SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Mathematics Area - PhD course in
Mathematical Analysis, Modelling, and Applications

Sard properties for polynomial maps in infinite dimension and applications to sub-Riemannian geometry

Candidate:
Daniele Tiberio

Advisors:
Prof. Antonio Lerario
Prof. Luca Rizzi

Academic Year 2023-24

Abstract

Some classical and fundamental results in analysis and geometry fail to hold in spaces of infinite dimension, leading to striking phenomena that require new investigations. We address two distinct questions of this type. To each of them is dedicated a part of the thesis: the first part contains the main work of my Ph.D. studies, while the second one is a complementary work that I did during the same years. Here we give a brief summary, in order to give a flavour of the topics. Each part contains a detailed introduction.

Part I

Sard properties for polynomial maps in infinite dimension and applications to sub-Riemannian geometry.

It is well-known that the classical Morse-Sard theorem is false for smooth maps from an infinite dimensional Hilbert space to $\mathbb{R}$, even under the assumption that the map is "polynomial", and a general theory is still missing. In this part we address this issue, providing sharp quantitative criteria for the validity of Sard-type theorems for polynomial maps from an infinite dimensional Hilbert space to $\mathbb{R}^m$. As an application, we present new advances on the sub-Riemannian Sard conjecture in Carnot groups.

The research presented in this part appears in the following preprints:

- A. Lerario, L. Rizzi, D. Tiberio, Sard properties for polynomial maps in infinite dimension, *arxiv:2407.02296*
- A. Lerario, L. Rizzi, D. Tiberio, Quantitative approximate definable choices, *arxiv:2409.14869*

Part II

Vanishing geodesic distances and the Michor-Mumford conjecture in Hilbertian H-type groups.

It is well-known that for weak Riemannian metrics on infinite dimensional manifolds the geodesic distance may not be a genuine distance, indeed it can be zero on distinct points. In their 2005 paper, Michor and Mumford conjectured that the degeneracy of the geodesic distance is related to the local unboundedness of the sectional curvature. In this part of the thesis, we introduce Heisenberg-type Lie groups modelled on Hilbert spaces, and we show that in this setting the degeneracy of the geodesic

distance and the local unboundedness of the sectional curvature coexist for a large class of weak Riemannian metrics.

The research presented in this part appears in the following preprints and publications:

- V. Magnani, D. Tiberio, On the Michor-Mumford phenomenon in the infinite dimensional Heisenberg group, *Revista Matematica Complutense*
- V. Magnani, D. Tiberio, The Michor-Mumford conjecture in Hilbertian H-type groups, *arxiv:2404.04583*

Contents

Part I

Sard properties for polynomial maps in infinite dimension and applications to sub-Riemannian geometry

Introduction

i Motivations: the Morse-Sard problem in infinite dimension and the Sard conjecture

The Morse-Sard theorem [77, 84] states that the set of critical values of a smooth map $f : N \to M$, where N and M are finite dimensional manifolds, has measure zero in M. Smale [88] proved a version of the Morse-Sard theorem in the case both M and N are infinite dimensional Banach manifolds and the map f is Fredholm: in this case the conclusion is that the set of regular values of f is residual in N. However, when N is infinite-dimensional and M is finite dimensional, f cannot be Fredholm and Smale's result cannot be applied. In fact, there are smooth surjective maps without regular points from any infinite-dimensional Banach space to $\mathbb{R}^2$, see [15]. Even under the hypothesis that f is a *polynomial map*, the set of its critical values can be of positive measure: Kupka [58] constructed an example of a smooth map $f : \ell^2 \to \mathbb{R}$, whose restriction to each finite dimensional space is a polynomial of degree 3, and with the property that the set of its critical values is the segment $[0, 1]$. However, Yomdin proved that if a map $f : \ell^2 \to \mathbb{R}$ can be approximated well-enough with finite dimensional polynomials, then the set of its critical values has measure zero, see [95]. (Kupka's map does not have this property.) Following this approach, Yomdin constructed a family of polynomials on ℓ^2 with the *Sard property*, that is, the set of their critical values has measure zero (here by polynomial on ℓ^2 we mean that the restriction to any finite dimensional linear subspace is a polynomial of some fixed degree). To sum up, already for polynomials on ℓ^2 the situation in subtle: on the one hand Kupka's counterexample shows that the Sard property does not hold in general for these maps, on the other hand Yomdin proved that it holds under the further assumption that the map is well-approximated with finite-dimensional polynomials. A general theory in this framework is still missing. Furthermore, in Yomdin's proof, the fact that the codomain is one-dimensional is essential and new technical difficulties arise for maps with values in $\mathbb{R}^m$, $m > 1$.

An interesting class of polynomial maps from a Hilbert space to $\mathbb{R}^m$ comes from sub-Riemannian geometry. They are the *endpoint maps* of Carnot groups (the model spaces of sub-Riemannian geometry), and they play a fundamental role in the study of non-holonomic geometries. One of the main open questions in sub-Riemannian geometry is the *Sard conjecture*: it claims that the endpoint maps have the Sard property, that is, that their sets of critical values have measure zero. This question is still wide open in Carnot groups and motivates a better investigation of the Sard problem in the context of polynomial maps from a Hilbert space to $\mathbb{R}^m$, $m > 1$. This is the main scope of this thesis.

In this thesis we provide Sard-type theorems for polynomial maps from a Hilbert space to $\mathbb{R}^m$, for general $m \geq 1$. We present a comprehensive framework that encompasses Kupka's counterexample, the maps contructed by Yomdin, the endpoint maps of Carnot groups, and other examples discussed in this thesis. As an application, we present new results on the Sard conjecture in Carnot groups. These results have been obtained in [62], using also as a key tool a result in semialgebraic geometry obtained in [61].

ii Overview of main contributions and chapters

In Chapter 2 we recall the classical versions of the Morse-Sard theorem both in finite and infinite dimension, we sketch Yomdin's approach in both cases, and we discuss the Morse-Sard problem in infinite dimension.

In the subsequent chapters of the thesis we present the research contributions. We give here a quick glance in order to describe the structure of the chapters.

Sard-type theorems in infinite dimension and for the endpoint maps of Carnot groups

In Chapter 3 we present Sard-type theorems for polynomial maps f from a Hilbert space H to $\mathbb{R}^m$, obtained in [62]. The main theorems are stated in terms of the Kolmogorv n-width, that we introduce in this context. We prove that on any infinite dimensional subspace $V \subset H$ satisfying a quantitative assumption on its Kolmogorov n-width, the restriction $f|_V$ as the Sard property, i.e the set of critical values has measure zero. We prove various quantitative versions of this result, allowing to study the Sard property also for the case $V = H$. We prove the sharpness of our n-width assumption constructing suitable counterexamples. A detailed presentation of these results is given in Section iii.1.

In Chapter 4 we apply our results to sub-Riemannian geometry: we prove the Sard conjecture for the restriction of the endpoint maps of Carnot groups to the set of piece–wise real–analytic controls with large enough radius of convergence. A detailed presentation of these results is given in Section iii.3.

Quantitative approximate definable choices in semialgebraic geometry

The Sard-type theorems in Chapter 3 and Chapter 4 rest on a key result in semialgebraic geometry, that we prove in Chapter 5, and that now we describe.

A *definable choice* is a semialgebraic selection of one point in every fiber of the projection of a semialgebraic set. Definable choices exist by semialgebraic triviality, but their complexity depends exponentially on the number of variables. By allowing the selection to be approximate (in the Hausdorff sense), we quantitatively improve this result in Chapter 5. More precisely, we construct an approximate selection with degree that is linear in the complexity of the original set, and independent on the number of variables. To prove these results, we develop a general quantitative theory for Hausdorff approximations in semialgebraic geometry. See Section iii.4 for a detailed presentation of the results.

iii Statement of the results

iii.1 Sard properties for polynomial maps in infinite dimension

In this section we introduce the content of Chapter 3, where we investigate the Sard property for polynomial maps defined on Hilbert spaces and with values in $\mathbb{R}^m$, for general $m \geq 1$.

Definition 1. Given a Hilbert space H and $d, m \in \mathbb{N}$, we define the class of maps

$$\mathscr{P}_d^m(H) := \left\{ f : H \to \mathbb{R}^m \,\middle|\, \dim(E) < \infty \implies f|_E \text{ is a polynomial map of degree } d \right\},$$

where by polynomial map we mean that each component of $f|_E$ is a polynomial of degree d. We require that each element of $\mathscr{P}_d^m(H)$ is of class C^1 (in the Fréchet sense), and that its differential $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is weakly continuous and locally Lipschitz.

The map constructed by Kupka belongs to this family, as well as the Endpoint maps for horizontal path spaces on Carnot groups (see Section iii.3).

We need to introduce a notion of *quantitative compactness* for subsets of Hilbert spaces. More precisely, given $K \subset H$ and $n \in \mathbb{N}$, we denote by $\Omega_n(K, H)$ its *Kolmogorov n–width*:

$$\Omega_n(K, H) := \inf_{\dim(E)=n} \sup_{u \in K} \inf_{v \in E} \|v - u\|, \tag{1}$$

where $\|\cdot\|$ denotes the norm of the Hilbert space H, and the infimum ranges over all vector subspaces $E \subset H$ of dimension n (see Definition 3.18). A set $K \subset H$ is compact if and only if it is bounded and its n–width goes to zero as n goes to infinity. For a compact set $K \subset H$ we also define the quantity:

$$\omega(K, H) := \limsup_{n \to \infty} \Omega_n(K, H)^{1/n} \in [0, 1]. \tag{2}$$

Roughly speaking, smaller $\omega(K, H)$ means that K is "more compact" in H, and better approximated by finite-dimensional subspaces.

Our results are stated in terms of *entropy dimension* of the set of *ν-critical values* of a map $f : H \to \mathbb{R}^m$, namely $f(\mathrm{Crit}_\nu(f))$, where $\mathrm{Crit}_\nu(f)$ is the set of points where Df has rank at most $\nu \leq m - 1$. The classical set of critical values corresponds to setting $\nu = m - 1$, and it is denoted by $f(\mathrm{Crit}(f))$, where $\mathrm{Crit}(f)$ is the set of critical points. We refer to Definition 3.16 of entropy dimension, noting that it is larger than the Hausdorff one so that in the forthcoming estimates one can replace the former with the latter for simplicity.

Our first theorem is a sufficient condition for the validity of the Sard property (see Theorem 3.23). We denote by μ the Lebesgue measure on $\mathbb{R}^m$.

Theorem A (Sard under n–width assumptions)**.** *Let $d, m \in \mathbb{N}$. There exists $\beta_0 = \beta_0(d, m) > 0$ such that the following holds. Let H be a Hilbert space, $f \in \mathscr{P}_d^m(H)$ and $K \subset H$ be a compact set such that*

$$\omega(K, H) = \limsup_{n \to \infty} \Omega_n(K, H)^{1/n} \leq q^{-1} \in (0, 1).$$

Then, for every $\nu = 1, \dots, m - 1$ we have

$$\dim_e \Big(f(\mathrm{Crit}_\nu(f) \cap K) \Big) \leq \nu + \frac{\ln \beta_0}{\ln q}.$$

In particular, if $q > \beta_0$, then the Sard property holds on K:

$$\mu\Big(f(\mathrm{Crit}(f) \cap K) \Big) = 0.$$

Remark 2. The constant $\beta_0(d,m)$ of Theorem A comes from semialgebraic geometry, and its origin is discussed in Section iii.2. For this reason, we sometimes call it the *semialgebraic constant*.

By building upon Kupka's counterexamples [58], we prove that the quantitative assumption $q > \beta_0$ in Theorem A cannot be dispensed of. The following result corresponds to Theorem 3.27. We denote by $B_X(r)$ the closed ball of radius r in a Banach space X.

Theorem B (Counterexamples to Sard)**.** *Let $d, m \in \mathbb{N}$, with $d \geq 3$, and set $q := (d-1)^{1/d}$. There exist a Hilbert space H, and $f \in \mathscr{P}_d^m(H)$ such that $K = \mathrm{Crit}(f) \cap B_H(r)$ is compact for all $r > 0$, with*

$$\omega(K,H) = \limsup_{n\to\infty} \Omega_n(K,H)^{1/n} \leq q^{-1},$$

and $f : H \to \mathbb{R}^m$ does not verify the Sard property, namely $\mu(f(\mathrm{Crit}(f) \cap K)) > 0$. Therefore, the semialgebraic constant $\beta_0(d,m)$ of Theorem A satisfies

$$\beta_0(d,m) \geq (d-1)^{1/d}, \qquad \forall\, m \in \mathbb{N},\ d \geq 3.$$

Remark 3. The case $d = 2$, which is left out from Theorem B, is special. This is related to the fact that the complexity of semialgebraic sets defined by quadratic equations has a different behaviour compared to the case $d \geq 3$. In this case we currently expect that *any* $f \in \mathscr{P}_2^m(H)$ satisfies the Sard property globally on H. This will be the subject of a future work.

As a consequence of Theorems A and B, we see that there is a quantitative threshold on the n–width of compact sets for the validity of the Sard property. Indeed, we can define $\omega_0(d,m)$ as the supremum of the set of $\omega \in [0,1]$ such that for every $f \in \mathscr{P}_d^m(H)$ and for every compact set $K \subset H$ with $\omega(K,H) = \omega$ it holds $\mu\,(f\,(\mathrm{Crit}(f) \cap K)) = 0$. We obtain the following statement (see Theorem 3.29).

Theorem C (Sard threshold on compacts)**.** *For all $d, m \in \mathbb{N}$, there exists $\omega_0(d,m) \in (0,1]$ such that*

(i) for every $f \in \mathscr{P}_d^m(H)$ and for every compact set $K \subset H$ with $\omega(K,H) < \omega_0(d,m)$,

$$\mu\Big(f\,(\mathrm{crit}(f) \cap K)\Big) = 0;$$

(ii) for every $\omega > \omega_0(d,m)$, with $\omega \in (0,1]$, there exist $f \in \mathscr{P}_d^m(H)$ and a compact set $K \subset H$ with $\omega(K,H) = \omega$ and such that

$$\mu\Big(f\,(\mathrm{crit}(f) \cap K)\Big) > 0.$$

Remark 4. By construction, and Theorem A, $\omega_0 \geq \beta_0^{-1}$. Furthermore, if $d \geq 3$, Theorem B yields $\omega_0 < 1$, so that Item (ii) is non-vacuous in these cases.

We can strengthen the conclusions from Item (i) of Theorem C as follows (see Theorem 3.30).

Theorem D (Sard threshold on linear subspaces)**.** *Let $f \in \mathscr{P}_d^m(H)$ and $K \subset H$ be a compact set such that $\omega(K,H) < \omega_0(d,m)$. Consider the linear subspace*

$$V := \mathrm{span}(K).$$

Then the restriction $f|_V : V \to \mathbb{R}^m$ satisfies the Sard property:

$$\mu\Big(f(\mathrm{Crit}(f|_V))\Big) = 0.$$

In particular, $\mu(f(\mathrm{Crit}(f) \cap V)) = 0$.

Remark 5. Since $V \subset H$ is a (possibly non-closed) linear subspace, the restriction $f|_V : V \to \mathbb{R}^m$ is a C^1 map in the Fréchet sense from the normed vector space $(V, \|\cdot\|)$ to $\mathbb{R}^m$, with $D_u(f|_V) = (D_u f)|_V$ for all $u \in V$.

Example 6. Let $\eta \in (1, \infty)$ with $\eta > \omega_0(d,m)^{-1}$, where $\omega_0(d,m)$ is the number in Theorems C and D, for $d, m \in \mathbb{N}$. Let $H = \ell^2$ with the usual norm. Consider the subset $K \subset B_H(1)$ given by

$$K = \left\{ u \in \ell^2 \,\middle|\, \sum_{j=1}^{\infty} |u_j|^2 \eta^{2j} \leq 1 \right\}.$$

One can easily verify that K is compact. Furthermore, the linear space $V = \operatorname{span}(K)$ is dense in H. We can estimate the n–width of K as follows. For $n \in \mathbb{N}$ consider the n-dimensional subspaces $E_n = \{x \in \ell^2 \mid x_j = 0, \ \forall\, j \geq n+1\} \subset H$. From (1), we obtain

$$\begin{aligned}
\Omega_n(K,H)^2 &\leq \sup_{u \in K} \inf_{v \in E_n} \|u - v\|^2 \\
&= \sup_{u \in K} \sum_{j=n+1}^{\infty} |u_j|^2 \\
&\leq \eta^{-2(n+1)} \sup_{u \in K} \sum_{j=1}^{\infty} |u_j|^2 \eta^{2j} \leq \eta^{-2(n+1)}.
\end{aligned}$$

Hence $\omega(K,H) = \limsup_n \Omega_n(K,H)^{1/n} \leq \eta^{-1} < \omega_0(d,m)$. Hence, by Theorem D it holds

$$\mu\Big(f(\operatorname{Crit}(f|_V))\Big) = \mu\Big(f(\operatorname{Crit}(f) \cap V)\Big) = 0, \qquad \forall\, f \in \mathscr{P}^m_d(H).$$

Note that even if $V \subset H$ is dense, the unrestricted map $f : H \to \mathbb{R}^m$ may not have the Sard property (see e.g. the Kupka counterexamples in Section 3.2.3).

Theorems A, C and D are deduced from a more general result for maps that are "well-approximated" by polynomials. This can be regarded as our main result concerning the Sard property, and does not make use of the concept of n–width. We report here the statement (see Theorem 3.17).

Theorem E (Sard criterion for well-approximated maps)**.** *Let $d, m \in \mathbb{N}$. There exists a constant $\beta_0 = \beta_0(d,m) > 0$ such that the following holds. Let H be a Hilbert space, and let $f : H \to \mathbb{R}^m$ be a C^1 map such that its differential $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is weakly continuous. Let $K \subset H$ be a bounded set with this approximation property: there exist a sequence $E_n \subset H$ of linear subspaces,* $\dim(E_n) = n$*, and polynomial maps $f_n : E_n \to \mathbb{R}^m$ with uniformly bounded degree:*

$$\sup_{n \in \mathbb{N}} \deg f_n \leq d < \infty,$$

such that for some $q > 1$, $c \geq 0$, and all large enough n it holds

$$\sup_{x \in K} \Big(\|f(x) - f_n \circ \pi_{E_n}(x)\| + \|(D_x f)|_{E_n} - D_{\pi_{E_n}(x)} f_n\|_{\mathrm{op}} \Big) \leq c q^{-n}. \tag{3}$$

Then
$$\dim_e\Big(f(\operatorname{Crit}_\nu(f)\cap K)\Big)\le \nu+\frac{\ln\beta_0}{\ln q},\qquad \forall\,\nu=1,\dots,m-1.$$
In particular, if $q>\beta_0$, then f satisfies the Sard property on K:
$$\mu\Big(f(\operatorname{Crit}(f)\cap K)\Big)=0.$$

Remark 7. The constant $\beta_0(d,m)$ of Theorem E is the same appearing in Theorem A, and comes from semialgebraic geometry, see Section iii.2.

By means of the general Theorem E we single out a class of maps satisfying the Sard property on the whole domain of definition. This is the content of Theorem 3.34, that we report here. The case $m=1$ corresponds to [96, Thm. 10.12].

Theorem F (Global Sard for special maps)**.** *Let H be a separable Hilbert space. For all $k\in\mathbb{N}$, let $p_k: E_k\to\mathbb{R}^m$ be polynomial maps with $\sup_{k\in\mathbb{N}}\deg p_k\le d$ for some $d\in\mathbb{N}$, and such that*
$$\sup_{x\in B_{E_k}(1)}\|p_k(x)\|\le q^{-k},\qquad \forall\,k\in\mathbb{N},$$
for some $q>1$. Then the map $f:H\to\mathbb{R}^m$ defined by
$$f(x):=\sum_{k=1}^{\infty}p_k(x_1,\dots,x_k),\qquad \forall\,x\in H,$$
is well–defined, $f\in\mathscr{P}^m_d(H)$ (see Definition 1), and for all $\nu=1,\dots,m-1$ and $r>0$ it holds
$$\dim_e\Big(f\left(\operatorname{Crit}_\nu(f)\cap B_H(r)\right)\Big)\le\nu+\frac{\ln\beta_0}{\ln q},$$
where $\beta_0=\beta_0(d,m)$ is the same constant given by Theorem E. In particular, if $q>\beta_0$, then f satisfies the Sard property globally on H:
$$\mu\Big(f(\operatorname{Crit}(f))\Big)=0.$$

iii.2 The role of quantitative semialgebraic geometry

The proof of Theorem E needs some fine properties of semialgebraic sets. The strategy of the proof is conceptually similar to Yomdin's [95], and uses the theory of *variations*, introduced by Vitushkin [91, 92] and developed in [96].

Let us denote with $V_i(S)$ the i-th variation of a semialgebraic set, which is a sort of i-dimensional volume, (see Definition 3.6). Furthermore, recall that for a C^1 map $F:\mathbb{R}^n\to\mathbb{R}^m$, and $\Lambda=(\Lambda_1,\dots,\Lambda_m)\in\mathbb{R}^m_+$ the set of almost-critical values of F is defined by
$$C^\Lambda(F):=\Big\{x\in\mathbb{R}^n\;\Big|\;\sigma_i(D_xF)\le\Lambda_i,\quad\forall\,i=1,\dots,m\Big\},$$
where $\sigma_1(D_xF)\ge\cdots\ge\sigma_m(D_xF)$ are the singular values of D_xF, see Definition 3.8 (here we assume $n\ge m$). In [96, Cor. 7.4] a quantitative estimate on the variations of the almost-critical values of polynomial maps has been obtained. In that estimate the dimensional parameter n does not appear explicitly. We obtain Theorem G below, which makes this dependence explicit (see Theorem 3.15).

Theorem G (Quantitative variations estimates)**.** *Let $n \geq m$, and $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d. For $i = 0, \dots, m$, $\Lambda = (\Lambda_1, \dots, \Lambda_m) \in \mathbb{R}_+^m$ and $r > 0$, we have*

$$V_i(p(C^\Lambda(p) \cap B_{\mathbb{R}^n}(r))) \leq \mathrm{cst}(m, r) n^m \beta_0^n \Lambda_0 \cdots \Lambda_i,$$

where $\beta_0 = \beta_0(d, m)$ depends only on d and m, $\mathrm{cst}(m, r)$ *depends only on m, r, and we set $\Lambda_0 = 1$.*

A key technical step in the proof of Theorem G is a quantitative substitute for the *definable choice* theorem in semialgebraic geometry from [61], see Theorem 3.5. The latter is proved in Chapter 5, indeed it follows from Theorem M and Remark 10.

iii.3 Applications to the Endpoint maps of Carnot groups

We discuss now the implications of Theorem C and Theorem D in the context of sub-Riemannian geometry. These results are proved in Chapter 4.

Recall that an m-dimensional *Carnot group* of step $s \in \mathbb{N}$ is a connected and simply connected Lie group $(\mathbb{G}, \cdot)$ of dimension m, whose Lie algebra $\mathfrak{g}$ admits a stratification of step s, that is

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s,$$

where $\mathfrak{g}_i \neq \{0\}$, $\mathfrak{g}_{i+1} = [\mathfrak{g}_1, \mathfrak{g}_i]$ for all $i = 1, \dots, s-1$ and $[\mathfrak{g}_1, \mathfrak{g}_s] = \{0\}$. The group exponential map yields an identification $\mathbb{G} \simeq \mathbb{R}^m$, and the first stratum $\mathfrak{g}_1$ of the Lie algebra defines a smooth, totally non-holonomic distribution $\Delta \subseteq T\mathbb{R}^m$ of rank $k := \dim \mathfrak{g}_1$.

Fix a global trivialization of Δ:

$$\Delta = \mathrm{span}\{X_1, \dots, X_k\},$$

where each X_i is a left-invariant vector field. Let $I := [0, 1]$ be the unit interval and $H := L^2(I, \mathbb{R}^k)$. We say that a curve $\gamma : I \to \mathbb{G}$ is *horizontal* if it is absolutely continuous and there exist $u \in H$, called *control* such that for a.e. $t \in I$ it holds

$$\dot{\gamma}(t) = \sum_{i=1}^{k} u_i(t) X_i(\gamma(t)), \tag{4}$$

Furthermore, for any given $u \in H$ there exists a unique $\gamma_u : I \to \mathbb{G}$ satisfying (4) and such that $\gamma_u(0) = e$ (the identity $e \in \mathbb{G}$, identified with $0 \in \mathbb{R}^m$). Note that the class of horizontal curves does not depend on the choice of the global trivialization of Δ.

The *Endpoint map* is the map that sends a control u to the the corresponding final point $\gamma_u(1)$:

$$\mathcal{E} : L^2(I, \mathbb{R}^k) \to \mathbb{G}.$$

The *Sard conjecture* is the conjecture that Endpoint map has the Sard property:

Sard conjecture: the set $\mathcal{E}(\mathrm{Crit}(\mathcal{E}))$ has zero measure.

The conjecture was introduced by Zhitomirskii and Montgomery, see [76, Sec. 10.2] for general sub-Riemannian structures, where the Endpoint map can be defined in a similar way, on a suitable domain of horizontal paths, which has the structure of a Hilbert manifold. For the specific case of Carnot groups, it is known to be true for step ≤ 2, see [7, 60]. Furthermore, the conjecture has been verified in [29] for filiform Carnot groups, in [30] for Carnot groups of rank 2 and step 4, and rank 3 and step 3, and for a handful of other specific examples described in [60].

Assume that Δ is also endowed with a left-invariant norm. This choice induces a length structure on the space of horizontal curves, and correspondingly, a distance, which gives $\mathbb{G}$ the structure of *sub-Finsler Carnot group* (or *sub-Riemannian Carnot group* if the norm is induced by a scalar product). In this case, within the set critical points of $\mathcal{E}$ (also called singular horizontal paths), the energy-minimizing ones play a significant role. At the corresponding critical values, the distance loses regularity: it is never smooth and it can even lose local semiconcavity [11, Sec. 4.2]. For this reason, the understanding of the "abundance" of such critical values is crucial. Denoting with $H_{\min} \subseteq L^2(I, \mathbb{R}^k)$ the set of minimizing horizontal paths starting from the identity, the *minimizing Sard conjecture* can be formulated as follows:

Minimizing Sard conjecture: the set $\mathcal{E}(\mathrm{Crit}(\mathcal{E}) \cap H_{\min})$ has zero measure.

The most general result to date is that the above set is a closed nowhere dense set [3, 83], which of course does not imply the minimizing Sard conjecture. The above conjecture is one of the main open problems in sub-Riemannian geometry [4, Prob. 3], [83, Conj. 1]. In general, the problem is settled in the following cases:

- Thanks to the work of Agrachev-Sarychev [9] and Agrachev-Lee [8], in absence of so-called Goh singular minimizing paths, all sub-Riemannian minimizing paths (resp. sub-Finsler, under suitable smoothness assumptions on the norm) are solution of a finite-dimensional Hamiltonian flow so that the corresponding set of critical values has zero measure by the finite-dimensional Sard theorem. Recently, Rifford proved that, more in general, the minimizing Sard conjecture holds for structures where all non-trivial Goh paths have Goh-rank ≤ 1 almost everywhere, see [82]. For Carnot groups, these results can be applied for example when the distribution is *pre-medium fat*, see [82, (1.2)]. In passing, we remark that for generic (so, typically not Carnot) sub-Riemannian structures with distribution of rank ≥ 3, there are no non-trivial singular Goh minimizing paths, see [34, Cor. 2.5] and [6, Thm. 8]. As a consequence the minimizing Sard conjecture holds true. Thanks to [82, Cor. 1.4], the latter result extends the case of rank ≥ 2.
- The minimizing Sard conjecture is true also for Carnot groups of step ≤ 3, see [60]. In this case, the problem is reduced (in a non-trivial way) to a finite-dimensional one. First, by noting that singular minimizing curves are non-singular in some proper subgroup, and then exploiting the fact that Carnot subgroups are parametrized by a finite-dimension manifold.

Finally, we refer to [97, 26, 23, 25, 24, 80, 83] for further details and discussions of various forms of the Sard conjecture on general sub-Riemannian structures.

The following result (Theorem 4.3) connects this framework to the previous sections.

Theorem H (Polynomial properties of the Endpoint map)**.** *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Then the Endpoint map $\mathcal{E} \in \mathscr{P}^m_s(H)$, for $H = L^2(I, \mathbb{R}^k)$*

As a consequence, we can apply Theorem D to the study of Endpoint maps of Carnot groups. We record here a first immediate corollary (see Theorem 4.4).

Theorem I (Sard criterion for Endpoint maps)**.** *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Let $K \subset L^2(I, \mathbb{R}^k)$ be a compact subset with*

$$\omega(K, L^2(I, \mathbb{R}^k)) < \omega_0(m, s).$$

Then, letting $V := \mathrm{span}(K)$, the restriction $\mathcal{E}|_V : V \to \mathbb{G}$ has the Sard property, namely

$$\mu\Big(\mathcal{E}(\mathrm{Crit}(\mathcal{E}|_V))\Big) = 0.$$

In particular $\mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}) \cap V)) = 0$.

We can apply Theorem I to find large sets on which the Sard property holds for the Endpoint map of Carnot groups. This is a more functional-analytic approach to the Sard conjecture that, unlike previous ones, does not resort to reduction to finite-dimensional cases.

An interesting case in which the assumptions of Theorem I can be effectively checked is given by sets of controls that are "sufficiently regular". We introduce the notation to state our result, see Section 4.3 for additional details. Given $r > 0$, and a closed interval $I \subset \mathbb{R}$, consider the set:

$$C^\omega(I, \mathbb{R}^k; r) := \left\{ u : I \to \mathbb{R}^k \,\middle|\, u \text{ is real-analytic with radius of convergence} > r \right\},$$

endowed with the sup norm over the r-neighbourhood of I in $\mathbb{C}$. More in general, let $C^\omega(I, \mathbb{R}^k; r, \ell)$ be the set of piecewise analytic controls:

$$C^\omega(I, \mathbb{R}^k; r, \ell) = \left\{ u : I \to \mathbb{R}^k \,\middle|\, u|_{I_j} \in C^\omega(I_j, \mathbb{R}^k; r), \text{ for all } j = 1, \dots, \ell \right\},$$

where $I_j = \inf I + \left[\frac{(j-1)|I|}{\ell}, \frac{j|I|}{\ell}\right]$, for all $j = 1, \dots, \ell$, which we equip with a suitable norm (see Section 4.3). In Theorem 4.5 we prove that the unit ball K of $C^\omega(I, \mathbb{R}^k; r, \ell)$, with $I = [0, 1]$, is compact in H and its n–width satisfies:

$$\Omega_n(K, H) \leq \frac{(k\ell)^{1/2}}{\ln r} \left(\frac{1}{r}\right)^{\lfloor \frac{n}{k\ell} \rfloor}, \qquad \forall\, n \in \mathbb{N}, \quad r > 1.$$

By applying Theorem I we can therefore deduce the following result (see Theorem 4.6).

Theorem J (Sard property on piecewise real-analytic controls)**.** *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Given $\ell \in \mathbb{N}$, there exists $r = r(m, s, k, \ell) > 0$ such that, letting $V = C^\omega(I, \mathbb{R}^k; r, \ell)$, with $I = [0, 1]$, it holds*

$$\mathcal{E}(V) = \mathbb{G} \qquad \text{and} \qquad \mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}|_V)) = \mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}) \cap V)) = 0.$$

Namely, the Sard property holds on the space of piecewise real-analytic controls with radius of convergence > r and with ℓ pieces.

Remark 8 (Regularity and Sard). Sussmann established that for real-analytic sub-Riemannian structures (such as Carnot groups), minimizing horizontal paths are real-analytic on an open and dense set of their interval of definition [89]. In light of that result, Theorem J hints at a unexpected link between the regularity problem of sub-Riemannian geodesics and the minimizing Sard conjecture.

Remark 9 (Surjectivity of the Endpoint map). Of course a subset of controls $V \subset L^2(I, \mathbb{R}^k)$ must satisfy $\mathcal{E}(V) = \mathbb{G}$ to be relevant for the theory. By the Rashevskii-Chow theorem, $\mathcal{E} : L^2(I, \mathbb{R}^k) \to \mathbb{R}^m$ is surjective. We observe that $\mathcal{E}$ is surjective also on $C^\omega(I, \mathbb{R}^k; r, \ell)$ if $\ell \geq 4m$, since from the proof of the Rashevskii-Chow theorem any pair of points can be joined by a concatenation of $4m$ horizontal paths with constant controls, see [5, Sec. 3.2]. Actually, the Endpoint map is surjective when restricted on any set that is dense in $L^2(I, \mathbb{R}^k)$, see e.g. [22, Sec. 2.5]. It follows that $\mathcal{E}$ is surjective on $C^\omega(I, \mathbb{R}^k; r)$, since the latter contains all (restrictions on I of) polynomial controls.

Finally, we prove in Theorem 4.10 that the Endpoint map on Carnot groups is surjective already when restricted to polynomial controls of large enough degree (depending on $\mathbb{G}$). This is a particular case of the following more general result (see Theorem 4.11).

Theorem K (Quantitative surjectivity)**.** *Let $\mathbb{G}$ be a Carnot group with topological dimension m and rank k. Let $S \subset L^2(I, \mathbb{R}^k)$ be a dense set. Then there exist $u_0, u_1, \dots, u_m \in S$ such that*

$$\mathcal{E}(\mathrm{span}\{u_0, u_1, \dots, u_m\}) = \mathbb{G}.$$

In other words, one can find an $(m+1)$-dimensional vector subspace $E \subset L^2(I, \mathbb{R}^k)$ of controls (depending on $\mathbb{G}$) such that the restriction of the Endpoint map to E is surjective. Furthermore, E can be assumed to have generators in any prescribed dense set $S \subset L^2(I, \mathbb{R}^k)$. Theorem K can be seen as a more refined version (for Carnot groups) of the well–known fact that the Endpoint map is surjective when restricted on any set of controls that is dense in $L^2(I, \mathbb{R}^k)$, see e.g. [22, Sec. 2.5].

Closing thoughts on the Sard conjecture

We do not venture in a guess in favor of Sard conjecture on the whole $H = L^2(I, \mathbb{R}^k)$, instead we propose a direction of future investigation that can be approached with our techniques. Recall that $H_{\min} \subset H$ is the set of energy-minimizing controls. It is known that $H_{\min}$ is boundedly compact in H i.e., the intersection $H_{\min} \cap B_H(r)$ is compact for all $r > 0$, see [5, Thm. 8.66], [2]. In fact, from the very recent result by Lokutsievskiy and Zelikin [63, Cor. 1 ($\mathfrak{G}$)], one can deduce the following estimate for the n–width on Carnot groups

$$\Omega_{2n+1}(H_{\min} \cap B_H(1), H) \leq C n^{-\frac{1}{2s}}, \qquad \forall\, n \in \mathbb{N}, \tag{5}$$

for some $C > 0$ depending on $\mathbb{G}$. Estimate (5) can be understood as a quantitative compactness property for $H_{\min}$. Unfortunately, the polynomial decay (w.r.t. n) of (5) is too weak to apply Theorem A, since the latter requires an exponential one.

If one could prove that, for some $q^{-1} < \omega_0(s, m)$, it holds

$$\Omega_n(H_{\min} \cap B_H(1), H) \leq q^{-n}, \qquad \text{as } n \to \infty, \tag{6}$$

then Theorem C would settle the minimizing Sard conjecture on Carnot groups.

We also note that it would be sufficient to prove (6) for the set of the so-called *strictly abnormal* energy-minimizing controls $H_{\min}^{\mathrm{str.abn}} \subsetneq H_{\min}$. Unfortunately, we were not able to produce a direct estimate of the n–width of (bounded subsets of) $H_{\min}$ or $H_{\min}^{\mathrm{str.abn}}$, so that we have no further evidence to support this idea towards a proof of the minimizing Sard conjecture.

In the opposite direction, and to conclude, our results can be used to identify spaces where the (general) Sard conjecture can be violated. If, for a given Carnot group $\mathbb{G}$, there is a linear subset of controls $V \subset H$ such that $\mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}) \cap V)) > 0$, then the following necessary condition must hold:

$$\limsup_{n\to\infty} \Omega_n(V \cap B_H(r), H)^{1/n} \geq \omega_0(s, m).$$

iii.4 Definable choices in semialgebraic geometry

In this section we introduce the content of Chapter 5, where we prove a quantitative approximate version of the *definable choice* theorem, [96, Thm. 4.10]. Before stating our main results, let us explain the name and the context of results of this type.

Let $S \subset \mathbb{R}^n$ be a compact semialgebraic set, presented as

$$S = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \left\{ x \in \mathbb{R}^n \mid \mathrm{sign}(p_{ij}(x)) = \sigma_{ij} \right\}, \tag{7}$$

where $\sigma_{ij} \in \{0, +1, -1\}$, the p_{ij} are polynomials of degree at most d. We condense this information into the *diagram* $D(S) := (n, c, d)$ of the representation (7), where $c = a \max b_i$.

We denote by

$$\pi : \mathbb{R}^n \to \mathbb{R}^\ell$$

the linear projection onto the last ℓ coordinates. It follows from Tarski–Seidenberg that the set $\pi(S)$ is semialgebraic and a natural problem is to choose, for every element $y \in \pi(S)$, an element $x(y) \in \pi^{-1}(y) \cap S$ so that the resulting set $A := \{x(y)\}_{y\in\pi(S)}$ is semialgebraic and of dimension at most ℓ. The set A is called a *definable choice* over $\pi(S)$.

The fact that this can be done follows from a result in semialgebraic geometry called *semialgebraic triviality* (see Corollary 5.8). However, as a result of this process there is no good control on the geometry of A in terms of the data defining S (see Remark 5.9).

For many geometric applications one does not really need that A is a choice over $\pi(S)$, but it is enough that it is close to it. More precisely, given $\epsilon > 0$, denote by $\mathcal{U}_\epsilon(S)$ the Euclidean ϵ–neighbourhood of S. Then one can relax the requirements for the definable choice and ask, given $\epsilon > 0$, for a set $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$, with $\pi(A_\epsilon)$ "close" to $\pi(S)$ (in the Hausdorff metric, denoted by dist_H), and possibly with a control on the diagram of A_ϵ.

In this direction we prove two related results. The first one deals specifically with the problem that we have just discussed (see Theorem 5.42).

Theorem L (Quantitative approximate definable choice, first version)**.** *For every $c \in \mathbb{N}$ there exist $\kappa \in \mathbb{N}$ such that the following holds. Let $n, \ell, d \in \ell$, with $1 \leq \ell \leq n$. Let $\pi : \mathbb{R}^n \to \mathbb{R}^\ell$ be the projection onto the last ℓ coordinates and let $S \subset \mathbb{R}^n$ be a bounded closed semialgebraic set with*

$$D(S) = (n, c, d).$$

Then, for every $\epsilon > 0$ there exists a closed semialgebraic set $A_\epsilon \subset \mathbb{R}^n$ such that:

(i) $\dim(A_\epsilon) \leq \ell$*;*

(ii) $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$*;*

(iii) $\mathrm{dist}_H(\pi(A_\epsilon), \pi(S)) \leq \epsilon$*;*

(iv) $D(A_\epsilon) = (n, \kappa, \kappa d)$*.*

The set A_ϵ from the statement is therefore an approximate (in the Hausdorff metric) definable choice over $\pi(S)$, with a quantitative control on its diagram.

In view of applications, it is also useful to have the following alternative version of the previous result (Theorem 5.43), which essentially follows from Theorem L applied to the case of the graph of a semialgebraic map.

Theorem M (Quantitative approximate definable choice, second version)**.** *For every* $c, d, \ell \in \mathbb{N}$ *there exists* $\beta > 1$ *satisfying the following statement. Let* $n \in \mathbb{N}$ *and let* $K \subset \mathbb{R}^n$ *be a closed semialgebraic set contained in the ball* $B_{\mathbb{R}^n}(\rho)$ *and let* $F : \mathbb{R}^n \to \mathbb{R}^\ell$ *be a locally Lipschitz semialgebraic map such that*

$$D(\mathrm{graph}(F|_K)) = (n + \ell, c, d).$$

Then for every $\epsilon \in (0, \rho)$ *there exists a closed semialgebraic set* $C_\epsilon \subset \mathbb{R}^n$ *such that:*

(i) $\dim(C_\epsilon) \leq \ell$*;*

(ii) $C_\epsilon \subseteq \mathcal{U}_\epsilon(K)$*;*

(iii) $\mathrm{dist}_H(F(C_\epsilon), F(K)) \leq L(F, \rho) \cdot \epsilon$*, where* $L(F, \rho) := 2 + \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho))$*;*

(iv) for every $e = 1, \ldots, n$ *and every affine space* $\mathbb{R}^e \simeq E \subseteq \mathbb{R}^n$*, the number of connected components of* $E \cap C_\epsilon$ *is bounded by*

$$b_0(E \cap C_\epsilon) \leq \beta^e. \qquad (8)$$

Remark 10 (The case of polynomial maps). Theorem 5.43 can be applied to any polynomial map $F : \mathbb{R}^n \to \mathbb{R}^\ell$ with components of degree bounded by d, assuming that the diagram of the original set satisfies $D(K) = (n, c, d)$. In fact, in this case

$$\mathrm{graph}(F|_K) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^\ell \mid x \in K,\ y = F(x)\}$$

is a bounded and closed semialgebraic set with $D(\mathrm{graph}(F|_K)) = (n + \ell, c + 1, d)$.

Theorem M corresponds to [96, Thm. 4.10 and Ex. 4.11], where it is proved the existence of a semialgebraic set C_ϵ as in Theorem M, except for (8), which for them has the shape

$$b_0(E \cap C_\epsilon) \leq f(n, \ell, c, d, e),$$

for some (non–explicit) function $f : \mathbb{N}^5 \to \mathbb{N}$. The most notable conclusion from Theorem M is therefore the explicit dependence of the bound on the dimension of the affine space of Item (iv), which says that we can take $f(n, \ell, c, d, e) = \beta(c, d, \ell)^e$.

Similarly, the difficult part from Theorem L is proving that the diagram of A_ϵ has the explicit shape $D(A_\epsilon) = (m, \kappa, \kappa d)$ with κ depending only on the combinatorial data c of the diagram of the original set (and not on the number of variables, for instance).

Obtaining this explicit dependence is non–trivial. Compared to [96, Thm. 4.11], a conceptual novelty is the use of ideas from [12], which in turn involves the study of approximation of semialgebraic sets in the Hausdorff metric. A large part of the chapter is devoted to expanding and developing these ideas, see Section 5.1.3.

Hausdorff approximations

Section 5.1.3 contains new ideas that have their own interest and that are based on the observation that the Hausdorff distance between semialgebraic sets in $\mathbb{R}^n$ can be studied within the framework of semialgebraic geometry. In particular, this notion can be defined also on a real closed extension R of the real numbers (see Section 5.1.1), leading to a notion of "Hausdorff distance" between semialgebraic sets in R^n. The main result from Section 5.1.3 is then a technique to produce Hausdorff approximations of semialgebraic sets using infinitesimals, i.e. working in the real closed field of algebraic Puiseux series (see Definition 5.2). We allow multiple infinitesimals, which makes the technique handful and practical, but which requires nontrivial extensions of the ideas from [12]. What is important for us is that this technique allows to keep control on the combinatorial part and the degrees of the approximating set: for example, in Proposition 5.36, we show how to approximate a closed semialgebraic set with a *closed basic* semialgebraic set (see Definition 5.4) whose combinatorial data and degree are controlled in terms of the diagram of the original set (see the beginning of Section 5.1.4 to appreciate the subtlety of this statement).

Chapter 1

Table of notations

$\mathbb{R}^m_+$	m-tuples of positive real numbers
$B_X(x,r)$	closed ball of radius $r > 0$ centered at x of a Banach space X
$B_X(r)$	closed ball of radius $r > 0$ centered at 0 of a Banach space X
B_X	closed unit ball centered at 0 of a Banach space X
$\mathcal{L}(X,Y)$	linear continuous maps between banach spaces X, Y
$\|\cdot\|_{\mathrm{op}}$	operator norm on $\mathcal{L}(X,Y)$, see (3.19)
$\|\cdot\|_X$	norm for a Banach space X (the X is omitted when there is no confusion)
$\mathcal{B}(X,Y)$	continuous bounded maps between Banach spaces X, Y
$C^\omega(I,\mathbb{R}^k;r,\ell)$	piece-wise real-analytic functions, Section 4.3
μ	Lebesgue measure
$\Omega_n(K,H)$	Kolmogorov n–width of $K \subset H$, Definition 3.18
$\omega(K,H)$	asymptotic Kolmogorov width of $K \subset H$, Eq. (2)
$\mathcal{P}^m_d(H)$	set of well-approximable maps, Definition 1
$\beta_0(d,m)$	semialgebraic constant, Remark 2
$\omega(d,m)$	threshold constant, Theorem C
$\mathcal{H}^i(A)$	i-dimensional Hausdorff measure of a rectifiable Borel set A
$V_i(A)$	Vitushkin variation of a bounded semialgebraic set A, Definition 3.6
$b_0(S)$	number of connected components of a set S
$\sigma_k(L)$	k-th singular value of a linear map $L : H \to \mathbb{R}^m$ on a Hilbert space, ordered in decreasing order $\sigma_1(M) \geq \cdots \geq \sigma_m$, Definition 3.8
$\mathrm{Crit}(f)$	critical points of a map, Eq. (3.18)
$\mathrm{Crit}_\nu(f)$	ν-critical points of a map, Eq. (3.18)
$C^\Lambda(f)$	Λ-critical (or almost-critical) points of a map, Definition 3.11
$M(\epsilon,S)$	ϵ-entropy of a set $S \subset \mathbb{R}^m$, Definition 3.16
$\dim_e(S)$	entropy dimension of a set $S \subset \mathbb{R}^m$, Definition 3.16
$\mathcal{U}_\varepsilon(S)$	ϵ-neighbourhood of a subset $S \subset \mathbb{R}^n$
dist_H	Hausdorff distance
β_{tm}	Thom-Milnor bound constant, Theorem 3.4
β_{dc}	definable choice constant, Theorem 3.5

β_{yc}	Yomdin-Comte constant, Theorem 3.10
$D(S)$	diagram of a semialgebraic set S, Definition 3.2
$\mathbb{G}$	Carnot group (of step s, rank k, dimension m), Section 4.1
$\mathcal{E}$	Endpoint map, Definition 4.2

Chapter 2

Overview on the Morse-Sard problem in infinite dimension

The aim of this chapter is to present the state of the art on the Morse-Sard problem in infinite dimension, focusing on Yomdin's contributions. We start presenting Yomdin's quantitative version of the Morse-Sard theorem in finite dimension, in Section 2.1. Then, we discuss in Section 2.2 the infinite dimensional case, showing how the quantitative Yomdin's approach in finite dimension is suitable to obtain Sard theorems in infinite dimension.

We start recalling the classical Morse-Sard theorem. It was proved in [77, 84], and it is nowadays a fundamental tool in geometry and analysis.

Theorem 2.1 (Morse-Sard)**.** *Let $f : N \to M$ be a C^k map between smooth, finite-dimensional manifolds. If $k \geq \max\{\dim N - \dim M + 1, 1\}$, then the set $f(\mathrm{Crit}(f))$ has measure zero in M.*

A finer version of this theorem was proved in [40, Theorem 3.4.3], obtaining a bound on the Hausdorff dimension of sets of critical values (see also [85, 86]). To state the result, given U an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ a C^1 map, for any $\nu = 0, \dots, n-1$ we set

$$\mathrm{Crit}_\nu(f) = \{x \in U \mid \mathrm{rank}(D_x f) \leq \nu\}.$$

Theorem 2.2. *Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ be a C^k map. Then, for every $\nu = 0, \dots, n-1$ we have*

$$\dim_{\mathcal{H}}(f(\mathrm{Crit}_\nu(f))) \leq \nu + \frac{n-\nu}{k}.$$

The sharpness of such estimate is discussed in [40, Section 3.4.4].

2.1 Yomdin's quantitative version of the Morse-Sard theorem in finite dimension

In this section we present the quantitative version of the classical Morse-Sard theorem in finite dimension, proved by Yomdin. We first sketch the proof of the real-valued case in order to illustrate the philosophy of Yomdin's approach in the simplest case. Then we sketch the proof of the vector-valued case, which follows the same approach of the real-valued case, but it requires additional tools. To state these results we need the following definition.

Given a relatively compact $S \subset \mathbb{R}^m$, for any $\epsilon > 0$ we denote with $M(\epsilon, S)$ the *ϵ-entropy* of S, which is the minimum number of closed balls of radius ϵ that we need to cover S.

Definition 2.3 (Entropy dimension)**.** The *entropy dimension* of S is defined as

$$\dim_e(S) = \limsup_{\epsilon \to 0^+} \frac{\ln M(\epsilon, S)}{\ln(\frac{1}{\epsilon})}.$$

We refer to [96, Ch. 2] for a more detailed treatment of the properties of the entropy dimension. For instance, it is proved that

$$\dim_{\mathcal{H}}(S) \leq \dim_e(S). \tag{2.1}$$

We point out that the inequality can be strict, since there are sequences of real numbers with positive entropy dimension. Indeed, by [96, Theorem 2.9], given $a > 0$ we have

$$\dim_e\left(\left\{\frac{1}{j^a}\right\}_{j\in\mathbb{N}}\right) = \frac{1}{a+1}. \tag{2.2}$$

2.1.1 One-dimensional codomain

In this section we present Yomdin's theorem for real-valued functions defined on a closed ball of $\mathbb{R}^n$. The main result of this section is Theorem 2.7.

Definition 2.4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1 map. Given $\lambda \in \mathbb{R}_+$, the set of *almost-critical points* of f is defined as

$$C^\lambda(f) := \left\{x \in \mathbb{R}^n \,\middle|\, \|\nabla_x f\| \leq \lambda\right\}.$$

We denote by $B_{\mathbb{R}_n}(x_0, \rho)$ the closed ball of $\mathbb{R}^n$ centered at x_0 and of radius $\rho > 0$; when x_0 is the origin we denote it as $B_{\mathbb{R}_n}(\rho)$.

The first key step of Yomdin's approach is to prove a finer version of the Morse-Sard theorem in the case of polynomials, providing quantitative estimates for their almost-critical points.

Theorem 2.5 (Quantitative Morse-Sard theorem for polynomials)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. Then, for any $\lambda \geq 0$ the set $p(C^\lambda(p) \cap B_{\mathbb{R}^n}(r))$ can be covered by $N(n, d)$ intervals of length λr, where $N(n, d)$ depends only on n and d.*

The proof of this result is based on semialgebraic geometry and is given in [96, Theorem 1.8]. We stress that the fact that $N(d, n)$ depends only on n and d comes from semialgebraic geometry, and it is important for the rest of the proof.

The next step in Yomdin's approach is to estimate the number of intervals that we need to cover the critical values of a C^k-function $g : B_{\mathbb{R}_n}(x_0, \rho) \to \mathbb{R}$, which corresponds to [96, Theorem 1.9]. In order to state it here, we introduce the following notation. For $x \in B_{\mathbb{R}^n}(x_0, \rho)$ we denote by $D^k_x g$ the differential of order k of g at x. For any $r > 0$, we set

$$R_k(g, x_0, r) := r^k \frac{1}{k!} \max_{x \in B_{\mathbb{R}^n}(x_0, r)} \left\|D^k_x g\right\|.$$

When x_0 is the origin we simply write $R_k(g, r)$. This quantity bounds the remainders in the Taylor approximations of g over a ball of radius r.

