 function $h : G \to \mathbb{R}$ is said to be* **arc-steep** *at a point $I \in G$ along an affine subspace V which contains I if exist constants $C > 0, \widetilde{\xi} > 0$ and $\delta > 0$ such that along any continous*

curve $\gamma : I \subseteq \mathbb{R} \to G$ that connects I to a point at a distance $\xi \leq \widetilde{\xi}$, the norm of the projection of the gradient $\nabla h(I)$ on the direction of V is greater than $C\widetilde{\xi}^\delta$ at some point $\gamma(t_*)$, that is, $\nabla h(I)$ does not vanish identically along any continous curve contained in V.
$(C, \widetilde{\xi})$ and δ are respectively called the **steepness coefficients** *and the* **steepness index**.
With the same assumptions, a C^1 function h is said to be arc-steep at a point $I \in G$ if I is not a critical point for h and if, for every $k \in \{1, ..., n-1\}$, there exist positive constants $C_k, \widetilde{\xi}_k$ and δ_k such that h is arc-steep at I along any affine subspace of dimension k containing I and h is arc-steep uniformly with respect to the coefficients $(C_k, \widetilde{\xi}_k)$ and the index δ_k.
Finally, a C^1 function h is arc-steep over a domain $G \subseteq \mathbb{R}^n$ if h has no critical points in G and if there exists steepness coefficients $(C_1, ..., C_{n-1}, \xi_1, ..., \xi_{n-1})$ and the steepness indices $(\delta_1, ..., \delta_{n-1})$ such that h is arc-steep at any point $I \in G$ along any k-dimensional subspace V for every k uniformly with respect to these coefficients and indices.

Remark 2.1.1. In general, definitions 2.1.1 and 2.1.4 are not equivalent. See section 2.1.1.

We now define a class of functions known as the class of "steepest" functions; steepest implies that their steepness index, as we will see, is minimum.

Definition 2.1.5. *A C^2 function $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$ with G open is said to be* **convex** *at the point $I \in G$ if the only real solutions of the equation*

$$\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial I_i \partial I_j}(I) x_i x_j = 0 \tag{2.3}$$

is $x = (x_1, ..., x_n) = 0$.

Definition 2.1.6. *A C^2 function $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$ with G open is said to be* **quasi-convex** *at the point $I \in G$ if the following properties are satisfied:*

i) $\nabla h(I) \neq 0$

ii) The system

$$\begin{cases} \displaystyle\sum_{i=1}^{n} \frac{\partial h}{\partial I_i}(I) x_i = 0 \\ \displaystyle\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial I_i \partial I_j}(I) x_i x_j = 0 \end{cases} \tag{2.4}$$

has only the trivial solution $x = (x_1, ..., x_n) = 0$.

A function is **quasi-convex** *(respectively* **convex***) on G if it is quasi-convex (respectively convex) at every $I \in G$.*

The class of quasiconvex functions is extensively used in the study of perturbed dynamical systems and the first results on exponential stability involved mainly this kind of unperturbed Hamiltonians (See [15] or [23]).

The following definition concerns another class of steep functions that generalizes the class of quasi-convex functions. Such generalization is not trivial as we will see in Section 2.3.

Definition 2.1.7. *A C^3 function $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be* **three-jet non-degenerate** *at a point $I \in G$ if $\nabla h(I) \neq 0$ and if the only solution $x = (x_1, ..., x_n)$ of the system*

$$\begin{cases} \displaystyle\sum_{i=1}^{n} \frac{\partial h}{\partial I_i}(I) x_i = 0 \\ \displaystyle\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial I_i \partial I_j}(I) x_i x_j = 0 \\ \displaystyle\sum_{i,j,k=1}^{n} \frac{\partial^3 h}{\partial I_i \partial I_j \partial I_k}(I) x_i x_j x_k = 0 \end{cases} \tag{2.5}$$

is $x = (x_1, ..., x_n) = 0$.

The following is one of the hypothesis of theorem 1.2.1 stated in the first chapter.

Definition 2.1.8. *A funcion $h : G \subseteq \mathbb{R}^n \Rightarrow \mathbb{R}$ is called* **iso-energetically non-degenerate** *at a point $I \in G$ if, called*

$$\hat{H} := \begin{pmatrix} h''(I) & \nabla h(I)^T \\ \nabla h(I) & 0 \end{pmatrix} \tag{2.6}$$

it is:

$$\det \hat{H} \neq 0 \tag{2.7}$$

Finally, we have the following

Definition 2.1.9. *Let G be an open domain in $\mathbb{R}^n$. A C^1 function $h : G \to \mathbb{R}$ is said to be* **NNI-non-degenerate**[2] *on G if it is regular on G (i.e. $\nabla h(I) \neq 0$ for every $I \in G$) and if its restriction to any subset G_k where $G_k = D \cap V^k$ and V^k is a k-dimensional subspace of $\mathbb{R}^n$, admits only isolated critical points.*

[2]NNI stands for Nekhoroshev-Neishtadt-Il'yashenko.

We will have to talk about the importance of this definition again it the next section.

2.1.1 Comments and remarks

i) We see that the definition 2.1.1 implies that the function

$$f(I, \eta, \bar{u}) := \|P_{V^k} \nabla h(I + \eta \bar{u})\|$$

defined on $G \times [0, \xi] \times V^k$ reaches both its minimum and then its maximum on the sets V^k and $[0, \xi]$ respectively. One may ask if these points are always reached or if it could happen that one of them is not: the answer is that they are both always reached. The minimum is clearly reached since $f(I, \eta, \bar{u})$ is continous in $\bar{u}$ on the compact set

$$K = \left\{ V^k \text{ such that } \|\bar{u}\| = 1 \right\}$$

so we can apply the Weierstrass's theorem. Once we know that the function

$$\mu(\eta) := \min_{\substack{\|\bar{u}\|=1 \\ \bar{u} \in V^k}} f(I, \eta, \bar{u}) \tag{2.8}$$

is well defined, we investigate if the maximum is reached on $[0, \xi]$.
We prove in Appendix B that (2.8) is an upper semi-continous function and since they reach their maximum on a compact set, then the definition is well-stated.

ii) In definitions 2.1.2 and 2.1.3 we expand the concept of steep function to all the subspaces contained in $\nabla h(I)^\perp$. We then can talk about steepness of a function at a point indipendently from the dimension of the subspace where the function verifies the definition.
We are interested in functions that verify the definition on all the subspaces of dimension k, for all k.

iii) With the aim of giving a geometrical meaning or an intuitive description of the steepness property, we consider a (at least) C^1 function h defined on $\mathbb{R}^n$ steep at I on a plane $V^k \subseteq \mathbb{R}^n$ with coefficients and indices C_k, ξ_k, δ_k. Then, we consider a curve γ joining the point I with another point I' at a distance ξ such that $0 < \xi \leq \xi_k$. Then on this curve a point exists such that the lenght of $\nabla h_{|_{V^k}}$ at this point is greater than $C_k \xi_k^{\delta_k}$.
We also want to remark that our definitions consider only planes orthogonal to $\nabla h(I)$, while the definition can also be given with respect to a

generic subspace of dimension $k < n$ containing the point I, as it can be seen in many papers.
We prefer this last definition since if V^k is not orthogonal to $\nabla h(I)$ and if h is not constant, then $\|P_{V^k} \nabla h(I)\| \neq 0$ and steepness coefficients and indices can always be found. This point is shown in the next section.
This last case of planes, not orthogonal to $\nabla h(I)$, is not interesting since we focus on functions that are uniformly steep in the sense of definition 2.1.3, that is, we look for uniform indices and coefficients for all the k-dimensional subspaces.
On the other hand, if V^k is orthogonal to $\nabla h(I)$, then $\|P_{V^k} \nabla h(I)\| = 0$ and h could be nonsteep along some subspaces. This is the reason why for us the only interesting situation is the case $V^k \perp \nabla h(I)$.

iv) As outlined in the previous section, when a subspace V^k is not orthogonal to the gradient $\nabla h(I)$ then the function always verifies the steepness condition at that point on that plane, since it can always be found a point close to I where the projection is not null. Here in this section we recall just few general properties of the orthogonal projection to explain this trivial case.

1. Let $h : G \subseteq \mathbb{R}^n \to \mathbb{R}$.
   Then for each V^k vector subspace and $\forall\, \bar{x} \in V^k \setminus \{0\}$ it is

$$\|P_{V^k} \nabla h(I + \bar{x})\| \geq \frac{\|\nabla h(I + \bar{x}) \cdot \bar{x}\|}{\|\bar{x}\|} \tag{2.9}$$

   Proof. It is

$$\begin{aligned} \|\nabla h(I + \bar{x}) \cdot \bar{x}\| &= \|\nabla h(I + \bar{x}) \cdot P_{V^k} \bar{x}\| = \\ = \|P_{V^k} \nabla h(I + \bar{x}) \cdot \bar{x}\| &\leq \|P_{V^k} \nabla h(I + \bar{x})\| \cdot \|\bar{x}\| \end{aligned}$$

   The result follows also from the characterization of P_V in fact

$$P_V u = w \iff w \in V \text{ is such that } \|w\| = \max_{\substack{v \in V \\ \|v\|=1}} u \cdot v$$

   □

2. Let now V be the straight line

$$V := \{tv \text{ such that } \; v \cdot \nabla h(I) \neq 0 \text{ and } \|v\| = 1;\; t \in \mathbb{R}\}$$

Let fix $0 < \alpha < \|v \cdot \nabla h(I)\|$. Then for every $\bar{x} \in V$ small enought, it is

$$\|P_V \nabla h(I + \bar{x})\| \geq \alpha$$

Proof. Let $\bar{x} \in V \setminus \{0\}$ then $\bar{x} = tv$. It follows from the previous property and by the continuity of the gradient that

$$\|P_V \nabla h(I + \bar{x})\| \geq \left\| \nabla h(I + \bar{x}) \cdot \frac{\bar{x}}{\|\bar{x}\|} \right\| = \|\nabla h(I + \bar{x}) \cdot v\| \geq \alpha$$

$\square$

Remark 2.1.2. If V is a linear vector space such that $\nabla h(I) \cdot \bar{x} \neq 0$ for every $\bar{x} \in V \setminus \{0\}$, then $\dim V = 1$.
Indeed, if by contradiction it is $\dim V = k \geq 2$, then $V = \text{span}[u_1, ..., u_k]$ for some vectors u_i linearly indipendent such that $u_i \cdot \nabla h(I) \neq 0$ for every i, but we find that

$$-\frac{u_2 \cdot \nabla h(I)}{u_1 \cdot \nabla h(I)} u_1 + u_2 \in \nabla h(I)^{\perp} \setminus \{0\}$$

which is a contradiction.

v) We introduced the notions of quasi-convex and three-jet non-degenerate functions following the same definition provided in [17]. We highlight that if a function h is steep at a point I on a space V^k, then the index is $\delta_k \geq 1$, and it can not be smaller because of the regularity of h.
We will prove that the quasi-convex functions are the class of "steepest" functions, that is, their steepnes index is equal to 1 and they are immediately followed by the three-jet non-degenerate ones that have index 1 or 2, depending on the dimension of V^k. The proofs and the computation of the indices are postponed to the next section.

vi) We introduced NNI-non-degenerate functions in Definition 2.1.9. This class of functions is very important for our work.
Indeed, we prove that a real analytic function is NNI-non-degenerate over a compact subset of the domain of the function if and only if the function is steep on that subset. The label "NNI" stays for Nekhoroshev-Neishtadt-Il'yashenko. They were the first mathematicians who guessed the connection between the steepness property and the numer of critical points of the restriction.
Actually, A.I. Neishtadt was the first who thought about this connection, since he presented it as a conjecture, while Nekhoroshev in 1979 found

an alternative proof for the steepness condition, stated and proved in [18] and [20][3]. Finally in 1984, Yu.S. Il'yashenko, for the first time in [13], provided a proof of Neishtadt's conjecture and proved that a NNI-non-degenerate function is steep with steepness index equal to μ, which is the number of isolated critical points of the restriction counted with their multiplicity. The proof uses results for functions of several complex variables and of Riemann's geometry.
In our work, in the next Chapter, we prove a new theorem showing that the steepness index can be taken equal to a number called the Lojasiewicz's exponent of two particular functions (See Definition 3.1.1 below). This result is local.
On the other hand, Nekhoroshev's work was followed by G. Schirinzi in [25]. In these paper she showed that over a certain order of non-degeneration ($r = 4$) Nekhoroshev's conditions become implicit since an elaborate computation of the clousure of certain sets is required.

2.2 Quasi-convex functions

Let us now we recall some importat results that characterize quasi-convex functions in order to recognize them easily and to compute their steepness indices. In particular, we provide a tool to compute in low dimensions ($n = 2, 3$) one of which is the quasi-convexity region of a function.
The definition 2.1.6 implies that the restriction of the second order term of the Taylor expansion of h to the surface $\{x \in \mathbb{R}^n \ : \ \nabla h(I) \cdot x = 0\}$, tangential to the level surface of h at I, has fixed sign.
This is the reason for the term "quasi" that characterizes these functions.

1. Let $V := \nabla h(I)^{\perp}$ and let P_V be the usual orthogonal projection onto V. Then Definition 2.1.6 is equivalent to say that the linear map

$$H_{|_V} := P_V \circ H : V \to V$$

is strictly positive or negative definite. This means that the restriction of the Hessian matrix of h to the hypersurface V has a fixed sign and then, as we said, the level surface of h is a convex set.

2. If h is iso-energetically degenerate then $\hat{H}$ defined in (2.6) has a vanishing eigenvalue, that is, exist $\bar{x} \in \mathbb{R}^n \setminus \{0\}$ and $\lambda \in \mathbb{R} \setminus \{0\}$ such that

$$\begin{cases} \nabla h \cdot \bar{x} = 0 \\ H\bar{x} = \lambda \nabla h \end{cases}$$

[3] See Appendix C

which, in turn, means that either $\nabla h = 0$ or $\nabla h \neq 0$ and there is a $\bar{x} \in V \setminus \{0\}$ and $\lambda \in \mathbb{R}^n$ such that $H\bar{x} = \lambda \nabla h$.

3. If h is quasi-convex, then it is iso-energetically non-degenerate.

   Proof. We assume that $\det \hat{H} = 0$. If $\nabla h = 0$ then h is not quasi-convex. Now, let $\nabla h \neq 0$. By property 2, exists $\bar{x} \in V \setminus \{0\}$ such that $H\bar{x} = \lambda \nabla h$ for some $\lambda \in \mathbb{R}$ not null, so if we take the inner product with $\bar{x}$ we get $H\bar{x} \cdot \bar{x} = 0$, showing that h is not quasi-convex. □

4. Let $h : G \subseteq \mathbb{R}^2 \to \mathbb{R}$. Then, h is quasi-convex if and only if $\det \hat{H} \neq 0$, that is, h is iso-energetically non-degenerate.

   Proof. In view of property 3 it is enough to check that if h is not quasi-convex, then it is iso-energetically degenerate.
   If $\nabla h \neq 0$, then h is iso-energetically degenerate. So let now $\nabla h \neq 0$. Then it exists $\bar{x} \in V \setminus \{0\}$ such that $H\bar{x} \cdot \bar{x} = 0$, that is, $H\bar{x} \in \bar{x}^\perp$.
   But, since $n = 2$, then $\bar{x}^\perp = \left\{ \lambda \nabla h \,\middle|\, \lambda \in \mathbb{R} \right\}$, hence, $H\eta = \lambda \nabla h$, for some $\lambda \in \mathbb{R}$ then the definition of iso-energetically degeneration holds. □

5. Let $h : G \subseteq \mathbb{R}^3 \to \mathbb{R}$. Then, h is quasi-convex if and only if $\det \hat{H} < 0$.

   Proof. By property 1, we shall prove that $\det \hat{H} < 0$ if and only if $H_{|_V}$ is strictly definite.
   Let $U \in O(3)$ be such that $U\nabla h = e_3 := (0, 0, 1)$ and observe that $H_{|_V}$ strictly definite is equivalent to

$$UHU^T|_{(U\nabla h)^\perp} = UHU^T|_{\mathbb{R}^2 \times \{0\}}$$

   strictly definite.
   Now $\hat{U} := \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \in O(4)$, so

$$\begin{aligned} \det \hat{H} &= \det \left(\hat{U} \hat{H} \hat{U}^T \right) = \det \begin{pmatrix} UHU^T & U\nabla h \\ U\nabla h & 0 \end{pmatrix} = \\ &= \det \begin{pmatrix} UHU^T & e_3 \\ e_3 & 0 \end{pmatrix} = -\det \left(UHU^T \right)_{3,3} \end{aligned}$$

   where $\left(UHU^T \right)_{3,3}$ denotes the (2×2)-matrix obtained by deleting the third row and the third column from UHU^T. □

2.2.1 Steepness indices of quasi-convex functions

The steepness index is a foundamental number in the study of perturbed Hamiltonian systems with steep unperturbed Hamiltonian.
We saw that Nekhoroshev proved in theorem 1.2.2 that the distance between the action variable of the perturbed and unperturbed Hamiltonians and their time of closeness depend both on the steepness indices: the more the function is steep (so the smaller the indices are), the more the action variables are close to each other and their time of closeness increases.
The class of "steepest" functions consists of quasi-convex functions that have index exactly equal to one.
We remind that in perturbed system with quasiconvex unperturbed Hamiltonians the numbers a and b defined in (1.6) characterizing theorem 1.2.2 assume the simple form[4]:

$$a = b = \frac{1}{2n}$$

this means that with such a and b the estimates (1.7) and (1.8) are the best.

We have the following

Proposition 2.2.1. *Let $h : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a regular C^2 function. If the restriction of $h''(I)$ to any linear space V^k orthogonal to $\nabla h(I)$ with $\dim V^k = k$ is non-degenerate[5], then it is $\delta_k = 1$.*

Proof. Since $P_{V^k} \nabla h(I) = 0$, then there are positive constants $c_2 > 0$ and $\widetilde{\xi} > 0$ such that we can calculate the Taylor Series and obtain

$$P_{V^k} \nabla h\,(I + \xi u) = \xi P_{V^k} h''(I) u + h_3$$

where it is that $|h_3| < c_2 \xi^2$ for every V^k, for every $u \in V^k$ with unit norm and for every $\xi \leq \widetilde{\xi}$. Now, if the restriction of $h''(I)$ to any $k-$dimensional subspace V^k is non-degenerate, then it exists $C > 0$ such that $|P_{V^k} h''(I) u| \geq C$ for any u and for any V^k. Then

$$|P_{V^k} \nabla h(I + \xi u)| \geq \xi \,|P_{V^k} h''(I) u| - |h_3| \geq \frac{C}{2} \xi$$

is verified as soon as $\xi \leq \frac{C}{2c_2}$ and to let it be true we just have to choose $\widetilde{\xi} \leq \frac{C}{2c_2}$.
Therefore, it is $\delta_k = 1$ □

[4] See [15] or [23]

[5] A quadratic form q on the field of real numbers and defined on a finite-dimensional vector space V is said to be non-degenerate if the determinant of the matrix associated to the form q is nonzero, that is, the restriction of $h''(I)$ does not have null eigenvalues.

The following is the result for quasi-convex functions which can be seen as a trivial consequence of Proposition 2.2.1.

Corollary 2.2.2. *Let $h : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a C^2 and almost everywhere regular function and assume that exists $\lambda_2 > 0$ such that*

$$\lambda_2 := \min_{\substack{u \in \nabla h(I)^\perp \\ \|u\|=1}} |h''(I)u \cdot u| > 0$$

that is, h is quasi-convex. Then h is steep with uniform steepness index equal to 1.

Proof. Apply Theorem 2.2.1 to h where V^k is the $(n-1)$-dimensional subspace $\nabla h(I)^\perp$. □

Then the quasi-convex functions are the steepest, that is, their index is minimal. We could now think that if the function loose regularity then the steepness index increases. In this context we can state the following

Lemma 2.2.3. *Let $h : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function differentiable $\alpha \in \mathbb{N}_{>0}$ times at a point $I \in A$. Let*

$$\hat{h}_{\alpha+1}(\eta\bar{u}) := \sum_{m=0}^{\alpha+1} \frac{1}{m!}\eta^m \mathrm{d}^m h(I)[\bar{u}]^m \tag{2.10}$$

where $[\bar{u}]^m = \bar{u} \cdot \bar{u} ... \cdot \bar{u}$ m times, be the Taylor polynomial[6] of order $\alpha + 1$ of h at the point I. If $\hat{h}_{\alpha+1}$ is steep, then h is steep with steepness index at most equal to α.

