ARENBERG DOCTORAL SCHOOL
Faculty of Science

Parametrizations of power-subanalytic sets

Siegfried Van Hille

Supervisor:
Prof. dr. R. Cluckers
(KU Leuven)
Co-supervisor:
Prof. dr. W. Veys
(KU Leuven)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Science (PhD): Mathematics

July 2021

Parametrizations of power-subanalytic sets

Siegfried VAN HILLE

Examination committee:
Prof. dr. J. Van der Veken, chair (KU Leuven)
Prof. dr. R. Cluckers, supervisor (KU Leuven)
Prof. dr. W. Veys, co-supervisor (KU Leuven)
Prof. dr. N. Budur (KU Leuven)
Prof. dr. M. Zambon (KU Leuven)
Prof. dr. G.O. Jones (University of Manchester)
Prof. dr. P. Speissegger (McMaster University)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Science (PhD): Mathematics

July 2021

Uitgegeven in eigen beheer, Siegfried Van Hille, Celestijnenlaan 200B box 2400, B-3001 Leuven (Belgium)

Preface

juni 2021

Na 9 jaar studeren komt er een einde aan het mooie studentenleven met het schrijven van deze doctoraatsthesis, in het bijzonder van dit voorwoord. De ideale gelegenheid voor een korte terugblik op mijn studententijd en de uitgelezen kans om enkele mensen in de bloemetjes te zetten.

Eind augustus 2017, klaar om leerkracht wiskunde te worden, kreeg ik alsnog een bericht van Wim of we "even konden bellen". Al snel werd wereldstad Antwerpen ingewisseld voor wereldstudentenstad Leuven, medestudenten voor collega's en pendelen eindelijk voor het verblijven op kot. Ik had geen idee waar ik aan begonnen was, maar voelde me al snel als een vis in het water. Personen die mij in het bijzonder welkom deden voelen waren: mede-assistenten Robin en Saskia, mijn kantoorgenoot Haopeng en Almateamverantwoordelijke Jan. Ik zal ook niet snel vergeten dat Wim het niet kon laten om op mijn eerste POC te zeggen dat ik eindelijk het licht gezien had met mijn komst naar Leuven.

Als ik alle leuke verhalen hier zou moeten neerpennen, dan zou dit voorwoord mogelijks langer worden, misschien zelfs interessanter, dan de inhoud van de thesis. Laat ik er daarom enkele opsommen, zij die aanwezig waren weten daarmee genoeg. Memorabel waren: de workshop in Marseille, de teambuildings van de afdeling algebra, de pizzalezingen, de conferentie in Santander, de vele recepties, enkele cantussen, sportactiviteiten... Over sommige zaken zou ik hier nog uitgebreid kunnen vertellen, maar het is misschien beter om dat niet te doen.

Mijn grootste dank gaat uit aan Raf, mijn promotor, en aan Wim, mijn co-promotor en dagelijks leidinggevende (of kortweg: baas). Op de vraag waarom ze deze inwoner van de provincie Antwerpen aangenomen hebben, zal ik allicht nooit een goed antwoord vinden. Ik ben hen zeer dankbaar voor de kans die ze mij gegeven hebben en kijk met een tevreden blik terug op de voorbije 4 jaar aan de KU Leuven. Ik dank Raf voor het interessante en mij toen nog onbekende onderwerp voor mijn onderzoek, zijn toegankelijkheid en zijn bereidheid om uitgebreid op mijn vragen te antwoorden. Mijn bewondering voor Wim om niet enkel mijn technische en wellicht saaie bewijzen te doorgronden, maar vooral om meer collega dan baas te zijn.

Voor de formaliteit, maar daarom niet minder gemeend, bedank ik de volgende collega's voor de vele, al dan niet ongevraagde, koffiepauzes en babbels: Haopeng, Marlies, Robin, Lena, Naud, Irene, Saskia en Philip. Het was zeer jammer dat dit in de laatste twee jaar niet meer face-to-face kon verlopen. Dat hield me echter niet tegen om dat ook digitaal te doen, waarvoor ik Anne, Kristof, Irene, Jolien en Naud hartelijk bedank. Ook een eervolle vermelding voor de harde kern van het Almateam: Jan, Eva en Wim.

Vervolgens dank ik natuurlijk enkele vrienden. Pauline, Stef en Zara zouden wel een oscar kunnen winnen voor het geïnteresseerd luisteren naar de status en de details van mijn onderzoek. Ik dank ook Stef en Zara voor de vele digitale contacten in het "nieuwe normaal", dat hopelijk snel terugkeert naar het oude normaal. Het was steeds een plezier om calorieën te verbranden met Anne en Lena, die ik beiden ook dank voor de vele wandelingen tijdens dat nieuwe normaal. Ik kan Anne niet genoeg bedanken voor de vele uren die we samen hebben doorgebracht tijdens de lockdowns.

Last but not least, I would like to say that I have met many people during the past 4 years, which were all really nice. I thank all members of the Examination Committee for their commitment, especially Gareth and Patrick.

Met deze woorden sluit ik dit avontuur nu af. Ik kijk uit naar het volgende aan het Fields Institute in Canada.

Siegfried

Abstract

The main results of this thesis are two uniform parametrization theorems for families of bounded real power-subanalytic sets. The first one is a more precise statement of a recent C^r-parametrization result by Cluckers, Pila and Wilkie. They have shown that it is possible to construct such a parametrization with an amount of charts that is polynomial in r. I show that this exponent is polynomial in the dimension of the set that is being parametrized.

The second result is a uniform mild parametrization result for curves. More precisely, for any $C > 0$, a family of bounded power-subanalytic curves has a uniform C-mild parametrization. This improves a recent result by Binyamini and Novikov, who have shown that a family of bounded subanalytic curves have a uniform 2-mild parametrization.

Both results are deduced from a pre-parametrization result by Cluckers, Pila and Wilkie, when combined with precise results on mild functions. Most notably are a result on compositions, that can be deduced by translating a similar result on the composition of Gevrey functions or analytic functions, and a result that a particular function, that is used in the construction of the mild parametrization, is suitably mild.

Some knowledge of o-minimality is required to fully understand the proofs and is included in the thesis. The final chapter consists of a brief summary of the well known applications in diophantine geometry.

Samenvatting

De hoofdresultaten van deze thesis zijn twee uniforme parametrisatiestellingen voor families van begrensde reële power-subanalytische verzamelingen. De eerste is een meer precieze versie van een recent C^r-parametrisatieresultaat van Cluckers, Pila en Wilkie. Zij hebben bewezen dat het mogelijk is om een parametrizatie te construeren bestaande uit een aantal kaarten dat polynomiaal in r is. Ik toon aan dat de exponent in deze veelterm polynomiaal is in de dimensie van de verzameling die geparametriseerd wordt.

Het tweede resultaat is een uniform milde parametrisatiestelling voor krommen. Meer precies heeft een familie van begrensde power-subanalytische krommen een C-milde parametrisatie voor elke $C > 0$. Dit verbetert een recent resultaat van Binyamini en Novikov, namelijk dat een familie van begrensde subanalytische krommen een uniforme 2-milde parametrisatie heeft.

Beide resultaten worden afgeleid uit een pre-parametrisatiestelling van Cluckers, Pila en Wilkie via nauwkeurige resultaten over milde functies. Meest noemenswaardig zijn een resultaat over samenstellingen, dat volgt uit een gelijkaardig resultaat voor Gevreyfuncties of analytische functies, en de mildheid van een specifieke functie, die gebruikt wordt in de constructie van de milde parametrizatie.

Voorkennis van o-minimaliteit is vereist om de bewijzen goed te kunnen begrijpen en is daarom in de thesis opgenomen. Het laatste hoofdstuk bestaat uit een kort overzicht van de welbekende toepassingen in diophantische meetkunde.

Contents

Introduction

There are two main results in this thesis. The first is a uniform C^r-parametrization theorem for bounded power-subanalytic sets, where we cover a set with finitely many images of C^r-maps that have bounded derivatives. The second is a uniform mild parametrization theorem for bounded power-subanalytic curves, where the functions are C^∞ and there is a bound on all derivatives. Precise statements can be found below. I also provide a short overview of applications of parametrizations in diophantine geometry in the last chapter.

Main results

Consider a set $\mathcal{X} \subset \mathbb{R}^k \times \mathbb{R}^n$, where the coordinates in $\mathbb{R}^k$ should be considered as parameters, so $\mathcal{X}$ is a family of subsets of $\mathbb{R}^n$. Denote T for the projection of $\mathcal{X}$ onto the first k coordinates, i.e., T is the parameter space. For each $t \in T$, denote

$$X_t = \{x \in \mathbb{R}^n \mid (t, x) \in \mathcal{X}\}.$$

Suppose the largest dimension among the family members X_t is m, which is of course at most n. A **uniform parametrization** of $\mathcal{X}$ is a *finite* set of functions

$$\{\phi_i : T \times (0,1)^m \to \mathcal{X} \mid i \in \{1, \dots, N\}\}$$

for some $N \in \mathbb{N}$, such that for each $t \in T$

$$\bigcup_{i=1}^{N} \mathrm{Im}(\phi_{i,t}) = X_t,$$

where $\phi_{i,t} : (0,1)^m \to X_t$ is defined by $\phi_{i,t}(x) = \phi_i(t, x)$.

If no condition is put on the maps ϕ_i, which we call **charts**, then this is not an interesting notion. We consider two possible conditions. The first one states that all derivatives up to a certain order $r \in \mathbb{N}$ should be bounded. Therefore, the charts are of course required to be suitably differentiable. Imposing this condition on the charts, we obtain the notion of **C^r-parametrizations**, see Definition 3.1.3.

Theorem (Main Theorem 1)**.** Suppose $\mathcal{X}$ is a power-subanalytic family of subsets of $(0,1)^n$ of dimension at most m. Then $\mathcal{X}$ has a uniform C^r-parametrization, moreover the amount of charts N is at most cr^d, where c is some constant that depends on $\mathcal{X}$ and d is polynomial in m.

A precise definition of "power-subanalytic" is given in Section 1.2. Roughly speaking it says $\mathcal{X}$ can be defined in terms of polynomials, analytic functions on a compact domain and (possibly irrational) power functions. The theorem is proved in Section 3.3. Actually more is shown, see Theorem 3.3.1, but these details are not so important. My contribution to Theorem 3.3.1, which is a result by Cluckers, Pila and Wilkie, is that d is polynomial in m, i.e., this makes the amount of charts N more explicit. This is the main result of my article [VH21b] and was the initial goal of my PhD project.

The second main result of this thesis is a uniform parametrization result where one imposes an upper bound on *all* order derivatives of ϕ_i. More precisely, for $A, B > 0$ and $C \geq 0$, say that a function $f : U \subset \mathbb{R}^m \to \mathbb{R}$ is **(A, B, C)-mild** if f is infinitely differentiable on U and for any $\alpha \in \mathbb{N}^m$ and $x \in U$

$$\left| f^{(\alpha)}(x) \right| \leq (BA^{|\alpha|} \, |\alpha|!)^{1+C}.$$

See Section 2.1 for definitions and notations regarding multivariate calculus. A map $f : U \to \mathbb{R}^n$ is (A, B, C)-mild if all of its component functions are. A uniform (A, B, C)-**mild parametrization** is a uniform parametrization such that for any $i \in \{1, \ldots, N\}$ and $t \in T$, the chart $\phi_{i,t}$ is (A, B, C)-mild, see also Definition 3.1.5.

Theorem (Main Theorem 2)**.** Suppose $\mathcal{X}$ is a power-subanalytic family of curves in $(0,1)^n$. Then for any $0 < C \leq 1$, there exist A and B, depending on $\mathcal{X}$ only, such that $\mathcal{X}$ has a uniform $(A/C, B, C)$-mild parametrization.

The proof of this theorem is given in Section 3.5, see Theorem 3.5.1. The proof strategy is completely the same as the proof of Main Theorem 1. A key ingredient is the following result.

Proposition. Let $\kappa \geq 1$ and consider the function $f : (0,1) \to (0,1)$ defined by

$$f(x) = e^{1 - \frac{1}{x^\kappa}}.$$

Then f is $(6\kappa, e, 1/\kappa)$-mild.

This is shown in Section 2.3. It is a new result, in the sense that other results only obtain $C = 1+1/\kappa$, which makes the bounds on the derivatives significantly larger. Together with the mild parametrization theorem, this is the main content of my article [VH21a].

The actual introduction

If a set X has a parametrization, one could say that X is a manifold with a finite atlas. This notion is already quite old and manifolds are of course still extensively studied. However, in the field of differential geometry one is interested in the geometry of X, rather independent of (the cardinality of) the atlas, while we are exactly interested in the latter.

In [Yom87b], Yomdin proved Shub's entropy conjecture for C^∞-maps. That is, for a smooth compact C^∞-manifold M and C^∞-map $f : M \to M$, the logarithm of the spectral radius of $f_* : H_*(M, \mathbb{R}) \to H_*(M, \mathbb{R})$ is a lower bound for the topological entropy of f (see [BN19, Section 1.3], which also contains a brief introduction on this subject, based on the exposition in [BLY15]). A key ingredient in his proof was a parametrization theorem, now often referred to as the "Yomdin-Gromov Algebraic Lemma". Yomdin showed this result in [Yom87a], see also the article [Gro87] of Gromov on the results of Yomdin. In Section 3.2 this technique is applied to a simple example, following Yomdin's exposition in [Yom15].

Yomdin considered in [Yom91] the case of a real analytic function $f : M \to M$, where M is a compact analytic surface, and proved a result similar to [Yom87b]. The improvement in the analytic case is an upper bound on the rate of convergence to zero of the "ϵ-tail entropy". He conjectured that it would also hold in higher dimensions. His result was based on a holomorphic version of the Algebraic Lemma, see [BN19, Section 1.3].

So, although maybe by coincidence, this first motivation comes from smooth dynamics. Since I am not familiar with these results, I won't give many more details, but we will revisit it shortly later in this introduction. The details here are mostly taken from [BN19, Section 1.3], which is based on the content of [BLY15]. The second motivation comes from number theory and was being developed around the same time. In this case however, as far as I know, it took a little bit longer before a parametrization result was used.

In [BP89], Pila and Bombieri proved some results on the distribution of rational points on curves, say $y = f(x)$ for some smooth or analytic function f. Precise statements can be found in the introduction of Chapter 4. Let me summarize them here by saying that if f is algebraic, then there is polynomial growth of rational points on the curve, and if f is transcendental, it is subpolynomial. It was already mentioned in their work that their result could be generalized to higher dimensions (for affine and projective varieties) by "simple slicing arguments" and Pila did so in [Pil95]. Note that the parametrization was assumed to be given and was not constructed.

Pila generalized his work with Bombieri to compact subanalytic surfaces in [Pil04] and [Pil05]. A closer look at these papers shows that a parametrization result is used, that is deduced from the uniformization theorem from [BM88]. This uniformization theorem states that every closed subanalytic subset X of an analytic manifold M is the image under a proper real analytic map $\psi : N \to M$, where N is an analytic manifold of the same dimension as X.

In [Wil04], Wilkie showed a similar result as in [BP89], using different methods, for curves that are definable in o-minimal expansions of the real field. See Chapter 1 for a short introduction to o-minimality. A very fruitful collaboration between Pila and Wilkie in [PW06] yielded a theorem that holds for a large class of sets, namely definable in an o-minimal structure, that generalized their two results. It is now well known as the Pila-Wilkie Counting Theorem, we discuss it in Section 4.1. Roughly speaking, it states that rational points on the transcendental part of a definable set grow subpolynomially.

The main new ingredient in the work of Pila and Wilkie is a uniform C^r-parametrization theorem for (bounded) definable sets that holds in any expansion of the real ordered field. The fact that is it uniform, in contrast to the parametrization used by Pila in [Pil04, Pil05], makes it possible to prove the result for sets of any arbitrary dimension, see the discussion below [PW06, Theorem 1.8].

The construction of this uniform C^r-parametrization theorem uses the construction by Yomdin and Gromov. Indeed, this construction can be adapted to this general setting since sets that are definable in an o-minimal structure share many of the nice geometric properties of semi-algebraic sets. In fact, semi-algebraic sets are the role model of o-minimal structures on $\mathbb{R}$. The most basic results in o-minimality are stated in Section 1.3.

However, there is an important benefit when just working with semi-algebraic sets. Since these sets are defined by polynomial equalities and inequalities, these sets naturally have a notion of "complexity". The method developed by Pila and Bombieri consists of intersecting the set X with algebraic hypersurfaces of suitable degree, such that all rational points (up to a certain "height") lie on at least one of these hypersurfaces. One then considers all the intersections and can proceed by induction, at least if there is some information about these intersections. For instance if it is expressed in terms of complexity, as in the semi-algebraic case as in [Pil95], or if one can parametrize it again, where the amount of charts of the parametrization is independent of the intersection, i.e., a uniform parametrization (this is explained below [PW06, Theorem 1.8]).

There is a larger class of functions that also have a natural notion of complexity, playing the role of degree for polynomials, namely Pfaffian functions, see [Kho91]. I will not formally define them here. Shortly stated these are functions whose derivatives satisfy a polynomial relation in their lower order derivatives. For instance, the exponential function, for which $f' - f = 0$. By [Spe99], the structure of Pfaffian sets, say sets defined by Pfaffian functions, is o-minimal. Pila successfully obtained results for Pfaffian curves, in the same spirit as before, in [Pil06a, Pil06b, Pil07].

In [Pil06a, Pil07], Pila did not use the C^r-parametrization theorem, but put the existence of a new type of parametrization, a mild parametrization, as an assumption. In particular he conjectured that every Pfaffian curve has a mild parametrization. As far as I know, this conjecture is still open in general. In fact, the theory of mild parametrizations seems to be more delicate than the theory of C^r-parametrizations. We will further discuss this in the introduction of Chapter 3.

Let me point out that the second main result of this thesis poses that this conjecture is true for power-subanalytic curves by the main result of [VH21a], but this is still far away from Pfaffian curves. Also, remark that power-subanalytic curves are not a subset of Pfaffian curves or vice versa.

The advantage of working with functions that have a notion of complexity is that one can express the amount of charts N of the C^r-parametrization in terms of this complexity, or at least it is natural to ask this question. This is indeed the case for the C^r-parametrization theorem by Yomdin and Gromov. Say X is an m-dimensional semi-algebraic subset of $(0,1)^n$ of complexity β (whatever complexity exactly is), then N is a constant depending on the complexity only. In particular, it does not depend on the coefficients of the polynomials that define X.

It turns out that it is interesting to investigate how this constant depends on these data. In [BLY15], they also prove the result of Yomdin for analytic maps in [Yom91], but only using C^r-parametrizations, so not using holomorphic extensions. They show that if $m = 1$, then N can be taken polynomially in r and n [BLY15, Proposition 3.7], and this allows them to recover the result of Yomdin. In [BN19] it is shown that for semi-algebraic sets X, one can take $N = \mathrm{poly}_n(m, \beta) r^m$, which allows them to deduce the result of Yomdin for compact analytic manifolds of arbitrary dimension, see [BN19, Theorem 5]. In fact, they show in [BN19] that a subanalytic set has a uniform C^r-

parametrization, where $N = cr^m$ for some constant $c > 0$. If X is semi-algebraic, one additionally has that $c = \mathrm{poly}_n(m, \beta)$.

A similar C^r-parametrization theorem was obtained by Cluckers, Pila and Wilkie in [CPW20]. Their result is slightly more general, in the sense that it holds for power-subanalytic sets (see Section 1.2 for a definition of power-subanalytic). However, the polynomial dependence on r is not as explicit as in the result in [BN19]. More precisely, their result yields cr^d charts, for some constants $c, d > 0$, but there is no information on the constants c and d, even if one additionally assumes that X is semi-algebraic. According to [BN19, Section 1.3], one can not deduce the results in smooth dynamics without this information, but this was not the goal of [CPW20], where a further refinement of their C^r-parametrization theorem in this fashion is raised as a question, see [CPW20, Remark 4.4.3]. Both papers provide further improvements of the earlier work by Pila, which we will further discuss in Section 4.3.

The first main result of this thesis (partially) answers this question in [CPW20], by showing that $d = m^3$. This is part of my article [VH21b]. This result is obtained by carefully checking the methods in [CPW20], combined with results on mild functions, which we prove in Chapter 2.

About the chapters of the thesis

The thesis is build up as follows. In principle, each chapter can be read mostly independently. With the exception of Chapter 1, each of them has its own thorough introduction. The introduction of a chapter is tailored to the subject there, in the sense that I mostly pretend that it has nothing to do with the other chapters and, if not, refer for the details to another chapter.

In principle, the basics of calculus, in particular differentiation and analytic functions, are sufficient to understand most of this thesis. Chapter 1 gives a brief overview of the field of o-minimality that should be sufficient to understand the arguments in Chapter 3. It is intended for readers that are not familiar with the theory. There are no proofs in Chapter 1 and it is kept to a minimum on purpose. Chapter 2 contains many results on mild functions, in particular on compositions and substitutions. It is independent of Chapter 1.

The heart of the thesis is Chapter 3 on parametrizations. The main results, which we have stated above, are proved in this chapter. Although it heavily relies on the results of the first two chapters, it should be accessible even if one did not read those. Chapter 4 is a short overview of the applications of parametrizations in diophantine geometry. It uses parametrization results, but it is not necessary to have read Chapter 3 in order to understand it.

Chapter 1

O-minimality

This short chapter is intended for readers that are not familiar with the field of o-minimality. It contains the necessary definitions and results that will be used in Chapter 3 on parametrizations. In Section 1.1 we start with a short recap of model theory in order to give a precise definition of semi-algebraic, subanalytic and power-subanalytic sets in Section 1.2. Finally, we review some important geometric results of o-minimality in Section 1.3.

O-minimality is a framework that suites Grothendieck's idea of "Topologie Modérée" (tame topology). In his famous note "Esquisse d'un programme" from 1984, he describes this idea as follows [Gro97].

C'est quelques années plus tard que j'étais informé de la théorie de Hironaka des ensembles qu'il appelle, je crois, "semi-analytiques" (réels), qui satisfont à certaines des conditions de stabilité essentielles (sans doute même à toutes) nécessaires au développement d'un contexte utilisable de "topologie modérée".

Roughly summarizing his idea, one could say that he wanted to find an axiomatic way to obtain a collection of sets that has geometrical properties similar to semi-algebraic and semi-analytic sets. Thus, a call to model theorists to develop such a theory. It turns out that o-minimality is indeed a general framework that yields tame geometry.

Geometrical properties of o-minimal structures, although not yet called this way, were first studied by Lou van den Dries in [vdD84], and a series of papers called *Definable sets in ordered structures* [PS86, KPS86, PS88]. Many of these geometric results can be found in [vdD98], which is the standard reference in the field. It is written in a "geometric" way, avoiding the use of model theory.

1.1 Terminology of model theory

This section is based on the first chapter, "Structures and Theories", of the book "Model Theory: An introduction" by Marker [Mar02].

The aim of this section is to understand what it means that a set $X \subset \mathbb{R}^n$ is definable in a language $\mathcal{L}$. In this way, one can give a very elementary and precise definition of some o-minimal structures in Section 1.2. Our examples and motivation merely focus on what we need in this thesis.

Definition 1.1.1 ([Mar02, Definition 1.1.1])**.** A **language** $\mathcal{L}$ consists of the following data:

1. a set of function symbols $\mathcal{F}$ and positive integers n_f for each $f \in \mathcal{F}$;
2. a set of relation symbols $\mathcal{R}$ and positive integers n_R for each $R \in \mathcal{R}$;
3. a set of constant symbols $\mathcal{C}$.

The numbers n_f are called the arity of the function f and the numbers n_R indicate that R is an n_R-ary relation.

Example 1.1.2.

1. $\mathcal{L}_< = \{<\}$, the language of ordered sets, where $<$ is a binary relation.
2. $\mathcal{L}_r = \{+, -, \cdot, 0, 1\}$, the language of rings, where $+, -$ and $\cdot$ are binary functions and 0 and 1 are constant symbols.
3. $\mathcal{L}_{or} = \{+, -, \cdot, <, 0, 1\}$ the **language of ordered rings**.

The idea is that X is definable in a language $\mathcal{L}$ if there is a formula for X that only consists of symbols of that language. Let us first define what a formula is and then explain how to define sets with it. The building blocks of formulas are called terms and these are defined as follows.

Definition 1.1.3 ([Mar02, Definition 1.1.4])**.** The set of **$\mathcal{L}$-terms** is the smallest set $\mathcal{T}$ such that

1. $c \in \mathcal{T}$ for each $c \in \mathcal{C}$;
2. all variable symbols $x_i \in \mathcal{T}$ for $i = 1, 2, \ldots$, and
3. if $t_1, \ldots, t_{n_f} \in \mathcal{T}$ and $f \in \mathcal{F}$, then $f(t_1, \ldots, t_{n_f}) \in \mathcal{T}$.

For example, $\cdot(x_1, x_2)$ and $-(x_1, 1)$ are $\mathcal{L}_r$-terms, which we will write more compactly as x_1x_2 and $x_1 - 1$ respectively.

Definition 1.1.4 ([Mar02, Definition 1.1.5]). We say that ϕ is an **atomic $\mathcal{L}$-formula** if ϕ is either

1. $t_1 = t_2$, where t_1 and t_2 are $\mathcal{L}$-terms, or
2. $R(t_1, \ldots, t_{n_R})$, where $R \in \mathcal{R}$ and $t_1, \ldots, t_{n_R}$ are $\mathcal{L}$-terms.

Finally, the set of **$\mathcal{L}$-formulas** is the smallest set $\mathcal{W}$ containing the atomic $\mathcal{L}$-formulas such that

1. if $\phi \in \mathcal{W}$, then $\neg\phi \in \mathcal{W}$;
2. if $\phi, \psi \in \mathcal{W}$, then $(\phi \wedge \psi)$ and $(\phi \vee \psi)$ are in $\mathcal{W}$, and
3. if $\phi \in \mathcal{W}$, then $\exists x_i\, \phi$ and $\forall x_i\, \phi$ are in $\mathcal{W}$.

An example of an $\mathcal{L}_r$-formula is $\exists x_2\; x_2x_2 = x_1$, which asserts that x_1 is a square. However, currently x_2x_2 is just an $\mathcal{L}_r$-term, it does not mean anything yet. To this end, we have to give these symbols an interpretation.

Definition 1.1.5 ([Mar02, Definition 1.1.2]). An **$\mathcal{L}$-structure** $\mathcal{M}$ is given by the following data:

1. a nonempty set M, the underlying set of $\mathcal{M}$;
2. a function $f^{\mathcal{M}} : M^{n_f} \to M$ for each $f \in \mathcal{F}$;
3. a set $R^{\mathcal{M}} \subset M^{n_R}$ for each $R \in \mathcal{R}$;
4. an element $c^{\mathcal{M}} \in M$ for each $c \in \mathcal{C}$.

To summarize: an $\mathcal{L}$-structure $\mathcal{M}$ is a set M where all symbols of $\mathcal{L}$ make sense in *some* way. Indeed, we did not impose that the interpretations of these symbols have to satisfy any conditions. So at this point, an $\mathcal{L}_r$-structure is not necessarily a ring. For example we could define $\cdot^{\mathcal{M}} : M \times M \to M : (x_1, x_2) \mapsto 1^{\mathcal{M}}$ and this is perfectly fine.

It is clear that $\mathbb{R}$ is naturally an $\mathcal{L}$-structure for any of the languages of Example 1.1.2 and we will assume it is always interpreted in this way. While there is, formally, a difference between f and $f^{\mathcal{M}}$, when given an $\mathcal{L}$-structure, we will not distinguish between these two and just write f for $f^{\mathcal{M}}$, and similarly for the relation symbols and constant symbols. Considering $\mathbb{R}$ naturally as an $\mathcal{L}_r$-structure, the formula $\exists x_2\, x_2 x_2 = x_1$ does really mean that x_1 is a square. We will often just say that $\mathbb{R}$ is an $\mathcal{L}$-structure and use the same notation for the underlying set of real numbers $\mathbb{R}$.

Definition 1.1.6. Let $\mathcal{L}$ be a language and suppose that $\mathbb{R}$ is an $\mathcal{L}$-structure. A set $X \subset \mathbb{R}^n$ is **$\mathcal{L}$-definable** if there exists some $\mathcal{L}$-formula φ such that

$$X = \{x \in \mathbb{R}^n \mid \varphi(x)\},$$

the set of points that satisfy formula φ.

Formally, we should have introduced what it means for a formula to be true in an $\mathcal{L}$-structure, but I think this definition suffices here. We say that a function is $\mathcal{L}$-definable if its graph is $\mathcal{L}$-definable and a relation is $\mathcal{L}$-definable if it is as a set. If the language $\mathcal{L}$ is clear from the context, we might just write definable.

Example 1.1.7.

1. $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 < x_2\}$ is $\mathcal{L}_<$-definable.
2. $\{x_1 \in \mathbb{R} \mid \exists x_2\, x_2^2 = x_1\}$ is $\mathcal{L}_r$-definable. Note that it is the projection of the $\mathcal{L}_r$-definable set $\{x \in \mathbb{R}^2 \mid x_1 = x_2^2\}$ onto the first coordinate.
3. $\{x_1 \in \mathbb{R} \mid \exists x_2\, (x_1 - x_2^2) = 1\}$ is $\mathcal{L}_r$-definable. It is the same set as $\{x_1 \in \mathbb{R} \mid x_1 \geq 1\}$, which is $\mathcal{L}_{or}$-definable.
4. If X is $\mathcal{L}_{or}$-definable, then its topological closure $\bar{X}$ is $\mathcal{L}_{or}$-definable [Mar02, Lemma 1.3.3].

Definable sets are closed under elementary geometric operations. For example, they are closed under intersection, since if $X = \{x \in \mathbb{R}^n \mid \varphi(x)\}$ and $Y = \{x \in \mathbb{R}^n \mid \psi(x)\}$, then we have that $X \cap Y = \{x \in \mathbb{R}^n \mid \varphi(x) \wedge \psi(x)\}$. In example 2 above we have seen a projection. The following proposition gives a geometric description of definable sets. It is essentially [Mar02, Proposition 1.3.4], but in the case that $\mathbb{R}$ is an $\mathcal{L}$-structure, with the conventions we have made before.

Proposition 1.1.8. *Let $\mathbb{R}$ be an $\mathcal{L}$-structure. Suppose that D_n is a collection of subsets of $\mathbb{R}^n$ for all $n \geq 1$ and $\mathcal{D} = \{D_n \mid n \geq 1\}$ is the smallest collection such that:*

1. *$\mathbb{R}^n \in D_n$;*
2. *for all n-ary function symbols f of $\mathcal{L}$, the graph of f is in D_{n+1};*

3. *for all n-ary relation symbols R of $\mathcal{L}$, $R \in D_n$;*
4. *for all $i, j \leq n$, $\{(x_1, \ldots, x_n) \mid x_i = x_j\} \in D_n$;*
5. *if $X \in D_n$, then $\mathbb{R} \times X \in D_{n+1}$;*
6. *each D_n is closed under complement, union, and intersection;*
7. *if $X \in D_{n+1}$ and $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the projection map $(x_1, \ldots, x_{n+1}) \mapsto (x_1, \ldots, x_n)$, then $\pi(X) \in D_n$;*
8. *if $X \in D_{n+m}$ and $b \in \mathbb{R}^m$, then $\{a \in \mathbb{R}^n \mid (a, b) \in X\} \in D_n$.*

Then $X \subset \mathbb{R}^n$ is $\mathcal{L}$-definable if and only if $X \in D_n$.