Proposition 2.6. *Let $g : B_{\mathbb{R}_n}(x_0, \rho) \to \mathbb{R}$ be C^k. Then the set $g(\mathrm{Crit}(g))$ can be covered with $N_2(n, k)$ intervals of length $R_k(g, \rho, x_0)$, with $N_2(n, k) = 3 \cdot k \cdot N(n, k-1)$.*

Proof. The idea of the proof is to apply Theorem 2.5 to the Taylor polynomial of g, observing that the property is stable under C^1 perturbations. Let P be the Taylor polynomial of degree $k-1$ centered at x_0. We have the following estimates

$$\max_{x \in B_{\mathbb{R}_n}(x_0, \rho)} |g(x) - P(x)| \leq R_k(g, x_0, \rho), \tag{2.3}$$

$$\max_{x \in B_{\mathbb{R}_n}(x_0, \rho)} \|\nabla_x g - \nabla_x P\| \leq R_k(g, x_0, \rho)\frac{k}{\rho}. \tag{2.4}$$

We set $\lambda := R_k(g, x_0, \rho)\frac{k}{\rho}$. By (2.4) we get

$$\mathrm{Crit}(g) \subset C^\lambda(P) \cap B_{\mathbb{R}_n}(x_0, \rho).$$

By (2.3) the set $g(C^\lambda(P) \cap B_{\mathbb{R}_n}(x_0, \rho))$ is contained in a $R_k(g, x_0, \rho)$-neighborhood of

$$P(C^\lambda(P) \cap B_{\mathbb{R}_n}(x_0, \rho)).$$

Hence, $g(\mathrm{Crit}(g))$ is contained in the $R_k(g, x_0, \rho)$-neighborhood of

$$P(C^\lambda(P) \cap B_{\mathbb{R}_n}(x_0, \rho)).$$

The rest of the proof consists in estimating the number of intervals of length $R_k(g, x_0, \rho)$ covering the latter. By Theorem 2.5 the set $P(C^\lambda(P) \cap B_{\mathbb{R}_n}(x_0, \rho))$ is covered by $N(n, k-1)$ intervals of length

$$\lambda\rho = kR_k(g, x_0, \rho).$$

Hence it is covered by $kN(n, k-1)$ intervals of length $R_k(g, x_0, \rho)$. Hence, its $R_k(g, x_0, \rho)$-neighborhood is covered by $3kN(n, k-1)$ intervals of length $R_k(g, \rho)$. Setting $N_2(n, k) := 3kN(n, k-1)$ we conclude the proof. □

We can now state and prove Yomdin's version of the Morse-Sard theorem for real-valued functions [96, Corollary 1.11].

Theorem 2.7 (Yomdin)**.** *Let $f : B_{\mathbb{R}^n}(r) \to \mathbb{R}$ be a C^k function, then*

$$\dim_e(f(\mathrm{Crit}(f))) \leq \frac{n}{k}.$$

In particular, if $k > n$, then $f(\mathrm{Crit}(f))$ has measure zero.

We recall that by (2.1) this theorem implies the estimate on the Hausdorff dimension given in Theorem 2.2 on balls.

Proof. The proof consists in estimating $M(\epsilon, f(\mathrm{Crit}(f)))$ for any $\epsilon > 0$. For fixed $\epsilon > 0$ we apply the previous proposition on small enough balls. The radii are chosen in such a way the Taylor remainder is less than ϵ, hence we can estimate the number of intervals of radius ϵ that cover the critical values on such small balls. Then, we globalize the result to the whole domain. We choose $\rho_\epsilon > 0$ such that on every ball $B_{\mathbb{R}^n}(y, \rho_\epsilon) \subset B_{\mathbb{R}^n}(r)$ we have $R_k(g, y, \rho_\epsilon) \leq 2\epsilon$. To find a suitable expression for ρ_ϵ, we observe that on every ball $B_{\mathbb{R}^n}(y, \rho) \subset B_{\mathbb{R}^n}(r)$ we have

$$R_k(g, y, \rho) = \frac{1}{k!} \max_{x \in B_{\mathbb{R}^n}(y,\rho)} \left\| D^k_x g \right\| \rho^k \leq \frac{1}{k!} \max_{x \in B_{\mathbb{R}^n}(r)} \left\| D^k_x g \right\| r^k \frac{\rho^k}{r^k} = R_k(g, r) \left(\frac{\rho}{r} \right)^k.$$

Hence, we take ρ_ϵ such that $R_k(g, r) \left(\frac{\rho_\epsilon}{r} \right)^k = 2\epsilon$, more precisely we set

$$\rho_\epsilon := r \left(\frac{2\epsilon}{R_k(g, r)} \right)^{\frac{1}{k}}.$$

By Proposition 2.6 for every such ball $B_{\mathbb{R}^n}(y, \rho_\epsilon) \subset B_{\mathbb{R}^n}(r)$, we cover $f(\mathrm{Crit}(f) \cap B_{\mathbb{R}^n}(y, \rho_\epsilon))$ with at most $N_2(n, k)$ intervals of length $R_k(g, y, \rho_\epsilon) \leq 2\epsilon$, in particular with at most $N_2(n, k)$ intervals of length 2ϵ. Now we estimate the number of intervals of length 2ϵ needed to cover the whole set of critical values of f on $B_{\mathbb{R}^n}(r)$. We cover $B_{\mathbb{R}^n}(r)$ with balls of radius ρ_ϵ. The number of such balls that we need is at least

$$c(n) \left(\frac{r}{\rho_\epsilon} \right)^n = c(n) \left(\frac{R_k(g, r)}{2\epsilon} \right)^{n/k}.$$

Therefore, we can cover $f(\mathrm{Crit}(f))$ with $N_2(n, k) \cdot c(n) \left(\frac{R_k(g,r)}{2\epsilon} \right)^{n/k}$ intervals of length 2ϵ (hence with the same number of balls of radius ϵ). Hence, setting $N_3(n, k) = N_2(n, k) \cdot c(n)$, we have proved that

$$M(\epsilon, f(\mathrm{Crit}(f))) \leq N_3(n, k) \left(\frac{R_k(g, r)}{2\epsilon} \right)^{n/k}.$$

We conclude the proof observing that

$$\dim_e(f(\mathrm{Crit}(f))) \leq \limsup_{\epsilon \to 0} \frac{\ln(N_3(n, k)) + \frac{n}{k} \ln \left(\frac{R_k(g,r)}{2\epsilon} \right)}{\ln \left(\frac{1}{\epsilon} \right)} = \frac{n}{k}.$$

□

Combining Theorem 2.7 and (2.2) we obtain the following result.

Corollary 2.8 (Yomdin)**.** *Fix $a > 0$. The sequence $\{1/j^a\}_{j \in \mathbb{N}}$ cannot be the set of critical values of a function $f : B_{\mathbb{R}^n}(r) \to \mathbb{R}$ which is C^k with $k > n(a + 1)$.*

We point out that from Theorem 2.2 it would not be possible to exclude that a sequence of real numbers is the set of critical values of a smooth real-valued function on a ball. This is a new aspect of Yomdin theorem compared to previous versions of the Morse-Sard theorem.

2.1.2 Higher-dimensional codomains

In this section we present Yomdin theorem for higher-dimensional codomains proved in [94], and we outline the main steps of the proof. The main result of this section is Theorem 2.13.

The first ingredient that we need is a notion of almost-critical points for C^1 maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, which is stable under C^1 perturbations. We need the following definition.

Given a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ we consider its *singular values*, see [96, Chapter 6] and [90]. Geometrically they correspond to the semiaxes of the ellipsoid $L(B_{\mathbb{R}^n})$, labeled in decreasing order: $\sigma_1(L) \geq \cdots \geq \sigma_{\min\{n,m\}}(L) \geq 0$.

Definition 2.9. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be C^1 and set $q := \min\{n, m\}$. Given $\Lambda = (\Lambda_1, \ldots, \Lambda_q) \in \mathbb{R}^q_+$, the set of *almost-critical points* of f is defined as

$$C^\Lambda(f) := \left\{ x \in \mathbb{R}^n \middle| \sigma_i(D_x f) \leq \Lambda_i,\ \forall\, i = 1, \ldots, q \right\}.$$

Yomdin's approach to the Morse-Sard theorem for higher-dimensional codomains follows the same philosophy of the real-valued case: the first key step is to prove a quantitative version of the Morse-Sard theorem for almost-critical points of polynomial maps (in the real-valued case it corresponds to Theorem 2.5). For higher-dimensional codomains, we estimate the ϵ-entropy of images of almost-critical points of polynomial maps in a different way, introducing in this setting new tools. Indeed, the ϵ-entropy of a bounded set $A \subset \mathbb{R}^n$ is controlled in terms of its *Vitushkin variations*, denoted by $V_i(A)$ for every $i = 0, \ldots, n$. They represent a sort of i-dimensional volume, defined through an integral-geometric approach (see Definition 3.6). From [96, Theorem 3.5] we have the following estimate for any $\epsilon > 0$,

$$M(\epsilon, A) \leq \mathrm{cst}(n) \sum_{i=0}^{n} \frac{1}{\epsilon^i} V_i(A), \tag{2.5}$$

where $\mathrm{cst}(n)$ is a constant depending only on n.

In order to estimate the ϵ-entropy, we provide quantitative estimates on the Vitushkin variations of images of almost-critical points of polynomial maps. This is the content of the next result, which corresponds to [96, Corollary 7.4]. The proof strongly relies on semialgebraic geometry, and a self contained presentation is given in [96].

Theorem 2.10 (Yomdin's variations estimates)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d, and let $q := \min\{n, m\}$. For $r > 0$, $x_0 \in \mathbb{R}^n$, $i = 0, \ldots, q$ and $\Lambda = (\Lambda_1, \ldots, \Lambda_q) \in \mathbb{R}^q_+$ we have*

$$V_i(p(C^\Lambda(p) \cap B_{\mathbb{R}^n}(x_0, r))) \leq \mathrm{cst}(m, d, n)\Lambda_0 \cdots \Lambda_i r^i,$$

where $\mathrm{cst}(m, d, n)$ *depends only on d, m, n, and we set $\Lambda_0 = 1$.*

From (2.5) and Theorem 2.10 we immediately obtain the first key result of Yomdin's approach to the Morse-Sard theorem for higher-dimensional codomains, it corresponds to [96, Theorem 7.5].

Theorem 2.11 (Quantitative Morse-Sard theorem for polynomial maps)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d, and let $q := \min\{n, m\}$. For $r > 0$, $x_0 \in \mathbb{R}^n$, $i = 0, \ldots, q$ and $\Lambda = (\Lambda_1, \ldots, \Lambda_q) \in \mathbb{R}^q_+$ we have*

$$M(\epsilon, p(C^\Lambda(p) \cap B_{\mathbb{R}^n}(x_0, r))) \leq \mathrm{cst}(m, d, n) \sum_{i=0}^{q} \frac{r^i}{\epsilon^i} \Lambda_0 \cdots \Lambda_i,$$

where $\mathrm{cst}(m, d, n)$ *depends only on* d, m, n*, and we set* $\Lambda_0 = 1$.

Form this theorem we obtain Yomdin's quantitative version of the Morse-Sard theorem for higher-dimensional codomains, proved in [94]. The new feature of this result, compared with the classical Morse-Sard theorems, is that it provides estimates for almost-critical points, and they are quantitative.

Theorem 2.12 (Yomdin's quantitative version of the Morse-Sard theorem)**.** *Let* $f : B_{\mathbb{R}^n}(r) \to \mathbb{R}^m$ *be a* C^k *map and set* $q := \min\{n, m\}$*. Then, for any* $\Lambda = (\Lambda_1, \dots, \Lambda_q) \in \mathbb{R}^q_+$*, and for any* $0 < \epsilon \le R_k(f, r)$ *we have*

$$M(\epsilon, f(C^\Lambda(f))) \le \mathrm{cst}(n, m, k) \sum_{i=0}^{q} \Lambda_0 \cdots \Lambda_i \frac{r^i}{\epsilon^i} \left(\frac{R_k(f, r)}{\epsilon}\right)^{\frac{n-i}{k}},$$

where $\mathrm{cst}(n, m, k)$ *is a constant depending only on n, m and k.*

The proof follows the same philosophy of the real-valued case: we apply Theorem 2.11 to the Taylor polynomials on small balls, then we globalize the estimate on the whole domain. For the detailed proof see [96, Theorem 9.2].

From Yomdin's quantitative Morse-Sard theorem, we obtain in particular the following finer version of the classical Morse-Sard theorem, see [96, Theorem 9.3].

Theorem 2.13 (Yomdin)**.** *Let* $f : B_{\mathbb{R}^n}(r) \to \mathbb{R}^m$ *a* C^k *map. Then, for any* $\nu = 0, \dots, m-1$ *we have*

$$\dim_e(f(\mathrm{Crit}_\nu(f))) \le \nu + \frac{n-\nu}{k}.$$

We recall that $\dim_{\mathcal{H}}(f(\mathrm{Crit}_\nu(f))) \le \dim_e(f(\mathrm{Crit}_\nu(f)))$, hence, in particular, Yomdin theorem provides the estimate on the Hausdorff dimension obtained in Theorem 2.2.

2.2 Morse-Sard theorems in infinite dimension

A version of the Morse-Sard theorem for maps between infinite-dimensional smooth Banach manifolds was proved by Smale in [88]. In order to state it here, we recall that a *Fredholm operator* between Banach spaces is a linear and continuous map whose kernel is finite dimensional, the range is closed, and the cokernel is finite dimensional. A *Fredholm map* between smooth Banach manifolds is a smooth map whose differential at any point is a Fredholm operator.

We can now state Smale theorem for C^∞ Fredholm maps, the statement for C^k functions can be found in [88, Theorem 1.3].

Theorem 2.14 (Smale)**.** *Let* N, M *be smooth Banach manifolds. Let* $f : N \to M$ *be a smooth Fredholm map. Then, the set of regular values of* f *is residual in* M*.*

We point out that when N is infinite dimensional and M is finite dimensional is not possible to apply Smale theorem, since there are no Fredholm maps in this case (the kernel of the differential is necessarily infinite dimensional). In fact, there are smooth surjective maps without regular points from any infinite-dimensional Banach space to $\mathbb{R}^2$, see [15]. Hence, for maps from an infinite dimensional space to a finite dimensional one smoothness is not sufficient for the validity of the Sard property. Even if the map is "polynomial" the Sard property may fail to hold, as proved by Kupka in [58] (see also [96, Section 10.2.3]), more precisely we have the following result.

Theorem 2.15 (Kupka)**.** *There exists a smooth function $f : \ell^2 \to \mathbb{R}$ such that the set of its critical values is the whole interval* $[0, 1]$. *Furthermore, f has the following polynomial property: for any linear subspace E with* $\dim E < \infty$ *the restriction $f|_E$ is a polynomial of degree* 3.

From these results the question whether there are Morse-Sard type theorems for maps from infinite dimensional spaces to finite dimensional spaces is left open: the classical result by Smale cannot be applied, furthermore, regularity assumptions as smoothness and even polynomiality do not guarantee that the set of critical values is negligible. Hence, the Morse-Sard problem in infinite dimension can be stated as the problem of finding sufficient conditions under which maps from infinite dimensional spaces to finite dimensional ones have the Morse-Sard property, that is, the set of critical values is negligible.

2.2.1 Yomdin theorem for polynomials from infinite-dimensional spaces to $\mathbb{R}$

The first contribution to the Morse-Sard problem in infinite dimension was given by Yomdin for real-valued maps. It was proved in [95], see also [96, Section 10.2]. We state it here, and we sketch the proof. The main result of this section is Theorem 2.18.

As discussed at the beginning of the previous section, for functions from infinite dimensional spaces to $\mathbb{R}$, the validity of the Sard property is not guaranteed under smoothness assumptions, nor under the condition of being a "polynomial". However, Yomdin found a sufficient condition which implies the Sard property in this setting (we state it in Theorem 2.18) and he provided an explicit class of maps from ℓ^2 to $\mathbb{R}$ with such property (see Corollary 2.21).

Yomdin's assumption is that the map is well approximated in C^1 norm by a sequence of polynomials depending on an increasing (and finite) number of coordinates.

The strategy of the proof is to provide a refined version of the quantitative Morse-Sard theorem for polynomials (Theorem 2.5), where the dependence of the estimate on the number of variables n is explicit, see Theorem 2.16. Then, we can pass to the limit as n goes to ∞ thanks to the good approximation hypothesis, obtaining the Sard property for the original function, see Theorem 2.18.

Now we state the first key step of the proof, which was proved in [95, Theorem 1.5], see also the discussion in [96, page 21].

Theorem 2.16 (Yomdin)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. Then, for any $\lambda \geq 0$ the set $f(C^\lambda(f) \cap B_{\mathbb{R}^n}(r))$ can be covered with $N(n, d) = (2d)^n$ intervals of length λr.*

As a consequence of Theorem 2.16, we obtain the following estimate on the ϵ-entropy of almost-critical points of finite dimensional polynomials, proved in [95, Proposition 4.1].

Proposition 2.17. *Let X be a Banach space. Given $p : \mathbb{R}^n \to \mathbb{R}$ a polynomial of degree d and given $L : X \to \mathbb{R}^n$ a linear, continuous and surjective map, setting $\tilde{p} : X \to \mathbb{R}$ as $\tilde{p} = p \circ L_n$, we have for any $\epsilon > 0$*

$$M(\epsilon, \tilde{p}(C^\epsilon(\tilde{p}) \cap B_X)) \leq n(2d)^n.$$

We can now state and prove the main result of this section. In the work of Yomdin this statement does not appear explicitly, however it follows from his construction in [95] (see also [96, Section 10.2]).

Given a Banach space X and a C^1 function $g : B_X \to \mathbb{R}$ bounded and with bounded differential $Dg : B_X \to X^*$, we consider its C^1 norm

$$\|g\|_{C^1_{B_X}} = \sup_{x \in B_X} |g(x)| + \sup_{x \in B_X} \|D_x g\|_{\mathrm{op}},$$

where $\|D_x g\|_{\mathrm{op}}$ denotes the operator norm of the differential of g at x.

Theorem 2.18 (Yomdin)**.** *Let X be a Banach space and $f : B_X \to \mathbb{R}$ be a C^1 map. Suppose that given $d \in \mathbb{N}$ there exist $q > 1$, $c > 0$, a sequence $L_n : X \to \mathbb{R}^n$ of linear, continuous and surjective maps, and a sequence of polynomials $p_n : \mathbb{R}^n \to \mathbb{R}$ of degree d such that for any $n \in \mathbb{N}$ we have*

$$\|f - p_n \circ L_n\|_{C^1_{B_X}} \leq cq^{-n}. \tag{2.6}$$

Then

$$\dim_e(f(\mathrm{Crit}(f))) \leq \frac{\ln(2d)}{\ln q}.$$

In particular, if $q > 2d$, then the set of critical values of f has measure zero.

Proof. We observe that if $x \in \mathrm{Crit}(f)$, then $\|D_x(p_n \circ L_n)\| \leq cq^{-n}$ for any $n \in \mathbb{N}$. Hence, setting $\lambda_n := cq^{-n}$ we have

$$\mathrm{Crit}(f) \subset C^{\lambda_n}(p_n \circ L_n) \cap B_X. \tag{2.7}$$

For any $\epsilon > 0$ we take

$$n_\epsilon := \left\lceil \log_q \frac{c}{\epsilon} \right\rceil.$$

Hence, $\lambda_{n_\epsilon} \leq \epsilon$. We get the following estimates

$$\begin{aligned} M(2\epsilon, f(\mathrm{Crit}(f))) &\leq M(2\epsilon, \mathcal{U}_\epsilon(p_{n_\epsilon} \circ L_{n_\epsilon}(\mathrm{Crit}(f)))) && \text{by (2.6)} \\ &\leq M(\epsilon, p_{n_\epsilon} \circ L_{n_\epsilon}(\mathrm{Crit}(f))) && \text{by definition} \\ &\leq M(\epsilon, p_{n_\epsilon} \circ L_{n_\epsilon}(C^{\lambda_{n_\epsilon}}(p_{n_\epsilon} \circ L_{n_\epsilon}) \cap B_X)) && \text{by (2.7).} \end{aligned}$$

By Proposition 2.17 we get

$$M(2\epsilon, f(\mathrm{Crit}(f))) \leq M(\epsilon, p_{n_\epsilon} \circ L_{n_\epsilon}(C^\epsilon(p_{n_\epsilon} \circ L_{n_\epsilon}) \cap B_X)) \leq n_\epsilon (2d)^{n_\epsilon}.$$

Hence, by definition of entropy dimension we have

$$\dim_e f(\mathrm{Crit}(f)) \leq \limsup_{\epsilon \to 0} \frac{\ln n_\epsilon + n_\epsilon \ln(2d)}{\ln \frac{1}{2\epsilon}} = \frac{\ln(2d)}{\ln q}.$$

□

Motivated by Kupka counterexample and Theorem 2.18, we introduce a class of polynomials on Hilbert spaces.

Definition 2.19. Let H be a Hilbert space, and let U be an open subset of H. We say that a function $f : U \to \mathbb{R}$ is a polynomial of degree d on U if it is C^1 and for any linear subspace E with $\dim E < \infty$ the restriction $f|_E$ is a polynomial of degree d.

The validity of the Morse-Sard property in this class of maps is subtle. Indeed, the map constructed by Kupka is a polynomial of degree 3 on ℓ^2. Hence, the Morse-Sard property does not hold in this class of maps. However, Yomdin found a sufficient condition for the validity of the Morse-Sard property for polynomial maps on ℓ^2, indeed we have the following corollary of Theorem 2.18, proved in [95].

Corollary 2.20. *Let $f : B_{\ell^2} \to \mathbb{R}$ be a polynomial of degree d. Suppose that there exist $q > 1$, $c > 0$ and a sequence of linear subspaces $E_n \simeq \mathbb{R}^n$ such that, denoting by $\pi_n : \ell^2 \to E_n$ the projection, for any $n \in \mathbb{N}$ we have*

$$\|f \circ \pi_n - f\|_{C^1_{B_{\ell^2}}} \leq cq^{-n}.$$

Then,

$$\dim_e(f(\mathrm{Crit}(f))) \leq \frac{\ln(2d)}{\ln q}$$

In particular, if $q > 2d$, then the set of critical values of f has measure zero.

We also have the following corollary of Theorem 2.18, which provides an explicit class of polynomial maps on ℓ^2 with the Morse-Sard property, it was proved in [95]. (The fact that these maps are polynomials on ℓ^2 can be proved as in Proposition 3.33.)

Corollary 2.21 (Yomdin)**.** *Let us consider a sequence of polynomials $p_i : \mathbb{R}^i \to \mathbb{R}$ of degrees at most d and such that*

$$\max_{y \in B_{\mathbb{R}^i}} |p_i(y)| \leq 1. \tag{2.8}$$

Then, given $q > 1$, the function $f : B_{\ell^2} \to \mathbb{R}$ defined as

$$f(x) = \sum_{i=1}^{\infty} \frac{1}{q^i} p_i(x_1, \ldots, x_i)$$

is C^1, and we have

$$\dim_e(f(\mathrm{Crit}(f))) \leq \frac{\ln(2d)}{\ln q}.$$

In particular, if $q > 2d$, then f has the Morse-Sard property: the set of its critical values has meaure zero.

Proof. First we observe that the funtion f is well-defined, thanks to the hypothesis (2.8). Furthermore, f is C^1, indeed, by Markov inequality (see [52, Thm. VI]) we have $\max_{y \in B_{\mathbb{R}^i}} \left\|D_y p_i\right\| \leq d^2$. To conclude, we prove that f satisfies the assumption (2.6) of Theorem 2.18. We consider for any $y \in \mathbb{R}^n$ the polynomial $f_n(y) := \sum_{i=1}^{n} \frac{1}{q^i} p_i(y_1, \ldots, y_i)$, which has degree at most d. For any $n \in \mathbb{N}$ we denote by $\pi_n : \ell^2 \to \mathbb{R}^n$ the projection on the first n entries of sequences in ℓ^2. We have

$$\|f - f_n \circ \pi_n\|_{C^1_{B_{\ell^2}}} \leq (1 + d^2) \sum_{i=n+1}^{\infty} \frac{1}{q^i} \leq c(d, q) q^{-n},$$

where $c(d, q)$ is a constant depending only on d and q. Hence we conclude by Theorem 2.18, setting $p_n := f_n$ and $L_n := \pi_n$. □

2.2.2 Contributions for higher-dimensional codomains

In Yomdin's version of the Morse-Sard theorem in infinite dimension, the fact that the codomain is one-dimensional is fundamental. A key aspect of the proof is that in Theorem 2.16 the dependence of all the quantities on n is explicit. This allows us to pass to the limit in the estimates as n goes to ∞.

In [62] we prove Morse-Sard theorems for polynomial *maps* from an infinite-dimensional Hilbert space to $\mathbb{R}^m$. The new aspect is that the dimension of the codomain is arbitrary. Following Yomdin's approach for the real-valued case, our results for higher-dimensional codomains are based on a refined version of Theorem 2.11. Indeed, we prove a version of the latter where the dependence on the dimensional parameter n is explicit, see Theorem G. This result has been obtained using new tools from semialgebraic geometry from [61], see Theorem M.

Using this key result, in [62] we give a criterion for the validity of the Morse-Sard poperty for polynomial *maps* from an infinite dimensional Hilbert space to $\mathbb{R}^m$, under the assumption that they are well-approximated by finite dimensional polynomial maps, see Theorem E. Starting from this result, we prove various quantitative Morse-Sard type theorems for polynomial maps from a Hilbert space to $\mathbb{R}^m$, presented in Section iii.1. In particular, we also prove Morse-Sard type theorems for polynomial maps without any approximation hypothesis: in this case, the Morse-Sard property holds on linear subspaces with small enough Kolmogorov width (but still infinite dimensional), see Theorem D.

Chapter 3

Sard properties for polynomial maps in infinite dimension

3.1 Quantitative variations estimates

3.1.1 Semialgebraic sets and maps

The goal of this section is to prove Theorem G. We begin by recalling some basic notions from semialgebraic geometry.

Definition 3.1 (Semialgebraic sets and maps)**.** We say that a set $S \subset \mathbb{R}^n$ is *semialgebraic* if

$$S = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \left\{ x \in \mathbb{R}^n \mid \operatorname{sign}(p_{ij}(x)) = \sigma_{ij} \right\}, \tag{3.1}$$

for some finite set of polynomials $p_{ij} \in \mathbb{R}[x_1, \ldots, x_n]$ and $\sigma_{ij} \in \{0, +1, -1\}$, where

$$\operatorname{sign}(r) := \begin{cases} +1 & r > 0, \\ -1 & r < 0, \\ 0 & r = 0. \end{cases}$$

We call the description (3.1) a *representation* of S.

A map $f : A \to B$ between semialgebraic sets $A \subset \mathbb{R}^n, B \subset \mathbb{R}^\ell$ is said to be *semialgebraic* if its graph is a semialgebraic set in $\mathbb{R}^n \times \mathbb{R}^\ell$.

The representation of a semialgebraic set S as in (3.1) is not unique. However, a representation is useful to quantify the complexity of S, using the following notion.

Definition 3.2 (Diagram of a semialgebraic set)**.** Let $S \subset \mathbb{R}^n$ be a semialgebraic set represented as in (3.1). We will say that the triple

$$\left(n, a \cdot \max_i \{b_i\}, \max_{i,j} \{\deg(p_{ij})\} \right) \in \mathbb{N}^3$$

is a *diagram* for S. Below, the equation "$D(S) = (m, c, d)$" will mean that there exists a representation of S as in (3.1) with $n \leq m$, $a \cdot \max_i\{b_i\} \leq c$ and $\max_{i,j}\{\deg(p_{ij})\} \leq d$.

Remark 3.3. There are several ways of quantifying the complexity of a semialgebraic set, depending on the way it is presented. Essentially, these notions should contain three pieces of information: the number of variables, the "combinatorics" of the presentation, and the degrees of the defining polynomials. For instance, every semialgebraic set $S \subset \mathbb{R}^n$ can be presented as

$$S = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} E_{ij},$$

where each E_{ij} has the form $\{x \in \mathbb{R}^n \mid p_{ij}(x) \leq 0\}$ or $\{x \in \mathbb{R}^n \mid p_{ij}(x) < 0\}$, where p_{ij} is a polynomial with $\deg(p_{ij}) \leq d$ for all i, j. Given a presentation of this type, it is immediate to see that S also admits a diagram $D(S) = (n, c, d)$ (i.e. it can be presented as in (3.1) with $c = c(a, b_1, \ldots, b_a)$.

Dimension and stratifications

Every semialgebraic subset of $\mathbb{R}^n$ can be written as a finite union of semialgebraic sets, each of them semialgebraically homeomorphic to an open cube $(0, 1)^m \subset \mathbb{R}^m$, for some $m \leq n$ ([31, Thm. 2.3.6]). This allows to define the *dimension* of a semialgebraic set as the maximum of the dimensions m of these cubes. Every semialgebraic set can be written as a finite union of smooth, semialgebraic disjoint manifolds called *strata* ([31, Prop. 9.1.8]).

Notice that the dimension of a semialgebraic set is preserved by semialgebraic homeomorphisms and behaves naturally under product structures:

$$\dim(A \times B) = \dim(A) + \dim(B).$$

Moreover, if $f : A \to B$ is a continuous semialgebraic map, then

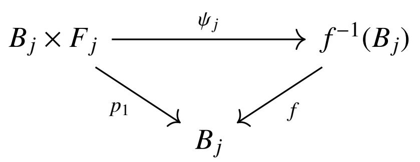
$$\dim(f(A)) \leq \dim(A). \tag{3.2}$$

Semialgebraic triviality

Continuous semialgebraic maps $f : A \to B$ are "piecewise" trivial fibrations: there exists a partition of B into finitely many semialgebraic sets

$$B = \bigsqcup_{j=1}^{b} B_j$$

and, for every $j = 1, \ldots, b$ there exist fibers $F_j := f^{-1}(b_j)$, with $b_j \in B_j$ and a semialgebraic homeomorphism $\psi_j : f^{-1}(B_j) \to B_j \times F_j$ that make the following diagram commutative:



This result is called *semialgebraic triviality*, see [13, Thm. 5.46].

Projections

The image of a semialgebraic set $S \subset \mathbb{R}^n \times \mathbb{R}^\ell$ under the projection map $\pi : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}^n$ is semialgebraic ([31, Thm. 2.2.1]). Moreover, it follows from [13, Thm. 14.16, Notation 8.2][1] that there exists $C_1 > 0$ such that if $(n + \ell, c, d)$ is a diagram for S, we can write

$$\pi(S) = \bigcup_{i \in I} \bigcap_{j \in J_i} \bigcup_{k \in N_{ij}} \left\{ x \in \mathbb{R}^n \mid \operatorname{sign}(p_{ijk}(x)) = \sigma_{ijk} \right\},$$

with

$$\#I \le c^{(n+1)(\ell+1)} d^{C_1 \ell n}, \quad \#J_i \le c^{\ell+1} d^{C_1 \ell}, \quad \#N_{ij} \le d^{C_1 \ell},$$

and with

$$\deg(p_{ijk}) \le d^{C_1 \ell}.$$

In particular, there exists $C_2 > 0$ such $\pi(S)$ admits a diagram with:

$$D(\pi(S)) = \left(n, (cd)^{C_2 \ell n} (d^{C_2 \ell})^{(cd)^{C_2 \ell}}, d^{C_2 \ell} \right).$$

In particular, the diagram of the projection of a semialgebraic set is controlled explicitly by the diagram of the original set.

Thom-Milnor bound

We will need the following qualitative bound on the number of connected components of a semialgebraic set, that follows from [13, Thm. 7.50].

Theorem 3.4 (Thom-Milnor)**.** *Let $c, d \in \mathbb{N}$. There exists a constant $\beta_{\mathrm{tm}} = \beta_{\mathrm{tm}}(c, d) > 1$, such that for any semialgebraic set $S \subset \mathbb{R}^n$ with diagram $D(S) = (n, c, d)$ it holds*

$$b_0(S) \le \beta_{\mathrm{tm}}^n, \tag{3.3}$$

where $b_0(S)$ denotes the number of connected components of S.

3.1.2 Definable choice

We recall from [61] the following *quantitative approximate definable choice* result, which will be a key tool in the proof of Theorem G. We state it here in the form we need in this chapter, which follows from Theorem M and Remark 10. We will prove it in Chapter 5 (see Theorem 5.43).

Theorem 3.5 (Quantitative approximate definable choice [61])**.** *For every $c, d, \ell \in \mathbb{N}$ there exists $\beta_{\mathrm{dc}} > 1$ satisfying the following statement. Let $n \in \mathbb{N}$ and let $A \subset \mathbb{R}^n$ be a closed semialgebraic set contained in the ball $B_{\mathbb{R}^n}(r)$ with diagram*

$$D(A) = (n, c, d).$$

Let also $F = (F_1, \ldots, F_\ell) : \mathbb{R}^n \to \mathbb{R}^\ell$ be a polynomial map with components of degree bounded by d. Then for every $\epsilon \in (0, r)$ there exists a closed semialgebraic set $C_\epsilon \subset \mathbb{R}^n$ such that

[1] Note that the role of k and ℓ for us is swapped with respect to [13, Thm. 14.16]: we eliminate ℓ variables from a list of $n + \ell$, they eliminate k variables from a list of $k + \ell$.

(i) $\dim(C_\epsilon) \le \ell$;

(ii) $C_\epsilon \subset \mathcal{U}_\epsilon(A)$;

(iii) $\mathrm{dist}_H(F(C_\epsilon), F(A)) \le L(F,r)\cdot\epsilon$, *where* $L(F,r) := 3 + \mathrm{Lip}(F, B_{\mathbb{R}^n}(2r))$;

(iv) *for every* $e = 1,\dots,n$ *and every affine space* $\mathbb{R}^e \simeq E \subset \mathbb{R}^n$, *the number of connected components of* $E \cap C_\epsilon$ *is bounded by*

$$b_0(E\cap C_\epsilon) \le \beta^e_{\mathrm{dc}}.$$

3.1.3 Variations and their behaviour under polynomial maps

We recall the definition of *variations*, introduced by Vitushkin [91, 92] and developed in the semialgebraic context by Comte and Yomdin [96]. The variations encode the size of a set, through an integral-geometric approach. Given a polynomial map $p : \mathbb{R}^n \to \mathbb{R}^m$ and a semialgebraic set $A \subset \mathbb{R}^n$, it is possible to estimate the variations of the set $p(A)$, in a quantitative way, see [96, Thm. 7.2]. However, these estimates are not quantitative with respect to the dimensional parameter n. The main result of this section is Theorem 3.10, where we prove a version of [96, Thm. 7.2], where the dependence on n is explicit, thanks to the use of Theorem 3.5.

For $i \le n$, we denote by $G_i(\mathbb{R}^n)$ the Grassmannian of all i-dimensional linear subspaces of $\mathbb{R}^n$. We endow it with the standard probability measure $\gamma_{i,n}$ defined through the action of the orthogonal group of $\mathbb{R}^n$. For any $Z \in G_i(\mathbb{R}^n)$ we the denote the orthogonal projection on Z by

$$\pi_Z : \mathbb{R}^n \to Z.$$

Definition 3.6 (Vitushkin variations)**.** Let $A \subset \mathbb{R}^n$ be a bounded semialgebraic set. We define the 0–th *variation* of A as

$$V_0(A) := b_0(A),$$

where $b_0(A)$ denotes the number of connected components of A. For $i = 1,\dots,n$ we define the i–th *variation* of A as

$$V_i(A) := c(n,i)\int_{G_i(\mathbb{R}^n)}\left(\int_Z b_0(A\cap\pi_Z^{-1}(x))\,\mathcal{H}^i(dx)\right)\gamma_{i,n}(dZ), \tag{3.4}$$

where $\mathcal{H}^i$ denotes the i-dimensional Hausdorff measure.

$$c(n,i) := \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{i+1}{2})\Gamma(\frac{n-i+1}{2})}. \tag{3.5}$$

We note that the integrand in (3.4) is measurable, see [96, Sec. 3].

A list of properties of variations is given in [96, page 35]. Here we prove the following one.

Lemma 3.7. *If* $A \subset \mathbb{R}^n$ *is a bounded semialgebraic set of dimension* ℓ, *then the* ℓ*-variation coincides with the* ℓ*-dimensional Hausdorff volume:*

$$V_\ell(A) = \mathcal{H}^\ell(A).$$

We also have $V_i(A) = 0$ *for all* $i > \ell$.

Proof. We recall a classical result in Integral Geometry, called Cauchy–Crofton formula [39, Thm. 5.14]: if A is a i-rectifiable and Borel subset of $\mathbb{R}^n$, for $i \leq n$, then it holds

$$\mathcal{H}^i(A) = c(n,i) \int_{G_i(\mathbb{R}^n)} \left(\int_Z \#(A \cap \pi_Z^{-1}(x))\, \mathcal{H}^i(dx) \right) \gamma_{i,n}(dZ). \tag{3.6}$$

If $A \subset \mathbb{R}^n$ is semialgebraic, the set A can be written as a finite union of smooth, semialgebraic disjoint manifolds called strata (see Section 3.1.1). We can relabel the strata so that

$$A = \bigsqcup_{j=0}^{\ell} A_j,$$

where each A_j is a finite union of smooth j-dimensional manifolds. For any fixed $Z \in G_\ell(\mathbb{R}^N)$ and for almost every $x \in Z$ the set $\pi_Z^{-1}(x)$ does not intersect[2] A_j for $j < \ell$. Hence, $V_0(A \cap \pi_Z^{-1}(x)) = V_0(A_\ell \cap \pi_Z^{-1}(x))$ for almost every $x \in Z$. Furthermore, for almost every $x \in Z$ the intersection $A_\ell \cap \pi_Z^{-1}(x)$ is transverse, and $\dim A + \dim \pi_Z^{-1}(x) = n$, so that $A_\ell \cap \pi_Z^{-1}(x)$ has dimension zero. Thus, its cardinality coincides with its number of connected components. We conclude by (3.6). □

Let M be a real $m \times m$ symmetric and positive semidefinite matrix. We denote by

$$\lambda_1(M) \geq \cdots \geq \lambda_m(M) \geq 0,$$

its eigenvalues, ordered in decreasing order.

Definition 3.8 (Singular values)**.** Let H be a Hilbert space with $\dim H \geq m$, and let $L : H \to \mathbb{R}^m$ be a linear and continuous map. For any $k = 1, \ldots, m$ the k-th *singular value* of L is

$$\sigma_k(L) = \lambda_k(LL^\top)^{\frac{1}{2}}.$$

In finite dimension, the following Weyl inequality holds for the singular values of the difference of two matrices (see [90, Ex. 1.3.22 (iv)]): for $n \geq m$ and linear maps $L_1, L_2 : \mathbb{R}^n \to \mathbb{R}^m$ we have

$$|\sigma_k(L_1) - \sigma_k(L_2)| \leq \|L_1 - L_2\|_{\mathrm{op}}, \qquad \forall\, k = 1, \ldots, m.$$

The Weyl inequality generalizes to infinite-dimensional spaces. Since we are not able to find a reference, we provide below a self-contained statement and its proof in the form that we will need later.

Lemma 3.9. *If $L_1, L_2 \in \mathcal{L}(H, \mathbb{R}^m)$, then $|\sigma_k(L_1) - \sigma_k(L_2)| \leq \|L_1 - L_2\|_{\mathrm{op}}$ for all $k = 1, \ldots, m$.*

Proof. Let $L : H \to \mathbb{R}^m$ be linear and continuous. Let $W \subset H$ a finite dimensional subspace such that $(\ker L)^\perp \subseteq W$, and denote by $L_W : W \to \mathbb{R}^m$ the restriction of L to W. More precisely, $L_W = L \circ i_W$ where $i_W : W \hookrightarrow H$ is the inclusion. We now prove that for all $k \leq m$

$$\sigma_k(L) = \sigma_k(L_W). \tag{3.7}$$

[2]We note that in this step of the proof it is enough to use the Semialgebraic Sard's theorem [31, Thm. 9.6.2], which is logically independent from the result in the smooth category.

Indeed, since the adjoint $i_W^\top$ coincides with the orthogonal projection $\pi_W : H \to W \subset H$, we have

$$\lambda_k(L_W \circ L_W^\top) = \lambda_k(L \circ i_W \circ i_W^\top \circ L^\top) = \lambda_k(L \circ \pi_W \circ L^\top) = \lambda_k(L \circ L^\top).$$

In the last equality, we used the fact that $L \circ \pi_W = L$, indeed $L(v) = L(\pi_W v + \pi_{W^\perp} v) = L(\pi_W v)$ since $W^\perp \subseteq \ker L$. Now given $L_1, L_2 \in \mathcal{L}(H, \mathbb{R}^m)$ let $W := (\ker L_1)^\perp + (\ker L_2)^\perp$, which is a finite dimensional linear subspace of H. By (3.7) and the finite dimensional Weyl inequality, we have

$$|\sigma_k(L_1) - \sigma_k(L_2)| = |\sigma_k((L_1)_W) - \sigma_k((L_2)_W)| \leq \|(L_1)_W - (L_2)_W\|_{\mathrm{op}} \leq \|L_1 - L_2\|_{\mathrm{op}},$$

concluding the proof. □

In the next result, given a semialgebraic set $A \subset \mathbb{R}^n$ and a polynomial map $p : \mathbb{R}^n \to \mathbb{R}^m$, we bound the i-th variation $V_i(p(A))$ in terms of the singular values of p and the diagram of A. Its proof follows the blueprint of [96, Thm. 7.2]. The main novelty (which is key for our work) is that the dependence on the dimension n is explicit.

Theorem 3.10. *For every $c, d, m \in \mathbb{N}$ there exists $\beta_{\mathrm{yc}} = \beta_{\mathrm{yc}}(c, d, m) > 1$ with the following property. For $n \geq m$, let $A \subset \mathbb{R}^n$ be a closed and bounded semialgebraic set with $A \subset B_{\mathbb{R}^n}(r)$ and diagram $D(A) = (n, c, d)$. Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degrees at most d. For all $i = 1, \ldots, m$ set*

$$\overline{\sigma}_i := \sup_{x \in A} \Big(\sigma_1(D_x p) \cdots \sigma_i(D_x p)\Big), \qquad \overline{\sigma}_0 := 1. \tag{3.8}$$

Then, for all $i = 0, \ldots, m$ it holds

$$V_i(p(A)) \leq \mathrm{cst}(m, r) n^m \beta_{\mathrm{yc}}^n \overline{\sigma}_i$$

where $\mathrm{cst}(m, r) > 0$ *is a constant that depends only on m, r.*

Remark 1 (Origin of β_{yc}). An inspection of the proof shows that it holds $\beta_{\mathrm{yc}} = \beta_{\mathrm{tm}}^2 \beta_{\mathrm{dc}}$ where $\beta_{\mathrm{tm}} = \beta_{\mathrm{tm}}(c(m{+}1), d)$ is the constant from the Thom-Milnor-type bound of Theorem 3.4, and $\beta_{\mathrm{dc}} = \max_{i \leq m} \beta_{\mathrm{dc}}(c, d, i)$ is the definable choice constant from Theorem 3.5.

Remark 2. In the course of the proof, see (3.14), we obtain the following estimate, reminiscent of an area formula for variations, which is an improved version of [96, Thm. 7.2], with the explicit dependence of the constants on n: there exist a constant $\beta_{\mathrm{yc}} = \beta_{\mathrm{yc}}(c, d, m)$, such that

$$V_i(p(A)) \leq c(m, i) \beta_{\mathrm{yc}}^n \overline{\sigma}_i V_i(A), \qquad i = 1, \ldots, m,$$

where $c(m, i)$ is the constant in (3.5).

Proof. By Definition 3.6, for $i = 1, \ldots, m$, we have

$$V_i(p(A)) = c(m, i) \int_{G_i(\mathbb{R}^m)} \left(\int_Z b_0(p(A) \cap \pi_Z^{-1}(x)) \mathcal{H}^i(dx) \right) \gamma_{i,m}(dZ)\,.$$

For any $x \in Z$, the set $A \cap (\pi_Z \circ p)^{-1}(x)$ is semialgebraic with controlled diagram (more precisely, $D(A \cap (\pi_Z \circ p)^{-1}(x)) = (n, c(1 + m), d)$. Using Theorem 3.4 we obtain

$$b_0(p(A) \cap \pi_Z^{-1}(x)) \leq b_0(A \cap (\pi_Z \circ p)^{-1}(x))) \leq \beta_{\mathrm{tm}}^n, \tag{3.9}$$

where $\beta_{\rm tm} = \beta_{\rm tm}(c(1+m), d)$ depends only on c, d, m. Note that, since by definition $V_0 = b_0$, (3.9) yields the thesis of the theorem for the case $i = 0$. We continue then assuming $i = 1, \dots, m$. Denoting by $\mathbb{1}_S$ the characteristic function of a set S, we get

$$\begin{aligned} V_i(p(A)) &\le c(m,i)\beta_{\rm tm}^n \int_{G_i(\mathbb{R}^m)} \left(\int_Z \mathbb{1}_{\pi_Z \circ p(A)}(x)\, \mathcal{H}^i(dx) \right) \gamma_{i,m}(dZ) \\ &= c(m,i)\beta_{\rm tm}^n \int_{G_i(\mathbb{R}^m)} \mathcal{H}^i(\pi_Z \circ p(A))\, \gamma_{i,m}(dZ). \end{aligned} \tag{3.10}$$

We would like to estimate $\mathcal{H}^i(\pi_Z \circ p(A))$ applying area formula to the map

$$p_Z := \pi_Z \circ p : \mathbb{R}^n \to \mathbb{R}^i.$$

However, exactly as in [96, Thm. 7.2], we cannot apply directly the area formula to A, since the dimension of A can be larger than i. In [96, Thm. 7.2] this problem is solved using [96, Ex. 4.11]. Here we overcome the obstacle using Theorem 3.5, replacing the set A with a suitable approximation.

For sufficiently small $\epsilon > 0$ let C_ϵ the semialgebraic set obtained by applying Theorem 3.5 to the set A and the polynomial map $F = p_Z : \mathbb{R}^n \to \mathbb{R}^i$. We then proceed using the area formula to estimate $\mathcal{H}^i(\pi_Z \circ p(C_\epsilon))$. (We will see at the end of the proof how $\mathcal{H}^i(\pi_Z \circ p(C_\epsilon)) \to \mathcal{H}^i(\pi_Z \circ p(A))$ as $\epsilon \to 0$). By Item (i) of Theorem 3.5, $\dim(C_\epsilon) \le i$. We can assume without loss of generality that $C_\epsilon \subset \mathbb{R}^n$ is a smooth embedded submanifold with dimension equal to i, since the other strata give zero contribution to $\mathcal{H}^i$. Denoting with $\bar{p}_Z : C_\epsilon \to \mathbb{R}^i$ the restriction, we obtain

$$\begin{aligned} \mathcal{H}^i(p_Z(C_\epsilon)) &\le \int_{C_\epsilon} |\det(D_x \bar{p}_Z)| \mathcal{H}^i(dx) && \text{by the area formula} \\ &\le \int_{C_\epsilon} \sigma_1(D_x \bar{p}_Z) \cdots \sigma_i(D_x \bar{p}_Z) \mathcal{H}^i(dx). && \text{by definition of singular values} \end{aligned}$$

Note that, denoting with $j : C_\epsilon \hookrightarrow \mathbb{R}^n$ the inclusion, we have

$$D_x \bar{p}_Z : T_x C_\epsilon \xrightarrow{D_x j} \mathbb{R}^n \xrightarrow{D_x p} \mathbb{R}^m \xrightarrow{D_{p(x)}\pi_Z} \mathbb{R}^i.$$

Thus, since $D_{p(x)}\pi_Z$ is an orthogonal projection and $D_x j$ is a linear inclusion, we have

$$\langle z, (D_x \bar{p}_Z)(D_x \bar{p}_Z)^\top z \rangle \le \langle z, (D_x p)(D_x p)^\top z \rangle, \qquad \forall z \in Z = \mathbb{R}^i.$$

It follows that $\sigma_k(D_x \bar{p}_Z) \le \sigma_k(D_x p)$ for all $k = 1, \dots, i$. Continuing the above inequalities we obtain

$$\begin{aligned} \mathcal{H}^i(p_Z(C_\epsilon)) &\le \int_{C_\epsilon} \sigma_1(D_x p) \cdots \sigma_i(D_x p) \mathcal{H}^i(dx) \\ &\le \sup_{x \in \mathcal{U}_\epsilon(A)} \sigma_1(D_x p) \cdots \sigma_i(D_x p) \mathcal{H}^i(C_\epsilon) && \text{by Item (ii) of Theorem 3.5} \\ &\le (\overline{\sigma}_i + \eta(\epsilon))\, \mathcal{H}^i(C_\epsilon), && \text{by continuity of singular values} \end{aligned}$$

where $\overline{\sigma}_i$ is defined in (3.8), and $\eta(\epsilon) \to 0$ as $\epsilon \to 0$. Since C_ϵ is a bounded semialgebraic set of dimension i, then $\mathcal{H}^i(C_\epsilon) = V_i(C_\epsilon)$ by Lemma 3.7. Hence

$$\begin{aligned}
\mathcal{H}^i(C_\epsilon) &= c(n,i)\int_{G_i(\mathbb{R}^n)}\left(\int_Z b_0(C_\epsilon \cap \pi_Z^{-1}(x))\,\mathcal{H}^i(dx)\right)\gamma_{i,n}(dZ) \\
&\le c(n,i)\beta_{\mathrm{dc}}^{n-i}\int_{G_i(\mathbb{R}^n)}\left(\int_Z \mathbb{1}_{\pi_Z(C_\epsilon)}(x)\,\mathcal{H}^i(dx)\right)\gamma_{i,n}(dZ) && \text{by Item (iv) of Theorem 3.5} \\
&\le c(n,i)\beta_{\mathrm{dc}}^{n-i}\int_{G_i(\mathbb{R}^n)}\left(\int_Z \mathbb{1}_{\pi_Z(\mathcal{U}_\epsilon(A))}(x)\,\mathcal{H}^i(dx)\right)\gamma_{i,n}(dZ) && \text{by Item (ii) of Theorem 3.5} \\
&\le c(n,i)\beta_{\mathrm{dc}}^{n-i}\int_{G_i(\mathbb{R}^n)}\left(\int_Z b_0(\mathcal{U}_\epsilon(A) \cap \pi_Z^{-1}(x))\,\mathcal{H}^i(dx)\right)\gamma_{i,n}(dZ) \\
&= \beta_{\mathrm{dc}}^{n-i}V_i(\mathcal{U}_\epsilon(A)),
\end{aligned}$$

where $\beta_{\mathrm{dc}} = \beta_{\mathrm{dc}}(c,d,i) > 1$ is the constant from Theorem 3.5, which (up to taking the maximum for $i \le m$) depends only on c, d, m. Summing up, we have proved that for all $i = 1, \ldots, m$ it holds

$$\mathcal{H}^i(p_Z(C_\epsilon)) \le (\overline{\sigma}_i + \eta(\epsilon))\beta_{\mathrm{dc}}^n V_i(\mathcal{U}_\epsilon(A)). \tag{3.11}$$

In order to take the limit in (3.11), we recall two properties of semialgebraic sets. The first one is the continuity of the Lebesgue measure in the Hausdorff topology in the semialgebraic category (see [96, Thm. 5.10]): if $\{S_\epsilon\}_{\epsilon\ge 0} \subset \mathbb{R}^i$ is a one-parameter family of bounded semialgebraic sets then

$$\lim_{\epsilon\to 0}\mathrm{dist}_H(S_\epsilon, S_0) = 0 \qquad \Longrightarrow \qquad \lim_{\epsilon\to 0}\mathcal{H}^i(S_\epsilon) = \mathcal{H}^i(S_0). \tag{3.12}$$

The second property is the behaviour of variations for ϵ-neighbourhoods (see [96, Thm. 5.11]): if $A \subset \mathbb{R}^n$ is a bounded semialgebraic set, then for all $i = 1, \ldots, n$ it holds

$$\lim_{\epsilon\to 0} V_i(\mathcal{U}_\epsilon(A)) \le V_i(A). \tag{3.13}$$

Therefore, using (3.12) (and recalling that by Item (iii) of Theorem 3.5 it holds $p_Z(C_\epsilon) \to p_Z(A)$ in the Hausdorff topology) and (3.13) we can pass to the limit in (3.11) and obtain

$$\mathcal{H}^i(p_Z(A)) \le \overline{\sigma}_i\beta_{\mathrm{dc}}^n V_i(A).$$

Hence from (3.10) we get for all $i = 1, \ldots, m$

$$V_i(p(A)) \le c(m,i)\beta_{\mathrm{tm}}^n\beta_{\mathrm{dc}}^n\overline{\sigma}_i V_i(A). \tag{3.14}$$

To conclude, we estimate $V_i(A)$. Similarly as in (3.10), and recalling that $A \subset B_{\mathbb{R}^n}(r)$ we have

$$V_i(A) \le c(n,i)\beta_{\mathrm{tm}}^n\int_{G_i(\mathbb{R}^n)}\mathcal{H}^i(\pi_Z(A))\gamma_{i,n}(dZ) \le c(n,i)\beta_{\mathrm{tm}}^n\mathcal{H}^i(B_{\mathbb{R}^i}(r)) = c(n,i)\beta_{\mathrm{tm}}^n\frac{\pi^{i/2}}{\Gamma\left(\frac{i}{2}+1\right)}r^i. \tag{3.15}$$

Putting together (3.14) and (3.15), we obtain

$$V_i(p(A)) \le \left[c(m,i)c(n,i)\frac{\pi^{i/2}}{\Gamma\left(\frac{i}{2}+1\right)}\right](\beta_{\mathrm{tm}}^2\beta_{\mathrm{dc}})^n\overline{\sigma}_i r^i.$$

Set $\beta_{\mathrm{yc}} = \beta_{\mathrm{tm}}^2 \beta_{\mathrm{dc}}$. Using the form of the constants $c(m,i), c(n,i)$ in (3.5) and elementary estimates we obtain that for all $i = 0, \dots, m$ there exists a constant $\mathrm{cst}(i) > 0$ such that

$$V_i(p(A)) \le \mathrm{cst}(i) n^i \beta_{\mathrm{yc}}^n \overline{\sigma}_i r^i.$$

Taking the maximum over $i = 1, \dots, m$ we obtain the result with $\mathrm{cst}(m,r) = \max_{i=0,\dots,m} \mathrm{cst}(i) r^i$. □

3.1.4 Variations of almost-critical values of polynomial maps

In this section we recall the notion of almost-critical points $C^\Lambda(p)$ of a polynomial map $p : \mathbb{R}^n \to \mathbb{R}^m$. The set $p(C^\Lambda(p))$ is the set of its almost-critical values. The main result of this section is Theorem 3.15, which provides a version of [96, Cor. 7.4] where the dependence of the parameters with respect to n is explicit. This result is based on Theorem 3.10 and on the specific semialgebraic structure of almost-critical points of polynomial maps, that we prove in Proposition 3.14.