Proof. We firstly remark that if h is differentiable $\alpha + 1$ times at I and for sufficiently small $1 > \eta > 0$, then

$$h\left(I + \eta\bar{u}\right) = \hat{h}_{\alpha+1}\left(\eta\bar{u}\right) + o\left(\|\eta\bar{u}\|^{\alpha+1}\right)$$

We now consider as usual a k-dimensional affine plane V^k orthogonal to $\nabla h(I)$. We set $C_m := P_{V^k}\frac{\mathrm{d}^m h(I)}{m!}$ with $m \in \{1, ..., \alpha+1\}$, then we find

$$\begin{aligned}
\|P_{V^k}\nabla h(I+\eta\bar{u})\| &\geq \left\|P_{V^k}\nabla\hat{h}_{\alpha+1}(\eta\bar{u})\right\| - o\left(\|\eta\bar{u}\|^{\alpha+1}\right) = \\
&= \left\|P_{V^k}\sum_{m=0}^{\alpha+1}\frac{1}{m+1!}\eta^m \mathrm{d}^{m+1}h(I)[\bar{u}]^m\right\| - o\left(\|\eta\bar{u}\|^{\alpha+1}\right) = \\
&= \left\|\eta C_2\bar{u} + \frac{\eta^2}{2}C_3\bar{u}\cdot\bar{u} + ... + \eta^\alpha C_{\alpha+1}[\bar{u}]^\alpha\right\| - \left\|o\left(\|\eta\bar{u}\|^{\alpha+1}\right)\right\| \geq \\
&\geq \|\eta^\alpha(C_2\bar{u} + ... + C_{\alpha+1}[\bar{u}]^\alpha)\| - \eta^{\alpha+1} \geq C\eta^\alpha - \bar{C}\eta^a = \widetilde{C}\eta^\alpha
\end{aligned}$$

[6]See Appendix A for details.

where we called $C = \|C_2\bar{u} + ... + C_{\alpha+1}[\bar{u}]^\alpha\|$ and for $\bar{C} > 0$ small enough and this is true $\forall\, \bar{u} \in V^k$ such that $\|\bar{u}\| = 1$ and for all the affine planes of dimension k. Then in particular

$$\max_{0<\eta\leq\xi} \min_{\substack{\bar{u}\in V^k \\ \|\bar{u}\|=1}} \|P_{V^k}\nabla h\,(I+\eta\bar{u})\| \geq \widetilde{C}\xi^\alpha$$

□

2.3 Three-jet non-degenerate functions

We are now ready to focus our attention on three-jet non-degenerate functions, that is, on functions which verify the definition (2.5). Non degeneracy of the three-jet of h is a sort of natural generalization of quasi-convexity and moreover it does not procede further, namely the non-degeneracy of the four-jet does not imply steepness and more conditions are needed, except for the special case $n = 2$ (See [25] or [24]).

In this section, following [8] we provide an original proof that if at a point $I \in \mathbb{R}^n$ a function is three-jet non-degenerate then it is steep.

The proof is divided into two steps. The first step is the proof of the result in the unidimensional case. In the general theorem we used simple inequalities while in the second step we bring the problem back to a unidimensional case. We need the following

Lemma 2.3.1. *Let $h : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a C^3 function that satisfies the three-jet non-degeneracy condition (2.5) at any point I contained in a compact set $B \subset D$. Then there exists $\beta > 0$ such that, for any $I \in B$, for any $0 \neq V^k \subseteq \nabla h(I)^\perp$ k-dimensional subspace orthogonal to $\nabla h(I)$ and for any versor $\bar{u} \in V^k$ at least one of the following inequalities is satisfied:*

$$\|P_{V^k}h''(I)\bar{u}\| \geq \beta \qquad \left\|P_{V^k}h^{(3)}[\cdot,\bar{u},\bar{u}]\right\| \geq \beta \tag{2.11}$$

Proof. The function

$$M(I,u) := \max\left\{\, |h'(I)\bar{u}|\,, |h''(I)\bar{u}\cdot\bar{u}|\,, \left|\sum_{i,j,k}\frac{\partial^3 h}{\partial I_i\partial I_j\partial I_k}(I)u_iu_ju_k\right|\right\}$$

is continous and the domain is compact since M is defined in

$$B \times \left\{\bar{u} \in V^k \text{ such that } \|\bar{u}\| = 1\right\}$$

then it has a minimum β and by the three-jet non-degeneracy condition it is $\beta > 0$ with the same β that appears in equation (2.11). Indeed, for any linear space $V^k \subseteq \nabla h(I)^\perp$ and for every $\bar{u} \in V^k$ with unit norm, it is $\nabla h(I)\bar{u} = 0$ and

$$\|P_{V^k} h''(I)\bar{u}\| \geq \|P_{V^k} h''(I)\bar{u} \cdot \bar{u}\| \geq \|h''(I)\bar{u} \cdot \bar{u}\|$$

and

$$\left\|P_{V^k} h^{(3)}[\cdot, \bar{u}, \bar{u}]\right\| \geq \left\|P_{V^k} h^{(3)}[\cdot, \bar{u}, \bar{u}] \cdot u\right\| \geq \left\|h^{(3)}[\bar{u}, \bar{u}, \bar{u}]\right\|$$

□

Now we use the result of Lemma 2.2.3, that is, we prove that the third order expansion of a three-jet non-degenerate function around a point $\bar{I}$ is steep and to obtain the following

Theorem 2.3.2. *Let $h(I_1, ..., I_n)$ be a function that satisfies the three-jet non-degenerate condition (2.5) at a point $\bar{I} \in G \subseteq \mathbb{R}^n$ with G open. Then h is steep at $\bar{I}$ with steepness indices equal to 2, that is, exists $\widetilde{\xi} > 0$ such that for any $0 < \xi \leq \widetilde{\xi}$ and any linear space $0 \neq V^k \subseteq \nabla h(I)^\perp$ of dimension $k \in \{1, ..., n-1\}$ we have*

$$\max_{0<\eta\leq\xi} \min_{\substack{\bar{u}\in V^1 \\ \|\bar{u}\|=1}} \left\|P_{V^k} \nabla h(\bar{I} + \eta\bar{u})\right\| \geq \kappa\xi^2 \tag{2.12}$$

Proof. By Lemma (2.3.1) we know that for the three-jet non degenerate function h exists the positive constant

$$\beta = \min_{\substack{\bar{u}\in V^1 \\ |\bar{u}|=1}} \max\left\{\left\|h''(\bar{I})u \cdot u\right\|, \left\|\sum_{i,j,k=1}^{n} \frac{\partial^3 h}{\partial I_i \partial I_j \partial I_k}(\bar{I}) u_i u_j u_k\right\|\right\}$$

Let us denote by $\widetilde{h}(\bar{I})$ the third order expansion of h around $\bar{I}$, that is

$$\begin{aligned}\widetilde{h}(I) = h(\bar{I}) + \nabla h(\bar{I})(I - \bar{I}) + \frac{1}{2} h''(\bar{I})(I - \bar{I}) \cdot (I - \bar{I}) + \\ + \frac{1}{6} \sum_{i,j,k=1}^{n} \frac{\partial^3 h}{\partial I_i \partial I_j \partial I_k}(\bar{I})(I - \bar{I})_i (I - \bar{I})_j (I - \bar{I}) k\end{aligned}$$

we will show that the function $\widetilde{h}$ verifies equation (2.12) and then by Lemma 2.2.3 we have the result.

1. <u>The case k=1</u>
   We know that for any linear subspace $V^k \subseteq \nabla h(I)^\perp$ and for any vector u it is

$$P_{V^k} \widetilde{h}\left(\bar{I} + \eta u\right) = \eta P_{V^k} h''\left(\bar{I}\right) u + \frac{\eta^2}{2} P_{V^k} h^{(3)}\left(\bar{I}\right)[\cdot, u, u]$$

then if we consider a subspace of dimension $k = 1$ and a vector u with unit norm, let

$$\begin{aligned} A(\eta, u) &:= \left\| \eta P_{V^1} h''(\bar{I})u + \frac{\eta^2}{2} P_{V^1} h^{(3)}(\bar{I})[\cdot, u, u] \right\| \\ B(\eta, u) &:= \left\| \eta P_{V^1} h''(\bar{I})u - \frac{\eta^2}{2} P_{V^1} h^{(3)}(\bar{I})[\cdot, u, u] \right\| \end{aligned}$$

then equation (2.12) can be explicitly written as

$$\max_{0<\eta\leq\xi} \min \{A(\eta, u), B(\eta, u)\} \geq \kappa\xi^2 \tag{2.13}$$

and the direction of V^1 can give the following results

$$\left\|P_{V^1} h''(\bar{I})u\right\| \geq \beta \quad \text{or} \quad 0 \leq \left\|P_{V^1} h''(\bar{I})u\right\| < \beta$$

We will prove the steepness of h doing estimates for all the possible directions of the unidimensional subspaces.
First of all we start taking a positive constant $M > 0$ which will appear throughout the proof, uniform with respect to u and $\bar{I}$, such that

$$M \geq \left\|h^{(3)}(\bar{I})[\cdot, u, u]\right\|$$

and in particular we can take

$$M \geq \sqrt{\sum_{i,j,k=1}^{n} \left| \frac{\partial^3 h}{\partial_i I \partial_j I \partial_k I}(\bar{I}) \right|^2} \geq \left\|h^{(3)}(\bar{I})[\cdot, u, u]\right\|$$

(a) <u>case</u> $\left\|P_{V^1} h''(\bar{I})u\right\| \geq \beta$
With these conditions on $\bar{u}$ it is

$$\begin{aligned} \min_{\substack{\bar{u}\in V^1 \\ |\bar{u}|=1}} \left\|P_{V^1} \nabla h(\bar{I} + \eta u)\right\| &= \min_{\substack{\bar{u}\in V^1 \\ |\bar{u}|=1}} \{A(\eta, \bar{u}), B(\eta, \bar{u})\} \geq \\ \geq &\eta \left\|P_{V^1} h''(\bar{I})u\right\| - \frac{\eta^2}{2} \left\|P_{V^1} h^{(3)}(\bar{I})[\cdot, u, u]\right\| \geq \eta\beta - \frac{\eta^2}{2} M \end{aligned}$$

where we took the constant $M > 0$ introduced before. Then if

$$\frac{\eta^2 M}{2} \leq \frac{\eta\beta}{2}$$

that is if $\eta \leq \frac{\beta}{M}$, then we now want that the last inequality is satisfied by every $\eta \leq \xi \leq \widetilde{\xi}$ so we have to take

$$\widetilde{\xi} \leq \frac{\beta}{M} \tag{2.14}$$

and then

$$\min_{\substack{\bar{u}\in V^1\\|\bar{u}|=1}} \left\|P_{V^1}\nabla h(\bar{I}+\eta u)\right\| \geq \frac{\eta\beta}{2}$$

and we have $\frac{\eta\beta}{2} \geq \kappa\xi^2$ as soon as

$$\eta \geq \frac{2\kappa}{\beta}\xi^2$$

If we fix $\eta = \xi$ then the last condition becomes $\xi \geq \frac{2\kappa}{\beta}\xi^2$, so in particular we have

$$\widetilde{\xi} \leq \frac{\beta}{2\kappa} \tag{2.15}$$

We now can state that if we are in the condition (1a) then equation (2.12) is satisfied with

$$\eta \in \left(\frac{2\kappa}{\beta}\xi^2, \xi\right] \tag{2.16}$$

for every

$$\xi \leq \widetilde{\xi} \leq \beta \cdot \min\left\{\frac{1}{M}, \frac{1}{2\kappa}\right\} \tag{2.17}$$

and the maximum is reached when $\eta_\xi = \xi$.

(b) <u>case</u> $\left\|P_{V^1}h''(\bar{I})u\right\| = 0$
We know by Lemma 2.3.1 that in this case and in the following one, by the three-jet non-degeneracy, it is

$$\left\|P_{V^1}h^{(3)}(\bar{I})[\cdot,u,u]\right\| \geq \left\|h^{(3)}(\bar{I})[\cdot,u,u]\cdot u\right\| = \left\|h^{(3)}(\bar{I})[u,u,u]\right\| \geq \beta$$

Then in this case we have

$$\min_{\substack{\bar{u}\in V^1\\|\bar{u}|=1}} \{A(\eta,u), B(\eta,u)\} \geq \frac{\eta^2}{2}\left\|P_{V^1}h^{(3)}(\bar{I})[\cdot,u,u]\right\| \geq \frac{\eta^2}{2}\beta$$

where as usual

$$A(\eta,u) := \left\|\eta P_{V^1}h''(\bar{I})u + \frac{\eta^2}{2}P_{V^1}h^{(3)}(\bar{I})[\cdot,u,u]\right\|$$
$$B(\eta,u) := \left\|\eta P_{V^1}h''(\bar{I})u - \frac{\eta^2}{2}P_{V^1}h^{(3)}(\bar{I})[\cdot,u,u]\right\|$$

that is

$$\min_{\substack{\bar{u}\in V^1\\|\bar{u}|=1}} \left\|P_{V^1}\nabla h(\bar{I}+\eta u)\right\| \geq \frac{\eta^2}{2}\beta$$

and we obtain

$$\frac{\eta^2}{2}\beta \geq \kappa\xi^2$$

as soon as

$$\eta \geq \sqrt{\frac{2\kappa}{\beta}}\xi \tag{2.18}$$

and this must be satisfied by every $\eta \leq \xi$ then we have to take

$$\kappa \leq \frac{\beta}{2} \tag{2.19}$$

then we can state that in case (1b) we have to take (2.18) and (2.19) to verify equation (2.12).

(c) <u>case</u> $0 < \left\|P_{V^1}h''(\bar{I})u\right\| < \beta$
We have to do the usual estimate and recalling that

$$A(\eta, u) = \left\|\eta P_{V^1}h''(\bar{I})u + \frac{\eta^2}{2}P_{V^1}h^{(3)}(\bar{I})[\cdot, u, u]\right\|$$
$$B(\eta, u) = \left\|\eta P_{V^1}h''(\bar{I})u - \frac{\eta^2}{2}P_{V^1}h^{(3)}(\bar{I})[\cdot, u, u]\right\|$$

in this case we have

$$\min_{\substack{\bar{u}\in V^1\\ |\bar{u}|=1}} \{A(\eta, u, B(\eta, u)\}$$
$$\geq \left|\eta \left\|P_{V^1}h''(\bar{I})u\right\| - \frac{\eta^2}{2}\left\|P_{V^1}h^{(3)}(\bar{I})[\cdot, u, u]\right\|\right| \geq \kappa\xi^2$$

and then, named $a = \left|P_{V^1}h''(\bar{I})u\right|$ and $b = \left|P_{V^1}h^{(3)}(\bar{I})[\cdot, u, u]\right|$, the previous equation becomes

$$\left|\eta a - \frac{\eta^2}{2}b\right| \geq \kappa\xi^2 \tag{2.20}$$

with $0 < a < \beta \leq b \leq M$ so we now have to study this parabola and its coefficients which will provide us the estimates we need.
The function h is steep if we find at least one η smaller than ξ such that the inequality (2.20) is satisfied. The proof is divided in two further subcases because we might have that

$$\kappa\xi^2 \geq \frac{a^2}{2b} \text{ or it is } \kappa\xi^2 < \frac{a^2}{2b}$$

that is, the line $\kappa\xi^2$ might be above or below the local maximum $\eta_+ = \frac{a^2}{2b}$ of the function $F(\eta) := \left|\eta a - \frac{\eta^2}{2}b\right|$.

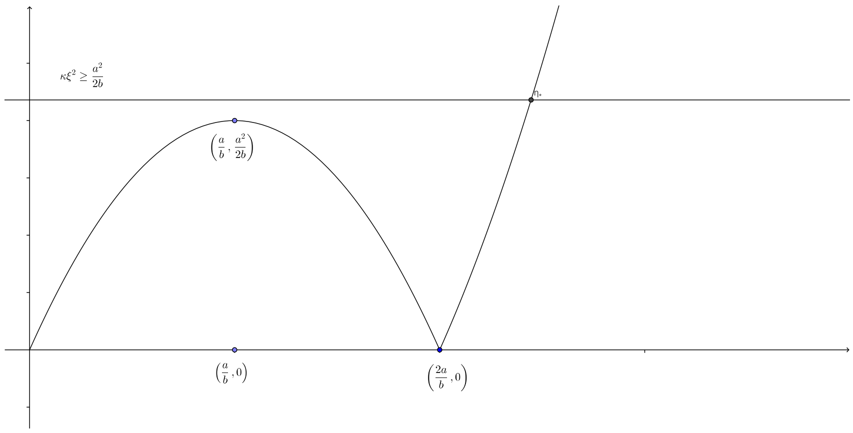

Figure 2.1: Case A

- Case A $\kappa\xi^2 \geq \frac{a^2}{2b}$ [Figure 2.1] In the first case, as one can see from figure 2.1, the first η such that equation (2.20) is satisfied is

$$\eta_\xi := \eta_*(\xi) = \frac{a + \sqrt{a^2 + 2\kappa b\xi^2}}{b} = \frac{a}{b} + \sqrt{\left(\frac{a}{b}\right)^2 + \alpha^2\xi^2}$$

where we put $\alpha = \sqrt{\frac{2\kappa}{b}}$.
So knowing that we are in case A, that is $\frac{a}{b} \leq \alpha\xi$, we find out

$$\eta_\xi < \alpha\xi + \sqrt{2\alpha^2\xi^2} = \alpha\xi\left(1 + \sqrt{2}\right) \leq \xi$$

if and only if $\alpha\left(1 + \sqrt{2}\right) \leq 1$, that is if and only if

$$\kappa \leq \frac{b}{2\left(1 + \sqrt{2}\right)^2} \tag{2.21}$$

On the other side we find

$$\eta_\xi \geq \alpha\xi$$

so in case A, under the assumption (2.21), we find the following estimate for η

$$\alpha\xi < \eta_\xi < \alpha\left(1 + \sqrt{2}\right)\xi \leq \xi \tag{2.22}$$

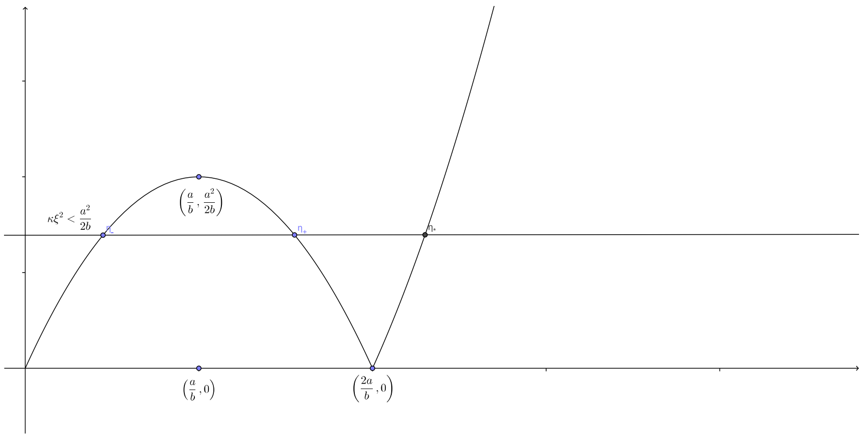

Figure 2.2: Case B

- <u>Case B</u> $\kappa\xi^2 < \frac{a^2}{2b}$ [Figure 2.2]
  In this case we can write

$$\kappa\xi^2 < \frac{a^2}{2b} \Rightarrow \frac{a}{b} \geq \sqrt{\frac{2\kappa}{b}}\xi = \alpha\xi$$

and equation (2.20) is satisfied for $\eta_- \leq \eta_\xi \leq \eta_+$ when $\eta_\xi \leq \frac{2a}{b}$ or for $\eta_\xi > \eta_*$ when $\eta > \frac{2a}{b}$ and where η_* is the same as Case A and where

$$\eta_\pm := \eta_\pm(\xi) = \frac{a \pm \sqrt{a^2 - 2\kappa b\xi^2}}{b}$$

We still have to do the last two subcases that are those depending on the width of ξ.
Indeed, if $\xi > \frac{a}{b}$, that is if ξ is larger than the local maximum of $F(\eta)$, then we obtain immediately that

$$\alpha\xi \leq \eta_\xi \leq \xi$$

if and only if $\alpha \leq 1$ that is $\kappa \leq \frac{b}{2}$ and we know that $\beta \leq b$ so we can finally choose

$$\kappa \leq \frac{\beta}{2}$$

and we fix as η_ξ that verifies the definition of steepness the number

$$\eta_\xi = \frac{a}{b}$$

On the other hand if

$$\xi \leq \frac{a}{b} \tag{2.23}$$

then we can put $\eta_\xi := \xi$ and have to show that $\eta_\xi \geq \eta_-$. Then

$$\eta_- \leq \xi^2 \iff 0 \leq a - b\xi \leq \sqrt{\left(\frac{a}{b}\right)^2 - \alpha^2\xi^2}$$

and squaring the last inequality, it is

$$(1+\alpha)^2\,\xi^2 \leq \frac{2a}{b}\xi$$

but because of (2.23), it is

$$(1+\alpha)^2\,\xi^2 \leq (1+\alpha)^2\,\frac{a}{b}\xi \leq \frac{2a}{b}\xi$$

which provides $\alpha^2 \leq 1$, so as before we choose

$$\kappa \leq \frac{\beta}{2}$$

We verified the steepness condition on unidimensional subspaces of all directions and therefore we proved the following

Lemma 2.3.3. *For any $u \in \nabla h(\bar{I})^\perp$ vector with unit norm and for any $0 < \xi \leq \widetilde{\xi}$ with $\widetilde{\xi} \leq \beta \cdot \min\left\{\frac{1}{M}, \frac{1}{2\kappa}\right\}$ and with κ depending on β and on the direction of u, there exists $\eta_{\xi,u}$*

$$\min\left\{\sqrt{\frac{2\kappa}{\beta}}, 1\right\}\xi \leq \eta_{\xi,u} \leq \xi$$

such that

$$\left\|\nabla\widetilde{h}(\bar{I} \pm \eta_{\xi,u}u)\cdot u\right\| \geq \kappa\xi^2 \tag{2.24}$$

2. The case $k > 1$.
   To prove the theorem on subspaces of higher dimension, we want to show that we are able to bring the high-dimensional problem back to a unidimensional one and to do this we firstly have to state a simple Lemma. It is known that the restriction of the Hessian operator of a three-jet non-degenerate function to any linear space $V^k \subseteq \nabla h(\bar{I})^\perp$ may be degenerate only on a space of dimension one. Furthermore, $k-1$ of its eighenvalues satisfy uniformly in k the estimate provided by the following

Lemma 2.3.4. *Let $V^k \subseteq \nabla h(\bar{I})^{\perp}$ and let $\{\bar{e}_1, ..., \bar{e}_k\}$ be an orthonormal set of eigenvectors of the restriction of the Hessian operator $h''(\bar{I})$ to the space*

$$P_{V^k} h''(\bar{I}) P_{V^k} : V^k \to V^k$$

Let also $\mu_1, ..., \mu_k$ be the respective eigenvalues such that

$$|\mu_1| \leq |\mu_2| \leq ... \leq |\mu_k|$$

Then, it is

$$|\mu_2|, ..., |\mu_k| \geq \beta$$

that is, at most one eigenvalue is smaller than β in modulus.