Using this proposition, it is possible to talk about definable sets, without using the terminology of model theory. For instance, this definition is used in [vdD98], which is the main source of the results in Section 1.3. Using the above axioms as a definition might be more down to earth, but in this way it is quite hard to show that a set is definable. Using model theory, all you need to know is that one can write some first order formula for the set, which is usually easy to see. For instance, the last example of Example 1.1.7.

1.2 Important structures

This short section contains the languages and the induced structures on $\mathbb{R}$ we will consider later on. First, we make the following three assumptions on our language $\mathcal{L}$:

1. $+ \in \mathcal{F}$ and $\cdot \in \mathcal{F}$;
2. $\mathcal{R} = \{<\}$;
3. $\mathcal{C} = \mathbb{R}$.

Assumptions (1) and (2) state that we will use “at least” the language of ordered rings. The third assumption says that any real number may appear in a formula. In model theoretic terms, we call this “definable with parameters”. Since $\mathcal{R}$ and $\mathcal{C}$ are fixed, we will simply define a language by defining its function symbols. For example, if $\mathcal{L}$ is defined by $(+, \cdot)$, we mean the language $\mathcal{L}_{or}$ of ordered rings. Formally, it is not the same language as the language of the previous section. To be precise, if X is $\mathcal{L}_{or}$-definable in this setting, it means $\mathcal{L}_{or}$-definable with parameters, using the definition of Section 1.1.

Definition 1.2.1. A set $X \subset \mathbb{R}^n$ is **semi-algebraic** if and only if it is $\mathcal{L}_{or}$-definable.

Let us now define the subanalytic language $\mathcal{L}_{an}$. A **restricted analytic function** is a function defined by

$$\begin{cases} f(x) & \text{if } x \in [0,1]^n, \\ 0 & \text{elsewhere,} \end{cases}$$

where f is an analytic function on an open (in this thesis, always in the Euclidean topology) neighborhood of $[0,1]^n$. The language $\mathcal{L}_{an}$ is defined by $(+, \cdot, \mathcal{F}_{an})$, where $\mathcal{F}_{an}$ is the set of all restricted analytic functions. At this point, it might be unclear why we only allow restricted analytic functions. We will explain in the next section why we avoid entire analytic functions, see also Remark 1.2.4. We will denote the $\mathcal{L}_{an}$-structure $\mathbb{R}$ by $\mathbb{R}_{an}$.

Definition 1.2.2. A set $X \subset \mathbb{R}^n$ is **subanalytic** if and only if it is $\mathcal{L}_{an}$-definable. We may also simply say definable in $\mathbb{R}_{\mathrm{an}}$.

Finally, if one adds to the language $\mathcal{L}_{an}$ a function symbol for each **power function**, that are the unary functions defined by

$$\begin{cases} x^r & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

for some $r \in \mathbb{R}$, then we obtain a language which we will denote by $\mathcal{L}_{an}^{\mathbb{R}}$. These power functions were studied by C. Miller in [Mil94]. Note that if $r \in \mathbb{Q}$, these functions are already $\mathcal{L}_{or}$-definable, but if $r \in \mathbb{R} \setminus \mathbb{Q}$, it is not. Moreover, since they are not analytic at the origin, they are not restricted analytic functions.

Definition 1.2.3. A set $X \subset \mathbb{R}^n$ is **power-subanalytic** if and only if it is $\mathcal{L}_{an}^{\mathbb{R}}$-definable.

Finally, all of these languages can be packed together into a family of languages, which were studied by D. Miller in [Mil06]. In this article, one considers a set $\mathcal{F}$ of restricted analytic functions, whose germs at the origin are most importantly closed under compositions and Weierstrass preparation. Given such a set $\mathcal{F}$, one can consider its field of exponents K, which consists of all functions of the form $(1+x)^r$ for $r \in \mathbb{R}$ that are in $\mathcal{F}$. See [Mil06] for more details. One can then consider the language $\mathcal{L}_{\mathcal{F}}^{K}$ defined by $+$, $\cdot$, all functions in $\mathcal{F}$ and all power functions x^r for each $r \in K$, defined as above. For the smallest Weierstrass system, one just recovers the structure of semi-algebraic sets, and for the largest choice of $\mathcal{F}$, we obtain $\mathcal{L}_{an}^{\mathbb{R}}$. We shall also use **power-subanalytic** for the $\mathcal{L}_{\mathcal{F}}^{K}$-definable sets.

If one adds to the language $\mathcal{L}_{or}$ a symbol for the *entire* exponential function, one obtains the language $\mathcal{L}_{\exp}$. We will denote the corresponding $\mathcal{L}_{\exp}$-structure $\mathbb{R}$ by $\mathbb{R}_{\exp}$. One may also add the entire exponential function to the subanalytic language to obtain the $\mathcal{L}_{an,\exp}$. Instead of saying $\mathcal{L}_{\exp}$-definable, we will say definable in $\mathbb{R}_{\exp}$ and similarly definable in $\mathbb{R}_{\mathrm{an},\exp}$.

Remark 1.2.4. It is important to mention that the definitions stated here are actually results. At first sight, it is not clear that $\mathcal{L}_{or}$-definable sets are the same as semi-algebraic sets, where the latter is classically defined. The Tarski-Seidenberg Theorem on projections of semi-algebraic sets is a key ingredient of this proof. Similarly, Gabrielov's theorem of the complement is a key result for the structure of subanalytic sets, which in fact consists of the *globally* subanalytic sets. For example, the graph of the sine function is subanalytic, but not globally subanalytic.

1.3 Important results

We assume that the language $\mathcal{L}$ is at least the language of ordered rings, as explained in the beginning of Section 1.2. We assume that $\mathbb{R}$ is an $\mathcal{L}$-structure in a natural way. If we say definable in $\mathbb{R}$, it means $\mathcal{L}$-definable in this section. Via Proposition 1.1.8, the collection of definable sets is equivalent to a collection of subsets of $\mathbb{R}^n$ that are closed under simple geometric operations such as intersections, unions or projections. We will simply call this a **structure**.

Definition 1.3.1. A structure is **o-minimal** if all definable subsets of $\mathbb{R}^1$ are finite unions of intervals and points.

It should now be clear why we cannot allow entire analytic functions in our language $\mathcal{L}_{an}$. If $\mathcal{L}$ is some language such that the sine function is definable, the $\mathcal{L}$-structure $\mathbb{R}$ is not o-minimal since the set $\{x \in \mathbb{R} \mid \sin(\pi x) = 0\}$ defines the integers. All structures defined in Section 1.2 are o-minimal. In particular $\mathbb{R}_{\exp}$, by Wilkie's Theorem [Wil96], and also $\mathbb{R}_{\mathrm{an},\exp}$ [vdDM94].

As we have explained in the introduction, o-minimal structures have a rich and "tame" geometry. We will now provide some results of o-minimality, which we took from the basic reference in the field [vdD98]. We start with a theorem on unary functions in o-minimal structures. It is called the **monotonicity theorem**.

Proposition 1.3.2 ([vdD98, Chapter 3, 1.2])**.** *Let $f : (a, b) \to \mathbb{R}$ be a definable function on the interval (a, b). Then there are points $a_1 < \ldots < a_k$ in (a, b)*

such that on each subinterval (a_j, a_{j+1})*, with* $a_0 = a$*,* $a_{k+1} = b$*, the function is either constant, or strictly monotone and continuous.*

The next result is called the **finiteness lemma**.

Lemma 1.3.3 ([vdD98, Chapter 3, 1.7])**.** *Let* $A \subset \mathbb{R}^2$ *be definable and suppose that for each* $x \in \mathbb{R}$*, the fiber* $A_x = \{y \in \mathbb{R} \mid (x, y) \in A\}$ *is finite. Then there is* $N \in \mathbb{N}$ *such that* $|A_x| \leq N$ *for all* $x \in \mathbb{R}$*.*

So finiteness for each x actually implies "uniform" finiteness. This lemma and the previous proposition are essential ingredients to a key result in o-minimality, namely the cell decomposition theorem. First we have to explain what a cell is.

Definition 1.3.4. A **cell** in $\mathbb{R}^n$ is defined inductively as follows.

1. A cell in $\mathbb{R}^1$ is either a point or an open interval, which may be $(-\infty, a)$, $(a, +\infty)$ or $(-\infty, +\infty) = \mathbb{R}$.

2. Let C be a cell in $\mathbb{R}^n$, then we can construct a cell in $\mathbb{R}^{n+1}$, in any of the following two ways.

   - *Graphs of definable continuous functions.* Let $f : C \to \mathbb{R}$ be a definable continuous function. Then
   $$\{(x, y) \in C \times \mathbb{R} \mid y = f(x)\}$$
   is a cell in $\mathbb{R}^{n+1}$.

   - *A region bounded by graphs of definable continuous functions.* Let f be either constant $-\infty$ on C or a definable continuous function and g be either constant $+\infty$ on C or a definable continuous function such that for all $x \in C : f(x) < g(x)$. Then
   $$\{(x, y) \in C \times \mathbb{R} \mid f(x) < y < g(x)\}$$
   is a cell in $\mathbb{R}^{n+1}$.

All of the functions f, or f and g, used to construct the cell are called the **walls** of the cell.

There is an obvious way to define the **dimension** of a cell. For cells in $\mathbb{R}^1$, the points have dimension zero and the intervals have dimension 1. Then, when inductively constructing the cell, if one uses the first way, the dimension does not increase, if one uses the second way, the dimension increases by one. In particular, an n-dimensional cell in $\mathbb{R}^n$ is an open connected subset of $\mathbb{R}^n$.

Next, we have to define what a decomposition is. It is a special type of a partition of $\mathbb{R}^n$ using cells.

Definition 1.3.5. A **decomposition** of $\mathbb{R}^n$ is inductively defined as follows.

1. A decomposition of $\mathbb{R}^1$ is a set of the form
$$\{(-\infty, a_1), (a_1, a_2), \ldots, (a_k, +\infty), \{a_1\}, \ldots, \{a_k\}\},$$
where $a_1 < \ldots < a_k$ are points in $\mathbb{R}$.

2. A decomposition of $\mathbb{R}^{n+1}$ is a finite partition of $\mathbb{R}^{n+1}$ into cells C such that the set of projections $\pi(C)$ is a decomposition of $\mathbb{R}^n$.

A decomposition **partitions** a set $A \subset \mathbb{R}^n$ if each cell of the decomposition is either contained in A or disjoint from A. Thus, A is a finite union of cells.

The following theorem is the **cell decomposition theorem**.

Theorem 1.3.6 ([vdD98, Chapter 3, 2.11])**.**

1. *Given any definable sets $A_1, \ldots, A_k \subset \mathbb{R}^n$, there is a decomposition of $\mathbb{R}^n$ partitioning each of $A_1, \ldots, A_k$.*

2. *For each definable function $f : A \to \mathbb{R}$, $A \subset \mathbb{R}^n$, there is a decomposition of $\mathbb{R}^n$ partitioning A such that the restriction of f to each cell contained in A is continuous.*

This theorem can be used to prove many stronger results. Since this is not the subject of this thesis, we will not do so here. In Chapter 3, we will frequently have to perform operations with definable functions and sets to obtain that some desired conditions are satisfied. For example, if $f : U \to \mathbb{R}$ is a definable function, then we may "suppose" that f is continuous, or more generally r-times continuously differentiable, on (a cell contained in) U. This means that we use the above theorem, finitely many times, to obtain a partition of U into cells and then consider the restriction of f to each cell separately. Sometimes this argument is also invoked by writing "by o-minimality".

Remark 1.3.7. By work of Le Gal and Rolin [LGR08], it is known that we cannot expect that U can be decomposed into finitely many cells such that the restriction of f to these cells is C^∞.

Finally, since cells are inductively defined and because every definable set is a finite union of cells, we often use induction to prove statements in o-minimality. When partitioning a set, it is likely that some cells are not open, but in that case they have lower dimension. Thus, when using induction (on the dimension of the cells), we can ignore these lower dimensional cells. An example of this strategy is the proof of Theorem 3.3.4, the pre-parametrization theorem, in Section 3.4.

Chapter 2

Mild functions

With the exception of the results in Section 2.6, all results in this Chapter are based on the results on mild functions in [VH21a] and [VH21b].

In this chapter we study various properties of a class of differentiable functions, whose derivatives can be bounded by

$$(BA^n n!)^{1+C}$$

for some real numbers $A, B > 0$, $C \geq 0$, where n is the order of the derivative. We call such a function (A, B, C)-mild. In this introduction I will first give a summary of important results in this section, then I will provide some history and motivation. Throughout we point out the relevance of this section for other parts of the thesis.

Section 2.1 includes the notation for multivariate analysis, the precise definition of an (A, B, C)-mild function and contains examples that will be used later in this chapter and the thesis. We study properties of these functions in Section 2.2. In particular we derive formulas for the parameters A, B and C for the addition, multiplication and composition of (A, B, C)-mild functions. In Section 2.3, we prove that a certain function, which is an example given in Section 2.1, is (A, B, C)-mild for some specific $A, B > 0$ and $C \geq 0$. This function will be used in Section 2.5.

In Section 2.4, we introduce weakly (A, B, C)-mild functions. As the reader might guess, these functions are not (A, B, C)-mild. The difference is that in the case of weakly (A, B, C)-mild functions, there is an additional factor in the above upper bound that depends on the argument of the function. We state similar results for the addition and multiplication of weakly (A, B, C)-mild functions, and show some result on the composition of such a function with an (A, B, C)-mild function. In general it is not true that the composition of weakly (A, B, C)-mild functions is weakly (A, B, C)-mild (for some possibly different A, B and C). In this section we also provide an important class of examples of weakly (A, B, C)-mild functions, namely monomials.

The key results of this chapter, at least for the thesis, are in Section 2.5. There we study how the parameters A, B and C behave under three types of substitutions: linear substitutions, power substitutions and exponential substitutions. Roughly speaking, a substitution is a composition of a (weakly) (A, B, C)-mild function with some map, such that the composition has sharper upper bounds on the derivatives. Using the power and exponential substitution, one can construct many new examples of (A, B, C)-mild functions. In fact, this is how (A, B, C)-mild functions will arise later in the thesis. Finally, Section 2.6 contains some more examples and related notions.

The definition of a mild function was first given by Pila in [Pil06a]. The main result of this article is a result in number theory, diophantine geometry to be more precise, and we will discuss this in Chapter 4. The idea is that if some set X is the image of an (A, B, C)-mild map, one can derive this number theoretic result. In [Pil06a], there was some specific class of sets X, and it was conjectured that they satisfy the condition that they are the image of some (A, B, C)-mild function. If some set X is the finite union of images of (A, B, C)-mild maps, we say it admits mild parametrization. In [Pil10] (see also [Pil17]), it is shown that some specific set X admits a mild parametrization by explicitly constructing it and similar number theoretic results are deduced. The theory of parametrizations is the content of Chapter 3, which contains the main results of this thesis, so we will not further discuss it now. It suffices to say here that admitting a mild parametrization is a strong condition and therefore it is required to relax it.

For this reason, Cluckers, Pila and Wilkie have introduced in [CPW20] the definition of an (A, B, C)-mild function up to order r, where $r \in \mathbb{N} \cup \{+\infty\}$. If $r = +\infty$, this coincides with the definition above, but if $r \in \mathbb{N}$, then the above upper bound on the derivatives of the function should only hold as long as $n \leq r$. Now any function whose derivatives up to order r are all bounded, is (A, B, C)-mild up to order r, so this condition seems rather weak. However, if one applies the techniques of [Pil06a] to a set X, the result is significantly more interesting if X is the image of an (A, B, C)-mild function up to order r, rather than a function whose derivatives up to order r are just bounded, see Remark 2.5.2.

This section contains all results on (A, B, C)-mild functions (up to order r) of my articles [VH21b] and [VH21a]. One could say that it is just a bunch of formulas on the parameters A, B and C for various operations involving (A, B, C)-mild functions, but eventually it will be these formulas that make it possible to deduce the main results of my thesis on parametrizations. The proofs are, in my opinion, not hard, but are quite technical and might discourage the reader. The formulas for addition and multiplication are not difficult to verify and can be considered as a warm up. The harder proofs eventually reduce to Lemma 2.2.6. This lemma seems like a complicated combinatorial result, but its proof is actually easily verified.

I believe it is my work to have explicitly computed these formulas, that were not known before for (A, B, C)-mild functions. Also, I think it is good to have them grouped together here in this thesis for the first time, as far as I am aware. However, that does not mean that they are really new. For addition and multiplication, I would say anyone could come up with this, of course. The result on compositions can be attributed to Gevrey. In his work [Gev18] from 1918, he proves that a certain class of functions, which he calls "functions of class α" for some $\alpha \geq 0$, are closed under composition if $\alpha \geq 1$. His proof actually contains explicit formulas, basically Lemma 2.2.6, and it turns out that if $\alpha \geq 1$, the function is $(A, B, \alpha - 1)$-mild. Conversely, an (A, B, C)-mild function is of class $C + 1$. It only requires a small effort to translate his formula to (A, B, C)-mild functions. More on this is in the end of Section 2.6.

My contribution to this field would be that I have further implemented the formulas in the techniques of [CPW20], which was not always straightforward. This is the main content of [VH21b]. In Section 2.3, I show that a certain function is (A, B, C)-mild. This function also appears in [Pil10], but my result is sharper, due to a trick that makes it possible to use Lemma 2.2.6. This is the main result of [VH21a] and I would say that it is new, especially the functions that arise as an application of Proposition 2.5.10.

Finally, for this thesis I have tried to provide more details, both in the proofs and the statements, than I did in my articles. I hope this significantly increases the readability of this section. I have also included more examples, some of them are new, and so are some of the results in Section 2.6.

2.1 Notation and definitions

We will work with functions $f : U \to \mathbb{R}$, where U is open in $\mathbb{R}^m$. For $r \in \mathbb{N} \cup \{+\infty\}$ we will write $f \in \mathcal{C}^r(U)$ to say that f is r times continuously differentiable on U. For $\alpha \in \mathbb{N}^m$, we denote $|\alpha| = \alpha_1 + \ldots + \alpha_m$ and if $|\alpha| \leq r$, we write $f^{(\alpha)}$ for

$$\left(\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}}\right) f,$$

the derivative of f with respect to α. For a map $f : U \to \mathbb{R}^n$, we define $f^{(\alpha)} = \left(f_1^{(\alpha)}, \ldots, f_n^{(\alpha)}\right)$. For $\mu \in \mathbb{R}^m$, we define $x^\mu = x_1^{\mu_1} \cdots x_m^{\mu_m}$. Furthermore

$$\alpha! = \alpha_1! \cdots \alpha_m!$$

and by definition we put $0^0 = 1$.

Definition 2.1.1. Let $A, B > 0$ and $C \geq 0$ be real numbers and $r \in \mathbb{N} \cup \{+\infty\}$. A function $f : U \to \mathbb{R}$ is **(A, B, C)-mild up to order r** if $f \in \mathcal{C}^r(U)$ and if for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in U$:

$$\left|f^{(\alpha)}(x)\right| \leq \left(BA^{|\alpha|}|\alpha|!\right)^{1+C}.$$

A map $f : U \to \mathbb{R}^n$ is (A, B, C)-mild up to order r if all of its component functions are.

Remark 2.1.2. Suppose that $f \in \mathcal{C}^\infty(U)$. It is well known [KP02, Proposition 2.2.10] that f is analytic if and only if for every $x \in U$, there is some $A > 0$ such that for all $\alpha \in \mathbb{N}^m$:

$$\left|f^{(\alpha)}(x)\right| \leq A^{|\alpha|+1} |\alpha|!$$

Therefore, one could say that an $(A, B, 0)$-mild function up to order $+\infty$ is a "uniform" analytic function, because there is a constant A that works for every $x \in U$ in that case.

Remark 2.1.3. The upper bounds on the derivatives in this definition are not completely the same as in my articles [VH21b] and [VH21a]. The difference is that here, the constants A and B are also raised to the power $1 + C$. I decided to use this definition here since it gives the least messy result in Proposition 2.2.5 if $C > 0$. The downside is that if f is (A, B, C)-mild up to order r, it is not automatically (A, B, C')-mild up to order r for some $C' > C$ since $A < A^{1+C}$ is false if $A < 1$. For this reason, when proving results on manipulations with (A, B, C)-mild functions, we will assume the mildness parameters C are equal. If not, one can still derive a similar result, carefully checking how to adjust A and B, or one assumes that $A, B \geq 1$ for simplicity.

Example 2.1.4. All of the following examples are used throughout this thesis.

1. If f is analytic on an open neighborhood of the topological closure $\bar{U}$ of a bounded set U, then f is $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$. This follows by compactness and [KP02, Proposition 2.2.10].

2. In particular, if $a > 0$, the function $\sqrt{x} : (a, b) \to (\sqrt{a}, \sqrt{b})$ is $(1/a, \sqrt{b}, 0)$-mild up to order $+\infty$. If $a = 0$, the function is $(1, \sqrt{b}, 0)$-mild up to order 0 (bounded). We will revisit this example in Example 2.4.3.

3. The map $(0, 1)^m \to (0, 1)^m$ given by $(x_1, \ldots, x_m) \mapsto (x_1^k, \ldots, x_m^k)$, with $k \in \mathbb{N}$, is $(k, 1, 0)$-mild up to order $+\infty$.

4. Let $\kappa \geq 1$. In Section 2.3 we will show that the map $(0, 1)^m \to (0, 1)^m$ defined coordinatewise by $x_i \mapsto \exp(1 - 1/x_i^\kappa)$ is $(6\kappa, e, 1/\kappa)$-mild up to order $+\infty$.

2.2 Properties of mild functions

In this section we establish formulas for the mildness parameters of the addition, multiplication and composition of (A, B, C)-mild functions up to order r. Let us start with the easiest one, addition.

Proposition 2.2.1. *Suppose that for $i = 1, \ldots, \ell$ the function $f_i : U \to \mathbb{R}$ is (A_i, B_i, C)-mild up to order r. Then their sum is $(A, \ell B, C)$-mild up to order r with $A = \max(A_i)$ and $B = \max(B_i)$.*

Proof. Take $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in U$. Using linearity of the derivative and the triangle inequality, we find that

$$\left|(f_1 + \ldots + f_\ell)^{(\alpha)}(x)\right| \leq \left|f_1^{(\alpha)}(x)\right| + \ldots + \left|f_\ell^{(\alpha)}(x)\right| \leq \left(\ell B A^{|\alpha|} |\alpha|!\right)^{1+C}.$$

□

Next, we turn to the product of (A, B, C)-mild functions up to order r. Its proof relies on a combinatorial result in [CPW20].

Proposition 2.2.2 ([VH21b, Proposition 2.8])**.** *Suppose that for $i = 1, \ldots, \ell$ the function $f_i : U \to \mathbb{R}$ is (A_i, B_i, C)-mild up to order r. Then their product is $(\ell A, B^\ell, C)$-mild up to order r with $A = \max(A_i)$ and $B = \max(B_i)$.*

Proof. Take $\alpha \in \mathbb{N}^m$ with $|\alpha| \le r$ and $x \in U$. By the product rule and using the triangle inequality, we find that

$$\left|(f_1 \cdots f_\ell)^{(\alpha)}(x)\right| \le \sum C(\alpha_1, \ldots, \alpha_\ell) \prod_{i=1}^{\ell} \left|f_i^{(\alpha_i)}(x)\right|,$$

where the sum runs over all $\alpha_1, \ldots, \alpha_\ell \in \mathbb{N}^m$ such that $\alpha_1 + \ldots + \alpha_\ell = \alpha$ and $C(\alpha_1, \ldots, \alpha_\ell)$ is some combinatorial constant (that does not depend on the functions). It is shown in [CPW20, Section 3.2] that $\sum C(\alpha_1, \ldots, \alpha_\ell) \le \ell^{|\alpha|} \le \ell^{|\alpha|(1+C)}$. Finally, by the mildness of the functions $f_1, \ldots, f_\ell$, we have that

$$\prod_{i=1}^{\ell} \left|f_i^{(\alpha_i)}(x)\right| \le \prod_{i=1}^{\ell} \left(BA^{|\alpha_i|}|\alpha_i|!\right)^{1+C} \le \left(B^\ell A^{|\alpha|}\,|\alpha|!\right)^{1+C}.$$

Together with the bound on the constants, this yields the proof. ☐

Remark 2.2.3. Suppose for simplicity that $f_1, \ldots, f_\ell$ are unary functions. In that case the constants in the proof are the same as the constants in the expansion of

$$(x_1 + \ldots + x_\ell)^{|\alpha|}.$$

Evaluating at $x_1 = \ldots = x_\ell = 1$ gives that the sum of all these constants is equal to $\ell^{|\alpha|}$.

Finally we will show a result on the composition of mild functions. It is part of the article [VH21b], although our presentation here will be slightly different. More precisely, we will not translate (A, B, C)-mild functions up to order r to Gevrey functions. One then obtains a proof that is equivalent to the proof of the fact that the composition of real analytic functions is real analytic (see [KP02, Proposition 1.4.2]). Indeed, remember that, if $C = 0$, we can think of (A, B, C)-mild functions up to order $+\infty$ as uniform analytic functions (see Remark 2.1.2). If $C > 0$ we will use that the map $x \mapsto x^{1+C}$ is convex to reduce to the case $C = 0$.

We will require a formula that gives us an expression for an arbitrary derivative of the composition of two functions. This is known as the Faà di Bruno formula, which is formula FdB in the proposition below.

Proposition 2.2.4 ([CS96])**.** *Suppose that $f : U \to \mathbb{R}$ is C^r on U and $g : V \subset \mathbb{R}^p \to U$ is C^r on V. Suppose that $\alpha \in \mathbb{N}^p$ with $|\alpha| \le r$ and $x \in V$, then we have that*

$$(f \circ g)^{(\alpha)}(x) = \sum_{1 \le |\lambda| \le |\alpha|} f^{(\lambda)}(g(x)) \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(g^{(l_j)}(x))^{k_j}}{k_j!(l_j!)^{|k_j|}}, \tag{FdB}$$

where $p_s(\alpha, \lambda)$ is the set of all $k_1, \ldots, k_s \in \mathbb{N}^m$ with $|k_j| > 0$ and $l_1, \ldots, l_s \in \mathbb{N}^p$ with $0 \prec l_1 \prec \ldots \prec l_s$ such that

$$\sum_{j=1}^{s} k_j = \lambda$$

and

$$\sum_{j=1}^{s} |k_j|\, l_j = \alpha.$$

Finally, $l_j \prec l_{j+1}$ if $|l_j| < |l_{j+1}|$ or, if $|l_j| = |l_{j+1}|$, then l_j comes lexicographically before l_{j+1}.

We now show the result on compositions of (A, B, C)-mild functions up to order r.

Proposition 2.2.5 ([VH21b, Corollary 2.5.1])**.** *Suppose that $f : U \to \mathbb{R}$ is (A_f, B_f, C)-mild up to order r and $g : V \subset \mathbb{R}^p \to U$ is (A_g, B_g, C)-mild up to order r. Then their composition is (A, B, C)-mild up to order r with*

$$A = A_g(mB_gA_f + 1),$$

$$B = \frac{mA_fB_fB_g}{mA_fB_g + 1} < B_f.$$

Proof. Let $\alpha \in \mathbb{N}^p$ with $|\alpha| \leq r$ and $x \in U$. By the Faà di Bruno formula FdB, the triangle inequality and the bounds on the derivatives on f and g (by their mildness) we have that

$$\left|(f \circ g)^{(\alpha)}(x)\right| \leq \sum_{1 \leq |\lambda| \leq |\alpha|} \left(B_f A_f^{|\lambda|}\, |\lambda|!\right)^{1+C} \left(\sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(B_g A_g^{|l_j|}\, |l_j|!)^{|k_j|(1+C)}}{k_j!(l_j!)^{|k_j|}} \right). \tag{MLD}$$

Note that

$$\frac{\alpha!}{\prod_{j=1}^{s} k_j!(l_j!)^{|k_j|}}$$

is a positive integer (since it counts how many times some derivative appears). Using that for all $x, y > 0$, $x^{1+C} + y^{1+C} \leq (x+y)^{1+C}$,we have that

$$\left|(f \circ g)^{(\alpha)}(x)\right| \leq \left(\sum_{1 \leq |\lambda| \leq |\alpha|} B_f A_f^{|\lambda|}\, |\lambda|! \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(B_g A_g^{|l_j|}\, |l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}} \right)^{1+C}.$$

Then we conclude by using Lemma 2.2.6 below. $\square$

Lemma 2.2.6. *Let a, A, b and B be strictly positive real numbers, then for all $\alpha \in \mathbb{N}^p$ we have that*

$$\sum_{1 \leq |\lambda| \leq |\alpha|} ba^{|\lambda|} \, |\lambda|! \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(BA^{|l_j|} \, |l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}}$$

$$= \frac{mabB}{maB+1} (A(maB+1))^{|\alpha|} \, |\alpha|!,$$

where the left hand side follows the notation of Proposition 2.2.4.

Proof. Consider the functions f and g given by:

$$f(x_1, \ldots, x_m) = \frac{b}{1 + maB - a(x_1 + \ldots + x_m)},$$

$$g_i(x_1, \ldots, x_p) = \frac{B}{1 - A(x_1 + \ldots + x_p)} \quad (i = 1, \ldots, m).$$

Note that $g(0, \ldots, 0) = (B, \ldots, B)$. One verifies that for all $\lambda \in \mathbb{N}^m$:

$$f^{(\lambda)}(g(0, \ldots, 0)) = ba^{|\lambda|} \, |\lambda|!$$

and that for all $\lambda \in \mathbb{N}^p$:

$$g^{(\lambda)}(0, \ldots, 0) = (BA^{|\lambda|} \, |\lambda|!, \ldots, BA^{|\lambda|} \, |\lambda|!).$$

Finally, one computes directly, not using the Faà di Bruno formula FdB, that for all $\lambda \in \mathbb{N}^p$:

$$(f \circ g)^{(\lambda)}(0, \ldots, 0) = \frac{mabB}{maB+1} (A(maB+1))^{|\lambda|} \, |\lambda|!$$

The result now follows by plugging this data into the Faà di Bruno formula FdB. $\square$

One can summarize this section as follows. For a fixed $C \geq 0$, the set consisting of all functions f such that f is (A, B, C)-mild up to order r for some $A, B > 0$, is a ring that is closed under composition.

2.3 A particular mild function

The goal of this section is to prove the claim made in Example 2.1.4 on the mildness of the map $(0,1)^m \to (0,1)^m$ given coordinatewise by

$$x_i \mapsto e^{1-1/x_i^\kappa}$$

for some $\kappa \geq 1$. Note that $\exp(1-1/x_i^\kappa)$ is order preserving and a bijection on $(0,1)$. Since all coordinate functions are the same function $(0,1) \to (0,1)$, we may assume that $m = 1$. Below, we will more generally assume that $\kappa > 0$. The cases $\kappa < 1$ and $\kappa \geq 1$ have the same proof method, but a slightly different result. In the proof of Proposition 2.5.10 in Section 2.5, which generalizes this result, we will reduce to the proof here.