Definition 3.11. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a C^1 map with $n \ge m$. Given $\Lambda = (\Lambda_1, \dots, \Lambda_m) \in \mathbb{R}_+^m$, the set of *almost-critical points* of f is defined as

$$C^\Lambda(f) := \left\{ x \in \mathbb{R}^n \middle| \sigma_i(D_x f) \le \Lambda_i, \ \forall\, i = 1, \dots, m \right\}.$$

The following lemma is fundamental to study the semialgebraic structure of the set of almost-critical points of a polynomial map.

Lemma 3.12. *Let $A \subset \mathbb{R}^s$ be a semialgebraic set and let $f : A \to \mathbb{R}$ be a semialgebraic function. Then, for any $t \in \mathbb{R}$, the sublevel set $\{x \in A \mid f(x) \le t\}$ admits a semialgebraic description whose diagram does not depend on t.*

Proof. The graph of f is semialgebraic, hence we can write it as

$$\mathrm{graph}(f) = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} E_{ij},$$

where E_{ij} is of the form $\{p_{ij} < 0\}$ or $\{p_{ij} \le 0\}$ and $p_{ij} : \mathbb{R}^s \times \mathbb{R} \to \mathbb{R}$ is a polynomial. Denoting by $\pi_1 : \mathbb{R}^s \times \mathbb{R} \to \mathbb{R}^s$ the projection, we have

$$\{x \in A \mid f(x) \le t\} = \pi_1\Big(\mathrm{graph}(f) \cap \{(x,y) \in \mathbb{R}^s \times \mathbb{R} \mid y - t \le 0\}\Big). \tag{3.16}$$

The set $\mathrm{graph}(f) \cap \{(x,y) \in \mathbb{R}^s \times \mathbb{R} \mid y - t \le 0\}$ is semialgebraic, with a description given by

$$\mathrm{graph}(f) \cap \{(x,y) \mid y - t \le 0\} = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i+1} E_{i,j},$$

where $E_{i,b_i+1} = \{(x,y) \in \mathbb{R}^s \times \mathbb{R} \mid y - t \le 0\}$. In particular, this set admits a diagram that does not depend on t, since the degree of the polynomial $y - t$ is equal to 1. As explained in Section 3.1.1, the projection of a semialgebraic set is semialgebraic and its diagram is controlled explicitly by the diagram of the original set only. In particular, the diagram of the projection (3.16) admits a description with a diagram (s, I, J), that does not depend on t. □

Let Sym_m^+ denote the set of positive semidefinite matrices of size m. This is clearly a semialgebraic subset of all $m \times m$ real matrices. The following lemma is elementary and we omit its proof.

Lemma 3.13. *The functions $\lambda_i : \mathrm{Sym}_m^+ \to \mathbb{R}$ and are continuous and semialgebraic.*

In the following proposition we provide an estimate on the diagram of a description of the (semialgebraic) set of almost-critical points of polynomial maps.

Proposition 3.14. *Let $d, m \in \mathbb{N}$. Then there are $c' = c'(d, m)$ and $d' = d'(d, m)$ such that for all $n \geq m$ and any polynomial map $p : \mathbb{R}^n \to \mathbb{R}^m$ with components of degrees at most d, and all $\Lambda = (\Lambda_1, \dots, \Lambda_m) \in \mathbb{R}_+^m$, the set $C^\Lambda(p)$ is closed, semialgebraic, with*

$$D(C^\Lambda(p)) = (n, c', d').$$

(In particular, the diagram does not depend on Λ and it depends on n only in the number of variables.)

Proof. By Lemma 3.13, for every $k = 1, \dots, m$, the function $\lambda_k : \mathrm{Sym}_m^+ \to \mathbb{R}$ is semialgebraic, hence from Lemma 3.12 applied to $A = \mathrm{Sym}_m^+$ and the semialgebraic functions $\lambda_k : A \to \mathbb{R}$, we deduce the following fact: for any $t \in \mathbb{R}$ we can write

$$\{M \in \mathrm{Sym}_m^+ \mid \lambda_k(M) \leq t^2\} = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \{M \in \mathbb{R}^{m\times m} \mid P_{k,i,j,t}(M) \leq 0\}$$

for some polynomials $P_{k,i,j,t} : \mathbb{R}^{m\times m} \to \mathbb{R}$, $a, b_1, \dots, b_a \in \mathbb{N}$, and with $a \max_i b_i$ and $\max_{i,j} \deg(P_{k,i,j,t})$ depending only on m and on k (and not on t). We can assume that a does not depend on k, and all the b_i's do not depend on a, k, simply by adding empty or trivial sets.

Now, for any $k = 1, \dots, m$

$$\{N \in \mathbb{R}^{m\times n} \mid \sigma_k(N) \leq t\} = \{N \in \mathbb{R}^{m\times n} \mid \lambda_k(NN^\top) \leq t^2\},$$

therefore we can write

$$\{x \in \mathbb{R}^n \mid \sigma_k(D_x p) \leq t\} = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \{x \in \mathbb{R}^n \mid P_{k,i,j,t}(D_x p D_x p^\top) \leq 0\}.$$

Observe that $N \to NN^\top$ is polynomial with components of degree 2. Therefore, the map $q : \mathbb{R}^n \to \mathbb{R}^{m\times m}$, given by $q(x) := D_x p D_x p^\top$ is polynomial, with components of degree at most $2d$. Hence, the polynomial

$$P_{k,i,j,t} \circ q : \mathbb{R}^n \to \mathbb{R}$$

has degree at most $2d \max_{i,j} \deg P_{k,i,j,t}$, and in particular its degree has a bound that depends on only d, m, k (and not on t). The set of almost-critical points of p can be described as

$$C^\Lambda(p) = \bigcap_{k=1}^{m} \{x \in \mathbb{R}^n \mid \sigma_k(D_x p) \leq \Lambda_k\} = \bigcap_{k=1}^{m} \bigcup_{i=1}^{a} \bigcup_{j=1}^{b_i} \{x \mid P_{k,i,j,\Lambda_k}(D_x p D_x p^\top) \leq 0\}.$$

It is clear that $C^\Lambda(p)$ is closed. Note that $d' := \max_{k,i,j} \deg(P_{k,i,j,\Lambda_k} \circ q)$ depends only on d, m. Using the distributivity of intersection, we can rewrite $C^\Lambda(p)$ in the form (3.1), where the number of unions and intersections depends only on m, and the maximum degree of the polynomials is $c' = c'(d, m)$, concluding the proof. □

Now we state and prove the main result of this section (Theorem G in the Introduction).

Theorem 3.15. *Let $n \geq m$, and $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d. For $i = 0, \dots, m$, $\Lambda = (\Lambda_1, \dots, \Lambda_m) \in \mathbb{R}^m_+$ and $r > 0$, we have*

$$V_i(p(C^\Lambda(p) \cap B_{\mathbb{R}^n}(r))) \leq \mathrm{cst}(m, r) n^m \beta_0^n \Lambda_0 \cdots \Lambda_i, \tag{3.17}$$

where $\beta_0 = \beta_0(d, m)$ depends only on d and m, $\mathrm{cst}(m, r)$ depends only on m, r, and we set $\Lambda_0 = 1$.

Proof. By Proposition 3.14 the set $A = C^\Lambda(p) \cap B_{\mathbb{R}^n}(r)$ is a closed and bounded semialgebraic set whose diagram is $D(A) = (n, c', d')$ where $c' = c'(d, m)$, $d' = d'(d, m)$ (which in particular do not depend on n, Λ). Furthermore, for the polynomial map $p : \mathbb{R}^n \to \mathbb{R}^m$ it holds, by construction

$$\overline{\sigma}_i := \sup_{x \in A} \Big(\sigma_1(D_x p) \cdots \sigma_i(D_x p) \Big) \leq \Lambda_0 \cdots \Lambda_i.$$

We can then apply Theorem 3.10, yielding (3.17) with $\beta_0 = \beta_{\mathrm{yc}}(c', \max\{d, d'\}, m)$. □

3.2 Sard properties for infinite-dimensional maps

Let H be a Hilbert space. We consider a map $f : U \to \mathbb{R}^m$ where U is an open subset of H. If f is C^1 we denote by $\mathrm{Crit}(f)$ the set of its critical points. For a fixed $\nu \in \mathbb{N}$, we also consider the set

$$\mathrm{Crit}_\nu(f) = \{x \in U \mid \mathrm{rank}(D_x f) \leq \nu\}. \tag{3.18}$$

Given a relatively compact $S \subset \mathbb{R}^m$, for any $\epsilon > 0$ we denote with $M(\epsilon, S)$ the *ϵ-entropy* of S, which is the minimum number of closed balls of radius ϵ that we need to cover S.

Definition 3.16 (Entropy dimension)**.** The *entropy dimension* of S is defined as

$$\dim_e(S) = \limsup_{\epsilon \to 0^+} \frac{\ln M(\epsilon, S)}{\ln(\frac{1}{\epsilon})}.$$

In [96, Ch. 2], it is proved that

$$\dim_{\mathcal{H}}(S) \leq \dim_e(S).$$

The inequality can be strict, since there are sequences of real numbers with positive entropy dimension.

For Banach spaces X, Y and linear continuous maps $L \in \mathcal{L}(X, Y)$, we denote by

$$\|L\|_{\mathrm{op}} := \sup_{\|x\|_X \leq 1} \|Lx\|_Y \tag{3.19}$$

the operator norm. With some abuse of notation the same symbol $\|L\|_{\mathrm{op}}$ is used for different X and Y, but there should be no confusion as the domain and codomain of L are clear from the context. We also use without risk of confusion the symbol $\|\cdot\|$ to denote the usual Euclidean norm of $\mathbb{R}^m$.

3.2.1 Sard-type theorem for well approximated maps

The purpose of this section is to prove the following theorem, which is our main Sard-type result (Theorem E in the Introduction).

Theorem 3.17. *Let $d, m \in \mathbb{N}$. There exists a constant $\beta_0 = \beta_0(d, m) > 0$ such that the following holds. Let H be a Hilbert space, and let $f : H \to \mathbb{R}^m$ be a C^1 map such that its differential $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is weakly continuous. Let $K \subset H$ be a bounded set with this approximation property: there exist a sequence $E_n \subset H$ of linear subspaces,* $\dim(E_n) = n$*, and polynomial maps $f_n : E_n \to \mathbb{R}^m$ with uniformly bounded degree:*

$$\sup_{n\in\mathbb{N}} \deg f_n \leq d < \infty,$$

such that for some $q > 1$, $c \geq 0$, and all large enough n it holds

$$\sup_{x\in K} \Big(\|f(x) - f_n \circ \pi_{E_n}(x)\| + \|(D_x f)|_{E_n} - D_{\pi_{E_n}(x)} f_n\|_{\mathrm{op}} \Big) \leq c q^{-n}. \tag{3.20}$$

Then

$$\dim_e \Big(f(\,\mathrm{Crit}_\nu(f) \cap K) \Big) \leq \nu + \frac{\ln \beta_0}{\ln q}, \qquad \forall\, \nu = 1, \dots, m-1.$$

In particular, if $q > \beta_0$, then f satisfies the Sard property on K:

$$\mu\Big(f(\mathrm{Crit}(f) \cap K) \Big) = 0.$$

Remark 3. As it will be clear from the proof, the upper bound cq^{-n} in the r.h.s. of (3.20) can be replaced by any $C_n \geq 0$, such that $\limsup_{n\to\infty} C_n^{1/n} = q^{-1}$, yielding the same results.

For the benefit of the reader we provide a general outline of the strategy of the proof. Our goal is to estimate the "size" of the set $f(\mathrm{Crit}_\nu(f) \cap K)$. More precisely, we bound the *ϵ-entropy* of $f(\mathrm{Crit}_\nu(f) \cap K)$ for any $\epsilon > 0$, and then pass to the limit $\epsilon \to 0$.

We observe that the ϵ-entropy of $f(\mathrm{Crit}_\nu(f) \cap K)$ is controlled with the one of the set of almost-critical values $f(C^\Lambda(f) \cap K)$, see Definition 3.8. In turn, this is controlled by the ϵ-entropy of suitable almost-critical values of the approximating maps f_n, namely $f(C^{\Lambda_\epsilon}(f_n) \cap K)$, see (3.25).

To proceed, we use the relation between the ϵ-entropy of sets in $\mathbb{R}^m$ and their Vitushkin variations (Definition 3.6), provided by [96, Thm. 3.5].

Finally, we estimate the Vitushkin variations of $f_n(C^{\Lambda_\epsilon}(f_n) \cap K)$ applying Theorem G, exploiting the hypothesis that the maps f_n are polynomial. This is the connection with quantitative semialgebraic geometry from Section 3.1. The crucial point of Theorem G is that the bound is quantitative with respect to parameters $\Lambda = \Lambda_\epsilon$ and n. This allows to chose a large enough $n = n_\epsilon$, see (3.31), to ensure convergence when passing to the limit $\epsilon \to 0$.

Proof. We divide the proof in steps.

Step 1. In the first step we show that for any $\nu = 1, \dots, m-1$ the set $\mathrm{Crit}_\nu(f) \cap K$ is contained in a suitable set of almost-critical points of f. By assumption, K is weakly pre-compact. Since $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is weakly continuous, and by Lemma 3.9 the function $\sigma_i : \mathcal{L}(H, \mathbb{R}^m) \to \mathbb{R}$ is continuous for every $i = 1, \dots, m$. Therefore we have

$$\Sigma_i := \sup_K \sigma_i(D_x f) < \infty.$$

Hence, defining $\Lambda := (\Sigma_1, \dots, \Sigma_\nu, 0 \dots, 0) \in \mathbb{R}^m$, we get

$$\mathrm{Crit}_\nu(f) \cap K \subset C^\Lambda(f) \cap K. \tag{3.21}$$

Step 2. We now relate critical points of f to almost-critical points of the approximating polynomials f_n. Fix $\epsilon > 0$. Recall that, by Lemma 3.9, the singular values are 1-Lipschitz with respect to the operator norm. Then for any sufficiently large $n \in \mathbb{N}$ such that $cq^{-n} \leq \epsilon$ and for all $x \in K$ we have

$$|\sigma_i\left(D_x f|_{E_n}\right) - \sigma_i(D_{\pi_{E_n}(x)} f_n)| \leq \|\left(D_x f\right)|_{E_n} - D_{\pi_{E_n}(x)} f_n\|_{\mathrm{op}} \leq \epsilon.$$

Hence, for all $x \in K \cap C^\Lambda(f)$ we have the following estimate for all $i = 1, \dots, m$

$$\sigma_i\left(D_{\pi_{E_n}(x)} f_n\right) \leq \epsilon + \sup_{y \in C^\Lambda(f) \cap K} \sigma_i\left(D_y f|_{E_n}\right) \leq \epsilon + \sup_{y \in C^\Lambda(f) \cap K} \sigma_i\left(D_y f\right) \leq \epsilon + \Lambda_i, \tag{3.22}$$

where, in the second inequality, we used that singular values are monotone with respect to restriction to subspaces. Hence, defining

$$\Lambda^\epsilon := (\Sigma_1 + \epsilon, \dots, \Sigma_\nu + \epsilon, \epsilon, \dots, \epsilon),$$

by (3.22) we get

$$\pi_{E_n}\left(C^\Lambda(f) \cap K\right) \subset C^{\Lambda^\epsilon}(f_n) \cap \pi_{E_n}(K) \subset C^{\Lambda^\epsilon}(f_n) \cap B_{E_n}(r), \tag{3.23}$$

where $r > 0$ is such that $K \subseteq B_H(r)$, and thus $B_{E_n}(r) = B_H(r) \cap E_n$ denotes the ball in E_n defined by the restriction of the Hilbert norm.

Step 3. Now we use the inclusion (3.23) to estimate the entropy dimension of $f(\mathrm{Crit}_\nu(f) \cap K)$ with the one of almost-critical values of f_n. By (3.20), for sufficiently large n, we have

$$\sup_{x \in K} \|f(x) - f_n \circ \pi_{E_n}(x)\| \leq \epsilon. \tag{3.24}$$

Finally, we get

$$\begin{aligned}
M(2\epsilon, f(\mathrm{Crit}_\nu(f) \cap K) &\leq M(2\epsilon, f(C^\Lambda(f) \cap K)) && \text{by (3.21)} \\
&\leq M\left(2\epsilon, \mathcal{U}_\epsilon\left(f_n \circ \pi_{E_n}\left(C^\Lambda(f) \cap K\right)\right)\right) && \text{by (3.24)} \\
&\leq M\left(\epsilon, f_n \circ \pi_{E_n}\left(C^\Lambda(f) \cap K\right)\right) && \text{by definition of } \mathcal{U}_\epsilon \\
&\leq M(\epsilon, f_n(C^{\Lambda^\epsilon}(f_n) \cap B_{E_n}(r))), && \text{by (3.23)}
\end{aligned} \tag{3.25}$$

where we recall that $M(\varepsilon, A)$ is the minimal number of closed balls of $\mathbb{R}^m$ of radius ϵ needed to cover $A \subset \mathbb{R}^m$, and $\mathcal{U}_\epsilon(A)$ denotes the ϵ neighbourhood of A. This concludes the proof of step 3.

The rest of the proof consists in estimating (3.25). We will use the results from Section 3.1.4, for this reason in the next step we discuss how to bring the problem to the Euclidean space $\mathbb{R}^n$.

Step 4. We fix a basis for E_n which is orthonormal with respect to the inner product of H. This yields a linear isometry of Hilbert spaces $\ell_n : \mathbb{R}^n \to E_n$. In particular, for the transpose, it holds $\ell_n^\top = \ell_n^{-1}$. Consider then the polynomial map $\tilde{f}_n : \mathbb{R}^n \to \mathbb{R}^m$ defined by $\tilde{f}_n := f_n \circ \ell_n$, $n \geq m$. We have

$$\sigma_i(D_v \tilde{f}_n) = \lambda_i(D_{\ell_n(v)} f_n \circ \ell_n \circ \ell_n^\top \circ D_{\ell_n(v)} f_n^\top)^{1/2} = \sigma_i(D_{\ell_n(v)} f_n), \qquad \forall\, v \in \mathbb{R}^n.$$

It follows that, for sufficiently large n,

$$\ell_n(C^{\Lambda^\epsilon}(\tilde{f}_n)) = C^{\Lambda^\epsilon}(f_n).$$

Hence, we have

$$f_n(C^{\Lambda^\epsilon}(f_n) \cap B_{E_n}(r)) = f_n(\ell_n(C^{\Lambda^\epsilon}(\tilde{f}_n)) \cap B_{E_n}(r)) = \tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r)). \tag{3.26}$$

By [96, Thm. 3.5] we have

$$M(\epsilon, \tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r))) \leq C(m) \sum_{i=0}^{m} \frac{1}{\epsilon^i} V_i\Big(\tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r))\Big), \tag{3.27}$$

where $C(m) > 0$ is a constant that depends only on m. We now apply the results from Section 3.1.4 in order to estimate the right-hand side of the last inequality.

Step 5. Take n sufficiently large so that it also satisfies $n \geq m$. Then, by Theorem G, for every $i = 0, \ldots, m$ we have

$$V_i\Big(\tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r))\Big) \leq \mathrm{cst}(m, r) n^m \beta_0^n \Lambda_0^\epsilon \cdots \Lambda_i^\epsilon, \tag{3.28}$$

where we recall $\beta_0 = \beta_0(d, m)$ and $\Lambda_0^\epsilon = 1$ by definition. Now, we have for all $i = 1, \ldots, m$

$$\begin{aligned}
\Lambda_1^\epsilon \cdots \Lambda_i^\epsilon &= (\Lambda_1 + \epsilon) \cdots (\Lambda_i + \epsilon) \\
&= \sum_{h=0}^{i} \epsilon^{i-h} \sum_{0=j_0<j_1<\cdots<j_h \leq i} \Lambda_{j_0} \cdots \Lambda_{j_h} \\
&\leq \sum_{h=0}^{i} \epsilon^{i-h} \sum_{0=j_0<j_1<\cdots<j_h \leq i} \Lambda_0 \cdots \Lambda_h \\
&= \binom{i}{h} \sum_{h=0}^{i} \epsilon^{i-h} \Lambda_0 \cdots \Lambda_h \\
&\leq i! \sum_{h=0}^{i} \epsilon^{i-h} \Lambda_0 \cdots \Lambda_h.
\end{aligned} \tag{3.29}$$

Therefore, by (3.27), (3.28) and (3.29), for large n we have

$$\begin{aligned}
M(\epsilon, \tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r))) &\leq \mathrm{cst}(m, r) C(m) n^m \beta_0^n \sum_{i=0}^{m} \frac{i!}{\epsilon^i} \sum_{h=0}^{i} \epsilon^{i-h} \Lambda_0 \cdots \Lambda_h \\
&\leq \widetilde{\mathrm{cst}}(m, r) n^m \beta_0^n \sum_{h=0}^{\nu} \frac{\Lambda_0 \cdots \Lambda_h}{\epsilon^h},
\end{aligned}$$

where we note that, by definition, $\Lambda_h = 0$ for all $h > \nu$, and $\widetilde{\mathrm{cst}}(m, r)$ denotes a constant that depends only on m, r. From this we deduce that

$$\ln M(\epsilon, \tilde{f}_n(C^{\Lambda^\epsilon}(\tilde{f}_n) \cap B_{\mathbb{R}^n}(r))) \leq \ln(\widetilde{\mathrm{cst}}(m, r)) + n \ln \beta_0 + m \ln n + \ln\left(\sum_{h=0}^{\nu} \frac{\Lambda_0 \cdots \Lambda_h}{\epsilon^h}\right), \tag{3.30}$$

for all $\epsilon > 0$ and $n \in \mathbb{N}$ such that $cq^{-n} \leq \epsilon$ and $n \geq m$.

Step 6. We now show how to use (3.30) in order to get the estimate on $\dim_e(f(\mathrm{Crit}_\nu(f) \cap K))$. For $\epsilon > 0$ we choose $n = n_\epsilon$ where

$$n_\epsilon := \left\lceil \log_q \frac{c}{\epsilon} \right\rceil, \tag{3.31}$$

in such a way that (3.25), (3.26) and (3.30) hold when $\epsilon \to 0$. By (3.25) and (3.26) we have

$$\dim_e f(\mathrm{Crit}_\nu(f) \cap K) \leq \limsup_{\epsilon \to 0} \frac{\ln M(\epsilon, \tilde{f}_{n_\epsilon}(C^{\Lambda^\epsilon}(\tilde{f}_{n_\epsilon}) \cap B_{\mathbb{R}^{n_\epsilon}}(r)))}{\ln \frac{1}{2\epsilon}}.$$

Hence now we estimate the right-hand side through (3.30). As ϵ goes to zero we have

$$\limsup_{\epsilon \to 0} \frac{n_\epsilon \ln \beta_0}{\ln \frac{1}{2\epsilon}} = \lim_{\epsilon \to 0} \frac{\left\lceil \log_q \frac{c}{\epsilon} \right\rceil \ln \beta_0}{\ln \frac{1}{2\epsilon}} = \frac{\ln \beta_0}{\ln q}.$$

We also have, taking into account (3.31), that

$$\lim_{\epsilon \to 0} \frac{m \ln n_\epsilon}{\ln(\frac{1}{2\epsilon})} = 0.$$

Furthermore

$$\lim_{\epsilon \to 0} \frac{1}{\ln(\frac{1}{2\epsilon})} \ln \left(\sum_{h=0}^{\nu} \frac{\Lambda_0 \cdots \Lambda_h}{\epsilon^h} \right) = \nu.$$

Therefore,

$$\dim_e f(\mathrm{Crit}_\nu(f) \cap K) \leq \nu + \frac{\ln \beta_0}{\ln q},$$

concluding the proof. □

3.2.2 Sard-type theorems and Kolmogorov n-width

In this section, we recall the definition of Kolmogorov n–width, and we show its role in Sard-type theorems.

Definition 3.18 (Kolmogorov n–width)**.** Let X be a normed space and $S \subset X$ be a subset. For $n \in \mathbb{N}$, the *Kolmogorov n–width* of S in X is

$$\Omega_n(S, X) = \inf_{\dim E = n} \sup_{u \in S} \inf_{y \in E} \|u - y\|,$$

where the infimum is taken over all n-dimensional linear subspaces E of X.

The asymptotics of n–width measures in a quantitative way the compactness of a set, in fact the following holds (see [81, Prop. 1.2]).

Proposition 3.19. *S is compact if and only if S is bounded and* $\lim_{n\to\infty} \Omega_n(S) = 0$.

For various properties of the n–width we refer the reader to [81]; we recall here those that we will need in the sequel, see [81, Thm. 1.1].

Theorem 3.20. *Let X be a normed space and $S \subset X$. Then for all $n \in \mathbb{N}$:*

(i) $\Omega_n(S, X) = \Omega_n(\bar{S}, X)$*, where* $\bar{S}$ *is the closure of* S*;*

(ii) For every $\alpha \in \mathbb{R}$

$$\Omega_n(\alpha S, X) = |\alpha|\Omega_n(S, X);$$

(iii) Let $\mathrm{co}(S)$ *be the convex hull of* S*. Then*

$$\Omega_n(\mathrm{co}(S), X) = \Omega_n(S, X);$$

(iv) Let $\mathrm{b}(S) = \{\alpha x \mid x \in S,\ |\alpha| \leq 1\}$ *be the balanced hull of* S*. Then*

$$\Omega_n(\mathrm{b}(S), X) = \Omega_n(S, X).$$

Remark 3.21. As a consequence, in all our results, up to enlarging S, one can assume without loss of generality that S is convex and centrally symmetric, without changing its n–width.

We establish now a connection between n–width of compact sets and locally Lipschitz functions.

Lemma 3.22. *Let H be a Hilbert space, $(Y, \|\cdot\|_Y)$ a Banach space, and let $f : H \to Y$ be locally Lipschitz. Let $K \subset H$ be a compact subset. Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there exist a linear subspace $E_n \subset H$ of dimension n such that*

$$\sup_{x \in K} \|f(x) - f(\pi_{E_n}(x))\|_Y \leq c(f, K)\Omega_n(K, H).$$

Proof. Let $\{B_H(x_i, r/2)\}_{i \in I}$ a finite cover of K by balls, with centers $x_i \in K$ and radii $r/2 > 0$. We can assume that f is L-Lipschitz on each $B_H(x_i, r)$, for some $L > 0$. Since K is compact $\Omega_n(K, H) \to 0$ as $n \to \infty$, hence there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $\Omega_n(K, H) < r/4$. By definition of n–width, for all $n \geq n_0$ there exists a n-dimensional subspace $E_n \subset H$ such that

$$\sup_{x \in K} \|x - \pi_{E_n}(x)\| < 2\Omega_n(K, H) < \frac{r}{2}. \tag{3.32}$$

Let $x \in K$. Then by construction $x \in B_H(x_i, r/2)$ for some $i \in I$. By (3.32) we have $\pi_{E_n}(x) \in B_H(x_i, r)$. Since f is L-Lipschitz on every $B_H(x_i, r)$ we have

$$\sup_{x \in K} \|f(x) - f(\pi_{E_n}(x))\|_Y \leq L \sup_{x \in K} \|x - \pi_{E_n}(x)\| \leq 2L\Omega_n(K, H),$$

where we used again (3.32). This proves the result, with $c(f, K) = 2L$. □

We can now prove the following Sard-type theorem, corresponding to Theorem A in the Introduction.

Theorem 3.23. *Let $d, m \in \mathbb{N}$. There exists $\beta_0 = \beta_0(d,m) > 0$ such that the following holds. Let H be a Hilbert space, $f \in \mathscr{P}_d^m(H)$ and $K \subset H$ be a compact set such that*

$$\omega(K,H) = \limsup_{n\to\infty} \Omega_n(K,H)^{1/n} \leq q^{-1} \in (0,1).$$

Then, for every $\nu = 1,\ldots,m-1$ we have

$$\dim_e\Big(f(\operatorname{Crit}_\nu(f)\cap K)\Big) \leq \nu + \frac{\ln \beta_0}{\ln q}.$$

In particular, if $q > \beta_0$, then the Sard property holds on K:

$$\mu\Big(f(\operatorname{Crit}(f)\cap K)\Big) = 0.$$

Proof. Let $q_\epsilon \in (0,q)$. By assumption, $\Omega_n(K,H) \leq q_\epsilon^{-n}$ for sufficiently large n. We prove that f satisfies the hypothesis of Theorem E. Since $f \in \mathscr{P}_d^m(H)$ (see Definition 1), the map

$$(f, Df) : H \to \mathbb{R}^m \times \mathcal{L}(H,\mathbb{R}^m)$$

is locally Lipschitz. The codomain $Y = \mathbb{R}^m \times \mathcal{L}(H,\mathbb{R}^m)$ is a Banach space equipped with the norm $\|(v,A)\|_Y = \|v\| + \|A\|_{\mathrm{op}}$ for $v \in \mathbb{R}^m$ and $A \in \mathcal{L}(H,\mathbb{R}^m)$. Hence we can apply Lemma 3.22 to get $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there exists a n-dimensional linear subspace $E_n \subset H$ such that

$$\sup_{x\in K} \|f(x) - f(\pi_{E_n}(x))\| + \|D_x f - D_{\pi_{E_n(x)}} f\|_{\mathrm{op}} \leq c(f,K) q_\epsilon^{-n},$$

where $\pi_{E_n} : H \to E_n$ denotes the orthogonal projection. Let $f_n : E_n \to \mathbb{R}^m$ defined by the restriction $f_n := f|_{E_n}$, which are polynomial maps with degree uniformly bounded above by d. It holds

$$\|(D_x f)|_{E_n} - D_{\pi_{E_n}(x)} f_n\|_{\mathrm{op}} = \|(D_x f)|_{E_n} - (D_{\pi_{E_n}(x)} f)|_{E_n}\|_{\mathrm{op}} \leq \|D_x f - D_{\pi_{E_n}(x)} f\|_{\mathrm{op}}.$$

As a consequence of the above two estimates, assumption (3) of Theorem E holds, yielding

$$\dim_e\Big(f(\operatorname{Crit}_\nu(f)\cap K)\Big) \leq \nu + \frac{\ln \beta_0}{\ln q_\epsilon}.$$

Letting $q_\epsilon \uparrow q$ we obtain the thesis. □

3.2.3 Counterexamples to the Sard theorem in infinite dimension

In this section we provide examples of maps $f \in \mathscr{P}_d^m(H)$ as in Theorem A for which there exists a compact set $K \subset H$ such that the set $f(\operatorname{Crit}(f)\cap K)$ has not measure zero, however it holds $\Omega_n(K,H) \leq cq^{-n}$. This shows the necessity of the quantitative assumption $q > \beta_0(m,d)$. As a consequence, we get lower bounds on the semialgebraic constant β_0.

We start with the following example, which is a minor modification of Kupka's one [58].

Example 3.24 (Kupka revisited). Let $q > 1$. Let $f : \ell^2 \to \mathbb{R}$ defined by

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \phi(q^{k-1} x_k),$$

where ϕ is the Kupka polynomial as in [96, Sec. 10.2.3]. Namely ϕ has degree 3, with $\phi(0) = \phi'(0) = \phi(1) - 1 = \phi'(1) = 0$. If we assume $q^3/2 < 1$ then f is C^∞. All critical points have the form

$$\mathrm{Crit}(f) = \left\{ \left(x_1, \frac{x_2}{q}, \dots, \frac{x_k}{q^{k-1}}, \dots \right) \,\middle|\, x_i \in \{0, 1\} \right\},$$

and the set of critical values is $[0, 1]$. Furthermore, given $y = \left(x_1, \frac{x_2}{q}, \dots, \frac{x_k}{q^{k-1}}, \dots \right) \in \mathrm{Crit}(f)$, we consider its projection y_n on $\mathrm{span}\{e_1, \dots, e_n\}$, that is

$$y_n = \left(x_1, \frac{x_2}{q}, \dots, \frac{x_n}{q^{n-1}}, 0, 0, 0, \dots \right).$$

We have

$$\|y - y_n\|_{\ell^2}^2 \le \sum_{k=n+1}^{\infty} \frac{1}{q^{2(k-1)}} \le \int_n^{\infty} \frac{1}{q^{2t}} = \frac{q^{-2n}}{2 \ln q}.$$

Hence,

$$\Omega_n(\mathrm{Crit}(f), \ell^2) \le \sup_{y \in \mathrm{Crit}(f)} \|y - y_n\|_{\ell^2} \le \frac{q^{-n}}{\sqrt{2 \ln q}} \qquad \Longrightarrow \qquad \omega(\mathrm{Crit}(f), \ell^2) \le q^{-1}.$$

Since f and $\mathrm{Crit}(f)$ satisfy all the hypothesis of Theorem A (provided that $q^3/2 < 1$), we get a consequence that

$$\beta_0(1, 3) \ge 2^{1/3}.$$

We now get the lower bound on $\beta_0(1, d)$ for $d \ge 3$ by improving on the above construction.

Example 3.25 (Kupka revisited, degree $d \ge 3$). We need the following classical fact, which follows from [57]: for any $d \ge 2$ there exists a polynomial $\psi : \mathbb{R} \to \mathbb{R}$ of degree d such that

$$\psi(\mathrm{Crit}(\psi)) = \{0, \dots, d - 2\}. \tag{3.33}$$

(See also [56, Thm. 1] which gives a version of this result taking into account also possible multiplicities.) We fix $d \ge 3$ and we consider a polynomial ψ of degree d such that (3.33) holds. For $q > 1$ we consider the map $f_d : \ell^2 \to \mathbb{R}$

$$f_d(x) := \sum_{k=1}^{\infty} \frac{1}{(d-1)^k} \psi(q^{k-1} x_k).$$

If $q \in (1, (d - 1)^{1/d})$, then the map f_d is C^∞. The set of critical points of f_d is

$$\mathrm{Crit}(f_d) = \left\{ \left(y_1, \frac{y_2}{q}, \dots, \frac{y_k}{q^{k-1}}, \dots \right) \in \ell^2 \,\middle|\, \psi'(y_k) = 0,\ \forall\, k \in \mathbb{N} \right\}.$$

Hence we have

$$f_d(\mathrm{Crit}(f_d)) = \left\{ \sum_{k=1}^{\infty} \frac{1}{(d-1)^k} c_k \,\middle|\, c_k \in \{0, \dots, d-2\} \right\} = [0,1].$$

Let $\zeta = \max\{|z| \mid \psi'(z) = 0\}$. Exactly as in Example 3.24 we get for every $n \in \mathbb{N}$

$$\Omega_n(\mathrm{Crit}(f_d), \ell^2) \le \frac{\zeta}{\sqrt{2\ln q}} q^{-n} \qquad \Longrightarrow \qquad \omega(\mathrm{Crit}(f_d), \ell^2) \le q^{-1}.$$

For any $q \in (1, (d-1)^{1/d})$ the function f_d and the set $K = \mathrm{Crit}(f_d)$ satisfy the hypothesis of Theorem 3.23, but $f_d(\mathrm{Crit}(f_d)) = [0,1]$. It follows that

$$\beta_0(1,d) \ge (d-1)^{1/d}, \qquad \forall\, d \ge 3.$$

In the next example we extend the construction to arbitrary dimension of the codomain.

Example 3.26 (Kupka revisited, $d \ge 3$, arbitrary codomain). Let f_d be as in Example 3.25, and consider $g_d : \mathbb{R}^{m-1} \times \ell^2 \to \mathbb{R}^m$ defined as

$$g_d(x,y) := (x, f_d(y)).$$

We have

$$\mathrm{Crit}(g_d) = \mathbb{R}^{m-1} \times \mathrm{Crit}(f_d).$$

Taking $E_{n+m-1} = \mathrm{span}\{(e_i, 0), (0, e_j) \mid i = 1, \dots, m-1,\ j = 1, \dots, n\} \subset \mathbb{R}^{m-1} \times \ell^2$, for all $n \in \mathbb{N}$, we can estimate the n–width of $\mathrm{Crit}(g_d)$ as

$$\Omega_{n+m-1}(\mathrm{Crit}(g_d), \mathbb{R}^{m-1} \times \ell^2) \le \frac{\zeta}{\sqrt{2\ln q}} q^{-n} \qquad \Longrightarrow \qquad \omega(\mathrm{Crit}(g_d), \mathbb{R}^{m-1} \times \ell^2) \le q^{-1}.$$

We observe, provided that $q \in (1, (d-1)^{1/d})$, the function g_d satisfies all the hypothesis of Theorem 3.23, but the set $g_d(\mathrm{Crit}(g_d)) = \mathbb{R}^{m-1} \times [0,1]$ has not measure zero. Hence

$$\beta_0(d,m) \ge (d-1)^{1/d}, \qquad \forall\, m \in \mathbb{N},\ d \ge 3.$$

We collect the content of Examples 3.24 to 3.26 in a unified statement (see Theorem B).

Theorem 3.27. *Let $d, m \in \mathbb{N}$, with $d \ge 3$, and set $q := (d-1)^{1/d}$. There exist a Hilbert space H, and $f \in \mathscr{P}^m_d(H)$ such that $K = \mathrm{Crit}(f) \cap B_H(r)$ is compact for all $r > 0$, with*

$$\omega(K, H) = \limsup_{n\to\infty} \Omega_n(K,H)^{1/n} \le q^{-1},$$

and $f : H \to \mathbb{R}^m$ does not verify the Sard property, namely $\mu(f(\mathrm{Crit}(f) \cap K)) > 0$. Therefore, the semialgebraic constant $\beta_0(d,m)$ of Theorem E satisfies

$$\beta_0(d,m) \ge (d-1)^{1/d}, \qquad \forall\, m \in \mathbb{N},\ d \ge 3.$$

In all the examples above we have taken K to be set of all critical points of a given function in a ball. We now show a different construction, where the set K is strictly contained in the set of critical points, more precisely is the set of points where the differential has rank zero.

Example 3.28 (Rank zero counterexample to Sard). We consider $f_d : \ell^2 \to \mathbb{R}$ as in Example 3.25, and $m \in \mathbb{N}$. We define $h : (\ell^2)^m \to \mathbb{R}^m$ as

$$h(x^1, \dots, x^m) := (f_d(x^1), \dots, f_d(x^m)).$$

For any $x = (x^1, \dots, x^m) \in (\ell^2)^m$ we have

$$\operatorname{im}(D_x h) = \operatorname{im}(D_{x^1} f_d) \times \cdots \times \operatorname{im}(D_{x^m} f_d).$$

Hence $x \in \operatorname{Crit}_\nu(h)$ (the set of points where the rank of the differential is $\leq \nu$) if and only if at least $m - \nu$ components of x are critical points of f_d. In particular, we consider the compact set K of all critical points of h of rank zero, namely

$$K := \operatorname{Crit}_0(h) = \operatorname{Crit}(f_d)^m.$$

We have $K \subsetneq \operatorname{Crit}(h)$, and

$$h(\operatorname{Crit}(h) \cap K) = h(\operatorname{Crit}_0(h)) = [0, 1]^m.$$

As in the previous examples, K is compact with exponential n–width. By taking finite-dimensional spaces $E_{nm} = \operatorname{span}\{\underbrace{(0, \dots, e_i, \dots, 0)}_{j\text{-th component}} \in (\ell^2)^m \mid i = 1, \dots, n,\ j = 1, \dots, m\}$, for $n \in \mathbb{N}$ we have

$$\Omega_{nm}(K, (\ell^2)^m) \leq m\Omega_n(\operatorname{Crit}(f_d), \ell^2) \leq \frac{m\zeta}{\sqrt{2 \ln q}} q^{-n} \qquad \Longrightarrow \qquad \omega(K, (\ell^2)^m) \leq q^{-1/m},$$

where ζ is the same constant appearing in Example 3.25.

3.2.4 Sard threshold theorems

We can now prove Theorem C, of which we recall the statement.

Theorem 3.29. *For every $d, m \in \mathbb{N}$, there exists $\omega_0(d, m) \in (0, 1]$ such that*

(i) for every $f \in \mathscr{P}_d^m(H)$ and for every compact set $K \subset H$ with $\omega(K, H) < \omega_0(d, m)$,

$$\mu\Big(f(\operatorname{crit}(f) \cap K)\Big) = 0;$$

(ii) for every $\omega > \omega_0(d, m)$, with $\omega \in (0, 1]$, there exist $f \in \mathscr{P}_d^m(H)$ and a compact set $K \subset H$ with $\omega(K, H) = \omega$ and such that

$$\mu\Big(f(\operatorname{crit}(f) \cap K)\Big) > 0.$$

Proof. By definition, for any compact set K, it holds $\omega(K, H) \in [0, 1]$, see (2). We define $\omega_0(d, m)$ as the supremum of the set of $\omega \in [0, 1]$ such that for every $f \in \mathscr{P}_d^m(H)$ and for every compact set $K \subset H$ with $\omega(K, H) = \omega$ it holds:

$$\mu\Big(f(\operatorname{Crit}(f) \cap K)\Big) = 0.$$

By Theorem A this set is non-empty and $\omega_0(d, m) \geq \beta_0(d, m)^{-1}$. Hence $\omega_0(d, m) \in (0, 1]$. Both items of Theorem 3.29 follow. Note that by Theorem B (more specifically, the construction of Example 3.26), provided that $d \geq 3$, one must have $\omega_0(d, m) \leq (d - 1)^{-1/d} < 1$, so that Item (ii) is non-vacuous. □

Next, we prove Theorem D, of which we recall the statement.

Theorem 3.30. *Let $f \in \mathscr{P}^m_d(H)$ and $K \subset H$ be a compact set such that $\omega(K, H) < \omega_0(d, m)$. Consider the linear subspace*

$$V := \operatorname{span}(K).$$

Then the restriction $f|_V : V \to \mathbb{R}^m$ satisfies the Sard property:

$$\mu\Big(f(\operatorname{Crit}(f|_V))\Big) = 0.$$

In particular, $\mu(f(\operatorname{Crit}(f) \cap V)) = 0$.

Proof. Let $\tilde{H}$ denote the closure of V in H. Observe first that

$$\Omega_n(K, H) = \Omega_n(K, \tilde{H}).$$

The inequality $\Omega_n(K, H) \leq \Omega_n(K, \tilde{H})$ is clear since every n-dimensional subspace of $\tilde{H}$ is also a subspace of H. For the other inequality, since $\tilde{H}$ is closed in H, we can write $H = \tilde{H} \oplus \tilde{H}^\perp$ and denote by $\pi : H \to \tilde{H}$ the orthogonal projection. Then, using that $K \subset V \subset \tilde{H}$, we obtain

$$\begin{aligned}
\Omega_n(K, H) &= \inf_{\substack{Y \subset H \\ \dim Y = n}} \sup_{u \in K} \inf_{y \in Y} \|u - y\| \geq \inf_{\substack{Y \subset H \\ \dim Y = n}} \sup_{u \in K} \inf_{y \in Y} \|u - \pi(y)\| \\
&= \inf_{\substack{Y \subset H \\ \dim Y = n}} \sup_{u \in K} \inf_{z \in \pi(Y)} \|u - z\| \\
&= \inf_{\substack{Z \subset \tilde{H} \\ \dim Z \leq n}} \sup_{u \in K} \inf_{z \in Z} \|u - z\| \\
&= \inf_{\substack{Z \subset \tilde{H} \\ \dim Z = n}} \sup_{u \in K} \inf_{z \in Z} \|u - z\| = \Omega_n(K, \tilde{H}).
\end{aligned}$$

We can apply Item (i) of Theorem C to the map $\tilde{f} \in \mathscr{P}^m_d(\tilde{H})$, defined by $\tilde{f} := f|_{\tilde{H}}$, with $f \in \mathscr{P}^m_d(H)$, and the compact set $K \subset \tilde{H}$, which satisfies $\omega(K, \tilde{H}) < \omega_0(d, m)$, obtaining

$$\mu\Big(\tilde{f}(\operatorname{Crit}(\tilde{f}) \cap K)\Big) = 0.$$

Observe now that if u is a critical point for $g := f|_V$, then u is also critical for $\tilde{f} = f|_{\tilde{H}}$. In fact, $D_u g = (D_u f)|_V$ for all $u \in V$. Thus we have $\lambda \in (\mathbb{R}^m)^* \setminus \{0\}$ such that $0 \equiv \lambda \circ D_u g = (\lambda \circ D_u \tilde{f})|_V$ and, since V is dense in $\tilde{H}$, it follows that $\lambda \circ D_u \tilde{f} \equiv 0$. This means that

$$\operatorname{Crit}(g) \cap K \subseteq \operatorname{Crit}(\tilde{f}) \cap K.$$

In particular, since on $K \subset V$ we have $\tilde{f} = g$ by definition, this implies that

$$0 = \mu\Big(\tilde{f}(\operatorname{Crit}(\tilde{f}) \cap K)\Big) = \mu\Big(g(\operatorname{Crit}(g) \cap K)\Big). \tag{3.34}$$

Let K' be the balanced and convex hull of K. It holds:

$$V = \operatorname{span}(K) = \operatorname{span}(K') = \bigcup_{j=1}^{\infty} jK'.$$

By Theorem 3.20 the n–width of the compact, convex, centrally symmetric sets jK' satisfy

$$\Omega_n(jK', H) = j\Omega_n(K', H) = j\Omega_n(K, H).$$

Thus $\omega(jK', H) = \omega(K, H)$ for all $j \in \mathbb{N}$. Then, (3.34) holds also for jK' in place of K. Therefore,

$$\mu\Big(g(\mathrm{Crit}(g))\Big) \leq \sum_{j\geq 1} \mu\Big(\tilde{f}\left(\mathrm{Crit}(\tilde{f}) \cap jK'\right)\Big) = 0.$$

Finally, since $\mathrm{Crit}(f) \cap V \subset \mathrm{Crit}(g)$, and $g = f|_V$, it also holds $\mu(f(\mathrm{Crit}(f) \cap V)) = 0$. □

3.2.5 A class of maps with the global Sard property

Using Theorem E, we study a class of maps from a Hilbert space to $\mathbb{R}^m$ for which the Sard property holds true. The case $m = 1$ is not new, and it was proved in [95], see also [96]. We need the following preliminary facts.

Lemma 3.31 (Markov inequality)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d. Then for every $r > 0$ it holds*

$$\sup_{x\in B_{\mathbb{R}^n}(r)} \|D_x p\|_{\mathrm{op}} \leq \frac{\sqrt{m}d^2}{r} \sup_{x\in B_{\mathbb{R}^n}(r)} \|p(x)\|.$$

Proof. From [52, Thm. VI] we get the thesis for $r = 1$. Then, we conclude considering the rescaled polynomial $x \to p(rx)$. □

Lemma 3.32 (Estimates on balls at different radii)**.** *Let $p : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map with components of degree at most d. Then for all $r > 0$ it holds*

$$\sup_{x\in B_{\mathbb{R}^n}(r)} \|p(x)\| \leq \alpha(d, m)n^d(1 + r)^d \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(x)\|,$$

where $\alpha(d, m) = m^{d/2}d^{2d}(d + 1)$ is a constant depending only on d and m.

Proof. We can write the polynomial map as

$$p(x) = \sum_{|\alpha|\leq d} c_\alpha x^\alpha,$$

where $c_\alpha \in \mathbb{R}^m$, and the sum is over the multi-indices $\alpha \in \mathbb{N}^n$ with $|\alpha| = \sum_i \alpha_i \leq d$. Since $|\alpha|!c_\alpha = \partial_{x^\alpha} p|_{x=0}$, we can iterate Markov inequality (for $r = 1$, see Lemma 3.31) and obtain

$$\|c_\alpha\| \leq \frac{1}{|\alpha|!} m^{|\alpha|/2} d^{2|\alpha|} \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(x)\| \leq m^{|\alpha|/2} d^{2|\alpha|} \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(x)\|.$$

Therefore, for $r > 1$, we have

$$\begin{aligned}\sup_{x\in B_{\mathbb{R}^n}(r)} \|p(x)\| &= \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(rx)\| \\ &= \sup_{x\in B_{\mathbb{R}^n}(1)} \left\| \sum_{i=1}^{d} r^i \left(\sum_{|\alpha|=i} c_\alpha x^\alpha \right) \right\| \\ &\leq (1+r)^d \left(\sum_{|\alpha|\leq d} m^{|\alpha|/2} d^{2|\alpha|} \right) \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(x)\| \\ &\leq (1+r)^d m^{d/2} d^{2d} (d+1) n^d \sup_{x\in B_{\mathbb{R}^n}(1)} \|p(x)\|,\end{aligned}$$

where we used the fact that the number of multi-indices $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq d$ is equal to $(1 + n + n^2 + \cdots + n^d) \leq (d+1)n^d$ □

For the rest of this section, we assume that the Hilbert space H is separable, and we fix a Hilbert basis $\{e_i\}_{i\in\mathbb{N}} \subset H$. For $k \in \mathbb{N}$, we set $E_k := \mathrm{span}\{e_1, \ldots, e_k\}$ and we denote by $\pi_k : H \to E_k$ the corresponding orthogonal projection, that is if $x = \sum_i x_i e_i$, then $\pi_k(x) = (x_1, \ldots, x_k)$. We now define a special class of maps in this setting.

Proposition 3.33 (Construction of special maps)**.** *Let H be a separable Hilbert space. For all $k \in \mathbb{N}$, let $p_k : E_k \to \mathbb{R}^m$ be polynomial maps with $\sup_{k\in\mathbb{N}} \deg p_k \leq d$ for some $d \in \mathbb{N}$, and such that*

$$\sum_{k=1}^{\infty} \sup_{x\in B_{E_k}(r)} \|p_k(x)\| < \infty, \qquad \forall\, r > 0. \tag{3.35}$$

Then the map $f : H \to \mathbb{R}^m$ given by

$$f(x) := \sum_{k=1}^{\infty} p_k(x_1, \ldots, x_k), \qquad \forall\, x \in H, \tag{3.36}$$

is well–defined, $f \in \mathscr{P}^m_d(H)$ (see Definition 1) and its differential is given by

$$D_x f = \sum_{k=1}^{\infty} D_x(p_k \circ \pi_k). \tag{3.37}$$

Proof. For a Banach spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, and a closed set $U \subseteq X$, we denote by $\mathcal{B}(U, Y)$ the space of all continuous and bounded functions $f : U \to Y$. It is a Banach space when endowed with the norm $\sup_{x\in U} \|f(x)\|_Y$.

We first note that the assumption (3.35) implies that the sequence of maps $f_n : H \to \mathbb{R}^m$ given by $f_n := \sum_{k=1}^{n} p_k \circ \pi_k$ is Cauchy in $\mathcal{B}(B_H(r), \mathbb{R}^m)$ for all $r > 0$. In particular the right hand side of (3.36) converges locally uniformly. For any $n \in N$ we have

$$D_x f_n = \sum_{k=1}^{n} D_x(p_k \circ \pi_k). \tag{3.38}$$

Furthermore, $x \mapsto D_x f_n \in \mathcal{B}(B_H(r); \mathcal{L}(H, \mathbb{R}^m))$ for all $n \in \mathbb{N}$, $r > 0$. We now prove that it is a Cauchy sequence. Indeed, using Lemma 3.31, for any $n_2 \geq n_1$

$$\sup_{x \in B_H(r)} \left\| D_x f_{n_2} - D_x f_{n_1} \right\|_{\text{op}} \leq \frac{\sqrt{m} d^2}{r} \sum_{k > n_1} \sup_{x \in B_{E_k}(r)} \|p_k(x)\| \to 0,$$

as $n_2, n_1 \to \infty$. Therefore, $\{Df_n\}_{n \in \mathbb{N}}$ has a limit in $\mathcal{B}(B_H(r), \mathcal{L}(H, \mathbb{R}^m))$, for all $r > 0$. It follows that f is C^1 and (3.37) holds, see e.g. [59, Thm. 9.1].