Proof. Let suppose that there exists two eigenvalues $|\mu_1| \leq |\mu_2|$ such that they are both smaller than β in modulus.
Then, for any $\gamma \in [0, 2\pi]$, we see that the vector

$$\bar{u}_\gamma = \cos(\gamma)\bar{e}_1 + \sin(\gamma)\bar{e}_2$$

satisfies

$$\begin{aligned} &\left\| h''(\bar{I})[u_\gamma, u_\gamma] \right\| = \left\| P_{V^k} h''(\bar{I}) P_{V^k} u_\gamma \cdot u_\gamma \right\| = \\ &= \left| \cos^2(\gamma)\mu_1 + \sin^2(\gamma)\mu_2 \right| \leq |\mu_2| \leq \beta \end{aligned}$$

and therefore by Lemma 2.3.1 it is $\left\| h^{(3)}(\bar{I})[u_\gamma, u_\gamma, u_\gamma] \right\| \geq \beta$ for every γ. But it is a contradiction. In fact $h^{(3)}(\bar{I})[u_\gamma, u_\gamma, u_\gamma]$ is continous in γ and if we have $\left\| h^{(3)}(\bar{I})[u_\gamma, u_\gamma, u_\gamma] \right\| > 0$ for some γ then it is also true that

$$\begin{aligned} &h^{(3)}\left(\bar{I}\right)[u_{\gamma+\pi}, u_{\gamma+\pi}, u_{\gamma+\pi}] = h^{(3)}\left(\bar{I}\right)[u_{-\gamma}, u_{-\gamma}, u_{-\gamma}] = \\ &\quad = -h^{(3)}\left(\bar{I}\right)[u_\gamma, u_\gamma, u_\gamma] < 0 \end{aligned}$$

and by continuity there exists a point γ_* such that

$$h^{(3)}\left(\bar{I}\right)[u_{\gamma_*}, u_{\gamma_*}, u_{\gamma_*}] = 0$$

$\square$

This lemma allows us to state that the direction of $\bar{e}_1$ related to the minimun eigenvalue is the only possible direction of degeneration for $P_{V^k} h''(\bar{I}) P_{V^k}$.
Then following the notation of Lemma 2.3.3 we advance the proof setting for any

$$0 < \xi < \bar{\xi} \qquad \eta := \eta_{\xi, \bar{e}_1} \in (c\xi, \xi]$$

with $c = \min\left\{\sqrt{\frac{2\kappa}{\beta}}, 1\right\}$.
We remark that we are working with the truncated series $\widetilde{h}$; for sake of simplicity of notation we omitted the symbol $\tilde{}$ so we write h instead of $\widetilde{h}$.

For any $u = \sum_{j=1}^{k} x_j \bar{e}_j \in V^k$ and for any $j \in \{1, ..., k\}$it is

$$\left\|P_{V^k} \nabla h(\bar{I} + \eta u)\right\| \geq \|\nabla h(\bar{I} + \eta u) \cdot \bar{e}_j\| \geq \eta |\mu_j||x_j| - \frac{\eta^2}{2} M \tag{2.25}$$

as well as

$$\left\|P_{V^k} \nabla h(\bar{I} + \eta u)\right\| \geq \frac{\eta}{k} \sum_{j=1}^{k} |\mu_j||x_j| - \frac{\eta^2}{2k} M \tag{2.26}$$

We now fix a positive constant $N \geq \frac{M}{\beta}$ and we firstly suppose that if $u = \sum_{j=1}^{k} x_j \bar{e}_j$ satisfies

$$\sum_{j=2}^{k} x_j^2 \geq N^2 \eta^2$$

then we have

$$\sum_{j=2}^{k} |\mu_j||x_j| \geq \beta N \eta$$

and this last inequality follows from the following

$$\left(\sum_{j=2}^{k} |\mu_j||x_j| \right)^2 \geq \sum_{j=2}^{k} |\mu_j|^2 |x_j|^2 \geq \beta^2 \sum_{j=2}^{k} x_j^2 \geq \beta^2 N^2 \eta^2$$

after having computed the square root on both sides.
Then from Equation 2.26 and summing from $j = 2$ to k we obtain

$$\left\|P_{V^k} \nabla h(\bar{I} + \eta u)\right\| \geq \frac{\eta}{k} \sum_{j=2}^{k} |\mu_j||x_j| - \frac{\eta^2}{2k} M \geq \left(\frac{\beta N}{k} - \frac{M}{2k} \right) \eta^2 \geq \frac{\beta N}{2k} \eta^2$$

and the last inequality follows from our choice of N. Now we apply this inequality with $\eta = \eta_\xi$.
We now go on with the other case

$$\sum_{j=2}^{k} x_j^2 < N^2 \eta^2$$

which means in particular that

$$\begin{aligned} &|x_2|, ..., |x_j| < N\eta \\ &(1 - |x_1|) \leq \frac{N^2\eta^2}{1+|x_1|} \leq N^2\eta^2 \end{aligned} \tag{2.27}$$

where the first inequality is an obvious consequence of our choice of the coefficients x_j while the second one is true because we have

$$x_1^2 = 1 - \sum_{j=2}^{k} x_j^2 \geq 1 - N^2\eta^2$$

which provides

$$(1 - x_1)(1 + x_1) \leq N^2\eta^2 \Rightarrow (1 - |x_1|) \leq \frac{N^2\eta^2}{1+|x_1|} \leq N^2\eta^2$$

Now recalling (2.25) we have in particular

$$\left\|P_{V^k}\nabla h(\bar{I}+\eta u)\right\| \geq \left\|\nabla h\left(\bar{I}+\eta u\right)\cdot \bar{e}_1\right\|$$

We attempt to find an estimate along the degenerate direction $\bar{e}_1$.
It is

$$\begin{aligned}
&\nabla h\left(\bar{I}+\eta u\right)\cdot\bar{e}_1 - \nabla h\left(\bar{I}+\eta x_1\bar{e}_1\right)\cdot\bar{e}_1 = \\
&= \eta h''\left(\bar{I}\right) u\cdot\bar{e}_1 + \frac{\eta^2}{2}h^{(3)}\left(\bar{I}\right)[\bar{e}_1, u, u] - \eta x_1 h''(\bar{I})\bar{e}_1\cdot\bar{e}_1 - \frac{\eta^2}{2}x_1^2 h^{(3)}[,\bar{e}_1,\bar{e}_1,\bar{e}_1] = \\
&= \eta x_1 h''(\bar{I})\bar{e}_1\cdot\bar{e}_1 + \eta\sum_{j=2}^{k} x_j h''(\bar{I})\bar{e}_j\cdot\bar{e}_1 + \frac{\eta^2}{2}h^{(3)}(\bar{I})[\bar{e}_1, u, u] \\
&- \eta x_1 h''(\bar{I})\bar{e}_1\cdot\bar{e}_1 - \frac{\eta^2}{2}x_1^2 h^{(3)}[\bar{e}_1,\bar{e}_1,\bar{e}_1] = \\
&= \frac{\eta^2}{2}h^{(3)}(\bar{I})[\bar{e}_1, u, u] - \frac{\eta^2}{2}x_1^2 h^{(3)}[\bar{e}_1,\bar{e}_1,\bar{e}_1] = \\
&= \frac{\eta^2}{2}x_1\sum_{j=2}^{k} x_j h^{(3)}(\bar{I})[\bar{e}_1,\bar{e}_1\bar{e}_j] + \frac{\eta^2}{2}\sum_{i,j=2}^{k} x_i x_j h^{(3)}(\bar{I})[\bar{e}_1,\bar{e}_i,\bar{e}_j]
\end{aligned}$$

so we have

$$\begin{aligned} &\left\|\nabla h(\bar{I}+\eta u)\cdot\bar{e}_1 - \nabla h(\bar{I}+\eta x_1\bar{e}_1)\right\| \leq \frac{\eta^2}{2}kN\eta M + \frac{\eta^2}{2}kN^2\eta^2 M \\ &\leq \eta^3 kNM + \eta^4 kN^2 M \leq 2\eta^3 kNM \end{aligned} \tag{2.28}$$

as soon as $\eta \leq \frac{1}{N}$ so we have to take $\widetilde{\xi} \leq \frac{1}{N}$ and recalling the choice we did of $N \geq \frac{M}{\beta}$ we obtain that we must take

$$\widetilde{\xi} \leq \frac{\beta}{M} \tag{2.29}$$

Let us now consider σ such that $\sigma = 1$ if $x_1 \geq 0$ and $\sigma = -1$ if $x_1 < 0$ and knowing (2.27) do the following estimate

$$\begin{aligned} &\left\|\nabla h\left(\bar{I} + \eta x_1 \bar{e}_1\right) \cdot \bar{e}_1 - \nabla h\left(\bar{I} + \eta \sigma \bar{e}_1\right) \cdot \bar{e}_1\right\| \leq \\ &\leq \eta\left(1 - |x_1|\right)\mu_1 + \frac{\eta^2}{2}\left(1 - x_1^2\right) M \\ &\leq \eta^3 N^2 \beta + \eta^4 N^2 M \leq 2\eta^3 N^2 \beta \end{aligned}$$

where we supposed as before that $\widetilde{\xi} \leq \frac{\beta}{M}$.
Now using (2.28) and $N > \frac{M}{\beta}$ we obtain[7]

$$\begin{aligned} &\left\|P_{V^k} \nabla h\left(\bar{I} + \eta u\right)\right\| \geq \left\|\nabla h\left(\bar{I} + \eta u\right) \cdot \bar{e}_1\right\| \geq \\ &\geq \left\|\nabla h\left(\bar{I} + \eta x_1 \bar{e}_1\right) \cdot \bar{e}_1\right\| - 2\eta^3 k N M \geq \\ &\geq \left\|\nabla h\left(\bar{I} + \eta \sigma \bar{e}_1\right) \cdot \bar{e}_1\right\| - 2\eta^3 N (k+1) M \end{aligned}$$

and for Lemma 2.3.3 when $\eta = \eta_{\xi,u}$ it is

$$\left\|P_{V^k} \nabla h\left(\bar{I} + \eta_{\xi,\bar{e}_1} u\right)\right\| \geq \kappa \xi^2 - 2\xi^3 N\left(k+1\right) M \geq \frac{\kappa}{2}\xi^2$$

as soon as

$$\xi \leq \frac{\kappa}{2N\left(k+1\right) M}$$

which is always satisfied if we choose

$$\widetilde{\xi} \leq \frac{\kappa}{2N\left(n+1\right) M} \tag{2.30}$$

Summarizing the proof in the higher-dimensional case which we bring to a unidimensional case (direction $\bar{e}_1$) we proved that for $\eta = \eta_{\xi,\bar{e}_1} \in (c\xi, \xi]$ and for $u \in V^k \subseteq \nabla h(\bar{I})^{\perp}$ such that $u = \sum_{j=1}^k x_j \bar{e}_j$ with $\sum_{j=1}^k x_j^2 = 1$, setting $N > \frac{M}{\beta}$ then it may happen that

$$\sum_{j=2}^{k} x_j^2 \geq N^2 \eta^2$$

[7] We recall that if $|x - y| \leq c$ with $c > 0$ then by reverse triangle inequality it is $|x| \geq |y| - c$

and then

$$\left\|P_{V^k}\nabla h\left(\bar{I}+\eta_{\xi,\bar{e}_1}u\right)\right\| \geq \frac{\beta N}{2k}\eta^2 \geq \frac{\beta N}{2k}c^2\xi^2 \tag{2.31}$$

otherwise it may happen that

$$\sum_{j=2}^{k} x_j^2 < N^2\eta^2$$

and then

$$\left\|P_{V^k}\nabla h\left(\bar{I}+\eta_{\xi,\bar{e}_1}u\right)\right\| \geq \frac{\kappa}{2}\xi^2 \tag{2.32}$$

and the steepness definition is verified.

$\square$

Chapter 3

Real analytic steep functions

In this chapter we mainly want to enunciate and prove a theorem that joins regular real analytic functions and the steepness property.
The connection occurs when the restriction of the analytic function to a subspace of lower dimension admits only isolated critical points.
The original idea and the first proof that provided the connection between this apparently different properties was Il'yashenko's in [13]. He proved for the first time that the presence of only isolated critical points to the restriction implied the steepness of a function. He also provided a lower estimate for the steepness indices depending on the numer of critical points for the restriction counted with their multiplicity. To be exact, he proved that the index δ_k related to the k-dimensional subspace can be taken greater than or equal to the number of critical points of the restriction to that subspace, counted with multiplicity.

In the subsequent work [21], that we decided to follow, the approach to the proof is different and it is also improved the estimate on the steepness indices. This aim is achieved since the steepness index is taken as a number, know as *Lojasiewicz's exponent* of two functions (see [5], [6] and [10]). This exponent is a real number that in some specific cases is rational [10] and it is related to two functions: in our context the first will be the projection of the gradient and the second will be the distance function to the isolated critical point.
We will not calculate these exponents in this thesis and in general their computation is very difficult. In the real analytic context the computation is a bit easier since it can be exploit the Taylor series of the function and the exponent is related to the series.
In [21] a concrete example can be found where the estimates on the indices

following the two different methods are strictly different. This show that it is convenient to consider, when computable, the Lojasiewicz's exponent instead of the number of isolated critical points.

3.1 Main theorem and geometric tools

To prove the main theorem, as we have already observed in the Introduction, we use properties and results of subanalytic geometry (See [5], [26] and [12] for details).
We refer to the previous chapter for all the definitions.

Remark 3.1.1. Definition 2.1.3 implies definition 2.1.4.

Proof. If definition 2.1.3 is verified then it is true at some point $\bar{\xi} \in (0, \widetilde{\xi_k}]$, so an arbitrary continous curve $\gamma : I \subseteq \mathbb{R} \to G$ connecting I to a point at a distance $\xi \leq \widetilde{\xi_k}$ crosses a sphere of radius $0 < \eta \leq \xi$ centered at I at time t_* such that $\gamma(t_*) = I + \bar{\xi}\bar{u}$ and this is equivalent to definition 2.1.4.

□

It is important to emphasize that definition 2.1.4 does not imply definition 2.1.1 if the real analyticity condition is omitted.
In section 4.3 we provide a concrete example of a C^1 but not real analytic function that is arc-steep in its domain but does not verify the original definition 2.1.1. In this context of real analyticity, we prove that 2.1.3 and 2.1.4 are equivalent.
We state the main

Theorem 3.1.1. *Let $D \subseteq \mathbb{R}^n$ be an open subset and let $h : D \to \mathbb{R}$ be a real analytic function. Then on any compact subset $K \subset D$, h is steep on K in the sense of Definition 2.1.3 if and only if h is NNI non-degenerate in the sense of Definition 2.1.9.*

We introduce now those geometric tools which will attend in the proof of Theorem 3.1.1.
The following definitons and results are taken from subanalytic geometry so they are valid in general manifolds. In this context we will always suppose that the manifolds are real analytic, that is, that the charts and the transition maps[1] are real analytic.

Subanalytic geometry is based on the concept of subanalytic sets which are a sort of extension of semi-analytic sets (everything will be defined in a while). More precisely the class of subanalytic sets has good properties and it is the

[1]See Appendix D for the definitions of charts and transition maps.

smallest one that contains all semi-analytic sets that is closed under the operation of taking an image by proper real analytic maps[2]. This property is essential since we need to keep real analyticity.

The first definition is the number that will provide the steepness index

Definition 3.1.1. *Let M be a real analytic manifold and let K be a compact subset of M and let f, g be two vector-valued functions continous over K. We set:*

$$E_K(f,g) = \Big\{ \alpha \in \mathbb{R}_+ \textit{ such that } \exists C > 0 \\ \textit{such that } \|f(u)\|^\alpha \leq C\|g(u)\|, \forall u \in K \Big\}$$

$$\alpha_K(f,g) = \inf \{E_K(f,g)\}$$

$\alpha_K(f,g)$[3] is called the Lojasiewicz's exponent of f with respect to g over K.

Remark 3.1.2. In this context, we are interested in the case where f is defined on a compact subset K of $\mathbb{R}^n$ and admits an isolated zero at a point $\bar{x} \in K$ then we set:

$$\alpha_{\bar{x}}(f) = \inf \Big\{ \alpha \in \mathbb{R}_+ \text{ such that } \exists C > 0,\ r > 0 \\ \text{with } \|f(x)\| \geq C\|x - \bar{x}\|^\alpha \text{ if } \|x - \bar{x}\| \leq r \Big\}$$

that is, $\alpha_{\bar{x}}(f) = \alpha_K(f, d(\cdot, \bar{x}))$ where $d(\cdot, \bar{x})$ is the distance function from the point $\bar{x}$. From now on, when we talk about the Lojasiewicz's exponent of f at a point $\bar{x}$[4] we will assume it is calculated with respect to the distance function from the point $\bar{x}$.

It is natural to presume that the Lojasiewicz's exponent is candidated to be the steepness index of a steep function. More precisely, we prove that the Lojasiewicz's exponent for the function $f(\eta u) = \|P_{V^k} \nabla h(I + \eta u)\|$ is the index δ_k for the function h.

Let now M be a real analytic manifold and let U be an open subset of M[5]

[2] A map $f : X \to Y$ between two topological spaces is said to be proper if the preimage of every compact set in Y is compact in X.

[3] We set $\inf\{\emptyset\}$ defined as $+\infty$

[4] It is often implied that $\bar{x} = 0$

[5] The definitions come from [5].

Definition 3.1.2. *A subset $X \subseteq M$ is* **semianalytic** *if each $a \in M$ has a neighbourhood U such that*

$$X \cap U = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} X_{i,j}$$

where each set $X_{i,j}$ is defined by either $\{f_{i,j} > 0\}$ or $\{f_{i,j} = 0\}$ for some $f_{i,j}$, real analytic functions defined on U. It is said that X is described by the functions $\{f_{i,j}\}$.

Definition 3.1.3. *A subset $X \subset M$ is said to be* **subanalytic** *if each point of M admits a neighbourhood U such that $X \cap U$ is a projection of a relatively compact semianalytic set A, i.e., there exists a real analytic manifold N and a relatively compact semianalytic set $A \subset M \times N$ such that $X \cap U = \Pi(A)$ with the canonical projection Π from $M \times N$ to N.*

Remark 3.1.3. Basically, a subset X of a real analytic manifold M is subanalytic if it is a projection of a relatively compact semianalytic set.
Equivalently, it is the image of the projection of a semianalytic subset A of an $(m+n)$-dimensional manifold to an n-dimensional manifold.

Remark 3.1.4. All such definitions are local.

Definition 3.1.4. *Let $X \subset M$ and let N be a real analytic manifold. A map $f : X \to N$ is subanalytic if its graph $G(f) := \{(x, f(x)) | x \in X\}$ is a subanalytic subset of $M \times N$.*

We now need some properties and results for subanalytic functions.
We refer to [5] and [12] for the proofs.

Proposition 3.1.2. *Being subanalytic is closed under the following properties:*

1. *Finite union.*
2. *Finite intersection.*
3. *The difference of any two.*
4. *The image of a relatively compact set by a subanalytic mapping.*

Proposition 3.1.3. *If A is subanalytic in a real analytic space X, then its clousure in X is also subanalytic and also every connected component of A.*

Theorem 3.1.4. *i)* (Theorem of the complement) *If M is a real analytic manifold and X is a subanalytic set of M then $M \setminus X$ is subanalytic.*

ii) For a numerical function f continous subanalytic over a real analytic manifold M, the set $X = \{x \in M \mid f(x) > 0\}$ is a subanalytic set. Indeed, X is the projection of the intersection of the graph of f with the subset $\{(x, y) \in M \times \mathbb{R} \mid y > 0\}$.

Remark 3.1.5. Let X be a subanalytic subset of $\mathbb{R}^n$. Then the distance function $d(x, X) = \min_{z \in \bar{X}} |x - z|$ is continous and subanalytic while it is not analytic even if the set X is analytic.

We have this result that characterize the clousure of this "set of sets" under proper analytic maps [26]

Theorem 3.1.5.

i) If A and B are subanalytic subsets of real analytic manifolds M and N respectively, then $A \times B$ is a subanalytic subset of $M \times N$.

ii) Let $g : M \to N$ be an analytic map and let A be a subanalytic subset of M such that $g : \overline{A} \to N$ is proper. Then $g(A)$ is a subanalytic subset of N.

Lemma 3.1.6. *Let $X \subseteq M$ and $K \subset N$ be two subanalytic subsets of the real analytic manifolds M and N with K compact. If $f : X \times K \to \mathbb{R}$ is a continous subanalytic function then*

$$m(x) := \min_{u \in K} (f(x, u)) \quad and \quad M(x) := \max_{u \in K} (f(x, u))$$

are continous subanalytic.

Proof. From Theorem 3.1.4 ii), the set

$$A = \{(x, u, v) \in M \times K \times K \mid f(x, u) > f(x, v)\}$$

is subanalytic since it can be written as

$$A = \{(x, u, v) \in M \times K \times K \mid g(x, u, v) > 0\}$$

where $g(x, u, v) = f(x, u) - f(x, v)$ is a continous subanalytic function . Let $\Pi : M \times K \times K \to M \times K$ be the projection on the first two components,i.e. $\Pi(x, u, v) = (x, u)$, then according to Theorem 3.1.4 $\Pi(A)$ is subanalytic with its complement $B = M \times K \setminus \Pi(A)$. B has been defined in a way such that if $(x, u) \in B$ then $f(x, u) = m(x)$ and the graph of m is subanalytic since it is the image of B through the map $F(x, u) = (x, f(x, u))$ which has subanalytic components.
The proof that $M(x)$ is subanalytic is analogous. □

We now state two classical results of subanalytic geometry. The first Lemma leads to a part of the proof of Theorem 3.1.1 while the second one ensures the existance of the Lojasiewicz's exponent for a subanalytic function. We prove only theorem 3.1.8. For other proofs see [12] or [16].