The rest of this section follows the proof of [VH21a, Proposition 3.4]. We consider the function $\exp(1-1/x^\kappa)$ as the composition of the functions f and g given by $f(x) = \exp(x)$ and $g(x) = 1 - 1/x^\kappa$. For all $\nu \in \mathbb{N} \setminus \{0\}$ we have that

$$\left| g^{(\nu)}(x) \right| \leq x^{-(\kappa+\nu)} \kappa(\kappa+1)\cdots(\kappa+\nu-1).$$

To simplify notation, set $c(\nu,\kappa) = \kappa(\kappa+1)\cdots(\kappa+\nu-1)$. Using the Faà di Bruno formula FdB and the triangle inequality, we obtain for all $\alpha \in \mathbb{N} \setminus \{0\}$ that

$$\left| (f \circ g)^{(\alpha)}(x) \right| \leq \sum_{1 \leq \lambda \leq \alpha} e^{1-1/x^\kappa} \sum_{s=1}^{\alpha} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(x^{-(\kappa+l_j)} c(l_j,\kappa))^{k_j}}{k_j!(l_j!)^{k_j}}.$$

By the relations between α, λ, k_j and l_j, the right hand side of this inequality is equal to

$$\sum_{1 \leq \lambda \leq \alpha} x^{-(\lambda\kappa+\alpha)} e^{1-1/x^\kappa} \sum_{s=1}^{\alpha} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{c(l_j,\kappa)^{k_j}}{k_j!(l_j!)^{k_j}},$$

which we will rewrite as

$$e x^{-\alpha} e^{-1/(2x^\kappa)} \sum_{1 \leq \lambda \leq \alpha} x^{-\lambda\kappa} e^{-1/(2x^\kappa)} \sum_{s=1}^{\alpha} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{c(l_j,\kappa)^{k_j}}{k_j!(l_j!)^{k_j}}. \tag{2.1}$$

Next, it is routine to check that

$$\max_{x \in (0,1)} x^{-\alpha} e^{-1/(2x^\kappa)} = \left(\frac{2\alpha}{e\kappa} \right)^{\alpha/\kappa} \leq \left(\frac{2}{\kappa} \right)^{\alpha/\kappa} (\alpha!)^{1/\kappa}$$

and

$$\max_{x\in(0,1)} x^{-\lambda\kappa}e^{-1/(2x^\kappa)} = \left(\frac{2\lambda\kappa}{e\kappa}\right)^{(\lambda\kappa)/\kappa} \leq 2^\lambda \lambda!,$$

where the last inequality follows from Stirling's formula. Finally, we need the following estimate on the constants $c(l_j,\kappa)$, which is not so hard to find.

$$c_{l_j,\kappa} \leq \begin{cases} \kappa^{l_j} l_j! & \text{if } \kappa \geq 1, \\ l_j! & \text{if } \kappa < 1. \end{cases}$$

Let us suppose $\kappa \geq 1$. Plugging in these bounds into 2.1 gives us that

$$\left|(f\circ g)^{(\alpha)}(x)\right| \leq e\left(\frac{2}{\kappa}\right)^{\alpha/\kappa}(\alpha!)^{1/\kappa}\sum_{1\leq\lambda\leq\alpha} 2^\lambda\lambda! \sum_{s=1}^{\alpha}\sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(\kappa^{l_j} l_j!)^{k_j}}{k_j!(l_j!)^{k_j}}.$$

By Lemma 2.2.6, we can compute the sum and find the following upper bound:

$$\left|(f\circ g)^{(\alpha)}(x)\right| \leq e\left(\frac{2}{\kappa}\right)^{\alpha/\kappa}(\alpha!)^{1/\kappa}\frac{2}{3}(3\kappa)^\alpha\alpha!.$$

Since we have assumed that $\kappa \geq 1$, one has that $(2/\kappa)^{1/\kappa} \leq 2$. Therefore, we can conclude that

$$\left|(f\circ g)^{(\alpha)}(x)\right| \leq (e(6\kappa)^\alpha\alpha!)^{1+1/\kappa}$$

and thus, it follows that $f\circ g$ is $(6\kappa, e, 1/\kappa)$-mild up to order $+\infty$. Now, if $\kappa < 1$, one has to use the other upper bound on the constants $c(l_j,\kappa)$ and there is no suitable upper bound for $(2/\kappa)^{1/\kappa}$, which becomes very large if κ is small. In this case, one finds the upper bound

$$\left|(f\circ g)^{(\alpha)}(x)\right| \leq \left(e3^\alpha\left(\frac{2}{\kappa}\right)^{\alpha/\kappa}\alpha!\right)^{1+1/\kappa}$$

and concludes that the function is $(3(2/\kappa)^{1/\kappa}, e, 1/\kappa)$-mild up to order $+\infty$.

2.4 Weakly mild functions

In this section U is an open subset of $(0,1)^m$ for the following two reasons. Firstly, in the definition of a weakly mild function, we want to use negative powers of variables, therefore we have to make sure they cannot be zero. Secondly, in Section 2.5 we will modify their domain with functions that are bijections $(0,1)^m \to (0,1)^m$, for example the map we studied in the previous section.

Definition 2.4.1. Let $A, B > 0$ and $C \geq 0$ be real numbers and $r \in \mathbb{N} \cup \{+\infty\}$. A function $f : U \to \mathbb{R}$ is **weakly (A, B, C)-mild up to order r** if $f \in \mathcal{C}^r(U)$ and if for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in U$:

$$\left| f^{(\alpha)}(x) \right| \leq x^{-\alpha} \left(B A^{|\alpha|} |\alpha|! \right)^{1+C}.$$

A map $f : U \to \mathbb{R}^n$ is weakly (A, B, C)-mild up to order r if all of its component functions are.

An example of a class of functions that are weakly (A, B, C)-mild up to order $+\infty$ is given by the following lemma.

Lemma 2.4.2 ([VH21b, Lemma 3.2])**.** *Let $b : U \to \mathbb{R}$ be given by $b(x) = x^\mu$ for some $\mu \in \mathbb{R}^m$ and suppose that b is bounded. Then b is weakly $(A, B, 0)$-mild up to order $+\infty$ with $A = \max(|\mu_1|, \ldots, |\mu_m|, 1)$ and $B = \sup_U b(x)$.*

Proof. Clearly we have that

$$b^{(\alpha)}(x) = c(\alpha, \mu) x^{-\alpha} b(x)$$

for some constant $c(\alpha, \mu)$ that depends on α and μ, because of the form of the function b. Let $B = \sup_U b(x)$, then we have that

$$\left| b^{(\alpha)}(x) \right| \leq x^{-\alpha} B \left| c(\alpha, \mu) \right|.$$

It remains to bound $|c(\alpha, \mu)|$. Let $A = \max(|\mu_1|, \ldots, |\mu_m|, 1)$, then one computes that

$$|c(\alpha, \mu)| \leq A(A+1) \cdots (A + |\alpha| - 1) \leq A^{|\alpha|} |\alpha|!$$

□

Example 2.4.3. This lemma tells us that the function $\sqrt{x} : (0, 1) \to (0, 1)$ is weakly $(1, 1, 0)$-mild up to order $+\infty$, which is more accurate than the poor conclusion we had to make on its mildness in Example 2.1.4. Another example, that does not fit into this class of functions, is the function $x \ln(x)$ on $(0, 1)$ because we have that $|\ln(x)| < 1/x$ on $(0, 1)$. It is also weakly $(1, 1, 0)$-mild up to order $+\infty$.

For completeness, let us give the following result. The proof is completely analogous to the results on addition and multiplication of (A, B, C)-mild functions up to order r in propositions 2.2.1 and 2.2.2.

Proposition 2.4.4. *Suppose that for $i = 1, \ldots, \ell$ the function $f_i : U \to \mathbb{R}$ is weakly (A_i, B_i, C)-mild up to order r. Let $A = \max(A_i)$ and $B = \max(B_i)$.*

1. *Their sum is weakly $(A, \ell B, C)$-mild up to order r.*
2. *Their product is weakly $(\ell A, B^\ell, C)$-mild up to order r.*

Thus we can conclude that, for a fixed $C \geq 0$, the set consisting of all functions f such that f is weakly (A, B, C)-mild up to order r for some $A, B > 0$, is a ring. Obviously, it contains the (A, B, C)-mild functions up to order r as a subring. However, it is *not* closed under compositions. Indeed, the composition of weakly (A, B, C)-mild functions up to order r is not (necessarily) weakly (A, B, C)-mild up to order r (for some possibly different A and B). However, one type of compositions can be made. The notation of the following proposition follows the notation of propositions 2.2.4 and 2.2.5.

Proposition 2.4.5 ([VH21b, Proposition 2.7])**.** *Suppose that $f : U \to \mathbb{R}$ is (A_f, B_f, C)-mild up to order r and that $g : V \to U$ is weakly (A_g, B_g, C)-mild up to order r. Then their composition is weakly (A, B, C)-mild up to order r with*

$$A = A_g(mB_gA_f + 1),$$

$$B = \frac{mB_fB_g}{mB_g + A_f^{-1}} < B_f.$$

Proof. The only modification that has to be made to the proof of Proposition 2.2.5 is that inequality MLD now additionally has negative powers of x coming from upper bounds on the derivatives of g. More precisely, in this case we have that

$$\begin{aligned}
&\left|(f \circ g)^{(\alpha)}(x)\right| \\
&\leq \sum_{1 \leq |\lambda| \leq r} \left(B_f A_f^{|\lambda|} \, |\lambda|!\right)^{1+C} \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{x^{-|k_j| l_j} (B_g A_g^{|l_j|} \, |l_j|!)^{|k_j|(1+C)}}{k_j!(l_j!)^{|k_j|}} \\
&\leq x^{-\alpha} \sum_{1 \leq |\lambda| \leq r} \left(B_f A_f^{|\lambda|} \, |\lambda|!\right)^{1+C} \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(B_g A_g^{|l_j|} \, |l_j|!)^{|k_j|(1+C)}}{k_j!(l_j!)^{|k_j|}},
\end{aligned}$$

since $\sum_{j=1}^{s} |k_j| \, l_j = \alpha$. Thus, up to the factor $x^{-\alpha}$, we obtain the same upper bound as in the right hand side of inequality MLD in Proposition 2.2.5. The proof now continues by bounding the sum in the same way. $\square$

2.5 Substitutions

The subject of this section is the following. Given some function $f : U \to \mathbb{R}$ which is (weakly) (A, B, C)-mild function up to order r, we want to compose it with some bijection $P : V \to U$ such that the upper bounds on the derivatives of $f \circ P$ are even better. So we replace the variables of f by some functions in these variables and this should improve f from this point of view. For this reason, this technique is often referred to as a substitution.

Linear substitutions

The first kind of substitution, a linear substitution, improves the bounds on the derivatives of (A, B, C)-mild functions up to order r.

Lemma 2.5.1 ([VH21b, Lemma 2.1])**.** *Suppose $f : U \to \mathbb{R}$ is (A, B, C)-mild up to order $r \in \mathbb{N}$. Let $P_\ell : \mathbb{R}^m \to \mathbb{R}^m$ be the map given coordinatewise by*

$$x_i \mapsto \frac{1}{(Ar)^{1+C}} x_i.$$

Then for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_\ell^{-1}(U)$: $\left|(f \circ P_\ell)^{(\alpha)}(x)\right| \leq B^{1+C}$.

Proof. Let $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_\ell^{-1}(U)$. By the chain rule, we have that

$$(f \circ P_\ell)^{(\alpha)}(x) = \left(\frac{1}{(Ar)^{1+C}}\right)^{|\alpha|} f^{(\alpha)}(P_\ell(x)).$$

Since $P_\ell(x) \in U$, f is (A, B, C)-mild up to order r on U and $|\alpha| \leq r$, we have that

$$\left|f^{(\alpha)}(P_\ell(x))\right| \leq (BA^{|\alpha|}\,|\alpha|!)^{1+C}.$$

Since $|\alpha| \leq r$, we have that $|\alpha|! \leq r^{|\alpha|}$, thus:

$$\left|(f \circ P_\ell)^{(\alpha)}(x)\right| \leq \left(\frac{1}{(Ar)^{1+C}}\right)^{|\alpha|} (BA^{|\alpha|}\,|\alpha|!)^{1+C} \leq B^{1+C}.$$

□

Remark 2.5.2.

1. Of course, if a function is (A, B, C)-mild up to order $+\infty$, it is also (A, B, C)-mild up to order r for all $r \in \mathbb{N}$. In that way, one can apply the lemma to that function.

2. Clearly, if $f : U \to \mathbb{R}$ is a function such that all derivatives up to order $r < +\infty$ are bounded by some formula, one can find some constant $L > 0$, define P_ℓ coordinatewise by $x_i \mapsto Lx_i$ and also find such a result. The point here is that we have an explicit formula for L in this lemma that is related to the mildness parameters A, B and C in a convenient way. This lemma will allow us to count the number of charts of a parametrization in Chapter 3, see Lemma 3.1.7.

Power substitutions

For weakly (A, B, C)-mild functions up to order r satisfying a condition on their first order derivatives, power substitutions transform them into an (A', B', C)-mild function up to some finite order for some $A', B' > 0$ for which we will derive a formula, using the results of Section 2.2. In fact, this property could be regarded as the motivation for the definition of weakly (A, B, C)-mild functions up to order r (see Lemma 3.2.3 in Chapter 3). It also gives us a way to construct more examples of (A, B, C)-mild functions up to order $r \in \mathbb{N}$. Remember that for weakly (A, B, C)-mild functions up to order r, we suppose that their domain U is an open subset of $(0, 1)^m$.

Proposition 2.5.3 ([VH21b, Proposition 2.6]). *Suppose that for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$, the function $f^{(\beta)} : U \to \mathbb{R}$ is weakly (A, B, C)-mild up to order $r \in \mathbb{N}$. Let $N \in \mathbb{N}^m$ such that for all $i \in \{1, \ldots, m\}$, $r \leq N_i$ and $P_r : (0, 1)^m \to (0, 1)^m$ be the map given coordinatewise by*

$$x_i \mapsto x_i^{N_i}.$$

Then $f \circ P_r$ is $(M(mA + 1), \max(B, B/A), C)$-mild up to order r on $P_r^{-1}(U)$, where $M = \max_i N_i$.

Proof. Of course, we start with the Faà di Bruno formula FdB and apply the triangle inequality to obtain that for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_r^{-1} : (U)$:

$$\left|(f \circ P_r)^{(\alpha)}(x)\right| \leq \sum_{1 \leq |\lambda| \leq |\alpha|} \left|f^{(\lambda)}(P_r(x))\right| \left(\sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{\left|\left(P_r^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}} \right). \tag{2.2}$$

Now fix a term of the sum in this upper bound, i.e., fix some $\lambda \in \mathbb{N}^m$, $s \in \{1, \dots, |\alpha|\}$ and $k_1, \dots, k_s, l_1, \dots, l_s \in p_s(\alpha, \lambda)$. We now focus on

$$\left|f^{(\lambda)}(P_r(x))\right| \alpha! \prod_{j=1}^{s} \frac{\left|\left(P_r^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}}. \tag{2.3}$$

Let $k \in \{1, \dots, m\}$ be such that $x_k = \min\{x_i \mid \lambda_i \neq 0\}$ and let $\lambda' \in \mathbb{N}^m$ be such that $f^{(\lambda)} = (\partial f/\partial x_k)^{(\lambda')}$. We will now use that for *all* $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$, the function $f^{(\beta)}$ is weakly (A, B, C)-mild up to order r. Since $P_r(x) \in U$, we have that

$$\left|f^{(\lambda)}(P_r(x))\right| = \left|\left(\frac{\partial}{\partial x_k} f\right)^{(\lambda')} (P_r(x))\right| \leq x^{-N\lambda'}(BA^{|\lambda'|} \, |\lambda'|!)^{1+C},$$

where $N\lambda' = (N_i \lambda'_i)_i$. Note that $x^{-N\lambda'}$ is the only factor in this upper bound that depends on k. This dependance will vanish later on. Using this upper bound, we can bound 2.3 by:

$$x^{-N\lambda'}(BA^{|\lambda'|} \, |\lambda'|!)^{1+C} \alpha! \prod_{j=1}^{s} \frac{\left|\left(P_r^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}}. \tag{2.4}$$

Recall that $M = \max_i N_i$. Let $i \in \{1, \dots, m\}$ be arbitrary. We will find an upper bound on the contribution of x_i to 2.4. Since $(P_r)_i(x) = x_i^{N_i}$, we have that

$$\left|(P_r)_i^{(l_j)}(x)\right| \leq M^{|l_j|} x_i^{N_i - l_{j,i}} \leq x_i^{N_i - l_{j,i}}(M^{|l_j|} \, |l_j|!)^{1+C}.$$

From this, using the relations that α, λ, $k_1, \dots, k_s$ and $l_1, \dots, l_s$ satisfy, we deduce that

$$\begin{aligned}
& x_i^{-N_i\lambda'_i} \prod_{j=1}^{s} \left|((P_r)_i^{(l_j)}(x))^{k_{j,i}}\right| \\
& \leq x_i^{-N_i\lambda'_i} \prod_{j=1}^{s} \left(x_i^{N_i - l_{j,i}}(M^{|l_j|} \, |l_j|!)^{1+C}\right)^{k_{j,i}} \\
& \leq x_i^{-N_i\lambda'_i} x_i^{\lambda_i N_i - \sum_{j=1}^{s} k_{j,i} l_{j,i}} \prod_{j=1}^{s} \left((M^{|l_j|} \, |l_j!|)^{1+C}\right)^{k_{j,i}}.
\end{aligned}$$

If $i \neq k$, then

$$x_i^{-N_i\lambda'_i} x_i^{\lambda_i N_i - \sum_{j=1}^{s} k_{j,i} l_{j,i}} = x_i^{N_i(\lambda_i - \lambda'_i) - \sum_{j=1}^{s} k_{j,i} l_{j,i}} = x_i^{- \sum_{j=1}^{s} k_{j,i} l_{j,i}}.$$

If $\lambda_i \neq 0$, then we have that $x_k \leq x_i$ and thus we can further bound this as follows:

$$x_i^{-\sum_{j=1}^s k_{j,i} l_{j,i}} \leq x_k^{-\sum_{j=1}^s k_{j,i} l_{j,i}} \leq x_k^{-\alpha_i},$$

where we have used that $\sum_{j=1}^s |k_j| \, l_j = \alpha$ and that $x \in P_r^{-1}(U) \subset (0,1)^m$. Now if $\lambda_i = 0$, we have that $k_{j,i} = 0$ for $j = 1, \ldots, s$, since $\sum_{j=1}^s k_j = \lambda$, and therefore we may use the same upper bound.

If $i = k$, then

$$x_i^{-N_i \lambda_i'} x_i^{\lambda_i N_i - \alpha_i} = x_k^{N_k(\lambda_k - \lambda_k') - \alpha_k} \leq x_k^{N_k - \alpha_k}.$$

We can now further bound 2.4, additionally using that $|\alpha| \leq r \leq N_k$:

$$x^{-N\lambda'}(BA^{|\lambda'|}\,|\lambda'|!)^{1+C} \alpha! \prod_{j=1}^s \frac{\left|\left(P_r^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}}$$

$$\leq x_k^{N_k - |\alpha|}(BA^{|\lambda'|}\,|\lambda'|!)^{1+C} \alpha! \prod_{j=1}^s \frac{\left(M^{|l_j|}\,|l_j|!\right)^{(1+C)|k_j|}}{k_j!(l_j!)^{|k_j|}}$$

$$\leq (BA^{|\lambda'|}\,|\lambda'|!)^{1+C} \alpha! \prod_{j=1}^s \frac{\left(M^{|l_j|}\,|l_j|!\right)^{(1+C)|k_j|}}{k_j!(l_j!)^{|k_j|}}$$

$$\leq \left(\frac{B}{A} A^{|\lambda|}\,|\lambda|!\right)^{1+C} \alpha! \prod_{j=1}^s \frac{\left(M^{|l_j|}\,|l_j|!\right)^{(1+C)|k_j|}}{k_j!(l_j!)^{|k_j|}}.$$

This upper bound no longer depends on x (and thus neither on k).

Finally, putting everything together, we can now further bound 2.2 and obtain that

$$\left|(f \circ P_r)^{(\alpha)}(x)\right| \leq \sum_{1 \leq |\lambda| \leq |\alpha|} \left(\frac{B}{A} A^{|\lambda|}\,|\lambda|!\right)^{1+C} \left(\sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^s \frac{\left(M^{|l_j|}\,|l_j|!\right)^{(1+C)|k_j|}}{k_j!(l_j!)^{|k_j|}} \right).$$

The right hand side of this inequality is exactly of the form of the right hand side in inequality MLD. Hence, the proof continues as the proof of Proposition

2.2.5 and we can conclude that $f \circ P_r$ is $(M(mA+1), \max(B, B/A), C)$-mild up to order r.

Remark that one has to take $\max(B, B/A)$ since if $A \geq 1$ the best upper bound for $|(f \circ P_r)(x)|$ remains B, of course. $\square$

Remark 2.5.4.

1. If f is weakly (A, B, C)-mild up to order $+\infty$, it is also weakly (A, B, C)-mild up to order r for all $r \in \mathbb{N}$. Thus we can apply this proposition to these functions. Also, if $k \leq r$ and P_r is defined by some m-tuple N with $k \leq N_i$, then the composition will be $(M(mA+1), \max(B, B/A), C)$-mild up to order k, where M is again the maximum over all N_i.

2. Unless stated otherwise, we will assume that P_r is the map defined as above, with $N = (r, \dots, r) \in \mathbb{N}^m$.

3. One can relax the condition on $f^{(\beta)}$ for $|\beta| = 1$ and also obtain a slightly cleaner result, namely that the composition is $(M(mA+1), B, C)$-mild up to order r, if one asks instead that

$$\left| \left(f^{(\beta)} \right)^{(\lambda)} (x) \right| \leq (BA^{|\lambda|+1}(|\lambda|+1)!)^{1+C} \frac{1}{x^\lambda}$$

for all $\beta \in \mathbb{N}^m$ with $|\beta| = 1$. In that case, we don't have to do the additional step

$$(BA^{|\lambda'|} \, |\lambda'|!)^{1+C} \leq \left(\frac{B}{A} A^{|\lambda|} \, |\lambda|! \right)^{1+C}$$

in the end of the proof. Using this trick, the result on the mildness parameters A, B and C is the same as in Proposition 2.2.5, i.e., completely agrees with the formula for composition.

Example 2.5.5. Let $U = \{(x, y) \in (0, 1)^2 \mid x < y\}$ and let $f : U \to \mathbb{R}$ be given by $f(x, y) = x^2 y^{-1}$. One verifies that

$$\sup_{\substack{|\beta| \leq 1 \\ (x,y) \in U}} \left| f^{(\beta)}(x) \right| \leq 2.$$

By Lemma 2.4.2, we can conclude that for all $\beta \in \mathbb{N}^2$ with $|\beta| \leq 1$, the function $f^{(\beta)}$ is weakly $(2, 2, 0)$-mild up to order $+\infty$. Let $r \in \mathbb{N}$. By Proposition 2.5.3

the composition $f \circ P_r$ is $(6r, 2, 0)$-mild up to order r on $P_r^{-1}(U)$. Note that $(f \circ P_r)(x, y) = x^{2r}y^{-r}$ and that $P_r^{-1}(U) = U$, so one can check directly that $f \circ P_r$ is $(2r, 1, 0)$-mild up to order r.

Clearly, in the case of monomials, it is not necessary to use the Faà di Bruno formula to compute an upper bound on the composition with P_r. By avoiding this, one can obtain a sharper and more precise result on the mildness parameters A and B, similar to the result in Lemma 2.4.2. The next proposition is a slightly simpler version of [VH21b, Proposition 3.4].

Proposition 2.5.6. *Let $b : U \to \mathbb{R}$ be given by $b(x) = x^\mu$ for some $\mu \in \mathbb{R}^m$ and suppose that there exists some $B > 0$ such that for all $x \in U : |b(x)| \leq B$ and*

$$\left| \frac{1}{\mu_i} \left(\frac{\partial b}{\partial x_i} \right) (x) \right| \leq B$$

for all $i \in \{1, \dots, m\}$ such that $\mu_i \neq 0$. Let P_r and M be defined as in Proposition 2.5.3. Then $b \circ P_r : P_r^{-1}(U) \to \mathbb{R}$ is $(AM, B, 0)$-mild up to order r, where $A = \max(|\mu_1|, \dots, |\mu_m|, 1)$.

Note that the condition on the first order derivatives of b is equivalent to saying that they are bounded.

Proof. Let $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_r^{-1}(U)$. Since $b \circ P_r$ is also a monomial, we have that

$$(b \circ P_r)^{(\alpha)}(x) = c(\alpha, N\mu)x^{-\alpha}(b \circ P_r)(x)$$

as in Lemma 2.4.2, where $N\mu = (N_i\mu_i)_i$. Following the proof strategy of Proposition 2.5.3, for some suitable $k \in \{1, \dots, m\}$, we have that

$$\left|x^{-\alpha}(b \circ P_r)(x)\right| \leq \left|x_k^{-N_k}(b \circ P_r)(x)\right| = \left| \frac{1}{\mu_k} \left(\frac{\partial}{\partial x_k} b \right) (P_r(x)) \right|.$$

From the proof of Lemma 2.4.2, we see that $|c(\alpha, N\mu)| \leq (AM)^{|\alpha|} |\alpha|!$ $\square$

The exponential substitution

The exponential substitution has a similar property as the power substitution, namely if f is weakly (A, B, C)-mild up to order $+\infty$ and satisfies the same additional condition on its first order derivatives, then the composition will be (A', B', C')-mild up to order $+\infty$ for some $A', B' > 0$ and $C' > 0$. The map we will use, is the map of Section 2.3.

Proposition 2.5.7 ([VH21a, Proposition 4.3])**.** *Suppose that for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$, the function $f^{(\beta)} : U \to \mathbb{R}$ is weakly (A, B, C)-mild up to order $r \leq +\infty$. Let $\kappa > 0$ and $P_\infty : (0,1)^m \to (0,1)^m$ be the map given coordinatewise by*

$$x_i \mapsto e^{1-1/x_i^\kappa}.$$

Then $f \circ P_\infty$ is $(\tilde{A}, \tilde{B}, 1 + C + 1/\kappa)$-mild up to order r on $P_\infty^{-1}(U)$, where

$$\tilde{A} = \begin{cases} 8\kappa(mA+1) & \text{if } \kappa \geq 1, \\ 2(1+1/\kappa)^{1+1/\kappa}(mA+1) & \text{if } \kappa < 1, \end{cases}$$

and $\tilde{B} = \max(eB/A, B, 1)$.

Proof. The proof follows the same strategy as the proof of Proposition 2.5.3. Let us suppose $\kappa \geq 1$ for the bounds on the derivatives of P_∞, the other case is completely analogous.

Let $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_\infty^{-1}(U)$. To bound $\left|(f \circ P_\infty)^{(\alpha)}(x)\right|$, we use the Faà di Bruno formula FdB and apply the triangle inequality. Next, we consider a fixed term of this sum, thus we fix $\lambda \in \mathbb{N}^m$, $s \in \{1, \ldots, |\alpha|\}$ and $k_1, \ldots, k_s, l_1, \ldots, l_s \in p_s(\alpha, \lambda)$ and will found an upper bound for

$$\left|f^{(\lambda)}(P_\infty(x))\right| \alpha! \prod_{j=1}^{s} \frac{\left|\left(P_\infty^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}}.$$

Let $k \in \{1, \ldots, m\}$ be such that $x_k = \min\{x_i \mid \lambda_i \neq 0\}$ and let $\lambda' \in \mathbb{N}^m$ be such that $f^{(\lambda)} = ((\partial/\partial x_k)(f))^{(\lambda')}$. Then we have that

$$\left|f^{(\lambda)}(P_\infty(x))\right| = \left|\left(\frac{\partial}{\partial x_k} f\right)^{(\lambda')} (P_\infty(x))\right| \leq (P_\infty(x))^{-\lambda'} \left(\frac{B}{A} A^{|\lambda|} |\lambda|!\right)^{1+C}.$$

For the derivatives of P_∞, we use the upper bound on the coordinate functions from Expression 2.1 in Section 2.3. In this way, we obtain that:

$$\left|(P_\infty)_i^{(l_j)}(x)\right| \leq x_i^{-(\kappa+1)l_{j,i}} e^{1-1/x_i^\kappa} \sum_{1 \leq \tilde{\lambda} \leq l_{j,i}} \sum_{\tilde{s}=1}^{l_{j,i}} \sum_{p_{\tilde{s}}(l_{j,i},\tilde{\lambda})} l_{j,i}! \prod_{\tilde{j}=1}^{\tilde{s}} \frac{c(\tilde{l}_{\tilde{j}}, \kappa)^{\tilde{k}_{\tilde{j}}}}{\tilde{k}_{\tilde{j}}!(\tilde{l}_{\tilde{j}}!)^{\tilde{k}_{\tilde{j}}}} \quad \text{(exp)}$$

$$\leq x_i^{-(\kappa+1)l_{j,i}} (P_\infty)_i(x) (2\kappa)^{|l_j|} |l_j|!,$$

where we have used that $\kappa \geq 1$ to bound the constants $c(\tilde{l}_{\tilde{j}}, \kappa)$. Putting these two bounds together, using the relations between λ, $k_1, \ldots, k_s$ and $l_1, \ldots, l_s$, and by our choice of k, we find that

$$\begin{aligned}
&\left|f^{(\lambda)}(P_\infty(x))\right| \alpha! \prod_{j=1}^{s} \frac{\left|\left(P_\infty^{(l_j)}(x)\right)^{k_j}\right|}{k_j!(l_j!)^{|k_j|}} \\
&\leq x_k^{-(\kappa+1)|\alpha|}(P_\infty)_k(x) \left(\frac{B}{A} A^{|\lambda|} |\lambda|!\right)^{1+C} \alpha! \prod_{j=1}^{s} \frac{((2\kappa)^{|l_j|} |l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}} \\
&\leq x_k^{-(\kappa+1)|\alpha|}(P_\infty)_k(x) \left(\frac{B}{A} A^{|\lambda|} |\lambda|!\right)^{1+C} \alpha! \prod_{j=1}^{s} \frac{((2\kappa)^{|l_j|} |l_j|!)^{|k_j|(1+C)}}{k_j!(l_j!)^{|k_j|}}.
\end{aligned}$$

The factor $x_k^{-(\kappa+1)|\alpha|}(P_\infty)_k(x)$ no longer depends on λ and thus can be factored out. More precisely, we have that

$$\begin{aligned}
\left|(f \circ P_\infty)^{(\alpha)}(x)\right| \leq x_k^{-(\kappa+1)|\alpha|}(P_\infty)_k(x) \Bigg(\\
\sum_{1 \leq |\lambda| \leq |\alpha|} \left(\frac{B}{A} A^{|\lambda|} |\lambda|!\right)^{1+C} \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{((2\kappa)^{|l_j|} |l_j|!)^{|k_j|(1+C)}}{k_j!(l_j!)^{|k_j|}} \Bigg).
\end{aligned}$$

Finally, as in Section 2.3, one checks that

$$\begin{aligned}
x_k^{-(\kappa+1)|\alpha|}(P_\infty)_k(x) &\leq e \left(\frac{(\kappa+1)|\alpha|}{e\kappa}\right)^{(\kappa+1)|\alpha|/\kappa} \\
&\leq e \left(\frac{\kappa+1}{\kappa}\right)^{(\kappa+1)|\alpha|/\kappa} (|\alpha|!)^{1+1/\kappa}.
\end{aligned}$$

Since we supposed $\kappa \geq 1$, it is true that $((\kappa+1)/\kappa)^{(\kappa+1)/\kappa} \leq 4$. As in the proof of Proposition 2.2.5, using Lemma 2.2.6, one can bound the sum by

$$\left(\frac{B}{A}(2\kappa(mA+1))^{|\alpha|} |\alpha|!\right)^{1+C}.$$

Thus we conclude that

$$\left|(f \circ P_\infty)^{(\alpha)}(x)\right| \leq \left(\tilde{B}(8\kappa(mA+1))^{|\alpha|} |\alpha|!\right)^{2+C+1/\kappa}.$$

$\square$

Remark 2.5.8. The mildness parameter C of the composition $f \circ P_\infty$ in this result is too large, it would be better to obtain $C + 1/\kappa$ instead. We will obtain this in a particular case in Proposition 2.5.10. This problem is caused by the upper bound in Inequality exp, which deviates from the proof strategy in Section 2.3 to show that P_∞ has $C = 1/\kappa$. More precisely, in Formula 2.1 in Section 2.3, we have only factored out a factor that can be bounded by $(|\alpha|!)^{1/\kappa}$, while here we have factored out a factor that can only be bounded by $(|\alpha|!)^{1+1/\kappa}$. It is necessary to factor out $P_\infty(x)_i$ in order to cancel the negative powers of $P_\infty(x)$ coming from upper bounds on the weakly (A, B, C)-mild function f. But then one can not use the full strength of Lemma 2.2.6 in Inequality exp, since there is no $\tilde{\lambda}!$ left in the summation.