Thanks to (3.37), we prove that $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is weakly continuous. Let us consider a sequence $\{x^j\}_{j \in \mathbb{N}}$ weakly convergent to $x \in B_H(r)$; we prove that $D_{x^j} f \to D_x f$ as $j \to \infty$. Indeed, given $\epsilon > 0$ we fix $N_\epsilon \in \mathbb{N}$ such that $\sum_{k > N_\epsilon} \sup_{B_{E_k}(r)} \|p_k\| \leq \epsilon$ and, by Lemma 3.31, we get

$$\begin{aligned}
\left\| D_{x^j} f - D_x f \right\|_{\text{op}} &\leq \sum_{k=1}^{\infty} \left\| D_{x^j}(p_k \circ \pi_k) - D_x(p_k \circ \pi_k) \right\|_{\text{op}} \\
&= \sum_{k=1}^{\infty} \left\| (D_{\pi_k(x^j)} p_k) \circ \pi_k - (D_{\pi_k(x)} p_k) \circ \pi_k \right\|_{\text{op}} \\
&= \sum_{k=1}^{\infty} \left\| D_{\pi_k(x^j)} p_k - D_{\pi_k(x)} p_k \right\|_{\text{op}} \\
&\leq \sum_{k=1}^{N_\epsilon} \left\| D_{\pi_k(x^j)} p_k - D_{\pi_k(x)} p_k \right\|_{\text{op}} + \sum_{k > N_\epsilon} \left\| D_{\pi_k(x^j)} p_k \right\|_{\text{op}} + \sum_{k > N_\epsilon} \left\| D_{\pi_k(x)} p_k \right\|_{\text{op}} \\
&\leq \sum_{k=1}^{N_\epsilon} \left\| D_{\pi_k(x^j)} p_k - D_{\pi_k(x)} p_k \right\|_{\text{op}} + \frac{\sqrt{m} d^2}{r} 2\epsilon,
\end{aligned}$$

where with some abuse of notation we denoted by the same symbol $\|\cdot\|_{\text{op}}$ either the norm in $\mathcal{L}(H, \mathbb{R}^m)$ or the one for the finite-dimensional subspaces $\mathcal{L}(E_k, \mathbb{R}^m)$. Since $\pi_k(x^j) \to \pi_k(x)$ as $j \to \infty$ for all $k \in \mathbb{N}$, we conclude by the continuity of Dp_k on E_k.

We prove now that $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is locally Lipschitz. We will prove that the sequence $\{Df_n\}_{n \in \mathbb{N}}$ defined by (3.38) is uniformly Lipschitz on balls. For $k \in \mathbb{N}$, fix $z, w \in E_k$. Then it holds

$$D_z p_k - D_w p_k = D_z p_k - D_z \tilde{p}_k = D_z(p_k - \tilde{p}_k), \tag{3.39}$$

where $\tilde{p}_k : E_k \to \mathbb{R}^m$, defined by $\tilde{p}_k(\cdot) := p_k(\cdot - z + w)$, is polynomial depending on the fixed choices of

z, w, with degree bounded by d. Thus for all fixed $z, w \in B_{E_k}(r)$, $r > 0$, it holds

$$\begin{aligned}
\|D_z p_k - D_w p_k\|_{\mathrm{op}} &\le \sup_{a \in B_{E_k}(r)} \|D_a(p_k - \tilde{p}_k)\|_{\mathrm{op}} && \text{by (3.39)} \\
&\le \frac{\sqrt{m} d^2}{r} \sup_{a \in B_{E_k}(r)} \|p_k(a) - \tilde{p}_k(a)\| && \text{by Lemma 3.31} \\
&\le \frac{\sqrt{m} d^2}{r} \sup_{a \in B_{E_k}(r)} \|p_k(a) - p_k(a - z + w)\| && \text{by definition of } \tilde{p}_k \\
&\le \left(\frac{\sqrt{m} d^2}{r} \sup_{a \in B_{E_k}(3r)} \|D_a p_k\|_{\mathrm{op}} \right) \|z - w\| && p_k \text{ is Lipschitz} \\
&\le \left(\frac{m d^4}{r^2} \sup_{a \in B_{E_k}(3r)} \|p_k(a)\| \right) \|z - w\|. && \text{by Lemma 3.31}
\end{aligned} \tag{3.40}$$

Note that if $x \in B_H(r)$ then $\pi_k(x) \in B_{E_k}(r)$. Thus, for all $x, y \in B_H(r)$ it holds

$$\begin{aligned}
\|D_x f_n - D_y f_n\|_{\mathrm{op}} &\le \sum_{k=1}^{n} \|D_x(p_k \circ \pi_k) - D_y(p_k \circ \pi_k)\|_{\mathrm{op}} && \text{by (3.38)} \\
&\le \sum_{k=1}^{n} \|D_{\pi_k(x)} p_k - D_{\pi_k(y)} p_k\|_{\mathrm{op}} && \text{by Leibniz rule} \\
&\le \frac{m d^4}{r^2} \sum_{k=1}^{n} \sup_{a \in B_{E_k}(3r)} \|p_k(a)\| \|\pi_k(x) - \pi_k(y)\| && \text{by (3.40)} \\
&\le \frac{m d^4}{r^2} \|x - y\| \sum_{k=1}^{\infty} \sup_{a \in B_{E_k}(3r)} \|p_k(a)\| && \pi_k \text{ is 1-Lipschitz} \\
&\le c(r) \|x - y\|,
\end{aligned}$$

where $c(r) > 0$ does not depend on n, and in the last inequality we used (3.35) for the convergence of the series. Therefore the sequence $\{D f_n\}_{n \in \mathbb{N}}$, is uniformly Lipschitz on every ball $B_H(r)$. Thus its limit $Df : H \to \mathcal{L}(H, \mathbb{R}^m)$ is Lipschitz on any ball $B_H(r)$, and a fortiori locally Lipschitz.

Finally, we show that for every finite dimensional space $E \subset H$, the map $f|_E$ is a polynomial map with components of degree bounded by d. Let $n = \dim(E)$ and fix a linear isometry $L : \mathbb{R}^n \to E$. The statement is equivalent to show that $f \circ L$ is a polynomial map with components of degree bounded by d. Denoting by $(t_1, \ldots, t_n)$ the standard coordinates in $\mathbb{R}^n$, we have

$$\begin{aligned}
f(L(t_1, \ldots, t_n)) &= \sum_{k=1}^{\infty} p_k(x_1(t_1, \ldots, t_n), \ldots, x_k(t_1, \ldots, t_n)) \\
&= \sum_{k=1}^{\infty} \tilde{p}_k(t_1, \ldots, t_n),
\end{aligned}$$

where each $\tilde{p}_k$ is a polynomial in $(t_1, \ldots, t_n)$ of degree at most d (since the x_i depend linearly on the t_j). Condition (3.35) guarantees that, for each monomial $t_1^{\alpha_1} \cdots t_n^{\alpha_n}$, its coefficient in the expansion $\sum_{k \le s} \tilde{p}_k$

converges as $s \to \infty$. Note that there are only a finite number of such coefficients, determined by the upper bound d on the degree and the dimension n. We conclude by noting that the uniform limit of polynomials with degree bounded by d is a polynomial with degree bounded by d. □

With the following we exhibit a special class of maps, obtained via the construction of Proposition 3.33, for which the Sard property holds true *globally*. This is the content of Theorem F in the introduction, of which we recall the statement here.

Theorem 3.34. *Let H be a separable Hilbert space. For all $k \in \mathbb{N}$, let $p_k : E_k \to \mathbb{R}^m$ be polynomial maps with $\sup_{k\in\mathbb{N}} \deg p_k \leq d$ for some $d \in \mathbb{N}$, and such that*

$$\sup_{x\in B_{E_k}(1)} \|p_k(x)\| \leq q^{-k}, \qquad \forall\, k \in \mathbb{N}, \tag{3.41}$$

for some $q > 1$. Then the map $f : H \to \mathbb{R}^m$ defined by

$$f(x) := \sum_{k=1}^{\infty} p_k(x_1, \dots, x_k), \qquad \forall\, x \in H,$$

is well–defined, $f \in \mathscr{P}^m_d(H)$ (see Definition 1), and for all $\nu = 1, \dots, m-1$ and $r > 0$ it holds

$$\dim_e\Big(f\,(\mathrm{Crit}_\nu(f) \cap B_H(r))\Big) \leq \nu + \frac{\ln \beta_0}{\ln q},$$

where $\beta_0 = \beta_0(d, m)$ is the same constant given by Theorem E. In particular, if $q > \beta_0$, then f satisfies the Sard property globally on H:

$$\mu\Big(f(\,\mathrm{Crit}(f))\Big) = 0.$$

Proof. We show first that the map f is well–defined. The upper bound (3.41), together with Lemma 3.32 imply that for all $r > 0$

$$\sup_{x\in B_{E_k}(r)} \|p_k(x)\| \leq \alpha(d, m)(1 + r)^d \frac{k^d}{q^k}, \qquad \forall\, k \in \mathbb{N}, \tag{3.42}$$

with $q > 1$. In turn, this implies that

$$\sum_{k=1}^{\infty} \sup_{x\in B_{E_k}(r)} \|p_k(x)\| < \infty, \qquad \forall\, r > 0.$$

We can then apply Proposition 3.33 obtaining that $f \in \mathscr{P}^m_d(H)$.

Fix $r > 0$. We now prove that f satisfies the hypothesis of Theorem E with $K = B_H(r)$ and polynomial approximating maps $f_n : E_n \to \mathbb{R}^m$

$$f_n(x) := \sum_{k=1}^{n} p_k(x_1, \dots, x_k), \qquad \forall\, x \in H.$$

Note that, by our assumptions, $\sup_{n\in\mathbb{N}} \deg f_n \le d$. It remains to check the validity of (3). Firstly, we have for all $n \in \mathbb{N}$

$$\sup_K \|f(x) - f_n \circ \pi_{E_n}(x)\| \le \alpha(d,m)(1+r)^d \sum_{k=n+1}^{\infty} \frac{k^d}{q^k},$$

where we used (3.42).

We now estimate the derivatives. We have

$$\begin{aligned}
\sup_{x\in K} \left\|(D_x f)\,|_{E_n} - D_{\pi_{E_n}(x)} f_n\right\|_{\mathrm{op}} &\le \sup_{x\in K} \left\|D_x f - D_x(f_n \circ \pi_{E_n})\right\|_{\mathrm{op}} \\
&\le \sup_{x\in K} \sum_{k=n+1}^{\infty} \|D_x(p_k \circ \pi_k)\|_{\mathrm{op}} \qquad \text{by (3.37)} \\
&= \sup_{x\in K} \sum_{k=n+1}^{\infty} \left\|(D_{\pi_k(x)} p_k) \circ \pi_k\right\|_{\mathrm{op}} \\
&\le \sum_{k=n+1}^{\infty} \sup_{x\in B_{E_k}(r)} \|D_x p_k\|_{\mathrm{op}}\,.
\end{aligned}$$

Note that, in the last line, the supremum is taken on a finite-dimensional Euclidean ball. To proceed, by Lemma 3.31 and (3.42) we have

$$\sup_{x\in B_{E_k}(r)} \|D_x p_k\|_{\mathrm{op}} \le \frac{d^2\sqrt{m}}{r} \alpha(d,m)(1+r)^d \frac{k^d}{q^k}.$$

Therefore, continuing the previous estimate, we obtain

$$\sup_{x\in K} \left\|(D_x f)\,|_{E_n} - D_{\pi_{E_n}(x)} f_n\right\|_{\mathrm{op}} \le \frac{d^2\sqrt{m}}{r} \alpha(d,m)(1+r)^d \sum_{k=n+1}^{\infty} \frac{k^d}{q^k}.$$

We have therefore proved that for all $n \in \mathbb{N}$, and $r > 0$ it holds

$$\sup_{x\in K} \left(\|f(x) - f_n \circ \pi_{E_n}(x)\| + \|(D_x f)\,|_{E_n} - D_{\pi_{E_n}(x)} f_n\|_{\mathrm{op}}\right) \le \left(1 + \frac{d^2\sqrt{m}}{r}\right) \alpha(d,m)(1+r)^d \sum_{k=n+1}^{\infty} \frac{k^d}{q^k}.$$

To conclude, note that by elementary estimates it holds

$$\left(1 + \frac{d^2\sqrt{m}}{r}\right) \alpha(d,m)(1+r)^d \sum_{k=n+1}^{\infty} \frac{k^d}{q^k} \le c q^{-n},$$

for sufficiently large n, where $c \ge 1$ is a constant depending only on the fixed parameters r, d, m. Therefore, the main assumption (3) of Theorem E is satisfied, yielding

$$\dim_e\left(f(\mathrm{Crit}_\nu(f) \cap B_H(r))\right) \le \nu + \frac{\ln \beta_0}{\ln q}, \qquad \forall\, \nu = 1, \ldots, m-1.$$

In particular, if $q > \beta_0$ then f satisfies the Sard property on $B_H(r)$, namely

$$\mu\Big(f(\mathrm{Crit}(f) \cap B_H(r))\Big) = 0.$$

Since $r > 0$ is arbitrary, we obtain that $\mu(f(\mathrm{Crit}(f))) = 0$. □

Chapter 4

Applications to the Endpoint maps of Carnot groups

In this chapter we apply the results from the previous chapter to the study of the Sard property for Endpoint maps for Carnot groups, proving the statements of Section iii.3.

4.1 Carnot groups

An m-dimensional Carnot group of step $s \in \mathbb{N}$ is a connected and simply connected Lie group $(\mathbb{G}, \cdot)$ of dimension m, whose Lie algebra of left-invariant vector fields $\mathfrak{g}$ admits a stratification of step s, that is

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s, \tag{4.1}$$

where $\mathfrak{g}_i \neq \{0\}$, $\mathfrak{g}_{i+1} = [\mathfrak{g}_1, \mathfrak{g}_i]$ for all $i = 1, \dots, s-1$ and $[\mathfrak{g}_1, \mathfrak{g}_s] = \{0\}$. We set $k_i = \dim \mathfrak{g}_i$, $a_i = k_1 + \cdots + k_i$, and $k = k_1$, which is called the *rank* of the Carnot group. As usual, we identify elements of $\mathfrak{g}$ with vectors of $T_e\mathbb{G}$. Since $\mathfrak{g}$ is nilpotent and $\mathbb{G}$ is simply connected, the group exponential map $\exp_{\mathbb{G}} : \mathfrak{g} \to \mathbb{G}$ is a smooth diffeomorphism.

The stratification (4.1) allows the definition of a family of automorphisms $\delta_\lambda : \mathbb{G} \to \mathbb{G}$, for $\lambda > 0$, called *dilations* such that

$$\delta_\lambda(\exp_{\mathbb{G}}(v)) = \exp_{\mathbb{G}}(\lambda^i v), \qquad \forall\, v \in \mathfrak{g}_i, \quad i = 1, \dots, s.$$

Fix an *adapted basis* of $\mathfrak{g}$, namely a set of vectors $v_1, \dots, v_m \in \mathfrak{g}$ such that for all $i = 1, \dots, s$ the list of vectors $v_{k_{i-1}+1}, \dots, v_{k_i}$ is a basis of $\mathfrak{g}_i$, with the convention $k_0 = 0$. For any $j = 1, \dots, m$ the *weight* of v_j is the unique $w_j \in \{1, \dots, s\}$ such that $v_j \in \mathfrak{g}_{w_j}$. We define a diffeomorphism $\varphi_{\mathbb{G}} : \mathbb{R}^m \to \mathbb{G}$, by

$$\varphi_{\mathbb{G}}(x_1, \dots, x_m) := \exp_{\mathbb{G}}\left(\sum_{i=1}^m x_i v_i\right).$$

The map $\varphi_{\mathbb{G}}$ defines global coordinates on $\mathbb{G}$, called *exponential coordinates* (of first kind). For any $x \in \mathbb{G}$ we denote by $\tau_x : \mathbb{G} \to \mathbb{G}$ the left translation on $\mathbb{G}$, that is $\tau_x(y) = x \cdot y$. We denote by

$X_1, \ldots, X_m$ the left-invariant vector fields on $\mathbb{G}$ corresponding to the adapted basis. They are given by $X_j(x) = D_e\tau_x(v_j)$ for any $x \in \mathbb{G}$, where $D_e\tau_x$ denotes the differential of τ_x at the unit element.

We recall the definition of (homogeneous) polynomials on Carnot groups. We say that $P : \mathbb{G} \to \mathbb{R}$ is a polynomial if for some (and then any) set of exponential coordinates $\varphi_{\mathbb{G}}$, the map $P \circ \varphi_{\mathbb{G}} : \mathbb{R}^m \to \mathbb{R}$ is a polynomial. Any polynomial $P : \mathbb{G} \to \mathbb{R}$ can be written as

$$P(\varphi_{\mathbb{G}}(x_1, \ldots, x_m)) = \sum_{\alpha \in \mathbb{N}^m} c_\alpha x_1^{\alpha_1} \cdots x_m^{\alpha_m}, \qquad \forall\, x \in \mathbb{R}^m,$$

for some $c_\alpha \in \mathbb{R}$, which are non-zero only for a finite set of multi-indices. For $\alpha \in \mathbb{N}^m$, we denote by

$$|\alpha|_{\mathbb{G}} := \sum_{i=1}^m w_i \alpha_i,$$

the weighted degree of the multi-index α. The weighted degree of a polynomial P is

$$\deg_{\mathbb{G}}(P) := \max\{|\alpha|_{\mathbb{G}} \mid c_\alpha \neq 0\}.$$

A polynomial P is homogeneous of weighted degree w if $P(\delta_\lambda x) = \lambda^w P(x)$ for any $x \in \mathbb{G}$ and $\lambda > 0$.

Remark 4. In other words, P is homogeneous of weighted degree w if and only if

$$P(\varphi_{\mathbb{G}}(x_1, \ldots, x_m)) = \sum_{\substack{\alpha \in \mathbb{N}^m \\ |\alpha|_{\mathbb{G}} = w}} c_\alpha x_1^{\alpha_1} \cdots x_m^{\alpha_m}.$$

In particular any such polynomial depends only on the variables x_i with $w_i \leq w$, namely $x_1, \ldots, x_{k_w}$.

It is easy to see that the concepts we just introduced do not depend on the choice of exponential coordinates. For this reason, in the following, we fix such a choice, and we omit $\varphi_{\mathbb{G}}$ from the notation, identifying $\mathbb{G} \simeq \mathbb{R}^m$. Note that $e \in \mathbb{G}$ is identified with $0 \in \mathbb{R}^m$.

We recall from [22, Prop. 5.18] the following result on the structure of the vector fields X_j.

Proposition 4.1. *In exponential coordinates $(x_1, \ldots, x_m) \in \mathbb{R}^m$, the vector fields X_j, for $j = 1, \ldots, k$ have the following form:*

$$X_j(x_1, \ldots, x_m) = \partial_j + \sum_{h>k} Q_{jh}(x_1, \ldots, x_{k_{w_h-1}})\partial_h,$$

where $Q_{jh} : \mathbb{G} \simeq \mathbb{R}^m \to \mathbb{R}$ is a homogeneous polynomial of weighted degree $w_h - 1$.

4.2 Endpoint maps

In this section we introduce the Endpoint map associated to a Carnot group. We work with the separable Hilbert space

$$H := L^2(I, \mathbb{R}^k),$$

for some fixed interval I, say $I = [0, 1]$, and where k is the rank of $\mathbb{G}$. For any $u \in H$ with $u = (u_1, \ldots, u_k)$ we consider the following Cauchy problem for a curve $\gamma : I \to \mathbb{G}$:

$$\dot\gamma(t) = \sum_{j=1}^k u_j(t) X_j(\gamma(t)), \qquad \gamma(0) = e. \tag{4.2}$$

It is well–known that (4.2) admits a unique absolutely continuous solution, $\gamma_u : I \to \mathbb{G}$, for any $u \in H$.

Definition 4.2. Let $\mathbb{G}$ be a Carnot group. The *Endpoint map* of $\mathbb{G}$ is the function

$$\mathcal{E} : H \to \mathbb{G},$$

given by $\mathcal{E}(u) := \gamma_u(1)$. We call the function $u \in H$ a *control*. We use the same notation to denote the Endpoint map $\mathcal{E} : H \to \mathbb{R}^m \simeq \mathbb{G}$ with an identification in exponential coordinates.

Thanks to the special form of the vector fields X_j given by Proposition 4.1, we can rewrite the system (4.2) in exponential coordinates. Adopting the notation $\gamma = (\gamma_1, \dots, \gamma_m)$, we have

$$\begin{cases} \gamma_i(0) = 0 & i = 1, \dots, m, \\ \dot\gamma_i(t) = u_i(t) & i = 1, \dots, k, \\ \dot\gamma_i(t) = \displaystyle\sum_{j=1}^{k} u_j(t) Q_{ji}(\gamma_1(t), \dots, \gamma_{k_{w_i-1}}(t)) & i = k+1, \dots, m, \end{cases} \tag{4.3}$$

for a.e. $t \in I$. Similarly, we write $\gamma_u = (\gamma_{u,1}, \dots, \gamma_{u,m})$ for the solution to (4.3).

The next result (Theorem H in the Introduction) connects this framework to the previous sections.

Theorem 4.3. *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Then the Endpoint map $\mathcal{E} \in \mathscr{P}^m_s(H)$, for $H = L^2(I, \mathbb{R}^k)$.*

Proof. The fact that $\mathcal{E}$ is C^1 (in fact smooth) and that the differential $D\mathcal{E}$ is locally Lipschitz follows from [5, Prop. 8.5]. The fact that the $D\mathcal{E}$ is weakly continuous is proved in [28, Thm. 23]. It remains to show that for all finite dimensional linear subspaces E of H, the restriction $\mathcal{E}|_E : E \to \mathbb{R}^m$ is a polynomial map of degree at most s.

We actually prove a stronger claim: for any $n \in \mathbb{N}$, for any n-dimensional linear subspace $E \subset L^2(I, \mathbb{R}^k)$, for any $t \in I$ and $i = 1, \dots, m$, the function on E defined by $E \ni u \mapsto \gamma_{u,i}(t)$ is a polynomial of degree w_i, and its coefficients (when u is written in terms of some basis of $E \simeq \mathbb{R}^n$) are continuous functions of $t \in I$. We prove the claim by induction on w. We prove the base case, that is $w_i = 1$. In this case $i = 1, \dots, k$. For these values of i, by (4.3) we obtain

$$\gamma_{u,i}(t) = \int_0^t u_i(s)\, ds,$$

which is clearly a polynomial function of $u \in E$, of degree 1. Moreover its coefficients (when u is written in terms of a basis of $E \simeq \mathbb{R}^n$) are continuous functions of $t \in I$. This concludes the base case.

Let then $1 \leq \theta < s$. The induction assumption is that for any $t \in I$ the function $E \ni u \to \gamma_{u,i}(t)$ is a polynomial of degree w_i for all i such that $w_i \leq \theta$, with coefficients that depends continuously on t.

Now we prove that the above property holds for any i such that $w_i = \theta + 1$. In particular $i > k$. In this case, by (4.3), the curve $\gamma_{u,i} : I \to \mathbb{R}$ is determined by

$$\gamma_{u,i}(t) = \sum_{j=1}^{k} \int_0^t u_j(s) Q_{ji}(\gamma_{u,1}(s), \dots, \gamma_{u,k_{w_i-1}}(s))\, ds. \tag{4.4}$$

Recall that Q_{ji} is a homogeneous polynomial of weighted degree $w_i - 1$ and as already highlighted in the notation, depends only on the functions $\gamma_{u,\ell}(s)$ with $w_\ell \leq w_i - 1 = \theta$. By the induction assumption, all the functions $E \ni u \mapsto \gamma_{u,\ell}(s)$ are polynomials of degree $w_\ell \leq \theta$, depending continuously on s. By using the definition of weighted degree (see Remark 4), it follows that the function

$$E \ni u \mapsto Q_{ji}(\gamma_{u,1}(s), \dots, \gamma_{u,k_{w_i-1}}(s)),$$

is a polynomial of degree $w_i - 1$, with coefficients depending continuously on $s \in I$. Then, the integrand in (4.4), namely the function

$$E \ni u \mapsto \sum_{j=1}^{k} u_j(s) Q_{ji}(\gamma_{u,1}(s), \dots, \gamma_{u,k_{w_i-1}}(s)),$$

is a polynomial of degree $w_i = \theta + 1$, with coefficients depending continuously on $s \in I$. We conclude easily using (4.4). □

Thanks to Theorem H we can apply the results of Section 3.2 to Endpoint maps of Carnot groups, obtaining the following statement (see Theorem I in the Introduction).

Theorem 4.4. *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Let $K \subset L^2(I, \mathbb{R}^k)$ be a compact subset with*

$$\omega(K, L^2(I, \mathbb{R}^k)) < \omega_0(m, s).$$

Then, letting $V := \operatorname{span}(K)$, the restriction $\mathcal{E}|_V : V \to \mathbb{G}$ has the Sard property, namely

$$\mu\Big(\mathcal{E}(\operatorname{Crit}(\mathcal{E}|_V))\Big) = 0.$$

In particular $\mu(\mathcal{E}(\operatorname{Crit}(\mathcal{E}) \cap V)) = 0$.

Proof. Since $\mathcal{E} \in \mathscr{P}^d_m(H)$ by Theorem H, we can apply Theorem D. □

4.3 Sard property for real-analytic controls

We specialize the result in Theorem I to some concrete set of controls K, estimating its n–width.

Fix a closed interval $I \subset \mathbb{R}$ and let $\mathcal{D}(r, I) := \{z \in \mathbb{C} \mid d(z, I) \leq r\}$. The radius of convergence of a real-analytic function $u : I \to \mathbb{R}$ is the largest $r \in (0, \infty]$ such that u extends to a (unique) holomorphic function on $\mathcal{D}(r, I)$. For fixed $r > 0$ we denote by $C^\omega(I, \mathbb{R}^k; r) \subset L^2(I, \mathbb{R}^k)$ the set of real-analytic controls whose components have radius of convergence strictly greater than r:

$$C^\omega(I, \mathbb{R}^k; r) := \left\{ u \in L^2(I, \mathbb{R}^k) \,\middle|\, u \text{ is real-analytic with radius of convergence strictly greater than } r \right\},$$

which we endow with the supremum norm:

$$\|u\|_{C^\omega(I, \mathbb{R}^k; r)} := \sup_{z \in \mathcal{D}(r, I)} \|u(z)\|.$$

Similarly, we consider finite concatenations of real-analytic controls as above. More precisely, for $\ell \in \mathbb{N}$ divide the interval I in equispaced sub-intervals:

$$I = I_1 \cup \dots \cup I_\ell, \qquad I_a = \inf I + \left[\frac{(a-1)|I|}{\ell}, \frac{a|I|}{\ell} \right], \tag{4.5}$$

and we let $C^{\omega}(I,\mathbb{R}^k;r,\ell) \subset L^2(I,\mathbb{R}^k)$ be the set of piecewise analytic controls $u : I \to \mathbb{R}^k$ that are real-analytic on each set $I_1,\dots,I_\ell$, with radius of convergence strictly greater than r, namely

$$C^{\omega}(I,\mathbb{R}^k;r,\ell) = \left\{u \in L^2(I,\mathbb{R}^k) \,\middle|\, u|_{I_a} \in C^{\omega}(I_a,\mathbb{R}^k;r) \subset L^2(I_a,\mathbb{R}^k), \text{ for all } a = 1,\dots,\ell\right\},$$

endowed with the norm

$$\|u\|_{C^{\omega}(I,\mathbb{R}^k;r,\ell)} = \max_{a=1,\dots,\ell} \|u\|_{C^{\omega}(I_a,\mathbb{R}^k;r,\ell)}. \tag{4.6}$$

Theorem 4.5. *Fix $I = [0,1]$, $\ell \in \mathbb{N}$ and $r > 1$. Let $V = C^{\omega}(I,\mathbb{R}^k;r,\ell)$. Then the set $K_V := \{u \in V \mid \|u\|_V \le 1\}$ is compact in $L^2(I,\mathbb{R}^k)$, and for its n–width it holds:*

$$\Omega_n(K_V, L^2(I,\mathbb{R}^k)) \le \frac{(k\ell)^{1/2}}{\ln r}\left(\frac{1}{r}\right)^{\lfloor \frac{n}{k\ell} \rfloor}.$$

Proof. We start by proving the case $\ell = 1$. Fix $u = (u_1,\dots,u_k) \in K_V$. Since $r > 1$, any u_j for $j = 1,\dots,k$ has an expansion in power series on the complex unit ball centered at the origin, denoted by $B_{\mathbb{C}}(1)$. In particular, we have for any $t \in I$

$$u_j(t) = \sum_{h=0}^{\infty} \frac{u_j^{(h)}(0)}{h!} t^h.$$

For any $N \ge 1$ we consider the truncated sum, given for any $t \in I$ as

$$u_{j,N}(t) = \sum_{h=0}^{N-1} \frac{u_j^{(h)}(0)}{h!} t^h.$$

By the Cauchy integral formula, and since $r > 1$, we have

$$u_j^{(h)}(0) = \frac{h!}{2\pi i} \int_{\partial B_{\mathbb{C}}(r)} \frac{u_j(z)}{z^{h+1}}\, dz,$$

where $B_{\mathbb{C}}(r)$ is the complex unit ball of radius one centred at zero. Therefore,

$$|u_j^{(h)}(0)| \le \frac{h!}{r^h}.$$

Hence we have

$$\|u_j - u_{j,N}\|_{L^2(I)} \le \left\|u_j - u_{j,N}\right\|_{L^\infty(I)} \le \sum_{h=N}^{\infty} \frac{|u_j^{(h)}(0)|}{h!} \le \sum_{h=N}^{\infty} \frac{1}{r^h} \le \int_N^\infty \frac{1}{r^x}\, dx = \frac{r^{-N}}{\ln r}.$$

It follows that for all $u \in K_V$ it holds

$$\|u - u_N\|_{L^2(I,\mathbb{R}^k)} \le \left(\sum_{j=1}^{k} \|u_j - u_{j,N}\|_{L^2(I)}^2\right)^{1/2} \le \frac{k^{1/2} r^{-N}}{\ln r}. \tag{4.7}$$

Thanks to this result we can find an approximating subspace. Consider the standard basis $e_1, \dots, e_k$ of $\mathbb{R}^k$, and for any $k \in \mathbb{N}$ we denote by $f_h \in L^2(I)$ the function $f_h(t) = t^h$. We have

$$u_N = \sum_{h=0}^{N-1} \frac{u^{(h)}(0)}{h!} f_h = \sum_{j=1}^{k} \sum_{h=0}^{N-1} \frac{u_j^{(h)}(0)}{h!} f_h \otimes e_j.$$

In particular, $u_N \in V_N := \mathrm{span}\{f_h \otimes e_j \mid j = 1, \dots, k,\ h = 0, \dots, N-1\}$ and $\dim(V_N) = kN$. Hence from (4.7) it follows that for any $N \in \mathbb{N}$

$$\Omega_{kN}(K_V, L^2(I, \mathbb{R}^k)) \leq \frac{k^{1/2}}{\ln r} \left(\frac{1}{r}\right)^N. \tag{4.8}$$

For $n \in \mathbb{N}$ we consider $N = \left\lfloor \frac{n}{k} \right\rfloor$. By Definition 3.18, the n–width is non-increasing as a function of $n \in \mathbb{N}$, hence by (4.8) we have

$$\Omega_n(K_V, L^2(I, \mathbb{R}^k)) \leq \Omega_{kN}(K_V, L^2(I, \mathbb{R}^k)) \leq \frac{k^{1/2}}{\ln r} \left(\frac{1}{r}\right)^{\lfloor \frac{n}{k} \rfloor}.$$

This proves the estimate on the n–width per $\ell = 1$.

We sketch the argument for general ℓ. In this case, for any $u \in C^\omega(I, \mathbb{R}^k; r, \ell)$ we can write

$$u = \sum_{a=1}^{\ell} u|_{I_a} \mathbb{1}_{I_a},$$

where I_a are the intervals of the decomposition (4.5), and $\mathbb{1}_{I_a}$ are the corresponding characteristic functions. By definition, each $u|_{I_a}$ is the restriction to I_a of a real-analytic function with radius of convergence $r_a \geq r > 1$. We can expand each $u|_{I_a}$ in Taylor series centered at the lower bound of I_a. Then, in order to obtain the finite-dimensional approximation, we truncate the series as in the previous case, repeating analogous estimates for the remainder (taking into account the length of the intervals $|I_a| = 1/\ell$). This concludes the estimate on the n–width for general ℓ.

To prove the compactness assertion, observe that there is a linear immersion (given by the inclusion) $C^\omega(I, \mathbb{R}^k; r, \ell) \hookrightarrow L^2(I, \mathbb{R}^k)$ and it holds $\|u\|_{L^2(I,\mathbb{R}^k)} \leq \|u\|_{C^\omega(I,\mathbb{R}^k;r,\ell)}$ for all $u \in C^\omega(I, \mathbb{R}^k; r, \ell)$. In particular K_V is bounded in $L^2(I, \mathbb{R}^k)$ and since its n–width tends to zero as $n \to \infty$ it is also compact by Proposition 3.19. □

We can now prove Theorem J in the Introduction, of which we recall the statement.

Theorem 4.6. *Let $\mathbb{G}$ be a Carnot group of topological dimension m, step s, and rank k. Given $\ell \in \mathbb{N}$, there exists $r = r(m, s, k, \ell) > 0$ such that, letting $V = C^\omega(I, \mathbb{R}^k; r, \ell)$, with $I = [0, 1]$, it holds*

$$\mathcal{E}(V) = \mathbb{G} \qquad \textit{and} \qquad \mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}|_V)) = \mu(\mathcal{E}(\mathrm{Crit}(\mathcal{E}) \cap V)) = 0.$$

Namely, the Sard property holds on the space of piecewise real-analytic controls with radius of convergence $> r$ and with ℓ pieces.

Proof. The surjectivity of $\mathcal{E}$ on $C^\omega(I, \mathbb{R}^k; r, \ell)$, for any value of r, ℓ, is discussed in Section 4.4, where we prove in particular that $\mathcal{E}$ is surjective when restricted on polynomial controls (see Theorem 4.10). To prove the Sard property, observe that $\mathcal{E} \in \mathscr{P}_s^m(L^2(I, \mathbb{R}^k))$ by Theorem H. Let $V = C^\omega(I, \mathbb{R}^k; r, \ell)$, and K_V be its unit ball with respect to the norm (4.6). By Theorem 4.5, for any $r > 1$ we have

$$\omega(K_V, L^2(I, \mathbb{R}^k)) := \limsup_{n\to\infty} \Omega_n(K_V, L^2(I, \mathbb{R}^k))^{1/n} \le r^{-\frac{1}{k\ell}},$$

which is smaller than $\omega_0(s, m)$ if $r > 1$ is large enough. The result follows now from Theorem D, observing that $\operatorname{span}(K_V) = V$. □

4.4 Surjectivity of the Endpoint map on finite-dimensional spaces of controls

The Endpoint maps of sub-Riemannian manifolds are surjective when restricted to piecewise constant controls: this follows from the proof of the Rashevskii-Chow theorem in [5]. In this section, we prove that in Carnot groups the Endpoint maps are surjective also when they are restricted to the space of controls which consists in the set of polynomial maps of some large enough fixed degree (which depends on the Carnot group). The proof is obtained by using a quantitative version of the inverse function theorem, as it can be found in [35]. For completeness we present and prove in our setting the statements we need.

Given $M \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ we denote by $\sigma(M)$ the smallest singular value of M, namely

$$\sigma(M) := \inf_{\|v\|=1} \|Mv\|.$$

From the definition it follows that for all $M_1, M_2 \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ it holds

$$|\sigma(M_1) - \sigma(M_2)| \le \|M_1 - M_2\|_{\text{op}}. \tag{4.9}$$

From [35, Lemma 3] we obtain the following technical lemma.

Lemma 4.7. *Let $M_0 \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ be invertible. Then, for every $v \in \mathbb{R}^m$ with $\|v\| = 1$ there exists $w \in \mathbb{R}^m$ with $\|w\| = 1$ such that*

$$\langle w, Mv \rangle \ge \frac{\sigma(M_0)}{2},$$

for all M such that $\|M - M_0\|_{\text{op}} \le \frac{\sigma(M_0)}{2}$.

Proof. Given $v \in \mathbb{R}^m$ with $\|v\| = 1$, the set $C_v \subset \mathbb{R}^m$

$$C_v := \left\{ Mv \in \mathbb{R}^m \,\middle|\, \|M - M_0\|_{\text{op}} \le \frac{\sigma(M_0)}{2} \right\},$$

is convex. We now prove that its distance from 0 is at least $\frac{\sigma(M_0)}{2}$. Indeed, by (4.9), for any M with $\|M - M_0\|_{\text{op}} \le \frac{\sigma(M_0)}{2}$ we obtain $\sigma(M) > \frac{\sigma(M_0)}{2}$. Therefore,

$$\|Mv\| \ge \inf_{\|v\|=1} \|Mv\| = \sigma(M) > \frac{\sigma(M_0)}{2}.$$

Hence, we have proved that C_v is separated from the ball $B_{\mathbb{R}^m}\left(\frac{\sigma(M_0)}{2}\right)$. We conclude the proof thanks to the usual separation theorem for convex sets. Indeed, it directly provides $w \in \mathbb{R}^m$ with $\|w\| = 1$ such that for all M with $\|M - M_0\|_{\text{op}} \leq \frac{\sigma(M_0)}{2}$ we have

$$\langle w, Mv \rangle \geq \frac{\sigma(M_0)}{2},$$

concluding the proof. □

From [35, Lemma 4] we obtain the following "quantitative injectivity" lemma.

Lemma 4.8. *Let $W \subset \mathbb{R}^m$ be an open set, and $f \in C^1(W, \mathbb{R}^m)$. Assume $x_0 \in W$ is such that $D_{x_0} f$ is invertible. Set $\sigma_{f,x_0} := \sigma(D_{x_0} f) > 0$. There exists $r_{f,x_0} > 0$ such that for any $g \in C^1(W, \mathbb{R}^m)$ with*

$$\sup_{z \in W} \|D_z f - D_z g\|_{\text{op}} < \frac{\sigma_{f,x_0}}{4}, \tag{4.10}$$

and all $x, y \in B_{\mathbb{R}^m}(x_0, r_{f,x_0})$ it holds

$$\|g(x) - g(y)\| \geq \frac{\sigma_{f,x_0}}{2} \|x - y\|.$$

Proof. Since f is C^1 there exists $r = r_{f,x_0} > 0$ such that $B_{\mathbb{R}^m}(x_0, r) \subset W$ and

$$\sup_{z \in B_{\mathbb{R}^m}(x_0, r)} \left\| D_z f - D_{x_0} f \right\|_{\text{op}} < \frac{\sigma_{f,x_0}}{4}.$$

Let $g \in C^1(W, \mathbb{R}^m)$ as in the statement. By (4.10) we get

$$\left\| D_z g - D_{x_0} f \right\|_{\text{op}} < \frac{\sigma_{f,x_0}}{2}, \tag{4.11}$$

for all $z \in B_{\mathbb{R}^m}(x_0, r)$. Now we fix $x \neq y$ in $B_{\mathbb{R}^m}(x_0, r)$, and we have

$$g(x) - g(y) = \|x - y\| \int_0^1 D_{(1-t)y+tx} g \cdot \frac{x - y}{\|x - y\|} \, dt.$$

We define $v_{x,y} := \frac{x-y}{\|x-y\|} \in \mathbb{R}^m$, and we apply Lemma 4.7 to $v = v_{x,y}$ and $M_0 = D_{x_0} f$, which is invertible by hypothesis. We denote by $w_{x,y} \in \mathbb{R}^m$ the provided unit vector, such that

$$\left\langle w_{x,y}, D_z g \cdot \frac{x - y}{\|x - y\|} \right\rangle \geq \frac{\sigma_{f,x_0}}{2},$$

for all $z \in B_{\mathbb{R}^m}(x_0, r)$. We conclude the proof by the Cauchy-Schwarz inequality, which gives

$$\|g(x) - g(y)\| \geq \|x - y\| \left\langle w_{x,y}, \int_0^1 D_{(1-t)y+tx} g \cdot \frac{x - y}{\|x - y\|} \, dt \right\rangle \geq \frac{\sigma_{f,x_0}}{2} \|x - y\|,$$

concluding the proof. □

We conclude with the following inclusion, analogous to [35, Lemma 5], providing the "quantitative surjectivity" counterpart of the previous statement. We report the proof for completeness.

Lemma 4.9. *In the same setting of Lemma 4.8, the following inclusion holds*

$$g(B_{\mathbb{R}^m}(x_0, r_{f,x_0})) \supset B_{\mathbb{R}^m}\left(g(x_0), \frac{r_{f,x_0}\sigma_{f,x_0}}{8}\right).$$

Proof. Set $r = r_{f,x_0}$ and $\sigma = \sigma_{f,x_0}$. Let $y \in B_{\mathbb{R}^m}\left(g(x_0), \frac{r\sigma}{8}\right)$. Let x be a minimum of $\|y - g(\cdot)\|^2$ on $B_{\mathbb{R}^n}(x_0, r)$. We claim that $x \in \mathrm{int} B_{\mathbb{R}^n}(x_0, r)$. Otherwise, using Lemma 4.8, we have

$$\begin{aligned}
\frac{r\sigma}{8} &\geq \|y - g(x_0)\| \\
&\geq \|g(x) - g(x_0)\| - \|y - g(x)\| \\
&\geq \frac{\sigma}{2}\|x - x_0\| - \|y - g(x)\| \\
&\geq \frac{\sigma r}{2} - \|y - g(x_0)\| \\
&\geq \frac{\sigma r}{2} - \frac{\sigma r}{8} = \frac{3r\sigma}{8},
\end{aligned}$$

which is a contradiction. Thus x yields a local minimum for the function $\|y - g(\cdot)\|^2$, and consequently $\nabla_x\|y - g(\cdot)\|^2 = -2D_x g \cdot (y - g(x)) = 0$. Note that by (4.11) and the fact that the singular values are 1-Lipschitz (see Lemma 3.9) it follows that $\sigma(D_x g) > \frac{\sigma}{2}$, hence $D_x g$ is invertible. Hence $y = g(x)$. □

We can now state and prove the surjectivity property of the Endpoint maps of Carnot groups.

Theorem 4.10. *For every Carnot group $\mathbb{G}$ there exists $d_{\mathbb{G}} \in \mathbb{N}$ such that*

$$\mathcal{E}\left(\{u \in L^2(I, \mathbb{R}^k) \mid u \text{ is a polynomial map with components of degrees at most } d_{\mathbb{G}}\}\right) = \mathbb{G}.$$

Proof. The classical proof of the Rashevskii-Chow theorem provides $u_0 \in L^2(I, \mathbb{R}^k)$ such that $D_{u_0}\mathcal{E}$ is surjective and $\mathcal{E}(u_0) = 0$ (see e.g. [5, Sec. 3.2]). In particular there exist $w_1, \ldots, w_m \in \ker(D_{u_0}\mathcal{E})^\perp$ such that the vectors $D_{u_0}\mathcal{E}(w_1), \ldots, D_{u_0}\mathcal{E}(w_m)$ are linearly independent. Now we consider the map $f : \mathbb{R}^m \to \mathbb{G} \simeq \mathbb{R}^m$ defined as

$$f(s_1, \ldots, s_m) = \mathcal{E}\left(u_0 + \sum_{i=1}^m s_i w_i\right).$$

The differential $D_0 f$ is invertible by construction, in particular f covers a neighbourhood of $f(0) = 0$. The idea of the proof is to construct a perturbation $g : \mathbb{R}^m \to \mathbb{G} \simeq \mathbb{R}^m$ of the form

$$g(s_1, \ldots, s_m) = \mathcal{E}\left(q_0 + \sum_{i=1}^m s_i p_i\right),$$

where $q_0, p_1, \ldots, p_m \in L^2(I, \mathbb{R}^k)$ are suitably chosen polynomial maps, in such a way that the image of g stills contains a neighbourhood of 0. This will be done by applying Lemma 4.9, as we now explain.

Given $\epsilon_1, \eta_1 > 0$ by the density of polynomial maps in $L^2(I, \mathbb{R}^k)$ and by the continuity of $\mathcal{E}$ and $D\mathcal{E}$, there exists a polynomial map $q_0 \in L^2(I, \mathbb{R}^k)$ such that

$$\|\mathcal{E}(q_0)\| < \epsilon_1 \qquad \text{and} \qquad \left\|D_{u_0}\mathcal{E} - D_{q_0}\mathcal{E}\right\|_{\text{op}} < \eta_1.$$

Now we prove that we can find polynomial maps $p_1, \dots, p_m$ and a neighbourhood $W \subset \mathbb{R}^m$ of 0 such that, setting $\sigma_{f,0} = \sigma(D_{x_0} f)$, it holds

$$\sup_W \|D_s f - D_s g\|_{\text{op}} \leq \frac{\sigma_{f,0}}{4}. \tag{4.12}$$

By the density of polynomial maps, given $\eta_2 > 0$ we consider $p_1, \dots, p_m$ such that for any $i = 1, \dots, m$

$$\|w_i - p_i\| < \eta_2.$$

We get the following estimate

$$\begin{aligned} \|D_0 f - D_0 g\|_{\text{op}} &\leq \sum_{i=1}^m \|D_{u_0}\mathcal{E}(w_i) - D_{q_0}\mathcal{E}(p_i)\| \\ &\leq \sum_{i=1}^m \|D_{u_0}\mathcal{E}(w_i) - D_{q_0}\mathcal{E}(w_i)\| + \|D_{q_0}\mathcal{E}(w_i) - D_{q_0}\mathcal{E}(p_i)\| \\ &\leq \left\|D_{u_0}\mathcal{E} - D_{q_0}\mathcal{E}\right\|_{\text{op}} \sum_{i=1}^m \|w_i\| + \left\|D_{q_0}\mathcal{E}\right\|_{\text{op}} \sum_{i=1}^m \|w_i - p_i\| \\ &\leq \eta_1 \sum_{i=1}^m \|w_i\| + \left\|D_{q_0}\mathcal{E}\right\|_{\text{op}} m\eta_2. \end{aligned}$$

Given $\eta_3 > 0$, by the continuity of Df and Dg there exists a neighbourhood $W \subset \mathbb{R}^m$ of 0 such that

$$\sup_{s \in W} \|D_s f - D_0 f\|_{\text{op}} + \|D_s g - D_0 g\|_{\text{op}} < \eta_3.$$

Hence, by the triangle inequality we get the following estimate

$$\sup_{s \in W} \|D_s f - D_s g\|_{\text{op}} \leq \eta_3 + \eta_1 \sum_{i=1}^m \|w_i\| + \left\|D_{q_0}\mathcal{E}\right\|_{\text{op}} m\eta_2.$$

To obtain (4.12) it is enough to choose η_1, η_2, η_3 small enough. We have thus found g in such a way $\|f(0) - g(0)\| < \epsilon_1$ and g satisfies the hypothesis of Lemma 4.9. By the latter, we find $r_{f,0} > 0$ such that

$$g(B_{\mathbb{R}^m}(r_{f,0})) \supset B_{\mathbb{R}^m}\left(g(0), \frac{r_{f,0}\sigma_{f,0}}{8}\right).$$

Since $\|g(0)\| < \epsilon_1$ by construction, taking $\epsilon_1 = \frac{r_{f,0}\sigma_{f,0}}{16}$ we also get

$$g(B_{\mathbb{R}^m}(r_{f,0})) \supset B_{\mathbb{R}^m}\left(\frac{r_{f,0}\sigma_{f,0}}{16}\right).$$

Let then $d_{\mathbb{G}}$ be the maximum degree of the polynomial maps $q_0, \ldots, p_m$. We have proved that

$$B_{\mathbb{R}^m}\left(\frac{r_{f,0}\sigma_{f,0}}{16}\right) \subset \mathcal{E}\Big(\{u \in L^2(I, \mathbb{R}^k) \mid u \text{ is a polynomial map of degree at most } d_{\mathbb{G}}\}\Big). \tag{4.13}$$

In other words, the Endpoint map is surjective on a small ball when restricted to polynomial controls of degree $\leq d_{\mathbb{G}}$. Now recall that, for Carnot groups, dilations have the following property:

$$\delta_\lambda(\mathcal{E}(u)) = \mathcal{E}(\lambda u), \qquad \forall\, \lambda > 0, \quad u \in L^2(I, \mathbb{R}^k). \tag{4.14}$$

From (4.13) and (4.14) it follows that $\mathcal{E}$, when restricted to polynomial controls of degree $\leq d_{\mathbb{G}}$ is surjective on the whole $\mathbb{G} \simeq \mathbb{R}^m$. □

With the same proof we can obtain the following statement, corresponding to Theorem K in the Introduction, of which Theorem 4.10 is a special case.

Theorem 4.11. *Let $\mathbb{G}$ be a Carnot group with topological dimension m and rank k. Let $S \subset L^2(I, \mathbb{R}^k)$ be a dense set. Then there exist $u_0, u_1, \ldots, u_m \in S$ such that*

$$\mathcal{E}(\mathrm{span}\{u_0, u_1, \ldots, u_m\}) = \mathbb{G}.$$

Chapter 5

Definable choices in semialgebraic geometry

In this chapter we prove the results described in Section iii.4.

5.1 Hausdorff approximations in semialgebraic geometry

5.1.1 Semialgebraic sets and maps

In order to study Hausdorff approximations of semialgebraic sets in $\mathbb{R}^n$, it will be convenient to work with an extension R of the field of real numbers, where real "infinitesimals" will be themselves elements of the field (below we will take for R the field of algebraic Puiseux series with coefficients in $\mathbb{R}$). For this reason, in this section we recall the basic notions from semialgebraic geometry over a general *real closed field* and we refer the reader to the monographs [13, 31] for more details.

Remark 5.1 (On the use of the language of real closed fields). The language of real closed fields might be unfamiliar for some readers, but it is especially useful in this context. We want to stress that using this language is not strictly necessary and it is possible that the proofs below can be formulated only using semialgebraic geometry in $\mathbb{R}^n$. However, it has become quite standard in semialgebraic geometry since, once it is introduced, the proofs and the statements become much shorter. Moreover, we are quoting some technical constructions from [12, 13, 14] that are stated using this language, and translating them to the classical semialgebraic language would make this part more technical – the precise point where Puiseux series will be used instead of classical real algebraic geometry is Proposition 5.40. From here, the tools of Hausdorff approximation that we develop in Section 5.1.3 will allow to get back to real sets, keeping control of this process.

Recall that a real closed field R is an ordered field whose positive cone is the set of squares of elements from R, and such that every polynomial in $R[x]$ of odd degree has a root in R. The order of R allows to define the sign of an element $r \in R$ as:

$$\operatorname{sign}(r) := \begin{cases} +1 & r > 0, \\ -1 & r < 0, \\ 0 & r = 0. \end{cases}$$

We denote by $R_+ := \{r \in R \,|\, \mathrm{sign}(r) \geq 0\}$ the cone of non-negative elements of R.

The field $\mathbb{R}$ is real closed. The main example of real closed field we use is the following.

Definition 5.2 (Algebraic Puiseux series)**.** Let R be a real closed field and $R(\zeta)$ be the field of rational functions in the variable ζ. The real closed field $R\langle\zeta\rangle$ of *algebraic Puiseux series with coefficients in R* is defined by:

$$R\langle\zeta\rangle := \left\{ f = \sum_{k=k_0}^{\infty} a_k \zeta^{\frac{k}{q}} \,\middle|\, a_k \in R,\ k_0 \in \mathbb{Z},\ m \in \mathbb{N},\ f \text{ is algebraic over } R(\zeta) \right\}.$$

In order to define a ring structure on the set $R\langle\zeta\rangle$, we insist that for $r_1, r_2 \in \mathbb{Q}$ it holds

$$\zeta^{r_1}\zeta^{r_2} = \zeta^{r_1+r_2}, \quad (\zeta^{r_1})^{r_2} = \zeta^{r_1 r_2}, \quad \zeta^0 = 1.$$

Therefore, two Puiseux series $a = \sum_{k \geq k_1} a_k \zeta^{k/q_1}$ and $b = \sum_{k \geq k_2} b_k \zeta^{k/q_2}$ can be written as a formal power series in $\zeta^{1/q}$, where q is the least common multiple between q_1 and q_2. This allows to add and multiply Puiseux series. If $a = a_1\zeta^{r_1} + a_2\zeta^{r_2} + \cdots \in R\langle\zeta\rangle$, with $a_1 \neq 0$ and $r_1 < r_2 < \cdots$, we say that $a > 0$ if $a_1 > 0$. From this definition it follows that $0 < \zeta < r$ for every $r \in R$ with $r > 0$, and this is why ζ is called *infinitesimal*.