Lemma 3.1.7. Curve selection Lemma
Let X be a subanalytic set of a real analytic manifold M and let $x \in X$ be an accumulation point (i.e. $x \in \bar{X}$). Then there exists $\varepsilon > 0$ and a non-constant real analytic map

$$\gamma : (-\varepsilon, \varepsilon) \to M$$

such that $\gamma(0) = x$ and $\gamma\left((0, \varepsilon)\right) \subset X$.

Theorem 3.1.8. Lojasiewicz's inequality
Let M be a real analytic manifold and let K be a compact subset of M. Let then $f, g : K \to \mathbb{R}$ be subanalytic functions. If $f^{-1}(0) \subseteq g^{-1}(0)$ then there exist constants $C, r > 0$ such that for every $x \in K$

$$\|f(x)\| \geq C\|g(x)\|^r \tag{3.1}$$

As we have already remarked, if $M = \mathbb{R}^n$ and if we set $Z = f^{-1}(0)$ and $g(x) := dist(x, Z)$ with $x \in K$, then we get

$$\|f(x)\| \geq C dist(x, Z)^r$$

If we suppose that f has an isolated zero (at the origin for simplicity) we get $g(x) = \|x\|$ and the inequality becomes

$$\|f(x)\| \geq C\|x\|^r$$

Proof. We consider the set

$$L := \left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = |g(x)|, x_2 = |f(x)| \; x \in K\right\}$$

and $\pi_1(x_1, x_2) = x_1$ the projecion on the first component.
We first need a Lemma. For the proof see [5].

Lemma 3.1.9. *Let M be a real analytic manifold*

1. *Let L be a subanalytic subset of M. If $\dim L \leq 1$ then X is semianalytic.*

2. *Let $L \subset M$ be a one-dimensional semianalytic subset and let $a \in \bar{L}$. Assume that $L \setminus \{a\}$ is locally connected. Then there exist $\varepsilon > 0$ and a real analytic map $\gamma : (-\varepsilon, \varepsilon) \to M$ such that $\gamma(0) = a$ and $\gamma((0, \varepsilon))$ is a neighbourhood of a in $L \setminus \{a\}$.*

By this Lemma the set L that we defined above is a semianalytic subset of $\mathbb{R}^2$. We can assume that $0 \in \pi_1(L)$ and that 0 is not isolated in $\pi_1(L)$. Then by the second point Lemma 3.1.9, exist $\varepsilon > 0$ and a parametrized real analytic curve $\gamma(s) = (x_1(s), x_2(s))$ with $s \in (-\varepsilon, \varepsilon)$ such that $x_1(0) = 0$ and $x_1(s) > 0$ for $s > 0$ and such that $L_* := L \cap \{[0, x_1(\varepsilon)) \times \mathbb{R}\}$ is bounded by $\gamma([0, \varepsilon))$.
By a change of parameter we can assume that $x_1(s) = s^k$ for some $k > 0$. Then, by hypothesis, $x_2(s) > 0$ for every $s \in (0, \varepsilon)$ since for every $x_1 \in (0, \varepsilon^k)$ the set $\{x \in K : |g(x)| = x_1\}$ is compact and $|f(x)|$ does not vanish on it. In particular, $v(s)$ has a nonzero minimum.
Now, since L_* is bounded below by γ, then $|f(x)| \geq v(s)$ and $s^k = |g(x)|$ for some $x \in K$ and then $|f(x)| \geq v\left(|g(x)|^{\frac{1}{k}}\right) > 0$.
Thank to the analitycity of v, we verify that exist $C, r > 0$ such that $|f(x)| \geq C|g(x)|^r$, whenever $0 < |g(x)| \leq \varepsilon^k$.

□

3.2 Steepness and subanaytic functions

Now that we have a general idea of subanalytic sets, we have to give estimates on the growth of a subanalytic function that has an isolated zero. We consider a positive subanalytic function for which we prove an important lower estimate that leads to the proof of Theorem 3.1.1. The main idea is to obtain general estimates and apply them to the function $\|P_{V^k} \nabla h(I + \eta u)\|$.

For $\widetilde{\xi} > 0$ let us consider a compact analytic manifold K, the closed ball $\overline{B_{\widetilde{\xi}}(0)} \subset \mathbb{R}^n$ and a continous subanalytic function $\Phi : \overline{B_{\xi}(0)} \times K \to \mathbb{R}^+$ such that, for every fixed $y \in K$ and set

$$\Phi_y := \Phi(\cdot, y) \tag{3.2}$$

then either $\Phi_y(0) = 0$ and this zero is isolated[6] or $\Phi_y(0) > 0$. For this kind of functions we want to find constants $C > 0$ and $\delta > 0$ such that $\Phi_y(x) \geq C\|x\|^\delta$ for every $x \in \overline{B_{\widetilde{\xi}}(0)}$ and for a fixed parameter $y \in K$.

[6] a is said to be an isolated zero for f if $f(a) = 0$ and if there exists an open neighbourhood U for a such that $f(x) \neq 0$ for every $x \in U \setminus \{a\}$.

Proposition 3.2.1. *The functions*

$$m(\xi, y) = \min_{\|x\|=\xi} (\Phi(x, y)) \tag{3.3}$$

and

$$M(\xi, y) = \max_{0 \leq \eta \leq \xi} (m(\eta, y)) \tag{3.4}$$

are continous subanalytic on the set $[0, \widetilde{\xi}] \times K$.

Proof. Let us consider the function[7] $\varphi : [0, \widetilde{\xi}] \times K \times S^{n-1} \to \mathbb{R}^+$ such that $\varphi(\xi, y, \theta) = \Phi(\xi\theta, y)$. Then

$$m(\xi, y) := \min_{\theta \in S^{n-1}} (\varphi(\xi, y, \theta))$$

is continous and subanalytic by Lemma 3.1.6.
Then we set

$$f(\xi, y, t) : [0, \widetilde{\xi}] \times K \times [0, 1] \to \mathbb{R}^+$$

such that $f(\xi, y, t) = m(t\xi, y)$ and then always by Lemma 3.1.6 is continous subanalytic the function

$$M(\xi, y) := \max_{t \in [0,1]} (f(\xi, t, y)) = \max_{\eta \in [0,\xi]} (m(\eta, y))$$

□

We are now ready to state and prove a key theorem that allows a clear understanding of the main Theorem 3.1.1.

Theorem 3.2.2. *Let* $K \subseteq \mathcal{M}$ *be a compact subset of a real analytic manifold* $\mathcal{M}$ *and let* $\Phi : \overline{B}_\xi \times K \to \mathbb{R}^+$ *be a continuous subanalytic function. Then the following conditions are equivalent:*

i) For any fixed $y \in K$, *the function (3.2) is such that either* $\Phi_y(0) > 0$ *or* $0 \in \overline{B_{\widetilde{\xi}}(0)}$ *is an isolated zero.*

ii) There exist constants $C, \delta > 0$ *such that for every* $(\xi, y) \in [0, \widetilde{\xi}] \times K$ *the following inequality is satisfied*

$$\max_{\eta \in [0,\xi]} \left(\min_{\|x\|=\eta} \Phi(x, y) \right) \geq C\xi^\delta \tag{3.5}$$

[7]We denote by S^{n-1} the unit sphere in $\mathbb{R}^n$, i.e. $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$

iii) *There exist constants $C, \delta > 0$ such that for every fixed $y \in K$ and for any continuous curve $\gamma : I \subset \mathbb{R} \to \overline{B_{\widetilde{\xi}}(0)}$ that connects the origin to a point at a distance $d \leq \widetilde{\xi}$ there exists a point $t_* \in I$ such that $\Phi_y(\gamma(t_*)) \geq Cd^\delta$.*

Proof.

i)⇒ii) We observe that the term on the left in equation (3.5) is equivalent to (3.4) which is continuous subanalytic for Proposition 3.2.1 and if we denote with $M^{-1}(0)$ the set of zeros of M then it is that $M^{-1}(0) \subseteq \{0\} \times K$. Furthermore, $M(\cdot, y)$ is clearly nondecreasing with respect to ξ and if $M(0, y) > 0$ then $M(\xi, y) > 0$ for every $\xi \in [0, \widetilde{\xi}]$ and then ii) is verified. But if $M(0, y) = 0$ then $\Phi(0, y) = 0$ and since the origin is an isolated zero for Φ then there exists a neighbourhood $U \subseteq \overline{B_{\widetilde{\xi}}(0)}$ of the origin such that $\Phi(x, y) > 0$ for every $x \in U \setminus \{0\}$ and since M is nondecreasing it is also that $M(\xi, y) > 0$ for every $\xi \in (0, \widetilde{\xi}]$. Let then consider the continous subanalytic function $\pi_1(\xi, y) = \xi$ which is such that

$$M^{-1}(0) \subseteq \pi_1^{-1}(0) = \{0\} \times K$$

and then since $\overline{B_{\widetilde{\xi}}} \times K$ is compact we can apply Theorem 3.1.8 that ensures the existance of constants $C > 0$ and $\delta > 0$ such that

$$M(\xi, y) \geq C\xi^\delta$$

for every $(\xi, y) \in [0, \widetilde{\xi}] \times K$.

ii) ⇒ iii) See Remark 3.1.1.

iii) ⇒ i) Let suppose by contradiction that for a certain $y \in K$ there exists an accumulation of zeros at the origin and then for point 4) of Proposition 3.1.2 we have that

$$Z_y := \left\{ x \in \overline{B_{\widetilde{\xi}}(0)} \text{ such that } \Phi_y(x) = 0 \right\}$$

is a subanalytic set so we can apply Lemma 3.1.7 to find $\varepsilon > 0$ and a nonconstant analytic curve $\gamma : (-\varepsilon, \varepsilon) \to \overline{B_{\widetilde{\xi}}(0)}$ such that $\gamma(0) = 0$ and $\gamma(0, \varepsilon) \subset Z_y$ which is a contradiction since we supposed iii).
Hence iii) ⇒ i) and the theorem is proved.

□

3.3 Proof of the main theorem

We are now ready to prove Theorem 3.1.1.
Let consider a real analytic function $h : D \subseteq \mathbb{R}^n \to \mathbb{R}$ defined on an open domain D and for every $k \in \{1, ..., k\}$ we denote by $G_k(\mathbb{R}^n)$ the k-dimensional Grassmannian manifold[8], that is, the set of all the k-dimensional vector subspaces of $\mathbb{R}^n$. We denote by V^k an element[9] in $G_k(\mathbb{R}^n)$.
We set

$$\begin{aligned} &\mathcal{H}_k : \mathbb{R}^n \times D \times G_k(\mathbb{R}^n) \to \mathbb{R}^+ \\ &\mathcal{H}_k(x, x_0, V^k) := \|P_{V^k} \nabla h(x_0 + x)\| \end{aligned} \tag{3.6}$$

where $x \in V^k$ and $x_0 + x \in D$. We introduce this notation since that if an isolated critical point for the restriction of h on the affine subspace $x_0 + V^k$ coincides with an isolated zero for $\mathcal{H}_k$, then we apply Theorem 3.2.2 to $\mathcal{H}_k$ to obtain the thesis.
By hypothesis, the restriction of h to any affine subspace has only isolated critical points. Then following our new notations, if $x_0 \in D$ is a critical point for the restriction of h to the affine subspace $x_0 + V^k$, it follows that $\mathcal{H}_k$ has an isolated zero at the point $(0, x_0, V^k)$.
Let now consider a compact subset $K \subset D$ such that $x_0 \in K$. Then $\mathcal{H}_k$ is defined also on the restriction

$$\left\{ \left(x, x_0, V^k\right) \in \overline{B}_{\widetilde{\xi}} \cap V^k \times K \times G_k(\mathbb{R}^n) \right\} \tag{3.7}$$

where $\overline{B}_{\widetilde{\xi}}$ is the closed ball centered at the origin of $\mathbb{R}^n$ and $\widetilde{\xi} = dist(K, \mathbb{R}^n \setminus D)$ is its radius. We highlight that this choice of $\widetilde{\xi}$ ensure us that $x + x_0 \in D$ for every $x \in \overline{B}_{\widetilde{\xi}}$
We know (See Appendix D) that the Grassmannian is a compact smooth manifold and that a fiber over any element $V^k \in G_k(\mathbb{R}^n)$ is given by all the k-tuples of linearly indipendent orthonormal vectors in $\mathbb{R}^n$, that is, by[10] the open set in $\mathbb{R}^{nk}$ denoted by $V_k^0(\mathbb{R}^n)$. Then, since $G_k(\mathbb{R}^n)$ is compact, there exists for every $V^k \in G_k(\mathbb{R}^n)$ a compact neighbourhood $\overline{\Omega}_k \subseteq G_k(\mathbb{R}^n)$ and a real analytic map $\mathcal{T} : \overline{\Omega}_k \to \mathbb{R}^{nk}$ such that

$$\mathcal{T}\left(V^k\right) := \left(\mathcal{T}_1\left(V^k\right), ..., \mathcal{T}_k\left(V^k\right)\right) \tag{3.8}$$

[8] $G_k(\mathbb{R}^n)$ is a smooth manifold of dimension $k(n-k)$. See Appendix D for all the details.
[9] Do not confuse the element $V^k \in G_k(\mathbb{R}^n)$ with $V_k(\mathbb{R}^n)$ the set of all linearly indipendent k-tuples in $\mathbb{R}^n$.
[10] See D.2

is an orthonormal basis for every $V^k \in G_k(\mathbb{R}^n)$, that is, from now on we suppose that V^k has an orthonormal basis[11].

Then $x \in \overline{B}_{\tilde{\xi}} \cap V^k$ can be written as

$$x = \sum_{j=1}^{k} x_j \mathcal{T}_j(V^k) \qquad \text{for some} \quad x_j \in \mathbb{R}$$

and then

$$\mathcal{H}_k(x, x_0, V^k) = \mathcal{H}_k\left(\sum_{j=1}^{k} x_j \mathcal{T}_j(V^k), x_0, V^k\right)$$

and then, recalling how we defined $\mathcal{H}_k$ in (3.6), we can write[12]

$$\mathcal{H}_k(x, x_0, V^k) = \|P_{V^k}\nabla h(x + x_0)\| =$$

$$= \left\|\nabla h\left(x_0 + \sum_{j=1}^{k} x_j \mathcal{T}_j(V^k)\right) \cdot \mathcal{T}_1(V^k), ..., \nabla h\left(x_0 + \sum_{j=1}^{k} x_j \mathcal{T}_j(V^k)\right) \cdot \mathcal{T}_k(V^k)\right\| =$$

$$= \left[\sum_{i=1}^{k} \left\langle \nabla h\left(x_0 + \sum_{j=1}^{k} x_j \mathcal{T}_j\left(V^k\right)\right) \middle| \mathcal{T}_i\left(V^k\right)\right\rangle^2\right]^{\frac{1}{2}}$$

For Theorem 3.1.5, we have that $\mathcal{H}_k$ is continous and subanalytic (See Definitions 3.1.3 and 3.1.4) since now the function is defined on the compact set $\overline{B}_{\tilde{\xi}} \times K \times \overline{\Omega}_k$ as we reported in (3.7)[13].
Now since $G_k(\mathbb{R}^n)$ is compact and the union of all the neighbourhoods $\overline{\Omega}_k$, depending on the element $V^k \in G_k(\mathbb{R}^n)$, covers the whole space $G_k(\mathbb{R}^n)$, then we can consider a finite subcover $\left\{\overline{\Omega}_k^{(i)}\right\}_i$ such that

$$\bigcup_{i=1}^{n_k} \overline{\Omega}_k^{(i)} = G_k(\mathbb{R}^n)$$

for some $n_k > 0$. Then every element $\overline{\Omega}_k^{(i)}$ of the cover has its corresponding section on $V_k^0(\mathbb{R}^n)$ that we can define, following (3.8), as $\mathcal{T}^{(i)} : \overline{\Omega}_k^{(i)} \to \mathbb{R}^{nk}$.

---

11 We highlight that $T_j(V^k)$ is an unit vector in $\mathbb{R}^n$.
12 We denote by $< \cdot | \cdot >$ the standard scalar product in $\mathbb{R}^n$.
13 See [21].

We want to apply Theorem 3.2.2 to the function

$$\mathcal{H}_k^{(i)}\left(x, x_0, V^k\right) = \mathcal{H}_k\left(\sum_{j=1}^{k} x_j \mathcal{T}_j^{(i)}\left(V^k\right), x_0, V^k\right) \tag{3.9}$$

for every (x_0, V^k) contained in the compact set $K\times\overline{\Omega}_k^{(i)}$, for every $i \in \{1, ..., n_k\}$. Following our construction, the restriction of h admits only isolated critical points if and only if, fixed $(x_0, V^k) \in K \times \overline{\Omega}_k^{(i)}$, $\mathcal{H}_k^{(i)}$ verifies the first condition of theorem 3.2.2.
If $\mathcal{H}_k^{(i)}(0, x_0, V^k) > 0$, there is nothing to prove since it means that x_0 is regular for every restriction of h and then the function is clearly steep. Conversely, if $\mathcal{H}_k^{(i)}(0, x_0, V^k) = 0$, then it means that x_0 is a critical point for $h_{|_{x_0+V^k}}$ and by hypothesis it is isolated.
Then, for every $k \in \{1, ..., n-1\}$ and for every $i \in \{1, ..., n_k\}$, for Theorem 3.2.2 there exist $C_k^{(i)} > 0$ and $\delta_k^{(i)} > 0$ such that, along any continous curve

$$\gamma : I := (-a, a) \subseteq \mathbb{R} \to V^k \cap \overline{B}_{\widetilde{\xi}} \qquad \text{such that } \gamma(0) = x_0 \qquad a > 0 \tag{3.10}$$

that joins x_0 to a point at a distance $\xi \leq \widetilde{\xi}$, there exists a point $t_* \in I$ such that

$$\mathcal{H}_k^{(i)}\left(\gamma(t_*), x_0, V^k\right) \geq C_k^{(i)} \xi^{\delta_k^{(i)}}$$

We recall that, as we showed in the proof of Theorem 3.2.2, $\delta_k^{(i)}$ is the Lojasiewicz's exponent of $\mathcal{H}_k^{(i)}$ with respect to the distance function (See Definition 3.1.1 and Theorem 3.1.8). We now set

$$C_k = \min_{i\in\{1,...,n_k\}} C_k^{(i)} \tag{3.11}$$

$$\delta_k = \max_{i\in\{1,...,n_k\}} \delta_k^{(i)} \tag{3.12}$$

and with this notation it is clear that for every curve γ defined as in (3.10) there exists a point t_* such that these coefficients and indices are such that

$$\|P_{V^k}\nabla h(x_0 + \gamma(t_*))\| \geq C_k \xi^{\delta_k}$$

Now for every $k \in \{1, ..., n-1\}$ and for every $i \in \{1, ..., n_k\}$ the couple $(x_0, V^k) \in K \times \overline{\Omega}_k^{(i)}$ describes all the subspaces that intersect K then we can state that h is steep on K with steepness coefficients $(C_1, ..., C_{n-1})$ and steepness indices $(\delta_1, ..., \delta_{n-1})$ for curves of length $\widetilde{\xi} = dist(K, \mathbb{R}^n \setminus D)$.
The original Definition 2.1.1 of steep function is satisfied since we apply $ii)$ of Theorem 3.2.2 to every $\mathcal{H}_k^{(i)}$ for every i and every k and take coefficients and indices as in (3.11) and (3.12).
The proof is completed.

Chapter 4

Examples and counterexamples

We dedicate this chapter to various kinds of examples and counterexamples. For instance, we compute explicitly, for easy functions, the steepness indices. We will do it in different cases, focusing the attention on functions that are steep only on subspaces of a certain dimension k, that is, not uniformly.
This means that, in general, the definition can not be extended to all the subspaces. If we want to verity the steepness condition with the original definition 2.1.1, then we have to consider all the $k \in \{1, ..., , n-1\}$, see Section 4.2.
Another important question upon which we wish to draw attention is that a function can change its nature and its behavior in different parts of the domain and it obviously influences the steepness coefficients and the indices. For example, this situation happens when we consider a function that is quasiconvex only in a part of its domain while is only three-jet nondegenerate in another part or it is neither, as we show in Section 4.1.
We show that a non-steep unperurbed Hamiltonian has perturbed actions that drift away fastly, that is, linearly with time.
We finally prove with a counterexample that real analyticity is a necessary condition for the equivalence between Definitions 2.1.1 and 2.1.4. Indeed, the equivalence fails if the function in only C^k.

4.1 Quasiconvex and three-jet non degenerate functions

For a review, in Chapter 2, Sections 2.2 and 2.3, can be found the properties and the results on quasi-convex and three-jet non-degenerate functions.

1. Let us consider a parameter $m \in \mathbb{R}$ and a function $h : \mathbb{R}^3 \to \mathbb{R}$ defined as

$$h(x, y, z) = \frac{x^2}{2} - \frac{y^2}{2} + m\frac{y^3}{3} + z \tag{4.1}$$

We study the steepness property for h when the parameter m changes.