Example 2.5.9. Let $U = \{(x, y) \in (0, 1)^2 \mid x < y\}$ and let $f : U \to \mathbb{R}$ be given by $f(x, y) = x^4 y^{-1} \ln(x)$. One can check that for all $\beta \in \mathbb{N}^2$ with $|\beta| \leq 1$ the function f is weakly $(5, 1, 0)$-mild up to order $+\infty$. By Proposition 2.5.7 for $\kappa \geq 1$, we have that $f \circ P_\infty$ is $(88\kappa, 1, 1 + 1/\kappa)$-mild up to order $+\infty$ on $P_\infty^{-1}(U) = U$. Note that

$$(f \circ P_\infty)(x, y) = e^{3-4/x^\kappa+1/y^\kappa}(1 - 1/x^\kappa).$$

We will improve the mildness parameter $C = 1 + 1/\kappa$ to $1/\kappa$ in Section 2.6.

Finally, we show the stronger result that we have mentioned in Remark 2.5.8 above. The proof generalizes the method of Section 2.3.

Proposition 2.5.10 ([VH21a, Proposition 4.5])**.** *Let $b : U \to \mathbb{R}$ be given by $b(x) = x^\mu$ for some $\mu \in \mathbb{R}^m$ and $P_\infty : (0, 1)^m \to (0, 1)^m$ be the map of Proposition 2.5.7. Suppose that for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$ the function $b^{(\beta)}$ is bounded. Then $b \circ P_\infty$ is $(A, B, 1/\kappa)$-mild up to order $+\infty$ on $P_\infty^{-1}(U)$, where*

$$A = \begin{cases} 2\kappa(2mM + 1) & \text{if } \kappa \geq 1, \\ (2/\kappa)^{1/\kappa}(2mM + 1) & \text{if } \kappa < 1, \end{cases}$$

with $M = \max(|\mu_1|, \ldots, |\mu_m|) + 1$ and $B = Ne^M$ with

$$N = \max_i(\sup_{x\in U} |b(x)|, \sup_{x\in U} |(1/\mu_i)(\partial b/\partial x_i)(x)|, 1).$$

Proof. To start, note that $b \circ P_\infty$ is given by:

$$(b \circ P_\infty)(x) = e^{\sum_{i=1}^m \mu_i(1-1/x_i^\kappa)}.$$

We consider this as the composition of the functions f and g, where

$$f(x) = e^{x_1+\ldots+x_m},$$

and
$$g(x) = (\mu_1(1 - 1/x_1^\kappa), \ldots, \mu_m(1 - 1/x_m^\kappa)) .$$
For the computations below, we may suppose that $\mu_1, \ldots, \mu_m$ are non-zero. Let $i \in \{1, \ldots, m\}$ and $\nu \in \mathbb{N}^m$ with $|\nu| \geq 1$. We have that
$$\left|g_i^{(\nu)}(x)\right| \leq \kappa(\kappa+1)\cdots(\kappa+|\nu|-1)\mu_i x_i^{-(\kappa+\nu_i)} \leq c(\nu,\kappa) M x_i^{-(\kappa+\nu_i)},$$
where $c(\nu, \kappa) = \kappa(\kappa+1)\cdots(\kappa+|\nu|-1)$ as in Section 2.3. Now let $\alpha \in \mathbb{N}^m$ with $|\alpha| \geq 1$ and $x \in (P_\infty)^{-1}(U)$. As before, we use the Faà di Bruno formula FdB and the triangle inequality together with this bound to find that
$$\begin{aligned}
&\left|(b \circ P_\infty)^{(\alpha)}(x)\right| = \left|(f \circ g)^{(\alpha)}(x)\right| \\
&\leq \sum_{1\leq|\lambda|\leq|\alpha|} \left|f^{(\lambda)}(g(x))\right| \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{\left|(g^{(l_j)}(x))^{k_j}\right|}{k_j!(l_j!)^{|k_j|}} \\
&\leq \sum_{1\leq|\lambda|\leq|\alpha|} |f(g(x))| \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! x^{-(\alpha+\kappa\lambda)} \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}} \\
&= \sum_{1\leq|\lambda|\leq|\alpha|} x^{-(\alpha+\kappa\lambda)} |f(g(x))| \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}},
\end{aligned} \tag{2.5}$$
where we have used the relations between α, λ, $k_1, \ldots, k_s$ and $l_1, \ldots, l_s$ to obtain the factor $x^{-(\alpha+\kappa\lambda)}$, and where $c(\kappa)$ is κ if $\kappa \geq 1$ and $c(\kappa)$ is 1 if $\kappa < 1$.

Let $k \in \{1, \ldots, m\}$ be such that $x_k = \min_i(x_i)$. Then we have that
$$x^{-(\alpha+\kappa\lambda)} \leq x_k^{-(|\alpha|+\kappa|\lambda|)}.$$
Note that it might be possible that x_k does not occur in the left hand side of this inequality or there might not even be differentiation with respect to x_k, but that is not important here.

Next, define $\mu' \in \mathbb{R}^m$ by $\mu'_i = \mu_i$ if $i \neq k$ and $\mu'_k = \mu_k - 1$, then we have that $\mu_k x^{\mu'} = (\partial b/\partial x_k)(x)$. We now rewrite $f \circ g$ as follows.
$$\begin{aligned}
f(g(x)) &= e^{\sum_{i=1}^m \mu_i(1-1/x_i^\kappa)} \\
&= e^{\mu_k(1-1/x_k^\kappa)} e^{\sum_{i=1}^m \mu'_i(1-1/x_i^\kappa)} \\
&= e^{\mu_k(1-1/x_k^\kappa)} \left(\left(\frac{1}{\mu_k}(\partial b/\partial x_k)\right) \circ P_\infty\right)(x)
\end{aligned}$$

Since $P_\infty(x) \in U$ and $b^{(\beta)}$ is bounded for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$, we have that $(1/\mu_k)(\partial b/\partial x_k)(P_\infty(x)) \leq N$. We can now further bound 2.5:

$$\sum_{1\leq|\lambda|\leq|\alpha|} x^{-(\alpha+\kappa\lambda)}\,|f(g(x))| \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}}$$

$$\leq \sum_{1\leq|\lambda|\leq|\alpha|} x_k^{-(|\alpha|+\kappa|\lambda|)} N e^{\mu_k(1-1/x_k^\kappa)} \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}}$$

$$\leq Ne^M x_k^{-|\alpha|} e^{-1/(2x_k^\kappa)} \sum_{1\leq|\lambda|\leq|\alpha|} x_k^{-\kappa|\lambda|} e^{-1/(2x_k^\kappa)} \left(\sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}} \right). \quad (2.6)$$

The proof now follows the last part of Section 2.3. First one checks that

$$x^{-|\alpha|} e^{-1/(2x_k^\kappa)} \leq \left(\frac{2\,|\alpha|}{e\kappa}\right)^{|\alpha|/\kappa} \leq \left(\frac{2}{\kappa}\right)^{|\alpha|/\kappa} (|\alpha|!)^{1/\kappa}$$

$$x^{-\kappa|\lambda|} e^{-1/(2x_k^\kappa)} \leq \left(\frac{2\,|\lambda|}{e}\right)^{\kappa|\lambda|/\kappa} \leq 2^{|\lambda|}\,|\lambda|!$$

and plugs this into 2.6:

$$Ne^M \left(\frac{2}{\kappa}\right)^{|\alpha|/\kappa} (|\alpha|!)^{1/\kappa} \sum_{1\leq|\lambda|\leq|\alpha|} 2^{|\lambda|}\,|\lambda|! \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}}.$$

Now apply Lemma 2.2.6 to get that

$$\sum_{1\leq|\lambda|\leq|\alpha|} 2^{|\lambda|}\,|\lambda|! \sum_{s=1}^{|\alpha|} \sum_{p_s(\alpha,\lambda)} \alpha! \prod_{j=1}^{s} \frac{(Mc(\kappa)^{|l_j|}\,|l_j|!)^{|k_j|}}{k_j!(l_j!)^{|k_j|}} \leq (c(\kappa)(2mM+1))^{|\alpha|}\,|\alpha|!$$

Putting everything together we find that

$$\left|(b\circ P_\infty)^{(\alpha)}(x)\right| \leq \left(Ne^M \left(c(\kappa) \left(\frac{2}{\kappa}\right)^{1/\kappa} (2mM+1) \right)^{|\alpha|} |\alpha|! \right)^{1+1/\kappa}.$$

Finally, if $\kappa \geq 1$, we have that

$$c(\kappa)\left(\frac{2}{\kappa}\right)^{1/\kappa} \leq 2\kappa$$

and if $\kappa < 1$ we have that

$$c(\kappa)\left(\frac{2}{\kappa}\right)^{1/\kappa} \leq \left(\frac{2}{\kappa}\right)^{1/\kappa}.$$

□

Remark 2.5.11. Unless stated otherwise, we will always use $\kappa \geq 1$ from now on, since the mildness parameters A and C are large if $\kappa < 1$.

Example 2.5.12. Consider the function of Example 2.5.5, which was $f(x, y) = x^2y^{-1}$ on the domain $U = \{(x, y) \in (0, 1)^2 \mid x < y\}$. Using this proposition, we find that $f \circ P_\infty$ is $(26\kappa, 2e^3, 1/\kappa)$-mild up to order $+\infty$.

Now this example might make you wonder: if I want to use a substitution to improve the derivatives of a weakly (A, B, C)-mild function up to order $+\infty$ (or suitably large $r \in \mathbb{N}$), with regard to obtaining the smallest mildness parameters A, B and C, should I use the power substitution or the exponential substitution?

Let us give some answer in the case that $f(x) = x^\mu$ with bounded first order derivatives. In that case we can use the stronger result of Proposition 2.5.10 instead of Proposition 2.5.7 for the exponential substitution, which makes a difference of a factor $|\alpha|!$. Using a power substitution, the function becomes $(rA, B, 0)$-mild up to order r and using the exponential substitution, it becomes $(\kappa A, B, 1/\kappa)$-mild up to order $+\infty$ for some $A, B > 0$, which we assume to be the same in both cases for simplicity. Thus, the bounds on the derivatives are

$$B(rA)^{|\alpha|}\,|\alpha|!$$

for the power substitution, and

$$(B(\kappa A)^{|\alpha|}\,|\alpha|!)^{1+1/\kappa}$$

for the exponential substitution. Now let us assume that $A \geq 1$ such that the bounds increase as $|\alpha|$ increases. We have that $r^r \geq r!$. Therefore, if r is large, it seems slightly better to use the exponential substitution for some $\kappa > 1$, from this point of view.

The mildness parameters A and B have a significant impact if the order of the derivatives is small. The formulas from Proposition 2.5.3 (or 2.5.6) seem better than in Proposition 2.5.10 (in terms of A and B, if not assumed equal) in that case.

2.6 Additional results

We generalize or improve some results on (A, B, C)-mild functions up to order r, introduce a variant that allows more flexibility and discuss a more general class of functions.

Improving Example 2.5.9

We will improve the mildness parameter C from $1 + 1/\kappa$ to $1/\kappa$ in Example 2.5.9, as we have mentioned there. To achieve this, we first show the following lemma.

Lemma 2.6.1. *Suppose $f : U \to \mathbb{R}$ is (A, B, C)-mild up to order $r \geq 1$. Then for all $\beta \in \mathbb{N}^m$ with $|\beta| = 1$, $f^{(\beta)}$ is $(2A, AB, C)$-mild up to order $r - 1$.*

Proof. Let $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r - 1$ and $x \in U$. Then we have that

$$\left|(f^{(\beta)})^{(\alpha)}(x)\right| = \left|f^{(\alpha+\beta)}(x)\right| \leq (BAA^{|\alpha|}(|\alpha| + 1)\,|\alpha|!)^{1+C},$$

where we used that $|\beta| = 1$. One finishes the proof using that for all $n \in \mathbb{N}$: $n + 1 \leq 2^n$. □

We now use this lemma to show that $f \circ P_\infty$ is $(A, B, 1/\kappa)$-mild up to order $+\infty$, where $f : U \to \mathbb{R}$ is given by $x^2 \ln(x)$, with $U = \{(x, y) \in (0, 1)^2 \mid x < y\}$. Note that $f \circ P_\infty$ is given by

$$e^{2(1-1/x^\kappa)} \left(1 - \frac{1}{x^\kappa}\right).$$

Recall from Section 2.3 that the function $\exp(1-1/x^\kappa)$ is $(6\kappa, e, 1/\kappa)$-mild up to order $+\infty$ if $\kappa \geq 1$. By the lemma above, we have that its first order derivative is $(12\kappa, 6e\kappa, 1/\kappa)$-mild up to order $+\infty$. Using Proposition 2.2.2, we conclude that the function

$$e^{1-1/x^\kappa}\, \frac{x}{\kappa} \left(\frac{\partial}{\partial x} e^{1-1/x^\kappa}\right) = \frac{1}{x^\kappa} e^{2(1-1/x^\kappa)}$$

is $(36\kappa, (6e\kappa)^3, 1/\kappa)$-mild up to order $+\infty$. Using Proposition 2.2.1, we conclude that $f \circ P_\infty$ is $(36\kappa, 2(6e\kappa)^3, 1/\kappa)$-mild up to order $+\infty$.

This computation shows that in Example 2.5.9, one could indeed improve the mildness parameter C from $1 + 1/\kappa$ to $1/\kappa$ (possibly enlarging the mildness parameters A and B) if one does not appeal to Proposition 2.5.7. Indeed, in that example we considered the function $x^4y^{-1}\ln(x)$ defined on the same domain U. We have that $x^4y^{-1}\ln(x) = (x^2y^{-1})(x^2\ln(x))$. The factor x^2y^{-1} satisfies the conditions of Proposition 2.5.3. Together with the above computations, this shows the claim on the mildness parameter C of the composition with P_∞.

Remark 2.6.2. Consider the function $x\ln(x)$ on $(0,1)$. Its first order derivative, $\ln(x)+1$, is unbounded. However, the computations above show that its composition with P_∞, given by:

$$e^{1-\frac{1}{x^\kappa}}\left(1-\frac{1}{x^\kappa}\right),$$

is $(A,B,1/\kappa)$-mild up to order $+\infty$ for some $A,B>0$. Of course, the monomial appearing in front of $\ln(x)$, just x in this case, does obviously have bounded first order derivative.

In this way, it seems that the exponential substitution for monomials, i.e., Proposition 2.5.10, can be extend to some class of functions where logarithms are involved.

Non-examples

Firstly, if $f\in\mathcal{C}^\infty(U)$ has bounded derivatives, this does not imply that f is (A,B,C)-mild up to order $+\infty$ for some $A,B>0$ and $C\geq 0$. To construct an example of such a function, consider the sequence $((i!)^i)_i$ and fix some $x\in U$. By Borel's Theorem (see for instance [Bes14]), there is some smooth function f, defined on an open neighborhood of x, whose i-th order derivative at x is given by $(i!)^i$. For any choice of $A,B>0$ and $C\geq 0$, $(BA^i i!)^{1+C}$ will eventually be smaller than $(i!)^i$, thus this function f cannot be (A,B,C)-mild up to order $+\infty$ for any $A,B>0$ and $C\geq 0$.

While I am not aware of an "easy example" of a smooth function that has bounded derivatives, but that is not (A,B,C)-mild up to order $+\infty$ for some $A,B>0$ and C, let me mention a function that I have in mind and an indication why I think so. Consider the function $f:(0,1)\to(0,1)$ given by $f(x)=x^{-\ln(x)}=e^{-\ln(x)^2}$. One can check that the derivatives of $-\ln(x)^2$ are given by:

$$\left(-\ln(x)^2\right)^{(k)} = -\frac{a_k+b_k\ln(x)}{x^k},$$

with $b_{k+1}=-kb_k$, $a_{k+1}=b_k-ka_k$, $b_1=2$ and $a_1=0$. After some computations using the Faà di Bruno formula FdB for the functions e^x and $-\ln(x)^2$, one finds that the derivative of order α is a sum over λ, where each summand has a factor of the form

$$\frac{\ln(x)^\lambda}{x^\alpha} e^{-\ln(x)^2}.$$

These functions have a critical point at $x = e^{(-\alpha-\sqrt{\alpha^2+8\lambda})/4}$. In this way, we end up with a factor of order e^{α^2}, which cannot be bounded by $(\alpha!)^{1+C}$. While this does not yet prove that this sum could not be bounded in a suitable way such that f is (A, B, C)-mild up to order $+\infty$, I think it seems rather unlikely for this reason. Finally, note that all derivatives of f vanish at 0 and that it is a bijection $(0, 1) \to (0, 1)$. Therefore it, could potentially be used as another type of exponential substitution.

More on power substitutions

Let $P : U \subset (0, 1)^2 \to \operatorname{Im}(P)$ be given by

$$P(x, y) = (x^4, x^3 y^{-1}),$$

where $U = \{(x, y) \in (0, 1)^2 \mid x < y\}$. This map is $(4, 1, 0)$-mild up to order 2. Now let $r \in \mathbb{N}$ and recall that the map $P_r : (0, 1)^2 \to (0, 1)^2$ is given by $P_r(x, y) = (x^r, y^r)$. Clearly, we have that $(P \circ P_r)(x, y) = (x^{4r}, x^{3r} y^{-r})$, which is $(4r, 1, 0)$-mild up to order r by Proposition 2.5.6. In fact, the mildness is up to order $2r$ here.

Now let $f : V \subset \operatorname{Im}(P) \to \mathbb{R}$ be such that for all $\beta \in \mathbb{N}^2$ with $|\beta| \leq 1$ the function $f^{(\beta)}$ is weakly (A, B, C)-mild up to order $+\infty$. Then $f \circ (P \circ P_r)$ is (A', B', C)-mild up to order $2r$ for some $A', B' > 0$. To show this, one has to slightly modify the proof of Proposition 2.5.3. More precisely, let $\alpha \in \mathbb{N}^2$ with $|\alpha| \leq 2r$ and $(x, y) \in U$. As usual, we fix a term in the Faà di Bruno formula FdB. We now have

$$\left|(f \circ (P \circ P_r))^{(\lambda)}(x)\right| \leq (P \circ P_r)(x)^{-\lambda'}((B/A)A^{|\lambda|}\,|\lambda|!)^{1+C},$$

where $\lambda' \in \mathbb{N}^2$ such that $\lambda = \lambda' + \beta$ for some $\beta \in \mathbb{N}^2$ with $|\beta| = 1$, which we can pick later. One checks that

$$\left|\prod_{j=1}^{s}((P \circ P_r)^{(l_j)}(x))^{k_j}\right| \leq (P \circ P_r)^\lambda(x, y)^{-\alpha}((4r)^{|\alpha|}\,|\alpha|!).$$

We now have to choose β such that $(P \circ P_r)^\beta(x, y)^{-\alpha}$ is bounded. If $\alpha_2 \neq 0$, then we may suppose $\lambda_2 \neq 0$. Indeed, if $\lambda_2 = 0$ and $\alpha_2 \neq 0$, this term will be zero. Thus in this case, $\alpha_2 \neq 0$, we pick $\beta = (0, 1)$. Then we obtain a factor

$x^{3r}y^{-r}$ which can be used to bound $(x,y)^{-\alpha}$ since the derivatives up to order $2r$ of $x^{3r}y^{-r}$ are bounded. If $\alpha_2 = 0$, the choice of β doesn't matter since both x^{4r} and $x^{3r}y^{-r}$ can be used to bound $(x,y)^{-\alpha}$. We conclude that in each term, we can bound all powers of (x,y) independent of (x,y) and then we can apply Lemma 2.2.6 to find an expression for A' and B'.

In this way, one can come up with more maps P that could be used as a power substitution. More precisely, one could look at maps whose component functions are of the form

$$x^{\mu}F(x)$$

for some $\mu \in \mathbb{R}$ such that all derivatives up to order r of x^{μ} are bounded and where F is non-vanishing and (A,B,C)-mild up to order r for some $A,B > 0$ and $C \geq 0$. Since F is non-vanishing, it will only contribute to the mildness parameters of the composition, it is the factor x^{μ} that makes it, possibly, a power substitution. Indeed, if P is of this form, it might not be completely clear that it is suitable to be a power substitution. One sees that, in order to be able to argue as above, one should be able to do the following. For all x in the domain of P, all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$, $\lambda \in \mathbb{N}^m$ with $1 \leq |\lambda| \leq |\alpha|$, $s \in \{1, \ldots, |\alpha|\}$ and $k_1, \ldots, k_s, l_1, \ldots, l_s \in p_s(\alpha, \lambda)$, if the corresponding term in the Faà di Bruno formula is non-zero, there should be a component function P_i, with $\lambda_i \neq 0$, such that

$$P_i(x)x^{-\alpha}$$

is bounded independently of x. A stronger condition, but that is easier to see, is that for all monomials in the component functions, "all derivatives up to order r correspond (up to a nonzero factor) to division by variables" or, equivalently, "is of order r". Let us put this in a formal definition.

Definition 2.6.3. Let $b : U \subset (0,1)^m \to \mathbb{R}$ be a function of the form $b(x) = x^{\mu}$ for some $\mu \in \mathbb{R}^m$. Then we say that b **has order r** $\in \mathbb{N}$ if for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ such that $\alpha_i = 0$ if $\mu_i = 0$ $(i = 1, \ldots, m)$, and $x \in U$:

$$b^{(\alpha)}(x) = c(\alpha, \mu)x^{-\alpha}b(x),$$

for some *nonzero* constant $c(\alpha, \mu)$ that depends on α and μ and there exists a $B > 0$ such that for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$: $\left|b^{(\alpha)}(x)\right| \leq B$.

So to obtain a power map, we need a map, all of whose component functions have order r, and that is bijective.

For example, let U be some subset of $(0,1)^2$ such that $x^{3/2}y^{-1}$ has bounded derivatives up to order r, then $x^{3/2}y^{-1}$ has order r. But for instance, if one would consider instead the function x^2y^{-1} and $r \geq 3$, then $(\partial/\partial x)^3(x^2y^{-1})$ is not, up to a nonzero factor, equal to $x^{-1}y^{-1}$. Note that the component functions of the power substitution P_r satisfy this condition, thus P_r is a power map in this more general sense.

More on compositions

One can prove a result similar to Proposition 2.2.5, where there is a separate mildness parameter A for each variable x_i. This idea is also in the paper of Gevrey [Gev18]. To this end, we slightly adjust the definition of these functions.

Definition 2.6.4. Let $A_1, \dots, A_m, B > 0$ and $C \geq 0$ be real numbers, denote $A = (A_1, \dots, A_m)$ and let $r \in \mathbb{N} \cup \{+\infty\}$. A function $f : U \to \mathbb{R}$ is (A, B, C)-mild up to order r if $f \in \mathcal{C}^r(U)$ and if for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in U$:

$$\left| f^{(\alpha)}(x) \right| \leq (BA^\alpha \, |\alpha|!)^{1+C}.$$

A map $f : U \to \mathbb{R}^n$ is (A, B, C)-mild up to order r if all of its component functions are.

We have the following result on compositions of these functions, which is a more precise version of Proposition 2.2.5.

Proposition 2.6.5. *Suppose that $f : U \to \mathbb{R}$ is (A_f, B_f, C)-mild up to order r and $g : V \subset \mathbb{R}^p \to U$ is (A_g, B_g, C)-mild up to order r. Then their composition is (A, B, C)-mild up to order r with*

$$A_i = A_{g,i}(|A_f| \, B_g + 1),$$

$$B = \frac{|A_f| \, B_f B_g}{|A_f|_g \, B_g + 1} < B_f,$$

where $|A_f| = A_{f,1} + \ldots + A_{f,m}$.

Proof. We only have to show a refinement of Lemma 2.2.6. Define F and G as follows:

$$F(x_1, \dots, x_m) = \frac{b}{1 + a_1(B - x_1) + \ldots + a_m(B - x_m)},$$

$$G_i(x_1, \dots, x_p) = \frac{B}{1 - A_1 x_1 - \ldots - A_p x_p} \quad (i = 1, \dots, m).$$

One checks that for all $\nu \in \mathbb{N}^m$

$$F^{(\nu)}(B, \dots, B) = F^{(\nu)}(G(0, \dots, 0)) = ba^\nu \, |\nu|!$$

and that for all $\nu \in \mathbb{N}^p$

$$G^{(\nu)}(0, \dots, 0) = BA^\nu \, |\nu|!.$$

Using that $|a| = a_1 + \ldots + a_m$, we have that

$$(F \circ G)(x) = \frac{b(1 - A_1 x_1 - \ldots - A_p x_p)}{(1 + |a| B)(1 - A_1 x_1 - \ldots - A_p x_p) - |a| B}.$$

One then computes directly that for all $\alpha \in \mathbb{N}^p$

$$(F \circ G)^{(\alpha)}(0, \ldots, 0) = \frac{|a| bB}{|a| B + 1} \tilde{A}^{\alpha} |\alpha|!,$$

with $\tilde{A}_i = A_i(|a| B + 1)$. □

Our last result is a refinement of Lemma 2.5.1, which now rescales the variable x_i with respect to the mildness parameter A_i of that variable. The proof is completely analogous.

Lemma 2.6.6. *Suppose $f : U \to \mathbb{R}$ is (A, B, C)-mild up to order $r \in \mathbb{N}$. Let $P_\ell : \mathbb{R}^m \to \mathbb{R}^m$ be the map given coordinatewise by*

$$x_i \mapsto \frac{1}{(A_i r)^{1+C}} x_i.$$

Then for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and $x \in P_\ell^{-1}(U)$: $\left|(f \circ P_\ell)^{(\alpha)}(x)\right| \leq B^{1+C}$.

Denjoy-Carleman classes

The (A, B, C)-mild functions up to order $+\infty$ are part of a more general class of functions studied by Denjoy [Den21] and Carleman [Car26]. They are defined as follows.

Definition 2.6.7. Let $M = (M_n)_n$ be a sequence of positive real numbers and suppose U is open in $\mathbb{R}^m$. The **Denjoy-Carleman class $\mathcal{C}(M, U)$** associated to M is the set of all functions f such that $f \in \mathcal{C}^\infty(U)$ and there exist $A, B > 0$ such that for all $x \in U$ and $\alpha \in \mathbb{N}^m$:

$$\left| f^{(\alpha)}(x) \right| \leq B A^{|\alpha|} M_{|\alpha|}.$$

Remark 2.6.8. Sometimes A and B are "absorbed" into M_n and one could say we have absorbed $|\alpha|!$ into $M_{|\alpha|}$ already.

For example, if $M_n = n!$, then $f \in \mathcal{C}(M, U)$ if and only if f is $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$, in particular it is analytic. Note that for functions in more variables, it is also natural to consider a sequence indexed over $\alpha \in \mathbb{N}^m$, this is called "G-mild" in [Tho11], where G is a function $\mathbb{N}^m \to (0, +\infty)$.

In general, a Denjoy-Carleman class is not closed under, for instance, composition and differentiation. One has to put some conditions on the sequence M to obtain stability under these operations. In particular, in [RS14] they show that these are closed under composition if and only if M "satisfies the Faà di Bruno Property", say a stability property for sums as in Lemma 2.2.6. Clearly, (A, B, C)-mild functions up to order r satisfy this property, whatever it precisely means.

If one takes $M_n = (n!)^\alpha$, one obtains the **Gevrey functions of class $\boldsymbol{\alpha}$**. Clearly, if $\alpha \geq 1$, the function is (A, B, C)-mild up to order $+\infty$ for some A, B, where $C = \alpha - 1$, this is [VH21b, Proposition 2.4]. We have also briefly mentioned this in the introduction of this chapter. It was already indicated by Gevrey in [Gev18] that if $\alpha \in (0, 1)$, this class of functions is not closed under composition.

Chapter 3

Parametrizations

In this chapter, sections 3.3 and 3.4 correspond to the results in [VH21b, Section 3]. Section 3.5 corresponds to the proof of [VH21a, Theorem 4.7].

In this chapter we prove the main results of this thesis. More precisely, we construct two kinds of parametrizations of a bounded power-subanalytic set $X \subset \mathbb{R}^n$. A parametrization is a finite set of functions such that the union of their images cover X. There are different types of parametrizations, depending on the conditions you put on the functions that parametrize X. The main results are a C^r-parametrization theorem (Theorem 3.3.1), where the functions are r times continuously differentiable for some $r \in \mathbb{N}$ and where we also bound the C^r-norm, and a mild parametrization theorem (Theorem 3.5.1), where we parametrize curves with C^∞-functions with a bound on all their derivatives.

C^r-parametrizations

In [Yom87b], Yomdin sketched a proof of a C^r-parametrization theorem for semi-algebraic sets, which he made more explicit in an addendum to this article [Yom87a]. The relevance of the work by Yomdin, in smooth dynamics, has been elaborated on and further refined by Gromov in [Gro87]. In that paper, the C^r-parametrization theorem of Yomdin, now often referred to as an “Algebraic Lemma”, is further elaborated on. While the proofs by Yomdin and Gromov are correct, a complete proof of the Algebraic Lemma has been written down two decades later by Burguet in [Bur08]. The precise statement is given in Theorem 3.2.1 in Section 3.2, where we will apply this construction to a simple example of algebraic curves by Yomdin.