Remark 5.3 (Algebraic Puiseux series as germs). The field $R\langle\zeta\rangle$ is isomorphic, as a real closed field, to the field of continuous semialgebraic functions germs $f : (0, \delta) \to R$, where $(0, \delta)$ is an interval in R, [13, Thm. 3.17]. Addition and multiplication translates to the usual ones. The order is defined as follows: the germ of a semialgebraic function $f : (0, \delta) \to R$ is positive if and only if there exists $0 < \delta' < \delta$ such that $f(\zeta) > 0$ for every $\zeta \in (0, \delta')$. Given a germ of a continuous semialgebraic function $f : (0, \delta) \to R$, in order to get an expression like $f = \sum_{k \geq k_0} a_k \zeta^{k/q}$, i.e. the corresponding element in the field of algebraic Puiseux series, one uses the fact that the graph of f lies on a *branch* of an algebraic curve $P(x, y) = 0$ in the plane R^2, where P is a polynomial with coefficients in R. In particular, this branch admits a parametrization near the origin as

$$x(\tau) = \tau^q, \quad y(\tau) = \sum_{k \geq k_0} a_k \tau^k,$$

where $q \in \mathbb{N}, k_0 \in \mathbb{Z}$ and $a_k \in R$, [93, Thm. 2.2]. This means that $P(x(\tau), y(\tau)) = 0$ as formal power series, and one gets a Puiseux series for f by the substitution $\tau = \zeta^{1/q}$ in $y(\tau)$.

If R is a real closed field, then so is $R\langle\zeta\rangle$, see [13, Cor. 2.98]. Using this observation, letting $K = R\langle\zeta_1\rangle$, then also $R\langle\zeta_1, \zeta_2\rangle = K\langle\zeta_2\rangle$ is a real closed field. Repeating this argument inductively, we define the real closed field of *algebraic Puiseux series with multiple infinitesimals and coefficients in R* as

$$R\langle\zeta_1, \ldots, \zeta_m\rangle := R\langle\zeta_1\rangle \cdots \langle\zeta_m\rangle.$$

Let R be a real closed field. The *norm* of an element $x = (x_1, \ldots, x_n) \in R^n$ is defined as

$$\|x\| := \sqrt{x_1^2 + \cdots + x_n^2} \in R.$$

For every $x = (x_1, \ldots, x_n) \in R^n$ and $r \in R_+$, we denote by $B_{R^n}(x, r) \subset R^n$ the set

$$B_{R^n}(x, r) := \left\{ y \in R^n \,\middle|\, \|x - y\| \leq r \right\},$$

and call it the *closed ball around x of radius r*. The *Euclidean topology* on R^n is the topology generated by the closed balls. These notions are inspired by their classical counterparts in the case $R = \mathbb{R}$, with the caveat that in these definitions r is not necessary real. Similarly, the order of a real closed field R can be used to define the notion of semialgebraic sets.

Definition 5.4 (Semialgebraic sets and maps)**.** We say that a set $S \subset R^n$ is *semialgebraic* if

$$S = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \left\{ x \in R^n \mid \operatorname{sign}(p_{ij}(x)) = \sigma_{ij} \right\}, \tag{5.1}$$

for some finite set of polynomials $p_{ij} \in R[x_1, \dots, x_n]$ and $\sigma_{ij} \in \{0, +1, -1\}$. We call the description (5.1) a *representation* of S.

Given two sets of polynomials $\mathcal{P} = \{p_1, \dots, p_{a_1}\}, \mathcal{Q} = \{q_1, \dots, q_{a_2}\} \subset R[x_1, \dots, x_n]$, we define the *algebraic* set $Z(\mathcal{P}; R)$ and the *closed basic* semialgebraic set $\operatorname{Bas}(\mathcal{P}, \mathcal{Q}; R) \subset R^n$ by

$$Z(\mathcal{P}; R) := \{x \in R^n \mid p_1(x) = \cdots = p_{a_1}(x) = 0\},$$

$$\operatorname{Bas}(\mathcal{P}, \mathcal{Q}; R) := Z(\mathcal{P}; R) \cap \left(\bigcap_{j=1}^{a_2} \left\{ x \in R^n \mid q_j(x) \le 0 \right\} \right).$$

Finally, a map $f : A \to B$ between semialgebraic sets $A \subset R^n, B \subset R^\ell$ is said to be *semialgebraic* if its graph is a semialgebraic set in $R^n \times R^\ell$.

The representation of a semialgebraic set S as in (5.1) is not unique, however having such a representation quantifies the complexity of S, using the following notion.

Definition 5.5 (Diagram of a semialgebraic set)**.** Let $S \subset R^n$ be a semialgebraic set represented as in (5.1). We say that the triple

$$\left(n, a \cdot \max_i \{b_i\}, \max_{i,j} \{\deg(p_{ij})\} \right) \in \mathbb{N}^3$$

is a *diagram* for S. Below, the equation "$D(S) = (m, c, d)$" will mean that there exists a representation of S as in (5.1) with $n \le m$, $a \cdot \max_i\{b_i\} \le c$ and $\max_{i,j}\{\deg(p_{ij})\} \le d$.

It is often useful to use an alternative (equivalent) description of semialgebraic sets, using the notion of *first–order formulas*.

Definition 5.6 (First–order formula)**.** Let R be a real closed field. A first–order formula of the language of ordered fields with coefficients in R is a formula written with a finite number of conjunctions, disjunctions, negations, and universal or existential quantifiers on variables, starting from atomic formulas which are formulas of the kind $p(x_1, \dots, x_n) = 0$ or $q(x_1, \dots, x_n) < 0$, where p and q are polynomials with coefficients in R. The free variables of a formula are those variables of the polynomials appearing in the formula which are not quantified.

Semialgebraic sets are precisely those defined by first–order formulas [31, Prop. 2.2.4].

Definition 5.7 (Set defined by a formula, see [13, Sect. 1.1])**.** Let ψ be a first–order formula of the language of ordered fields, with coefficients in R, and with n free variables. The *set defined by* ψ in R^n (or the *realization* of ψ in R^n) is the semialgebraic set

$$\mathrm{Reali}(\psi; R) \subseteq R^n$$

defined by induction on the construction of the formula, starting from atoms:

$$\mathrm{Reali}(p = 0; R) := \{x \in R^n \mid p(x) = 0\}, \quad \mathrm{Reali}(p < 0; R) := \{x \in R^n \mid p(x) < 0\},$$

(p is a polynomial with coefficients in R),

$$\begin{aligned}
\mathrm{Reali}(\phi_1 \wedge \phi_2; R) &:= \mathrm{Reali}(\phi_1; R) \cap \mathrm{Reali}(\phi_2; R),\\
\mathrm{Reali}(\phi_1 \vee \phi_2; R) &:= \mathrm{Reali}(\phi_1; R) \cup \mathrm{Reali}(\phi_2; R),\\
\mathrm{Reali}(\neg\phi; R) &:= R^n \setminus \mathrm{Reali}(\phi; R),\\
\mathrm{Reali}((\exists y)\,\phi; R) &:= \{x \in R^n \mid \exists y \in R,\ (x, y) \in \mathrm{Reali}(\phi; R)\},\\
\mathrm{Reali}((\forall y)\,\phi; R) &:= \{x \in R^n \mid \forall y \in R,\ (x, y) \in \mathrm{Reali}(\phi; R)\},
\end{aligned}$$

where ϕ_1, ϕ_2, ϕ are first–order formulas with an appropriate number of free variables.

5.1.2 Some properties of semialgebraic sets

We collect here some properties of semialgebraic sets over a real closed field.

Semialgebraic triviality

Continuous semialgebraic maps $f : A \to B$ are "piecewise" trivial fibrations: there exists a partition of B into finitely many semialgebraic sets

$$B = \bigsqcup_{j=1}^{b} B_j$$

and, for every $j = 1, \ldots, b$ there exist fibers $F_j := f^{-1}(y_j)$, for some $y_j \in B_j$, and a semialgebraic homemorphism $\varphi_j : B_j \times F_j \to f^{-1}(B_j)$ that makes the following diagram commutative:

$$\begin{array}{ccc}
B_j \times F_j & \xrightarrow{\ \varphi_j\ } & f^{-1}(B_j)\\
 & {\scriptstyle p_1}\searrow \quad \swarrow{\scriptstyle f} & \\
 & B_j &
\end{array}$$

This result is called *semialgebraic triviality*, see [31, Thm. 9.3.2].

Corollary 5.8. *Definable choices exist.*

Proof. Let $S \subset \mathbb{R}^n$ be a semialgebraic set and $\pi : \mathbb{R}^n \to \mathbb{R}^\ell$ the projection. Then, by semialgebraic triviality there exists a finite partition of $\pi(S) = \sqcup_{j=1}^{b} B_j$ into semialgebraic sets and for every $j = 1, \ldots, b$ there exists fibers $F_j = \pi^{-1}(y_j) \subset S$, with $y_j \in B_j$, and semialgebraic homeomorphisms $\varphi_j : B_j \times F_j \to \pi^{-1}(B_j) \cap S$ making the corresponding diagram commutative. For every $j = 1, \ldots, b$ choose an element $x_j \in F_j$. The definable choice of S over $\pi(S)$ is the set

$$\bigcup_{j=1}^{b} \varphi_j\left(B_j \times \{x_j\}\right),$$

concluding the proof. □

Remark 5.9. Note that the complexity of a set built as in the proof of Corollary 5.8 depends on the complexity of the objects involved in the semialgebraic triviality. This uses the so–called cylindrical algebraic decomposition and it is known to be doubly exponential in the number of variables of S (see [36, 32]).

Dimension and stratifications

Every semialgebraic subset of R^n can be written as a finite union of semialgebraic sets, each of them semialgebraically homeomorphic to an open cube $(0, 1)^m \subset R^m$, for some $m \leq n$ ([31, Thm. 2.3.6]). This allows to define the *dimension* of a semialgebraic set as the maximum of the dimensions m of these cubes. In fact, it is possible to introduce the notion of smoothness also over general real closed fields, and every semialgebraic set can be written as a finite union of smooth, semialgebraic disjoint manifolds called *strata* ([31, Prop. 9.1.8]). This is called a *Nash stratification*. We will use this last result only in the classical case $R = \mathbb{R}$.

The dimension of a semialgebraic set is preserved by semialgebraic homeomorphisms and behaves naturally under Cartesian product:

$$\dim(A \times B) = \dim(A) + \dim(B).$$

Moreover, if $f : A \to B$ is a continuous semialgebraic map, then

$$\dim(f(A)) \leq \dim(A). \tag{5.2}$$

We will also need a stronger notion of dimension, introduced in [14].

Definition 5.10 (Strong dimension)**.** Let $S \subset R^n$ be a semialgebraic set and $1 \leq \ell \leq n$. We say that S is *strongly of dimension* $\leq \ell$ if, letting $\pi_\ell : R^n \to R^\ell$ be the projection on the last ℓ coordinates, it holds

$$\forall y \in R^\ell \quad \#\left(S \cap \pi_\ell^{-1}(y)\right) < \infty. \tag{5.3}$$

It follows from semialgebraic triviality (see Section 5.1.2) that if (5.3) holds, then

$$\sup_{y \in R^\ell} \#\left(S \cap \pi_\ell^{-1}(y)\right) < \infty, \tag{5.4}$$

so that in Definition 5.10 one can replace (5.3) with the apparently stronger condition (5.4).

Moreover, if S is strongly of dimension $\leq \ell$, then $\dim(S) \leq \ell$. However, a semialgebraic set S of dimension ℓ may not be strongly of dimension $\leq \ell$, as this depends on the relative position between S and the projecting subspace R^ℓ. Nevertheless we have the following result.

Proposition 5.11 (Genericity of strong dimension)**.** *Let $S \subset R^n$ be a semialgebraic set of dimension ℓ. The set of invertible linear transformations $L : R^n \to R^n$ such that $L(S)$ is strongly of dimension $\leq \ell$ is semialgebraic and dense.*

Proof. We first prove the case $R = \mathbb{R}$. The set of such linear transformations is semialgebraic, as it admits a semialgebraic description in the coefficients of L.

To prove density, let $M \subset \mathbb{R}^n$ be a smooth p–dimensional manifold with $p \leq \ell$. Given $r \in \mathbb{N}$ and a Thom–Boardman manifold $\Sigma \subset J^r(M, \mathbb{R}^\ell)$, it follows from [70, Thm. 2] that the set of linear maps $T : \mathbb{R}^n \to \mathbb{R}^\ell$ that restrict to maps $T|_M : M \to \mathbb{R}^\ell$ which are transversal to Σ is dense. In particular, the set of linear maps $T : \mathbb{R}^n \to \mathbb{R}^\ell$ such that $T|_M$ is a stratified immersion is dense. The fibers of every such map are discrete and every such map can be obtained as $T = \pi_\ell \circ L$ for an appropriate $L : \mathbb{R}^n \to \mathbb{R}^n$. Therefore, for every smooth p–dimensional manifold $M \subset \mathbb{R}^n$, with $p \leq \ell$, the set $\mathcal{L}(M)$ of linear maps $L : \mathbb{R}^n \to \mathbb{R}^n$ such that the fibers of the map $\pi_\ell|_{L(M)} : L(M) \to \mathbb{R}^\ell$ are discrete is dense. When M is semialgebraic, the set $\mathcal{L}(M)$ is also semialgebraic and therefore it contains an open dense set.

Let $S = \sqcup_{i=1}^a M_i$ be a Nash stratification of S. Applying the above argument to each stratum M_i, we see that the set $\mathcal{L}(S) := \mathcal{L}(M_1) \cap \cdots \cap \mathcal{L}(M_a)$ is semialgebraic and contains an open dense set. For every $L \in \mathcal{L}(S)$ the fibers of $\pi_\ell|_{L(S)} : L(S) \to \mathbb{R}^\ell$ are discrete and semialgebraic, therefore they are finite. Hence $L(S)$ is strongly of dimension $\leq \ell$. This concludes the proof for the case $R = \mathbb{R}$.

The statement we just proved can be written as a first–order formula with coefficients in $\mathbb{R}$. Therefore, the same statement holds if $\mathbb{R}$ is replaced by a real closed extension R of it, by [31, Prop. 5.2.3]. (This is the so-called *transfer principle* from real algebraic geometry: every property that can be expressed in the first–order language of ordered fields with coefficients in R can be transferred to any real closed extension of R.) □

Thom-Milnor bound

For every $c \in \mathbb{N}$ there exists $\beta_{\text{tm}} = \beta_{\text{tm}}(c) > 1$ such that, if S is a semialgebraic set with diagram $D(S) = (n, c, d)$, then the number of connected components of S, denoted by $b_0(S)$, can be bounded by ([13, Thm. 7.50]):

$$b_0(S) \leq \beta_{\text{tm}} d^n. \tag{5.5}$$

5.1.3 Hausdorff approximations using infinitesimals

Using Puiseux series we can produce Hausdorff approximations of semialgebraic sets in $\mathbb{R}^n$ as "specializations" of semialgebraic sets defined on an extension R^n which contains some "infinitesimals", and this can be done in a controlled way. Let us explain this idea.

Example 5.12. Given a set $A \subset \mathbb{R}^n$, for $r > 0$ denote its r–neighbourhood in $\mathbb{R}^n$ by

$$\mathcal{U}_r(A) := \left\{ x \in \mathbb{R}^n \,\middle|\, \exists a = (a_1, \ldots, a_n) \in A \quad \text{s.t.} \quad \sum_{i=1}^n (a_i - x_i)^2 \leq r^2 \right\}.$$

If A is semialgebraic, defined by a first order formula ϕ, then $\mathcal{U}_r(A)$ is defined by the first order formula

$$\phi_r := \left(\exists a \left(\phi(a) \wedge \sum_{i=1}^n (a_i - x_i)^2 \leq r^2 \right) \right).$$

Instead of interpreting $\{\phi_r(x)\}_{r>0}$ as a *family* of first order formulas with coefficients in $\mathbb{R}$, we can interpret it as a single first order formula ψ with coefficients in $R = \mathbb{R}\langle\zeta\rangle$

$$\psi := \left(\exists a \left(\phi(a) \wedge \sum_{i=1}^{n}(a_i - x_i)^2 \leq \zeta^2\right)\right).$$

The set $S \subset R^n$ defined by the formula ψ encodes the family of neighbourhoods of the original semialgebraic set S in a semialgebraic way. Note that the r–neighbourhood of S is described by the formula obtained from S by “evaluating” the infinitesimal ζ at r.

The evaluation map and the map $\mathscr{L}_0$

The procedure described in Example 5.12 is a special case of a general construction of “evaluation” of a semialgebraic set of Puiseux series. First, we provide the following definition, inspired by [41, Def. 1.7], to formalize the concept of properties that are true for “sufficiently small” nested sequences of Puiseux series. The intricacy of the statements for the case $m > 1$ is a consequence of the iterative definition of Puiseux series $R\langle\zeta_1, \ldots, \zeta_m\rangle$ for several infinitesimals.

Definition 5.13 (Predicates for sufficiently small Puiseux series)**.** Let R be a real closed field, $1 \leq k \leq m$, and let $\mathscr{P} = \mathscr{P}(t_k, \ldots, t_m)$ be a property where $t_j \in R\langle\zeta_1, \ldots, \zeta_{j-1}\rangle$ for $k \leq j \leq m$. We say that the property $\mathscr{P}$ holds for

$$0 < t_k \ll \cdots \ll t_m \ll 1$$

if there exists $\delta_m \in R\langle\zeta_1, \ldots, \zeta_{m-1}\rangle$ with $\delta_m > 0$ such that for $t_m \in R\langle\zeta_1, \ldots, \zeta_{m-1}\rangle$ with $0 < t_m < \delta_m$ there exists $\delta_{m-1} \in R\langle\zeta_1, \ldots, \zeta_{m-2}\rangle$ with $\delta_{m-1} > 0$ such that for all $t_{m-1} \in R\langle\zeta_1, \ldots, \zeta_{m-2}\rangle$ with $0 < t_{m-1} < \delta_{m-1}$ there exists (...) $\delta_k \in R\langle\zeta_1, \ldots, \zeta_{k-1}\rangle$ with $\delta_k > 0$ such that for all $t_k \in R\langle\zeta_1, \ldots \zeta_{k-1}\rangle$ with $0 < t_k < \delta_k$ we have $\mathscr{P}(t_k, \ldots, t_m)$.

Definition 5.14 (The evaluation map)**.** Let R be a real closed field and let ψ be a formula with n free variables and with coefficients in $R\langle\zeta_1, \ldots, \zeta_m\rangle$. For every $1 \leq k \leq m$ and for every $(t_k, \ldots, t_m)$ with $t_j \in R\langle\zeta_1, \ldots, \zeta_{j-1}\rangle$ for $j = k, \ldots, m$, we denote by $\psi|_{\zeta_k=t_k,\ldots,\zeta_m=t_m}$ the formula with coefficients in $R\langle\zeta_1, \ldots, \zeta_{k-1}\rangle$ that is obtained from ψ by replacing iteratively in its coefficients each ζ_j by t_j, starting from $j = m$. This can be done provided that, at each step, t_j is sufficiently small so that the coefficients of the formula coming from the previous step (which are finitely many and which we see as germs of continuous semialgebraic functions in ζ_j) have a representative defined on a common interval containing t_j, yielding a Puiseux series in $j-1$ infinitesimals. If $S \subset R\langle\zeta_1, \ldots, \zeta_m\rangle^n$ is the semialgebraic set defined by ψ, we denote by $S|_{t_k,\ldots,t_m} \subset R\langle\zeta_1, \ldots, \zeta_{k-1}\rangle^n$ the semialgebraic set defined by the formula $\psi|_{\zeta_k=t_k,\ldots,\zeta_m=t_m}$, i.e.

$$S|_{t_k,\ldots,t_m} := \mathrm{Reali}\Big(\psi|_{\zeta_k=t_k,\ldots,\zeta_m=t_m}; R\langle\zeta_1, \ldots, \zeta_{k-1}\rangle\Big) \subset R\langle\zeta_1, \ldots, \zeta_{k-1}\rangle^n, \tag{5.6}$$

provided that the right hand side of (5.6) is defined. In particular, this is the case for $0 < t_k \ll \cdots \ll t_m \ll 1$. The set $S|_{t_k,\ldots,t_m}$ is called the *evaluation* of S (at $t_k, \ldots, t_m$).

Remark 5.15. Indeed, (5.6) depends on the formula ψ defining S; for us the formula will be clear and this abuse of notation will create no harm.

Remark 5.16 (Polynomial coefficients). If the coefficients of the formula ψ are *polynomials* in the infinitesimals (and not general algebraic Puiseux series), then (5.6) is defined for any $(t_k, \dots, t_m)$ with $t_j \in R[\zeta_1, \dots, \zeta_{j-1}] \subset R\langle\zeta_1, \dots, \zeta_{j-1}\rangle$. In this case, at the i–th step of the process one gets a new formula with coefficients in $R[\zeta_1, \dots, \zeta_{m+i-2}]$; therefore the process can be iterated and the concept of "sufficiently small" from Definition 5.13 can be dispensed of. This is the approach followed in [12] in the case of one infinitesimal ($m = 1$). Our general formulation, albeit defined only for $0 < t_k \ll \cdots \ll t_m \ll 1$, is more flexible and technically necessary. Notably, the key Lemmas 5.19 and 5.20 require evaluation at general Puiseux series.

Remark 5.17 (Evaluations and germs). Let $S = \mathrm{Reali}(\psi; R\langle\zeta\rangle) \subset R\langle\zeta\rangle^n$ be a semialgebraic set. Recall from Remark 5.3 that the germ of continuous semialgebraic function $g : (0, \delta) \to R$ can be identified as an element of $R\langle\zeta\rangle$ (we denote this germ still by g). Then, using the notation from Definition 5.14, the following property is true by construction:

$$\exists\delta > 0 \quad \forall r \in (0, \delta) \quad g(r) \in S|_r \quad \Longleftrightarrow \quad g \in S. \tag{5.7}$$

Lemma 5.18 (Evaluation preserves the diagram)**.** *In the same setting of Definition 5.14, for every* $1 \le k \le m$ *it holds*

$$D(S|_{t_k,\dots,t_m}) = D(S),$$

for all $(t_k, \dots, t_m)$ *such that the evaluation is defined.*

Proof. The statement follows from the fact that a representation of $S|_{t_1,\dots,t_m}$ is obtained from a representation of S by replacing ζ_j with t_j in the coefficients, in the prescribed order. □

The strong dimension condition of Definition 5.10 is not increased by small evaluations.

Lemma 5.19 (Small evaluations do not increase the strong dimension: one infinitesimal)**.** *In the same setting of Definition 5.14, with* $m = 1$*. Let* $\pi_\ell : R\langle\zeta\rangle^n \to R\langle\zeta\rangle^\ell$ *be the projection on the last* ℓ *coordinates,* $1 \le \ell \le n$*. Then, there exists* $\delta \in R$*,* $\delta > 0$ *such that for all* $t \in R$ *with* $0 < t < \delta$ *it holds*

$$\sup_{y \in R^\ell} \#S|_t \cap \pi_\ell^{-1}(y) \le \sup_{w \in R\langle\zeta\rangle^\ell} \#S \cap \pi_\ell^{-1}(w).$$

In particular, if $S \subset R\langle\zeta\rangle^n$ *is strongly of dimension* $\le \ell$*, then the set* $S|_t$ *is also strongly of dimension* $\le \ell$*.*

Proof. In the proof, set $\pi = \pi_\ell$. Let

$$N := \sup_{w \in R\langle\zeta\rangle^\ell} \#\left(S \cap \pi^{-1}(w)\right). \tag{5.8}$$

If $N = \infty$ there is nothing to prove, then assume $N < \infty$.

Assume by contradiction that for all $\delta \in R$, $\delta > 0$, there exists $t \in R$, $0 < t < \delta$, and $y \in R^\ell$ such that

$$\#S|_t \cap \pi^{-1}(y) > N.$$

Consider the semialgebraic set $D \subset R^\ell \times R_+$ defined by

$$D := \left\{(y, t) \in R^\ell \times R_+ \mid \#S|_t \cap \pi^{-1}(y) > N\right\}.$$

Let $P_2 : D \to R_+$ be the projection on the second factor. By our assumption the image $P_2(D) \subset R_+$ contains a sequence accumulating to 0. By semialgebraic triviality (Section 5.1.2) there exists $\delta_0 \in R_+$, $\delta_0 > 0$, a point $t \in (0, \delta_0)$ a fiber $F_0 = P_2^{-1}(t_0)$ and a semialgebraic homomorphism $\Phi : (0, \delta_0) \times F_0 \to P_2^{-1}((0, \delta_0))$ such that the following diagram is commutative

$$\begin{array}{ccc} (0,\delta_0) \times F_0 & \xrightarrow{\quad\Phi\quad} & P_2^{-1}((0,\delta_0)) \\ & {\scriptstyle p_1}\searrow \quad \swarrow{\scriptstyle P_2} & \\ & (0,\delta_0) & \end{array}$$

Fix an element $f_0 \in F_0$ and consider the function $y_0 : (0, \delta_0) \to R^\ell$ given by

$$y_0(t) := P_1 \circ \Phi(t, f_0),$$

where $P_1 : D \to R^\ell$ denotes the projection in the first factor. The (components of the) function y_0 are semialgebraic and continuous curves on the right of the origin, thus the corresponding germs (which we denote with the same symbol) define algebraic Puiseux series $y_0 \in R\langle\zeta\rangle^\ell$. Furthermore, we note that by construction $(y_0(t), t) \in D$ for all $t \in (0, \delta_0)$ so that

$$\forall t \in (0, \delta_0), \qquad \#\Big(S|_t \cap \pi^{-1}(y_0(t))\Big) > N. \tag{5.9}$$

Denoting by $(v, w) \in R\langle\zeta\rangle^n = R\langle\zeta\rangle^{n-\ell} \times R\langle\zeta\rangle^\ell$, the set $\pi^{-1}(y_0(t)) \subset R^n$ can be seen as the evaluation at t of the set $\pi^{-1}(y_0) \in R\langle\zeta\rangle^n$, defined by the formula $(w = y_0)$ with variables (v, w) and coefficients in $R\langle\zeta\rangle$.

Thus for $0 < t \ll 1$ (which means, according to Definition 5.13, up to taking a smaller δ_0, for all $0 < t < \delta_0$), it holds

$$\begin{aligned} \Big(S \cap \pi^{-1}(y_0)\Big)\Big|_t &= \mathrm{Reali}\Big((\psi \cap (w = y_0))|_{\zeta=t}; R\Big) \\ &= \mathrm{Reali}\Big(\psi|_{\zeta=t}; R\Big) \cap \mathrm{Reali}\Big((w = y_0)|_{\zeta=t}; R\Big) \\ &= S|_t \cap \pi^{-1}(y_0(t)). \end{aligned}$$

In particular from (5.9) we obtain

$$\forall t \in (0, \delta_0) \qquad \#\Big(S \cap \pi^{-1}(y_0)\Big)\Big|_t = \#\Big(S|_t \cap \pi^{-1}(y_0(t))\Big) > N. \tag{5.10}$$

Consider then the semialgebraic set $T \subset R^n \times R_+$ defined by

$$T := \left\{(z, t) \in R^n \times R_+ \,\middle|\, z \in \Big(S \cap \pi^{-1}(y_0)\Big)\Big|_t\right\}.$$

Denote by $P : T \to R_+$ the projection on the last factor. By semialgebraic triviality, for $\delta_0 > 0$ small enough, we find a semialgebraic homomorphism

$$\varphi : (0, \delta_0) \times H \to P^{-1}((0, \delta_0)),$$

where $H = P^{-1}(\bar{t})$ for some $\bar{t} \in (0, \delta_0)$. By (5.10) such a fiber must have cardinality $> N$. Pick $N + 1$ distinct points $h_1, \ldots, h_{N+1} \in H$. Denote by $Q : T \to R^n$ the projection on the first factor and define for $j = 1, \ldots, N + 1$ continuous semialgebraic curves $g_j : (0, \delta_0) \to R^n$ by

$$g_j(t) := Q(\varphi(t, h_j)).$$

By construction, these curves satisfy

$$\forall t \in (0, \delta_0) \quad g_j(t) \in \left(S \cap \pi^{-1}(y_0)\right)\Big|_t .$$

Therefore (the germ of) each $g_j \in S \cap \pi^{-1}(y_0) \subset R\langle\zeta\rangle^n$, by (5.7). Since the elements h_j are distinct, the germs of these curves near zero represent $N+1$ distinct points in $S \cap \pi^{-1}(y_0)$, contradicting (5.8).

We have proved that there exists $\delta > 0$ such that for all $t \in R$ with $0 < t < \delta$ it holds

$$\sup_{y \in R^\ell} \#S|_t \cap \pi^{-1}(y) \leq \sup_{w \in R\langle\zeta\rangle^\ell} \#S \cap \pi^{-1}(w).$$

which is the statement. □

We will need a version of Lemma 5.19 for multiple infinitesimals ($m > 1$).

Lemma 5.20 (Small evaluations do not increase the strong dimension: several infinitesimals)**.** *In the same setting of Definition 5.14, let* $\pi_\ell : R\langle\zeta_1, \dots, \zeta_m\rangle^n \to R\langle\zeta_1, \dots, \zeta_m\rangle^\ell$ *be the projection on the last* ℓ *coordinates,* $1 \leq \ell \leq n$*, and let* $1 \leq k \leq m$*. Then for* $0 < t_k \ll \cdots \ll t_m \ll 1$ *it holds*

$$\sup_{y \in R\langle\zeta_1,\dots,\zeta_{k-1}\rangle^\ell} \#S|_{t_k,\dots,t_m} \cap \pi_\ell^{-1}(y) \leq \sup_{w \in R\langle\zeta_1,\dots,\zeta_m\rangle^\ell} \#S \cap \pi_\ell^{-1}(w).$$

In particular, if $S \subset R\langle\zeta_1, \dots, \zeta_m\rangle^n$ *is strongly of dimension* $\leq \ell$*, then for* $0 < t_k \ll \cdots \ll t_m \ll 1$ *the set* $S|_{t_k,\dots,t_m}$ *is also strongly of dimension* $\leq \ell$*.*

Proof. Note that Lemma 5.19 corresponds to the case $m = 1$. To prove the case $m > 1$, recall that $R\langle\zeta_1, \dots, \zeta_m\rangle \simeq K\langle\zeta\rangle$, with $K := R\langle\zeta_1, \dots, \zeta_{m-1}\rangle$. We find $\delta_m \in R\langle\zeta_1, \dots, \zeta_{m-1}\rangle$, $\delta_m > 0$, such that for all $t_m \in R\langle\zeta_1, \dots, \zeta_{m-1}\rangle$ with $0 < t_m < \delta_m$ it holds

$$\sup_{y \in R\langle\zeta_1,\dots,\zeta_{m-1}\rangle^\ell} \#S|_{t_m} \cap \pi^{-1}(y) \leq \sup_{w \in R\langle\zeta_1,\dots,\zeta_m\rangle^\ell} S \cap \pi^{-1}(w).$$

We can now iterate the argument k-times, with $1 \leq k \leq m$ to get the statement, using the definition of $S|_{t_k,\dots,t_m}$. □

Definition 5.21 (Bounded elements and the limit homomorphism: one infinitesimal)**.** Let R be a real closed field. An element $s \in R\langle\zeta\rangle$ is called *bounded over* R if $\|s\| \leq r$ for some $r \in R_+$. The subring $R\langle\zeta\rangle_b$ of elements that are bounded over R consists of algebraic Puiseux series with non-negative exponents.

For all $n \in \mathbb{N}$, we define the *limit homomorphism*

$$\lambda_\zeta : R\langle\zeta\rangle_b^n \to R^n, \tag{5.11}$$

the ring homomorphism mapping $\sum_{k\geq 0}^{\infty} a_k \zeta^{\frac{k}{m}}$ to $a_0 \in R^n$.

Remark 5.22. Viewing $R\langle\zeta\rangle$ as the field of germs of semialgebraic functions continuous on the right of zero, the bounded elements corresponds to those germs that have a finite limit as $\zeta \to 0$ and

$$\lambda_\zeta(f) = \lim_{\zeta \to 0} f(\zeta).$$

From this we see that the map λ_ζ is order preserving, in the following sense: if $f_1, f_2 \in R\langle\zeta\rangle_b$ with $f_1 \leq f_2$, then $\lambda_\zeta(f_1) \leq \lambda_\zeta(f_2)$.

Definition 5.23 (Bounded elements: several infinitesimals)**.** Let R be a real closed field. We say that $f = (f_1, \dots, f_n) \in R\langle\zeta_1, \dots, \zeta_m\rangle^n$ is *bounded over* R if $\|f\| \le r$ for some $r \in R_+$.

Remark 5.24. The subring $R\langle\zeta_1, \dots, \zeta_m\rangle_b^n$ of elements that are bounded over R contains algebraic Puiseux series with non-negative exponents. However the inclusion is strict: the Puiseux series $t = \zeta_1^{-1}\zeta_2 \in R\langle\zeta_1, \zeta_2\rangle$ is such that $\|t\| \le r$ for all $r \in R_+$, $r > 0$.

Note that if R is real closed, then on the set of elements of $R\langle\zeta_1, \dots, \zeta_m\rangle^n$ that are bounded over R (i.e. with $\|f\| \le r$ with $r \in R_+$) the composition of maps

$$R\langle\zeta_1, \dots, \zeta_m\rangle_b^n \xrightarrow{\lambda_{\zeta_m}} R\langle\zeta_1, \dots, \zeta_{m-1}\rangle_b^n \xrightarrow{\lambda_{\zeta_{m-1}}} \cdots R\langle\zeta_1\rangle_b^n \xrightarrow{\lambda_{\zeta_1}} R^n$$

is well–defined, since at every step we get bounded elements.

Remark 5.25. We stress that the composition above is well–defined only if taken with the order prescribed by the infinitesimals. For instance, let $f \in \mathbb{R}\langle\zeta_1, \zeta_2\rangle_b$ be given by

$$f = \sum_{k \ge 0} a_k(\zeta_1)\zeta_2^{\frac{k}{q}} \quad \text{where} \quad a_k(\zeta_1) = \sum_{j \ge 0} b_{k,j}\zeta_1^{\frac{j}{q_k}}.$$

(Note that each $a_k(\zeta_1)$ has "its own" q_k.) Then

$$f = \sum_{k \ge 0}\left(\sum_{j \ge 0} b_{k,j}\zeta_1^{\frac{j}{m_k}}\right)\zeta_2^{\frac{k}{m}},$$

and $\lambda_{\zeta_1}(\lambda_{\zeta_2}(f)) = b_{0,0}$, whereas the composition in the other order is not well–defined.

Definition 5.26 (The map $\mathscr{L}_0$)**.** Let R be a real closed field and let $S \subset R\langle\zeta_1, \dots, \zeta_m\rangle^n$ be a semialgebraic set bounded over R. We denote by

$$\mathscr{L}_0(S) := \lambda_{\zeta_1} \cdots \lambda_{\zeta_m}(S).$$

Remark 5.27. The set $\mathscr{L}_0(S) \subset R^n$ is closed. In fact, if S is defined by a formula ψ with coefficients in $R\langle\zeta\rangle$, then it follows from [13, Prop. 12.43] that

$$\begin{aligned} \mathscr{L}_0(S) &= \overline{\left\{(x,t) \in R^{n+1} \,\middle|\, (x \in S|_t) \wedge (t > 0)\right\}} \cap R^n \\ &= \overline{\mathrm{Reali}\Big(\psi|_{\zeta = t} \wedge (t > 0); \mathbb{R}\Big)} \cap R^n, \end{aligned} \tag{5.12}$$

where we identify R^n with $\{u = 0\}$. A similar argument holds for $m > 1$. Notice, however, that deducing a presentation for $\mathscr{L}_0(S)$ as in Definition 5.4 is more complicated and requires quantifier elimination, which would bring us back to the problem mentioned in Remark 5.9.

Hausdorff limits

Going back to Example 5.12, assuming that $A \subset \mathbb{R}^n$ is bounded, then also the corresponding $S \subset \mathbb{R}\langle\zeta\rangle^n$ is bounded. Moreover, we notice that $\mathscr{L}_0(S) = A$. Then, the set $\mathcal{U}_r(A) = S|_r$ converges in the Hausdorff metric to $A = \mathscr{L}_0(S)$. To state the analogue result in general, we need to recall some more preliminary notions.

Let R be a real closed field. Given a semialgebraic set $S \subset R^n$, the distance from S is the function defined by

$$\delta_S(x) := \inf_{s\in S} \|x - s\|, \quad x \in R^n.$$

This is a continuous, semialgebraic function, vanishing on the closure of S and positive elsewhere, see [31, Prop. 2.2.8]. Given $S \subset R^n$ and $r \in R$, the r–neighbourhood of S in R^n is the set defined by

$$\mathcal{U}_r(S, R^n) := \left\{ x \in R^n \,\middle|\, \delta_S(x) \le r \right\}.$$

Since $\mathrm{dist}_S(\cdot)$ is semialgebraic, for every $r > 0$ the set $\mathcal{U}_r(S, R^n)$ is also semialgebraic.

Definition 5.28 (Semialgebraic Hausdorff distance)**.** The *Hausdorff distance* between two semialgebraic sets $S_1, S_2 \subset R^n$ is defined as

$$\mathrm{dist}_H(S_1, S_2) := \inf \left\{ \epsilon \in R \,\middle|\, S_1 \subseteq \mathcal{U}_\epsilon(S_2),\ S_2 \subseteq \mathcal{U}_\epsilon(S_2) \right\}.$$

Note that, if $R = \mathbb{R}$, this gives the usual Hausdorff distance. However, for general real closed fields R, this is not a "distance" in the sense of metric geometry, since the values of this function are elements of R. Still, given three closed semialgebraic sets S_1, S_2, S_3, we have

$$\mathrm{dist}_H(S_1, S_3) \le \mathrm{dist}_H(S_1, S_2) + \mathrm{dist}_H(S_2, S_3). \tag{5.13}$$

This is proved exactly as in the classical case $R = \mathbb{R}$.

The next result outlines a useful property related to the map (5.11).

Proposition 5.29. *Let $S_1, S_2 \subset R\langle\zeta\rangle^n_b$ be semialgebraic sets. Then*

$$\mathrm{dist}_H(\lambda_\zeta(S_1), \lambda_\zeta(S_2)) \le \lambda_\zeta\left(\mathrm{dist}_H(S_1, S_2)\right). \tag{5.14}$$

Proof. We first prove the following fact. If $x, y \in R\langle\zeta\rangle^n_b$, then

$$\|\lambda_\zeta(x) - \lambda_\zeta(y)\| = \lambda_\zeta\left(\|x - y\|\right). \tag{5.15}$$

In fact, since $\lambda_\zeta : R\langle\zeta\rangle^n_b \to R^n$ is a ring homomorphism, writing $u = \sum_{k\ge0} a_k \zeta^{\frac{k}{q}}$, with $a_k \in R^n$, it holds

$$\|\lambda_\zeta(u)\| = \|a_0\| = \lambda_\zeta(\|u\|).$$

Let us now prove (5.14). Let $\epsilon \in R\langle\zeta\rangle$ such that $\mathrm{dist}_H(S_1, S_2) \le \epsilon$. This is equivalent to the following statement: for every $a_1 \in S_1$ there exists $a_2(a_1) \in S_2$ such that $\|a_1 - a_2(a_1)\| \le \epsilon$ and for every $b_2 \in S_2$ there exists $b_1(b_2) \in S_1$ such that $\|b_1(b_2) - b_2\| \le \epsilon$.

Then, for every $r_1 = \lambda_\zeta(a_1) \in \lambda_\zeta(S_1)$ the element $r_2 := \lambda_\zeta(a_2(a_1)) \in \lambda_\zeta(S_2)$ is such that, using (5.15),

$$\|r_1 - r_2\| = \|\lambda_\zeta(a_1) - \lambda_\zeta(a_2(a_1))\| = \lambda_\zeta(\|a_1 - a_2\|) \leq \lambda_\zeta(\epsilon).$$

Similarly, for every $s_2 \in \lambda_\zeta(S_2)$ there exists and element $s_1 \in \lambda_\zeta(S_1)$ such that

$$\|s_1 - s_2\| \leq \lambda_\zeta(\epsilon).$$

This means that $\mathrm{dist}_H(\lambda_\zeta(S_1), \lambda_\zeta(S_2)) \leq \lambda_\zeta(\epsilon)$. □

Recall now the following result from [12].

Proposition 5.30 ([12, Prop. 2.7]). *Let $S = \mathrm{Reali}(\psi; \mathbb{R}\langle\zeta\rangle) \subset \mathbb{R}\langle\zeta\rangle^n$ be a bounded semialgebraic set. Then, in the usual Hausdorff metric,*

$$\lim_{r\to 0} S|_r = \mathscr{L}_0(S).$$

Remark 5.31. The previous result is stated in [12, Prop. 2.7] under the assumption that the formula ψ has coefficients in $\mathbb{R}[\zeta]$. In fact the proof in the general case goes exactly in the same way, provided that r is sufficiently small so that the evaluation is well–defined as explained in Definition 5.14.

Remark 5.32. Even if $S|_0$ is well–defined, in general, this may be far, in the Hausdorff metric, from $\mathscr{L}_0(S)$. For example let $\alpha = \left(x^2 - \zeta^2 < 0\right)$ and $A = \mathrm{Reali}(\alpha; R\langle\zeta\rangle)$. Then $A_0 = \emptyset$ and $\mathscr{L}_0(A) = \{0\}$, so that $A|_0 \subsetneq \mathscr{L}_0(A)$. On the other hand, if $\beta = (\zeta x = 0)$, and $B = \mathrm{Reali}(\beta; R\langle\zeta\rangle)$, then $B|_0 = R$ and $\mathscr{L}_0(B) = \{0\}$, so that $B|_0 \supsetneq \mathscr{L}_0(B)$.

We extend Proposition 5.30 to any real closed field using the transfer principle.

Proposition 5.33. *Let F be a real closed extension of $\mathbb{R}$. Let $S = \mathrm{Reali}(\psi; F\langle\zeta\rangle) \subset F\langle\zeta\rangle^n$ be a bounded semialgebraic set. Then,*

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall 0 < t < \delta \quad \mathrm{dist}_H(\mathrm{Reali}(\psi|_{\zeta=t}; F), \lambda_\zeta(S)) < \epsilon.$$

(Here all the variables ϵ, δ, t are in F, and dist_H denotes the semialgebraic Hausdorff distance.)

Proof. Given a semialgebraic set S as in the statement of Proposition 5.30, we can rephrase the content of its conclusion by saying that,

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall 0 < t < \delta \quad \mathrm{dist}_H(\mathrm{Reali}(\psi|_{\zeta=t}; \mathbb{R}), \lambda_\zeta(S)) < \epsilon.$$

All variables in this formula are in $\mathbb{R}$ and dist_H is the usual Hausdorff distance, which is written as a first order formula with coefficients in $\mathbb{R}$. Therefore, the same conclusion holds if $\mathbb{R}$ is replaced by a real closed extension F of it, by [31, Prop. 5.2.3]. □

We now generalize Proposition 5.33 to the case of multiple infinitesimals.

Theorem 5.34. *Let R be a real closed extension of $\mathbb{R}$. Let $S = \mathrm{Reali}(\psi; R\langle\zeta_1, \dots, \zeta_m\rangle) \subset R\langle\zeta_1, \dots, \zeta_m\rangle^n$ be a semialgebraic set, bounded over R. Then, for every $1 \leq k \leq m$ it holds*

$$\text{“}\lim_{t_k\to 0} \cdots \lim_{t_m\to 0} S|_{t_k,\dots,t_m} = \lambda_{\zeta_k} \cdots \lambda_{\zeta_m}(S)\text{”}, \tag{5.16}$$

where (5.16) means: for all $\epsilon \in R\langle\zeta_1, \dots, \zeta_{k-1}\rangle$, $\epsilon > 0$, for $0 < t_k \ll \cdots \ll t_m \ll 1$ it holds

$$\mathrm{dist}_H\Big(S|_{t_k,\dots,t_m}, \lambda_{\zeta_k} \cdots \lambda_{\zeta_m}(S)\Big) < \epsilon,$$

where dist_H is the semialgebraic Hausdorff distance (see Definition 5.28).

Proof. When $k = m$, the statement is given by Proposition 5.33 applied to the case $F = R\langle\zeta_1, \dots, \zeta_{m-1}\rangle$, so that $F\langle\zeta\rangle \simeq R\langle\zeta_1, \dots, \zeta_{m-1}, \zeta\rangle$, where $\zeta = \zeta_m$.

Assume that the statement is true for $k = j \le m$. We prove it for $k = j - 1$. By our working assumption, for every $\epsilon \in R\langle\zeta_1, \dots, \zeta_{j-1}\rangle$, $\epsilon > 0$, for $0 < t_j \ll t_{j+1} \ll \cdots \ll t_m \ll 1$ the following property holds:

$$\mathrm{dist}_H\left(S|_{t_j,\dots,t_m}, \lambda_{\zeta_j}\cdots\lambda_{\zeta_m}(S)\right) < \frac{\epsilon}{2}.$$

Recall that here $t_i \in R\langle\zeta_1, \dots, \zeta_{i-1}\rangle$ for every $j \le i \le m$. A fortiori we can take $\epsilon \in R\langle\zeta_1, \dots, \zeta_{j-2}\rangle$.

Denote by $A := S|_{t_j,\dots,t_m} \subset R\langle\zeta_1, \dots, \zeta_{j-1}\rangle^n$ and by $B := \lambda_{\zeta_j}\cdots\lambda_{\zeta_m}(S) \subset R\langle\zeta_1, \dots, \zeta_{j-1}\rangle^n$, which are both bounded over R. Given $\epsilon \in R\langle\zeta_1, \dots, \zeta_{j-2}\rangle$, $\epsilon > 0$ let $\delta_{j-1} \in R\langle\zeta_1, \dots, \zeta_{j-2}\rangle$, $\delta_{j-1} > 0$ be given by Proposition 5.33 such that for all $t_{j-1} \in R\langle\zeta_1, \dots, \zeta_{j-2}\rangle$ with $0 < t_{j-1} < \delta_{j-1}$ the following property holds:

$$\mathrm{dist}_H\left(A|_{t_{j-1}}, \lambda_{\zeta_{j-1}}(A)\right) < \frac{\epsilon}{2}. \tag{5.17}$$

Then, for $0 < t_{j-1} \ll t_j \ll \cdots \ll t_m \ll 1$ it holds

$$\begin{aligned}
\mathrm{dist}_H(S|_{t_{j-1},\dots,t_m}, \lambda_{\zeta_{j-1}}\cdots\lambda_{\zeta_m}(S)) &\overset{(5.13)}{\le} \mathrm{dist}_H(A|_{t_{j-1}}, \lambda_{\zeta_{j-1}}(A)) + \mathrm{dist}_H(\lambda_{\zeta_{j-1}}(A), \lambda_{\zeta_{j-1}}(B)),\\
&\overset{(5.17)}{<} \frac{\epsilon}{2} + \lambda_{\zeta_{j-1}}\left(\mathrm{dist}_H(A, B)\right),\\
&\overset{(5.14)}{\le} \frac{\epsilon}{2} + \lambda_{\zeta_{j-1}}\left(\frac{\epsilon}{2}\right) = \epsilon,
\end{aligned}$$

where in the last inequality we used the fact that ϵ was chosen in $R\langle\zeta_1, \dots, \zeta_{j-2}\rangle$. □

5.1.4 Hausdorff approximations of closed and bounded sets

Recall that every closed (in the Euclidean topology) semialgebraic set $S \subseteq R^n$ can be written as a finite union of closed basic semialgebraic sets ([31, Thm. 2.7.2]):

$$S = \bigcup_{i=1}^{a}\bigcap_{j=1}^{b_i}\{x \in R^n \mid p_{ij}(x) \le 0\}. \tag{5.18}$$

In general, for a closed semialgebraic set S with diagram $D(S) = (n, c, d)$, passing from a representation of the form (5.1) to one of the form (5.18), we cannot control the number of unions and intersections in (5.18) as a function of c, nor the degrees of the polynomials in (5.18) as a function of d. (On the other hand, in passing from (5.18) to (5.1) the process is controlled.)

However, here we use the tools from the previous section to show that, given a closed and bounded semialgebraic set with a representation as in (5.1), we can *approximate* it with a closed and bounded semialgebraic set described as in (5.18) in a controlled way (Proposition 5.36).

We start with the following elementary lemma.

Lemma 5.35. *Let $C \subset \mathbb{R}^n$ be of the form*

$$C = \left(\bigcap_{j\in J^=}\left\{x \in \mathbb{R}^n \,\middle|\, p_j(x) = 0\right\}\right) \cap \left(\bigcap_{j\in J^<}\left\{x \in \mathbb{R}^n \,\middle|\, q_j(x) < 0\right\}\right),$$

where the p_j and the q_j are polynomials and $J^=, J^<$ are finite sets. Then, denoting by (x,u) points in $\mathbb{R}^n \times \mathbb{R}$, and identifying $\{u=0\}$ with $\mathbb{R}^n$, the closure of C can be described as:

$$\overline{C} = \overline{\left\{(x,u) \,\middle|\, \Big(p_j(x)=0,\ \forall j \in J^=\Big) \wedge \Big(q_j(x)+u \le 0,\ \forall j \in J^<\Big) \wedge \Big(u>0\Big)\right\}} \cap \{u=0\}. \tag{5.19}$$

Proof. We prove the two inclusions separately. To prove the inclusion of the set on the left of (5.19) into the one on the right, by the monotonicity of the closure operation, it is enough to show that for any $z \in C$ there exists a sequence $\{(z_k,u_k)\}_{k\ge1}$ converging to $(z,0)$ satisfying for every $k \ge 1$

$$\Big(p_j(z_k)=0,\ \forall j \in J^=\Big) \wedge \Big(q_j(z_k)+u_k \le 0,\ \forall j \in J^<\Big) \wedge \Big(u_k>0\Big).$$

Since $z \in C$, we have

$$\Big(p_j(z)=0,\ \forall j \in J^=\Big) \wedge \Big(q_j(z)+U \le 0,\ \forall j \in J^<\Big),$$

with $U := -\max\{q_j(z) \,|\, j \in J^<\} > 0$. Hence, setting $u_k := \frac{U}{k} > 0$ and $z_k := z$, we observe that the sequence (z_k,u_k) satisfies all the above conditions and converges to $(z,0)$.

To prove the other inclusion, we take

$$(z,0) \in \overline{\left\{(x,u) \in \mathbb{R}^n \times \mathbb{R} \,\middle|\, \Big(p_j(x)=0,\ \forall j \in J^=\Big) \wedge \Big(q_j(x)+u \le 0,\ \forall j \in J^<\Big) \wedge \Big(u>0\Big)\right\}}.$$

Hence, there exists a sequence (z_k,u_k) converging to $(z,0)$ such that for any k

$$\Big(p_j(z_k)=0,\ \forall j \in J^=\Big) \wedge \Big(q_j(z_k)+u_k \le 0,\ \forall j \in J^<\Big) \wedge \Big(u_k>0\Big).$$

In particular $z_k \in C$ for all k, hence $z \in \overline{C}$. □

Proposition 5.36. *For every $\epsilon > 0$ and for every closed and bounded semialgebraic set $S \subset \mathbb{R}^n$ with diagram $D(S) = (n,c,d)$, there exists a closed semialgebraic set $S' \subset \mathbb{R}^n$ satisfying*

$$\mathrm{dist}_H(S,S') \le \epsilon$$

and such that

$$S' = \bigcup_{i=1}^{a} \mathrm{Bas}(\mathcal{P}_i, Q_i; \mathbb{R}),$$

with the property that, for every $i = 1,\dots,a$ we have $\#(\mathcal{P}_i \cup Q_i) \le b$, with $ab \le c$, and with each polynomial in $\mathcal{P}_i \cup Q_i$ of degree bounded by d.