   (a) For $m = 0$ the function is not steep. We firstly observe that it is neither quasi-convex nor three-jet non-degenerate. Indeed, we immediately see that $\det \hat{H} = 4 > 0$ and then by property 5 of Section 2.2 the function it is not quasi-convex.
   In addition, the function is three-jet degenerate since the three-tensor $h'''(x, y, z)$ of third order derivatives is the null three-tensor, then the system (2.5) has infinitely many solutions.
   However, as we will see, these conditions are sufficient for steepness, but not necessary.
   In this specific the definition of steepness is verified only on the vector subspace of dimension $k = 2$, that is, on the entire tangent plane $\nabla h(x, y, z)^{\perp}$ but not on all the subspaces of dimension $k = 1$ (see the next section 4.2).
   We consider a general point $\bar{x} = (x_0, y_0, z_0)$ so we can write

$$\nabla h(\bar{x}) = (x_0, -y_0, 1)$$

   For sake of simplicity and in order to avoid too many parameters, we take $\bar{x}$ on the z axis, that is in the form

$$\bar{x} = (0, 0, z_0)$$

   We obtain $\nabla h(\bar{x}) = (0, 0, 1)$ and then

$$\nabla h(\bar{x})^{\perp} = \{z = 0\} \subset \mathbb{R}^3$$

   To sum up, for $k = 2$ we have

$$V^2 = \nabla h(\bar{x})^{\perp} = \{z = 0\}$$

   and if $\bar{u}$ is a unit vector in V^2, then in polar coordinates, for $\eta, t \in \mathbb{R}$, it is

$$\eta\bar{u} = (\eta \cos t, \eta \sin t, 0) \in V^2$$

   so

$$\left\| P_{V^2} \nabla h\left(\bar{x} + \eta\bar{u}\right) \right\| = \left\| \left(2\eta \cos t, -2\eta \sin t, 0\right) \right\| = 2\eta$$

then when $k = 2$ the steepnes property is verified at $\bar{x}$ with steepness index equal to 1.
For $k = 1$, we consider the versor $\bar{u} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$, the subspace

$$V^1 = \text{span}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) \subset \nabla h(\bar{x})^\perp$$

and the vector $\bar{x} + \eta\bar{u} = \left(\eta\frac{1}{\sqrt{2}}, \eta\frac{1}{\sqrt{2}}, z_0\right)$. This time we obtain

$$\begin{aligned} &\|P_{V^1}\nabla h\left(\bar{x} + \eta\bar{u}\right)\| = \\ &= \left(\eta\frac{1}{\sqrt{2}}, -\eta\frac{1}{\sqrt{2}}, 1\right) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = \frac{\eta}{2} - \frac{\eta}{2} = 0 \end{aligned}$$

then on this line the property is not verified, that is, h is not steep if we consider subspaces of dimension $k = 1$. Then the function is not steep when $m = 0$.

(b) For $m \neq 0$ (without loss of generality $m > 0$), the function is steep since it is quasi-convex in a part of its domain and is three-jet non-degenerate in its complementary.
We have

$$\nabla h(\bar{x}) = \left(x_0, -y_0 + my_0^2, 1\right)$$

and

$$h''(\bar{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 + 2my_0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We use Proposition 2.2.1 to compute the steepness indices and then we look for the quasiconvexity regions.
It is

$$\det \hat{H} = \det \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & -1 + 2my_0 & 0 & -y_0 + my_0^2 \\ 0 & 0 & 0 & 1 \\ x & -y_0 + my_0^2 & 1 & 0 \end{pmatrix} = 1 - 2my_0$$

then

$$\det \hat{H} < 0 \iff y_0 > \frac{1}{2m}$$

and by property 5 in Section 2.2, we have that h is quasi-convex if and only if $y_0 > \frac{1}{2m}$.
It follows from Proposition 2.2.1 that the steepness index is $\delta_k = 1$

for every $k \in \{1,2\}$ whenever $y_0 > \frac{1}{2m}$.

Now the funcion is also three-jet non degenerate. When $y_0 = \frac{1}{2m}$ system (2.5) with unknown (x_1, x_2, x_3) is

$$\begin{cases} x_1 x_0 - \dfrac{x_2}{4} + x_3 = 0 \\ x_1^2 = 0 \\ 2x_2^3 = 0 \end{cases}$$

which has only the trivial solution $(x_1, x_2, x_3) = (0,0,0)$.
Similarly, for $y_0 < \frac{1}{2m}$ we obtain

$$\begin{cases} x_1 x_0 - y_0(1 - my_0)x_2 + x_3 = 0 \\ x_1^2 - (1 - 2my_0)x_2^2 = 0 \\ 2x_2^3 = 0 \end{cases}$$

which also has only the trivial solution $(x_1, x_2, x_3) = (0,0,0)$.
In this case the space orthogonal to $\nabla h(\bar{x}) = \nabla h(x_0, y_0, z_0)$ is the plane

$$\nabla h\left(\bar{x}\right)^{\perp} = \left\{(x,y,z) \in \mathbb{R}^3 \;\middle|\; x_0 x + \left(my_0^2 - y_0\right) y + z = 0\right\}$$

then the restriction of the quadratic form $h''(\bar{x})$ to that space has eigenvalues 1 and $-1 + 2my_0$, that is, the quadratic form $h''(\bar{x})$ is non-degenerate if and only if $-1 + 2my_0 \neq 0$.
We use Propositions 2.2.1 and 2.3.2 to state that h is steep on the orthogonal plane $\nabla h(\bar{x})^{\perp}$ with index $\delta_2 = 1$ when $y_0 \neq \frac{1}{2m}$, while is steep with index $\delta_2 = 2$ when $y_0 = \frac{1}{2m}$.
Finally, on the subspaces of dimension 1 contained in $\nabla h(\bar{x})^{\perp}$ we use Proposition 2.3.2 to state that $\delta_1 = 2$ in the non quasi-convexity but three-jet non-degenerate region $y_0 \leq \frac{1}{2m}$.

(c) This example shows that in the inequality (2.9) can be strict and the second term does not ensure the steepness condition.
Indeed, it is not true that for every fixed $y_0 \leq \frac{1}{2m}$ there exist constants $C, \delta > 0$ such that for every $\eta\bar{u} \in \nabla h(\bar{x})^{\perp}$

$$\frac{\|\nabla h(\bar{x} + \eta\bar{u}) \cdot \eta\bar{u}\|}{\|\eta\bar{u}\|} \geq C\|\eta\bar{u}\|^{\delta} \tag{4.2}$$

If we set $m=1$ we have $\nabla h(\bar{x}) = (x_0, y_0(y_0-1), 1)$ and the followings:

$$\begin{aligned}
\nabla h(\bar{x})^{\perp} &= \left\{(x,y,z)\in\mathbb{R}^3 |\, x_0x + y_0(y_0-1)y + z = 0\right\} \\
&= \left\{(x,y,z)\in\mathbb{R}^3 |\ \ z = (1-y_0)y_0y - x_0x\right\} \\
\nabla h(\bar{x}+\eta\bar{u}) &= (x_0+\eta u_1, (y_0+\eta u_2)\left(-1+y_0+\eta u_2\right), 1) \\
\Rightarrow \nabla h(\bar{x}+\eta\bar{u})\cdot\eta\bar{u} &= (\eta u_1)^2 - (1-2y_0)(\eta u_2)^2 + (\eta u_2)^3 = \\
&= (\eta u_1)^2 - (\eta u_2)^2(1-2y_0-\eta u_2)
\end{aligned}$$

with $\bar{u} = (u_1, u_2, u_3) \in \nabla h(\bar{x})^{\perp}$ with unit norm and $\eta\in\mathbb{R}$.

Now we consider the following curve in $\nabla h(\bar{x})^{\perp}$ that reaches the orgin at time $t=0$ defined for $t\geq 0$ in the non quasi-convex region $y_0 > \frac{1}{2}$:

$$\begin{aligned}
\bar{u}(t) &= \left(t\sqrt{(1-2y_0+t)}, -t, -t\,(1-y_0)\,y_0 - t\sqrt{(1-2y_0+t)}x_0\right) = \\
&= (u_1(t), u_2(t), (1-y_0)y_0u_2(t) - u_1(t)x_0)
\end{aligned}$$

We immediately check that

$$\nabla h(I+\bar{u})\cdot \bar{u}_{|_{\bar{u}(t)}} = t^2\left(1-2y+t\right) - t^2(1-2y+t)\big) \equiv 0 \qquad \forall t>0$$

However, as we have just proved, the function is steep with steepness index equal to 1.

2. We show a function that is steep but does not verify the three-jet non-degeneracy condition (2.5).
We consider the function $h:\mathbb{R}^3\to\mathbb{R}$

$$h(\bar{x}) = \frac{x_1^2}{2} + \frac{x_2^4}{4} + x_3$$

in a neighbourhood of $(0,0,x_3)$ for all $x_3\in\mathbb{R}$.
We have

$$\nabla h(\bar{x}) = \left(x_1, x_2^3, 1\right)_{|_{(0,0,x_3)}} = (0,0,1)$$

and then

$$\nabla h(\bar{x})^{\perp} = \{z=0\}$$

then we consider the canonical basis $\{e_1,e_2\}$ as a basis of $\nabla h(\bar{x}I)^{\perp}$.
The function is three-jet degenerate along the x_2 axis. We denote, as usual, with h'' the Hessian matrix of h at the point $\bar{x}$.

$$h'' := \begin{pmatrix}1 & 0 & 0\\ 0 & 3x_2^3 & 0\\ 0 & 0 & 0\end{pmatrix}\Bigg|_{(0,0,x_3)} = \begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$$

and with h''' the three-tensor

$$\frac{\partial^3 h}{\partial x_i \partial x_j \partial x_k}(\bar{x}) = 0$$

with $i, j, k \in \{1, 2, 3\}$ which is such that at the point $(0, 0, x_3)$ is the null three-tensor.
Then, for $v = (v_1, v_2, v_3) \in \mathbb{R}^3$, the system:

$$\begin{cases} \nabla h \cdot v = v_3 = 0 \\ h''(\bar{x})v \cdot v = v_1^2 = 0 \\ h'''(\bar{x})vv \cdot v = 0 \end{cases}$$

has infinite solutions $(0, v_2, 0)$ for every $v_2 \in \mathbb{R}$.
However the function is steep.
Indeed, if we take $v \in \nabla h(\bar{x})^{\perp}$ then v has the form $v = (v_1, v_2, 0)$ and we can verify the steepness condition applying the definition. Moreover, we compute the steep index of h on that plane.
Recalling that the steepness property is local and then $v_i \leq 1 \; \forall i$, named $V := \nabla h(I)^{\perp}$, we find

$$\begin{aligned} \|P_V \nabla h(I + v)\| &= \left\|P_V \left(v_1, v_2^3, 1\right)\right\| = \|(v_1, v_2^3, 0)\| = \\ &= \sqrt{v_1^2 + v_2^6} \geq \sqrt{v_1^6 + v_2^6} \geq \frac{1}{2^{\frac{3}{2}}}\|v\|^3 \end{aligned}$$

In the last inequality we used the general propery such that, for every $v \in \mathbb{R}^n$, the following inequality holds

$$\|v\|_p \geq \|v\|_\infty \geq \frac{1}{\sqrt{n}}\|v\|_2$$

where $\|\cdot\|_p$ is the usual L^p norm.

Remark 4.1.1. We chose this function just to avoid too many coefficients. We could have chosen also $h(x_1, x_2, x_3) = x_1^4 + x_2^4 + x_3$ to obtain the same results.

Remark 4.1.2. In general, we can easily show that around the point $(0, 0, x_3)$ the function

$$h(x_1, x_2, x_3) := \frac{x_1^p + x_2^p}{p} + x_3$$

never verifies the jet condition of order $p-1$ but, however, the function is steep with index $p-1$.
Indeed, with such functions we have

$$\left\|P_V(v_1^{p-1}, v_2^{p-1}, 1)\right\| = \left\|(v_1^{p-1}, v_2^{p-1}, 0)\right\| = \sqrt{v_1^{(p-1)2} + v_2^{(p-1)2}} =$$
$$= \|v\|_{2(p-1)}^{p-1} \geq \|v\|_\infty^{p-1} \geq \left(\frac{1}{\sqrt{n}}\right)^{p-1} \|v\|_2^{p-1}$$

4.2 Steepness and k-dimensional subspaces

We show by examples that the definition of steepness at a point, in general, depends on the dimension k of the subspace.
The following examples were suggested by Nekhoroshev itself, who highlighted the indipendence of the steepness condition for different dimensions of the subspace V^k.
Here we show two functions of three variables where the first satisfies the steepness condition when $k=2$ and not for $k=1$, while the second one verifies the definition in the opposite case.
In general it may happen that a function is steep along all the subspaces of dimension lower or equal to $n-2$ and it is not steep on a subspace of dimension equal or greater than $n-1$.

1. We consider the function $h(x, y, z) = x + y^2 - z^2$ defined on $\mathbb{R}^3$ and want to check the steepness conditions. The function h does not verify the conditions when $\dim V^k = 1$ on points in the form $I = (x_0, y_0, y_0)$.
   It is

   $$\nabla h(I) = (1, 2y_0, -2y_0)$$
   $$\Rightarrow \nabla h(I)^\perp = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; x + 2y_0 y - 2y_0 z = 0\right\}$$

   We consider, as usual, the subspace

   $$V^1 = \operatorname{span} \bar{u} = \operatorname{span}\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \subseteq \nabla h(I)^\perp$$

   and

   $$I + \eta\bar{u} = \left(x_0, y_0 + \frac{\eta}{\sqrt{2}}, y_0 + \frac{\eta}{\sqrt{2}}\right) \Rightarrow$$
   $$\Rightarrow \nabla h\left(I + \eta\bar{u}\right) = \left(1, 2\left(y_0 + \frac{\eta}{\sqrt{2}}\right), -2\left(y_0 + \frac{\eta}{\sqrt{2}}\right)\right)$$

so

$$\|P_{V^1}\nabla h(I+\eta\bar{u})\| = \nabla h(I+\eta\bar{u})\cdot\left(0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right) \equiv 0$$

which proves that h is not steep when $\dim V^k = 1$.
On the other hand, when $k = 2$, so when V^2 is the entire tangent plane, if we put for simplicity $y_0 = 0$, then it is $I = (x_0, 0, 0)$, so

$$\nabla h(I) = (1,0,0) \qquad \nabla h(I)^\perp = \{x = 0\}$$

and we can immediately compute the orthogonal projection of the gradient over the tangent space.
We express a unit vector on the tangent plane in the form $\bar{u} = (0, \cos t, \sin t)$ with $t \in \mathbb{R}$ and then we can write

$$I + \eta\bar{u} = (x_0, \eta\cos t, \eta\sin t) \Rightarrow \nabla h(I+\eta\bar{u}) = (1, 2\eta\cos t, -2\eta\sin t)$$

then

$$\|P_{V^2}\nabla h(I+\eta\bar{u})\| = \|(0, 2\eta\cos t, -2\eta\sin t)\| = 2\eta$$

which implies that the definition of steepness is verified when $k = 2$.

2. The function $h(x,y,z) = (x-y^2)^2 + z^2$ satisfies the steepness conditions on points in the form $x = y^2$ and $z \neq 0$ when $k = 1$, but not when $k = 2$. Indeed

$$\nabla h(x,y,z) = \big(2(x-y^2), -4xy(x-y^2), 2z\big)$$

and at the point $I = (y_0^2, y_0, z_0)$ it is

$$\nabla h(x,y,z) = (0,0,2z_0)$$

and $\nabla h^\perp$ is simply the plane $\{z = 0\}$. We will show the result in the case $y_0 = 0$, *i.e.* restricted to the z axis, but the result is true for all $y_0 \in \mathbb{R}$. We can now write, following the usual notation:

$$\bar{u} = (\cos t, \sin t, 0) \qquad \bar{u} \in \nabla h(I)^\perp \qquad \|\bar{u}\| = 1$$
$$I + \eta\bar{u} = (0,0,z_0) + \eta\,(\cos t, \sin t, 0) = (\eta\cos t, \eta\sin t, z_0)$$

then

$$\nabla h\,(I+\eta\bar{u}) = \big(2(\eta\cos t - \eta^2\sin^2 t), -4\eta^2\cos t\sin t(\eta\cos t - \eta^2\sin^2 t), 2z_0\big)$$

so when $k = 2$ it is

$$\begin{aligned}
&\|P_{V^2}\nabla h(I+\eta\bar{u})\| = \\
&= \left\|\left(2\left(\eta\cos t - \eta^2\sin^2 t\right), -4\eta^2\cos t\sin t\left(\eta\cos t-\eta^2\sin^2 t\right), 0\right)\right\| = \\
&= \left[4\eta^2\left(\cos t - \eta\sin^2 t\right)^2 + 16\eta^6\cos^2 t\sin^2 t\left(\cos t-\eta\sin^2 t\right)^2\right]^{\frac{1}{2}} = \\
&= 2\eta\left|\cos t-\eta\sin^2 t\right|\left[1+4\eta^4\cos^2 t\sin^2 t\right]^{\frac{1}{2}} = \\
&= 2\eta\left|\cos t-\eta+\eta\cos^2 t\right|\left[1+4\eta^4\cos^2 t\sin^2 t\right]^{\frac{1}{2}}
\end{aligned}$$

which shows that $\|P_{V^2}\nabla h(I+\eta\bar{u})\| = 0$ if

$$\eta\cos^2 t + \cos t - \eta = 0$$

and the last equation is satisfied for small $\eta > 0$ if and only if

$$\cos t = \frac{1-\sqrt{1+4\eta^2}}{2\eta} \Rightarrow t = \arccos\frac{1-\sqrt{1+4\eta^2}}{2\eta} =: t(\eta)$$

This shows that the function is not steep when $k = 2$. Indeed we found a curve $\gamma \subset V^2$ along which the projection of the gradient vanishes, in particular we found that

$$\min_{\bar{u}\in V^2}\|P_{V^k}\nabla h(I+\eta\bar{u})\| =: \mu_I(\eta) = 0$$

when $\bar{u} = (\cos t(\eta), \sin t(\eta), 0)$.
On the other hand, when $k = 1$, V^1 is generated by the versor $\bar{u} = (\cos\bar{t}, \sin\bar{t}, 0)$ with fixed $\bar{t} \in [0, 2\pi]$ so to verify the definition we just have to compute the lenght of the projection of the gradient along this versor.
We have

$$\|P_{V^1}\nabla h(I+\eta\bar{u})\| = \nabla h(I+\eta\bar{u})\cdot\bar{u}$$

that is

$$\left(\eta\cos\bar{t}-\eta^2\sin^2\bar{t}\right)\left[2\cos\bar{t}\left(1-2\eta^2\sin^2\bar{t}\right)\right]$$

4.3 Two counterexamples

4.3.1 A non-steep function

We saw in Section 4.1 that, when $m = 0$, the function (4.1) loose steepness. We now show a classical example of a non steep function to which the Nekhoroshev theorem can not be applied. The function is a unperturbed Hamiltonian $h : \mathbb{R}^2 \to \mathbb{R}$ such that

$$h(I_1, I_2) = I_1^2 - I_2^2 \tag{4.3}$$

We see that it looses steepness on the line $I_1 = I_2$.
Indeed, if $I = (I_1, I_1)$, it is

$$\nabla h(I) = 2(I_1, -I_1) \Rightarrow \nabla h(I)^{\perp} = \{(I_1, I_1)\}$$

then, if $\bar{u} = (u_1, u_1)$ is such that $\|\bar{u}\| = 1$, it is

$$\nabla h(I + \eta\bar{u}) = 2\,(I_1 + \eta u_1, -I_1 - \eta u_1)$$

and

$$\left\|P_{\nabla h(I)^{\perp}} \nabla h(I + \eta\bar{u})\right\| = 2\,(I_1 + \eta u_1, -I_1 - \eta u_1) \cdot \eta\bar{u} = 0$$

Now, recalling Section 1.2, if (4.3) is a unperturbed hamiltonian defined on a $2n = 4$ dimensional phase space and $(I_1, I_2) \in \mathbb{R}^2$ are the action variables, then if the perturbed Hamiltonian has the form

$$H(I_1, I_2, \varphi_1, \varphi_2) = h(I_1, I_2) + \varepsilon f(I_1, I_2, \varphi_1, \varphi_2)$$

where $f(I_1, I_2, \varphi_1, \varphi_2) = f(\varphi_1, \varphi_2) = \sin(\varphi_1 + \varphi_2)$ and ε is the small parameter, then, with initial condition in the origin, the solution is

$$I(t) = (\varepsilon t, \varepsilon t) \qquad \text{and} \qquad \varphi(t) = \left(-\varepsilon t^2, \varepsilon t^2\right)$$

that is

$$\|I(t) - I(0)\| = \sqrt{2}\varepsilon t$$

then Nekhoroshev theorem 1.2.2 clearly fails since perturbed actions drift away linearly with the time t.

4.3.2 Arcsteepness versus steepness

In the second chapter, we stated that the original definition 2.1.1 of steepness is stronger than definition 2.1.4 if the function is only C^1 without being analytic in its domain, while Teorem 3.1.1 ensures that in the real analytic case the two definitions are equivalent. Here in this section, we want to construct an explicit example of a C^1 and not real analytic function that is arc-steep at a point in the sense of definition 2.1.4 but does not verify the steepness condition 2.1.1.