Around the same time, Pila and Bombieri developed in [BP89] the so-called "determinant method", which requires a suitable parametrization theorem as an input. This suitable parametrization theorem did not exist at that time and therefore, to apply this method, an ad hoc parametrization was considered (for instance in [Pil10]) or it was stated as a conjecture ([Pil06a]). We further discuss these results and the determinant method in Chapter 4.

Note that in the same period, o-minimality was developed in the series of papers [vdD84, PS86, KPS86, PS88]. Definable sets in o-minimal structures share many of the properties of semi-algebraic sets. Hence, this had lead to a C^r-parametrization theorem (see Theorem 3.2.2) for definable sets in o-minimal structures in the well celebrated work of Pila and Wilkie [PW06]. While some model theory is used to prove the existence of C^r-parametrizations in arbitrary o-minimal structures, the techniques to construct the parametrization are essentially the same as the methods of Yomdin and Gromov. I try to indicate this in Section 3.2, while working through the example. Coupling this parametrization result with the determinant method, yields the fundamental Pila-Wilkie Counting Theorem, which is the main result of [PW06].

The results by Pila, Wilkie and Yomdin are existence results. For the applications, namely in smooth dynamics and number theory, it is interesting to know how many functions their method constructs, i.e., what is the number of charts that are used to parametrize a set X, as we have briefly mentioned in the introduction, more on this is also in Chapter 4. This number may of course be expected to depend on the required differentiability order r, but also on the dimension m of X and the dimension of the ambient space $\mathbb{R}^n$. Moreover, if X is semi-algebraic, it comes with a natural notion of "complexity" (think about the degree of an algebraic curve, see also below Theorem 3.2.1), and thus one might wonder how the construction depends on the complexity of X. None of these have been studied in the work by Yomdin or Pila and Wilkie.

For general o-minimal structures this question is, as far as I know, widely open. However, recently some results were obtained for subanalytic and power-subanalytic sets. In [BN19], Binyamini and Novikov prove a C^r -parametrization theorem for subanalytic sets. Their construction yields cr^m functions for some constant $c > 0$. Thus, polynomial in r, where the degree is the dimension m of the parametrized set X. Moreover, if X is semi-algebraic, they show that c is polynomial in the complexity of X. In [CPW20], there is a C^r-parametrization theorem for power-subanalytic sets, which also achieves polynomial dependence on r. However, the degree of this polynomial is not known, and they did not deduce an additional result on the constant c in the particular case that X is semi-algebraic.

The initial goal of my PhD was to make this polynomial dependence in the result in [CPW20] more explicit. This has been achieved in my paper [VH21b], where I show that the degree is m^3, where m is the dimension of the considered set X.

Mild parametrizations

The mild parametrizations have already been in the story of the C^r-parametrizations. More precisely, the parametrizations considered by Pila in [Pil06a, Pil10], are actually mild parametrizations, i.e., the set is parametrized with maps that are (A, B, C)-mild up to order $+\infty$, as defined in Chapter 2. From now on, I will call them C-mild parametrizations, where C should be understood as the mildness parameter C. If a set X is parametrized by a C-mild parametrization, one can deduce a C^r-parametrization from it, after a linear reparametrization, see Corollary 3.1.7. From this point of view, a C-mild parametrization is a C^r-parametrization for all $r \in \mathbb{N}$.

In [JMT11], Jones, D. Miller and Thomas show that any subanalytic set has a 0-mild parametrization. However, their result is not uniform. Indeed, there is a counter example by Yomdin [Yom08], the family of algebraic curves we have mentioned before, that does not have a uniform 0-mild parametrization. In the recent result [BN19] it is also shown that any family of subanalytic sets has a 2-mild parametrization. Very recently I have shown that any family of power-subanalytic curves has a C-mild parametrization for all $C > 0$ [VH21a].

Remark that there is a small catch in the recent C-mild parametrization results. In all C^r-parametrization results X is definable in some o-minimal structure, and so are the maps of the constructed C^r-parametrization of X. This also holds for the 0-mild parametrization result of [JMT11]. But this is not the case for the 2-mild parametrization in [BN19] and neither for my result in [VH21a] on curves. In the latter, the maps are definable in $\mathbb{R}_{\mathrm{an,exp}}$. This definability is an issue if one wants to use tools from model theory to construct uniform parametrizations "automatically". For instance, the C^r-parametrization theorem for general o-minimal structures in [PW06] uses such a model theoretic argument instead of carefully treating the parameters.

Finally, let us mention that there is another type of parametrizations that deals with the issue that there is no uniform 0-mild parametrization result. The idea is to exclude some small set, where small depends on some parameter $\delta > 0$, which then resolves the obstruction. This has been suggested by Yomdin in [Yom08], see also [Yom15, Section 3.3]. This type of parametrizations are called "analytic-δ-parametrizations" and are extended to the power-subanalytic setting in [CFY21]. The charts of this type of parametrizations are called "analytic K-charts", or simply a-charts, and are automatically 0-mild due to their nature.

Overview of this chapter

The chapter is organized as follows. We first introduce some additional notation, in particular the C^r-norm is defined. Most importantly, we will work with families of sets $\mathcal{X}$ instead of a single set X. Therefore, we have to introduce parameters, and the necessary conventions with respect to this are made in Section 3.1.

The next three sections are devoted to C^r-parametrizations, which we defined in the introduction and roughly in the previous paragraph. In Section 3.2, we explain the classical construction using a well known example by Yomdin, which is a family of algebraic curves. This should make the reader familiar with some concepts that will be generalized in sections 3.3 and 3.4. In Section 3.3 the C^r-parametrization theorem, using a so-called "pre-parametrization" and results of chapters 1 and 2. This pre-parametrization theorem is proved in Section 3.4. Finally, in Section 3.5, we prove Theorem 3.5.1, a mild parametrization theorem for curves. The main ingredient is also the pre-parametrization theorem of Section 3.4.

The pre-parametrization theorem (Theorem 3.3.4), is a key ingredient in the proof of the two main results. As its name suggests, it is some kind of parametrization. More precisely, the functions are continuously differentiable and have bounded first order derivatives. Moreover, these functions have a specific form, they are essentially monomials. This allows us to use the substitutions that we have studied in Chapter 2. For the C^r-parametrizations, this will be a power substitution, and for the mild parametrization, this will be the exponential substitution.

3.1 Notation and definitions

Throughout this chapter, X denotes an m-dimensional subset of $\mathbb{R}^n$. If $m = 1$, i.e., X is a curve, we denote C instead. We will study families of such sets, which are denoted by $\mathcal{X} = \{X_t \mid t \in T\}$, where $T \subset \mathbb{R}^k$ is some space of parameters. The dimension of a family $\mathcal{X}$ is the maximum of the dimension of its family members.

We will frequently work with families of maps $f : U \subset T \times \mathbb{R}^m \to \mathbb{R}^n$. We denote a family member of f by f_t, which is by definition

$$f_t : U_t \to \mathbb{R}^n : x \mapsto f(t, x),$$

where $U_t = \{x \in \mathbb{R}^m \mid (t,x) \in U\}$ is the corresponding family member of U. When we have to write more indices, we will write the parameter last. For instance, $f_{i,t}$ for the i-th component function of f_t, instead of $f_{t,i}$.

We will study two types of parametrizations: C^r-parametrizations and mild parametrizations. For the first one, we introduce the following norm.

Definition 3.1.1. Let $r \in \mathbb{N}$ and suppose that $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ is C^r. Then we define its ***C^r*-norm** by

$$|f|_r = \max_{i=1}^{n} \sup_{\substack{x \in U \\ \{\alpha \in \mathbb{N}^m \mid |\alpha| \le r\}}} \frac{\left| f_i^{(\alpha)}(x) \right|}{|\alpha|!}.$$

Remark 3.1.2.

1. Note that $|f|_r$ can be $+\infty$, for example, consider $|\sqrt{x}|_1$ for $U = (0,1)$. Thus, in the strict sense, it is not a norm.
2. If $1 < |\alpha| \le r$, then $|f|_r \le 1$ does not imply that $\left| f^{(\alpha)} \right|_0 \le 1$, due to the division by $|\alpha|!$ in the definition. Consider the C^r-norm as a weighted version of the C^r supremum norm.
3. We define the C^r-norm for a family of functions $f : U \subset T \times \mathbb{R}^m \to \mathbb{R}^n$ as follows:
$$|f|_r = \sup_{t \in T} |f_t|_r \,.$$

Definition 3.1.3. A ***C^r*-parametrization** of a set X is a finite collection of maps

$$\{f_j : (0,1)^m \to X \mid j \in \{1, \dots, N\}\}$$

such that

$$\bigcup_{j=1}^{N} \mathrm{Im}(f_j) = X$$

and $|f_j|_r \le 1$ for all $j \in \{1, \dots, N\}$. If $\mathcal{X}$ is an m-dimensional family of subsets of $\mathbb{R}^n$, a **uniform *C^r*-parametrization** is a finite collection of families of maps

$$\{f_j : T \times (0,1)^m \to \mathcal{X} \mid j \in \{1, \dots, N\}\}$$

such that for all $j \in \{1, \dots, N\}$ we have that $|f_j|_r$ is at most 1 and for all $t \in T$ the set $\{f_{j,t} \mid j \in \{1, \dots, N\}\}$ is a C^r-parametrization of X_t.

Remark 3.1.4.

1. Since we require that $|f_j|_r \leq 1$, it follows that $X \subset [-1,1]^n$. This is no obstruction for the applications in Chapter 4, where we will map X into $[-1,1]^n$ if necessary. In general, if X is bounded, one can also cover X with finitely many images of $(0,1)^m$ and then C^r-parametrize each of those pieces.

2. The number of maps N, which we also call **charts**, may be expected to depend on r. It is an important part of this thesis to study how a construction of a C^r-parametrization depends on r. It also depends on the chosen C^r-norm, see (2) of Remark 3.1.2.

3. We only impose an upper bound on the derivatives of the family members of f_j, so there is no condition with respect to partial differentiation with respect to parameter variables. Moreover, it is not required that f_j continuously depends on t.

4. It is not necessary that the family of functions f_j is defined for all $t \in T$. If not, one can trivially extend them or use the "empty function".

Convention. Recall from Chapter 2, Definition 2.1.1, that a map f is (A,B,C)-mild up to order $r \in \mathbb{N} \cup \{+\infty\}$ if it is C^∞ on U and if for all $i \in \{1,\dots,m\}$, $x \in U$ and $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$:

$$\left| f_i^{(\alpha)}(x) \right| \leq (BA^{|\alpha|}\,|\alpha|!)^{1+C}.$$

If f is a family of maps, we say that it is (A,B,C)-mild up to order r if f_t is (A,B,C)-mild up to order r for all $t \in T$, i.e., the same A, B and C hold for all $t \in T$.

Definition 3.1.5. Let $C > 0$. A ***C*-mild parametrization** of a set X is a finite collection of maps

$$\{f_j : (0,1)^m \to X \mid j \in \{1,\dots,N\}\}$$

such that

$$\bigcup_{j=1}^{N} \operatorname{Im}(f_j) = X$$

and there exist $A, B > 0$ such that f_j is (A,B,C)-mild up to order $+\infty$ for all $j \in \{1,\dots,N\}$. If $\mathcal{X}$ is an m-dimensional family of subsets of $\mathbb{R}^n$, a **uniform *C*-mild parametrization** is a finite collection of families of maps

$$\{f_j : T \times (0,1)^m \to \mathcal{X} \mid j \in \{1,\dots,N\}\}$$

such that there exist $A, B > 0$ such that f_j is (A, B, C)-mild up to order $+\infty$ for all $j \in \{1, \ldots, N\}$ and for all $t \in T$, the set $\{f_{j,t} \mid j \in \{1, \ldots, N\}\}$ is a C-mild parametrization of X_t.

Using a suitable linear substitution, as in Lemma 2.5.1, one can derive for all $r \in \mathbb{N}$ a C^r-parametrization from a C-mild parametrization, see Corollary 3.1.7 below. Wel call this a **linear reparametrization**.

Lemma 3.1.6. *Suppose $f : U \subset T \times \mathbb{R}^m \to \mathbb{R}$ is (A, B, C)-mild up to order $r < +\infty$. Let $P_\ell : T \times \mathbb{R}^m \to T \times \mathbb{R}^m$ be the map defined by by*

$$(t, x_1, \ldots, x_m) \mapsto \left(t, \frac{1}{A^{1+C} r^C} x_1, \ldots, \frac{1}{A^{1+C} r^C} x_m\right).$$

Then we have that $|f \circ P_\ell|_r \leq B^{1+C}$ on $P_\ell^{-1}(U)$.

The proof is completely analogous to the proof of Lemma 2.5.1. Now suppose that in this lemma $U = T \times (0,1)^m$. It follows that $P_\ell^{-1}(U) = T \times (0, A^{1+C} r^C)^m$. Clearly this set can be covered by at most $(A'^{1+C} r^C)^m$ translates of $T \times (0,1)^m$ for some A' slightly larger than A. In this way, we obtain the following result.

Corollary 3.1.7. Suppose $\mathcal{X} \subset T \times [-1,1]^n$ has a C-mild parametrization, say every chart is $(A, 1, C)$-mild up to order $+\infty$ for some $A > 0$. Then for every $r \in \mathbb{N}$, this C-mild parametrization induces via a linear reparametrization a C^r-parametrization consisting of cr^{Cm} charts for some constant $c > 0$ that only depends on A, C and m.

Note that we may assume $B = 1$, after possibly enlarging A, since $\mathcal{X} \subset T \times [-1,1]^n$, see Remark 3.1.4 (1).

Remark 3.1.8. Clearly, if one is only interested in some particular $r \in \mathbb{N}$, then it is not necessary that $\mathcal{X}$ is parametrized with charts that are mild up to order $+\infty$, instead, it suffices that $\mathcal{X}$ is parametrized by charts that are mild up to order r.

Finally we make the following convention regarding operations with (weakly) (A, B, C)-mild functions up to order r.

Convention. We will frequently use the formulas for addition (2.2.1), multiplication (2.2.2) and composition (2.2.5) of (A, B, C)-mild functions up to order r. When performing these operations via these results, we will only keep track of the dependence on r, and, by abuse of notation, use the same A, B and C again, even tough they might be different. For instance, the addition of two (Ar, B, C)-mild functions up to order r is then again (Ar, B, C)-mild up to order r, and the composition of these functions is (Ar^2, B, C)-mild up to order r. We make the same convention for weakly (A, B, C)-mild functions up to order r, including the use of Proposition 2.4.5 on the composition of an (A, B, C)-mild function up to order r with a weakly (A', B', C)-mild function up to order r. In the examples, the constants A, B and C will be computed more explicitly.

3.2 Yomdin-Gromov parametrizations

In this section we explain some classical techniques that were used by Yomdin and Gromov to construct C^r-parametrizations by working through a simple and well known example of algebraic curves $\mathcal{C}$, following the exposition in [Yom15, Section 4.3]. Throughout, we roughly explain how this can be achieved "in general", which means for any $\mathcal{X}$ that is definable in an o-minimal structure on $\mathbb{R}$, as defined in Chapter 1.

The Yomdin-Gromov parametrizations are C^r-parametrizations of semi-algebraic sets. Their result is frequently referred to as the **Yomdin-Gromov Algebraic Lemma**, which states the following.

Theorem 3.2.1 ([Bur08, Theorem 1]). *Let $A \subset [0,1]^n$ be a compact semi-algebraic set of dimension m. There exist an integer N and continuous semi-algebraic maps $\phi_1, \dots, \phi_N : [0,1]^m \to [0,1]^n$ such that*

1. *ϕ_i is analytic on $(0,1)^m$;*
2. *$\max_{\beta:|\beta|\leq r} \left|\left| \phi_i^{(\beta)}\Big|_{(0,1)^m} \right|\right|_\infty \leq 1$;*
3. *$\bigcup_{i=1}^N Im(\phi_i) = A$.*

Moreover, N and $\deg(\phi_i)$ are bounded by a function of $\deg(A)$, n and r.

Here, the degree $\deg(A)$ of a semi-algebraic set A is by definition the sum of the total degrees of all polynomials that are used to define A. Together with n, this is usually called the **complexity** of A. We will show that for our example, the construction leads to $N = 2 + 4^r$ charts. Therefore, we conclude that the number of charts N in this construction depends at least exponentially on r.

There is also a recent proof of the Yomdin-Gromov Algebraic Lemma by Binyamini and Novikov [BN20], using the methods of their paper [BN19]. The Algebraic Lemma has been generalized to sets definable in an o-minimal structure by Pila and Wilkie [PW06], with a small loss of some properties. In particular, there is no information on the number of charts N, it is just finitely many, and one cannot expect that the charts are analytic on $(0,1)^m$, see Remark 1.3.7. The version below is formulated for o-minimal structures on $\mathbb{R}$, as defined in Chapter 1, but their result holds for so-called "o-minimal expansions of the real field".

Theorem 3.2.2 ([PW06, Theorem 2.3]). *For all $r \in \mathbb{N}$ and bounded, definable set X, there exists a C^r-parametrization of X.*

Let us now focus on the example that will highlight many of the important tools in the proofs of these results. Consider the family of algebraic curves

$$\mathcal{C} = \{(t,x,y) \in (0,1) \times (-1,1)^2 \mid xy = t\},$$

where $t \in T = (0,1)$. This is basically the example by Yomdin in [Yom08, Proposition 3.3] that we have mentioned in the introduction. Clearly, $\mathcal{C}$ is a family of hyperbolas C_t given by $y = t/x$, intersected with $(-1,1)^2$. You can see some of the family members in Figure 3.1.

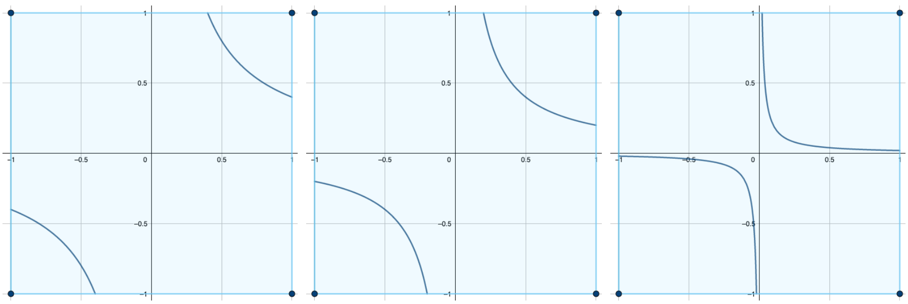

Figure 3.1: Some family members of $\mathcal{C}$.[1]

A C^0-parametrization of $\mathcal{C}$

We start with constructing a uniform C^0-parametrization of $\mathcal{C}$. We see that for all $t \in T = (0,1)$, C_t is the graph of the function $f_t : U_t \to (-1,1) : x \mapsto t/x$, where

$$U_t = \{x \in (-1,1) \mid t < |x| < 1\}.$$

Thus, we consider f as a family of functions $U \to \mathbb{R}$ defined by $(t,x) \mapsto t/x$, where $U = \{(t,x) \in T \times (-1,1) \mid t < |x| < 1\}$. We split U into the two parts

$$U_- = \{(t,x) \in T \times (-1,1) \mid -1 < x < -t\}$$

[1]Copyright © International GeoGebra Institute, 2021

and

$$U_+ = \{(t,x) \in T \times (-1,1) \mid t < x < 1\}.$$

We now have two families of functions, namely the restriction f_- of f to U_- and the restriction f_+ of f to U_+, such that C_t is the union of the graphs of $f_{+,t}$ and $f_{-,t}$. Moreover, the domain of $f_{+,t}$ and $f_{-,t}$ are intervals. This can be achieved in general using the cell decomposition theorem. In that case, we obtain finitely many families of functions $f : U \to (-1,1)^{n-m}$ such that for all $t \in T$, X_t is the union of the graphs of f_t and where the domain U_t of f_t is a cell in $(-1,1)^m$.

Our final step to construct the C^0-parametrization is to map $(0,1)$ onto the domain. To this end, consider the following maps:

$$\Phi_+ : T \times (0,1) \to U_+ : (t,x) \mapsto (t, t + (1-t)x),$$

$$\Phi_- : T \times (0,1) \to U_- : (t,x) \mapsto (t, -1 + (-t+1)x).$$

We then obtain a C^0-parametrization $\{f_1, f_2\}$, with

$$f_1(t,x) = (\Phi_+(t,x), (f \circ \Phi_+)(t,x))$$

and

$$f_2(t,x) = (\Phi_-(t,x), (f \circ \Phi_-)(t,x)).$$

In general, one maps $(0,1)^m$ onto a cell D using a composition of linear homotopies, more precisely, maps of the form

$$\alpha_i + (\beta_i - \alpha_i)x_i,$$

where α_i and β_i are the walls of the cell bounding the variables x_i form below and from above respectively. (If $x_i = \alpha_i(x_1, \ldots, x_{i-1})$, it is simply α_i.) The map $\Phi : (0,1)^m \to D$ constructed in this way is of course bounded since $D \subset [-1,1]^n$. However, there is no guarantee that this map has bounded derivatives and therefore should be treated as well in the following computations. We will not make this more precise in this section. Let us remark that these homotopies have bounded derivatives if the walls have bounded derivatives, but these walls depend on fewer variables than the map f, thus one can apply an induction argument.

A C^1-parametrization of $\mathcal{C}$

We will improve, or rather say refine, the C^0-parametrization to a uniform C^1- parametrization. In general, this requires an application of the inverse function theorem, but in this example, it immediately clear how to invert the function. Let us focus on the map $f_{+,t}$, the operation that should be applied to $f_{-,t}$ is completely analogous. Recall that $f_{+,t} : U_{+,t} \to (-1,1)$ is given by $f_{+,t}(x) = t/x$. We decompose U_+ into three parts, depending on the first order derivative of f_+ with respect to x.

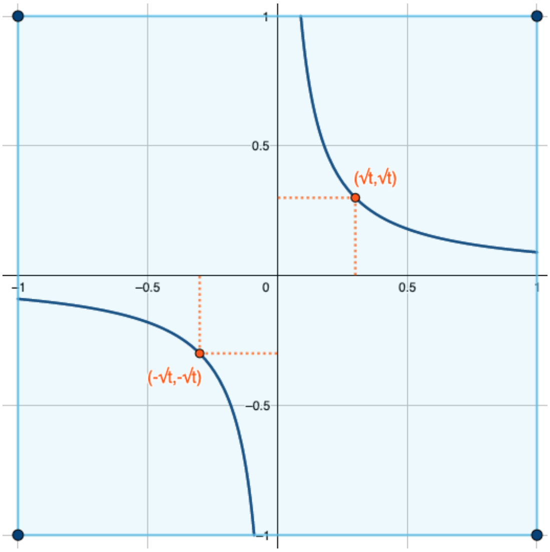

Figure 3.2: A C^1-parametrization of $\mathcal{C}$.[2]

1. We have that

$$\left|\frac{\partial}{\partial x} f_+(t,x)\right| = \frac{t}{x^2},$$

therefore

$$\left|\frac{\partial}{\partial x} f_+(t,x)\right| < 1 \iff \sqrt{t} < x < 1.$$

We then uniformly C^1-parametrize this part of the curve by setting

$$\Phi_1 : T \times (0,1) \to T \times (\sqrt{t},1) : (t,x) \mapsto (t, \sqrt{t} + (1-\sqrt{t})x)$$

and

$$f_1(t,x) = (\Phi_1(t,x), (f \circ \Phi_1)(t,x)).$$

[2]Copyright © International GeoGebra Institute, 2021

2. We have that
$$\left|\frac{\partial}{\partial x} f_+(t,x)\right| > 1 \iff t < x < \sqrt{t}.$$
We will invert $f_{+,t}$ on this domain. In this case, we end up in the previous case, namely we have that $x = t/y$, where $\sqrt{t} < y < 1$. This is also clear on Figure 3.2. Therefore, we can C^1-parametrize this part of the curve by f_2, which is obtained by composing f_1 with the map $(t,x,y) \mapsto (t,y,x)$.

3. Finally, we have to consider the case
$$\left|\frac{\partial}{\partial x} f_+(t,x)\right| = 1.$$
This occurs exactly once at $x = \sqrt{t}$. This point can be parametrized by a family of constant functions:
$$f_3 : T \times (0,1) \to T \times (-1,1)^2 : (t,x) \mapsto (t, \sqrt{t}, \sqrt{t}).$$

The collection $\{f_1, f_2, f_3\}$ is a uniform C^1-parametrization of $\mathcal{C} \cap (T \times (0,1)^2)$, i.e., the upper right hand side in Figure 3.2. Thus in total, we will end up with 6 charts. In general, due to the cell decomposition theorem, we know that we can decompose the domain into finitely many pieces, each of which belongs to one of the three cases above. To apply the inverse function theorem, we have to further decompose that part such that the conditions (in particular injectivity) are satisfied. This eventually boils down to the Monotonicity Theorem (Theorem 1.3.2). This method is more explicit in the proof of the pre-parametrization theorem in Section 3.4.

Higher order C^r-parametrizations of $\mathcal{C}$

Using the C^1-parametrization, we will now construct a uniform C^2- parametrization. To do so, we will use the power substitution given by
$$P_2 : T \times (0,1) \to T \times (0,1) : (t,x) \mapsto (t, x^2),$$
which is a family version of the power substitution P_2 in Proposition 2.5.3. In this simple example, the maps f_1 and f_2 of the C^1-parametrization constructed above already satisfy the assumptions of Proposition 2.5.3. In general, one should further decompose the domain such that the conditions of the following lemma (or suitable analogues) are satisfied. It is [Yom15, Lemma 4.1], where we have used the C^r-norm in the formulation instead.

Lemma 3.2.3. *Suppose that $f : (0,1) \to (0,1)$ with $|f|_{l-1} \leq 1$ for some $l \geq 2$. Suppose moreover that $f^{(l)}$ is positive and decreasing on $(0,1)$. Then $f^{(l)}(x) \leq \frac{l}{x}$.*

To study the upper bounds on the derivatives of this substitution, we will use the results of Chapter 2. We have that

$$f_1 : T \times (0,1) \to C_t : (t,x) \mapsto \frac{t}{\sqrt{t} + (1-\sqrt{t})x}$$

is weakly $(1,1,0)$-mild up to order $+\infty$. For its first order derivative (with respect to x) we may take the same mildness parameters to apply Proposition 2.5.3 by Remark 2.5.4. It follows that $f_1 \circ P_2$ is $(3,1,0)$-mild up to order 2, if you use that P_2 is $(1,2,0)$-mild up to order 2 and that $m = 1$ in our example.

Now we use a linear reparametrization (3.1.7) to obtain that the C^2-norm is at most 1. In this way, we obtain 4 charts that have C^2-norm at most 1. Since we apply this procedure to 4 of the 6 charts of the C^1-parametrization, our C^2-parametrization consists of 18 charts.

Remark 3.2.4. Below Lemma 3.1.6, we create maps that slightly overlap. That is why we slightly enlarge A there. One could also work in a different way. In this example, define the maps L_i for $i = 0,1,2$ by

$$L_i : T \times (0,1) \to \mathrm{Im}(L_i) \subset P_\ell^{-1}(T \times (0,1)) : (t,x) \mapsto (t, i+x).$$

In that case there is no overlap, but the points $(t,1)$ and $(t,2)$ are not in the image of any of the L_i. Therefore, one will obtain 5 charts for 4 of the charts of the C^1-parametrization and thus end up with 22 charts. The images of these charts can be seen in Figure 3.3.

Finally, we explain how to proceed to higher order parametrizations, that is how to construct a uniform C^{r+1}-parametrization, starting with a C^r-parametrization with $r \geq 1$. Let f be a chart of the uniform C^r-parametrization of $\mathcal{C}$. Therefore $|f|_r \leq 1$ and thus f is $(1,1,0)$-mild up to order r.

Next, one further decomposes the domain of f, if necessary, such that for each $t \in T$ the function $f_t^{(r+1)}$ does not change signs. In the example here, this is not necessary, in general one uses the cell decomposition theorem. This implies that one can use a suitable analogue of Lemma 3.2.3. Consequently, one uses the power substitution P_2. We now have that the family of functions $f \circ P_2$ is $(3,1,0)$-mild up to order $r+1$. Finally, one uses a linear reparametrization to obtain that $|f|_{r+1} \leq 1$.

In this way we see that the amount of charts N that is obtained via this inductive method depends on r as follows. In our example, except for the two points

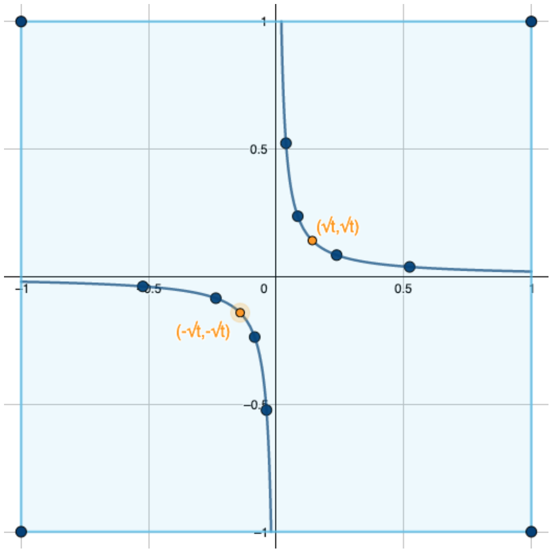

Figure 3.3: A C^2-parametrization of $\mathcal{C}$.[3]

$(-\sqrt{t}, \sqrt{t})$ and $(\sqrt{t}, \sqrt{t})$ that are parametrized by a family of constant functions, for each chart, one will construct 4 in the next stage. Therefore $N = 2 + 4^r$. In the next section, we will show that we can construct a C^r-parametrization of size $2 + 4(r+1)$ (for $r \geq 2$), see Example 3.3.10.

3.3 C^r-parametrization of power-subanalytic sets

One of the main results of [CPW20] is the following C^r-parametrization theorem for power-subanalytic sets (see Section 1.2 for the definition of power-subanalytic).

Theorem 3.3.1 ([CPW20, Theorem 2.1.3])**.** *Let n, k be positive integers and m be a nonnegative integer with $m \leq n$. Let $\mathcal{X}$ be a power-subanalytic family of m-dimensional subsets of $[-1,1]^n$, where T is some power-subanalytic subset of $\mathbb{R}^k$. Then there exist positive numbers c and d, depending only on the family $\mathcal{X}$, such that for each positive integer r, and for each $t \in T$, there exist analytic maps*

$$\phi_{r,i,t} : (0,1)^m \to X_t$$

for $i = 1, \ldots, cr^d$, whose C^r-norms are bounded by 1 and whose ranges cover X_t. Moreover, for each i and r, $\{\phi_{i,r,t} \mid t \in T\}$ is a power-subanalytic family of maps.

Remark 3.3.2. Recall from Section 1.2 that $\mathcal{X}$ is power-subanalytic if it is $\mathcal{L}_{\mathcal{F}}^K$-definable, where we assume that $\mathcal{F}$ and K are fixed. More precisely, in the theorem above, if $\mathcal{X}$ is $\mathcal{L}_{\mathcal{F}}^K$-definable, then the maps $\phi_{r,i,t}$ are $\mathcal{L}_{\mathcal{F}}^K$-definable for the same $\mathcal{F}$ and K.