Proof. Let us write

$$S = \bigcup_{i=1}^{a} \bigcap_{j=1}^{b_i} \left\{x \in \mathbb{R}^n \,\middle|\, \mathrm{sign}(p_{ij}(x)) = \sigma_{ij}\right\},$$

with $c = a \max_i\{b_i\}$. Let us examine the sets

$$C_i := \bigcap_{j=1}^{b_i} \left\{x \in \mathbb{R}^n \,\middle|\, \mathrm{sign}(p_{ij}(x)) = \sigma_{ij}\right\}.$$

By possibly relabelling the p_{ij} and multiplying them by ± 1, we can write each C_i as

$$C_i = \left(\bigcap_{j \in J_i^=} \left\{ x \in \mathbb{R}^n \,\middle|\, p_{ij}(x) = 0 \right\} \right) \cap \left(\bigcap_{j \in J_i^<} \left\{ x \in \mathbb{R}^n \,\middle|\, p_{ij}(x) < 0 \right\} \right),$$

where $\#\left(J_i^= \cup J_i^<\right) = b_i$. For each $i = 1, \dots, a$ and for every $i \in J_i^<$ we define the polynomial $\widetilde{p}_{ij} \in \mathbb{R}\langle\zeta\rangle[x_1, \dots, x_n]$ as follows:

$$\widetilde{p}_{ij}(x) := p_{ij}(x) + \zeta.$$

For any $i = 1, \dots, a$ we consider the semialgebraic set $\widetilde{S}_i \subset \mathbb{R}\langle\zeta\rangle^n$ defined by the formula:

$$\widetilde{\psi}_i := \left(\bigwedge_{j \in J_i^=} p_{ij}(x) = 0 \right) \wedge \left(\bigwedge_{j \in J_i^<} \widetilde{p}_{ij}(x) \leq 0 \right).$$

We claim that, if $S \subseteq B_{\mathbb{R}^n}(\rho)$ for $\rho > 0$, then $\widetilde{S}_i \subseteq B_{\mathbb{R}\langle\zeta\rangle^n}(\rho)$. In fact, let $g : (0, \delta) \to \mathbb{R}$ be a representative for an element of $\widetilde{S}_i$. Then, for every $\zeta \in (0, \delta)$ we have $p_{ij}(g(\zeta)) = 0$ for $j \in J_i^=$ and $p_{ij}(g(\zeta)) + \zeta \leq 0$ for $j \in J_i^<$. Since $\zeta < \delta$, the inequality $p_{ij}(g(\zeta)) + \zeta \leq 0$ implies $p_{ij}(g(\zeta)) < 0$, i.e. $g(\zeta) \in C_i \subseteq B_{\mathbb{R}^n}(\rho)$ for every $\zeta \in (0, \delta)$. Therefore $g \in B_{\mathbb{R}\langle\zeta\rangle^n}(\rho)$.

In particular $\widetilde{S}_i$ is bounded and, by Proposition 5.30, for any $\epsilon > 0$ we get $r > 0$ such that

$$\operatorname{dist}_H \left(\mathscr{L}_0(\widetilde{S}_i), \widetilde{S}_i|_r \right) \leq \epsilon. \tag{5.20}$$

We have by (5.12) and Lemma 5.35

$$\mathscr{L}_0(\widetilde{S}_i) = \overline{\left\{ (x, u) \in \mathbb{R}^{n+1} \,\middle|\, \left(x \in \widetilde{S}_i|_u \right) \wedge (u > 0) \right\}} \cap \mathbb{R}^n = \overline{C_i}.$$

Observe that $\widetilde{S}_i|_r = \operatorname{Bas}(\mathcal{P}_i, Q_i; \mathbb{R})$, where

$$\mathcal{P}_i = \left\{ p_{ij} \,\middle|\, j \in J_i^= \right\} \quad \text{and} \quad Q_i = \left\{ p_{ij} + r \,\middle|\, j \in J_i^< \right\}.$$

Note that $\#(\mathcal{P}_i \cup Q_i) \leq b := \max_i b_i$ so that $ab \leq c$, and each polynomial in $\mathcal{P}_i \cup Q_i$ has degree bounded by d. We now define

$$S' := \bigcup_{i=1}^{a} C_i', \qquad \text{with} \qquad C_i' := \operatorname{Bas}(\mathcal{P}_i, Q_i; \mathbb{R}).$$

Since S is closed, we have $S = \bigcup_{i=1}^a \overline{C_i}$. From (5.20) we get directly

$$\operatorname{dist}_H(S, S') \leq \max_i \operatorname{dist}_H \left(\overline{C_i}, C_i' \right) \leq \epsilon,$$

concluding the proof. □

5.2 Quantitative approximate definable choices

5.2.1 Preliminary constructions

Following a construction from [14], given a closed basic semialgebraic set

$$S = \mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R}) \subset \mathbb{R}^n$$

we construct a closed basic semialgebraic set $\widetilde{S} \subset R^n$, where R is a field of algebraic Puiseux series with coefficients in $\mathbb{R}$, such that $\mathscr{L}_0(\widetilde{S}) = S$ and for which a definable choice over a given projection can be made quantitatively.

Construction 5.37. The following construction is taken from [14]. Set

$$R = \mathbb{R}\langle \zeta_1, \zeta_2, \zeta_3 \rangle.$$

For a given closed basic and bounded semialgebraic set

$$S = \mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R}) \subset \mathbb{R}^n,$$

and for every $1 \leq k \leq n$ the construction provides new semialgebraic sets

$$\widetilde{S}^k \subseteq \widetilde{S} \subset R^n.$$

Assume that the polynomials in $\mathcal{P}, \mathcal{Q}$ have degree bounded by d. First, using a perturbation argument, one constructs new families of polynomials $\widetilde{\mathcal{P}}, \widetilde{\mathcal{Q}} \subset R[x_1, \ldots, x_n]$ with degrees bounded by $2d + 2$. The cardinality of $\widetilde{\mathcal{Q}}$ equals the one of $\mathcal{Q}$, whereas (for our purposes) the cardinality of $\widetilde{\mathcal{P}}$ can be assumed to be 1 (by taking the sum of the squares of its elements). In this way we get the set

$$\widetilde{S} := \mathrm{Bas}(\widetilde{\mathcal{P}}, \widetilde{\mathcal{Q}}; R).$$

Then, one constructs $\widetilde{S}^k$. First, set

$$g(x) := 1 + \sum_{j=1}^{n} j x_j^{2d+2},$$

and for every family $\mathcal{F} = \{f_1, \ldots, f_s\} \subset R[x_1, \ldots, x_n]$ define the algebraic set:

$$\mathrm{Crit}_k(\mathcal{F}; R) := \{x \in R^n \mid f_1(x) = \cdots = f_s(x) = 0,\ \mathrm{rank}(\widetilde{J}(x)) \leq k\}, \tag{5.21}$$

where $\widetilde{J}(x)$ is the matrix of size $(n-k) \times (s+1)$ whose columns are the partial derivatives of $f_1, \ldots, f_s, g$ with respect to $x_{k+1}, \ldots, x_n$. The set $\widetilde{S}^k$ is then defined by

$$\widetilde{S}^k := \bigcup_{\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}}} \mathrm{Crit}_k(\widetilde{\mathcal{P}} \cup \widetilde{\mathcal{Q}}'; R) \cap \widetilde{S}, \tag{5.22}$$

where the union runs over all subsets $\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}} \subset R[x_1, \ldots, x_n]$.

Remark 5.38 (On boundedness). Even though this is not explicitly stated in [14] (but indeed used in their arguments), by inspecting the form of $\widetilde{\mathcal{P}}$ and $\widetilde{\mathcal{Q}}$ provided there, one can see that, if $S \subset \mathbb{R}^n$ is bounded, then $\widetilde{S}^\ell \subseteq \widetilde{S} \subset R^n$ are bounded over $\mathbb{R}$.

The next result gives a control on the diagram of $\widetilde{S}^k$.

Lemma 5.39. *For every $a_1 \in \mathbb{N}$ there exist $a_2, a_3 \in \mathbb{N}$ such that, for a closed basic semialgebraic set $S = \mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R}) \subset \mathbb{R}^n$, with $\#\mathcal{P}, \#\mathcal{Q} \le a_1$ and with every polynomial in $\mathcal{P} \cup \mathcal{Q}$ of degree bounded by d, and any $1 \le k \le n$, it holds*

$$D(\widetilde{S}^k) = (n, a_2, a_3 d),$$

where $\widetilde{S}^k \subset R^n$ is the set associated with S by Construction 5.37.

Proof. Let $\widetilde{\mathcal{P}}, \widetilde{\mathcal{Q}} \subset R[x_1, \dots, x_n]$ be the family of polynomials of Construction 5.37 such that $\widetilde{S} = \mathrm{Bas}(\widetilde{\mathcal{P}}, \widetilde{\mathcal{Q}}; R)$. These polynomials have degree bounded by $2d + 2$.

Then, for every subset $\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}}$ (there are only finitely many such subsets, say at most $a_4 = a_4(a_1)$) one considers the family $\mathcal{F} = \widetilde{\mathcal{P}} \cup \widetilde{\mathcal{Q}}' = \{f_1, \dots, f_s\}$ (here $s \le a_4 + 1$) and defines the algebraic set $\mathrm{Crit}_k(\mathcal{F})$ as in (5.21). Since R is real closed, this algebraic set is defined by a single polynomial equation $P_{\widetilde{\mathcal{Q}}'}(x) = 0$, where

$$P_{\widetilde{\mathcal{Q}}'} = \sum_{i=1}^{s} f_i^2 + \sum m_{ij}^2.$$

Here, the second sum runs over all the $(k+1) \times (k+1)$ minors of $\widetilde{J}(x)$. Remember that, $\widetilde{J}(x)$ is a $(n-k) \times (s+1)$ matrix, hence if $k > s$ there are no such minors, so we can assume that $k \le s$. Therefore the polynomial $P_{\widetilde{\mathcal{Q}}'}$ has degree

$$\deg\left(P_{\widetilde{\mathcal{Q}}'}\right) \le (k+1)(4d+4) \le (s+1)(4d+4) \le a_5 d,$$

where $a_5 = a_5(a_1)$.

The set $\widetilde{S}^\ell$ from Construction 5.37 can therefore be written as:

$$\begin{aligned} \widetilde{S}^\ell &= \bigcup_{\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}}} \mathrm{Crit}_k(\widetilde{\mathcal{P}} \cup \widetilde{\mathcal{Q}}'; R) \cap \widetilde{S} = \widetilde{S} \cap \bigcup_{\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}}} Z(P_{\widetilde{\mathcal{Q}}'}; R) = \widetilde{S} \cap Z(\widetilde{F}; R) \\ &= Z(\widetilde{\mathcal{P}}; R) \bigcap_{\tilde{q} \in \widetilde{\mathcal{Q}}} \{\tilde{q} \le 0\} \cap Z(\widetilde{F}; R), \end{aligned} \tag{5.23}$$

where $\widetilde{\mathcal{P}}, \widetilde{\mathcal{Q}}$ are the families defining the basic set $\widetilde{S}$, and

$$\widetilde{F} := \prod_{\widetilde{\mathcal{Q}}' \subset \widetilde{\mathcal{Q}}} P_{\widetilde{\mathcal{Q}}'},$$

which is a polynomial of degree bounded by $a_4 a_5 d = a_3 d$, with a_3 depending only on a_1. To obtain a presentation as in (5.1) note that if $\{p_1, \dots, p_L\}$ is a family of polynomials, then

$$\bigcap_{i=1}^{L} \{p_i \le 0\} = \bigcap_{i=1}^{L} \Big(\{p_i < 0\} \cup \{p_i = 0\}\Big) = \bigcup_{\sigma} \bigcap_{i=1}^{L} \{\mathrm{sign}(p_i) = \sigma_i\}, \tag{5.24}$$

where σ runs over all possible choices in $\{0, -1\}^L$. We can apply identity (5.24) to (5.23). Since $\#\widetilde{\mathcal{P}}, \#\widetilde{\mathcal{Q}} \le a_1$, we get that $\widetilde{S}^\ell$ admits a representation as in (5.1) with $a = 2^L$, $b_i = L$, for $L = 2a_1 + 1$. In other words $D(\widetilde{S}^\ell) = (n, a_2, a_3 d)$, for $a_2 = (2a_1 + 1)2^{2a_1+1}$. □

The sets from Construction 5.37 enjoy the following properties. For $1 \leq k \leq n$ let

$$\pi_k : R^n \to R^k$$

be the projection onto the last k coordinates. Since $\mathbb{R} \subset R = \mathbb{R}\langle\zeta_1, \zeta_2, \zeta_3\rangle$, the same symbol $\pi_k : \mathbb{R}^n \to \mathbb{R}^k$ is used to denote the restriction to $\mathbb{R}^n$, without risk of confusion.

Proposition 5.40. *Let $S = \mathrm{Bas}(\mathcal{P}, Q; \mathbb{R}) \subset \mathbb{R}^n$ be a closed basic and bounded semialgebraic set such that, for some $1 \leq k \leq n$ it holds*

$$\forall y \in \mathbb{R}^k \quad \#\left(Z(\mathcal{P}; \mathbb{R}) \cap \pi_k^{-1}(y)\right) < \infty, \tag{5.25}$$

i.e. $Z(\mathcal{P}; \mathbb{R})$ is strongly of dimension $\leq k$. Then for any $\ell < k$, the sets $\widetilde{S}^\ell \subseteq \widetilde{S} \subset \mathbb{R}\langle\zeta_1, \zeta_2, \zeta_3\rangle^n$ from Construction 5.37 are closed and bounded over $\mathbb{R}$. Moreover they satisfy the following properties. For every $w \in \mathbb{R}\langle\zeta_1, \zeta_2, \zeta_3\rangle^\ell$

(i) the set $\widetilde{S}^\ell \cap \pi_\ell^{-1}(w)$ is finite, i.e. $\widetilde{S}^\ell$ is strongly of dimension $\leq \ell$;

(ii) if $\widetilde{S} \cap \pi_\ell^{-1}(w) \neq \emptyset$, the set $\widetilde{S}^\ell \cap \pi_\ell^{-1}(w)$ intersects every connected component of $\widetilde{S} \cap \pi_\ell^{-1}(w)$;

(iii) it holds $\mathscr{L}_0(\widetilde{S}) = S$.

Furthermore, for every $\epsilon \in \mathbb{R}$, $\epsilon > 0$, for $0 < t_1 \ll \cdots \ll t_m \ll 1$ the evaluation $S|_{t_1,\dots,t_m}$ is defined and the following properties hold:

(iv) $\mathrm{dist}_H(\widetilde{S}|_{t_1,\dots,t_m}, S) < \epsilon$;

(v) $\mathrm{dist}_H\left(\pi_\ell(\widetilde{S}|_{t_1,\dots,t_m}), \pi_\ell(S)\right) < \epsilon$.

Proof. The fact that $\widetilde{S}^\ell \subseteq \widetilde{S} \subset \mathbb{R}\langle\zeta_1, \zeta_2, \zeta_3\rangle^n$ are closed, and bounded follows directly by their construction. Items (i) to (iii) are proved in [14, Prop. 5.5 and Prop. 5.17], while Item (iv) follows from Item (iii) and Theorem 5.34. To conclude, observe that $\pi_\ell : \mathbb{R}^n \to \mathbb{R}^\ell$ is 1-Lipschitz with respect to dist_H, therefore Item (v) follows from Item (iv). □

5.2.2 Approximate definable choice: the case of a projection

The purpose of this section is to prove Theorem L. We begin with the following preliminary version of the latter, for the case of closed basic sets.

Proposition 5.41. *Let $\pi : \mathbb{R}^n \to \mathbb{R}^\ell$, with $1 \leq \ell \leq n$ be the projection onto the last ℓ coordinates. For every $a_1 \in \mathbb{N}$ there exist $a_2, a_3 \in \mathbb{N}$ such that the following holds. For every closed basic and bounded semialgebraic set*

$$S = \mathrm{Bas}(\mathcal{P}, Q; \mathbb{R}) \subset \mathbb{R}^n,$$

with $\#\mathcal{P}, \#Q \leq a_1$ and with every polynomial in $\mathcal{P} \cup Q$ of degree bounded by d, and for all $\epsilon > 0$ there exists a semialgebraic set $A_\epsilon \subset \mathbb{R}^n$ satisfying the following properties:

(i) $\dim(A_\epsilon) \leq \ell$;

(ii) $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$;

(iii) $\mathrm{dist}_H(\pi(A_\epsilon), \pi(S)) \leq \epsilon$;

(iv) $D(A_\epsilon) = (n, a_2, a_3 d)$.

Proof. If $\dim(S) \leq \ell$ and one can simply take $A_\epsilon = S$.

Assume then that $k = \dim(S) > \ell$. Let $\widetilde{S}^\ell \subseteq \widetilde{S} \subset R^n$, where $R = \mathbb{R}\langle\zeta_1, \zeta_2, \zeta_3\rangle$ be the set defined by Construction 5.37. The strategy is to define

$$A_\epsilon := \widetilde{S}^\ell|_{t_1,t_2,t_3} \tag{5.26}$$

for an appropriate choice of "sufficiently small" (t_1, t_2, t_3) and to use Proposition 5.40. However, in order to use it we need first to ensure that the condition (5.25) is satisfied. In order to do it, using Proposition 5.11, we can first perform a small linear change of variables $L : \mathbb{R}^n \to \mathbb{R}^n$ (i.e. we choose L sufficiently close to the identity) in the defining polynomials and get new families $\mathcal{P}', \mathcal{Q}'$ with the property that: $\#\mathcal{P}' = \#\mathcal{P}$, $\#\mathcal{Q}' = \#\mathcal{Q}$; the degrees of all elements of $\mathcal{P}', \mathcal{Q}'$ are bounded by d; the Hausdorff distance between $\mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R})$ and $\mathrm{Bas}(\mathcal{P}', \mathcal{Q}'; \mathbb{R}) = L(\mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R}))$ is arbitrarily small; and condition (5.25) is satisfied.

It is elementary to verify that a set $A'_\epsilon \subset \mathbb{R}^n$ satisfying the properties of the statement for the semialgebraic set $\mathrm{Bas}(\mathcal{P}', \mathcal{Q}'; \mathbb{R})$ will also satisfy the same properties for the original set $\mathrm{Bas}(\mathcal{P}, \mathcal{Q}; \mathbb{R})$, up to adjusting ϵ. Therefore, without loss of generality, we assume that (5.25) is verified for the set S.

Set A_ϵ as in (5.26). We show how to chose (t_1, t_2, t_3) so that the desired properties hold.

(i) By Item (i) of Proposition 5.40, the set $\widetilde{S}^\ell$ is strongly of dimension $\leq \ell$. By Lemma 5.20 for $0 < t_1 \ll t_2 \ll t_3 \ll 1$ also $\widetilde{S}^\ell|_{t_1,t_2,t_3}$ is strongly of dimension $\leq \ell$.

(ii) By Item (iv) of Proposition 5.40, for $0 < t_1 \ll t_2 \ll t_3 \ll 1$ we have

$$\mathrm{dist}_H(\widetilde{S}|_{t_1,t_2,t_3}, S) < \epsilon.$$

Since $\widetilde{S}^\ell \subseteq \widetilde{S}$ (in fact the formula defining $\widetilde{S}^\ell$ contains the formula defining $\widetilde{S}$ as a conjunction, by definition (5.22)), for such (t_1, t_2, t_3) we also have $\widetilde{S}^\ell|_{t_1,t_2,t_3} \subseteq \mathcal{U}_\epsilon(S)$.

(iii) By Item (v) of Proposition 5.40, for $0 < t_1 \ll t_2 \ll t_3 \ll 1$ we have

$$\mathrm{dist}_H(\pi(\widetilde{S}|_{t_1,t_2,t_3}), \pi(S)) < \epsilon.$$

(iv) By Lemma 5.18 and Lemma 5.39, for $0 < t_1 \ll t_2 \ll t_3 \ll 1$ we have

$$D(\widetilde{S}^\ell|_{t_1,t_2,t_3}) = D(\widetilde{S}^\ell) = (n, a_2, a_3 d),$$

where $a_2, a_3 \in \mathbb{N}$ depend only on a_1.

Since we are requiring only a finite number of properties, given $\epsilon > 0$, the quantifier $0 < t_1 \ll t_2 \ll t_3 \ll 1$ can be commonly chosen (see Definition 5.13) so that the set (5.26) satisfies simultaneously Items (i) to (iv). □

Extending the previous result to general (non-basic) semialgebraic sets, we obtain the following statement, that corresponds to Theorem L.

Theorem 5.42. *For every $c \in \mathbb{N}$ there exist $\kappa \in \mathbb{N}$ such that the following holds. Let $n, \ell, d \in \ell$, with $1 \leq \ell \leq n$. Let $\pi : \mathbb{R}^n \to \mathbb{R}^\ell$ be the projection onto the last ℓ coordinates and let $S \subset \mathbb{R}^n$ be a bounded closed semialgebraic set with*

$$D(S) = (n, c, d).$$

Then, for every $\epsilon > 0$ there exists a closed semialgebraic set $A_\epsilon \subset \mathbb{R}^n$ such that:

(i) $\dim(A_\epsilon) \leq \ell$;

(ii) $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$;

(iii) $\operatorname{dist}_H(\pi(A_\epsilon), \pi(S)) \leq \epsilon$;

(iv) $D(A_\epsilon) = (n, \kappa, \kappa d)$.

Proof. Applying Proposition 5.36 we find a closed semialgebraic set $S' \subset \mathbb{R}^n$ satisfying

$$\operatorname{dist}_H(S, S') \leq \frac{\epsilon}{2}, \tag{5.27}$$

(in particular, S' is also bounded) and such that

$$S' = \bigcup_{i=1}^{a} \operatorname{Bas}(\mathcal{P}_i, \mathcal{Q}_i; \mathbb{R}),$$

with the property that, for every $i = 1, \ldots, a$ we have $\#(\mathcal{P}_i \cup \mathcal{Q}_i) \leq b$, with $ab \leq c$, and with each polynomial in $\mathcal{P}_i \cup \mathcal{Q}_i$ of degree bounded by d. For every $i = 1, \ldots, a$, denote by $S'_i := \operatorname{Bas}(\mathcal{P}_i, \mathcal{Q}_i; \mathbb{R})$ and apply Proposition 5.41 to each S'_i to get sets $A_{i,\epsilon}$ satisfying the conclusions of Proposition 5.41 (with $\varepsilon/2$ in place of ε). Define

$$A_\epsilon := \bigcup_{i=1}^{a} A_{i,\epsilon}. \tag{5.28}$$

(i) Since for every $i = 1, \ldots, a$ we have $\dim(A_{i,\epsilon}) \leq \ell$, Item (i) follows.

(ii) As for Item (ii), this follows from the fact that $A_\epsilon \subseteq \mathcal{U}_{\frac{\epsilon}{2}}(S')$ and from (5.27).

(iii) Similarly, for Item (iii), we have

$$\operatorname{dist}_H(\pi(A_\epsilon), \pi(S)) \leq \operatorname{dist}_H(\pi(A_\epsilon), \pi(S')) + \operatorname{dist}_H(\pi(S'), \pi(S)) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq \epsilon.$$

(iv) By Item (iv) from Proposition 5.41, $D(A_{i,\epsilon}) = (n, a_2(b), a_3(b)d)$. Therefore $D(A_\epsilon) = (n, \kappa', \kappa' d)$ for some $\kappa' = \kappa'(a_2(b), a_3(b), a) \leq \kappa(c)$. This proves Item (iv).

The set A_ϵ defined by (5.28) satisfies Items (i) to (iv). □

5.2.3 Approximate definable choice: the case of a semialgebraic map

In this section we prove Theorem M, which we restate here.

Theorem 5.43. *For every $c, d, \ell \in \mathbb{N}$ there exists $\beta > 1$ satisfying the following statement. Let $n \in \mathbb{N}$ and let $K \subset \mathbb{R}^n$ be a closed semialgebraic set contained in the ball $B_{\mathbb{R}^n}(\rho)$ and let $F : \mathbb{R}^n \to \mathbb{R}^\ell$ be a locally Lipschitz semialgebraic map such that*

$$D(\mathrm{graph}(F|_K)) = (n + \ell, c, d).$$

Then for every $\epsilon \in (0, \rho)$ there exists a closed semialgebraic set $C_\epsilon \subset \mathbb{R}^n$ such that:

(i) $\dim(C_\epsilon) \leq \ell$;

(ii) $C_\epsilon \subseteq \mathcal{U}_\epsilon(K)$;

(iii) $\mathrm{dist}_H(F(C_\epsilon), F(K)) \leq L(F, \rho) \cdot \epsilon$, *where* $L(F, \rho) := 2 + \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho))$;

(iv) for every $e = 1, \dots, n$ and every affine space $\mathbb{R}^e \simeq E \subseteq \mathbb{R}^n$, the number of connected components of $E \cap C_\epsilon$ is bounded by

$$b_0(E \cap C_\epsilon) \leq \beta^e.$$

Proof. Let $\mathbb{R}^{n+\ell} = \mathbb{R}^n \times \mathbb{R}^\ell$ and denote by $\pi_1 : \mathbb{R}^{n+\ell} \to \mathbb{R}^n$ and $\pi_2 : \mathbb{R}^{n+\ell} \to \mathbb{R}^\ell$ the projections on the two factors.

Given $\epsilon > 0$, let $A_\epsilon \subset \mathbb{R}^{n+\ell}$ be the semialgebraic set obtained by applying Theorem 5.42 to $S = \mathrm{graph}(F|_K) \subset \mathbb{R}^{n+\ell}$, nothing that by assumption the latter is a bounded and closed semialgebraic set with $D(S) = (n + \ell, c, d)$. We define

$$C_\epsilon := \pi_1(A_\epsilon).$$

This set is semialgebraic and closed (it is a continuous image of a compact semialgebraic set).

We verify that the set C_ϵ has the desired properties.

(i) By Item (i) of Theorem 5.42, $\dim(A_\epsilon) \leq \ell$. Therefore, by (5.2), $\dim(C_\epsilon) \leq \ell$.

(ii) Let $x \in C_\epsilon$. Then, by Item (ii) of Theorem 5.42, there exists y such that $x = \pi_1(x, y)$ with $(x, y) \in A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$. Therefore there exists $(x', y') \in S$ such that

$$\|x - x'\| + \|y - y'\| \leq \epsilon.$$

This implies that $\|x - x'\| \leq \epsilon$ and, since $x' = \pi_1(x', y') \in \pi_1(S) = K$, we get $x \in \mathcal{U}_\epsilon(K)$. Since this is true for all $x \in C_\epsilon$, then $C_\epsilon \subseteq \mathcal{U}_\epsilon(K)$.

(iii) Recall that we have set $L(F, \rho) = 2 + \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho)) =: L$. We need to prove the two inclusions $F(C_\epsilon) \subseteq \mathcal{U}_{L\epsilon}(F(K))$ and $F(K) \subseteq \mathcal{U}_{L\epsilon}(F(C_\epsilon))$.

We first prove that $F(C_\epsilon) \subseteq \mathcal{U}_{L\epsilon}(F(K))$. Pick $c \in C_\epsilon$, then there exists $s \in A_\epsilon$ such that $c = \pi_1(s)$. Since $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$ (by Item (ii) of Theorem 5.42), there exists $z \in S$ such that $\|z - s\| \leq \epsilon$. Now, $\pi_1(z) \in K$ and we have $\|c - \pi_1(z)\| = \|\pi_1(s) - \pi_1(z)\| \leq \|s - z\| \leq \epsilon$. Hence, for $\epsilon \in (0, \rho)$ we have

$$\begin{aligned}\|F(c) - F(\pi_1(z))\| &\leq \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho))\,\|c - \pi_1(z)\| \\ &\leq \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho))\epsilon \leq L(F, \rho)\epsilon.\end{aligned}$$

This proves that $F(C_\epsilon) \subseteq \mathcal{U}_{L\epsilon}(F(K))$.

To prove the other inclusion we fix $v \in K$. By Item (iii) of Theorem 5.42 we have

$$\mathrm{dist}_H(\pi_2(A_\epsilon), F(K)) \leq \epsilon.$$

Hence, there exists $(x, y) \in A_\epsilon$ such that $\|F(v) - \pi_2((x, y))\| = \|F(v) - y\| \leq \epsilon$. We observe that for $\epsilon \in (0, \rho)$ we have $x \in C_\epsilon \subseteq \mathcal{U}_\epsilon(K) \subseteq B_{\mathbb{R}^n}(2\rho)$. To conclude the proof we estimate $\|F(v) - F(x)\|$. Since $A_\epsilon \subseteq \mathcal{U}_\epsilon(S)$, we find $(w, F(w)) \in S$ such that $\|(x, y) - (w, F(w))\| \leq \epsilon$, and in particular this implies $\|y - F(w)\| \leq \epsilon$ and $\|x - w\| \leq \epsilon$. Hence for $\epsilon \in (0, \rho)$ we have

$$\begin{aligned}\|F(v) - F(x)\| &\leq \|F(v) - y\| + \|y - F(w)\| + \|F(w) - F(x)\| \\ &\leq (2 + \mathrm{Lip}(F, B_{\mathbb{R}^n}(2\rho)))\epsilon = L(F, \rho)\epsilon.\end{aligned}$$

This proves that $F(K) \subseteq \mathcal{U}_{L\epsilon}(F(C_\epsilon))$.

(iv) Let now $e \in \mathbb{N}$ with $e \leq n$ and $E \simeq \mathbb{R}^e \subseteq \mathbb{R}^n$. Observe first that $E \cap C_\epsilon$ is the projection on $E \simeq \mathbb{R}^e$ of $(E \times \mathbb{R}^\ell) \cap A_\epsilon$.

Note that $D((E \times \mathbb{R}^\ell) \cap A_\epsilon) = (e + \ell, \kappa, \kappa d)$, where $\kappa = \kappa(c)$ is the constant of Item (iv) from Theorem 5.42. Therefore, the Thom–Milnor bound (5.5) yields

$$b_0(E \cap C_\epsilon) \leq b_0((E \times \mathbb{R}^\ell) \cap A_\epsilon) \leq \beta_{\mathrm{tm}}(2\kappa)(\kappa d)^{e+\ell} \leq \beta(c, d, \ell)^e,$$

for suitable $\beta = \beta(c, d, \ell) > 1$.

The proof is concluded. □

Part II

Vanishing geodesic distances and the Michor-Mumford conjecture in Hilbertian H-type groups

Introduction

iii Motivation

It is a well known fact that in a connected and finite dimensional Riemannian manifold taking the infimum among all lengths of curves connecting two points yields a distance, called geodesic distance. Following the same procedure, geodesic distances may be defined also on infinite dimensional manifolds. A recent introduction to infinite dimensional geometry and analysis can be found in the monograph [87] and the lecture notes [33, 69]. Some foundational works are [44, 55], and more specific information on infinite dimensional Lie groups can be found for instance in [79, 53] and in the surveys [75, 78, 42].

A new aspect in infinite dimensional geometry is that the geodesic distance may not define a genuine distance on the manifold, since it can be zero on distinct points. We call them *degenerate* geodesic distances. In general, the degeneracy of the geodesic distance may occur for certain Riemannian metrics, where no special conditions are assumed, namely for *weak Riemannian metrics*, [1, Definition 5.2.12]. These metrics are important, since they are the only possible metrics when the manifold is not modelled on a Hilbert space. Vanishing geodesic distances in Fréchet manifolds were first found in [38], [71], [72] and further examples were studied in [16], [17], [18], [46], [47] and [21]. A simple example of vanishing geodesic distance can be also constructed in a Hilbert manifold, [65].

Despite many known examples of this phenomenon, the geometrical reasons allowing the degeneracy of geodesic distances are still not well understood. In this direction, P. Michor and D. Mumford conjectured a relationship between the vanishing of the geodesic distance and the local unboundedness of the sectional curvature, see [71]. They proposed a fascinating interpretation behind this phenomenon: some parts of the infinite dimensional manifold "wrap up on theirselves" allowing to find curves connecting two distinct points and having shorter and shorter length, up to reaching vanishing infimum, see [71] and [21, Section 1.2]. In Mathematical terms, we may rephrase this phenomenon as follows: whenever a weak Riemannian metric admits a vanishing geodesic distance, then the sectional curvature must be unbounded on some special sequences of planes that stay in a neighborhood of some point. For infinite dimensional Lie groups, the homogeneity by translations allows this point to be the unit element.

In this part of the thesis we provide new degenerate geodesic distances for left-invariant weak Riemannian metrics in Heisenberg-type Lie groups modelled on Hilbert spaces, and we confirm the validity of the Michor-Mumford conjecture in this setting. We give here a quick overview of chapters and main results, a more detailed presentation is given in the next sections of this introduction.

In Chapter 6 we construct a left invariant weak Riemannian metric in the infinite dimensional Heisenberg group, with a degenerate geodesic distance (the same construction yields a degenerate sub-Riemannian distance). Then, we show how the standard notion of sectional curvature adapts to this framework, but it cannot be defined everywhere. We prove that the sectional curvature is unbounded on suitable sequences of planes, along which the vanishing of the distance precisely occurs. Hence, the degenerate Riemannian distance for this Riemannian metric appears in connection with an unbounded sectional curvature, confirming the phenomenon predicted by Michor and Mumford. These results have been obtained in [66], a more detailed introdution is given in Section iv.

In Chapter 7 we introduce Hilbertian H-type groups equipped with weak, graded, left invariant Riemannian metrics. In particular, this class contains the infinite dimensional Heisenberg group introduced in Chapter 6. For these Lie groups, we prove that degenerate geodesic distances appear for a large class of weak, left invariant Riemannian metrics. Their vanishing is rather surprisingly related to the infinite dimensional sub-Riemannian structure of Hilbertian H-type groups. We show that the vanishing of the geodesic distance and the local unboundedness of the sectional curvature coexist, as conjectured by Michor and Mumford. We also prove that the same class of weak Riemannian metrics yields the nonexistence of the Levi-Civita covariant derivative. These results have been obtained in [67], a more detailed introduction is given in Section v.

iv On the Michor-Mumford phenomenon in the infinite dimensional Heisenberg group

In order to introduce the content of Chapter 6, we observe that despite many examples of vanishing geodesic distances are known, one may still wonder whether replacing a weak Riemannian metric with a left invariant weak Riemannian metric with respect to a Hilbert Lie group structure might give a condition to have positive geodesic distance on distinct points. The answer to this question does not seem intuitively clear. For instance, we observe that connected, simply connected and commutative Banach Lie groups, equipped with a bi-invariant weak Riemannian metric have positive geodesic distance on distinct points. In short, their geodesic distance is actually a distance. The proof of this fact essentially follows from [78, Proposition IV.2.7], observing that the exponential mapping is a local Riemannian isometry.

Thus, the question is whether considering a left invariant weak Riemannian metric on a noncommutative, connected and simply connected Banach Lie group may prevent the vanishing of the geodesic distance. Our first result answers this question in the negative.

Theorem N. *There exists a left invariant weak Riemannian metric on the infinite dimensional Heisenberg group $\mathbb{H}$, whose Riemannian distance is not positive on all couples of distinct points.*

In analogy with the finite dimensional case, the infinite dimensional Heisenberg group $\mathbb{H}$ can be defined starting from the Heisenberg Lie algebra, due to the Baker–Campbell–Hausdorff formula. We use $\ell^2 \times \ell^2 \times \mathbb{R}$ to model $\mathbb{H}$ as a Hilbert manifold, where ℓ^2 denotes the standard linear space of real square-summable sequences. The Lie product associated to $\mathbb{H}$ is defined in (6.1). More information on the Lie group $\mathbb{H}$ is provided in Section 6.1.1.

Sub-Riemannian distances naturally appears also in infinite dimensional manifolds. We mention for instance [37, Theorem C.2], where strong sub-Riemannian metrics were considered in connection with Wiener spaces, lower bounds on the Ricci curvature and logarithmic Sobolev inequalities. In this case, the space of admissible velocities is strictly contained in the tangent space and we have an a priori smaller family of connecting curves. The possible vanishing of a sub-Riemannian distance between some distinct points of an in infinite dimensional manifold was explicitly mentioned in [43, Remark 2]. The following theorem seemingly provides a first example of such vanishing phenomenon for an infinite dimensional sub-Riemannian manifold.

Theorem O. *There exists a left invariant weak sub-Riemannian metric on the infinite dimensional Heisenberg group* $\mathbb{H}$ *such that its associated sub-Riemannian distance is not positive on all couples of distinct points.*

The sub-Riemannian metric and the sub-Riemannian distance that we consider on $\mathbb{H}$ are described in Section 6.1.3. Both Theorem N and Theorem O are contained in Theorem 6.3 and their proof rather surprisingly relies on the same sequence of length-minimizing curves. The proof of these results shows that both the Riemannian and sub-Riemannian distance are vanishing between points that have the same projection on the subspace $\ell^2 \times \ell^2 \times \{0\}$. Remark 6.2 completes the picture, showing that when the projections of two points on $\ell^2 \times \ell^2 \times \{0\}$ are different, then both their Riemannian and sub-Riemannian distance are positive. In sum, all distinct points with vanishing geodesic distance are characterized.

From another perspective, dealing with a left invariant weak Riemannian metric has the advantage to find the sectional curvature by more manageable formulas. Motivated by the Michor-Mumford conjecture, in Theorem P below we prove that the sectional curvature of the left invariant weak Riemannian metric σ defined in (6.9) is unbounded.

From the standard formula for the sectional curvature of Lie groups, see for instance [10] and [20], the sectional curvature of $\mathbb{H}$ with respect to σ can be defined on "many planes" of the Lie algebra Lie($\mathbb{H}$). We wish to point out that for general weak Riemannian metrics the existence of the Levi–Civita (and then of the sectional curvature) is not guaranteed a priori. An example of this fact can be found in [19], where more information on the problem is available. We also observe that the "finite dimensional formula" for the sectional curvature through the structure coefficients of Lie($\mathbb{H}$), [74, Lemma 1.1], converges on the previous planes to the same sectional curvature obtained by [10, Theorem 5]. Broadly speaking, we may think of the convergence of the sectional curvature in Milnor's paper [74] as a computation of the sectional curvature of $\mathbb{H}$ through a finite dimensional approximation by an orthonormal basis. On the other side, we also observe that this convergence does not hold on all 2-dimensional subspaces of Lie($\mathbb{H}$), as shown in Remark 6.4.

The next theorem shows that the sectional curvature with respect to the weak Riemannian metric σ is unbounded on a certain sequence of planes.

Theorem P. *Let* $\mathbb{H}$ *be the infinite dimensional Heisenberg group equipped with the left invariant weak Riemannian metric* σ*. Then there exists two sequences of orthonormal vectors* $a_{1j}, a_{2j} \in \mathrm{Lie}(\mathbb{H})$ *and* $b \in \mathrm{Lie}(\mathbb{H})$ *with* $j \geq 1$ *such that* $K_\sigma(a_{1j}, b) = K_\sigma(a_{2j}, b)$,

$$\lim_{j\to\infty} K_\sigma(a_{1j}, a_{2j}) = -\infty \quad \text{and} \quad \lim_{j\to\infty} K_\sigma(a_{1j}, b) = +\infty.$$

The numbers $K_\sigma(a_{1j}, a_{2j})$ *and* $K_\sigma(a_{1j}, b)$ *are the sectional curvatures of the planes of* Lie($\mathbb{H}$) *spanned by the orthonormal bases* (a_{1j}, a_{2j}) *and* (a_{1j}, b).

The proof of this theorem is provided in Section 6.3, where also more information on the vectors a_{1j}, a_{2j} and b can be found. Inspecting the proofs of Theorem 6.3 and Theorem P another interesting phenomenon appears. The curves whose lengths converge to zero and that connect two distinct points are precisely contained in the *span of the planes* where the sectional curvature blows-up. We finally mention Proposition 6.5, where we prove that the sectional curvature is unbounded also on some sequences of converging planes.

v The Michor-Mumford conjecture in Hilbertian H-type groups

In order to introduce the content of Chapter 7, we begin by highlightening Riemannian Hilbert manifolds, that constitute a class of infinite dimensional manifolds modeled on a Hilbert space. Their Riemannian metrics induce the manifold topology on tangent spaces, hence they are called *strong Riemannian metrics* and their geodesic distance is a distance function, that separates points. While the local geometry of Riemannian Hilbert manifolds shares some analogies with the finite dimensional case, for instance the Levi-Civita connection always exists, certain global properties fail to hold. For more information, we refer the reader to the interesting survey [27] and the references therein. Only for infinite dimensional manifolds, the so-called *weak Riemannian metrics* may induce on tangent spaces weaker topologies than the manifold topology, [1, 33, 69, 87]. If the manifold is not modeled on a Hilbert space, then the Riemannian metric must be weak. We focus our attention on Lie groups modeled on a Hilbert space, where both weak and strong Riemannian metrics can be defined. Therefore, we use the terminology "*strictly weak Riemannian metric*" to emphasize the cases where the weak Riemannian metric is not strong.

As we discussed above, for strictly weak Riemannian metrics, a striking phenomenon can occur, resulting in the vanishing of the geodesic distance between distinct points. In Chapter 6 we have proved that when a specific choice of strictly weak Riemannian metric on the infinite dimensional Heisenberg group is considered, then the following phenomenon conjectured by Michor and Mumford appears: the blow-up of the sectional curvature occurs along some planes that are related to the shrinking curves which connect the distinct points, where the geodesic distance vanishes, [66].

Taking into account the above comments and the subsequent Theorem Q, then we may interpret the Michor–Mumford conjecture in infinite dimensional Lie groups as follows. Considering an infinite dimensional Lie group, equipped with a weak, left invariant Riemannian metric and a degenerate geodesic distance, *then we expect that the sectional curvature at the unit element is positively unbounded*.

In this chapter, we introduce *Hilbertian H-type groups*, whose geometry validates the previous version of the conjecture, with respect to a large class of strictly weak, left invariant Riemannian metrics. Hilbertian H-type groups include the infinite dimensional Heisenberg group of [66] and in the finite dimensional case they exactly coincide with the well known H-type groups, that were discovered by A. Kaplan, [48, 49, 50], see also [51]. We notice that Kaplan's definition perfectly works also through the infinite dimensional interpretation. On the other hand, the effective existence of infinite dimensional H-type groups needs to be verified. In Section 7.1, we provide an infinite dimensional construction, from which one may notice that there are infinitely many infinite dimensional Hilbertian H-type Lie groups, see Remark 7.3.

We focus our attention on the "natural" *weak* Riemannian metrics on Hilbertian H-type groups, that are left invariant and make the subspaces $\mathbb{V}$ and $\mathbb{W}$ orthogonal. Borrowing the terminology from the finite dimensional case, we say that such metrics are *graded*. For instance, the Cameron-Martin subgroup of [37] is a two step, infinite dimensional Lie group equipped with a strong and graded Riemannian metric. Thus, the next statement validates our interpretation of the Michor–Mumford conjecture in Hilbertian H-type groups.

Theorem Q. *Let σ be a weak, graded Riemannian metric on a Hilbertian H-type group. If the metric σ yields a degenerate geodesic distance, then the sectional curvature at the unit element exists on a sequence of planes and it is positively unbounded.*

The starting point of the proof is that the degenerate geodesic distance forces the graded Riemannian metric to be strictly weak. Then we prove that for strictly weak, graded Riemannian metrics the blow-up of the sectional curvature always occurs. Extending Theorem Q to more general classes of infinite dimensional Lie groups seems an interesting open question.

It is also important to understand whether, and in which cases, the geodesic distance in a Hilbertian H-type group is actually degenerate. Here a rather striking fact appears, since infinite dimensional sub-Riemannian (sub-Finsler) Geometry enters the proof of Theorem Q. More generally, for *any* strictly weak, left invariant sub-Finsler metric on a Hilbertian H-type group, the sub-Finsler distance is *degenerate*, see Theorem 7.10. The idea behind the proof of this theorem is to use a sequence of vectors, where the weak and the strong topology differ. Then we use the map J associated with the structure of H-type group, which allows for the same "shrinking-space effect" that was first observed in [66]. As a consequence, we have the following result, corresponding to Theorem 7.13.

Theorem R (Characterization of points with vanishing distance)**.** *Let F be a strictly weak, left invariant Finsler metric on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$ and let us fix $x, y \in \mathbb{V}$, $z_1, z_2 \in \mathbb{W}$. Then we have*

$$d_F(x + z_1, y + z_2) = 0 \quad \text{if and only if} \quad x = y,$$

where d_F is the Finsler distance associated with F.

The subclass of strictly weak, graded Riemannian metrics on a Hilbertian H-type group gives rise to another singular phenomenon, i.e. the lack of the Levi-Civita covariant derivative.

Theorem S. *If σ is a strictly weak, graded Riemannian metric on a Hilbertian H-type group, then it does not admit the Levi-Civita covariant derivative.*

The proof of the previous theorem is given in Section 7.4. An example of nonexistence of the Levi-Civita connection was provided in [19]. We also notice that in [69, Example 2.26] the model of [65] is extended to a family of weak Riemannian metrics which do not possess Christoffel symbols, hence their associated Levi-Civita covariant derivative cannot exist.

Despite Theorem S, we observe that through the classical Arnold's formula [10] it is still possible to compute the sectional curvature of some planes in a Hilbertian H-type group. However, we cannot claim that the formula works for all planes. In fact, there are also planes for which the Arnold's formula does not apply, as it is shown for instance in [66, Remark 4.1].

The next theorem proves that in a Hilbertian H-type group equipped with a strictly weak, graded Riemannian metric, we can always find sequences of planes in $T_0\mathbb{M}$ where the sectional curvature is well defined, and also unbounded.

Theorem T (Unboundedness of the sectional curvature)**.** *Let $\mathbb{M}$ be a Hilbertian H-type group. If σ is a strictly weak, graded Riemannian metric on $\mathbb{M}$, then there exist two sequences of planes $\{P_n\}_{n\in\mathbb{N}}, \{Q_n\}_{n\in\mathbb{N}} \subset$*

$T_0\mathbb{M}$ whose sectional curvatures $K_\sigma(P_n)$ and $K_\sigma(Q_n)$ are well defined through the Arnold's formula and we have

$$\lim_{n\to\infty} K_\sigma(P_n) = -\infty \quad \textit{and} \quad \lim_{n\to\infty} K_\sigma(Q_n) = +\infty.$$

This theorem is a version of Theorem 7.24, where we provide an explicit form for the planes:

$$P_n = \operatorname{span}\{w_n, J_z w_n\} \quad \text{and} \quad Q_n = \operatorname{span}\{z, J_{Az} w_n\}.$$

We believe that there is a connection between these planes and the sequence of curves that progressively decrease their length, while connecting the fixed distinct points. In fact, we point out that the projection on $\mathbb{V}$ of these *horizontal curves* has the form

$$\gamma_1^n(t) = tc\sqrt{n}\, w_n + \frac{t^2 c}{2}\frac{1}{\sqrt{n}} J_z(w_n)$$

for $c \in \mathbb{R}$, see the proof of Theorem 7.24. In this sense, we surmise that the planes P_n and Q_n should be somehow related to the parts of the space where the curves γ^n "move", when their length reduces up to converging to zero. However, the precise relationship between the planes of the blow-up and the shrinking curves remains unclear to us.

We observe that Theorem T immediately gives Theorem Q, since the vanishing of the geodesic distance implies that the graded Riemannian metric is strictly weak. We hopefully expect that our remarks in Hilbertian H-type groups provide more insights to understand the Michor–Mumford phenomenon in other classes of infinite dimensional manifolds.

Chapter 6

On the Michor-Mumford phenomenon in the infinite dimensional Heisenberg group

6.1 Preliminary notions

In this section, we present an infinite dimensional version of the classical Heisenberg group equipped with either a Riemannian or a sub-Riemannian structure. The construction is well known and it has connections with different areas of Mathematics.

6.1.1 A short introduction to the infinite dimensional Heisenberg group

We denote by ℓ^2 the linear space of all real and square summable sequences. Its scalar product $\langle \cdot, \cdot \rangle$ has the associated norm $\|x\| = \sqrt{\sum_{j=1}^{\infty} x_j^2}$ for any element $x = \sum_{j=1}^{\infty} x_j e_j$. The set of unit vectors $\{e_j : j \geq 1\}$ denotes the canonical orthonormal basis of ℓ^2. For each integer $n \geq 1$, the element e_n of ℓ^2 has n-th entry equal to 1 and all the others are zero.

We consider $\ell^2 \times \ell^2 \times \mathbb{R}$ endowed with its standard structure of product of Hilbert spaces and we introduce the continuous and skew-symmetric function

$$\beta((h_1, h_2), (h_1', h_2')) = \langle h_1, h_2' \rangle - \langle h_2, h_1' \rangle.$$

defined on $(\ell^2 \times \ell^2) \times (\ell^2 \times \ell^2)$. Then we introduce a continuous Lie product on $\ell^2 \times \ell^2 \times \mathbb{R}$:

$$[(h_1, h_2, \tau), (h_1', h_2', \tau')] = 2\beta((h_1, h_2), (h_1', h_2'))\,(0, 0, 1) \in \ell^2 \times \ell^2 \times \mathbb{R}, \tag{6.1}$$

that makes $\ell^2 \times \ell^2 \times \mathbb{R}$ an *infinite dimensional Lie algebra*. Due to the Baker–Campbell–Hausdorff formula (in short BCH), we equip $\ell^2 \times \ell^2 \times \mathbb{R}$ with a noncommutative and analytic Lie group operation:

$$(h_1, h_2, \tau)(h_1', h_2', \tau') = (h_1 + h_1', h_2 + h_2', \tau + \tau' + \beta((h_1, h_2), (h_1', h_2'))) \tag{6.2}$$

for all elements $(h_1, h_2, \tau), (h_1', h_2', \tau') \in \ell^2 \times \ell^2 \times \mathbb{R}$. We denote by $\mathbb{H}$ the Hilbert Lie group arising from the previous group operation, that is the *infinite dimensional Heisenberg group* modelled on the Hilbert space $\ell^2 \times \ell^2 \times \mathbb{R}$.

From the viewpoint of Mathematical Physics, the group $\mathbb{H}$ naturally appears in the theory of representations of infinite dimensional Lie algebras, see [45] and the references therein. In the theory of infinite dimensional Lie groups, $\mathbb{H}$ can be seen as a special instance of more general BCH-Lie groups. They are infinite dimensional Lie groups with a local exponential mapping that is also a bianalytic diffeomorphism around the origin, [78], [42]. For infinite dimensional nilpotent Lie algebras, the BCH formula defines a global group operation, [79]. This viewpoint was followed in [64] to define infinite dimensional Banach homogeneous groups as suitable direct sums of Banach spaces, equipped with an analytic structure and an everywhere converging BCH formula. In the same work, several examples of Banach homogeneous groups were provided. We mention infinite producs of Engel groups using either ℓ^p or L^p spaces, an infinite product of Heisenberg groups modelled on $\ell^2 \times \ell^1$ and other analogous analytic constructions.

In relation to the understanding of hypoellipticity in infinite dimensions [68], infinite dimensional Heisenberg-like groups based on a Wiener space, along with their Brownian motion were introduced and studied in [37]. In this connection, also some Ricci curvature lower bounds are obtained, using a left invariant Riemannian metric.

6.1.2 Weak Riemannian metrics on $\mathbb{H}$

For each $p \in \mathbb{H}$, we denote by $L_p : \mathbb{H} \to \mathbb{H}$ the left multiplication by p, defined as $L_p(r) = p \cdot r$ for all $r \in \mathbb{H}$. The group operation (6.2) gives a simple formula for the differential of L_p at a point q, namely

$$(dL_p)_q(v) = \lim_{t\to 0} \frac{L_p(q + tv) - L_p(q)}{t} = (v_1, v_2, v_3 + \langle p_1, v_2\rangle - \langle p_2, v_1\rangle)$$

for every $v = (v_1, v_2, v_3) \in T_q\mathbb{H}$, with $p = (p_1, p_2, \tau)$. Notice that we have identified $T_q\mathbb{H}$ with $\mathbb{H}$, using the Hilbert space structure $\mathbb{H}$. We also notice that our formula for the differential $(dL_p)_q$ does not depend on the point q. We consider a scalar product

$$\sigma_0 : T_0\mathbb{H} \times T_0\mathbb{H} \to \mathbb{R}$$

on the tangent space $T_0\mathbb{H}$ of $\mathbb{H}$ at the origin, which is continuous with respect to the product topology of $T_0\mathbb{H} \times T_0\mathbb{H}$.

Then for every $p \in \mathbb{H}$ and $v, w \in T_p\mathbb{H}$ the following scalar product

$$\sigma_p(v, w) = \sigma_0\big((dL_{p^{-1}})_p v, (dL_{p^{-1}})_p w\big) = \sigma_0\big((dL_{-p})_p v, (dL_{-p})_p w\big) \tag{6.3}$$

defines a *left invariant weak Riemannian metric σ on* $\mathbb{H}$. The associated *Riemannian norm* is denoted by $\|\cdot\|_\sigma$.

If for any piecewise smooth curve $\gamma : [0, 1] \to \mathbb{H}$ we define its Riemannian length as

$$\ell_\sigma(\gamma) = \int_0^1 \|\dot\gamma(t)\|_\sigma \, dt,$$

then the associated geodesic distance $d : \mathbb{H} \times \mathbb{H} \to [0, +\infty)$ between $p, q \in \mathbb{H}$ is

$$d(p, q) = \inf\{\ell_\sigma(\gamma) : \gamma \text{ is a piecewise smooth curve with } \gamma(0) = p, \gamma(1) = q\}. \tag{6.4}$$

Clearly d is left invariant, symmetric and it satisfies the triangle inequality.