First of all, let consider in $\mathbb{R}^2$ a number $\bar{\xi} > 0$ and the closed ball $\overline{B_{\bar{\xi}}(0)}$ of radius $\bar{\xi}$ centered at the origin and we take sets R_j such that for $j \geq 0$ they are expressed in polar coordinates by

$$R_j = \left\{(\rho, \theta) \,\middle|\, \rho \in \left[\frac{\bar{\xi}}{2^{j+1}}, \frac{\bar{\xi}}{2^j}\right],\ \theta \in \left[-\frac{\pi}{3} + j\pi, \frac{\pi}{3} + j\pi\right]\right\} \tag{4.4}$$

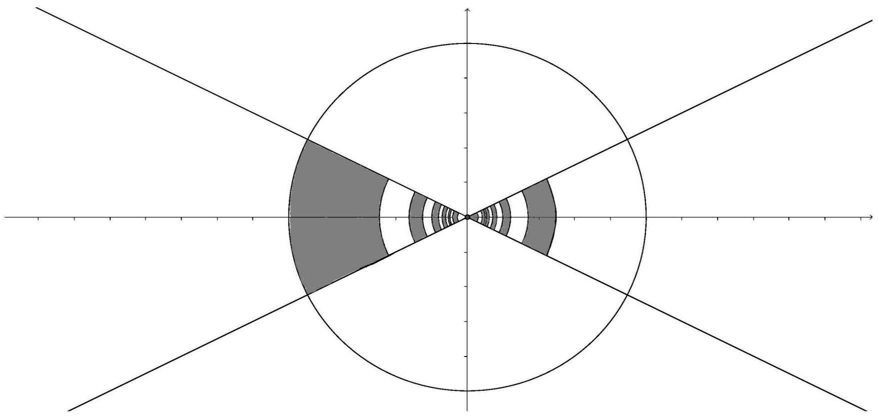

Figure 4.1: The set $R = \bigcup_{j=1}^{+\infty} R_j$

and we set

$$R = \bigcup_{j=0}^{+\infty} R_j$$

We observe that the last union is disjoint and afterwards we consider "thicker" sets $\widetilde{R}_j$ such that

$$R_j = \left\{ (\rho, \theta) \mid \rho \in \left[\frac{\bar{\xi}}{2^{j+1}} - \varepsilon_j, \frac{\bar{\xi}}{2^j} + \varepsilon_j \right], \ \theta \in \left[-\frac{\pi}{3} + j\pi, \frac{\pi}{3} + j\pi \right] \right\}$$

where

$$\varepsilon_j = \frac{1}{3} \left(\frac{\bar{\xi}}{2^{j+1}} - \frac{\bar{\xi}}{2^{j+2}} \right) = \frac{\bar{\xi}}{2^{j+2}3}$$

and where the quantity in brackets is the range of the length of the radius in the set R_{j+1}. We highlight that if the interval where ρ is contained has length

$$h_j = \frac{\bar{\xi}}{2^j} - \frac{\bar{\xi}}{2^{j+1}}$$

then $\varepsilon_j = \frac{h_{j+1}}{3}$. This choice of ε_j allow us to avoid overlap situations, that is, $\widetilde{R}_j \cap \widetilde{R}_{j+2} = \emptyset$ for every j.
Now we set

$$\widetilde{B} := \overline{B_{\bar{\xi}}(0)} \setminus \left(\bigcup_{j \geq 0} \widetilde{R}_j \right)$$

and consider a bump function[1] $\psi : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\psi(x_1, x_2) = \begin{cases} 1 & \text{if } (x_1, x_2) \in R \\ 0 & \text{if } (x_1, x_2) \in \left(\bigcup_{j \geq 0} \widetilde{R} \right)^c \end{cases}$$

If $\bar{x} = (x_1, x_2)$ we define

$$h(x_1, x_2) = x_2 + (1 - \psi(x_1, x_2)) \frac{|\bar{x}|^2}{2}$$

It is immediate that $h \in C^\infty(B_{\bar{\xi}}(0) \setminus \{0\}) \cap C^1(\mathbb{R}^2)$ without being analytic in the origin.
We state that this function does not verify the steepness condition 2.1.1 in the origin, although it is arc-steep. It should be noticed that

$$h_{|R}(x_1, x_2) = x_2$$
$$h_{|\widetilde{B}} = x_2 + \tfrac{|\bar{x}|^2}{2}$$

Computing the gradient of h we verify that in the regions we are interested in we obtain

$$\nabla h(x_1, x_2) = \begin{cases} (0, 1) & \text{if } (x_1, x_2) \in R \\ (x_1, 1 + x_2) & \text{if } (x_1, x_2) \in \widetilde{B} \end{cases}$$

and in particular

$$\nabla h(0, 0) = (0, 1) \ \Rightarrow V^1 := \nabla h(0, 0)^\perp = \mathbb{R} \times \{0\}$$

that is, the projection onto the orthogonal space to the gradient computed at the origin is just the projection on the first component.
Now, following definition 2.1.1, we consider a vector $u \in V^1$ such that $\|u\| = 1$ and a positive number $0 < \eta \leq \bar{\xi}$ so that the vector ηu is just $\eta u = (\pm \eta, 0)$ and the sign depends on the orientation of u.
We project onto the space V^1, obtaining

$$\|P_{V^1} \nabla h(\eta u)\| = \|\partial_{x_1} h(\eta u)\| = \begin{cases} 0 & \text{if } \eta u \in R \\ \eta & \text{if } \eta u \in \widetilde{B} \end{cases}$$

The interesting element is that if

[1]A bump function is a function $f : \mathbb{R}^n \to \mathbb{R}$ which is both smooth and compactly supported. The bump function is a sort of smoothing of the characteristic function and can be obtained by taking a convolution product with a mollifier. We remark that bump functions cannot be analytic unless they vanish identically.

$$\|P_{V^1}\nabla h(\eta,0)\| = 0 \ \text{ then } \|P_{V^1}(-\eta,0)\| = \eta$$

and vice versa. That is the minimum u is such that the projection of the gradient vanishes and from this result it is clear that the definition of steepness is not verifies.
Actually we just proved that

$$\forall\ 0 < \eta \leq \bar{\xi} \text{ there exists } u \in V^1 \text{ such that } \|P_{V^1}\nabla h(\eta u)\| = 0$$

On the other hand, the function is arc-steep. In fact on every continous curve

$$\gamma : I \to V^1 \cap B_{\bar{\xi}}(0)$$

joining the origin to a point $\bar{x}$ at a distance $d \leq \bar{\xi}$, there can always be found a point $\gamma(t_*)$ such that $\|P_{V^1}\nabla h(\gamma(t_*))\|$ is greater than $C\xi^\delta$ for some $\xi \leq \bar{\xi}$ and in particular it is that $\gamma(t_*) \in \widetilde{B}\cap V^1$. Then definition 2.1.4 with $f \in C^1$ is verified, so we conclude that the C^1 condition is not sufficient to make definition 2.1.1 and 2.1.4 equivalent.

Appendices

Appendix A

The Taylor's Formula in Several Variables

In this first Appendix we provide a review of the Taylor's Theorem in several variables. We consider only scalar-valued functions.

Definition A.0.1. *A* **multi-index** α *is an n-tuple of natural numbers, that is*

$$\alpha = (\alpha_1, \alpha_2, ..., \alpha_n) \qquad \alpha_i \in \mathbb{N} \qquad \forall i$$

If α is a multi-index, called $\bar{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, we can define

$$|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_n \qquad \alpha! = \alpha_1!\alpha_2!...\alpha_n!$$
$$\bar{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n}$$
$$\partial^\alpha f = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} ... \partial_{x_n}^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial {x_1}^{\alpha_1} \partial {x_2}^{\alpha_2} ... \partial x_n^{\alpha_n}}$$

It is known that if f is a function of class C^k, then the order of differentiation up to order k in a partial derivative is immaterial. Then, we can write the k-th order partial derivative of f simply as $\partial^\alpha f$ with $|\alpha| = k$.

Theorem A.0.1 (Multinomial Theorem)**.** *For any $\bar{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ and any positive integer k, it is*

$$(x_1 + x_2 + ... + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \bar{x}^\alpha$$

Proof. We proceed by induction on n. The case $n = 2$ is just the binomial formula

$$(x_1 + x_2)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} x_1^i x_2^{k-j} = \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \bar{x}^\alpha$$

where we have set $\alpha_1 = j$, $\alpha_2 = k - j$ and $\alpha = (\alpha_1, \alpha_2)$. We suppose then that the result is true for $N - 1 = n$ then, taken $\bar{x} = (x_1, ..., x_N)$ and using the inductive hypothesis, we obtain

$$(x_1 + ... + x_N)^k = \sum_{i+j=k} \frac{k!}{i!j!} (x_1 + ... + x_{N-1})^i x_N^j = \sum_{i+j=k} \frac{k!}{i!j!} \sum_{|\beta|=i} \frac{i!}{\beta!} \widetilde{x}^\beta x_N^j$$

where $\beta = (\beta_1, ..., \beta_{N-1})$ and $\widetilde{x} = (x_1, ..., x_{N-1})$.
If we now set $\alpha = (\beta_1, ..., b_{N-1})$, it is $\beta! j! = \alpha!$ and $\widetilde{x}^\beta x_N^j = \bar{x}^\alpha$ we obtain the result

$$\sum_{|\alpha|=k} \frac{k!}{\alpha!} \bar{x}^\alpha$$

□

We can extend the result to the product rule for higher-order partial derivatives

$$\partial^\alpha (fg) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \left(\partial^\beta f\right) (\partial^\gamma g)$$

and the proof is by induction on the number n of variables.

Suppose now to have $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ of class C^k defined on an open set S, with S convex. Then the Taylor expansion for $f(\bar{x})$ about a point $a \in S$ is

$$f(a+h) = \sum_{j=0}^{k} \frac{(h \cdot \nabla)^j f(a)}{j!} + R_{a,k}(h) \tag{A.1}$$

where we set $h = \bar{x} - a$, where $(h \cdot \nabla)^j f$ is the directional derivative

$$h \cdot \nabla = h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + ... + h_n \frac{\partial}{\partial x_n}$$

applied j times to f and where $R_{a,k}(h)$ is the reminder.
We now apply the multinomial theorem A.0.1 to the last expression to obtain

$$(h \cdot \nabla)^j = \sum_{|\alpha|=j} \frac{j!}{\alpha!} h^\alpha \partial^\alpha$$

and if we substitute this expression in equation (A.1), we obtain the following:

Theorem A.0.2 (Taylor's Theorem in Several Variables)**.** *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is of class* C^{k+1} *on an open convex set* S*. If* $a \in S$ *and* $a + h \in S$*, then*

$$f\,(a+h) = \sum_{|\alpha|\leq k} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha + R_{a,k}(h) \tag{A.2}$$

where the reminder is given in Lagrange's form by

$$R_{a,k}\,(h) = \sum_{|\alpha|=k+1} \partial^\alpha f(a+ch)\frac{h^\alpha}{\alpha!} \tag{A.3}$$

for some $c \in (0,1)$ *and in integral form by*

$$R_{a,k}(h) = (k+1) \sum_{|\alpha|=k+1} \frac{h^\alpha}{\alpha!} \int_0^1 (1-t)^k \partial^\alpha f(a+th)\mathrm{d}t \tag{A.4}$$

As in one variable, the following estimates for the reminder term follows from the Lagrange or integral formulas for it:

Corollary A.0.3. *If* f *is of class* C^{k+1} *on* S *and* $|\partial^\alpha f(x)| \leq M$ *for* $\bar{x} \in S$ *and* $|\alpha| = k+1$*, then*

$$|R_{a,k}(h)| \leq \frac{M}{(k+1)!}\|h\|^{k+1}$$

where

$$\|h\| = |h_1| + |h_2| + ... + |h_n|$$

Proof. It follows from either (A.3) or (A.4) that

$$|R_{a,k}(h)| \leq M \sum_{|\alpha|=k+1} \frac{|h|^\alpha}{\alpha!}$$

and the result follows from Theorem A.0.1 □

Appendix B

Upper semi-continous functions

In the definitions of steepness that we have provided (2.1.1, 2.1.2 and 2.1.3) it is implied that the function

$$\mu(\xi) := \min_{\substack{\|\bar{u}\|=1 \\ \bar{u}\in V^k}} \|P_{V^k}\nabla h(I+\xi\bar{u})\| \tag{B.1}$$

reaches its maximum value over the compact set $[0, \widetilde{\xi}]$. This property is always verified, as it is shown in this appendix, since the function (B.1) is upper semi-continous and these kinds of functions always reaches their maximum on a compact set.
It follows that, in definitions 2.1.1, 2.1.2 and 2.1.3, it is correct to use the notion of maximum of a function instead of the notion of supremum[1].
We prove that (B.1) is upper semi-continous and then that the maximum value is reached on a compact set.

Definition B.0.2. *A real valued function f defined on a metric space X is said to be upper semi-continous at a point $\bar{x} \in X$ if for every $\varepsilon > 0$ there exists a neighbourhood U of $\bar{x}$ such that $\forall x \in U \cap X$*

$$f(x) < f(\bar{x}) + \varepsilon$$

[1]The usage of the minimum instead of the inf in definition 2.1.1 obviously follows from the Weierstrass's Theorem, since $\nabla h(I+\xi u)$ is continous over the compact set

$$K = \{u \in V_k \cap G, \|\bar{u}\| = 1\}$$

A function is upper semi-continous on a space X if it is upper semi-continous at all the points of X.

It is well known that if $\bar{x}$ is an accumulation point for X then f is upper semi-continous at $\bar{x}$ if and only if

$$\limsup_{x \to \bar{x}} f(x) \le f(\bar{x})$$

We have

Theorem B.0.4. *If X is a compact set of $\mathbb{R}$ and if f is upper semi-continous on X then*

$$\sup_X f \in f(X)$$

that is, f is bounded above and reaches its max *on X.*

Proof. Let $L = \sup_X f$. L is in $\overline{\mathbb{R}}$ and in particular $L \in \overline{f(X)}$, then there exists a sequence $\{y_n\}_{n \in \mathbb{N}}$ with $y_n \in f(X)$ such that

$$\lim_{n \to \infty} y_n = L$$

Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in X such that $f(x_n) = y_n$ for every $n \in \mathbb{N}$. X is compact then we can extract a subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ which converges to a point $x_0 \in X$ and, on the other hand, $f(x_{n_k})$ is a subsequence extracted from $\{y_n\}$ then

$$\lim_{k \to \infty} f(x_{n_k}) = L$$

Now f is upper semi-continous in x_0 then, fixed $\varepsilon > 0$, there exists a neighbourhood U of x_0 such that

$$f(x) < f(x_0) + \varepsilon \qquad \forall x \in U \cap X$$

and for sufficiently big $\bar{k} > 0$ we have that x_{n_k} belongs to U for every $k > \bar{k}$ then for these k we have

$$f(x_{n_k}) < f(x_0) + \varepsilon$$

then finally it is

$$f(x_0) \le L \le f(x_0) + \varepsilon$$

and the statement follows from the arbitrariness of ε □

We show that (B.1) is upper semi-continous.

Proposition B.0.5. *Let $f \in C(\mathbb{R}^n, \mathbb{R})$ and, for $\xi \geq 0$, let*

$$\mu(\xi) := \min_{\bar{u} \in \partial B_\xi} f(\bar{u})^2$$

Then $\mu(\xi)$ is upper semi-continous.

Proof. By contradiction, we suppose that μ is not upper semi-continous at a point $\xi_0 > 0$. Then there exist $\varepsilon > 0$ and a sequence $\xi_n \to \xi_0$ such that

$$\mu(\xi_n) \geq \mu(\xi_0) + \varepsilon$$

for every $n \in \mathbb{N}$.
Let now $\{u_n\}_{n\in\mathbb{N}} \subseteq \partial B_\xi$ be a sequence such that

$$f(u_n) = \mu(\xi_n)$$

then we can extract a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} u_{n_k} = \bar{u} \qquad \|\bar{u}\| = \xi_0$$

and f is continous, then $f(u_{n_k}) \to f(\bar{u})$ when $k \to \infty$.
So

$$\mu(\xi_0) + \varepsilon \leq \mu(\xi_{n_k}) = f(u_{n_k})$$

and passing to the limit, we obtain

$$\mu(\xi_0) + \varepsilon \leq f(\bar{u}) \leq \mu(\xi_0)$$

which is a contradiction then f has to be upper semi-continous at ξ_0. $\square$

[2] ∂B_ξ denotes the usual sphere in $\mathbb{R}^n$ centered at the origin and of radius ξ.

Appendix C

A theorem on steepness by Nekhoroshev

We have premised in our work that we would have stated an general theorem on steep functions formulated by Nekhoroshev in [18]. The result is very general and basically it is not related to dynamical systems.
Indeed, steepness is a property of a general C^1 function which is not necessary a unperturbed Hamiltonian.
Roughly speaking, the author stated that a function $f(x)$ defined on a domain $G \subseteq \mathbb{R}^n$ is steep in a neighbourhood of a point $x \in G$ if $\nabla f(x)$ is nonzero and if the coefficients of the Taylor's expansion of f up to a order $r > 0$, namely the r-jets of $f(x)$, at the point x do not solve any of a certain number of equations and inequalities, together with other auxiliary parameters. It is also provided an estimate for the steepness indices.
The theorem asserts that a steep function lies outside a certain semialgebraic set[1] which is contained in the set of r-jets of smooth functions. However, it is not provided an explicit construction of such sets.
There is a recent paper [24] where is provided an explicit formulation of such sets for specific orders of differentiation ($r = 2, 3$) and dimension $n \geq 2$.
Let now review this classical result [18].

We start with the following

Definition C.0.3. *Let $f : G \subseteq \mathbb{R}^n \to \mathbb{R}$, we say that the r-jet of f at a point $x \in G$ is a vector $\bar{F}$ consisting of the coefficients f_μ of the Taylor polynomial of*

[1] See [7] for a complete definition of semialgebraic set.

order $r \geq 2$ of the function f at the point x with the exception of the constant term. We denote it by

$$\bar{F} = \{f_\mu, 1 \leq |\mu| \leq r\}$$

where $\mu = \mu_1, ..., \mu_n$ is a multi-index[2]*, $\mu_i \geq 0$ are integers for every i and $|\mu| = \sum_{i=1}^n \mu_i$ and where*

$$f_\mu = \frac{1}{\mu!} \frac{\partial^{|\mu|} f(x)}{\partial x^\mu}$$

Furthermore, we denote by $J^r(n)$ the space of the r-jets of all the smooth functions of n variables at the point x.

Now with this definitions there are sets denoted by $\sigma^r(n)$ contained in $J^r(n)$. Such sets are determined through a collection of systems of polynomial equalities and inequalities, depending the gradient ∇f and on other "parameters", that is numbers depending on the dimension n of the space G, on the order r and also on the dimension m of the subspaces contained in the space othogonal to the gradient of f. A r-jet of a function belongs to this set $\sigma^r(n)$ if and only if it is the solution of at least one of the inequalities.
The final result will be that the closure of these sets, denoted by $\Sigma^r(n)$, contains the r-jets of all the non-steep functions.
The construction of the sets $\sigma^r(n)$ is beyond the scope of this thesis. The complete construction can be found in [20] and [18][3].
To summarize, $\sigma := \sigma^r(n)$ is a subset of $J^r(n)$ and an r-jet $\bar{F}$ belongs to σ if and only if it verifies at least one of those inequalities.
The closure of $\sigma^r(n)$ in $J^r(n)$ is denoted by $\Sigma^r(n)$ and Nekhoroshev asserts that it coincides[4] with the closure $\overline{\sigma}^r(n)$ in $J^{2r-2}(n)$ and this equality leads to the final proof.

Recalling the definition of steepness (see Definition 2.1.1), we highlight that the result implies that the function is steep at a point, that is, it is steep on all subspaces of dimension $m \in \{1, ..., n-1\}$.
Indeed, as one can see below, we also have estimates for the steepness indices for all the m-dimensional subspaces.
The original theorem provides also an estimate for the codimension of $\Sigma^r(n)$ in $J^r(n)$.

We now can state the original

[2]See Appendix A for details.
[3]Sections 5.2.B, 5.2.C and 5.2.D.
[4]Lemma 5.3.B [18].

Theorem C.0.6. *For any $r \geq 2$ and $n \geq 2$ the semialgebraic set $\Sigma^r(n)$ contained in $J^r(n)$ has the following properties:*

1. *Let h be an arbitrary function of class C^{2r-1} in a neighbourhood of a point $\bar{I}$ where $\nabla h(\bar{I}) \neq 0$. If the r-jet of the function h lie outside of the set $\Sigma^r(n)$ then h is steep in a neighbourhood of $\bar{I}$.*

2. *For every $m \in \{1, ..., n-1\}$, the steepness index α_m of the function h in this neighbourhood is not greater than $\bar{\alpha}_m$ where*

$$\bar{\alpha}_m = \begin{cases} \max\left[1, 2r - 3 - \frac{n(n-2)}{2} + 2m(n-m-1)\right] & \textit{if } n \textit{ is even} \\ \max\left[1, 2r - 3 - \frac{(n-1)^2}{2} + 2m(n-m-1)\right] & \textit{if } n \textit{ is odd} \end{cases}$$

The difficulty of the theorem is determine the set $\Sigma^r(n)$. We know an explicit construction for $r = 2, 3$ but no explicit conditions are known[5] for $r \geq 4$.

[5]See [24].

Appendix D

The smooth Grassmann manifold

D.1 Introduction

In this appendix, we provide the main steps of the proof that a smooth structure can be given to the Grassmann manifold or Grassmannian. We essentially follow [14] and we take the opportunity to thank Dr. Domenico Monaco for his helpful suggestions on this geometric argument. In these proofs we use elementary tools easy to understand, but some of them may be pretty boring for the advanced reader. If the reader is familiar with the Lie groups, we suggest to see Example 7.22 in Chapter 7 in [14] where it is proved more concisely that the Grassmannian has a smooth structure.

From now on, it is always supposed that the space $\mathbb{R}^n$ is endowed with the euclidean topology.

Definition D.1.1. *For any integer $k \in \{1, ..., n\}$ the Grassmannian $G_k(\mathbb{R}^n)$ is the set of all k-dimensional linear vector subspaces contained in $\mathbb{R}^n$.*

We define some well known tools that are used to assign a good topology to $G_k(\mathbb{R}^n)$.