In [VH21b], I show that their construction yields $d = m^3$. In this section, we will derive the C^r-parametrization theorem, using the material of Chapter 2 and given a so-called "pre-parametrization". The pre-parametrization allows us to just use one power substitution to construct a C^r-parametrization, instead of the $r-1$ power substitutions in the construction of the previous section (for $r \geq 2$). This power substitution will be a good choice of exponents N_i in Proposition 2.5.3. In order to state the pre-parametrization theorem, we first need to give the following definition.

Definition 3.3.3. Call a family of functions $f : U \subset T \times (0,1)^m \to \mathbb{R}$ **prepared in x** if it is of the form

$$f(t,x) = b_j(t,x)F(b(t,x)),$$

where b_j is a component function of $b : U \to \mathbb{R}^N$ (for some $N \in \mathbb{N}$), which has bounded range and each component function of b is of the form

$$a(t)x^\mu$$

for some $\mu \in \mathbb{R}^m$ and power-subanalytic function a, and where F is an analytic and non-vanishing function on an open neighborhood of $\overline{\mathrm{Im}(b)}$. We call b the **associated bounded monomial map** of f. A map $f : U \subset T \times (0,1)^m \to \mathbb{R}^n$ is prepared in x if all of its component functions are.

In particular, b is also a family of functions. Recall that in that case $|b|_r \leq 1$ implies that for all $t \in T : |b_t|_r \leq 1$. We can now state the pre-parametrization theorem. The version below is from my work [VH21b], it is a small modification of the original statement in [CPW20, Theorem 4.3.1].

Theorem 3.3.4 ([VH21b, Theorem 3.11])**.** *Suppose that $\mathcal{X}$ is the graph of a family of functions $\varphi : U \subset T \times (0,1)^m \to [-1,1]^n$. Then there exist finitely many power-subanalytic maps*

$$f_l : C_l \to \mathcal{X}$$

such that:

1. *for each* $t \in T : \bigcup_l Im(f_{l,t}) = X_t$*;*
2. *for each* l*,* C_l *is an open cell in* $T_l \times (0,1)^m$*, where* $T_l \subset T$*;*
3. *for each* l*,* f_l *is prepared in* x *and its associated bounded monomial map has bounded* C^1*-norm;*
4. *for each* l*, the walls of* C_l *are prepared in* x *and their associated bounded monomial map have bounded* C^1*-norm.*

The proof of this theorem will be given in the next section. It tells us that $\mathcal{X}$ can be uniformly parametrized by functions that are prepared in x whose associated bounded monomial maps have bounded C^1-norm. By the form of these functions, this is some way of "reducing to monomials".

We will now start working towards a proof of Theorem 3.3.1. First, we analyse the mildness of functions that are prepared in x.

Lemma 3.3.5. *Suppose* $f : U \subset T \times (0,1)^m \to \mathbb{R}$ *is prepared in* x*. Then there exist* $A, B > 0$ *such that* f *is weakly* $(A, B, 0)$*-mild up to order* $+\infty$*.*

Proof. Since f is prepared in x, f_t is given by

$$f_t(x) = b_{j,t}(x)F(b_t(x))$$

as in Definition 3.3.3. By Proposition 2.4.4, it suffices to show that b_t and $F \circ b_t$ are weakly $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$ that do no depend on t. Because F is analytic and non-vanishing on an open neighborhood of $\overline{\mathrm{Im}(b)}$, it is $(A, B, 0)$-mild for some $A, B > 0$ (see Remark 2.1.4). By Proposition 2.4.5 it follows that $F \circ b_t$ is weakly $(A, B, 0)$-mild for some $A, B > 0$ independent of t if b_t is so.

Since b has bounded range, there exists some B' such that $|b|_0 \leq B'$. The proof of Lemma 2.4.2 shows that b_t is weakly $(A', B', 0)$-mild up to order $+\infty$ for all $t \in T$, where A' is the maximum of 1 and the largest absolute value of the powers of x in the component functions of b. ☐

By the chain rule and Proposition 2.4.4, one immediately deduces the following corollary.

Corollary 3.3.6. Suppose $f : U \subset T \times (0,1)^m \to \mathbb{R}$ is prepared in x and that its associated bounded monomial map has bounded C^1-norm. Then there exist $A, B > 0$ such that for all $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$ and $t \in T$ the function $f_t^{(\beta)}$ is weakly $(A, B, 0)$-mild up to order $+\infty$.

We conclude that the functions from the pre-parametrization theorem "uniformly" satisfy the conditions of Proposition 2.5.3. As stated before, one has to make a good choice of powers N_i in this proposition in order to deal with a problem that we will make clear with the following example.

Example 3.3.7. Consider the function $f : U \subset (0,1)^2 \to \mathbb{R}$ given by

$$f(x,y) = \frac{y^2}{\sqrt{x}},$$

where U is given by

$$\begin{cases} 0 < x < 1, \\ 0 < y < x^{3/2}. \end{cases}$$

Clearly, f is prepared in x and so are the walls of U. Moreover, they have bounded C^1-norm. Now let P_r be the map $(0,1)^2 \to (0,1)^2$ given by $(x_1, x_2) \mapsto (x_1^r, x_2^r)$ as in Proposition 2.5.3. By Proposition 2.5.6, we know that $f \circ P_r : P_r^{-1}(U) \to \mathbb{R}$ is $(2r, 1, 0)$-mild up to order r.

However, the function $x \mapsto x^{3/2}$ is not (A, B, C)-mild up to order r on (0,1) for any $A, B > 0$ and $C \geq 0$ for $r > 1$ on $P_r^{-1}(U) = U$. Therefore the linear map $(0,1)^2 \to P_r^{-1}(U)$ is not (A, B, C)-mild up to order r for any $A, B > 0$ and $C \geq 0$. This problem can be solved as follows. Consider the the map $P : (0,1)^2 \to (0,1)^2$ given by $(x_1, x_2) \mapsto (x_1^r, x_2)$. Then $P^{-1}(U)$ is given by

$$\begin{cases} 0 < x^r < 1 \\ 0 < y < (x^r)^{3/2} \end{cases} \iff \begin{cases} 0 < x < 1 \\ 0 < y < x^{3r/2}. \end{cases}$$

We see that $P^{-1}(U)$ is a cell and that its walls are also $(2r, 1, 0)$-mild up to order r. Therefore, instead of $f \circ P_r$, we will consider $f \circ P_r \circ P$. Note that $P_r \circ P : (0,1)^2 \to (0,1)^2$ is given by $(x_1, x_2) \mapsto (x_1^{r^2}, x_2^r)$, thus $f \circ P_r \circ P$ is $(2r^2, 1, 0)$-mild up to order r. The linear map $(0,1)^2 \to U$ is $(2r, 1, 0)$-mild up to order r.

We now generalize this example to the functions given by Theorem 3.3.4. The proof is long and technical, but it just proves the above example in general using the results of Chapter 2.

Proposition 3.3.8 ([VH21b, Proposition 3.10])**.** *Suppose $f : C \subset T \times (0,1)^m \to \mathbb{R}$ is a function defined on a cell C, where C_t is an open cell for each $t \in T$, such that f and all walls of C are prepared in x and all of their associated bounded monomial maps have bounded C^1-norm. Let P be the power map defined by*

$$P : T \times (0,1)^m \to T \times (0,1)^m : (t, x_1, \ldots, x_m) \mapsto (t, x_1^{r^m}, x_2^{r^{m-1}}, \ldots, x_m^r).$$

Then there exist $A, B > 0$ such that $f \circ P : P^{-1}(C) \to \mathbb{R}$ and all walls of the cell $P^{-1}(C)$ are $(Ar^m, B, 0)$-mild up to order r.

Proof. The claim on $f \circ P$ follows immediately by Proposition 2.5.3. We do not have to consider the walls of the t variables, since they are families of constant functions. Denote α_i and β_i for the walls of the cell C bounding the variable x_i from below and from above respectively. Then we extend these walls trivially to functions $C \to (0,1)$, i.e., $\alpha_i(t, x_1, \ldots, x_m) = \alpha_i(t, x_1, \ldots, x_{i-1})$. For $i = 1, \ldots, m$, we have that $P^{-1}(C)$ is given by inequalities

$$\alpha_i(P(t,x)) < x_i^{r^{m-i+1}} < \beta_i(P(t,x)),$$

which we rewrite as

$$\sqrt[r^{m-i+1}]{\alpha_i(P(t,x))} < x_i < \sqrt[r^{m-i+1}]{\beta_i(P(t,x))}$$

to obtain the proper form of a cell. In particular, this shows that it actually is a cell. We have to show that these walls are $(Ar^m, B, 0)$-mild up to order r for some $A, B > 0$. It suffices to show the claim for the wall of the cell $P^{-1}(C)$ bounding x_i form below.

Since α_i is prepared in x, it is of the form

$$\alpha_i(t,x) = b_j(t,x)F(b(t,x))$$

as in Definition 3.3.3. Since F is analytic and non-vanishing on an open neighborhood of $\overline{\mathrm{Im}(b)}$, there exists some $S \in (0,1)$ such that $\mathrm{Im}(F) \subset (S, 1/S)$ (we may suppose that F is positive). Moreover, it is $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$ (see Example 2.1.4). On $(S, 1/S)$ the function $x^{r^{m-i+1}}$ is $(1/S, 1/S, 0)$-mild up to order $+\infty$ (see Example 2.1.4). Finally, since $b(t,x)$ has bounded C^1-norm, it is $(Ar^m, B, 0)$-mild up to order r by Corollary 3.3.6, and Proposition 2.5.3 for some $A, B > 0$. Therefore

$$\sqrt[r^{m-i+1}]{F(b(P(t,x)))}$$

is $(Ar^m, B, 0)$-mild up to order r by Proposition 2.2.5 for some $A, B > 0$. Thus, to show that $\sqrt[r^{m-i+1}]{\alpha_i(P(t,x))}$ is $(Ar^m, B, 0)$-mild up to order r for some $A, B > 0$, it suffices to show that

$$\sqrt[r^{m-i+1}]{b_j(P(t,x))}$$

is $(Ar^m, B, 0)$-mild up to order r for some $A, B > 0$. By Definition 3.3.3, $b_j(t,x)$ is of the form

$$b_j(t,x) = a(t)x^{\mu}$$

for some power-subanalytic function a and $\mu \in \mathbb{R}^m$, and has bounded C^1-norm. We write

$$\sqrt[r^{m-i+1}]{b_j(P(t,x))} = \tilde{a}(t)x^{\mu'},$$

where $\tilde{a} = \sqrt[r^{m-i+1}]{a}$ and $\mu'_\ell = r^{i-\ell}\mu_\ell$. Note that $\mu_i = \mu_{i+1} = \ldots = \mu_m = 0$, since the wall α_i does not depend on $x_i, \ldots, x_m$. Now let $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$ and let $(t,x) \in P^{-1}(C) \subset T \times (0,1)^m$ be fixed but arbitrary. We will express the derivative of $\sqrt[r^{m-i+1}]{b_j(P(t,x))}$ with respect to α in terms of a first order derivative of b_j in $P(t,x) \in C$, from which we know that it is bounded on C. The idea of the proof is similar to the proof of Proposition 2.5.10 and a slightly more explicit version of [VH21b, Proposition 3.9].

We have that

$$\left| \left(\sqrt[r^{m-i+1}]{b_j(P(t,x))} \right)^{(\alpha)} \right| = \left| \left(\tilde{a}(t) x^{\mu'} \right)^{(\alpha)} \right| \leq |c(\alpha, \mu')| \left| \tilde{a}(t) x^{\mu' - \alpha} \right|, \tag{3.1}$$

where $c(\alpha, \mu')$ is a constant that depends on α and μ' (as in Lemma 2.4.2). Now let I be such that $x_I = \min\{x_\ell \mid \alpha_\ell \neq 0\}$. Then we may assume $\mu'_I \neq 0$, since if not, the derivative with respect to α is zero and therefore can trivially be bounded in absolute value by the upper bound we will compute now. For the same reason, we may also suppose that $1 \leq I < i$. We now further bound 3.1:

$$\begin{aligned} |c(\alpha, \mu')| \left| \tilde{a}(t) x^{\mu' - \alpha} \right| &\leq |c(\alpha, \mu')| \left| \tilde{a}(t) x^{\mu'} x_I^{-|\alpha|} \right| \\ &\leq |c(\alpha, \mu')| \left| \tilde{a}(t) x^{\mu'} x_I^{-r} \right| \\ &\leq |c(\alpha, \mu')| \left| \sqrt[r^{m-i+1}]{b_j(P(t,x)) x_I^{-r^{m-i+2}}} \right| \\ &\leq |c(\alpha, \mu')| \left| \sqrt[r^{m-i+1}]{b_j(P(t,x)) x_I^{-r^{m-I+1}}} \right|. \end{aligned} \tag{3.2}$$

Now, write $b_j(P(t,x)) = a(t)x^\nu$ with $\nu_\ell = r^{m-\ell+1}\mu_\ell$. Then we have that

$$b_j(P(t,x)) x_I^{-r^{m-I+1}} = a(t) x^\nu x_I^{-r^{m-I+1}}.$$

The power of x_I is: $r^{m-I+1}\mu_I - r^{m-I+1} = r^{m-I+1}(\mu_I - 1)$. It follows that

$$b_j(P(t,x)) x_I^{-r^{m-I+1}} = \mu_I^{-1} \left(\frac{\partial}{\partial x_I} b_j \right) (P(t,x)).$$

Since $P(t,x) \in C$ and the C^1-norm of b is bounded on C, this can be bounded independent of t, x and I. Clearly, the r^{m-i+1}-th root of this upper bound can be further bounded such that it does not depend on r, i and m. Finally, similar to the proof of Lemma 2.4.2, it is clear that

$$|c(\alpha, \mu')| \leq (Ar^m)^{|\alpha|} \, |\alpha|!$$

for some A that only depends on μ'.

Putting everything together, we conclude that there exist $A, B > 0$ such that

$$\left| \left(\sqrt[r^{m-i+1}]{b_j(P(t,x))} \right)^{(\alpha)} \right| \leq B(Ar^m)^{|\alpha|}\,|\alpha|!$$

for all $\alpha \in \mathbb{N}^m$ with $|\alpha| \leq r$, i.e., this function is $(Ar^m, B, 0)$-mild up to order r. Together with the same result on the other factor (containing the root of the analytic function F) of the wall bounding x_i from below, this finishes the proof by Proposition 2.2.2. □

Remark 3.3.9. Note that P can be seen as a composition of power maps of the form

$$(x_1, \ldots, x_m) \mapsto (x_1^r, \ldots, x_i^r, x_{i+1}, \ldots, x_m)$$

as explained in Example 3.3.7. More precisely, after the power substitution P_r, we obtain a domain $P_r^{-1}(C)$ that is of the same form as C. Then we can proceed by induction, so we then apply the power substitution defined by

$$(x_1, \ldots, x_m) \mapsto (x_1^r, \ldots, x_{m-1}^r, x_m),$$

which now ensures that the walls α_m and β_m of the cell will be $(Ar^2, B, 0)$-mild up to order r. This iteration of power maps yields the power map P in the proposition, and will ensure that all the walls will become $(Ar^m, B, 0)$-mild up to order r. In fact, from this point of view, one can deduce another proof of this proposition. In the proof above, we considered P as a single power substitution, not as a composition.

This proposition, together with the pre-parametrization theorem 3.3.4, yields the uniform C^r-parametrization theorem (Theorem 3.3.1) for power-subanalytic sets. Before we show this, let us return to the example of the previous section to show how this works and to compare the amount of charts that are obtained via this construction.

Example 3.3.10. Let $\mathcal{C}$ be the family of algebraic curves of Section 3.2. We have deduced there that it suffices to reparametrize the graph of the family of functions $f : U \subset (0,1)^2 \to \mathbb{R}$ given by $f(t,x) = t/x$, where $U = \{(t,x) \in (0,1)^2 \mid \sqrt{t} < x < 1\}$. In particular, f could be a potential outcome of the pre-parametrization theorem.

With the notation of Proposition 3.3.8, we have that $f \circ P : P^{-1}(U) \to \mathbb{R}$ is $(Ar, B, 0)$-mild up to order r for some $A, B > 0$. By Proposition 2.5.6, it follows that $f \circ P$ is $(r, 1, 0)$-mild up to order r. The cell $P^{-1}(U)$ is given by

$$\begin{cases} 0 < t < 1, \\ t^{1/(2r)} < x < 1. \end{cases}$$

Let $\Phi : (0,1)^2 \to P^{-1}(U)$ be defined by

$$(t,x) \mapsto (t, t^{1/(2r)} + (1 - t^{1/(2r)})x).$$

Therefore, a C^r-parametrization of this part of the family $\mathcal{C}$ is given by the map

$$(0,1)^2 \to \mathcal{C} : (t,x) \mapsto (\Phi(t,x), (f \circ P \circ \Phi)(t,x)).$$

In this case, one can compute directly that this map is $(r,1,0)$-mild. Finally, by a linear reparametrization (3.1.7), we obtain $r+1$ C^r-charts for this part of the family $\mathcal{C}$. It follows that we obtain $2 + 4(r+1)$ charts in total (for $r \geq 2$), which is significantly better than the $2 + 4^r$ charts obtained by the construction of Section 3.2.

Finally, we conclude this section with a proof of Theorem 3.3.1.

Proof. By the cell decomposition theorem (Theorem 1.3.6), we may suppose that $\mathcal{X}$ is the graph of a power-subanalytic function $\varphi : T \times (0,1)^m \to [-1,1]^{n-m}$ such that $\varphi(t,x) \in X_t$ for each $(t,x) \in T \times (0,1)^m$.

By the pre-parametrization theorem (Theorem 3.3.4), we obtain finitely many power-subanalytic charts $f_l : C_l \to \mathcal{X}$ that satisfy the conditions of Proposition 3.3.8. Thus, applying this proposition to each of these charts, we obtain finitely many charts, which we will also denote by $f_l : C_l \to \mathcal{X}$, satisfying the following properties:

1. for each $t \in T : \bigcup_l \mathrm{Im}(f_{l,t}) = X_t$;
2. for each l, C_l is an open cell in $T_l \times (0,1)^m$, where $T_l \subset T$;
3. there exist $A, B > 0$ such that for each l, f_l is $(Ar^m, B, 0)$-mild up to order r, moreover the walls of C_l are $(Ar^m, B, 0)$-mild up to order r.

Since $f_l : C_l \to \mathcal{X} \subset T \times (0,1)^n$ and $C_l \subset T_l \times (0,1)^m$, we may suppose that $B = 1$, after possibly enlarging A. Finally, we have to study the mildness of the linear map $\Phi : T_l \times (0,1)^m \to C_l$.

Denote α_m and β_m for the walls of C_l bounding the variable x_m from below and above respectively. Denote $\pi_{<m} : T_l \times (0,1)^m \to T_l \times (0,1)^{m-1}$ for the projection on the first $m-1$ coordinates of $(0,1)^m$. Denote $x_{<m} = (x_1, \ldots, x_{m-1})$. Let $\Phi_m : \pi_{<m}(C_l) \times (0,1) \to C_l$ be given by

$$\Phi_m(t,x) = (t, x_{<m}, \alpha_m(t, x_{<m}) + (\beta_m - \alpha_m)(t, x_{<m})x_m).$$

Since α_m and β_m are $(Ar^m, 1, 0)$-mild up to order r for some $A > 0$, it follows by propositions 2.2.1 and 2.2.2 that Φ_m is $(Ar^m, 1, 0)$-mild up to order r for some $A > 0$. Note that for the addition, we use that the image is contained in $T_l \times (0,1)^m$, such that B can be taken equal to 1, after possibly enlarging A.

By Proposition 2.2.5, we have that $f_l \circ \Phi_m$ is $(Ar^{2m}, 1, 0)$-mild up to order r for some $A > 0$. Repeating this proces, we obtain the map $\Phi = \Phi_1 \circ \ldots \circ \Phi_m : T_l \times (0,1)^m \to C_l$ and we have that $f \circ \Phi_l$ is $(Ar^{m^2}, 1, 0)$-mild up to order r for some $A > 0$. (Note that Φ_1 is always $(1, 1, 0)$-mild up to order $+\infty$.)

After we have treated every f_l in this way, and after applying a linear reparametrization (3.1.7) to each of them, this yields a C^r-parametrization of $\mathcal{X}$ with cr^{m^3} charts for some $c > 0$. Finally, since each f_l was power-subanalytic, it follows by this construction that all charts are power-subanalytic. Indeed, we have only composed with power functions with natural exponents and made linear combinations of these functions. By the form of functions that are prepared in x, each family member is indeed analytic. This concludes the proof. $\square$

3.4 Pre-parametrization

This section in entirely devoted to a proof of the pre-parametrization theorem (Theorem 3.3.4). Recall that $\mathcal{X}$ is an m-dimensional family of power-subanalytic sets in $[-1, 1]^n$. The proof uses induction on m. By the cell decomposition theorem, we may suppose $\mathcal{X}$ is the union of images of finitely many families of maps $f_l : T \times (0,1)^m \to \mathcal{X}$. If $m = 0$, there is nothing to show. In that case, X_t is a finite union of points and there is a uniform upper bound on the amount of points (see Lemma 1.3.3).

Suppose the theorem holds for families of dimension at most $m - 1$ for some $m \in \mathbb{N}$. We will (possibly) use the cell decomposition theorem multiple times. Each time, by induction, it suffices to consider the open cells of the partition.

Convention. In this section, a cell $C \subset T \times (0,1)^m$ is open if it is a family of open cells in $(0,1)^m$.

Throughout the proof, divided into three parts below, we will focus on one map $f_l : T \times (0,1)^m \to \mathcal{X}$. Therefore, we will drop the index l.

Preparation with center zero

To ensure that f is prepared in x, we will use the following *preparation theorem.*

Theorem 3.4.1 ([Mil06, Main theorem]). *Suppose $f : U \subset T \times (0,1)^m \to [-1,1]^n$ is power-subanalytic. Then U can be partitioned into finitely many cells C_i such that the restriction of f to each C_i is prepared with some center θ_i.*

We first have to explain what it means to be prepared on a cell with some center. Since we are only interested in the open cells of this partition, we will only explain it for open cells, as explained in the introduction of this section. This avoids some unnecessary extra definitions or terminology from [Mil06].

Definition 3.4.2 ([CPW20, Definition 4.4.1]). A function $f : C \subset T \times (0,1)^{m+1} \to [-1,1]$, where C is an open cell, is **prepared with center $\boldsymbol{\theta}$** if f is of the form

$$f(t,x,x_{m+1}) = b_j(t,x,x_{m+1})F(b(t,x,x_{m+1})),$$

where b_j is a component function of $b : C \to \mathbb{R}^N$ (for some $N \in \mathbb{N}$), which has bounded range, each component function of b is of the form

$$a(t,x)\left|x_{m+1} - \theta(t,x)\right|^r$$

for some power-subanalytic a and $r \in \mathbb{R}$, F is an analytic and non-vanishing function on an open neighborhood of $\overline{\operatorname{Im}(b)}$ and θ is either identically zero or a continuous power-subanalytic function, whose graph is disjoint of $\bar{C}$ or contained in $\bar{C} \setminus C$, such that there exists an $\epsilon \in (0,1)$ such that for all $(t,x,x_{m+1}) \in C : \epsilon x_{m+1} < \theta(t,x) < (1/\epsilon)x_{m+1}$.

We call b the associated bounded range map of f. A map $f : C \to [-1,1]^n$ is prepared with center θ if all of its component functions are (with the same θ).

The main idea of the proof of the pre-parametrization theorem is to redefine f such that $\theta = 0$, and then bound the partial derivative with respect to x_m. Consequently, we iteratively use Theorem 3.4.1 to achieve that f is prepared in x. We start with showing that we may assume that $\theta = 0$.

After applying Theorem 3.4.1 to $f : T \times (0,1)^m \to \mathcal{X}$, we obtain finitely many maps $C_i \subset T \times (0,1)^m \to \mathcal{X}$ that are prepared with center θ_i. Now, by abuse of notation, fix one such map $f : C \subset T \times (0,1)^m \to \mathcal{X}$ that is prepared with center θ and suppose that θ is not identically zero.

Since θ is continuous, and f is prepared with center θ on C, we may suppose that either $x_m < \theta(t, x_{<m})$ for all $(t,x) \in C$, where $x_{<m}$ denotes $(x_1, \ldots, x_{m-1})$,

or $\theta(t, x_{<m}) < x_m$ for all $(t, x) \in C$. Suppose that we are in the first case, the other one is similar. Consider the set

$$\tilde{C} = \{(t, x) \in T \times (0, 1)^m \mid (t, x_{<m}, Sx_m + \theta(t, x)) \in C\},$$

where we want to find a suitable choice of S such that $\tilde{C}$ is a cell in $T \times (0, 1)^m$. Denote α_m and β_m for the walls of C bounding the variable x_m from below and above respectively. Then $(t, x) \in \tilde{C}$ if and only if

$$\alpha_m(t, x_{<m}) < Sx_m + \theta(t, x_{<m}) < \beta_m(t, x_{<m}).$$

Since we supposed that $x_m < \theta(t, x_{<m})$ for all $(t, x) \in C$, it follows that $\beta_m(t, x_{<m}) - \theta(t, x_{<m}) \leq 0$. Therefore, we should pick some negative S, and we rewrite the inequality above in the form of a cell as follows:

$$(1/S)(\beta_m(t, x_{<m}) - \theta(t, x_{<m})) < x_m < (1/S)(\alpha_m(t, x_{<m}) - \theta(t, x_{<m})).$$

We should now determine S such that $(1/S)(\alpha_m(t, x_{<m}) - \theta(t, x_{<m})) < 1$. By Definition 3.4.2, there exists some $\epsilon \in (0, 1)$ such that for all $(t, x) \in C : \epsilon x_m < \theta(t, x_{<m}) < (1/\epsilon)x_m$, which implies that

$$0 < \epsilon\alpha_m(t, x_{<m}) < \theta(t, x_{<m}) < (1/\epsilon)\beta_m(t, x_{<m}) < (1/\epsilon).$$

It follows that $|\alpha(t, x_{<m}) - \theta(t, x_{<m})| < 1/\epsilon$. Therefore, we pick $S = -\epsilon$.

The composition of $f : C \to \mathcal{X}$ with the map $\tilde{C} \to C$ given by

$$(t, x) \mapsto (t, x_{<m}, Sx_m + \theta(t, x_{<m})),$$

is now prepared with center zero.

To summarize this part of the proof, we may replace each $f : T \times (0, 1)^m \to \mathcal{X}$ by finitely many maps $C \subset T \times (0, 1)^m$, such that the union of their images covers $\mathrm{Im}(f)$. Moreover, each of these maps is either defined on a lower dimensional cell, thus can be treated by induction, or if it is defined on an open cell, then it is prepared with center zero.

Bounding the first order derivative with respect to x_m

We only consider $f : C \subset T \times (0, 1)^m \to \mathcal{X}$ that are defined on an open cell and are prepared with center zero, as obtained at the end of the previous part of the proof.

Let b be the associated bounded range map of f, thus we have that

$$f(t, x) = b_j(t, x)F(b(t, x))$$

where b_j is a component function of b, and b and F are as in Definition 3.4.2. Adjusting b and F if necessary, we may assume that $\mathrm{Im}(b) \subset [-1,1]^N$. Using cell decomposition, we may suppose that b is continuously differentiable and that there is some i such that $|\partial b_i/\partial x_m| \geq |\partial b_{i'}/\partial x_m|$ for all component functions $b_{i'}$ of b. Then further decompose such that $|\partial b_i/\partial x_m| \leq 1$, in that case we are done, or $|\partial b_i/\partial x_m| > 1$ on C.

Suppose $|\partial b_i/\partial x_m| > 1$ on C. In that case, using o-minimality, further adjust b_i and F such that $\mathrm{Im}(b_i) \subset (0,1)$. Using cell decomposition once more, we may assume that for fixed $(t, x_1, \ldots, x_{m-1})$, the map

$$x_m \mapsto b_i(t, x_1, \ldots, x_m)$$

is injective. Now let $\tilde{C}$ be the image of C under the map

$$\phi : C \to \mathrm{Im}(\phi) : (t,x) \mapsto (t, x_{<m}, b_i(t,x)).$$

It is again a cell in $T \times (0,1)^m$. It follows that $f \circ \phi^{-1} : \tilde{C} \to \mathcal{X}$ has the same image as f, is prepared with center 0 and moreover, $|\partial b_i/\partial x_m| \leq 1$. Note that $f \circ \phi^{-1}$ remains prepared with center 0 due to the specific form of b. With more details, write $(t, x_1, \ldots, x_{<m}, y)$ for coordinates in $\tilde{C}$. Then we have that

$$y = b_i(t,x) = a(t, x_{<m}) x_m^r$$

for some power-subanalytic function a and $r \in \mathbb{R}$ (see Definition 3.4.2). Therefore one can express x_m as a function of $t, x_1, \ldots, x_{m-1}$ and y in the same form.

Thus, the conclusion of this part of the proof is that we may assume that f is prepared with center 0 and that the associated bounded range map b of f satisfies $|\partial b_j/\partial x_m| \leq 1$ for all component functions b_j of b.

The induction argument

Let $f : C \subset T \times (0,1)^m \to [-1,1]^n$ be prepared on an open cell C with center 0 and suppose that its associated bounded range map b satisfies $|\partial b_j/\partial x_m| \leq 1$ for all component functions b_j of b, i.e., the outcome of the previous part of the proof. Each component function b_j of b is of the form

$$a_j(t, x_{<m}) x_m^{r_j} \tag{3.3}$$

for some power-subanalytic function a_j and $r_j \in \mathbb{R}$.

Denote by α_m and β_m the walls of C bounding the variable x_m from below and above respectively. Up to further partitioning, using the cell decomposition theorem, we may suppose α_m is either identically zero, or non-zero on C. If α_m is identically zero, this will force $r \geq 1$ in the computations below, since $|\partial b_j / \partial x_m| \leq 1$. Moreover, in that case α_m is prepared in x.

We will now construct a map $F : \pi_{<m}(C) \to [-1, 1]^M$ for some $M \in \mathbb{N}$ that is constructed out of the component functions b_j of b, where $\pi_{<m}(C)$ denotes the image of C under the projection map $T \times (0, 1)^m \to T \times (0, 1)^{m-1}$ onto the first $m - 1$ coordinates. Since the domain has dimension $m - 1$, the graph $\Gamma(F)$ of this function F is a family of $m - 1$ dimensional subsets of $[-1, 1]^{M+m-1}$, and therefore we can apply the induction hypothesis to it. Essentially, this parametrizes the domain $\pi_{<m}(C)$ and reparametrizes F. We will show that after this reparametrization, each component function b_j of b is prepared in x and their associated bounded monomial maps have bounded C^1-norm.

The function F has as component functions the walls α_m and β_m, and also the following functions for each component function b_j of b, depending on the exponent r_j of x_m (see 3.3):

1. If $r_j = 0$, it is just b_j;
2. If $r_j > 0$, it is
$$g_{\beta_m,j} : \pi_{<m}(C) \to [-1, 1] : (t, x_{<m}) \mapsto \lim_{x_m \to \beta_m(t, x_{<m})} b_j(t, x_{<m}, x_m),$$
3. if $r_j < 0$, it is
$$g_{\alpha_m,j} : \pi_{<m}(C) \to [-1, 1] : (t, x_{<m}) \mapsto \lim_{x_m \to \alpha_m(t, x_{<m})} b_j(t, x_{<m}, x_m).$$

Note that the limits are power-subanalytic functions (by general theory of o-minimality).