6.1.3 Weak sub-Riemannian metrics on $\mathbb{H}$

Identifying $\mathbb{H}$ with $T_0\mathbb{H}$, the set $\ell^2 \times \ell^2 \times \{0\}$ can be seen as a closed subspace of $T_0\mathbb{H}$, that we denote by $H_0\mathbb{H}$. We may obtain a left invariant *horizontal subbundle*, denoted by $H\mathbb{H}$, introducing the fibers

$$H_p\mathbb{H} = (dL_p)_0(H_0\mathbb{H}) \subset T_p\mathbb{H}$$

for every $p = (p_1, p_2, \tau) \in \mathbb{H}$. We note that $v = (v_1, v_2, v_3) \in H_p\mathbb{H}$ if and only if

$$(dL_{-p})_p(v) = (v_1, v_2, v_3 - \langle p_1, v_2\rangle + \langle p_2, v_1\rangle) \in H_0\mathbb{H} \tag{6.5}$$

and the previous condition corresponds to the equality

$$v_3 - \langle p_1, v_2\rangle + \langle p_2, v_1\rangle = 0.$$

We have a precise formula to define the *horizontal curves* associated to $H\mathbb{H}$. They are continuous and piecewise smooth curves $\gamma : [0, 1] \to \mathbb{H}$ of the form $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in \mathbb{H}$, such that for almost every $t \in [0, 1]$ we have

$$\dot\gamma_3(t) - \langle\gamma_1(t), \dot\gamma_2(t)\rangle + \langle\gamma_2(t), \dot\gamma_1(t)\rangle = 0.$$

The previous differential constraint means that $\dot\gamma(t) \in H_{\gamma(t)}\mathbb{H}$.

On the horizontal fibers $H_p\mathbb{H}$ of $H\mathbb{H}$ we can fix a scalar product. A *left invariant weak sub-Riemannian metric* g on $H\mathbb{H}$ is defined by a continuous inner product

$$g_0 : H_0\mathbb{H} \times H_0\mathbb{H} \to \mathbb{R},$$

so that for all $p \in \mathbb{H}$ and $v, w \in H_p\mathbb{H}$ we have

$$g_p(v, w) = g_0\big((dL_{p^{-1}})_p v, (dL_{p^{-1}})_p w\big) = g_0\big((dL_{-p})_p v, (dL_{-p})_p w\big). \tag{6.6}$$

The associated *sub-Riemannian norm* is denoted by $\|\cdot\|_g$ and the length of a horizontal curve $\gamma : [0, 1] \to \mathbb{H}$ is defined by

$$\ell_g(\gamma) = \int_0^1 \|\dot\gamma(t)\|_g \, dt.$$

For any couple of points in $\mathbb{H}$, it is easy to construct a piecewise smooth horizontal curve that connects them, hence the following *sub-Riemannian distance*

$$\rho(p, q) = \inf\{\ell_g(\gamma) : \gamma \text{ is a horizontal curve with } \gamma(0) = p, \gamma(1) = q\} \tag{6.7}$$

is finite for every couple of points $p, q \in \mathbb{H}$, hence we have $\rho : \mathbb{H} \times \mathbb{H} \to [0, +\infty)$. One immediately notices that ρ is left invariant, symmetric and it satisfies the triangle inequality.

6.2 Degenerate geodesic distances in the infinite dimensional Heisenberg group

This section is devoted to the construction of special left invariant weak Riemannian and sub-Riemannian metrics that yield degenerate geodesic distances.

We introduce the linear and continuous operator $A : \ell^2 \to \ell^2$, which associates to each $x \in \ell^2$ of components $(x_k)_{k\geq 1}$ the element $Ax \in \ell^2$, whose k-th component is $(Ax)_k = x_k/k$. Then we define the scalar product $\eta : \ell^2 \times \ell^2 \to \mathbb{R}$ as

$$\eta(v, w) = \langle Av, w\rangle$$

for all $v, w \in \ell^2$ and its associated norm

$$\|v\|_\eta = \sqrt{\eta(v, v)} = \sqrt{\langle Av, v\rangle}. \tag{6.8}$$

We use η to define the new scalar product

$$g_0((v_1, v_2), (w_1, w_2)) = \eta(v_1, w_1) + \eta(v_2, w_2)$$

for every $(v_1, v_2), (w_1, w_2) \in \ell^2 \times \ell^2$. By our identification, g_0 can be seen as a scalar product on $H_0\mathbb{H}$, so that using (6.6) we obtain a left invariant weak sub-Riemannian metric g on $\mathbb{H}$. We follow the notation of the previous section, denoting by ρ the special sub-Riemannian distance associated to this choice of g through formula (6.7).

To obtain a left invariant weak Riemannian metric σ on $\mathbb{H}$, we extend g_0 as follows

$$\sigma_0((v_1, v_2, v_3), (w_1, w_2, w_3)) = g_0((v_1, v_2), (w_1, w_2)) + v_3w_3 \tag{6.9}$$

for every $(v_1, v_2, v_3), (w_1, w_2, w_3) \in T_0\mathbb{H}$, where $\sigma_0 : T_0\mathbb{H} \times T_0\mathbb{H} \to \mathbb{R}$. From (6.3), the scalar product in (6.9) immediately defines a left invariant weak Riemannian metric σ on $\mathbb{H}$. The Riemannian distance associated to σ through (6.4) will be denoted by d.

Remark 6.1. Let us consider the sub-Riemannian metric g, the Riemannian metric σ and their associated geodesic distances d and ρ defined above (the Riemannian distance and the sub-Riemannian distance, respectively). The family of piecewise smooth curves connecting two points also contains the horizontal curves connecting the same points. Since the restriction of σ to the horizontal subbundle $H\mathbb{H}$ coincides with g, the infimum defining d is taken over a larger family, hence $d \leq \rho$.

Remark 6.2. It is easy to notice that both d and ρ are not everywhere vanishing on $\mathbb{H}$. We consider $(p_1, p_2, \tau), (q_1, q_2, s) \in \mathbb{H}$ with $(p_1, p_2) \neq (q_1, q_2)$ and we choose any piecewise smooth curve $\gamma = (\gamma_1, \gamma_2, \gamma_3) : [0, 1] \to \mathbb{H}$ with $\gamma(0) = (p_1, p_2, \tau)$ and $\gamma(1) = (q_1, q_2, s)$. Let $i_0 \in \{1, 2\}$ be such that $p_{i_0} \neq q_{i_0}$ and let $k_0 \geq 1$ such that $p_{i_0k_0} \neq q_{i_0k_0}$, where

$$p_{i_0} = \sum_{j=1}^{\infty} p_{i_0j}e_j \quad \text{and} \quad q_{i_0} = \sum_{j=1}^{\infty} q_{i_0j}e_j.$$

We consider the component $\gamma_{i_0} = \sum_{j=1}^{\infty} \gamma_{i_0j}e_j$ and the following inequalities

$$\begin{aligned}\ell_\sigma(\gamma) &\geq \int_0^1 \sqrt{\|\dot\gamma_1\|_\eta^2 + \|\dot\gamma_2\|_\eta^2}dt \\ &\geq \int_0^1 \left\|\dot\gamma_{i_0}\right\|_\eta dt \geq \int_0^1 \frac{|\dot\gamma_{i_0k_0}|}{\sqrt{k_0}}dt \geq \frac{|p_{i_0k_0} - q_{i_0k_0}|}{\sqrt{k_0}} > 0.\end{aligned}$$

In particular, we have shown that

$$0 < \frac{|p_{i_0k_0} - q_{i_0k_0}|}{\sqrt{k_0}} \leq d((p_1, p_2, \tau), (q_1, q_2, s)) \leq \rho((p_1, p_2, \tau), (q_1, q_2, s)).$$

The previous computation also shows that both d and ρ are actually distances, if restricted to any hyperplane $\ell^2 \times \ell^2 \times \{\kappa\}$ with $\kappa \in \mathbb{R}$.

We are now in a position to prove the following theorem.

Theorem 6.3. *There exist a left invariant weak sub-Riemannian metric and a left invariant weak Riemannian metric on* $\mathbb{H}$ *such that their associated geodesic distances are not positive on all couples of distinct points.*

Proof. For each $p \in \mathbb{H}$, we denote the norm of a horizontal vector

$$v = (v_1, v_2, v_3) \in H_p\mathbb{H}$$

with respect to g as follows

$$\|v\|_g = \left\|(dL_{-p})_p v\right\|_g = \|(v_1, v_2, 0)\|_g\,,$$

where the last equality is due to (6.5) and $(v_1, v_2, 0)$ is identified with a vector of $H_0\mathbb{H}$.

Since the subspaces $\ell^2 \times \{0\} \times \{0\}$ and $\{0\} \times \ell^2 \times \{0\}$ of $H_0\mathbb{H}$ are orthogonal with respect to g_0, the previous equalities give

$$\|v\|_g^2 = \|v_1\|_\eta^2 + \|v_2\|_\eta^2\,,$$

where $\|\cdot\|_\eta$ is defined in (6.8). Thus, the length of a horizontal curve $\gamma : [0,1] \to \mathbb{H}$ with respect to g satisfies the formula

$$\ell_g(\gamma) = \int_0^1 \sqrt{\|\dot{\gamma}_1\|_\eta^2 + \|\dot{\gamma}_2\|_\eta^2}\,dt, \tag{6.10}$$

where $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$.

Next, we wish to show that whenever $(p_1, p_2, s_1), (p_1, p_2, s_2) \in \mathbb{H}$, then

$$\rho((p_1, p_2, s_1), (p_1, p_2, s_2)) = 0. \tag{6.11}$$

To do this, the main point is to prove that for all $s > 0$, we have $\rho((0,0,0),(0,0,s)) = 0$. We will construct a sequence of horizontal curves connecting $(0,0,0)$ to $(0,0,s)$, whose length converges to zero. Such sequence is obtained by gluing different sequences of horizontal curves. We fix $c > 0$ and consider $\gamma^n : [0,1] \to \mathbb{H}$ defined by

$$\gamma^n(t) = (\gamma_1^n(t), \gamma_2^n(t), \gamma_3^n(t)) = \Big(\frac{t^2}{2}ce_n, -te_nc, \frac{t^3}{6}c^2\Big),$$

where the unit vector e_n is the n-th vector of the fixed orthonormal basis $\{e_j : j \geq 1\}$ of ℓ^2. By definition (6.8), we get

$$\left\|\dot{\gamma}_1^n(t)\right\|_\eta^2 = \frac{t^2c^2}{n} \quad \text{and} \quad \left\|\dot{\gamma}_2^n(t)\right\|_\eta^2 = \frac{c^2}{n}. \tag{6.12}$$

From the form of γ^n, it is immediate to check that the differential constraint

$$\dot{\gamma}_3^n - \langle \gamma_1^n, \dot{\gamma}_2^n \rangle + \langle \gamma_2^n, \dot{\gamma}_1^n \rangle = 0$$

is satisfied for all $t \in [0,1]$, hence γ_n is horizontal. Thus, formula (6.10) holds and the expressions of (6.12) immediately prove that $\ell_g(\gamma^n) \to 0$ as $n \to +\infty$.

Now we define the sequence of curves $\alpha^n : [0,1] \to \mathbb{H}$ as

$$\alpha^n(t) = (\alpha_1^n(t), \alpha_2^n(t), \alpha_3^n(t)) = \left(c\left(\frac{1}{2} - \frac{t^2}{2}\right) e_n, c(t-1)e_n, c^2\left(\frac{1}{6} + \frac{t^3}{6} - \frac{t^2}{2} + \frac{t}{2}\right)\right).$$

We immediately obtain

$$\left\|\dot{\alpha}_1^n(t)\right\|_\eta^2 = \frac{t^2c^2}{n} \quad \text{and} \quad \left\|\dot{\alpha}_2^n(t)\right\|_\eta^2 = \frac{c^2}{n} \tag{6.13}$$

and the differential constraint

$$\dot{\alpha}_3^n - \langle \alpha_1^n, \dot{\alpha}_2^n \rangle + \langle \alpha_2^n, \dot{\alpha}_1^n \rangle = 0$$

is satisfied for all $t \in [0,1]$. All curves α^n are horizontal, hence combining (6.10) and (6.13), we conclude that $\ell_g(\alpha^n) \to 0$ as $n \to +\infty$. We note that

$$\alpha^n(0) = \left(\frac{c}{2}e_n, -ce_n, \frac{c^2}{6}\right) = \gamma^n(1)$$

for all $n \in \mathbb{N}$, hence we can consider the gluing $\alpha^n * \gamma^n : [0,1] \to \mathbb{H}$ of α^n and γ^n, that is a piecewise smooth curve. Clearly $\alpha^n * \gamma^n$ is a horizontal curve and for all $n \in \mathbb{N}$ we have

$$\alpha^n * \gamma^n(0) = \gamma^n(0) = (0,0,0) \quad \text{and} \quad \alpha^n * \gamma^n(1) = \alpha^n(1) = \left(0, 0, \frac{c^2}{3}\right)$$

and $\ell_g(\alpha^n * \gamma^n) = \ell_g(\alpha^n) + \ell_g(\gamma^n) \to 0$ as $n \to \infty$. We have proved that

$$\rho\left((0,0,0), \left(0,0,\frac{c^2}{3}\right)\right) = 0,$$

hence $\rho((0,0,0),(0,0,s)) = 0$ for all $s > 0$. By the left invariance of ρ, we have

$$\rho((0,0,0),(0,0,-s)) = \rho((0,0,s),(0,0,0)) = 0,$$

therefore $\rho((0,0,0),(0,0,t)) = 0$ for every $t \in \mathbb{R}$. We conclude that

$$\begin{aligned}
\rho((p_1,p_2,s_1),(p_1,p_2,s_2)) &= \rho((p_1,p_2,0)(0,0,s_1),(p_1,p_2,0)(0,0,s_2)) \\
&= \rho((0,0,s_1),(0,0,s_2)) \\
&= \rho((0,0,0),(0,0,s_2-s_1)) = 0,
\end{aligned}$$

that proves (6.11). According to Remark 6.1, the inequality $d \le \rho$ implies that for all $(p_1,p_2,s_1),(p_1,p_2,s_2) \in \mathbb{H}$, we have

$$d((p_1,p_2,s_1),(p_1,p_2,s_2)) = 0.$$

This concludes the proof. □

6.3 On the sectional curvature of a weak Riemannian Heisenberg group

In this section, we study the sectional curvature of $\mathbb{H}$ equipped with a specific left invariant weak Riemannian metric. Following Section 6.2, we define the unique left invariant weak Riemannian metric σ, such that

$$\sigma_0((v_1, v_2, v_3), (w_1, w_2, w_3)) = g_0((v_1, v_2), (w_1, w_2)) + v_3 w_3$$

for $(v_1, v_2, v_3), (w_1, w_2, w_3) \in T_0\mathbb{H}$, according to (6.9). We recall the formula

$$g_0((v_1, v_2), (w_1, w_2)) = \eta(v_1, w_1) + \eta(v_2, w_2) = \langle Av_1, w_1\rangle + \langle Av_2, w_2\rangle$$

and $Ax = \sum_{k=1}^{\infty} x_k/k$, $x = \sum_{k=1}^{\infty} x_k e_k \in \ell^2$. For every positive integer j, we use the notation

$$e_j^1 = (e_j, 0, 0), \quad e_j^2 = (0, e_j, 0) \quad \text{and} \quad e^3 = (0, 0, 1),$$

to indicate the standard orthonormal basis of $\mathbb{H}$ seen as the Hilbert space $\ell^2 \times \ell^2 \times \mathbb{R}$.

One easily realizes that the natural linear isomorphism between Lie($\mathbb{H}$) and $\mathbb{H}$ is also an isomorphism of Lie algebras, where we equip $\mathbb{H}$ with the Lie product (6.1). Thus, by slight abuse of notation, the left invariant vector fields of Lie($\mathbb{H}$) isomorphically associated with the basis e_j^1, e_j^2, e^3 are denoted by the same symbols. It follows that for all $i, j \geq 1$ and $l = 1, 2$ we have

$$[e_i^1, e_j^2] = 2\delta_{ij}e^3 \quad \text{and} \quad [e_i^l, e_j^l] = 0,$$

where e_i^1, e_j^1, e^3 are now understood as left invariant vector fields of Lie($\mathbb{H}$).

Now we consider a left invariant weak Riemannian metric υ on $\mathbb{H}$. The associated scalar product on Lie($\mathbb{H}$) is denoted by $\langle\cdot,\cdot\rangle_\upsilon$. We consider two orthonomal vectors $X, Y \in$ Lie($\mathbb{H}$) with respect to $\langle\cdot,\cdot\rangle_\upsilon$. By virtue of [10, Theorem 5], the sectional curvature $K_\upsilon(X, Y)$ of the plane in Lie($\mathbb{H}$) spanned by X and Y can be obtained by the adjoint operator $\mathrm{ad}(Y)^\top(X)$, that we now introduce. We define $\mathrm{ad}(Y)(Z) = [Y, Z]$ and consider (in case it exists) the unique vector $\mathrm{ad}(Y)^\top(X) \in$ Lie($\mathbb{H}$) that satisfies the equalities

$$\langle [Y, Z], X\rangle_\upsilon = \langle \mathrm{ad}(Y)(Z), X\rangle_\upsilon = \left\langle Z, \mathrm{ad}(Y)^\top(X)\right\rangle_\upsilon$$

for every $Z \in$ Lie($\mathbb{H}$). Then we define

$$B_\upsilon(X, Y) = \mathrm{ad}(Y)^\top(X) \in \mathrm{Lie}(\mathbb{H}).$$

In the case υ is a strong Riemannian metric, [1, Definition 5.2.12], the existence of $B_\upsilon(X, Y)$ is always ensured, but not for any weak Riemannian metric. For instance, in Remark 6.4 below, we show the nonexistence of $B_\sigma(X, Y)$ for a specific choice of X and Y, where σ is the weak Riemannian metric introduced at the beginning of this section.

From formula (53) of [10], we have

$$K_\upsilon(X, Y) = \langle\delta, \delta\rangle_\upsilon + 2\langle\alpha, \beta\rangle_\upsilon - 3\langle\alpha, \alpha\rangle_\upsilon - 4\langle B_X, B_Y\rangle_\upsilon, \tag{6.14}$$

where we define

$$\delta = \frac{1}{2}(B_\upsilon(X, Y) + B_\upsilon(Y, X)), \quad \beta = \frac{1}{2}(B_\upsilon(X, Y) - B_\upsilon(Y, X)), \quad \alpha = \frac{1}{2}[X, Y] \tag{6.15}$$

$$B_X = \frac{1}{2}B_\upsilon(X, X) \quad \text{and} \quad B_Y = \frac{1}{2}B_\upsilon(Y, Y). \tag{6.16}$$

The proof of Theorem P follows from the application of (6.14) with respect to σ on suitable choices of planes. We denote by $\langle\cdot,\cdot\rangle_\sigma$ the scalar product in Lie($\mathbb{H}$) induced by the left invariant weak Riemannian metric σ. The associated norm on Lie($\mathbb{H}$) is denoted by $\|\cdot\|_\sigma$. We assume that for $X, Y \in$ Lie($\mathbb{H}$) the adjoint

$$B_\sigma(X, Y) = \mathrm{ad}(Y)^\top(X)$$

with respect to σ exists. In this case, its scalar product with a vector $Z \in$ Lie($\mathbb{H}$) is assigned by the following formula

$$\left\langle \mathrm{ad}(Y)^\top(X), Z\right\rangle_\sigma = \langle [Y, Z], X\rangle_\sigma = 2\beta(\pi(Y), \pi(Z))x^3, \tag{6.17}$$

as a consequence of (6.1), where $\pi : \mathbb{H} \to \ell^2 \times \ell^2$ is the canonical projection defined by

$$X = (\pi(X), x^3) = (\pi(X), 0) + x^3 e^3.$$

We use the fixed orthonormal basis e^1_j, e^2_j, e^3 of $\mathbb{H}$ with respect to the standard Hilbert product of $\ell^2 \times \ell^2 \times \mathbb{R}$, getting

$$\mathrm{ad}(Y)^\top(X) = \sum_{j=1}^\infty [\mathrm{ad}(Y)^\top(X)]^1_j e^1_j + \sum_{j=1}^\infty [\mathrm{ad}(Y)^\top(X)]^2_j e^2_j + [\mathrm{ad}(Y)^\top(X)]^3 e^3.$$

Formula (6.17) yields

$$\sum_{j=1}^\infty \frac{1}{j}[\mathrm{ad}(Y)^\top(X)]^1_j Z^1_j + \sum_{j=1}^\infty \frac{1}{j}[\mathrm{ad}(Y)^\top(X)]^2_j Z^2_j + [\mathrm{ad}(Y)^\top(X)]^3 Z^3 = 2\beta(\pi(Y), \pi(Z))x^3 \tag{6.18}$$

for arbitrary $Z = Z^3 e^3 + \sum_{j=1}^\infty Z^1_j e^1_j + Z^2_j e^2_j$. In the case $X = \pi(X)$, formula (6.18) shows the existence of $\mathrm{ad}(Y)^\top(\pi(X))$ and yields

$$B_\sigma(\pi(X), Y) = \mathrm{ad}(Y)^\top(\pi(X)) = 0. \tag{6.19}$$

In the case $X = e^3$, again (6.18) for $Z = e^1_j$ and $Z = e^2_j$ respectively, gives

$$[\mathrm{ad}(Y)^\top(e^3)]^1_j = 2j\beta(\pi(Y), e^1_j) \quad \text{and} \quad [\mathrm{ad}(Y)^\top(e^3)]^2_j = 2j\beta(\pi(Y), e^2_j).$$

For $Z = e^3$, applying (6.18) we get

$$[\mathrm{ad}(Y)^\top(e^3)]^3 = 0.$$

Assuming the existence of $\mathrm{ad}(Y)^\top(e^3)$, we have shown that

$$B_\sigma(e^3, Y) = \mathrm{ad}(Y)^\top(e^3) = 2\sum_{j=1}^\infty j\beta(\pi(Y), e^1_j)e^1_j + 2\sum_{j=1}^\infty j\beta(\pi(Y), e^2_j)e^2_j.$$

Writing $Y = Y^3 e^3 + \sum_{j=1}^\infty (Y^1_j e^1_j + Y^2_j e^2_j)$, we finally get

$$B_\sigma(e^3, Y) = 2\sum_{j=1}^\infty j(Y^1_j e^2_j - Y^2_j e^1_j). \tag{6.20}$$

Then the assumption about the existence of $B_\sigma(e^3, Y)$ corresponds to the convergence of its series. The next remark shows a choice of Y for which the series (6.20) does not converge.

Remark 6.4. If we consider the vector

$$W = \sum_{j=1}^{\infty} \frac{e_j^1}{j} \in \mathrm{Lie}(\mathbb{H}),$$

then it is easy to check that the series (6.20) representing $B_\sigma(e^3, W)$ does not converge. As a consequence, the adjoint $\mathrm{ad}(W)^\top(e^3)$ cannot be defined. In addition, Arnold's formula (6.14) for the sectional curvature of the plane $\mathrm{span}\left\{W, e^3\right\}$ does not apply.

Proposition 6.5. *We consider the orthonormal elements* $W_k, e^3 \in \mathrm{Lie}(\mathbb{H})$ *with* $k \geq 1$ *and*

$$W_k = \left(\sum_{j=1}^{k} j^{-3}\right)^{-1/2} \sum_{j=1}^{k} \frac{e_j^1}{j} \in \mathrm{Lie}(\mathbb{H}).$$

As the subspace $\mathrm{span}\left\{W_k, e^3\right\}$ *converges to* $\mathrm{span}\{W_\infty, e^3\}$ *for* $k \to \infty$*, with*

$$W_\infty = \left(\sum_{j=1}^{\infty} j^{-3}\right)^{-1/2} \sum_{j=1}^{\infty} \frac{e_j^1}{j} \in \mathrm{Lie}(\mathbb{H}),$$

it follows that

$$K_\sigma(W_k, e^3) \to +\infty.$$

The convergence of $\mathrm{span}\left\{W_k, e^3\right\}$ *to* $\mathrm{span}\{W_\infty, e^3\}$ *is considered in the Grassmannian of the 2-dimensional planes contained in* $\mathrm{Lie}(\mathbb{H})$.

Proof. First of all, the pointwise convergence of W_k to W_∞ implies the convergence of $\mathrm{span}\left\{W_k, e^3\right\}$ to $\mathrm{span}\{W_\infty, e^3\}$. To compute $K_\sigma(W_k, e^3)$, we first apply (6.19), getting

$$B_\sigma(W_k, e^3) = \mathrm{ad}(e^3)^\top(W_k) = 0$$

for all $k \geq 1$. From (6.20), it follows that $B_\sigma(e^3, e_j^1) = 2je_j^2$, hence

$$B_\sigma\left(e^3, \frac{e_j^1}{j}\right) = 2e_j^2.$$

The bilinearity of $B_\sigma(\cdot,\cdot)$ yields

$$B_\sigma(e^3, W_k) = 2\left(\sum_{j=1}^{k} j^{-3}\right)^{-1/2} \sum_{j=1}^{k} e_j^2$$

From (6.15), taking $\delta = \left(B_\sigma(W_k, e^3) + B_\sigma(e^3, W_k)\right)/2$, we obtain

$$\langle \delta, \delta \rangle_\sigma = \frac{1}{4}\|B_\sigma(e^3, W_k)\|_\sigma^2 = \left(\sum_{j=1}^{\infty} j^{-3}\right)^{-1} \left\|\sum_{j=1}^{k} e_j^2\right\|_\sigma^2 = \left(\sum_{j=1}^{\infty} j^{-3}\right)^{-1} \sum_{j=1}^{k} j^{-1}$$

From (6.15), (6.16), (6.19) and (6.20), we find

$$\alpha = \frac{1}{2}B_\sigma(W_k, W_k) = \frac{1}{2}B_\sigma(e^3, e^3) = 0.$$

Finally, by formula (6.14), we have proved that

$$K_\sigma(W_k, e^3) = \langle \delta, \delta \rangle_\sigma = \left(\sum_{j=1}^{\infty} j^{-3} \right)^{-1} \sum_{j=1}^{k} j^{-1} \to +\infty$$

as $k \to \infty$. This concludes the proof. □

Proof of Theorem P. Following the notation of the present section, we define

$$a_{1j} = \sqrt{j} e_j^1 \quad \text{and} \quad a_{2j} = \sqrt{j} e_j^2$$

of Lie($\mathbb{H}$), that are orthonormal with respect to $\langle \cdot, \cdot \rangle_\sigma$ and do not commute. To apply (6.14) for finding $K_\sigma(a_{1j}, a_{2j})$, we use (6.15) and (6.16). Due to (6.19), we get

$$B_\sigma(a_{1j}, a_{2j}) = B_\sigma(a_{2j}, a_{1j}) = 0.$$

As a result, we have

$$K_\sigma(a_{1j}, a_{2j}) = -3 \langle \alpha, \alpha \rangle_\sigma = -\frac{3}{4} \|[a_{1j}, a_{2j}]\|_\sigma^2 = -3j^2.$$

Now we wish to compute $K_\sigma(a_{1j}, e^3)$ and $K_\sigma(a_{2j}, e^3)$. We first apply (6.19) and (6.20), getting

$$B_\sigma(e_j^l, e^3) = \mathrm{ad}(e^3)^\top(e_j^l) = 0, \quad B_\sigma(e^3, e_j^1) = 2je_j^2 \quad \text{and} \quad B_\sigma(e^3, e_j^2) = -2je_j^1$$

for all $l = 1, 2$ and $k \geq 1$. From (6.15), taking $\delta = \left(B_\sigma(a_{1j}, e^3) + B_\sigma(e^3, a_{1j}) \right)/2$, we obtain

$$\begin{aligned} \langle \delta, \delta \rangle_\sigma &= \frac{1}{4} \|B_\sigma(a_{1j}, e^3) + B_\sigma(e^3, a_{1j})\|_\sigma^2 = \frac{1}{4} \|\sqrt{j} B_\sigma(e^3, e_j^1)\|_\sigma^2 = \frac{j}{4} \|2je_j^2\|_\sigma^2 \\ &= j^3 \left\langle e_j^2, e_j^2 \right\rangle_\sigma = j^3 \left\langle A e_j^2, e_j^2 \right\rangle = j^2. \end{aligned}$$

From (6.15), (6.16), (6.19) and (6.20), we find

$$\alpha = \frac{1}{2}B_\sigma(e_j^1, e_j^1) = \frac{1}{2}B_\sigma(e^3, e^3) = 0.$$

Due to the formula for the sectional curvature (6.14), we have established that

$$K_\sigma(a_{1j}, e^3) = \langle \delta, \delta \rangle_\sigma = j^2. \tag{6.21}$$

In analogous setting $\delta = \left(B_\sigma(a_{2j}, e^3) + B_\sigma(e^3, a_{2j}) \right)/2$, we obtain

$$\langle \delta, \delta \rangle_\sigma = \frac{1}{4} \|B_\sigma(e^3, a_{2j})\|_\sigma^2 = \frac{j}{4} \|B_\sigma(e^3, e_j^2)\|_\sigma^2 = \frac{j}{4} \|2je_j^1\|_\sigma^2 = j^3 \|e_j^1\|_\sigma^2 = j^2.$$

Again (6.15), (6.16), (6.19) and (6.20) imply that

$$\alpha = \frac{1}{2}B_\sigma(e_j^2, e_j^2) = \frac{1}{2}B_\sigma(e^3, e^3) = 0.$$

Due to (6.14), we have also proved that

$$K_\sigma(a_{2j}, e^3) = \langle \delta, \delta \rangle_\sigma = j^2. \tag{6.22}$$

Taking into account (6.21) and (6.22), setting $b = e^3$, we have completed the proof. □

Remark 6.6. A direct verification shows that the computations of sectional curvature, to prove Theorem P, could be also carried out extending the finite dimensional formula of [74, Lemma 1.1] for the countable structure coefficients of Lie($\mathbb{H}$). These coefficients are obtained from the orthonormal vectors $\sqrt{j}e_j^1$, $\sqrt{j}e_j^2, e^3$ of Lie($\mathbb{H}$) with respect to $\langle \cdot, \cdot \rangle_\sigma$.

Following the notation of the this section, the sequence of curves whose length converges to zero in the proof of Theorem 6.3 can be written as

$$\gamma^j(t) = \frac{ct^2}{2}e_j^1 - cte_j^2 + \frac{c^2t^3}{6}e^3 \in \mathbb{H} \quad \text{and}$$

$$\alpha^j(t) = c\left(\frac{1}{2} - \frac{t^2}{2}\right)e_j^1 + c(t-1)e_j^2 + c^2\left(\frac{1}{6} + \frac{t^3}{6} - \frac{t^2}{2} + \frac{t}{2}\right)e^3 \in \mathbb{H}.$$

It is interesting to notice that all such curves are contained in the span of the planes

$$\text{span}\{e_j^1, e_j^2\}, \quad \text{span}\{e_j^1, e^3\} \quad \text{and} \quad \text{span}\{e_j^2, e^3\}.$$

When these planes are seen in the Lie algebra, Theorem P shows that their sectional curvature blows-up, as the length of the curves converges to zero.

Chapter 7

The Michor-Mumford conjecture in Hilbertian H-type groups

7.1 Infinite dimensional H-type groups

The approach of [64] can be used to construct specific classes of infinite dimensional Banach nilpotent Lie groups, starting from an infinite dimensional nilpotent Lie algebra. Indeed, the group operation is immediately provided by the Baker–Campbell–Hausdorff formula, which we abbreviate as "BCH formula". We will see that this simple viewpoint allows us to get the notion of a possibly infinite dimensional H-type group.

We fix some notions that we will use throughout the chapter. Let $\mathbb{M}$ be a Hilbert space, consider a continuous Lie product $[\cdot,\cdot] : \mathbb{M} \times \mathbb{M} \to \mathbb{M}$ and two orthogonal and nontrivial closed subspaces $\mathbb{V}$ and $\mathbb{W}$ such that $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$, with $\dim(\mathbb{W}) < +\infty$. We denote the scalar product on $\mathbb{M}$ by $\langle \cdot, \cdot \rangle$ and the associated norm by $|\cdot|$. The space of all linear continuous endomorphisms of a Banach space X is denoted by $\mathcal{E}(X)$.

We say that $\mathbb{M}$ is a *Hilbertian H-type group*, or simply an *H-type group*, if the following conditions hold.

(I) $[\mathbb{M}, \mathbb{M}] \subset \mathbb{W}$ and $[\mathbb{M}, \mathbb{W}] = \{0\}$,

(II) the unique linear and continuous operator $J : \mathbb{W} \to \mathcal{E}(\mathbb{V})$ defined by the formula

$$\langle J_z x, y \rangle = \langle z, [x, y] \rangle \tag{7.1}$$

for $z \in \mathbb{W}$, $x, y \in \mathbb{V}$, satisfies the additional condition

$$J_z^2 = -|z|^2 \, \mathrm{Id}_{\mathbb{V}}, \tag{7.2}$$

where $\mathrm{Id}_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$ is the identity mapping.

Notice that the existence of the linear and continuous operator J is a consequence of both the Riesz representation theorem applied on $\mathbb{V}$ and the continuity of the bilinear mapping $[\cdot,\cdot]$. Thus, (7.1) immediately follows.

The group operation is automatically obtained by the BCH formula:

$$x \cdot y = x + y + \frac{1}{2}[x, y]. \tag{7.3}$$

From the defining formula (7.1), we immediately notice that the adjoint operator J_z^* satisfies

$$J_z^* = -J_z.$$

As a consequence, using also (7.2), we may write

$$|J_z x|^2 = -\langle x, J_z^2 x\rangle = |z|^2|x|^2,$$

that gives

$$|J_z x| = |z||x|. \tag{7.4}$$

Therefore, using also the defining formula (7.1), we have

$$|z|^2|x|^2 = |J_z x|^2 = \langle z, [x, J_z x]\rangle. \tag{7.5}$$

For every $w \in \mathbb{W}$, we also notice that (7.4) implies

$$|\langle w, [x, y]\rangle| = |\langle J_w x, y\rangle| \le |w|\,|x|\,|y|,$$

therefore in any H-type group we have

$$|[x, y]| \le |x|\,|y|. \tag{7.6}$$

For $x, z \neq 0$, it follows that

$$\left|\frac{[x, J_z x]}{|x|^2|z|}\right| \le 1$$

and in addition (7.5) gives

$$\left\langle \frac{z}{|z|}, \frac{[x, J_z x]}{|x|^2|z|} \right\rangle = 1. \tag{7.7}$$

Combining (7.6) and (7.7), we have proved that

$$[x, J_z x] = |x|^2 z. \tag{7.8}$$

Notice that $[\mathbb{M}, \mathbb{W}] = \{0\}$ implies that $\mathbb{W}$ is contained in the center of $\mathbb{M}$, where we regard $\mathbb{M}$ as a Lie algebra. However, it is easy to notice that condition (7.8) shows that $\mathbb{W}$ exactly coincides with the center of $\mathbb{M}$.

Remark 7.1. Notice that in the case $\dim(\mathbb{V}) < +\infty$, the Hilbertian H-type group coincides with the well known (finite dimensional) H-type group, [48], hence motivating our terminology.

Next, we construct examples of (infinite dimensional) Hilbertian H-type groups. We fix an H-type group $\mathfrak{n} = \upsilon \oplus \zeta$, where υ and ζ are finite dimensional orthogonal subpaces of the Hilbert space $\mathfrak{n}$. We denote by $\langle \cdot, \cdot \rangle_{\mathfrak{n}}$ the scalar product of $\mathfrak{n}$ and by $|\cdot|_{\mathfrak{n}}$ its associated norm.

The endomorphism $J^{\mathfrak{n}} : \zeta \to \mathcal{E}(\upsilon)$ defines the H-type structure on $\mathfrak{n}$. We denote by $\mathbb{N}^+$ the set of positive integers and consider the space of square-summable sequences

$$\mathbb{P}_\upsilon = \left\{ (x_k)_k : \ x_k \in \upsilon, k \in \mathbb{N}^+, \sum_{k=1}^{\infty} |x_k|_{\mathfrak{n}}^2 < +\infty \right\}.$$

We set $\mathbb{M} = \mathbb{P}_\upsilon \times \zeta$ and identifying $\mathbb{P}_\upsilon$ and ζ with $\mathbb{P}_\upsilon \times \{0\}$ and $\{0\} \times \zeta$, respectively, we can also write

$$\mathbb{M} = \mathbb{P}_\upsilon \oplus \zeta.$$

For $(x, z), (x', z') \in \mathbb{M}$, we define the scalar product

$$\langle (x, z), (x', z') \rangle = \langle ((x_k)_k, z), ((x'_k)_k, z') \rangle = \langle z, z' \rangle_{\mathfrak{n}} + \sum_{k=1}^{\infty} \langle x_k, x'_k \rangle_{\mathfrak{n}} \tag{7.9}$$

that makes $\mathbb{M}$ a Hilbert space, where $\mathbb{P}_\upsilon$ and ζ are orthogonal closed subspaces. We denote by $|\cdot|$ the associated norm on $\mathbb{M}$. For $x = (x_k)_k \in \mathbb{P}_\upsilon$ and $z \in \zeta$, we define

$$J_z(x) = (J_z^{\mathfrak{n}} x_k)_k. \tag{7.10}$$

Thus, observing that

$$\sum_{k=1}^{\infty} |J_z^{\mathfrak{n}} x_k|_{\mathfrak{n}}^2 = |z|_{\mathfrak{n}}^2 \sum_{k=1}^{\infty} |x_k|_{\mathfrak{n}}^2 < +\infty,$$

the mapping $J_z : \mathbb{P}_\upsilon \to \mathbb{P}_\upsilon$ is well defined and

$$J_z^2 = -|z|_{\mathfrak{n}}^2 \, \mathrm{Id}_{\mathbb{P}_\upsilon},$$

since $(J_z^{\mathfrak{n}})^2 = -|z|_{\mathfrak{n}}^2 \, \mathrm{Id}_\upsilon$. The Lie product of $\xi + \eta, \xi' + \eta' \in \mathfrak{n} = \upsilon \oplus \zeta$ is given by a skew-symmetric continuous bilinear mapping

$$\beta : \upsilon \times \upsilon \to \zeta$$

such that

$$[\xi + \eta, \xi' + \eta'] = \beta(\xi, \xi').$$

By the property (7.6) for H-type groups, we get

$$|\beta(\xi, \xi')| = |[\xi, \xi']| \le |\xi|_{\mathfrak{n}} \, |\xi'|_{\mathfrak{n}}$$

for all $\xi, \xi' \in \upsilon$, therefore the Lie product

$$[(x, z), (x', z')] = \left(0, \sum_{k=1}^{+\infty} \beta(x_k, x'_k) \right), \tag{7.11}$$

is well defined for all $(x, z), (x', z') \in \mathbb{M}$. Cauchy–Schwarz inequality yields

$$|[(x, z), (x', z')]| \le \sum_{k=1}^{+\infty} |\beta(x_k, x'_k)| \le \sum_{k=1}^{\infty} |x_k|_{\mathfrak{n}} \, |x'_k|_{\mathfrak{n}} \le |x| \, |x'|,$$

hence the Lie product $[\cdot,\cdot]$ is continuous on $\mathbb{M}$. Finally, from definition (7.10) of $J_z : \mathbb{P}_\upsilon \to \mathbb{P}_\upsilon$, we obtain

$$\langle J_z x, y\rangle = \sum_{k=1}^{\infty}\langle J_z^{\mathfrak{n}} x_k, y_k\rangle_{\mathfrak{n}} = \sum_{k=1}^{\infty}\langle z, [x_k, y_k]\rangle_{\mathfrak{n}} = \sum_{k=1}^{\infty}\langle z, \beta(x_k, y_k)\rangle_{\mathfrak{n}} = \langle z, [x, y]\rangle$$

for all $x, y \in \mathbb{P}_\upsilon$ and $z \in \zeta$. We have proved the following result.

Theorem 7.2. *The linear space $\mathbb{M} = \mathbb{P}_\upsilon \oplus \zeta$ equipped with scalar product* (7.9)*, Lie product* (7.11) *and linear operator* (7.10) *is an infinite dimensional H-type group.*

Remark 7.3. By [48, Corollary 1], there exist infinitely many finite dimensional H-type groups, where there are no isomorphic couples. Indeed, these groups can be chosen to have centers of different dimensions. As a result, Theorem 7.2 also shows that there are infinitely many infinite dimensional H-type groups.

Remark 7.4. We point out that, when the finite dimensional H-type group $\mathfrak{n}$ coincides with the 3-dimensional Heisenberg group, in the construction of $\mathbb{M}$, then Theorem 7.2 yields the infinite dimensional Heisenberg group studied in [66].

7.2 Weak metrics on Hilbertian H-type groups

In the sequel, $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$ always denotes a Hilbertian H-type group, equipped with a scalar product $\langle\cdot,\cdot\rangle$ and its Hilbertian norm $|\cdot|$. This section presents various notions of weak metrics on $\mathbb{M}$. They include weak Finsler metrics and weak Riemannian metrics. Indeed, both of these metrics may induce a topology which is strictly weaker than the manifold topology. We will also follow the convention of identifying the tangent space $T_q\mathbb{M}$ with the group itself $\mathbb{M}$, $q \in \mathbb{M}$, due to the linear structure of $\mathbb{M}$.

For every $p \in \mathbb{M}$, the *left multiplication by* p is denoted by $L_p : \mathbb{M} \to \mathbb{M}$, with

$$L_p(q) = p \cdot q = p + q + \frac{1}{2}[p, q]$$

for all $q \in \mathbb{M}$. We define the skew-symmetric bilinear function $\beta : \mathbb{V} \times \mathbb{V} \to \mathbb{W}$ such that

$$[x, y] = \beta(x, y)$$

for every $x, y \in \mathbb{V}$. By definition of $\mathbb{M}$, we have two canonical projections $\pi_1 : \mathbb{M} \to \mathbb{V}$ and $\pi_2 : \mathbb{M} \to \mathbb{W}$ such that every $p \in \mathbb{M}$ can be written in a unique way as

$$p = \pi_1(p) + \pi_2(p)$$

where $\pi_1(p)$ and $\pi_2(p)$ are also orthogonal. We obviously have the isometric isomorphism

$$\mathbb{M} \to \mathbb{V} \times \mathbb{W}, \quad p \to (\pi_1(p), \pi_2(p))$$

with respect to the Hilbert structure of $\mathbb{M}$. We use the simplified notation $p_i = \pi_i(p)$ for $p \in \mathbb{M}$, so that we can write $p = p_1 + p_2$ with $p_1 \in \mathbb{V}$ and $p_2 \in \mathbb{W}$. Then the group operation (7.3) gives a simple formula for the differential of L_p at a point $q \in \mathbb{M}$ along $v = v_1 + v_2 \in \mathbb{M}$:

$$(dL_p)_q(v) = \frac{d}{dt}L_p(q + tv)\Big|_{t=0} = v + \frac{1}{2}[p, v] = v_1 + v_2 + \frac{\beta(p_1, v_1)}{2}. \tag{7.12}$$

Indeed, the linear structure of $\mathbb{M}$ allows us to identify $T_q\mathbb{M}$ with $\mathbb{M}$.

Weak Finsler metrics and Finsler distances

We fix a norm $F_0 : \mathbb{M} \to [0, +\infty)$ with respect to the linear structure of $\mathbb{M}$, which also yields a Finsler metric on $T\mathbb{M}$. We always assume that F_0 is continuous, namely $F_0(v) \le c_1|v|$ for some $c_1 > 0$ and for all $v \in \mathbb{M}$, where $|\cdot|$ is the fixed scalar product on $\mathbb{M}$. It is also natural to assume that the decomposition $\mathbb{V} \oplus \mathbb{W}$ is compatible with the Finsler norm, namely $\pi_1 : \mathbb{M} \to \mathbb{V}$ and $\pi_2 : \mathbb{M} \to \mathbb{W}$ are continuous with respect to F_0. In other words, there exists $C > 0$ such that

$$F_0(\pi_1(x)) \le CF_0(x) \quad \text{and} \quad F_0(\pi_2(x)) \le CF_0(x).$$

Thus, for each $p \in \mathbb{M}$, we set

$$F_p(v) = F_0((dL_{-p})_p(v))$$

for every $v \in T_p\mathbb{M}$. We say that the map F on $T\mathbb{M}$ arising from the norms F_p is a *weak, left invariant Finsler metric* on $T\mathbb{M}$. We say that F is a *strong, left invariant Finsler metric* if the topology induced by F_0 on $\mathbb{M}$ coincides with the already given Hilbert topology of $\mathbb{M}$. In different terms, there exist $\tilde{c}_1 > 0$ such that $F_0(v) \ge \tilde{c}_1|v|$ for all $v \in \mathbb{M}$. If a weak, left invariant Finsler metric F on $\mathbb{M}$ is not strong, then we say that F is a *strictly weak, left invariant Finsler metric* on $\mathbb{M}$.

Example 7.5. Let us consider the infinite dimensional Heisenberg group $\mathbb{H} = \ell^2 \times \ell^2 \times \mathbb{R}$ equipped with the product of the associated Hilbert structure and the group operation as defined in [66]. We have $\mathbb{H} = \mathbb{V} \oplus \mathbb{W}$, where $\mathbb{V} = \ell^2 \times \ell^2 \times \{0\}$ and $\mathbb{W} = \{0\} \times \{0\} \times \mathbb{R}$. We fix $\mathfrak{p} > 2$ and for an element $(h, k, t) \in \mathbb{M}$, we define the norm

$$F_0(h, k, t) = \|h\|_{\mathfrak{p}} + \|k\|_{\mathfrak{p}} + |t|,$$

where $\|x\|_{\mathfrak{p}} = (\sum_{k=1}^{\infty} |x_j|^{\mathfrak{p}})^{1/\mathfrak{p}} \le \|x\|_2 < +\infty$ for every $x \in \ell^2$. Clearly F_0 gives an example of strictly weak, left invariant Finsler metric. Indeed, it is also obvious that the projections π_1 and π_2 on $\mathbb{V}$ and $\mathbb{W}$ are F_0-continuous, respectively.

The length of a continuous, piecewise smooth curve $\gamma : [0, 1] \to \mathbb{M}$ is defined by the integral

$$\ell_F(\gamma) = \int_0^1 F_{\gamma(t)}(\dot{\gamma}(t))dt = \int_0^1 F_0\left((dL_{-\gamma(t)})_{\gamma(t)}\dot{\gamma}(t)\right) dt. \tag{7.13}$$

Then we can immediately define the associated *Finsler distance*

$$d_F(p, q) = \inf\{\ell_F(\gamma) : \gamma \text{ is continuous, piecewise smooth, } \gamma(0) = p \text{ and } \gamma(1) = q\}$$

for every $p, q \in \mathbb{M}$, hence $d_F : \mathbb{M} \times \mathbb{M} \to [0, +\infty)$. Clearly d_F is left invariant, symmetric and satisfies the triangle inequality.

Remark 7.6. Let us consider a weak, left invariant Finsler metric F on $\mathbb{M}$, and let d_F be the associated geodesic distance. We will prove that for $p, q \in \mathbb{M}$ with $\pi_1(p) = x \ne y = \pi_1(q)$, we have $C\, d_F(p, q) \ge F_0(x - y) > 0$. Indeed, for every continuous, piecewise smooth curve $\gamma : [0, 1] \to \mathbb{M}$ joining p to q, we get

$$\ell_F(\gamma) = \int_0^1 F_0\Big((dL_{-\gamma(t)})_{\gamma(t)}(\dot{\gamma}(t))\Big)\, dt \ge \frac{1}{C} \int_0^1 F_0(\dot{\gamma}_1(t))\, dt$$

in view of (7.15) and taking into account the F-continuity of the projections. Thus, if we consider the projected curve $\gamma_1 : [0, 1] \to \mathbb{V}$, we can piecewise integrate $\dot{\gamma}_1$ on the intervals where it is continuous.

Then we apply [44, Theorem 2.1.1 (ii)] and [44, Theorem 2.2.2] and the triangle inequality on a partition $t_0 = 0 < t_1 < \cdots < t_k = 1$. It follows that

$$\ell_F(\gamma) \geq \frac{1}{C}\int_0^1 F_0(\dot{\gamma}_1(t))\,dt \geq \frac{1}{C}\sum_{j=0}^{k-1} F_0\left(\int_{t_j}^{t_{j+1}} \dot{\gamma}_1(t)\right) = \frac{1}{C}\sum_{j=0}^{k-1} F_0\left(\gamma_1(t_{j+1}) - \gamma_1(t_j)\right),$$

hence

$$\ell_F(\gamma) \geq \frac{F_0(x-y)}{C} > 0.$$

Weak sub-Finsler metrics

Identifying $\mathbb{V} \oplus \mathbb{W}$ with $T_0(\mathbb{V} \oplus \mathbb{W})$, the subspace $\mathbb{V}$ can be seen as a closed subspace of $T_0\mathbb{M}$, that we denote by $H_0\mathbb{M}$ and we may introduce the *left invariant horizontal subbundle*, denoted by $H\mathbb{M}$, with fibers

$$H_p\mathbb{M} = (dL_p)_0(H_0\mathbb{M}) \subset T_p\mathbb{M}$$

for every $p \in \mathbb{M}$. For each $p \in \mathbb{M}$, on the horizontal fiber $H_p\mathbb{M}$ of $H\mathbb{M}$ we can fix a norm, which turns out to be continuous and left invariant. Precisely, a *weak, left invariant sub-Finsler metric S* on $H\mathbb{M}$ is defined by a norm

$$S_0 : \mathbb{V} \to [0, +\infty) \tag{7.14}$$

satisfying for some $c_0 > 0$ and for all $x \in \mathbb{V}$ the inequality

$$S_0(x) \leq c_0|x|.$$

The previous condition immediately yields the continuity of S_0 with respect to the fixed Hilbert topology on $\mathbb{M}$. Notice that the closed subspace $\mathbb{V}$ inherits a Hilbert structure from $\mathbb{M}$. With the previous identifications, for every $p \in \mathbb{M}$ and $v \in H_p\mathbb{M}$, we introduce the norm

$$S_p(v) = S_0\left((dL_{-p})_p(v)\right)$$

on the fiber $H_p\mathbb{M}$. If the topology defined by the norm S_0 on $\mathbb{V}$ coincides with the Hilbert one of $\mathbb{V}$, we say that S_0 defines a *strong, left invariant sub-Finsler metric*. This is equivalent to the existence of a constant $\tilde{c} > 0$ such that $\tilde{c}|x| \leq S_0(x)$ for all $x \in \mathbb{V}$. If this is not the case, we say that S_0 defines a *strictly weak, left invariant sub-Finsler metric*.

Example 7.7. Let us consider the infinite dimensional Heisenberg group $\mathbb{H} = \ell^2 \times \ell^2 \times \mathbb{R}$ equipped with the product of the associated Hilbert structure and the group operation as defined in [66]. We have $\mathbb{H} = \mathbb{V} \oplus \mathbb{W}$, where $\mathbb{V} = \ell^2 \times \ell^2 \times \{0\}$ and $\mathbb{W} = \{0\} \times \{0\} \times \mathbb{R}$. We fix $\mathfrak{p} > 2$ and for an element $(h, k, 0) \in \mathbb{V}$, we define the norm

$$S_0(h, k) = \|h\|_{\mathfrak{p}} + \|k\|_{\mathfrak{p}},$$

where $\|x\|_{\mathfrak{p}} = (\sum_{k=1}^{\infty} |x_j|^{\mathfrak{p}})^{1/\mathfrak{p}} \leq \|x\|_2 < +\infty$ for every $x \in \ell^2$. Clearly S_0 gives an example of strictly weak, left invariant sub-Finsler metric.

Horizontal curves and sub-Finsler distances

We notice that the expression of the differential of translations (7.12) proves that $v \in H_p\mathbb{M}$ if and only if

$$(dL_{-p})_p(v) = v - \frac{1}{2}[p, v] = v_1 + v_2 - \frac{\beta(p_1, v_1)}{2} \in H_0\mathbb{M} \tag{7.15}$$

and the previous condition corresponds to the equality

$$v_2 = \frac{\beta(p_1, v_1)}{2}. \tag{7.16}$$

Thus, we have a precise formula to define the *horizontal curves* associated with $H\mathbb{H}$. They are continuous and piecewise smooth curves $\gamma : [0, 1] \to \mathbb{M}$ of the form $\gamma = \gamma_1 + \gamma_2 \in \mathbb{M}$, such that for almost every $t \in [0, 1]$ we have

$$\dot\gamma_2(t) = \frac{\beta(\gamma_1(t), \dot\gamma_1(t))}{2}.$$

The previous differential constraint means that $\dot\gamma(t) \in H_{\gamma(t)}\mathbb{M}$. The length of a horizontal curve $\gamma : [0, 1] \to \mathbb{M}$ is defined by $\ell_S(\gamma) = \int_0^1 S_{\gamma(t)}(\dot\gamma(t))dt$, therefore

$$\ell_S(\gamma) = \int_0^1 S_0\left((dL_{-\gamma(t)})_{\gamma(t)}\dot\gamma(t)\right) dt = \int_0^1 S_0(\dot\gamma_1(t))\, dt.$$

It is not difficult to observe that all couple of points in $\mathbb{M}$ can be connected by horizontal curves. As a result, the *sub-Finsler distance*

$$\rho_S(p, q) = \inf\{\ell_S(\gamma) : \gamma \text{ is a horizontal curve with } \gamma(0) = p, \gamma(1) = q\}$$

is finite for every $p, q \in \mathbb{M}$, hence $\rho_F : \mathbb{M} \times \mathbb{M} \to [0, +\infty)$. The fact that ρ_F is left invariant, symmetric and satisfies the triangle inequality is straightforward.