Definition D.1.2. *Let $(X, \mathcal{T})$ be a topological Hausdorff space. A coordinate chart on X is the couple (U, φ) where $U \in \mathcal{T}$ is an open subset of X and $\varphi : X \to \mathbb{R}^n$ is a map such that:*

1. *$\varphi(U)$ is an open set in $\mathbb{R}^n$*
2. *φ is a homeomorphism on $\varphi(U)$.*

If we consider two charts (U, φ_U) and (V, φ_V) such that $U \cap V \neq \emptyset$ then we define the transition map from φ_U to φ_V as the map defined by

$$\varphi_V \circ \varphi_U^{-1} : \varphi_U(U \cap V) \to \varphi_V(U \cap V) \tag{D.1}$$

We remind that in our case the coordinate charts are smooth maps, that is, they are diffeomorphisms. Two charts (U, φ_U) and (V, φ_V) are called C^∞ compatible or smoothly compatible[1] *if $U \cap V = \emptyset$ or if the transition map is of class C^∞, that is, it is a diffeomorphism.*

Definition D.1.3. *A C^∞ atlas or a smooth atlas $\mathcal{A}$ is a collection of coordinate charts $\{U_\lambda, \varphi_l : U_\lambda \to \mathbb{R}^n\}$ such that:*

1. $\bigcup_\lambda U_\lambda = X$
2. *Any two coordinate charts in $\mathcal{A}$ are smoothly compatible with each other.*

Two atlases are equivalent if their intersection is still an atlas.

We remind that if the transition maps are real analytic, then the manifold is real analytic and essentially prove

Theorem D.1.1. *$G_k(\mathbb{R}^n)$ is a connected compact smooth*[2] *manifold of dimension $k(n-k)$.*

D.2 Topology of the Grassmannian

To prove Theorem D.1.1 we must provide a topology to $G_k(\mathbb{R}^n)$ and to do it we consider the set $V_k(\mathbb{R}^n)$ of all the k-tuples of linearly independent vectors in $\mathbb{R}^n$.

Lemma D.2.1. *$V_k(\mathbb{R}^n)$ is an open set in $\mathbb{R}^{nk}$*

Proof. A k-tuple of linearly independent vectors $\{v_1, ..., v_k\} \subseteq \mathbb{R}^{nk}$ is uniquely determined by a matrix $M \in \mathcal{M}_{n,k}(\mathbb{R})$ [3] and this matrix has maximum rank k, that is, there exists a not null minor from a $k \times k$ submatrix of M. We call this submatrix M' and consider the following diagram of continous functions:

$$V_k(\mathbb{R}) \xrightarrow{\alpha} M_{n,k}(\mathbb{R}) \xrightarrow{minor} M_{k,k}(\mathbb{R}) \xrightarrow{\det} \mathbb{R}$$

[1] We have to definte compatible functions because smoothness (and in general differentiability) is not invariant under homeomorphisms, this is why we need smooth charts

[2] With smooth we mean C^∞

[3] With $\mathcal{M}_{n,k}(\mathbb{R})$ we mean the space of $n \times k$ matrices with coefficients in $\mathbb{R}$

such that

$$\{v_1, ..., v_k\} \xrightarrow{\alpha} M \xrightarrow{minor} M' \xrightarrow{\det} \det M'$$

then $V_k(\mathbb{R}^n)$ is the preimage of the composition of continous functions of the open set $\mathbb{R} \setminus \{0\}$, which means that it is an open set in $\mathbb{R}^{nk}$. □

We now consider the surjection[4]

$$\begin{aligned} q: \quad V_k(\mathbb{R}) &\to \ G_k(\mathbb{R}^n) \\ \{v_1, ..., v_k\} &\to \operatorname{span}(v_1, ..., v_k) \end{aligned}$$

and endow the grassmanian $G_k(\mathbb{R}^n)$ with the quotient topology $\mathcal{T}$, setting

$$A \subseteq G_k(\mathbb{R}^n) \text{ is open in } \mathcal{T} \iff q^{-1}(A) \subseteq V_k(\mathbb{R}^n) \text{ is open in } \mathbb{R}^{nk}$$

However, if we consider the set $V_k^0(\mathbb{R}^n)$[5] of k-tuples of orthonormal vectors in $\mathbb{R}^n$ and the surjection $q_0 = q_{|_{V_k^0(\mathbb{R}^n)}}$, we likewise endow $G_k(\mathbb{R}^n)$ with the quotient topology $\mathcal{T}_0$, that is

$$A \subseteq G_k(\mathbb{R}^n) \text{ is open in } \mathcal{T}_0 \iff q_0^{-1}(A) \subseteq V_k^0(\mathbb{R}^n) \text{ is open in } \mathbb{R}^{nk}$$

We state omitting the proof the following

Lemma D.2.2. *The topological spaces $(G_k(\mathbb{R}^n), \mathcal{T})$ and $(G_k(\mathbb{R}^n), \mathcal{T}_0)$ are homeomorphic and an homeomorphism is given by the identitymap, that is, $\mathcal{T} = \mathcal{T}_0$*

Lemma D.2.3. *The maps $q : V_k(\mathbb{R}^n) \to G_k(\mathbb{R}^n)$ and $q_0 : V_k^0(\mathbb{R}^n) \to G_k(\mathbb{R}^n)$ are open.*

Now we can prove that the Grassmannian is a Hausdorff space and satisfies the second axiom of countability, that is, the topology of the Grassmannian has a countable base.

Lemma D.2.4. *Let $(X, \widetilde{\mathcal{T}})$ be a topological space. If for every $x \neq y$ such that $x, y \in X$ there exists a continous map $f : X \to \mathbb{R}$ such that $f(x) \neq f(y)$ then X is an Hausdorff space.*

Proof. Let $\varepsilon := |f(x) - f(y)| > 0$. If we set

$$U = \left(f(x) - \frac{\varepsilon}{3}, f(x) + \frac{\varepsilon}{3}\right) \quad \text{and } V = \left(f(y) - \frac{\varepsilon}{3}, f(y) + \frac{\varepsilon}{3}\right)$$

then it is clear that U and V are disjoint. Furthermore, f is continous then $f^{-1}(U)$ is an open neighbourhood of x and $f^{-1}(V)$ is an open neighbourhood

[4]By $\operatorname{span}(v_1, ..., v_k)$ we mean the space spanned by the vectors $v_1, ..., v_k$
[5]These structures are called Stiefel's manifolds

of y and they are disjoint, that is we found two disjoint open neighbourhoods of y. We found two disjoint open neighbourhoods for every pair $x, y \in X$ then X is an Hausdorff space.

□

With this Lemma we can immediately prove

Lemma D.2.5. *$G_k(\mathbb{R}^n)$ is a topological Hausdorff space.*

Proof. Let fix two subspaces $X, Y \in G_k(\mathbb{R}^n)$ such that $X \neq Y$ and choose $x \in X \setminus Y$. We define

$$\rho_w(X) : G_k(\mathbb{R}^n) \to \mathbb{R} \text{ such that } \ \rho_w(T) := dist(w, T)$$

then it is $\rho_w(X) = 0$ and $\rho_w(Y) \neq 0$. Let consider the diagram

$$\begin{tikzcd} V_k^0(\mathbb{R}^n) \arrow[r, "q_0"] \arrow[dr, "\widetilde{\rho_w}"'] & G_k(\mathbb{R}^n) \arrow[d, "\rho_w"] \\ & \mathbb{R} \end{tikzcd}$$

where

$$\widetilde{\rho_w} : V_k^0(\mathbb{R}^n) \to \mathbb{R} \text{ is such that } \ \widetilde{\rho_w}(v_1, ..., v_k) := \left\| w - \sum_{i=1}^{k} < w, v_i > \cdot v_i \right\|$$

The map $\widetilde{\rho_w}$ is continous and makes the previous diagram commute. By Lemma D.2.3 and by the commutativity of the diagram above, it is

$$A \subseteq \mathbb{R} \text{ open } \implies \widetilde{\rho_w}^{-1}(A) \text{ open in } V_k^0(\mathbb{R}^n)$$
$$\implies q_0\left(\widetilde{\rho_w}^{-1}(A)\right) \text{ open in } \ G_k(\mathbb{R}^n) \implies \widetilde{\rho_w}^{-1}(A) \text{ is open in } G_k(\mathbb{R}^n)$$

that is, ρ_w is continous and then by Lemma D.2.4 the proof is completed.

□

Now, to prove that $G_k(\mathbb{R}^n)$ satisfies the second axiom of countability we start with

Lemma D.2.6. *Let X and Y be two topological spaces such that X has a countable base and let $f : X \to Y$ be a continous surjective open map. Under these conditions Y has a countable base.*

Proof. Let $\mathcal{B}$ be a countable base of open sets of X. We now define $\mathcal{B}' := \{f(A) : A \in \mathcal{B}\}$ and show that this is a countable base for Y. Firstly, $f(A)$ is open in Y for every $A \in \mathcal{B}$ since f is an open map. Furthermore, if $\mathcal{B}$ is open in Y then by the continuity of f it is that $f^{-1}(B)$ is open in X and for the second axiom of countability there exist countably open sets $\{A_i\}_{i\in I}$ such that

$$f^{-1}(B) = \bigcup_{i\in I} A_i$$

and then recalling that f is surjective, we have

$$B = f\left(f^{-1}(B)\right) = f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f\left(A_i\right)$$

where the lasts are elements of $\mathcal{B}'$ and then $\mathcal{B}'$ is a countable base for Y. □

Corollary D.2.7. *$G_k(\mathbb{R}^n)$ is a topological space that satisfies the second axiom of countability.*

Proof. We use Lemma D.2.6 setting $X = V_k(\mathbb{R}^n)$ which has a countable base since it is a subspace of $\mathbb{R}^{nk}$, $Y = G_k(\mathbb{R}^n)$ and the map $q : V_k(\mathbb{R}^n) \to G_k(\mathbb{R}^n)$. □

D.2.1 Compactness and connectedness of $G_k(\mathbb{R}^n)$

i) **Compactness**

Lemma D.2.8. *$V_k^0(\mathbb{R}^n)$ is compact in $\mathbb{R}^{nk}$.*

Proof. An element of $V_k^0(\mathbb{R}^n)$ is a k-tuple of orthonormal vectors denoted as usual as $\{v_1, ..., v_k\}$, that is, it can be represented as a matrix $M \in \mathcal{M}_{n,k}(\mathbb{R})$ with v_i as column. These vectors are orthonormal so they have unit norm, that is, $V_k^0(\mathbb{R}^n)$ is a bounded set.
Let now $f : \mathcal{M}_{n,k}(\mathbb{R}) \to \mathcal{M}_{k,k}(\mathbb{R})$ be the continous function such that $f(A) = {}^tAA - \mathbb{1}$ where $\mathbb{1}$ is the $k \times k$ identity matrix. Since $\{v_1, ..., v_k\}$ are orthonormal then $v_i \cdot v_j = 0$ if $i \neq j$ and then, following the previous notation, it is ${}^tMM = \mathbb{1}$ which means that if 0_k is the $k \times k$ null matrix, then $V_k^0(\mathbb{R}^n) = f^{-1}(0_k)$, that is, $V_0^k(\mathbb{R}^n)$ is a closed set.

□

Corollary D.2.9. *$G_k(\mathbb{R}^n)$ is a topological compact space.*

Proof. $G_k(\mathbb{R}^n)$ is a continous image of a compact set through the surjection $q_0 : V_k^0(\mathbb{R}^n) \to G_k(\mathbb{R}^n)$. $\square$

ii) **Connectedness**

Lemma D.2.10. *$V_k(\mathbb{R}^n)$ is path-connected.*

Proof. Let consider two elements in $V_k(\mathbb{R}^n)$. They are two k-tuples of linearly indipendent vectors $\{v_1, ..., v_k\}$ and $\{w_1, ..., w_k\}$ that can be extended to two basis of $\mathbb{R}^n$ respectively $\bar{V} = \{v_1, ..., v_k, v_{k+1}, ..., v_n\}$ and $\bar{W} = \{w_1, ..., w_k, w_{k+1}, ..., w_n\}$. Let M be the matrix that has the vectors of $\bar{V}$ a s columns and let N be the matrix with the vectors of $\bar{W}$ as columns. Without loss of generality we can assume that

$$M, N \in GL_n(\mathbb{R})^+ := \{A \in \mathcal{M}_{n,n}(\mathbb{R}) \text{ such that } \det A > 0\}$$

and such a space is path-connected. It follows that there exists a continous path γ such that

$$\gamma : [0,1] \to GL_n(\mathbb{R})^+ \text{ such that } \gamma(0) = M \text{ and } \gamma(1) = N$$

If we now consider the projection $\pi : GL_n(\mathbb{R}) \to V_k(\mathbb{R}^n)$ that selects the first k columns of the matrix A, we have that π is clearly continous and then it is continous also the composition $\pi \circ \gamma : [0,1] \to V_k(\mathbb{R}^n)$. This is a continous path that joins the two elements of $V_k(\mathbb{R}^n)$, in fact for every $t \in [0,1]$ we have that $\gamma(t)$ is an invertible matrix, that is,its columns are linearly indipendent. $\square$

Corollary D.2.11. *$G_k(\mathbb{R}^n)$ is a connected topological space.*

Proof. $G_k(\mathbb{R}^n)$ is the continous image of the path-connected space $V_k(\mathbb{R}^n)$ through the map q. $\square$

To conclude, in this section we proved

Theorem D.2.12. *$(G_k(\mathbb{R}^n), \mathcal{T}) \simeq (G_k(\mathbb{R}^n), \mathcal{T}_0)$ is a compact, connected and Hausdorff topological space with countable basis.*

D.3 Construction of smooth charts and a smooth atlas

We now give to the Grassmannian the structure of smooth manifold.

i) **Construction of smooth coordinate charts**
Let $X \in G_k(\mathbb{R}^n)$ be a k-dimensional subspace of $\mathbb{R}^n$. We look for an open neighbourhood of X homeomorphic to an open set in $\mathbb{R}^{k(n-k)}$ and a coordinate chart between them. We define

$$U = U_X := \left\{ Y \in G_k(\mathbb{R}^n) : Y \cap X^{\perp} = 0 \right\}$$

where $X^{\perp} = \{ y \in \mathbb{R}^n \text{ such that } \ x \cdot y = 0, \forall \ x \in X \}$.

Lemma D.3.1. *U is open in $(G_k(\mathbb{R}^n), \mathcal{T})$.*

Proof. $\mathcal{T}$ is the quotient topology and then

$$U \text{ is open in } \mathcal{T} \iff q^{-1}(U) \text{ is open in } V_k(\mathbb{R}^n)$$

Let fix an orthonormal basis $\{x_1, ..., x_{n-k}\}$ of $X^{\perp}$ and it is

$$q^{-1}(U) = \left\{ \{v_1, ..., v_k\} \in V_k(\mathbb{R}^n) : \operatorname{span}(v_1, ..., v_k) \cap X^{\perp} = 0 \right\}$$
$$\left\{ \ \{v_1, ..., v_k\} \in V_k(\mathbb{R}^n) : \{v_1, ..., v_k, x_1, ..., x_{n-k}\} \text{ are linearly indipendent } \right\}$$

We define the map $f : V_k(\mathbb{R}^n) \to \mathbb{R}$ such that to an element $\{v_1, ..., v_k\} \in V_k(\mathbb{R}^n)$ associates the determinant of an $n \times n$ matrix that has

$$v_1, ..., v_k, x_1, ..., x_{n-k}$$

as columns. Such a function is continous and then $q^{-1}(U) = f^{-1}(\mathbb{R} \setminus \{0\})$ is an open set in $V_k(\mathbb{R}^n)$ and the proof is completed. □

Let denote by $\pi : \mathbb{R}^n \to X$ and by $\pi_{\perp} : \mathbb{R}^n \to X^{\perp}$ the orthogonal projections on X and $X^{\perp}$ respectively, that is

$$\pi(x) = \begin{cases} x & \text{if } x \in X \\ 0 & \text{if } x \in X^{\perp} \end{cases} \qquad \pi_{\perp}(x) = \begin{cases} x & \text{if } x \in X^{\perp} \\ 0 & \text{if } x \in (X^{\perp})^{\perp} = X \end{cases}$$

We have the following

Lemma D.3.2. *$\pi_{|_Y} : Y \to X$ is an isomorphism of vector spaces, for every $Y \in U$.*

Proof. The projection is a linear map and both Y and X have the same dimension, then we only have to prove that $\pi_{|_Y}$ is injective. Let $y \in Y$ be such that $\pi_{|_Y}(y) = 0$; then in particular $\pi(y) = 0$, that is, $y \in X^{\perp}$ but U is a set such that $Y \cap X = 0$ for every $Y \in U$ then it is $y = 0$ which means that $\ker(\pi_{|_Y})$ is trivial so the function is injective. □

With this Lemma we know that $\pi_{|_Y}$ is an invertible functio, then we define[6]

$$\begin{aligned} S : U &\to \mathrm{Hom}(X, X^{\perp}) \\ Y &\to \pi_{\perp_{|_Y}} \circ (\pi_{|_Y})^{-1} \end{aligned}$$

that is, S is a map that provides another linear map that "reads" the vector $y \in Y$ in terms of elements of the direct sum $X \oplus X^{\perp}$. We fix a basis $\mathcal{V} = \{v_1, ..., v_k\}$ for X and another basis $\mathcal{U} = \{u_1, ..., u_{n-k}\}$ for $X^{\perp}$, then we can define an isomorphism of vector spaces

$$\Phi_{\mathcal{V}}^{\mathcal{U}} : \ \mathrm{Hom}(X, X^{\perp}) \to \mathcal{M}_{n-k,k}(\mathbb{R}) \simeq \mathbb{R}^{k(n-k)}$$

where Φ sends a linear map $\phi : X \to X^{\perp}$ to itstransformation matrix $M_{\mathcal{U},\mathcal{V}}(\phi)$ that is to the matrix that represents ϕ.
We consider the composition

$$T = T_{\mathcal{V}}^{\mathcal{U}} := \Phi_{\mathcal{V}}^{\mathcal{U}} \circ S$$

This map is candidated to be chosen as coordinate chart.

Lemma D.3.3. *The map $T : U \to \mathbb{R}^{k(n-k)}$ is continous.*

Proof. We look for a map $\widetilde{T} : q^{-1} \to \mathbb{R}^{k(n-k)}$ such that the following diagram commutes.

$$\begin{tikzcd} q^{-1}(U) \subseteq V_k(\mathbb{R}^n) \arrow[r, "q"] \arrow[dr, "\widetilde{T}"] & U \subseteq G_k(\mathbb{R}^n) \arrow[d, "T"] \\ & \mathbb{R}^{k(n-k)} \end{tikzcd}$$

that is, $T \circ q = \widetilde{T}$.

[6] With $\mathrm{Hom}(X, X^{\perp})$ we mean the set of linear maps from X to $X^{\perp}$.

Let consider $Y \in U$ and let fix an orthonormal basis $\mathcal{Y} = \{y_1, ..., y_k\}$ for Y[7]. Such a basis is an element of $q^{-1}(U)$ and since every y_j is an element in $\mathbb{R}^n = X \oplus X^\perp$ then there exists unique coefficients $\lambda_{i,j}$ with $i, j \in \{1, ..., k\}$ and $\mu_{i,j}$ with $j \in \{1, ..., k\}$ and $i \in \{1, ..., n-k\}$ such that

$$y_j = \sum_{i=1}^{k} \lambda_{i,j} v_i + \sum_{i=1}^{n-k} \mu_{i,j} u_i \quad \text{for every } j = \{1, ..., k\}$$

Since the projection is a liner map, it is

$$\pi_{|_Y}(y_j) = \pi_{|_Y}\left(\sum_{i=1}^{k} \lambda_{i,j} v_i + \sum_{i=1}^{n-k} \mu_{i,j} u_i\right) =$$
$$= \sum_{i=1}^{k} \lambda_{i_j} \pi_{|_Y}(v_i) + \sum_{i=1}^{n-k} \mu_{i,j} \pi_{|_Y}(u_i) = \sum_{i=1}^{k} \lambda_{i,j} v_i$$

We observe that the transformation matrix for π_Y is

$$\Lambda = \Lambda_{\mathcal{V},\mathcal{Y}}(\pi_{|_Y}) = (\lambda_{i,j})_{i,j=1,\dots,k} \tag{D.2}$$

and by the previous Lemma that matrix is invertible. On the other hand we have

$$\pi_{\perp_{|_Y}}(y_j) = \pi_{\perp_{|_Y}}\left(\sum_{i=1}^{k} \lambda_{i,j} v_i + \sum_{i=1}^{n-k} \mu_{i,j} u_i\right) =$$
$$= \sum_{i=1}^{k} \lambda_{i_j} \pi_{\perp_{|_Y}}(v_i) + \sum_{i=1}^{n-k} \mu_{i,j} \pi_{\perp_{|_Y}}(u_i) = \sum_{i=1}^{k} \mu_{i,j} u_i$$

and in this case the transformation matrix for the map $\pi_{\perp_{|_Y}}$ is

$$M = \mathcal{M}_{\mathcal{U},\mathcal{Y}}(\pi_{\perp_{|_Y}}) = (\mu_{i,j})_{\substack{i=1,\dots,n-k \\ j=1,\dots,k}} \tag{D.3}$$

Now, once we fix the basis $\mathcal{Y}$ for Y, we have that the choice of the coefficients $\lambda_{i,j}$ and $\mu_{i,j}$ is unique then we can define a map

$$\widetilde{T} : q^{-1}(U) \to \mathbb{R}^{k(n-k)} \quad \text{such that} \quad \mathcal{Y} = \{y_1, ..., y_k\} \to M\Lambda^{-1} \tag{D.4}$$

that is, $M\Lambda^{-1}$ is the transformation matrix of

$$\pi_{\perp_{|_Y}} \circ \left(\pi_{|_Y}\right)^{-1} \in \text{ Hom}(X, X^\perp)$$

so $\widetilde{T}$ is continous. The thesis follows as the last part of the proof of Lemma D.2.4. □

[7] We recall that the choice of the basis is not restrictive, since the change-of-basis matrix is invertible and the multiplication by an invertible matrix is an endomorphism on $q^{-1}(U)$