We now apply the induction hypothesis to the graph $\Gamma(F)$ of F. This yields finitely many maps $\varphi_i : D_i \to \Gamma(F)$ that are prepared in $x_{<m}$ and their associated bounded monomial maps have bounded C^1-norm. This also holds for the walls of D_i.

For each i, consider the cell
$$E_i = \{(t, x) \in D_i \times (0, 1) \mid (\varphi_i(t, x_{<m})_{<m}, x_m) \in C\}.$$

The walls bounding the variables $x_1, \ldots, x_{m-1}$ are clearly as desired, since these are just the walls of D_i. The condition on x_m is given by the inequalities
$$\alpha_m((\varphi_i(t, x_{<m})_{<m})) < x_m < \beta_m(\varphi_i(t, x_{<m})_{<m}).$$

Since $\varphi_i(t, x_{<m}) \in \Gamma(F)$ it follows that $\alpha_m(\varphi_i(t, x_{<m})_{<m})$ is prepared in x and its associated bounded monomial map has bounded C^1-norm. Indeed, just explicitly write down the image of φ_i:

$$\varphi_i(t, x_{<m}) = (\underbrace{t, \varphi_i(t, x_{<m})_1, \ldots, \varphi_i(t, x_{<m})_{m-1}}_{\in \pi_{<m}(C)},$$
$$\underbrace{\varphi_i(t, x_{<m})_m, \varphi_i(t, x_{<m})_{m+1}}_{\alpha_m \text{ and } \beta_m}, \ldots) \in \Gamma(F), \quad (3.4)$$

which indeed shows that $\alpha_m(\varphi_i(t, x_{<m})_{<m}) = \varphi_i(t, x_{<m})_m$ is prepared in x and its associated bounded monomial map has bounded C^1-norm. The result on $\beta_m(\varphi_i(t, x_{<m})_{<m})$ is analogous. Note that the "..." represent the functions we have added for each b_j. It follows immediately that if b_j had $r_j = 0$, then $b_j(\varphi_i(t, x_{<m})_{<m})$ is as desired.

Finally, it remains to show that the composition of f with the maps

$$\psi_i : E_i \to C : (t, x) \mapsto (\varphi_i(t, x_{<m})_{<m}, x_m)$$

are prepared in x and their associated bounded monomial map have bounded C^1-norm. It is sufficient to show this for the bounded range map b of f. The proof is rather technical and I decided that it is more clear to explain it using a slight simplification that actually does not harm the generality of the proof.

For simplicity, we will assume that all analytic and non-vanishing functions F from definitions 3.3.3 and 3.4.2 that arise from this construction, are identically 1. For instance, this implies $f = b$. Fix some component function b_j of b. Recall that it is of the form

$$a_j(t, x_{<m}) x_m^{r_j}$$

for some power-subanalytic function a_j and $r_j \in \mathbb{R}$ (see 3.3). We assume $r_j > 0$. If $r_j < 0$, one should replace β_m with α_m. We have already explained that if $r_j = 0$, the composition $b_j \circ \psi_i = b_j(\varphi_i(t, x_{<m})_{<m})$ is as desired.

Since $\beta_m(\varphi_i(t, x_{<m})_{<m})$ is prepared in x, it is of the form

$$a_{\beta_m}(t) x_{<m}^{\mu}$$

for some power-subanalytic function a_{β_m} and $\mu \in \mathbb{R}^{m-1}$. Now, we have that

$$\begin{aligned} g_{\beta_m, j}(\psi_i(t, x)_{<m}) &= \lim_{x_m \to \beta_m(\psi_i(t, x_m)_{<m})} b_j(\psi_i(t, x)) \\ &= a_j(\psi_i(t, x)_{<m}) \left(a_{\beta_m}(t) x_{<m}^{\mu}\right)^{r_j}, \end{aligned}$$

but $g_{\beta_m,j}(\psi_i(t,x_{<m})_{<m}) = g_{\beta_m,j}(\varphi_i(t,x_{<m})_{<m})$ is also prepared in x (see 3.4), hence it is of the form

$$c_{\beta_m}(t)x_{<m}^{\nu}$$

for some power-subanalytic function c_{β_m} and $\nu \in \mathbb{R}^{m-1}$. We conclude that

$$a_j(\psi_i(t,x)_{<m}) = \frac{c_{\beta_m}(t)x_{<m}^{\nu}}{a_{\beta_m}(t)^{r_j}x_{<m}^{r_j\mu}}, \tag{3.5}$$

which shows that $b_j \circ \psi_i$ is indeed prepared in x.

Example. Suppose $b : U \to \mathbb{R}$ is given on a cell C defined by

$$\begin{cases} 0 & < & x_1 & < & 1, \\ x_1^{1/2} & < & x_2 & < & 1, \end{cases}$$

that b is of the form $a(x_1)x_2^{-1}$ and $\lim_{x_2 \to x_1^{1/2}} = 1$. Then clearly, $a(x_1) = x_1^{1/2}$.

Finally, we have to show that $b_j \circ \psi_i$ has bounded C^1-norm. Since we already bounded $\partial b_j/\partial x_m$ and by the form of the map ψ_i, there is nothing to show anymore for the first order derivative with respect to x_m.

Let $\ell \in \{1,\ldots,m-1\}$. If $b_j \circ \psi_i$ does not depend on x_ℓ, there is nothing to show. Suppose it does depend on x_ℓ. Because of the form of $b_j \circ \psi_i$, partial differentiation with respect to x_ℓ is (up to a constant) the same as division by x_ℓ. By 3.5, we see that $b_j \circ \psi_i$ depends on x_ℓ if and only if $g_{\beta_m}(\psi_i(t,x_{<m})_{<m})$ does, $\beta_m(\psi_i(t,x_{<m})_{<m})$ does or both. Since they are prepared in x and have bounded C^1-norm, they can absorb the division by x_ℓ, if they depend on x_ℓ.

More precisely, suppose $g_{\beta_m}(\psi_i(t,x_{<m})_{<m})$ depends on x_ℓ, equivalently: $\nu_\ell \neq 0$. Then we see that $(\partial/\partial x_\ell)(b_j \circ \psi_i)$ is bounded since:

$$\begin{aligned} \left|\frac{1}{x_\ell}(b_j \circ \psi_i)(t,x)\right| &= \left|\frac{1}{x_\ell}c_{\beta_m}(t)x_{<m}^{\nu}\frac{x_m^{r_j}}{a_{\beta_m}(t)^{r_j}x_{<m}^{r_j\mu}}\right| \\ &\leq \left|\frac{1}{x_\ell}g_{\beta_m}(\psi_i(t,x_{<m})_{<m})\right|\left|\frac{\beta_m(\psi_i(t,x_{<m})_{<m})^{r_j}}{\beta_m(\psi_i(t,x_{<m})_{<m})^{r_j}}\right|. \end{aligned}$$

The first factor is bounded by our assumptions on $g_{\beta_m}(\psi_i(t,x_{<m})_{<m})$.

Finally, suppose $g_{\beta_m}(\psi_i(t,x_{<m})_{<m})$ does not depend on x_ℓ, then the wall $\beta_m(\psi_i(t,x_{<m})_{<m})$ does (if not $b_j \circ \psi_i$ does not depend on x_ℓ and we were done).

We see that $(\partial/\partial x_\ell)(b_j \circ \psi_i)$ is bounded since:

$$\left|\frac{1}{x_\ell}(b_j \circ \psi_i)(t,x)\right| = \left|\frac{c_{\beta_m}(t)x^{\nu}_{<m}}{a_{\beta_m}(t)x^{\mu}_{<m}}\frac{1}{x_\ell}\frac{x_m^{r_j}}{(a_{\beta_m}(t)x^{\mu}_{<m})^{r_j-1}}\right|$$

$$\leq \left|\frac{g_{\beta_m}(\psi_i(t,x_{<m})_{<m})}{\beta_m(\psi_i(t,x_{<m})_{<m})}\right|\left|\frac{1}{x_\ell}\frac{\beta_m(\psi_i(t,x_{<m})_{<m})^{r_j}}{\beta_m(\psi_i(t,x_{<m}))^{r_j-1}}\right|.$$

By our assumption on $\beta_m(\psi_i(t,x_{<m})_{<m})$, it follows that the second factor is bounded. The first factor is bounded since $(\partial/\partial x_m)(b_j)$ is bounded. Indeed, up to a constant, we have that $(\partial/\partial x_m)(b_j) = b_j/x_m$ since $r_j \neq 0$. Taking the limit for $x_m \to \beta_m(\psi_i(t,x_{<m})_{<m})$ yields that the first factor above is bounded.

Example. Let us also consider an example of the last part of the proof. It will show that it is crucial that $(\partial/\partial x_m)(b_j)$ has to be bounded. Consider the cell C defined by

$$\begin{cases} 0 & < & x_1 & < & 1, \\ x_1 & < & x_2 & < & 1, \\ x_1^2x_2^{-1} & < & x_3 & < & 1, \end{cases}$$

suppose that $b(x)$ is of the form $a(x_1,x_2)x_3^{-2}$ and that $g_{\alpha_3}(x_1,x_2) = x_2$. It follows that $a(x_1,x_2) = x_1^4x_2^{-1}$, thus $b(x) = x_1^4x_2^{-1}x_3^{-2}$. We see that

$$|\partial b/\partial x_2| = x_1^4x_2^{-2}x_3^{-2} = (\alpha_3(x_1,x_2)/x_3)^2 < 1.$$

However, we see that

$$|\partial b/\partial x_1| = 4x_1^3x_2^{-1}x_3^{-2} < 4x_1^3x_2^{-1}(x_1^2x_2^{-1})^{-2} = 4\frac{x_2}{x_1}.$$

It follows that $|\partial b/\partial x_1|$ is unbounded. Note that $x_2/x_1 = g_{\alpha_3}(x_1,x_2)/x_1$, which is unbounded because

$$|\partial b/\partial x_3| = 2x_1^4x_2^{-1}x_3^{-3} < 2x_1^{-2}x_2^2$$

is not bounded.

3.5 C-Mild parametrization of power-subanalytic curves

In this section we show the following C-mild parametrization theorem for power-subanalytic curves for some arbitrary $C > 0$. The proof strategy is analogous to the proof of the C^r-parametrization theorem, one just has to replace the power substitution by an exponential substitution. We will see that this causes problems in higher dimensions.

Theorem 3.5.1 ([VH21a, Theorem 4.7]). *Let $\mathcal{C}$ be a power-subanalytic family of curves in $[-1,1]^n$, then for all $C > 0$, there exists a uniform C-mild parametrization of $\mathcal{C}$.*

Proof. By the cell decomposition theorem, we may suppose that $\mathcal{C}$ is the graph of a power-subanalytic function $\varphi : T \times (0,1) \to [-1,1]^{n-1}$ such that $\varphi(t,x) \in C_t$ for each $(t,x) \in T \times (0,1)^m$.

By the pre-parametrization theorem, we obtain finitely many charts $f_l : C_l \to \mathcal{C}$, where C_l is an open cell in $T_l \times (0,1)$ for some $T_l \subset T$, that are prepared in x. Let P be the exponential substitution $T_l \times (0,1) \to T_l \times (0,1)$ given by

$$(t,x) \mapsto (t, e^{1-1/x^{1/C}}).$$

Let b_l be the associated bounded monomial map of f_l, which has bounded C^1-norm (as a result of the pre-parametrization theorem). By Proposition 2.5.10, it follows that $b_l \circ P$ is (A,B,C)-mild up to order $+\infty$ on $P^{-1}(C_l)$, for some $A,B > 0$ (note that A depends on C). Since f_l is prepared in x with associated bounded monomial map b_l, we have that $f_l \circ P : P^{-1}(C) \to \mathcal{C}$ is given by

$$(f_l \circ P)(t,x) = b_{l,j}(P(t,x))F_l(b_l(P(t,x)),$$

where $b_{l,j}$ is a component function of b_l and F_l is analytic and non-vanishing on an open neighborhood of $\overline{\mathrm{Im}(b_l)}$. Then F_l is $(A',B',0)$-mild up to order $+\infty$ for some A',B'. Clearly, adjusting A' and B' if necessary, we may suppose F_l is (A',B',C)-mild up to order $+\infty$. Since $b_l \circ P$ is (A,B,C)-mild up to order $+\infty$, it follows by Proposition 2.2.5 that $F_l \circ b_l \circ P$ is (A,B,C)-mild up to order $+\infty$ for some larger $A,B > 0$. By Proposition 2.2.2, it follows that $f_l \circ P$ is (A,B,C)-mild up to order $+\infty$, for possibly larger A,B. Finally, linearly map $T_l \times (0,1)$ onto C_l to finish the proof. ☐

Remark 3.5.2.

1. Note that we use Proposition 2.5.10 for the choice $\kappa = 1/C$. Therefore, if $0 < C \leq 1$, we have $\kappa \geq 1$ and we may assume that the C-mild parametrization consists of functions that are $(A/C,B,C)$-mild for some $A,B > 0$ that only depend on $\mathcal{C}$.

2. Since the map P makes use of the unrestricted exponential function, it is not power-subanalytic. As a consequence, the charts in the theorem are not power-subanalytic. Moreover, it follows that also $P^{-1}(C_l)$ is not power-subanalytic. This is the main obstacle in higher dimensions. Additionally, the power substitution has a good interaction with functions that are prepared in x, since they are essentially monomials, but this is

not the case for the exponential substitution, which will be more clear in Example 3.5.3 below.

We conclude this section with applying this method to the family of hyperbolas of Section 3.2.

Example 3.5.3. Let $\mathcal{C}$ be the family of algebraic curves

$$\{(t,x,y) \in (0,1) \times (-1,1)^2 \mid xy = t\}$$

as in Section 3.2. In that section we have seen that is suffices to reparametrize the function $f : U = (0,1) \times (\sqrt{t}, 1) \to \mathbb{R}$ given by

$$f(t,x) = t/x.$$

Clearly, f is prepared in x and its associated bounded monomial map has C^1-norm 1. Let $C > 0$ and let P be defined as in the proof of Theorem 3.5.1. We have that $f \circ P$ is given by

$$(f \circ P)(t,x) = te^{1/x^{1/C}-1}$$

on $P^{-1}(U)$. Since the inverse of $e^{1-1/x^{1/C}}$ is given by $(1-\ln(x))^{-C}$, $P^{-1}(U)$ is equal to $(0,1) \times ((1-\ln(\sqrt{t}))^{-C}, 1)$. As we have mentioned in Remark 3.5.2, we see that the interaction of P with the walls of the cell is rather complicated.

To finish the C-mild parametrization of $\mathcal{C}$, consider the map $\Phi : (0,1) \times (0,1) \to P^{-1}(U)$ defined by

$$\Phi(t,x) = (t, (1-\ln(\sqrt{t}))^{-C} + (1-(1-\ln(\sqrt{t}))^{-C})x).$$

Now $f \circ P \circ \Phi$ is the desired reparametrization of f, i.e., the map $(\Phi, f \circ P \circ \Phi)$ is a C-mild chart of the curve $\mathcal{C}$ and the others are obtained by symmetry or were trivial. More precisely, two charts of the C^1-parametrization were families of constant functions and therefore are already as desired. The other 4 charts are reparametrized as above and become (A,B,C)-mild up to order $+\infty$ for some $A, B > 0$.

3.6 Remarks

We speculate on generalizations of the results of this section and give some additional remarks on the definability issues concerning C-mild parametrizations.

Preparation theorems

It is clear that the preparation theorem (Theorem 3.4.1) is the key ingredient of the pre-parametrization theorem. Theorem 3.4.1 is a strong result, because there is some information on the *unit*, that is the non-vanishing function F (hence a unit), as in Definition 3.4.2, that is analytic on an open neighborhood of $\overline{\mathrm{Im}(b)}$, equivalently it is $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$.

For so-called "polynomially bounded" o-minimal structures, there is a general preparation theorem.

Theorem ([vdDS02, Theorem 2.1])**.** Suppose that $\mathcal{S}$ is a polynomially bounded o-minimal structure on $\mathbb{R}$ and that $f : \mathbb{R}^{m+1} \to \mathbb{R}$ is definable. Then there is a finite covering $\mathcal{C}$ of $\mathbb{R}^{m+1}$ by definable sets, and for each set $S \in \mathcal{C}$, there is some $\lambda \in \mathbb{R}$ and functions $\theta, a : \mathbb{R}^m \to \mathbb{R}$ and $u : \mathbb{R}^{m+1} \to \mathbb{R}$, all definable, such that the graph of θ is disjoint from S and for all $(x, x_{m+1}) \in S$, we have that

$$f(x, x_{m+1}) = |x_{m+1} - \theta(x)|^{\lambda} a(x)u(x, x_{m+1})$$

with $|u(x, x_{m+1}) - 1| < 1/2$.

The function u in this theorem is the unit. Since it is just definable, we cannot say anything about its derivatives, in particular they could be unbounded or not (A, B, C)-mild up to order $+\infty$ for any $A, B > 0$ and $C \geq 0$. However the latter is not really an issue, as we have mentioned below Lemma 2.5.1.

Note that Theorem 3.4.1 holds for the structure of (globally) subanalytic sets. Its language, the subanalytic language, contains symbols for functions that are analytic on an open neighborhood of a compact domain. Therefore, each of them is $(A, B, 0)$-mild up to order $+\infty$ for some $A, B > 0$. Therefore one might ask: if the o-minimal structure is obtained from the semi-algebraic sets by adding function symbols for the restriction of functions that are (A, B, C)-mild up to order $+\infty$ for some $A, B > 0$ and $C \geq 0$, can we say the same about the unit? An example of such an o-minimal structure is the structure $\mathbb{R}_{\mathcal{G}}$, see [vdDS00] for the details. It is "generated" by some "Gevrey functions", as mentioned in the introduction of Chapter 2. As far as I know, the preparation theorem that I would claim in this case, is not known (or published) yet.

Another example of an o-minimal structure, rather say a class of o-minimal structures, that arises by adding functions with bounds on the derivatives is the structure $\mathbb{R}_{\mathcal{C}(M)}$, where $\mathcal{C}(M)$ is a Denjoy-Carleman class (see Definition 2.6.7). Recall that if $M_n = n!$, this is the same as $(A, B, 0)$-mild functions up to order $+\infty$ for some $A, B > 0$, and that additional assumptions are required on M such that it is closed under compositions. This structure is studied in [RSW03]. To show that $\mathbb{R}_{\mathcal{C}(M)}$ is o-minimal, they suppose additionally that $\mathcal{C}(M)$ is quasi-analytic, i.e., the map sending a function to its Taylor series (as formal power series) is injective. In fact, this technique is also used to show that $\mathbb{R}_{\mathcal{G}}$ is o-minimal and this approach is described in general in [RS15].

For the structure $\mathbb{R}_{\mathcal{C}(M)}$, it is natural to ask: is there a preparation theorem for this structure, where the unit belongs to $\mathcal{C}(M)$? The answer seems negative due to the fact that this is not true on the level of germs, see [ABB$^+$14]. However, this obstacle could perhaps be overcome. Indeed, the fact that these classes are not closed under Weierstrass preparation, does not mean that nothing could be said about the unit. Perhaps in belongs to some different class? Note that the class of C-mild functions with $C > 0$ is not quasi-analytic due to the Denjoy-Carleman Theorem (see for instance [KP02, Theorem 4.1.15]). Therefore, the result of [ABB$^+$14] does not apply to the C-mild functions with $C > 0$.

Suppose that in the preparation theorem above, the unit u is (A, B, C)-mild for some $A, B > 0$ and $C \geq 0$. Then it seems that one can deduce an analogue of the pre-parametrization theorem for that o-minimal structure and thus also deduce a uniform C^r-parametrization theorem as Theorem 3.3.1 and uniform C-mild parametrization theorem as Theorem 3.5.1.

More about definability

We have already mentioned in the introduction that for the C-mild parametrization theorem, there are some definability issues. If $\mathcal{X}$ is power-subanalytic, one obtains power-subanalytic charts by the pre-parametrization theorem. The composition with the power substitutions remains power-subanalytic. Clearly, this is not the case for the exponential substitution (see Remark 3.5.2).

It seems that for uniform C-mild parametrizations, one requires the exponential function or, more generally, the charts will be definable in an o-minimal structure that is not polynomially bounded. For instance, it follows from [BN19, Proposition 2] that the family of curves $\mathcal{C}$ in Section 3.2 cannot be uniformly parametrized with subanalytic functions that are (A, B, C)-mild up to order $+\infty$.

By work of Thomas in [Tho11], it follows that for all $C \geq 0$, there is a set X

definable in some polynomially bounded o-minimal structure on $\mathbb{R}$ that does not have a *definable* C-mild parametrization. Moreover, one can find such a set X in an o-minimal structure that has analytic cell decomposition (this is not the case in general, see Remark 1.3.7). In fact, she considers the more general setting of G-mild functions, as mentioned below Definition 2.6.7.

In light of these two results, it seems indeed required to study C-mild parametrizations in structures that are not polynomially bounded. In that case, we lose the preparation theorem in general. However, if a function is constructed using functions that are definable in some polynomially bounded o-minimal structure and compositions of the entire exponential function or logarithm, there is a preparation theorem for these type of functions in [vdDS02]. If in particular one takes the structure of subanalytic sets, this generalizes a result by Lion and Rolin [LR97]. In the latter case, the unit is analytic. Note that [Mil06] generalizes the result of [LR97] for subanalytic functions to power-subanalytic functions in the sense that one may use some subclass of the analytic functions rather than all analytic functions.

Chapter 4

Rational points on definable sets

In Chapters 2 and 3, we have mentioned that mild functions and parametrizations are used in diophantine geometry. In this chapter we make these statements more precise.

Most notably is a precise statement of the main result by Pila and Wilkie in [PW06], now well known as the Counting Theorem, in Section 4.1. Roughly speaking the theorem says that there are few rational points on transcendental sets. In that section we also provide some history of the results and more recent developments and applications. Furthermore, we discuss a conjecture by Wilkie that strengthens the Counting Theorem in the o-minimal structure $\mathbb{R}_{\exp}$.

Section 4.2 contains the key ingredients of the proof of the Counting Theorem, most importantly the determinant method. It is this method that requires a suitable parametrization theorem as an input. A complete proof of the Counting Theorem can be found in the original article [PW06]. Finally in Section 4.3 we give some more recent related results.

Nothing in this chapter is my work and there are several other articles or surveys that also provide on overview of this field. These will be clearly indicated in the text. There is one exception: I make a constant in Theorem 4.3.1, which is [CPW20, Theorem 2.3.1], more explicit using that we know how many charts their method yields by my work.

4.1 The Pila-Wilkie Counting Theorem

This section provides an overview of the developments towards the Counting Theorem and its applications afterwards. A lot of this information can also be found in the original article [PW06] by Pila and Wilkie and in the introduction of several articles, for instance [Pil10], [BN19] or [CPW20].

Some older results

Together with Bombieri, Pila started studying rational points on curves in [BP89]. They have two main results in that paper. Let $y = f(x)$ define a plane curve Γ. Firstly, if $f \in \mathcal{C}^D([0, N])$, $|f|_0 \leq N$, $|f'|_0 \leq 1$ and $f^{(D)} \neq 0$ on $[0, N]$, then there are at most

$$c(\epsilon_D)N^{\frac{1}{2}+\epsilon_D}$$

integer points on the curve, where $\epsilon_D \to 0$ if $D \to \infty$. Secondly, if f is a transcendental analytic function, then

$$\left|t\Gamma \cap \mathbb{Z}^2\right| \leq c(f, \epsilon)t^{\epsilon}$$

for every $\epsilon > 0$, where $t\Gamma = \{(x, y) \in \mathbb{R}^2 \mid ty = tf(x)\}$ for some $t \in \mathbb{N}$ and $c(f, \epsilon)$ is some constant that depends on f and ϵ. These type of results were studied earlier by Schmidt [Sch85], who improved earlier work by several other authors.

Note that in the first result, the exponent $1/2$ is the best possible, consider for example $f(x) = \sqrt{x}$. It seems that if a curve is transcendental, this phenomenon does not occur and thus one can obtain the better upper bound in the second result.

Pila has improved this work in several ways over the following two decades. [Pil91] and [Pil96] further refine the first result for plane curves. [Pil95] studies rational points on varieties, i.e., extends the work to higher dimensions. The results have been generalized to subanalytic surfaces in [Pil04] and [Pil05].

The Counting Theorem

In [Wil04], Wilkie showed similar results as in [BP89], for curves definable in an o-minimal structure. His method differs from the Bombieri-Pila strategy, he instead uses Siegel's Lemma. This method is also used to prove Theorem 4.3.3, one of the recent results in Section 4.3.

In [PW06] Pila and Wilkie proved a result that holds for a large class of sets, now known as the Counting Theorem. In order to precisely formulate this theorem, we need to introduce some notation.

Definition 4.1.1. Let $q = a/b \in \mathbb{Q}$ with $\gcd(a, b) = 1$. The **height** of q is given by:

$$h(q) = \max(|a|, |b|).$$

If $q \in \mathbb{Q}^n$, then $h(q) = \max(h(q_1), \ldots, h(q_n))$.

The notion of height measures the complexity of a rational number. There are several heights, we consider here the "multiplicative height". It can be extended to number fields and algebraic numbers, see [BG06]. The idea is that if we put some threshold $H \in \mathbb{N}$ on the height of the points of a set, there are only finitely many. More precisely, we study the set of **rational points of height at most H** on a set X:

$$X(\mathbb{Q}, H) = \{x \in X \cap \mathbb{Q}^n \mid h(x) \leq H\}.$$

Clearly, this set is finite for every $H \in \mathbb{N}$. The idea is that these sets approximate X as H tends to infinity, and that the cardinality of these sets yield information about X. The Counting Theorem is such a result. We need one more definition.

Definition 4.1.2. The **algebraic part** X^{alg} of X is the union of all semi-algebraic sets contained in X of dimension at least one. Its complement in X is called the **transcendental part** X^{trans} of X.

Note that X^{alg} might not be (semi-) algebraic. Actually, if X is definable in an o-minimal structure, then X^{alg} might not be definable. Moreover, it is in general not easy to compute X^{alg}. In [PW06] they consider as an example the set

$$X = \{(x, y, z) \in \mathbb{R}^3 \mid z = x^y, x, y \in [1, 2]\}.$$

If $y \in \mathbb{Q}$, then the equation $z = x^y$ yields a semi-algebraic curve contained in X, thus is part of X^{alg}. It is claimed that this is exactly the algebraic part of X. While X is subanalytic, X^{alg} is not definable in an o-minimal structure since it has infinitely many connected components. However, for some sets, $X = X^{\mathrm{alg}}$, for example, if X is open in $\mathbb{R}^n$. This shows that X^{alg} might be definable, but not semi-algebraic. The precise statement of the Counting Theorem deals with these complications.

Theorem 4.1.3 (Pila-Wilkie Counting Theorem, [PW06, Theorem 1.10])**.** *Let $\mathcal{X} \subset T \times \mathbb{R}^n$ be a definable family, where $T \subset \mathbb{R}^k$, and $\epsilon > 0$. There is a definable family $W(\mathcal{X}, \epsilon) \subset \mathcal{X}$ and a constant $c(\mathcal{X}, \epsilon)$ with the following property. For every $t \in T$, we have that $W_t \subset X_t^{alg}$, and*

$$|(X_t \setminus W_t)(\mathbb{Q}, H)| \leq c(\mathcal{X}, \epsilon) H^{\epsilon}.$$

This theorem generalizes the result above on curves defined by transcendental analytic functions to a much larger class of objects. Note that when X contains a semi-algebraic curve, one cannot expect such a result to hold since in the first result above, the exponent $1/2$ is sharp. This is the main reason why one considers the transcendental part of X.

A similar result on counting algebraic points has been proved by Pila in [Pil09]. In that case, the constant $c(\mathcal{X}, \epsilon)$ also depends on a threshold on the degree of the algebraic extension.

Wilkie's Conjecture

In [Pil04], Pila explains how he and Bombieri constructed examples to show that the upper bound of the Counting Theorem is the best possible for subanalytic curves, thus also for higher dimensional subanalytic sets. However, Wilkie conjectured that it can be improved for sets defined by "nice" analytic functions.

Conjecture (Wilkie, [PW06, Conjecture 1.11])**.** Suppose that X is definable in $\mathbb{R}_{\text{exp}}$. Then there are constants $c_1(X), c_2(X)$ such that (for $T \geq e$)

$$\left|X^{\text{trans}}(\mathbb{Q}, H)\right| \leq c_1(X)\log(H)^{c_2(X)}.$$

One could study Wilkie's Conjecture more generally in the structure $\mathbb{R}_{\text{Pfaff}}$ of Pfaffian sets, which is contained in the "Pfaffian closure" of $\mathbb{R}_{\text{exp}}$, see [Spe99]. There is a proof of this conjecture for the structure on $\mathbb{R}$ generated by the restricted exponential and restricted sine function by Binyamini and Novikov in [BN17]. They also show it in the more general setting of counting algebraic points over number fields. However, this does not deal with the *unrestricted* exponential function. One might consider more generally the o-minimal structure $\mathbb{R}_{\text{resPfaff}}$. This structure is constructed similar to $\mathbb{R}_{\text{an}}$, but adding restricted Pfaffian functions instead of restricted analytic functions (see Section 1.2). I am not aware of a proof of Wilkie's Conjecture in this setting, but it seems that it should hold there in view of the result of [BN17] and the result in [JMT11], which we will mention below.

The conjecture has been shown for Pfaffian curves in $\mathbb{R}^2$ by Pila in [Pil06a], under the assumption that these curves have a C-mild parametrization. We have seen in Chapter 3, that this conjecture is true for power-subanalytic curves in $\mathbb{R}^n$. However Wilkie's Conjecture is not true for subanalytic curves in general, as we have mentioned before. The reason for this is that they may have "oscillation" (see [Pil06a, Section 4.3]). More precisely, Pfaffian curves do not have many intersections with algebraic curves, while subanalytic function can have many intersections. This result by Pila has been generalized to functions definable in $\mathbb{R}_{\mathrm{resPfaff}}$ in [JMT11], using a 0-mild parametrization theorem.

In [Pil10], Pila considers the following surface in $\mathbb{R}^3$:

$$\{(x,y,z) \in (0,+\infty)^3 \mid \log(x)\log(y) = \log(z)\}.$$

He constructs an explicit C-mild parametrization of this surface (intersected with $(0,1)^3$) and thus was able to use the methods he had developed before. In particular, it follows that Wilkie's Conjecture is true for this surface, which is definable in $\mathbb{R}_{\exp}$. This result was generalized by Butler in [But12] to surfaces of the form

$$\{(x,y,z) \in (0,+\infty)^3 \mid \log(x)^a \log(y)^b \log(z)^c = 1\}$$

for some $a,b,c \in \mathbb{Q}$.

Independently, it is shown in [But12] and [JT12] that Wilkie's Conjecture holds for plane curves definable in $\mathbb{R}_{\exp}$. Moreover, in [JT12] the conjecture is shown for a curve defined by a function that is "existentially definable" in $\mathbb{R}_{\mathrm{Pfaff}}$, which is more general than Pfaffian, and for surfaces definable in $\mathbb{R}_{\mathrm{resPfaff}}$. Note that this does not include the surfaces in $\mathbb{R}^3$ above.