Remark 7.8. Let us consider a weak, left invariant sub-Finsler metric S on $\mathbb{M}$, and a weak, left invariant Finsler metric F on $\mathbb{M}$ such that $F_0|_{\mathbb{V}} = S_0$. We define ρ_S and d_F to be the associated sub-Finsler distance and Finsler distance, respectively. Taking into account (7.13), (7.15) and (7.16) we observe that $\ell_F(\gamma) = \ell_S(\gamma)$ for every horizontal curve. Then we immediately get

$$\rho_S(p, q) \geq d_F(p, q)$$

for every $p, q \in \mathbb{M}$. Taking into account Remark 7.6 we also have $\rho_S(p, q) \geq d_F(p, q) > 0$ whenever $\pi_1(p) \neq \pi_1(q)$. Notice that for any fixed weak sub-Finsler metric S_0 on $\mathbb{M}$, we can always find a weak Finsler metric F_0 such that $F_0|_{\mathbb{V}} = S_0$. It suffices to choose any Hilbert norm $|\cdot|$ on $\mathbb{W}$, defining

$$F_0(x + z) = S_0(x) + |z|$$

for every $x \in \mathbb{V}$ and $z \in \mathbb{W}$.

Weak Riemannian metrics and Riemannian distances

Following Section 7.1, we consider a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$ equipped with a Hilbert product $\langle \cdot, \cdot \rangle$ and the mapping J_z, $z \in \mathbb{W}$. We fix a continuous scalar product σ_0 on $\mathbb{M}$, namely

$$\|v\|_{\sigma_0} \leq c_0 |v| \tag{7.17}$$

for some $c_0 > 0$ and every $v \in \mathbb{M}$, where $\| \cdot \|_{\sigma_0}$ is the norm arising from σ_0. We also require that the canonical projections $\pi_1 : \mathbb{M} \to \mathbb{V}$ and $\pi_2 : \mathbb{M} \to \mathbb{W}$ are σ_0-continuous, that is

$$\|\pi_1(v)\|_{\sigma_0} \leq C\|v\|_{\sigma_0} \quad \text{and} \quad \|\pi_2(v)\|_{\sigma_0} \leq C\|v\|_{\sigma_0}$$

for all $v \in \mathbb{M}$ and some $C > 0$. Thus, σ_0 gives a scalar product

$$\sigma_p(v, w) = \sigma_0((dL_{p^{-1}})_p v, (dL_{p^{-1}})_p w) = \sigma_0((dL_{-p})_p v, (dL_{-p})_p w)$$

for each $p \in \mathbb{M}$ and $v, w \in T_p\mathbb{M}$. The corresponding Riemannian metric σ on $T\mathbb{M}$ is called *weak, left invariant Riemannian metric*. Notice that the Riemannian norm $\| \cdot \|_{\sigma_0}$ on $\mathbb{M}$ is also Finsler metric.

Let us consider the topology defined by σ_0 on $\mathbb{M}$. When it coincides with the topology determined by the Hilbert structure of $\mathbb{M}$, we say that σ is a *strong, left invariant Riemannian metric*. We say that σ is a *strictly weak, left invariant Riemannian metric* if it is not strong. Finally, a (strictly) weak, left invariant Riemannian metric σ on $\mathbb{M}$ such that $\mathbb{V}$ and $\mathbb{W}$ are σ_0-orthogonal is called *(strictly) weak, graded Riemannian metric*.

For a fixed weak, left invariant Riemannian metric σ on $\mathbb{M}$, we consider the linear and continuous operator $A : \mathbb{M} \to \mathbb{M}$ such that for all $v, w \in \mathbb{M}$ we have $\sigma_0(v, w) = \langle v, Aw \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the Hilbert product on $\mathbb{M}$. The operator A exists by the classical Riesz representation theorem and it is automatically self-adjoint and positive.

We denote by $A_{\mathbb{V}}$ its restriction to $\mathbb{V}$ and by $A_{\mathbb{W}}$ its restriction to $\mathbb{W}$. When σ_0 is graded, it is easy to notice that $A_{\mathbb{V}}(\mathbb{V}) \subset \mathbb{V}$ and $A_{\mathbb{W}}(\mathbb{W}) \subset \mathbb{W}$. Then we can consider the linear and continuous operators $A_{\mathbb{V}} : \mathbb{V} \to \mathbb{V}$ and $A_{\mathbb{W}} : \mathbb{W} \to \mathbb{W}$.

The following proposition is also standard.

Proposition 7.9. *If σ is a weak, left invariant Riemannian metric on $\mathbb{M}$, then the subspace $A(\mathbb{M})$ is dense in $\mathbb{M}$. Furthermore, σ is strong if and only if A is surjective.*

For any continuous and piecewise smooth curve $\gamma : [0, 1] \to \mathbb{M}$ its Riemannian length with respect to the weak Riemannian metric σ is defined as

$$\ell_\sigma(\gamma) = \int_0^1 \|\dot{\gamma}(t)\|_\sigma \, dt.$$

The *geodesic distance* associated with σ is the function $d_\sigma : \mathbb{M} \times \mathbb{M} \to [0, +\infty)$ defined as

$$d_\sigma(p, q) = \inf\{\ell_\sigma(\gamma) : \gamma \text{ is a continuous and piecewise smooth curve with } \gamma(0) = p, \gamma(1) = q\}.$$

Clearly d_σ is left invariant, symmetric and it satisfies the triangle inequality.

7.3 Degenerate geodesic distances

The next theorem proves the existence of degenerate sub-Finsler distaces in any Hilbertian H-type group equipped with a strictly weak, left invariant sub-Finsler metric.

Theorem 7.10 (Vanishinig of sub-Finsler distances). *Let $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$ be an infinite dimensional H-type group equipped with the canonical projections $\pi_1 : \mathbb{M} \to \mathbb{V}$ and $\pi_2 : \mathbb{M} \to \mathbb{W}$. Let ρ_S be the sub-Finsler distance arising from any strictly weak, left invariant sub-Finsler metric S on $\mathbb{M}$. Then for every $p, q \in \mathbb{M}$ with $\pi_1(p) = \pi_1(q)$, we have $\rho_S(p, q) = 0$.*

Proof. It suffices to prove that for all $c \in \mathbb{R}$ and all $z \in \mathbb{W}$ with $|z| = 1$, we have

$$\rho_S\left(0, \frac{c^2}{3} z\right) = 0. \tag{7.18}$$

Since the norm S_0 of (7.14) does not define the Hilbert topology of $\mathbb{V}$, there exists a sequence $\{w^n\}_n$ in $\mathbb{V}$ such that $|w^n| = 1$ and $S_0(w^n) \leq \frac{1}{n}$ for all $n \in \mathbb{N}^+$. We choose $z \in \mathbb{W}$ with $|z| = 1$, and for each $n \in \mathbb{N}$, define

$$\gamma_1^n(t) = tc\sqrt{n}w_n + \frac{t^2 c}{2}\frac{1}{\sqrt{n}}J_z(w_n).$$

Consider now the curve $\gamma^n = (\gamma_1^n, \gamma_2^n)$, where

$$\gamma_2^n(t) = \frac{1}{2}\int_0^t \beta(\gamma_1^n(s), \dot{\gamma}_1^n(s))\, ds \in \mathbb{W}.$$

By construction, the curve γ^n is horizontal, therefore $\ell_S(\gamma^n) = \int_0^1 S_0(\dot{\gamma}_1^n(t))\, dt$. Let us consider the following estimates

$$\begin{aligned}\ell_S(\gamma^n) = &\int_0^1 S_0(\dot{\gamma}_1^n(t))\, dt = \int_0^1 S_0\left(c\sqrt{n}w_n + \frac{ct}{\sqrt{n}}J_z(w_n)\right) dt \\ \leq &c\sqrt{n}\, S_0(w_n) + \frac{c}{\sqrt{n}}S_0(J_z(w_n)) \leq \frac{c}{\sqrt{n}} + \frac{cc_0}{\sqrt{n}}|J_z(w_n)| = \frac{c}{\sqrt{n}} + \frac{cc_0}{\sqrt{n}} \cdot |z|.\end{aligned}$$

It follows that

$$\lim_{n\to\infty} \ell_S(\gamma^n) = 0.$$

For each n, the endpoint of γ^n is

$$\gamma^n(1) = c\sqrt{n}w_n + \frac{c}{2}\frac{1}{\sqrt{n}}J_z(w_n) + \frac{c^2}{12}z.$$

Now, we define the curve $\alpha_1^n : [0, 1] \to \mathbb{V}$ as

$$\alpha_1^n(t) = c\sqrt{n}(1-t)w_n + \frac{c}{2}\frac{(1-t^2)}{\sqrt{n}}J_z(w_n)$$

and consider the lifting $\alpha^n = \alpha_1^n + \alpha_2^n$, where

$$\alpha_2^n(t) = \gamma_2^n(1) + \frac{1}{2}\int_0^t \beta(\alpha_1^n(s), \dot\alpha_1^n(s))\,ds \in \mathbb{W}$$

By construction, α^n is also horizontal and $\alpha^n(0) = \gamma^n(1)$, therefore the curve $\alpha^n \star \gamma^n$ obtained by joining γ_n and α_n is also horizontal. For each $n \in \mathbb{N}$, the curve $\alpha^n \star \gamma^n$ connects the origin $0 \in \mathbb{M}$ to the point $\frac{c^2 z}{6} \in \mathbb{W}$. We finally observe that

$$\ell_S(\alpha^n) = \int_0^1 S_0(\dot\alpha_1^n(t))\,dt = \int_0^1 S_0\left(c\sqrt{n}w_n + \frac{ct}{\sqrt{n}}J_z(w_n)\right)dt = \ell_S(\gamma^n) \to 0.$$

Therefore, $\ell_S(\alpha_n \star \gamma_n) = \ell_S(\gamma_n) + \ell_S(\alpha_n) \to 0$. We have proved that (7.18) holds for every $c \in \mathbb{R}$ and $z \in \mathbb{W}$. To conclude the proof, we consider $z_1, z_2 \in \mathbb{W}$, $z_1 \neq z_2$ and $x \in \mathbb{V}$. We notice that the left invariance of the sub-Finsler distance function yields

$$\rho_S(x + z_1, x + z_2) = \rho_S(xz_1, xz_2) = \rho_S(z_1, z_2) = \rho_S(0, z_2 - z_1).$$

Clearly, we can find $c \neq 0$ and $z \in \mathbb{W} \setminus \{0\}$ such that $z_2 - z_1 = c^2 z/6$, hence

$$\rho_S(x + z_1, x + z_2) = \rho_S(0, c^2 z/6) = 0,$$

concluding the proof. □

Corollary 7.11. *Let us fix a strictly weak, left invariant sub-Finsler metric S on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$. Then for $x, y \in \mathbb{V}$ and $z_1, z_2 \in \mathbb{W}$, we have*

$$\rho_S(x + z_1, y + z_2) = 0 \quad \textit{if and only if} \quad x = y,$$

where ρ_S is the sub-Finsler distance associated with S.

The main implication of this corollary follows by Theorem 7.10. The full characterization of the two conditions is obtained by showing that points with different projections on $\mathbb{V}$ must have positive Finsler distances. This is a consequence of combining Remark 7.6 and Remark 7.8.

Lemma 7.12. *If F be a strictly weak, left invariant Finsler metric F on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$. Then there exists a sequence $\{h_n\}_{n\in\mathbb{N}} \subset \mathbb{V}$ such that $F_0(h_n) \to 0$ and $|h_n| = 1$ for all $n \in \mathbb{N}$.*

Proof. The topology defined by F_0 on $\mathbb{M}$ is not the Hilbert one, therefore there exists a sequence u_n in $\mathbb{M}$ such that $|u_n| = 1$ for all n and $F_0(u_n) \to 0$. We can write

$$u_n = v_n + w_n = \pi_1(u_n) + \pi_2(u_n),$$

where $v_n \in \mathbb{V}$ and $w_n \in \mathbb{W}$. By the continuity of the projections, $CF_0(u_n) \geq F_0(w_n)$ therefore $F_0(w_n) \to 0$. Since $\mathbb{W}$ is finite dimensional, we also have $|w_n| \to 0$, therefore $|v_n| \to 1$. Again the continuity of the projections yields $F_0(v_n) \to 0$. To conclude the proof, we consider a subsequence v_n of nonzero vectors, and we observe that the renormalized sequence $h_n = \frac{v_n}{|v_n|}$ satisfies our claim. □

Theorem 7.13. *Let F be a strictly weak, left invariant Finsler metric on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$. Then for every $x, y \in \mathbb{V}$ and $z_1, z_2 \in \mathbb{W}$, we have $d_F((x, z_1), (y, z_2)) = 0$ if and only if $x = y$.*

Proof. The restriction of F_0 to $\mathbb{V}$ defines a weak sub-Finsler metric $S_0 : \mathbb{V} \to [0, +\infty)$. By Lemma 7.12, the corresponding left invariant sub-Finsler metric S is strictly weak. In view of Remark 7.8, we have $\rho_S \geq d_F$, so we can apply Theorem 7.10, obtaining that $d_F(p, q) = 0$, whenever $\pi_1(p) = \pi_1(q)$. By Remark 7.6, the proof is complete. □

Corollary 7.14. *Let σ be a strictly weak, left invariant Riemannian metric on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$. Then for every $x, y \in \mathbb{V}$ and $z_1, z_2 \in \mathbb{W}$, we have $d_\sigma((x, z_1), (y, z_2)) = 0$ if and only if $x = y$.*

The previous corollary follows by observing that a sitrictly weak, left invariant Riemannian metric also yields a strictly weak, left invariant Finsler metric.

7.4 Non-existence of the Levi-Civita covariant derivative

In this section, we fix a Hilbertian H-type group $\mathbb{M}$ with its Lie product $[\cdot, \cdot]$. We consider the Lie algebra Lie $\mathbb{M}$ of left invariant vector fields on $\mathbb{M}$. The associated Lie product is the skew-symmetric bilinear mapping $[\cdot, \cdot] : \text{Lie}\,\mathbb{M} \times \text{Lie}\,\mathbb{M} \to \text{Lie}\,\mathbb{M}$ its Lie product. In our setting, the linear structure of $\mathbb{M}$ allows us also consider the "identification" $\mathfrak{I} : \mathbb{M} \to \text{Lie}\,\mathbb{M}$, where $\mathfrak{I}(v) = X_v$ is the unique left invariant vector field of Lie($\mathbb{M}$) such that $X_v(0) = v$. In fact, there is the already mentioned identification between $T_0\mathbb{M}$ and $\mathbb{M}$. Throughout the section, the continuous linear and self-adjoint operator $A : \mathbb{M} \to \mathbb{M}$ is defined by the weak metric $\sigma_0(v, w) = \langle v, Aw \rangle$ for $v, w \in \mathbb{M}$.

The first result of this section is to prove that the Lie algebra Lie($\mathbb{M}$) is actually isomorphic to the starting Lie algebra $\mathbb{M}$, and the isomorphism is given by the map $\mathfrak{I}$.

Proposition 7.15. *Let $\mathbb{M}$ be an H-type group. Then the map $\mathfrak{I}$ is a Lie algebra isomorphism, that is, for every $x, y \in \mathbb{M}$ we have $\mathfrak{I}_{[x,y]} = [\mathfrak{I}_x, \mathfrak{I}_y]$.*

The proof of the previous proposition can be obtained by standard arguments, taking into account that the group operation in $\mathbb{M}$ is given by the BCH formula and the Lie product on $\mathbb{M}$. Actually, it holds in general Banach nilpotent Lie groups, [64, Proposition 2.1].

Theorem 7.16. *Let σ be a weak, graded Riemannian metric on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$. If σ admits the Levi-Civita covariant derivative ∇, then for every $x = x_1 + x_2 \in \mathbb{M}$ with $x_1 \in \mathbb{V}$ and $x_2 \in \mathbb{W}$ we have*

$$J_{Ax_2}x_1 \in \operatorname{im} A \quad \text{and} \quad \nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0) = -A^{-1}\left(J_{Ax_2}x_1\right).$$

Proof. Suppose that ∇ is the Levi-Civita covariant derivative. Since ∇ is torsion-free, we have $[\mathfrak{I}_x, \mathfrak{I}_y] = \nabla_{\mathfrak{I}_x}\mathfrak{I}_y - \nabla_{\mathfrak{I}_y}\mathfrak{I}_x$ for $x, y \in \mathbb{M}$. By the left invariance of $\mathfrak{I}_x$ and $\mathfrak{I}_y$, the function $\mathbb{M} \ni p \to \sigma_p(\mathfrak{I}_x(p), \mathfrak{I}_y(p))$ is constantly equal to $\sigma_0(x, y)$, by the identification of $\mathbb{M}$ with $T_0\mathbb{M}$. The key property of the Levi-Civita covariant derivative yields

$$0 = Z\sigma(\mathfrak{I}_x, \mathfrak{I}_y) = \sigma(\nabla_Z\mathfrak{I}_x, \mathfrak{I}_y) + \sigma(\mathfrak{I}_x, \nabla_Z\mathfrak{I}_y) \tag{7.19}$$

for every Z vector field on $\mathbb{M}$. Notice that the previous equations for $x = y$ yield

$$\sigma(\mathfrak{I}_x, \nabla_Z\mathfrak{I}_x) = 0.$$

As a consequence, using again (7.19), we get

$$\sigma([\mathfrak{I}_x,\mathfrak{I}_y],\mathfrak{I}_x) = \sigma(\nabla_{\mathfrak{I}_x}\mathfrak{I}_y,\mathfrak{I}_x) - \sigma(\nabla_{\mathfrak{I}_y}\mathfrak{I}_x,\mathfrak{I}_x) = \sigma(\nabla_{\mathfrak{I}_x}\mathfrak{I}_y,\mathfrak{I}_x)$$
$$= -\sigma(\mathfrak{I}_y,\nabla_{\mathfrak{I}_x}\mathfrak{I}_x) = -\sigma_0(y,\nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0)).$$

By Proposition 7.15, it follows that

$$\sigma([\mathfrak{I}_x,\mathfrak{I}_y],\mathfrak{I}_x) = \sigma(\mathfrak{I}_{[x,y]},\mathfrak{I}_x) = \sigma_0(\mathfrak{I}_{[x,y]}(0),\mathfrak{I}_x(0)) = \sigma_0([x,y],x).$$

Therefore, we have proved that

$$\sigma_0(y,\nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0)) = -\sigma_0([x,y],x),$$

which immediately leads us to the following equalities

$$\langle y, A\nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0)\rangle = -\langle [x,y],Ax\rangle = -\langle [x_1,y_1],Ax_2\rangle = -\langle y_1, J_{Ax_2}x_1\rangle. \tag{7.20}$$

In particular, formula (7.20) holds true for all $y \in \mathbb{W}$, hence $A\nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0) \in \mathbb{V}$. Now, taking $y \in \mathbb{V}$ in formula (7.20) we get

$$A\nabla_{\mathfrak{I}_x}\mathfrak{I}_x(0) = -J_{Ax_2}x_1,$$

which proves our claim. □

Theorem 7.17. *Let σ be a weak, graded Riemannian metric on an H-type group $\mathbb{M}$. If σ is strictly weak, then it does not admit the Levi-Civita covariant derivative.*

Proof. If σ is strictly weak, then its associated operator A is not surjective, by Proposition 7.9. Since $\mathbb{W}$ is finite dimensional and $A_{\mathbb{W}}$ is injective, then $A_{\mathbb{W}}$ is also surjective. As a consequence, $A_{\mathbb{V}}$ cannot be surjective, hence we have can choose $v \in \mathbb{V}$ such that $v \notin A_{\mathbb{V}}(\mathbb{V})$. We consider $x_2 \in \mathbb{W}$, $x_2 \neq 0$ and we define $x = J_{Ax_2}v + x_2$, $x_1 = J_{Ax_2}v$. By (7.2) we have

$$J_{Ax_2}x_1 = J_{Ax_2}(J_{Ax_2}v) = -|Ax_2|^2 v \notin \operatorname{im} A.$$

Hence, by Theorem 7.16 we get a contradiction, therefore the Levi-Civita covariant derivative does not exists for σ. □

Since strong Riemannian metrics always admit the Levi-Civita covariant derivative, the next corollary is straightforward.

Corollary 7.18. *Let σ be a weak, graded Riemannian metric on an H-type group $\mathbb{M}$. Then, σ admits the Levi-Civita covariant derivative if and only if it is a strong Riemannian metric.*

7.5 Blow-up of the sectional curvature

We consider a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$, endowed with a weak, graded Riemannian metric σ. If σ is strong, then the sectional curvature can be computed using the Riemann tensor and the Levi-Civita covariant derivative, [54]. This approach in general does not apply when σ is strictly weak, as a consequence of Theorem 7.17. We will also show how the Arnold's formula allows us to compute the sectional curvature for a special family of planes. Finally, we find a sequence of planes where the sectional curvatures blow-up.

7.5.1 The B-adjoint vector

We consider the adjoint representation $\mathrm{ad} : \mathbb{M} \to \mathcal{E}(\mathbb{M})$, where the endomorphism $\mathrm{ad}_x(y) = [x, y]$ is defined by the Lie product of $\mathbb{M}$. For a fixed couple of vectors $x, y \in \mathbb{M}$, we consider (in case it exists) the unique vector $B_{\sigma_0}(y, x) \in \mathbb{M}$ which satisfies the formula

$$\left\langle z, B_{\sigma_0}(y, x)\right\rangle_{\sigma_0} = \langle [x, z], y\rangle_{\sigma_0}$$

for every $z \in \mathbb{M}$. We say that $B_{\sigma_0}(y, x)$ is the *B-adjoint vector of* (y, x) *with respect to* σ_0. When this vector exists, it automatically satisfies

$$B_{\sigma_0}(ty, sx) = ts\, B_{\sigma_0}(y, x)$$

for every $t, s \in \mathbb{R}$. And also $B_{\sigma_0}(ty, sx)$ exists for some $t, s \neq 0$ if and only if $B_{\sigma_0}(y, x)$ exists. If σ_0 is a strong metric, then the classical Riesz representation theorem yields the existence of $B_\sigma(y, x)$ for all $x, y \in \mathbb{M}$. Precisely, in this case,

$$B_{\sigma_0}(y, x) = \mathrm{ad}_x^\top(y),$$

where $\mathrm{ad}_x^\top : \mathbb{M} \to \mathbb{M}$ is the adjoint operator of ad_x with respect to σ_0. For a strictly weak Riemannian metric, the existence of $B_\sigma(y, x) \in \mathbb{M}$ for fixed $x, y \in \mathbb{M}$ does not necessarily hold. A simple example can be found for instance in [66].

7.5.2 Arnold's formula

To compute the sectional curvature of planes in a Hilbertian H-type group $\mathbb{M}$, we use the Arnold's formula [10, Theorem 5], see also [73], [20] and [19].

Let us consider two σ_0-*orthonormal vectors* $x, y \in \mathbb{M}$, such that the B-adjoint vectors

$$B_{\sigma_0}(y, x), B_{\sigma_0}(x, y), B_{\sigma_0}(x, x), B_{\sigma_0}(y, y) \in \mathbb{M}$$

all exist. We introduce the notation $\Pi_{x,y}$ to denote the vector subspace spanned by x and y. The sectional curvature of $\Pi_{x,y}$ can be obtained by

$$K_\sigma(\Pi_{x,y}) = \langle \delta, \delta\rangle_{\sigma_0} + 2\,\langle \alpha, \beta\rangle_{\sigma_0} - 3\,\langle \alpha, \alpha\rangle_{\sigma_0} - 4\left\langle B_x, B_y\right\rangle_{\sigma_0}. \tag{7.21}$$

In the previous formula we have defined

$$\delta = \frac{1}{2}\left(B_{\sigma_0}(x, y) + B_{\sigma_0}(y, x)\right), \quad \beta = \frac{1}{2}\left(B_{\sigma_0}(x, y) - B_{\sigma_0}(y, x)\right), \quad \alpha = \frac{1}{2}[x, y]$$
$$B_x = \frac{1}{2}B_{\sigma_0}(x, x) \quad \text{and} \quad B_y = \frac{1}{2}B_{\sigma_0}(y, y).$$

It is a simple computation to verify that the sectional curvature of a plane defined through this formula does not depend on the choice of the σ_0-orthonormal basis for that plane.

First of all, we provide a condition for which the vector $B_\sigma(y, x)$ exists with $x, y \in \mathbb{M}$ fixed, see the following proposition.

Proposition 7.19 (Existence of the B-adjoint vector)**.** *Let σ_0 be a weak, graded Riemannian metric on a Hilbertian H-type group $\mathbb{M} = \mathbb{V} \oplus \mathbb{W}$ and let $x = x_1 + x_2$, $y = y_1 + y_2 \in \mathbb{M}$, with $x_1, y_1 \in \mathbb{V}$ and $x_2, y_2 \in \mathbb{W}$. It follows that*

$$\text{there exists } B_{\sigma_0}(y, x) \in \mathbb{M} \text{ if and only if } J_{Ay_2}x_1 \in A_{\mathbb{V}}(\mathbb{V}). \tag{7.22}$$

If one of the previous conditions holds, then

$$B_{\sigma_0}(y, x) = A^{-1}(J_{Ay_2}x_1). \tag{7.23}$$

Proof. Assume that $J_{Ay_2}x_1 \in A_{\mathbb{V}}(\mathbb{V})$. Thus, for all $z \in \mathbb{M}$ we have

$$\left\langle z, A^{-1}(J_{Ay_2}x_1)\right\rangle_{\sigma_0} = \langle z, J_{Ay_2}x_1\rangle = \langle [x, z], Ay_2\rangle = \langle [x, z], y\rangle_{\sigma_0},$$

hence there exists $B_{\sigma_0}(y, x) = A^{-1}(J_{Ay_2}x_1)$. If $B_{\sigma_0}(y, x) \in \mathbb{M}$ exists, then for all $z \in \mathbb{M}$ we have

$$\langle z, A(B_{\sigma_0}(y, x))\rangle = \langle z, B_{\sigma_0}(y, x)\rangle_{\sigma_0} = \langle [x, z], y\rangle_{\sigma_0} = \langle [x_1, z], Ay_2\rangle = \langle J_{Ay_2}x_1, z\rangle.$$

Therefore, $A(B_{\sigma_0}(y, x)) = J_{Ay_2}x_1$, concluding the proof. □

From (7.22) and (7.23) we get directly (1). From (7.22), (7.23) and (7.2) we get directly (2).

Remark 7.20. As a consequence of Proposition 7.19, precisely of (7.22), (7.23), for all

$$(y, x) \in (\mathbb{V} \times \mathbb{M}) \cup (\mathbb{M} \times \mathbb{W})$$

we have $J_{Ay_2}x_1 = 0$, hence the B-adjoint vector $B_{\sigma_0}(y, x)$ exists and it vanishes.

Remark 7.21. For all $z \in \mathbb{W}$ and $x \in \mathbb{V}$, we notice that

$$J_{Az}(J_{Az}(Ax)) = -|Az|^2 Ax \in A_{\mathbb{V}}(\mathbb{V}),$$

hence (7.22) yields the existence of the B-adjoint vector $B_{\sigma_0}(z, J_{Az}(Ax))$ and (7.23) gives

$$B_{\sigma_0}(z, J_{Az}(Ax)) = -|Az|^2 x. \tag{7.24}$$

We use Proposition 7.19 and the previous remarks to compute the sectional curvatures of specific planes, according to the following lemma.

Lemma 7.22. *Let σ be a weak graded Riemannian metric on $\mathbb{M}$.*

1. *If $x, y \in \mathbb{V}$ are σ_0-orthonormal, then the sectional curvature $K_\sigma(\Pi_{x,y})$ exists and*

$$K_\sigma(\Pi_{x,y}) = -\frac{3}{4}\, \|[x, y]\|^2_{\sigma_0}.$$

2. *For all $z \in \mathbb{W} \setminus \{0\}$ and $x \in \mathbb{V} \setminus \{0\}$, the vectors $J_{Az}(Ax)$ and z are orthogonal and*

$$K_\sigma(\Pi_{J_{Az}(Ax),z}) = \frac{1}{4} \frac{|Az|^4}{\|J_{Az}(Ax)\|^2_{\sigma_0}\, \|z\|^2_{\sigma_0}}\, \|x\|^2_{\sigma_0}.$$

Proof. Due to Remark 7.20, $B_\sigma(x,x)$, $B_\sigma(y,y)$, $B_\sigma(y,x)$, $B_\sigma(x,y)$ all exist and are null. Thus, (7.21) immediately gives the claim (1). The term α iof (7.21) obviously vanishes, and again Remark 7.20 gives the existence and the vanishing of $B_{J_{Az}(Ax)}$, B_z, and $B_{\sigma_0}(J_{Az}(Ax),z)$ in the corresponding Arnold's formula for the sectional curvature. From the property of the mapping J_Z, $Z\in\mathbb{W}$, of a Hilbertian H-type group, it is easy to notice that z and $J_{Az}(Ax)$ are σ_0-orthogonal.

Thus, by (7.21) applied to the σ_0-orthonormal basis $z/\|z\|_{\sigma_0}$ and $J_{Ay}(Ax)/\|J_{Ay}(Ax)\|_{\sigma_0}$, we get

$$\begin{aligned} K_\sigma(\Pi_{J_{Az}(Ax),z}) = \|\delta\|^2_{\sigma_0} &= \frac{1}{4}\left\|B_{\sigma_0}\left(\frac{z}{\|z\|_{\sigma_0}},\frac{J_{Az}(Ax)}{\|J_{Az}(Ax)\|_{\sigma_0}}\right)\right\|^2_{\sigma_0}\\ &= \frac{1}{4\|z\|^2_{\sigma_0}\|J_{Az}(Ax)\|^2_{\sigma_0}}\left\|B_{\sigma_0}(z,J_{Az}(Ax))\right\|^2_{\sigma_0}\\ &= \frac{1}{4\|z\|^2_{\sigma_0}\|J_{Az}(Ax)\|^2_{\sigma_0}}\left\||Az|^2x\right\|^2_{\sigma_0}, \end{aligned}$$

where the last equality also relied on (7.24) and immediately gives the claim (2). □

Lemma 7.23. *Let σ be a strictly weak graded Riemannian metric on a Hilbertian H-type group $\mathbb{M}=\mathbb{V}\oplus\mathbb{W}$. Then there exists $w_n\in A_{\mathbb{V}}(\mathbb{V})$ such that $|w_n|=1$ and $\|w_n\|_{\sigma_0}\to 0$ as $n\to+\infty$.*

Proof. We consider the sequence h_n given by Lemma 7.12, hence $\|h_n\|_{\sigma_0}\to 0$ and $|h_n|=1$. The image $A_{\mathbb{V}}(\mathbb{V})$ is dense in $\mathbb{V}$, as a consequence of Proposition 7.9. Therefore, for each $n\in\mathbb{N}\setminus\{0\}$ we may choose $v_n\in A_{\mathbb{V}}(\mathbb{V})$ such that $|v_n-h_n|\le\frac{1}{2n}$, and therefore $|v_n|\to 1$. We define the unit vectors $w_n=\frac{v_n}{|v_n|}$ and consider

$$\|v_n\|_{\sigma_0}\le\|h_n\|_{\sigma_0}+\|h_n-v_n\|_{\sigma_0}\le\|h_n\|_{\sigma_0}+c_0|v_n-h_n|\le\|h_n\|_{\sigma_0}+\frac{c_0}{2n}\to 0,$$

concluding the proof. □

Theorem 7.24. *Let σ be a strictly weak, graded Riemannian metric on a Hilbertian H-type group $\mathbb{M}=\mathbb{V}\oplus\mathbb{W}$. Then there exists a sequence $w_n\in\mathbb{V}$ such that for every $z\in\mathbb{W}\setminus\{0\}$ the following limits hold*

$$\lim_{n\to\infty}K_\sigma(\Pi_{w_n,J_zw_n})=-\infty\quad\text{and}\quad\lim_{n\to\infty}K_\sigma(\Pi_{z,J_{Az}w_n})=+\infty. \tag{7.25}$$

Proof. We consider the sequence $w_n\in A_{\mathbb{V}}(\mathbb{V})\subset\mathbb{V}$ of Lemma 7.23 and define the vector

$$\xi_n=J_zw_n-\frac{w_n}{\|w_n\|_{\sigma_0}}\left\langle J_zw_n,\frac{w_n}{\|w_n\|_{\sigma_0}}\right\rangle_{\sigma_0}\in\mathbb{V}.$$

By construction of ξ_n, the vectors $w_n/\|w_n\|_{\sigma_0}$ and $\xi_n/\|\xi_n\|_{\sigma_0}$ are σ_0-orthonormal and span the 2-dimensional subspace Π_{w_n,J_zw_n}. By Lemma 7.22 and (7.8), we have

$$K_\sigma(\Pi_{w_n,J_zw_n})=-\frac{3}{4}\left\|\left[\frac{w_n}{\|w_n\|_{\sigma_0}},\frac{\xi_n}{\|\xi_n\|_{\sigma_0}}\right]\right\|^2_{\sigma_0}=-\frac{3}{4}\frac{\|z\|^2_{\sigma_0}}{\|w_n\|^2_{\sigma_0}\left\|J_zw_n-\frac{w_n}{\|w_n\|^2_{\sigma_0}}\langle J_zw_n,w_n\rangle_{\sigma_0}\right\|^2_{\sigma_0}}.$$

We consider the estimates

$$\begin{aligned}\left\| J_z w_n - \frac{w_n}{\|w_n\|^2_{\sigma_0}} \langle J_z w_n, w_n\rangle_{\sigma_0} \right\|^2_{\sigma_0} &\le 2\left(\|J_z w_n\|^2_{\sigma_0} + \frac{\langle J_z w_n, w_n\rangle^2_{\sigma_0}}{\|w_n\|^2_{\sigma_0}} \right) \\ &\le 4\,\|J_z w_n\|^2_{\sigma_0} \le 4c_0^2 |J_z w_n|^2 \\ &\le 4c_0^2|z|^2,\end{aligned}$$

where we have applied both (7.17) and (7.4). It follows that

$$\left\|\left[\frac{w_n}{\|w_n\|_{\sigma_0}}, \frac{\xi_n}{\|\xi_n\|_{\sigma_0}} \right]\right\|^2_{\sigma_0} \ge \frac{1}{4c_0^2|z|^2} \frac{\|z\|^2_{\sigma_0}}{\|w_n\|^2_{\sigma_0}} \to +\infty,$$

proving the first limit of (7.25) To establish the second limit of (7.25), we consider the same previous sequence $w_n \in A_{\mathbb{W}}(\mathbb{W})$, along with $v_n \in \mathbb{W}$ such that $Av_n = w_n$. By Lemma 7.22 we have

$$K_\sigma(\Pi_{z,J_{Az}(Av_n)}) = \frac{1}{4} \frac{|Az|^4}{\|J_{Az}(Av_n)\|^2_{\sigma_0}\,\|z\|^2_{\sigma_0}}\, \|v_n\|^2_{\sigma_0}\,.$$

Again (7.17) and (7.4) give the inequalities

$$\frac{1}{\|J_{Az}(Av_n)\|_{\sigma_0}} \ge \frac{1}{c_0|J_{Az}(Av_n)|} = \frac{1}{c_0|Az||Av_n|} = \frac{1}{c_0|Az|} > 0,$$

where we have also use the condition $|w_n| = |Av_n| = 1$. By Cauchy-Schwarz inequality, we get

$$\|v_n\|_{\sigma_0} \ge \left\langle v_n, \frac{w_n}{\|w_n\|_{\sigma_0}} \right\rangle_{\sigma_0} = \langle Av_n, w_n\rangle \frac{1}{\|w_n\|_{\sigma_0}} = \frac{1}{\|w_n\|_{\sigma_0}} \to +\infty,$$

that immediately yields $K_\sigma(\Pi_{z,J_{Az}w_n}) \to +\infty$, concluding the proof. □

Bibliography

[1] R. H. Abraham, J. E. Marsden, and T. S. Ratiu. *Manifolds, tensor analysis, and applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988.

[2] A. Agrachev. Compactness for sub-Riemannian length-minimizers and subanalyticity. volume 56, pages 1–12 (2001). 1998. Control theory and its applications (Grado, 1998).

[3] A. Agrachev. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.

[4] A. Agrachev. Some open problems. In *Geometric control theory and sub-Riemannian geometry*, volume 5 of *Springer INdAM Ser.*, pages 1–13. Springer, Cham, 2014.

[5] A. Agrachev, D. Barilari, and U. Boscain. *A comprehensive introduction to sub-Riemannian geometry*, volume 181 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2020. From the Hamiltonian viewpoint, With an appendix by Igor Zelenko.

[6] A. Agrachev and J.-P. Gauthier. On the subanalyticity of Carnot-Caratheodory distances. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 18(3):359–382, 2001.

[7] A. Agrachev, A. Gentile, and A. Lerario. Geodesics and horizontal-path spaces in Carnot groups. *Geom. Topol.*, 19(3):1569–1630, 2015.

[8] A. Agrachev and P. Lee. Optimal transportation under nonholonomic constraints. *Trans. Amer. Math. Soc.*, 361(11):6019–6047, 2009.

[9] A. Agrachev and A. V. Sarychev. Abnormal sub-Riemannian geodesics: Morse index and rigidity. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 13(6):635–690, 1996.

[10] V. I. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16(fasc., fasc. 1):319–361, 1966.

[11] D. Barilari and L. Rizzi. Sub-Riemannian interpolation inequalities. *Invent. Math.*, 215(3):977–1038, 2019.

[12] S. Basu and A. Lerario. Hausdorff approximations and volume of tubes of singular algebraic sets. *Math. Ann.*, 387(1-2):79–109, 2023.

[13] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[14] S. Basu and M.-F. Roy. Divide and conquer roadmap for algebraic sets. *Discrete Comput. Geom.*, 52(2):278–343, 2014.

[15] S. M. Bates. On smooth rank-1 mappings of Banach spaces onto the plane. *J. Differential Geom.*, 37(3):729–733, 1993.

[16] M. Bauer, M. Bruveris, P. Harms, and P. W. Michor. Vanishing geodesic distance for the Riemannian metric with geodesic equation the KdV-equation. *Ann. Global Anal. Geom.*, 41(4):461–472, 2012.

[17] M. Bauer, M. Bruveris, P. Harms, and P. W. Michor. Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group. *Ann. Global Anal. Geom.*, 44(1):5–21, 2013.

[18] M. Bauer, M. Bruveris, and P. W. Michor. Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group. II. *Ann. Global Anal. Geom.*, 44(4):361–368, 2013.

[19] M. Bauer, M. Bruveris, and P. W. Michor. Homogeneous Sobolev metric of order one on diffeomorphism groups on real line. *J. Nonlinear Sci.*, 24(5):769–808, 2014.

[20] M. Bauer, M. Bruveris, and P. W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vision*, 50(1-2):60–97, 2014.

[21] M. Bauer, P. Harms, and S. C. Preston. Vanishing Distance Phenomena and the Geometric Approach to SQG. *Arch. Ration. Mech. Anal.*, 235(3):1445–1466, 2020.

[22] A. Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.

[23] A. Belotto da Silva, A. Figalli, A. Parusiński, and L. Rifford. Strong Sard conjecture and regularity of singular minimizing geodesics for analytic sub-Riemannian structures in dimension 3. *Invent. Math.*, 229(1):395–448, 2022.

[24] A. Belotto da Silva, A. Parusiński, and L. Rifford. Abnormal subanalytic distributions and minimal rank sard conjecture. *arXiv preprint*, 2022.

[25] A. Belotto da Silva, A. Parusiński, and L. Rifford. Abnormal singular foliations and the sard conjecture for generic co-rank one distributions. *arXiv preprint*, 2023.

[26] A. Belotto da Silva and L. Rifford. The Sard conjecture on Martinet surfaces. *Duke Math. J.*, 167(8):1433–1471, 2018.

[27] L. Biliotti and F. Mercuri. Riemannian Hilbert manifolds. In *Hermitian-Grassmannian submanifolds*, volume 203 of *Springer Proc. Math. Stat.*, pages 261–271. Springer, Singapore, 2017.

[28] F. Boarotto and A. Lerario. Homotopy properties of horizontal path spaces and a theorem of Serre in subriemannian geometry. *Comm. Anal. Geom.*, 25(2):269–301, 2017.

[29] F. Boarotto, L. Nalon, and D. Vittone. The Sard problem in step 2 and in filiform Carnot groups. *ESAIM Control Optim. Calc. Var.*, 28:Paper No. 75, 20, 2022.

[30] F. Boarotto and D. Vittone. A dynamical approach to the Sard problem in Carnot groups. *J. Differential Equations*, 269(6):4998–5033, 2020.

[31] J. Bochnak, M. Coste, and M.-F. Roy. *Géométrie algébrique réelle (Second edition in english: Real Algebraic Geometry)*, volume 12 (36) of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas].* Springer-Verlag, Berlin, 1987 (1998).

[32] C. W. Brown and J. H. Davenport. The complexity of quantifier elimination and cylindrical algebraic decomposition. In *ISSAC 2007*, pages 54–60. ACM, New York, 2007.

[33] M. Bruveris. Riemannian geometry on manifolds of maps. *Notes for a short course held at the summer school Mathematics of Shapes at the IMS in Singapore, July 2016.*

[34] Y. Chitour, F. Jean, and E. Trélat. Genericity results for singular curves. *J. Differential Geom.*, 73(1):45–73, 2006.

[35] F. H. Clarke. On the inverse function theorem. *Pacific J. Math.*, 64(1):97–102, 1976.

[36] J. H. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. *J. Symbolic Comput.*, 5(1-2):29–35, 1988.

[37] B. K. Driver and M. Gordina. Heat kernel analysis on infinite-dimensional Heisenberg groups. *J. Funct. Anal.*, 255(9):2395–2461, 2008.

[38] Y. Eliashberg and L. Polterovich. Bi-invariant metrics on the group of Hamiltonian diffeomorphisms. *Internat. J. Math.*, 4(5):727–738, 1993.

[39] H. Federer. The (φ, k) rectifiable subsets of n-space. *Trans. Amer. Math. Soc.*, 62:114–192, 1947.

[40] H. Federer. *Geometric Measure Theory*. Springer, 1969.

[41] A. Gabrielov and N. Vorobjov. Approximation of definable sets by compact families, and upper bounds on homotopy and homology. *J. Lond. Math. Soc. (2)*, 80(1):35–54, 2009.

[42] H. Glöckner. Fundamental problems in the theory of infinite-dimensional Lie groups. *J. Geom. Symmetry Phys.*, 5:24–35, 2006.

[43] E. Grong, I. Markina, and A. Vasil'ev. Sub-Riemannian geometry on infinite-dimensional manifolds. *J. Geom. Anal.*, 25(4):2474–2515, 2015.

[44] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[45] G. Hörmann. *Representations of the Infinite Dimensional Heisenberg Group*. PhD thesis, University of Vienna, 1993.

[46] R. L. Jerrard and C. Maor. Geodesic distance for right-invariant metrics on diffeomorphism groups: critical Sobolev exponents. *Ann. Global Anal. Geom.*, 56(2):351–360, 2019.

[47] R. L. Jerrard and C. Maor. Vanishing geodesic distance for right-invariant Sobolev metrics on diffeomorphism groups. *Ann. Global Anal. Geom.*, 55(4):631–656, 2019.

[48] A. Kaplan. Fundamental solutions for a class of hypoelliptic PDE generated by composition of quadratic forms. *Trans. Amer. Math. Soc.*, 258(1):147–153, 1980.

[49] A. Kaplan. Riemannian nilmanifolds attached to Clifford modules. *Geom. Dedicata*, 11(2):127–136, 1981.

[50] A. Kaplan. On the geometry of groups of Heisenberg type. *Bull. London Math. Soc.*, 15(1):35–42, 1983.

[51] A. Kaplan and F. Ricci. Harmonic analysis on groups of Heisenberg type. In *Harmonic analysis (Cortona, 1982)*, volume 992 of *Lecture Notes in Math.*, pages 416–435. Springer, Berlin, 1983.

[52] O. D. Kellogg. On bounded polynomials in several variables. *Math. Z.*, 27(1):55–64, 1928.

[53] B. Khesin and R. Wendt. *The geometry of infinite-dimensional groups*, volume 51 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2009.

[54] W. P. A. Klingenberg. *Riemannian geometry*, volume 1 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 1995.

[55] A. Kriegl and P. W. Michor. *The convenient setting of global analysis*. Mathematical Surveys and Monographs 053. AMS, 1997.

[56] G. K. Kristiansen. Characterization of polynomials by means of their stationary values. *Arch. Math. (Basel)*, 43(1):44–48, 1984.

[57] H. Kuhn. Interpolation vorgeschriebener Extremwerte. *J. Reine Angew. Math.*, 238:24–31, 1969.

[58] I. Kupka. Counterexample to the Morse-Sard theorem in the case of infinite-dimensional manifolds. *Proc. Amer. Math. Soc.*, 16:954–957, 1965.

[59] S. Lang. *Real and functional analysis*, volume 142 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 1993.

[60] E. Le Donne, R. Montgomery, A. Ottazzi, P. Pansu, and D. Vittone. Sard property for the endpoint map on some Carnot groups. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 33(6):1639–1666, 2016.

[61] A. Lerario, L. Rizzi, and D. Tiberio. Quantitative approximate definable choices. *arXiv preprint*, 2024.

[62] A. Lerario, L. Rizzi, and D. Tiberio. Sard properties for polynomial maps in infinite dimension. *arXiv preprint*, 2024.

[63] L. Lokutsievskiy and M. Zelikin. Derivatives of sub-Riemannian geodesics are L_p-Hölder continuous. *ESAIM Control Optim. Calc. Var.*, 29:Paper No. 70, 30, 2023.

[64] V. Magnani and T. Rajala. Radon-Nikodym property and area formula for Banach homogeneous group targets. *Int. Math. Res. Not. IMRN*, (23):6399–6430, 2014.

[65] V. Magnani and D. Tiberio. A remark on vanishing geodesic distances in infinite dimensions. *Proc. Amer. Math. Soc.*, 148(8):3653–3656, 2020.

[66] V. Magnani and D. Tiberio. On the Michor-Mumford phenomenon in the infinite dimensional Heisenberg group. *Rev. Mat. Complut.*, 36(3):973–989, 2023.

[67] V. Magnani and D. Tiberio. The Michor-Mumford conjecture in Hilbertian H-type groups. *arXiv preprint*, 2024.

[68] P. Malliavin. Hypoellipticity in infinite dimensions. In *Diffusion processes and related problems in analysis, Vol. I (Evanston, IL, 1989)*, volume 22 of *Progr. Probab.*, pages 17–31. Birkhäuser Boston, Boston, MA, 1990.

[69] C. Maor. Riemannian Geometry of Diffeomorphism Groups – Lecture notes. *Notes from a course held at the Hebrew University, January 2022*.

[70] J. N. Mather. Generic projections. *Ann. of Math. (2)*, 98:226–245, 1973.

[71] P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math.*, 10:217–245, 2005.

[72] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8(1):1–48, 2006.

[73] P. W. Michor and T. S. Ratiu. On the geometry of the Virasoro-Bott group. *J. Lie Theory*, 8(2):293–309, 1998.

[74] J. Milnor. Curvatures of left invariant metrics on Lie groups. *Advances in Math.*, 21(3):293–329, 1976.

[75] J. Milnor. Remarks on infinite dimensional Lie groups. In *Relativity, groups and topology: Proceedings, 40th Summer School of Theoretical Physics - Session 40: Les Houches, France, June 27 - August 4, 1983, vol. 2*, pages 1007–1057, 1984.

[76] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[77] A. P. Morse. The behavior of a function on its critical set. *Ann. of Math. (2)*, 40(1):62–70, 1939.

[78] K.-H. Neeb. Towards a Lie theory of locally convex groups. *Jpn. J. Math.*, 1(2):291–468, 2006.

[79] H. Omori. *Infinite-dimensional Lie groups*, volume 158 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1997. Translated from the 1979 Japanese original and revised by the author.

[80] A. Ottazzi and D. Vittone. On the codimension of the abnormal set in step two Carnot groups. *ESAIM Control Optim. Calc. Var.*, 25:Paper No. 18, 17, 2019.

[81] A. Pinkus. *n-widths in approximation theory*, volume 7 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1985.

[82] L. Rifford. Subdifferentials and minimizing Sard conjecture in sub-Riemannian geometry. *J. Éc. polytech. Math.*, 10:1195–1244, 2023.

[83] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.

[84] A. Sard. The measure of the critical values of differentiable maps. *Bull. Amer. Math. Soc.*, 48:883–890, 1942.

[85] A. Sard. Images of critical sets. *Annals of Mathematics*, 68(2):247–259, 1958.

[86] A. Sard. Hausdorff measure of critical images on banach manifolds. *American Journal of Mathematics*, 87(1):158–174, 1965.

[87] A. Schmeding. *An introduction to infinite-dimensional differential geometry*, volume 202 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2023.

[88] S. Smale. An infinite dimensional version of Sard's theorem. *Amer. J. Math.*, 87:861–866, 1965.

[89] H. J. Sussmann. A regularity theorem for minimizers of real-analytic subriemannian metrics. In *53rd IEEE Conference on Decision and Control*, pages 4801–4806. 2014.

[90] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[91] A. G. Vituškin. *O mnogomernyh variaciyah*. Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1955.

[92] A. G. Vituškin. *Theory of the transmission and processing of information*. Pergamon Press, New York-Oxford-London-Paris, 1961. Translated from the Russian by Ruth Feinstein; translation editor A. D. Booth.

[93] R. J. Walker. *Algebraic curves*. Springer-Verlag, New York-Heidelberg, 1978. Reprint of the 1950 edition.

[94] Y. Yomdin. The geometry of critical and near-critical values of differentiable mappings. *Math. Ann.*, 264(4):495–515, 1983.

[95] Y. Yomdin. Approximational complexity of functions. In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 21–43. Springer, Berlin, 1988.

[96] Y. Yomdin and G. Comte. *Tame geometry with application in smooth analysis*, volume 1834 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2004.

[97] I. Zelenko and M. Zhitomirskiĭ. Rigid paths of generic 2-distributions on 3-manifolds. *Duke Math. J.*, 79(2):281–307, 1995.

Acknowledgements

Thinking about the conclusion of this journey, the first people I want to thank are my thesis advisors, Professor Antonio Lerario and Professor Luca Rizzi. First of all, I thank them for introducing me to sub-Riemannian and semialgebraic geometry, and for believing in me and our project at every moment. I thank them for all the time, passion, and dedication they have invested in my education, and especially for being constant points of reference throughout the entire journey. I thank them for showing me how to conduct research. I am certain that what Luca and Antonio have imparted to me over these years will forever influence my work. Even after the end of my PhD, they will always remain models to follow.

I thank my master's thesis advisor, Professor Valentino Magnani, for always encouraging me over these years, for imparting his great passion and dedication for Mathematics to me, and above all, for continuing our collaboration even during the years of my PhD, continuing to contribute to my education.

A special thanks goes to Antonio, Ariel, Dario, Davide, Fabrizio, Harschit, Irene, Lorenzo, Maria Teresa, Michele, Roberto, Thomas. It has been a pleasure to share the office and most of the days at SISSA with you. Your presence has enriched this journey daily.

I thank the research groups of my supervisors. Feeling part of such passionate, close-knit, and motivated groups was fundamental during my PhD. In particular, I thank Giorgio and Samuel for involving me in a new sub-Riemannian research line, and for being excellent group mates and collaborators.

I thank the entire SISSA for the stimulating environment it represents for students.

I thank the European Research Council (ERC) for partially funding my research activities through the European Union's Horizon 2020 research and innovation programme (grant agreement No. 945655).

Ringrazio Adriano, Emanuele, Matteo e Michele per essere stati sempre presenti anche se distanti. Ringrazio Alessandro, Edoardo e Luca, e tutti coloro con cui ho condiviso questi anni a Trieste.

E soprattutto, ringrazio la mia famiglia per il supporto costante e per i valori che mi hanno trasmesso, i quali sono stati gli ingredienti principali di tutto il mio percorso.