Lemma D.3.4. *The map T is invertible and the inverse is continous.*

Proof. Recalling that $T : \Phi_{\mathcal{V}}^{\mathcal{U}} \circ S$ we firstly have to prove that $\Phi_{\mathcal{V}}^{\mathcal{U}}$ and S are bijective. We immediately see that $\Phi_{\mathcal{V}}^{\mathcal{U}}$ is invertible and its inverse

$$\left(\Phi_{\mathcal{V}}^{\mathcal{U}}\right)^{-1} : \mathcal{M}_{n-k,k}(\mathbb{R}) \to \mathrm{Hom}(X, X^{\perp})$$

is such that associates to a matrix the linear map that it represents. Indeed we know that $\Phi_{\mathcal{V}}^{\mathcal{U}}$ is an isomorphism between the space of linear maps from X to $X^{\perp}$ and the matrix space $\mathcal{M}_{n-k,k}(\mathbb{R})$.
Now we need an inverse for S. Let fix a basis $\mathcal{X} = \{x_1, ..., x_k\}$ for X and then we define

$$S^{-1} : \mathrm{Hom}(X, X^{\perp}) \to U \quad \text{such that } S^{-1}(\varphi) = \mathrm{span}(\mathcal{X} + \varphi(\mathcal{X}))$$

where

$$\mathrm{span}(\mathcal{X} + \varphi(\mathcal{X})) = \mathrm{span}(x_1 + \varphi(x_1), ..., x_k + \varphi(x_k))$$

Now fixing $\varphi : X \to X^{\perp}$, we set

$$Y := S^{-1}(\varphi) = \mathrm{span}(\mathcal{X} + \varphi(\mathcal{X}))$$

and in particular we observe that $Y \in U$ because every vector $y_j = x_j + \varphi(x_j)$ has a non-zero component x_j since $\mathcal{X}$ is a basis and then $Y \cap X^{\perp} = 0$. It is

$$S \circ S^{-1}(\varphi) = \pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}$$

but $\pi_{|Y}$ and $\pi_{\perp_{|Y}}$ act on the vectors in the form $x_j + \varphi(x_j)$ in the following way[8]

$$\pi_{|Y}\left(x_j + \varphi(x_j)\right) = x_j \qquad \text{since } \varphi(x_j) \in X^{\perp} \tag{D.5}$$

$$\pi_{\perp_{|Y}}\left(x_j + \varphi(x_j)\right) = \varphi(x_j) \qquad \text{since } x_j \in X \tag{D.6}$$

so we deduce that the map described by equation D.5 acts like the identity map on the set X, that is, if

$$\mathcal{B} = \{x_1 + \varphi(x_1), ..., x_k + \varphi(x_k), y_1, ..., y_{n-k}\}$$

is a basis for $\mathbb{R}^n$ (we completed the basis $\mathcal{Y}$) then the transformation matrix of π, that is, $M_{\mathcal{X},\mathcal{B}}(\pi)$ contains as submatrix the $k \times k$ identity

[8] These vectors are a basis for Y

matrix.
On the other hand we have

$$\pi_{\perp_{|Y}}(x_j + \varphi(x_j)) = \varphi(x_j) \;\text{ since } x_j \in X$$

then $\pi_{\perp_{|Y}}$ coincides with φ on X so we can conclude that

$$\pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1} = \varphi \circ \mathbb{1}_X = \varphi \Rightarrow S \circ S^{-1} = \varphi$$

that is, S^{-1} is a right inverse for S. To show that it is also a left inverse we prove that $S^{-1} \circ S : U \to U$ is the identity map on U.
Let then consider $Y \in U$. We already know by Lemma D.3.2 that $\pi_{|Y}$ is an isomorphism between the vector spaces Y and X then

$$\forall\, y \in Y \;\exists!\; x \in X \;:\; y = x + \pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}(x)$$

and in that way we obtain a basis for Y given by

$$y_j = x_j + \pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}(x_j) \;\text{ for every }\; j = 1, ..., k \tag{D.7}$$

But then we have

$$\begin{aligned} S^{-1} \circ S(Y) &= S^{-1}\left(\pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}\right) = \\ &= \operatorname{span}\left(x_1 + \pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}(x_1), ..., x_k + \pi_{\perp_{|Y}} \circ \left(\pi_{|Y}\right)^{-1}(x_k)\right) \\ &= \operatorname{span}(y_1, ..., y_k) = Y \end{aligned}$$

that is, S^{-1} is a left inverse for S.
We can choose $T^{-1} : \mathbb{R}^{k(n-k)} \to U$ where $T^{-1} := S^{-1} \circ \left(\Phi^{\mathcal{U}}_{\mathcal{V}}\right)^{-1}$ as an inverse for T and then the first part of the Lemma is proved.

Now, to show that T^{-1} is continous we prove the equivalent condition that $\widetilde{T}$ defined in D.4 is an open map. Indeed, if $\widetilde{T}$ is open then we can choose an open set $A \subseteq U$ and then $q^{-1}(A)$ is open in $V_k(\mathbb{R}^n)$ since q is continous and $\widetilde{T}\left(q^{-1}(A)\right) = T(A)$ is open in $\mathcal{M}_{n-k,k}(\mathbb{R})$ since $\widetilde{T}$ is open and if T is open then T^{-1} is continous. Following the usual notation we set for every $j \in \{1, ..., k\}$

$$y_j = \sum_{i=1}^{k} \lambda_{i,j} v_i + \sum_{i=1}^{n-k} \mu_{i,j} u_i \qquad Y = \operatorname{span}(y_1, ..., y_k) \in U$$

If we take $\{\mathcal{V},\mathcal{U}\}$ as a basis for $\mathbb{R}^n$ then the matrix that represents the vectors y_j is

$$N = \begin{pmatrix} \lambda_{1,1} & \cdots & \lambda_{1,k} \\ \vdots & \ddots & \vdots \\ \kappa_{k,1} & \cdots & \lambda_{k,k} \\ \mu_{1,1} & \cdots & \mu_{1,k} \\ \vdots & \ddots & \vdots \\ \mu_{n-k,1} & \cdots & \mu_{n-k,k} \end{pmatrix} \tag{D.8}$$

and following the notation used in the previous Lemma recalling (D.2) and (D.3) we write

$$N = \begin{pmatrix} \Lambda \\ M \end{pmatrix}$$

and for D.4 it is

$$\widetilde{T}(N) = M \cdot \Lambda^{-1}$$

We now consider the projections[9]

$$p_1 : \mathcal{M}_{n,k}(\mathbb{R}) \to \mathcal{M}_{n-k,k}(\mathbb{R}) \text{ such that } p_1(N) = p_1 \begin{pmatrix} \Lambda \\ M \end{pmatrix} = M$$

$$p_2 : \mathcal{M}_{n,k}(\mathbb{R}) \to GL_k(\mathbb{R}) \text{ such that } p_2(N) = p_2 \begin{pmatrix} \Lambda \\ M \end{pmatrix} = \Lambda$$

Let now consider the right action of the group $GL_k(\mathbb{R})$

$$g : \mathcal{M}_{n-k,k}(\mathbb{R}) \times GL_k(\mathbb{R}) \to \mathcal{M}_{n-k,k}(\mathbb{R}) \text{ such that } g(M, A) = M \cdot A^{-}1$$

this action is open.
If $N \in q^{-1}(U)$ is the matrix (D.8) then

$$\widetilde{T}(N) = g\left(p_1(N), p_2(N)\right)$$

that is, it is a composition of open maps and then it is open. □

From this Lemma it follows that every $k-$dimensional subspace X of $\mathbb{R}^n$ has a neighbourhood U homeomorphic to an open set in $\mathbb{R}^{k(n-k)}$ and the homeomorphism is given by the map T, that is, (U, T) is a coordinate chart for $G_k(\mathbb{R})$. We proved

[9]We could identify the matrix space $M_{n,k}(\mathbb{R})$ as the product $M_{n-k,k}(\mathbb{R}) \times M_k(\mathbb{R})$ so the following are the projections othe first and on the second componen, respectively.

Theorem D.3.5. *$G_k(\mathbb{R}^n)$ is a compact and connected $k(n-k)-$dimensional topological manifold.*

ii) **Construction of an atlas**
WE want to provide to $G_k(\mathbb{R}^n)$ an atlas as in Definition D.1.3 so we have to prove that two coordinate charts are smoothly compatible, that is, taht the transition map is smooth. we firstly remind that if U_X is an open neighbourhood of $X \in G_k(\mathbb{R}^n)$ and $\mathcal{V} = \{v_1, ..., v_k\}$ is a basis for X then given a coordinate chart $T : U_X \to \mathbb{R}^{k(n-k)}$ and a matrix $A \in \mathcal{M}_{n-k,k}(\mathbb{R}) \simeq \mathbb{R}^{k(n-k)}$ then the inverse map T^{-1} is defined as

$$T^{-1}(A) = \mathrm{span}(\mathcal{V} + A\mathcal{V}) \tag{D.9}$$

where A is a matrix that represents the linear map $L_A \in \mathrm{Hom}(X.X^{\perp})$. Recalling (D.7) we know that there exists a unique basis $\mathcal{Y} = \{y_1, ..., y_k\}$ of $Y \in U_X$ such that $\pi(y_j) = v_j$ for every $j \in \{1, ..., k\}$ and that holds the identity

$$y_j = v_j + S\left(Y\right)\left(v_j\right)$$

where $S(Y) = \pi_{\perp_{|_Y}} \circ \left(\pi_{|_Y}\right)^{-1}$.

Now let X_0 and X_1 be two elements of $G_k(\mathbb{R}^n)$ and let U_0 and U_1 be two open neighbourhoods of X_0 and X_1 respectively with non-empty intersection, i.e.:

$$U_0 = \left\{Y \in G_k(\mathbb{R}^n) \ : \ Y \cap X_0^{\perp} = \{0\}\right\}$$
$$U_1 = \left\{Y \in G_k(\mathbb{R}^n) \ : \ Y \cap X_1^{\perp} = \{0\}\right\} \qquad U_0 \cap U_1 \neq \emptyset$$

Let us now fix an element $Y \in U_0 \cap U_1$ and choose bases $\mathcal{V}_0 = \{v_1^0, ..., v_k^0\}$ and $\mathcal{U}_0 = \left\{u_1^0, ..., u_{n-k}^0\right\}$ for X_0 and for $X_0^{\perp}$ respectively[10]. We now denote by v_j^1 the projection on X_1 of the vector $v_j^0 + S\left(Y\right)\left(v_j^0\right)$. We state that

$$\mathcal{V}_1 = \left\{v_1^1, ..., v_k^1\right\}$$

[10]It is important to highlight that the choice of the bases does not change the result of our proof, that is, if two coordinate charts are smoothly compatible with two fixed bases V and V' then if A is the transformation matrix for some map $\varphi : X \to X^{\perp}$, i.e. $A = M_{V,U}(\varphi)$, then the transformation matrix $B = M_{V',U'}(\varphi)$ is simply $B = PAQ^{-1}$ where P is the matrix that changes basis from U to U' and Q changes basis from V to V'. Such a map is clearly smooth.

is a basis for X_1. Actually suppose that there exist coefficients $c_j \in \mathbb{R}$ such that

$$\sum_{j=1}^{k} c_j v_j^1 = 0$$

Let denote by π_1 the projection onto X_1 then it is

$$0 = \sum_{j=1}^{k} c_j v_j^1 = \sum_{j=1}^{k} c_j \pi_1 \left(v_j^0 + S\left(Y\right)\left(v_j^0\right)\right) = \pi_1 \left(c_j \sum_{j=1}^{k} [v_j^0 + S\left(Y\right)\left(v_j^0\right)] \right)$$

which means that

$$c_j \sum_{j=1}^{k} [v_j^0 + S\left(Y\right)\left(v_j^0\right)] \in X_1^{\perp}$$

but $X_1^{\perp}$ has trivial intersection with U_1 that contains Y so in particular $v_j \in U_1$ for every j and then it follows that

$$c_j \sum_{j=1}^{k} [v_j^0 + S\left(Y\right)\left(v_j^0\right)] = 0$$

If there exists one $c_{j_0} \neq 0$ then

$$v_{j_0}^0 = -S(Y)\left(v_{j_0}^0\right)$$

which is a contradiction. Indeed, since $v_{j_0} \in X_0$ and $S(Y)\left(v_{j_0}^0\right) \in X_0^{\perp}$, this implies that $v_{j_0}^0 = 0$ but this is impossible since v_j^0 is an element of the basis of X_0 so we conclude that $c_j = 0$ for every j, that is, v_j^1 are linearly indipendent.
It is well defined the isomorphism

$$\psi := \pi_{|_Y}^1 \circ \left(\pi_{|_Y}^0\right)^{-1}$$

since the maps π^0 and π^1 (the orthogonal projections on X_0 and X_1 respectively) are isomorphisms from Lemma D.3.2.
We set $\mathcal{V}_1 = \{v_1^1, ..., v_k^1\}$ as a basis for X_1 and $\mathcal{U}_1 = \left\{u_1^1, ..., u_{n-k}^1\right\}$ is an arbitrary orthonormal basis for $X_1^{\perp}$. We are now ready to prove the last

Theorem D.3.6. *The coordinate charts (U_0, T_0) and (U_1, T_1) are smoothly compatible.*

Proof. We consider two open sets $T_0(U_0 \cap U_1)$ and $T_1(U_0 \cap U_1)$ of $\mathbb{R}^{k(n-k)} \simeq \mathcal{M}_{n-k,k}(\mathbb{R})$ and prove that the transition map

$$\begin{aligned} T_1 \circ T_0^{-1} : \ & T_0(U_0 \cap U_1) \to T_1(U_0 \cap U_1) \\ & A \qquad\quad \to T_1(T_0^{-1}(A)) \end{aligned} \tag{D.10}$$

is smooth. The map (D.10) sends the matrix A to another matrix B that represents the linear map

$$L_B = \pi^1_{\perp|_{\text{span}(\mathcal{V}_0 + A\mathcal{V}_0)}} \circ \left(\pi^1_{|_{\text{span}(\mathcal{V}_0 + A\mathcal{V}_0)}}\right)^{-1}$$

In fact, since $A \in T_0(U_0 \cap U_1) \subseteq \mathcal{M}_{n-k,k}(\mathbb{R})$ then there exists a unique $Y \in U_0 \cap U_1$ such that $A = T(Y)$ and recalling (D.9) we have that $Y = \text{span}(\mathcal{V}_0 + A\mathcal{V}_0)$ and a basis for Y is given by

$$\mathcal{Y} = \left\{y_j = V_j^0 + Av_j^0\right\}_{j=1}^k$$

but since for Lemma D.3.2 we know that π^1 is an isomorphism between Y and X_1 it follows that for every j

$$y_j = v_j^1 + \pi^1_{\perp_{|_Y}} \circ \left(\pi^1_{|_Y}\right)^{-1} \left(v_j^1\right)$$

where we recall that $L_A = \pi^0_{\perp_{|_Y}} \circ \left(\pi^0_{|_Y}\right)^{-1}$ and where we defined

$$v_j^1 = \psi(v_j^0) = \pi^0_{\perp_{|_Y}} \circ \left(\pi^1_{|_Y}\right)^{-1} (v_j^0)$$

It follows that

$$y_j = v_j^0 + Av_j^0 = \psi(v_j^0) + \pi^1_{\perp_{|_Y}} \circ \left(\pi^1_{|_Y}\right)^{-1} \left(\psi(v_j^0)\right)$$

that is

$$\begin{aligned} &\psi^{-1}\left(v_j^1\right) + A\psi^{-1}(v_j^1) = v_j^1 + S_1\left(Y\right)\left(v_j^1\right) \\ \Rightarrow\ & S_1\left(Y\right)\left(v_j^1\right) = \psi^{-1}(v_j^1) + A\psi^{-1}(V_J^1) - v_j^1 \end{aligned}$$

and the last means that the map S_1 is the map

$$S_1 = \psi^{-1} + L_A \circ \psi^{-1} - \text{id}_X$$

If now we associate to these maps the diffeomorphism

$$\Phi^{\mathcal{V}'}_{\mathcal{U}'} : \text{Hom}(X_1, X_1^\perp) \to \mathcal{M}_{n-k,k}(\mathbb{R})$$

it is clear that the transition map $A \to T_1\left(T_0^{-1}(A)\right) = T_1(Y)$ acts in a smooth way on the coefficients of A since it is a composition of diffeomorphic maps.

$\square$

We found the atlas

$$\mathcal{A} = \left\{ \left(U_X, T_{\mathcal{V}}^{\mathcal{U}} : U_X \to \mathbb{R}^{k(n-k)} \right), X \in GL_k(\mathbb{R}^n) \right\} \tag{D.11}$$

where $X = \operatorname{span}(\mathcal{V})$ and $X^{\perp} = \operatorname{span}(\mathcal{U})$.
We proved

Theorem D.3.7. $(G_k(\mathbb{R}^n), [\mathcal{A}])$ *is a smooth connected and compact manifold of dimension* $k(n-k)$.

Bibliography

[1] Arnol'd V.I. *Instability of dynamical systems with several degrees of freedom*, Soviet Math. Doklady **5** (1964), 581-585.

[2] Arnol'd V.I., Kozlov V.V., Neishtadt A.I. *Mathematical Aspects of Classical and Celestial Mechanics* Third Edition. Springer-Verlag, Berlin (2007).

[3] Arnol'd V.I. *Mathematical Methods of Classical Mechanics* Nauka, Moscow (1974) translated by K. Vogtmann and A. Weinstein. Graduate texts in mathematics , Vol. 60, Springer-Verlag, New York (1978).

[4] Arnol'd V.I. *Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian*, Uspekhi Mat. Nauk **18** (1963), No.5 13-40; Russ. Math. Surv. **18** (1963), No.5, 9-36.

[5] Bierstone E., Milman P.D. *Semianalytic and subanalytic sets*, Publications mathématiques de l'I.H.É.S., **67** (1988), 5-42.

[6] Bochnak J. and Risler J.J. *Sur les exposants de Lojasiewicz.*, Commentarii mathematici Helvetici **50** (1975): 493-508.

[7] Bochnak J., Coste M. and Roy M.F. *Real albegraic geometry*, Springer-Verlang, Berlin, (1998).

[8] Guzzo M. *On the steepness of three-jet non-degenerate functions in* $\mathbb{R}^n$. Personal communucation., 2014.

[9] Guzzo M., Chierchia L., Benettin G. *The steep Nekhoroshev's Theorem and optimal stability exponents (an announcement)*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. **25** (2014), 293-299.

[10] Gwoździewicz J. *The Lojasiewicz exponent of an analytic function at an isolated zero*, Comment. Math. Helv. **74** (1999), no. 3, 364-375.

[11] Hatcher A. *Algebraic Topology*. Cambridge University Press (2002).

[12] Hironaka H. *Subanalytic sets* in Number Theory, Algebraic Geometry and Commutative Algebra, volume in honor of Y. Akizuki, Kinokuniya, Tokyo, 1973, pp. 453-493.

[13] Il'yashenko Y.S. *A steepness test for analytic functions*, Uspekhi Math. Nauk, **41** (1986), 193-194; Russian Math. Surveys **41** (1986), 229-230.

[14] Lee J. *Introduction to Smooth Manifolds (Graduate Texts in Mathematics, Vol. 218)* Second Edition, Springer Science & Business Media , August 2012.

[15] Lochak P. *Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian*, Chaos, **2** (1992), 495-499.

[16] Milnor J.W. *Singular Points of complex Hypersurfaces* Princeton University Press, 1968.

[17] Nekhoroshev N.N. *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems I*, Uspekhi Mat. Nauk **32** (1977), 5-66; Russian Math. Surveys **32** (1977), 1-65.

[18] Nekhoroshev N.N. *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems II*, Tr. Semin. Petrovsk. **5** (1979), 5-50; In *Topics in Modern Mathematics, Petrovskii Seminar*, No. 5, edited by O.A. Oleinik (Consultant Bureau, New York, 1985).

[19] Nekhoroshev N.N. *Behavior of Hamiltonian systems close to integrability*, Funct. Anal. Appl, **5** (1971), 338-339. (Funk. An. Ego Prilozheniya, **5** (1971),82-83).

[20] Nekhoroshev N.N. *Stable lower estimates for smooth mappings and for gradients of smooth functions*, Math USSR Sbornik, **19**, 1973, 425-467.

[21] Niederman L. *Hamiltonian stability and subanalytic geometry*, Annales de l'Institut Fourier, Grenoble **56**, 3(2006), 795-813.

[22] Pöschel J. *A Lecture on the Classical KAM Theorem* Proc. Symp. Pure Math **69** (2001), 707-732.

[23] Pöschel J. *Nekhoroshev estimates for quasi-convex Hamilronian systems*, Math. Z. **213** (1993), 187-217.

[24] Schirinzi G. and Guzzo M. *On the formulation of new explicit conditions for steepness from a former result of N.N.Nekhoroshev.* J. Math. Phys. **54**, (2013).

[25] Schirinzi G. *Investigation of new conditions for steepness from a former result by Nekhoroshev*, Ph. D. Thesis, Padova, 2014.

[26] Tamm M. *Subanalytic sets in the calculus of variation* Acta Math., **146** (1981), pp. 167-199.

Acknowledgments

Ringrazio per prima la mia famiglia, la quale mi ha permesso di studiare e mi ha sostenuta nel completare questo percorso.
Ringrazio le mie colleghe e colleghi matematici, ma in primo luogo grandi amici, con i quali ho condiviso momenti ed emozioni indimenticabili.
Ma ringrazio piú di tutti il mio relatore, il professor Luigi Chierchia, che mi ha sopportata con estrema calma e professionalitá per tutto questo tempo.
Questa tesi é per tutti voi.
Con affetto,
Martha
Roma, 26 Giugno 2015.