Consequences of the Counting Theorem

In [PZ08] Pila and Zannier gave another proof of the Manin-Mumford conjecture. Roughly speaking this result says that if a curve has infinitely many torsion points, then it is equal to its Jacobian variety. Their proof strategy consists of creating a contradiction between the upper bound on rational points on a set, coming from the Counting Theorem, and a lower bound coming from Galois theory based on the work of Masser [Mas84]. The torsion points are translated to rational points by a function that is definable in an o-minimal structure (usually $\mathbb{R}_{\mathrm{an,exp}}$), called the uniformization map.

This strategy has successfully been applied to other settings. Most notably, an unconditional proof of the André-Oort conjecture for the product of modular curves in [Pil11], for which Pila has received the Clay Research Award in 2011. A nice survey on this proof strategy, including the basics of o-minimality, can be found in [Sca17]. For more details, see the book [JW15].

In [PT14], together with Tsimerman, Pila showed the Ax-Lindemann Theorem for $\mathcal{A}_g$ and in [PT16] Ax-Schanuel for the j-function. Together with Mok and Tsimerman [MPT19], he proved the Ax-Schanuel conjecture for Shimura varieties. These results are expected to work in other similar settings as well. For instance the recent developments by Bakker, Klingler and Tsimerman [BKT20] and the consequences in Hodge theory in [BBKT20, BBT20].

The lower bounds that are used in the Pila-Zannier strategy, are hard to establish. The idea is that if Wilkie's Conjecture would be true for the considered application, then the desired lower bound can be relaxed. However, in the applications, one is often not working in $\mathbb{R}_{\exp}$. Therefore it would be required that the conjecture holds for an o-minimal structure containing $\mathbb{R}_{\exp}$, such that the maps involved in the application are definable in this structure. I am currently not aware of examples of such a structure, even in the case of curves.

4.2 The determinant method

The goal of this section is to explain the key ingredients in the proof of the Counting Theorem (Theorem 4.1.3). The proof consists of two main ingredients. Firstly, one requires a parametrization theorem, where the charts of the parametrization have bounded C^r-norm up to some order r. This order r depends on the choice of $\epsilon > 0$ in the statement of the Counting Theorem. Secondly, on each of these charts, one finds "few" hypersurfaces of degree d, such that all rational points of height at most H on X lie on at least one of these hypersurfaces. Here, "few" means that for $d \to \infty$, it is equal to cH^ϵ for some constant $c > 0$. This constant c depends on the parametrization of the considered set X. The proof of this second statement is often called the determinant method due to some determinant appearing in the proof.

More precisely, the second ingredient is the following proposition.

Proposition 4.2.1 ([PW06, Proposition 6.1])**.** *Suppose that $m < n$. Then there are, for each $d \in \mathbb{N}$, $d \geq 1$, a nonnegative integer $r(m,n,d)$ and positive constants $\epsilon(m,n,d)$, $C(m,n,d)$ with the following property.*

Suppose that $\phi : (0,1)^m \to \mathbb{R}^n$ is a $C^{r(m,n,d)}$-map with $|\phi|_{r(m,n,d)} \leq 1$. Let $X = Im(\phi)$ and $H \geq 1$. Then $X(\mathbb{Q}, H)$ is contained in the union of at most

$$C(m,n,d)H^{\epsilon(m,n,d)}$$

hypersurfaces of degree d. Furthermore, $\epsilon(m,n,d) \to 0$ as $d \to \infty$.

The additional assumption $m < n$ ensures that $\epsilon(m,n,d) \to 0$ for $d \to \infty$. If $m = n$, the statement still holds, without that last assertion. This is not a problem to prove the Counting Theorem, since if X is open in $\mathbb{R}^n$, then $X^{\text{alg}} = X$. Removing the interior of a set yields a lower dimensional set (by o-minimality), without the loss of rational points on the transcendental part.

This proposition gives the aforementioned result on "few" hypersurfaces for sets X such that $X = \mathrm{Im}(\phi)$, for some sufficiently differentiable map ϕ that has bounded derivatives up to order r. Therefore, to apply this proposition to a general set X, one requires a parametrization result of this type. It is not necessary that $|\phi|_{r(m,n,d)} \leq 1$, but imposing so masks the fact that $C(m,n,d)$ depends on the parametrization. For instance, if one has a C-mild parametrization, one can obtain a more explicit formula for $C(m,n,d)$ in terms of the mildness parameters A, B and C, see [Pil10, Corollary 3.3]. Also the functions that belong to some Denjoy-Carleman class (see Definition 2.6.7) could be considered, although I have not seen this in literature yet.

Combining a suitable parametrization result with Proposition 4.2.1 yields the following result.

Proposition 4.2.2 ([PW06, Main Lemma])**.** *Let X be a definable set of dimension $m < n$ and $\epsilon > 0$. There exists a constant $d(\epsilon, m, n)$ and a constant $c(X, \epsilon)$ with the following property. For all $H \geq 1$, the set $X(\mathbb{Q}, H)$ is contained in the union of a most*

$$c(X, \epsilon)H^{\epsilon}$$

hypersurfaces of degree d.

It is easily verified that this proposition holds uniform for a family $\mathcal{X}$ if one has a uniform parametrization. The constant $c(X, \epsilon)$ is obtained by multiplying $C(m,n,d)$ of Proposition 4.2.1 with the amount of charts of the parametrization. The required differentiability is also computed via Proposition 4.2.1, where one should pick d such that $\epsilon(m,n,d) < \epsilon$.

Finally, remark that the parametrization result required for Proposition 4.2.1 implies that $X \subset [-1,1]^n$. Now since rational points are mapped to rational points under the map $x \mapsto 1/x$, and because this map also preserves the height of a rational number, this is not an issue to prove the Counting Theorem. In fact, this is the first step of its proof.

Let us now give a sketch of the proof of Proposition 4.2.1, or stated differently: explain the determinant method. Precise details can be found in [PW06], [Pil10] or [CPW20].

Sketch of proof. Assume $X = \mathrm{Im}(\phi)$ for some C^r-map $\phi : (0,1)^m \to \mathbb{R}^n$, where the required differentiability on ϕ will be determined later. Let $d, H \in \mathbb{N}$. The goal is to cover $(0,1)^m$ with balls B_R of some radius R, also to be determined, such that $\mathrm{Im}(\phi|_{B_R})(\mathbb{Q}, H)$ is contained in a single algebraic hypersurface of degree d. Therefore, if we know R, we know how many hypersurfaces we will need at most.

If $q \in \mathrm{Im}(\phi)(\mathbb{Q}, H)$, it yields a linear equation on the coefficients of the hypersurface:

$$\sum a_\nu q^\nu = 0, \tag{4.1}$$

where the sum runs over all monomials with natural exponents in n variables of degree at most d. Denote the cardinality of this set by $D_n(d)$. Clearly, if there are less than $D_n(d)$ points in $\mathrm{Im}(\phi)(\mathbb{Q}, H)$, there is nothing to be done.

Suppose that there are more equations than variables. Consider any $D_n(d) \times D_n(d)$ subdeterminant Δ of the system 4.1. We will show that if R is sufficiently small, they are all zero, hence the system has a solution. (This explains the name determinant method.) We will use that $q = \phi(p)$ for some $p \in (0,1)^m$ and then use the Taylor approximation of ϕ^ν for each ν to reduce to monomials in m variables. We want a Taylor approximation of order $b+1$, where b is uniquely determined by:

$$D_m(b) \leq D_n(d) < D_m(b+1).$$

This number $b+1$, that depends on m, n and d, is the required order of differentiability $r(m, n, d)$ in Proposition 4.2.1. It does not depend on ϕ. Note that since $m < n$, we have that b is larger than d.

Now apply Taylor approximation to ϕ^ν with remainder term of degree $b+1$ and use linearity in columns to write the determinant in the same form of the system of equation 4.1. More precisely, for each column, we pick some exponent and split the determinant linearly accordingly. In this way, the original subdeterminant Δ becomes a sum of many determinants Δ'. Many of these determinants Δ' are automatically zero (for instance if one picks many terms of low degree, then the determinant will be zero as the columns will be linearly dependent since $D_m(b) \leq D_n(d)$).

A nonzero contribution Δ' can then be estimated (in absolute value) in terms of the derivatives of ϕ^ν, which can be bounded by some suitable power of R times a constant that depends on m, n, d and the upper bounds on the derivatives of ϕ. The estimate on the power of R uses the assumption that Δ' was nonzero. Since the amount of Δ' is also a constant depending on m, n and d, we have found an *upper bound* for $|\Delta|$, which is some constant times a power of R.

However, since Δ is a determinant with coefficients in $\mathbb{Q}$, one can find some integer Z such that $Z\Delta \in \mathbb{Z}$. Moreover, since we consider rational points up to height H, this Z can be bounded in terms of a power of H. This power depends on n and d. This yields a *lower bound* on Δ, in terms of a power of H. This lower bound is nonzero if Δ was nonzero.

Finally, if one chooses R small enough such that the upper bound is smaller than the lower bound, the determinant must be zero. $\square$

Remark 4.2.3.

1. This method only depends on the type of the parametrization in the step where the nonzero determinants Δ' are bounded in terms of the Taylor approximation of ϕ^ν. If the upper bounds are not too bad, for instance if the parametrization is C-mild, one can also use the method (see [Pil10, Corollary 3.3]).

2. To use Taylors theorem, it is required that the domain of ϕ, in this case $(0,1)^m$, is convex. Therefore, to use this method, it is not necessary that the domain of the charts of a parametrization is $(0,1)^m$, but they should have some convex domain.

3. For a precise analysis of the constants appearing in the proof, especially the part where some upper bound on the nonzero contribution Δ' is claimed (which is the most technical part of the proof), see [Pil10] and also [CPW20, p. 10-12].

4.3 Some recent results

In [Pil10] and [CPW20], Pila makes a particular choice of d, to achieve an upper bound on the amount of hypersurfaces that resembles the upper bound in Wilkie's Conjecture. More precisely, he picks $d = [\log(H)^{m/(n-m)}]$, where $[\cdot]$ denotes the unique integer such that $[x] \leq x < [x] + 1$. Using this particular choice of d, one deduces the following result.

Theorem 4.3.1 ([CPW20, Theorem 2.3.1])**.** *Suppose that X is an m-dimensional subset of $(0,1)^n$ with $m < n$. Then there exist positive constants $c_1(X), c_2(X)$ such that, for $H \geq e$, $X(\mathbb{Q}, H)$ is contained in the union of at most*

$$c_1(X) \log(H)^{c_2(X)}$$

algebraic hypersurfaces of degree at most $[\log(H)^{m/(n-m)}]$.

This result, and also its consequence below, actually hold for families $\mathcal{X}$ if the parametrization is uniform in families. But this is not important for the discussion here.

Let us make the constant $c_2(X)$ more explicit, using that I have shown that their construction leads to cr^{m^3} charts for some constant c that depends on X (see Theorem 3.3.1 and below). It is tempting to just follow the proof in [CPW20], but there is a catch. In [CPW20], the C^r supremum norm is used, not the weighted version that is defined in Definition 3.1.1, see (2) of Remark 3.1.4.

A first way to solve this problem is to improve the linear reparametrization method (Corollary 3.1.7), such that it bounds the C^r supremum norm by 1 instead of the C^r-norm. Recall that in the end of the proof of Theorem 3.3.1, we obtain a finite amount of charts $(0,1)^m \to X$, that are $(Ar^{m^2}, 1, 0)$-mild up to order r. To bound the C^r supremum norm, apply Lemma 2.5.1 instead of Lemma 3.1.6 in the proof of the linear reparametrization method. In that case, one should cover the hypercube

$$(0, Ar^{m^2} r)^m$$

with translates of $(0,1)^m$. This yields cr^{m^3+m} charts for some $c > 0$ that only depends on A and m.

Now the norm coincides with the norm in [CPW20], one has to use that

$$b + 1 \leq 2\left(\frac{m!d^n}{n!}\right)^{1/m}$$

for sufficiently large d, see [CPW20, p. 9]. Note that this means that H also has to be sufficiently large since $d = [\log(H)^{m/(n-m)}]$. Since $b+1$ is the required differentiability of the parametrization, the method will yield

$$c_1(X)\log(H)^{n(m^3+m)/(n-m)}$$

algebraic hypersurfaces of degree at most $[\log(H)^{m/(n-m)}]$, see [CPW20, p. 12] the end of the proof. We see that polynomial dependence in r for the amount of charts implies polynomial dependence in $\log(H)$ for the amount of algebraic hypersurfaces.

A second way to solve this problem caused by the use of two different norms, would be to go through the computations of [Pil10, Section 3], eventually ending up with a formula as in [Pil10, Corollary 3.3], which is more precise in the mildness parameters A and C, as we have mentioned in Remark 4.2.3.

In the end of Section 2.5, we have indicated that a C-mild parametrization, where $C \leq 1$ might be more suitable. Therefore, suppose that X is a curve. By the C-mild parametrization theorem (Theorem 3.5.1), there is some $A > 0$ such that for all $C \in (0,1)$, X has a $(A(1/C), 1, C)$-mild parametrization. Now perform a linear reparametrization to bound the C^r supremum norm as above. This yields $(\frac{A'}{C}r)^{1+C}$ charts, thus $c_1(X)C^{-(1+C)}\log(H)^{n(1+C)/(n-1)}$ algebraic hypersurfaces of degree at most $[\log(H)^{1/(n-1)}]$. In this case that doesn't really improve a lot, but if Theorem 3.5.1 were extended to higher dimensions, that would be a significant improvement, more precisely a linear dependence on m in the exponent instead of the polynomial of degree 3.

In [CPW20] they have deduced the following result from Theorem 4.3.1.

Proposition 4.3.2 ([CPW20, Proposition 2.3.6])**.** *Let r be a nonnegative integer and α, β positive integers. Let $X \subset (0,1)^3$ be a power-subanalytic surface, that is the intersection of $(0,1)^3$ with a Pfaffian surface of complexity (at most) (r, α, β). Then there are $c_3(X)$ and $c_4(X)$ such that for $H \geq e$*

$$\left|X^{trans}(\mathbb{Q}, H)\right| \leq c_3(X) \log(H)^{c_4(X)}.$$

We refer to [CPW20] for the precise definition of a Pfaffian surface of complexity (r, α, β). The result follows by following the paper [Pil10], using results on the amount intersections of such a surface with algebraic hypersurfaces by Gabrielov and Vorobjov [GV04].

Binyamini and Novikov also obtain results of this fashion in their paper [BN19]. Their methods allow to express this constants more explicitly if X is semi-algebraic. They also prove a "log" version of the Counting Theorem, see [BN19, Proposition 5], where one is allowed to take the *unrestricted* logarithm of a bounded subanalytic set. Their paper also contains more references on recent results of this type.

Finally, let us mention the following theorem by Habegger that is similar to the Counting Theorem, but counts rational points that are close to a set X, but not too close to X^{alg}. We state it here for rational points, the original statements allows algebraic points.

Theorem 4.3.3 ([Hab18, Theorem 2])**.** *Let $X \subset \mathbb{R}^n$ be closed and definable in a polynomially bounded o-minimal structure. Let $\epsilon > 0$. There exist $c(X, \epsilon) \geq 1$, $\lambda(X, \epsilon) > 0$ and $\theta(X, \epsilon) \in (0, 1]$ such that for all $H \geq 1$, we have*

$$\Big|\Big\{q \in \mathbb{R}^n(\mathbb{Q}, H) \mid \text{ there is } x \in X \text{ with } |x - q| \leq H^{-\lambda(X,\epsilon)}$$

$$\text{and } q \notin \mathcal{N}(X^{alg}, |x-q|^{\theta(X,\epsilon)})\Big\}\Big| \leq c(X, \epsilon) H^{\epsilon}$$

for all $H \geq 1$, where $\mathcal{N}(X^{alg}, |x-q|^{\theta(X,\epsilon)}) = \{z \in \mathbb{R}^n \mid \exists y \in X^{alg} : |z - y| \leq |x-q|^{\theta(X,\epsilon)}\}$ is the $|x-q|^{\theta(X,\epsilon)}$-neighborhood of X^{alg}.

The statement is simpler if $X^{\text{alg}} = \emptyset$, in the sense that no "neighborhood of the empty set" has to be removed, see [Hab18, Corollary 3]. In fact, this theorem follows from a more general statement in the paper. Many technical details are also explained with helpful examples in the introduction of the article.

The proof of the statement uses similar techniques as the proof of the Counting Theorem. However, the determinant method is replaced by the Siegel's Lemma, that Wilkie had proposed in [JW15], and which he had already used in [Wil04].

Conclusion

The two main results in this thesis were deduced from a pre-parametrization theorem by Cluckers, Pila and Wilkie in [CPW20]. Using precise results on mild functions and carefully analyzing the construction of the C^r-parametrization in their work, I have showed that the degree of the amount of charts, that was already known to be polynomial in the order of differentiability r, is itself polynomial in the dimension m of the parametrized set. More precisely, it is cr^{m^3}, for some constant c. This answers a question raised by them in light of the applications in smooth dynamics.

Purely for the sake of parametrizations, it would be interesting to optimize this construction of the C^r-parametrization to obtain cr^m charts. Furthermore, to express the constant c in terms of the complexity of the set if it is semi-algebraic, as in [BN19].

Using an exponential substitution, it is possible to deduce a mild parametrization theorem from the pre-parametrization theorem. However, due to definability issues, it currently only holds for curves. I expect that it also holds for higher dimensional sets. If so, it would be at least as good as the C^r-parametrizations with regard to the applications, especially with the added flexibility in the mildness parameter C.

The definability issue might lead to the following question: which sets, say definable in $\mathbb{R}_{\mathrm{an,exp}}$, can be parametrized using the exponential substitution? An example of such a set, definable in the smaller structure $\mathbb{R}_{\mathrm{exp}}$, was given by Pila [Pil10], namely:

$$\{(x, y, z) \in (0, 1)^3 \mid \log(x)\log(y) = \log(z)\}.$$

Maybe all sets in $\mathbb{R}_{\mathrm{an,exp}}$ have a uniform C-mild parametrization? Or some sufficiently interesting (for the applications in diophantine geometry) subclass? As far as I know, there are currently no results answering these questions.

Of course, one should not restrict to the exponential substitution, but perhaps some structural result for $\mathbb{R}_{\mathrm{an,exp}}$, that can play the role of the preparation Theorem 3.4.1, together with the exponential substitution could be sufficient. Such a theorem is available for certain functions definable in $\mathbb{R}_{\mathrm{an,exp}}$ [LR97], but it is rather complicated.

Bibliography

[ABB$^+$14] F. Acquistapace, F. Broglia, M. Bronshtein, A. Nicoara, and N. Zobin. Failure of the Weierstrass preparation theorem in quasi-analytic Denjoy-Carleman rings. *Adv. Math.*, 258:397–413, 2014.

[BBKT20] B. Bakker, Y. Brunebarbe, B. Klingler, and J. Tsimerman. Definability of mixed period maps. *arXiv*, arXiv:2006.12403, 2020.

[BBT20] B. Bakker, Y. Brunebarbe, and J. Tsimerman. Quasiprojectivity of images of mixed period maps. *arXiv*, arXiv:2006.13709, 2020.

[Bes14] A. Besenyei. Peano's unnoticed proof of Borel's theorem. *Amer. Math. Monthly*, 121(1):69–72, 2014.

[BG06] E. Bombieri and W. Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.

[BKT20] B. Bakker, B. Klingler, and J. Tsimerman. Tame topology of arithmetic quotients and algebraicity of Hodge loci. *J. Amer. Math. Soc.*, 33(4):917–939, 2020.

[BLY15] D. Burguet, G. Liao, and J. Yang. Asymptotic h-expansiveness rate of C^∞ maps. *Proc. Lond. Math. Soc. (3)*, 111(2):381–419, 2015.

[BM88] E. Bierstone and P. D. Milman. Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, (67):5–42, 1988.

[BN17] G. Binyamini and D. Novikov. Wilkie's conjecture for restricted elementary functions. *Ann. of Math. (2)*, 186(1):237–275, 2017.

[BN19] G. Binyamini and D. Novikov. Complex cellular structures. *Ann. of Math. (2)*, 190(1):145–248, 2019.

[BN20] G. Binyamini and D. Novikov. The yomdin-gromov algebraic lemma revisited. *arXiv:2009.02082*, 2020.

[BP89] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[Bur08] D. Burguet. A proof of Yomdin-Gromov's algebraic lemma. *Israel J. Math.*, 168:291–316, 2008.

[But12] L. A. Butler. Some cases of Wilkie's conjecture. *Bull. Lond. Math. Soc.*, 44(4):642–660, 2012.

[Car26] T. Carleman. Les fonctions quasi-analytiques. *Gauthier-Villars*, 1926.

[CFY21] R. Cluckers, O. Friedland, and Y. Yomdin. Doubling coverings via resolution of singularities and preparation. *Commun. Contemp. Math.*, 23(2):2050018, 23, 2021.

[CPW20] R. Cluckers, J. Pila, and A. Wilkie. Uniform parametrizations of subanalytic sets and diophantine applications. *Ann. Sci. Ecole Norm. Sup.*, 53(1):1–42, 2020.

[CS96] G. M. Constantine and T. H. Savits. A multivariate Faà di Bruno formula with applications. *Trans. Amer. Math. Soc.*, 348(2):503–520, 1996.

[Den21] A. Denjoy. Sur les fonctions quasi-analytiques de variable reélle. *C. R. Acad. Sci. Paris Ser. 1*, 173:1329–1331, 1921.

[Gev18] M. Gevrey. Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire. *Ann. Sci. École Norm. Sup. (3)*, 35:129–190, 1918.

[Gro87] M. Gromov. Entropy, homology and semialgebraic geometry. Number 145-146, pages 5, 225–240. 1987. Séminaire Bourbaki, Vol. 1985/86.

[Gro97] A. Grothendieck. Esquisse d'un programme. In *Geometric Galois actions, 1*, volume 242 of *London Math. Soc. Lecture Note Ser.*, pages 5–48. Cambridge Univ. Press, Cambridge, 1997. With an English translation on pp. 243–283.

[GV04] A. Gabrielov and N. Vorobjov. Complexity of computations with Pfaffian and Noetherian functions. In *Normal forms, bifurcations and finiteness problems in differential equations*, volume 137 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 211–250. Kluwer Acad. Publ., Dordrecht, 2004.

[Hab18] P. Habegger. Diophantine approximations on definable sets. *Selecta Math. (N.S.)*, 24(2):1633–1675, 2018.

[JMT11] G. O. Jones, D. J. Miller, and M. E. M. Thomas. Mildness and the density of rational points on certain transcendental curves. *Notre Dame J. Form. Log.*, 52(1):67–74, 2011.

[JT12] G. O. Jones and M. E. M. Thomas. The density of algebraic points on certain Pfaffian surfaces. *Q. J. Math.*, 63(3):637–651, 2012.

[JW15] G. O. Jones and A. J. Wilkie, editors. *O-minimality and diophantine geometry*, volume 421 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2015. Lecture notes from the LMS-EPSRC course held at the University of Manchester, Manchester, July 2013.

[Kho91] A. G. Khovanskiĭ. *Fewnomials*, volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Smilka Zdravkovska.

[KP02] S. G. Krantz and H. R. Parks. *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[KPS86] J. F. Knight, A. Pillay, and C. Steinhorn. Definable sets in ordered structures. II. *Trans. Amer. Math. Soc.*, 295(2):593–605, 1986.

[LGR08] O. Le Gal and J.-P. Rolin. Une structure o-minimale sans décomposition cellulaire C^∞. *C. R. Math. Acad. Sci. Paris*, 346(5-6):309–312, 2008.

[LR97] J.-M. Lion and J.-P. Rolin. Théorème de préparation pour les fonctions logarithmico-exponentielles. *Ann. Inst. Fourier (Grenoble)*, 47(3):859–884, 1997.

[Mar02] D. Marker. *Model theory*, volume 217 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002. An introduction.

[Mas84] D. W. Masser. Small values of the quadratic part of the Néron-Tate height on an abelian variety. *Compositio Math.*, 53(2):153–170, 1984.

[Mil94] C. Miller. Expansions of the real field with power functions. *Ann. Pure Appl. Logic*, 68(1):79–94, 1994.

[Mil06] D. J. Miller. A preparation theorem for Weierstrass systems. *Trans. Amer. Math. Soc.*, 358(10):4395–4439, 2006.

[MPT19] N. Mok, J. Pila, and J. Tsimerman. Ax-Schanuel for Shimura varieties. *Ann. of Math. (2)*, 189(3):945–978, 2019.

[Pil91] J. Pila. Geometric postulation of a smooth function and the number of rational points. *Duke Math. J.*, 63(2):449–463, 1991.

[Pil95] J. Pila. Density of integral and rational points on varieties. Number 228, pages 4, 183–187. 1995. Columbia University Number Theory Seminar (New York, 1992).

[Pil96] J. Pila. Density of integer points on plane algebraic curves. *Internat. Math. Res. Notices*, (18):903–912, 1996.

[Pil04] J. Pila. Integer points on the dilation of a subanalytic surface. *Q. J. Math.*, 55(2):207–223, 2004.

[Pil05] J. Pila. Rational points on a subanalytic surface. *Ann. Inst. Fourier (Grenoble)*, 55(5):1501–1516, 2005.

[Pil06a] J. Pila. Mild parameterization and the rational points of a Pfaff curve. *Comment. Math. Univ. St. Pauli*, 55(1):1–8, 2006.

[Pil06b] J. Pila. Note on the rational points of a Pfaff curve. *Proc. Edinb. Math. Soc. (2)*, 49(2):391–397, 2006.

[Pil07] J. Pila. The density of rational points on a Pfaff curve. *Ann. Fac. Sci. Toulouse Math. (6)*, 16(3):635–645, 2007.

[Pil09] J. Pila. On the algebraic points of a definable set. *Selecta Math. (N.S.)*, 15(1):151–170, 2009.

[Pil10] J. Pila. Counting rational points on a certain exponential-algebraic surface. *Ann. Inst. Fourier (Grenoble)*, 60(2):489–514, 2010.

[Pil11] J. Pila. O-minimality and the André-Oort conjecture for $\mathbb{C}^n$. *Ann. of Math. (2)*, 173(3):1779–1840, 2011.

[Pil17] J. Pila. A correction to "Counting rational points on a certain exponential-algebraic surface" [MR2667784]. *Ann. Inst. Fourier (Grenoble)*, 67(3):1277–1278, 2017.

[PS86] A. Pillay and C. Steinhorn. Definable sets in ordered structures. I. *Trans. Amer. Math. Soc.*, 295(2):565–592, 1986.

[PS88] A. Pillay and C. Steinhorn. Definable sets in ordered structures. III. *Trans. Amer. Math. Soc.*, 309(2):469–476, 1988.

[PT14] J. Pila and J. Tsimerman. Ax-Lindemann for A_g. *Ann. of Math. (2)*, 179(2):659–681, 2014.

[PT16] J. Pila and J. Tsimerman. Ax-Schanuel for the j-function. *Duke Math. J.*, 165(13):2587–2605, 2016.

[PW06] J. Pila and A. J. Wilkie. The rational points of a definable set. *Duke Math. J.*, 133(3):591–616, 2006.

[PZ08] J. Pila and U. Zannier. Rational points in periodic analytic sets and the Manin-Mumford conjecture. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 19(2):149–162, 2008.

[RS14] A. Rainer and G. Schindl. Composition in ultradifferentiable classes. *Studia Math.*, 224(2):97–131, 2014.

[RS15] J.-P. Rolin and T. Servi. Quantifier elimination and rectilinearization theorem for generalized quasianalytic algebras. *Proc. Lond. Math. Soc. (3)*, 110(5):1207–1247, 2015.

[RSW03] J.-P. Rolin, P. Speissegger, and A. J. Wilkie. Quasianalytic Denjoy-Carleman classes and o-minimality. *J. Amer. Math. Soc.*, 16(4):751–777, 2003.

[Sca17] T. Scanlon. O-minimality as an approach to the André-Oort conjecture. In *Around the Zilber-Pink conjecture/Autour de la conjecture de Zilber-Pink*, volume 52 of *Panor. Synthèses*, pages 111–165. Soc. Math. France, Paris, 2017.

[Sch85] W. M. Schmidt. Integer points on curves and surfaces. *Monatsh. Math.*, 99(1):45–72, 1985.

[Spe99] P. Speissegger. The Pfaffian closure of an o-minimal structure. *J. Reine Angew. Math.*, 508:189–211, 1999.

[Tho11] M. E. M. Thomas. An o-minimal structure without mild parameterization. *Ann. Pure Appl. Logic*, 162(6):409–418, 2011.

[vdD84] L. van den Dries. Remarks on Tarski's problem concerning $(\mathbf{R}, +, \cdot, \exp)$. In *Logic colloquium '82 (Florence, 1982)*, volume 112 of *Stud. Logic Found. Math.*, pages 97–121. North-Holland, Amsterdam, 1984.

[vdD98] L. van den Dries. *Tame topology and o-minimal structures*, volume 248 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1998.

[vdDM94] L. van den Dries and C. Miller. On the real exponential field with restricted analytic functions. *Israel J. Math.*, 85(1-3):19–56, 1994.

[vdDS00] L. van den Dries and P. Speissegger. The field of reals with multisummable series and the exponential function. *Proc. London Math. Soc. (3)*, 81(3):513–565, 2000.

[vdDS02] L. van den Dries and P. Speissegger. O-minimal preparation theorems. In *Model theory and applications*, volume 11 of *Quad. Mat.*, pages 87–116. Aracne, Rome, 2002.

[VH21a] S. Van Hille. On a family of mild functions. *International Journal of Number Theory*, jan 2021.

[VH21b] Siegfried Van Hille. Smooth parameterizations of power-subanalytic sets and compositions of gevrey functions. *Proceedings of the Edinburgh Mathematical Society*, pages 1–23, 2021.

[Wil96] A. J. Wilkie. Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *J. Amer. Math. Soc.*, 9(4):1051–1094, 1996.

[Wil04] A. J. Wilkie. Diophantine properties of sets definable in o-minimal structures. *J. Symbolic Logic*, 69(3):851–861, 2004.

[Yom87a] Y. Yomdin. C^k-resolution of semialgebraic mappings. Addendum to: "Volume growth and entropy". *Israel J. Math.*, 57(3):301–317, 1987.

[Yom87b] Y. Yomdin. Volume growth and entropy. *Israel J. Math.*, 57(3):285–300, 1987.

[Yom91] Y. Yomdin. Local complexity growth for iterations of real analytic mappings and semicontinuity moduli of the entropy. *Ergodic Theory Dynam. Systems*, 11(3):583–602, 1991.

[Yom08] Y. Yomdin. Analytic reparametrization of semi-algebraic sets. *J. Complexity*, 24(1):54–76, 2008.

[Yom15] Y. Yomdin. Smooth parametrizations in dynamics, analysis, diophantine and computational geometry. *Jpn. J. Ind. Appl. Math.*, 32(2):411–435, 2015.

FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS
SECTION OF ALGEBRA
Celestijnenlaan 200B box 2400
B-3001 Leuven