Proceedings of the Edinburgh Mathematical Society (2021) **64**, 532–554
doi:10.1017/S0013091521000298

SMOOTH PARAMETERIZATIONS OF POWER-SUBANALYTIC SETS AND COMPOSITIONS OF GEVREY FUNCTIONS

SIEGFRIED VAN HILLE

KU Leuven, Celestijnenlaan 200B, 3001 Leuven, Belgium
(siegfried.vanhille@kuleuven.be)

(Received 26 June 2019; first published online 4 June 2021)

Abstract We show that if X is an m-dimensional definable set in $\mathbb{R}_{\mathrm{an}}^{\mathrm{pow}}$, the structure of real subanalytic sets with real power maps added, then for any positive integer r there exists a C^r-parameterization of X consisting of cr^{m^3} maps for some constant c. Moreover, these maps are real analytic and this bound is uniform for a definable family.

Keywords: Gevrey functions; mild functions; subanalytic sets; C^r-parameterization; preparation of power-subanalytic functions; Weierstrass systems

2020 *Mathematics subject classification:* Primary 03C98; 14P15; 32B20; Secondary 26E10; 32B05

1. Introduction

A parameterization of a set $X \subset \mathbb{R}^n$ of dimension m is a finite set of maps $\{f_i : (0, 1)^m \to X \mid i \in \{1, \ldots, N\}\}$ for some $N \in \mathbb{N}$, such that the images of these maps cover X. Consider the following question: given any positive integer r, can one construct a parameterization of X consisting of r times continuously differentiable maps whose C^r-norm is bounded by 1, i.e. a C^r-parameterization?

The interest in this question first arose in 1987 in the study of entropy of dynamical systems in [16, 17]. In these papers, Yomdin sketched a C^r-parameterization result for semi-algebraic sets. A more refined explanation was given by Gromov in [7], a complete proof by Burguet in [3]. In their result, the number of maps of the parameterization is shown to depend on combinatorial data defining X, i.e. n, m, r and the complexity β, which is the maximum degree of the equations and inequalities necessary to define X. However, there is no explicit formula for the number of maps. Yomdin has written a nice survey on the study of parameterizations in [19].

Using o-minimality, a tool of model theory, Pila and Wilkie showed in 2006 that any bounded definable set in any o-minimal structure (containing the semi-algebraic sets) has a C^r-parameterization [13]. However, even in the special case that X is semi-algebraic, they did not deduce a formula for the number of maps of the parameterization in terms

© The Author(s) 2021. Published by Cambridge University Press on Behalf of The Edinburgh Mathematical Society

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

of the complexity of X. The main result of [13] bounds the number of rational points on a definable set X up to a certain height, known as the counting theorem. To obtain this result, they constructed C^r-parameterizations for arbitrarily large r and therefore it is interesting to know how large the parameterization becomes depending on r.

The dependence on r has recently been investigated in [1] by Binyamini and Novikov, who construct a C^r-parameterization consisting of cr^m maps for subanalytic sets. Here, c is a constant that depends on X. Moreover, they show that c depends polynomially on the complexity of X, if X is semi-algebraic. Around the same time, Cluckers, Pila and Wilkie proved in [4] a C^r-parameterization result for power-subanalytic sets using cr^d maps, where c and d are constants depending on X, and this result is uniform.

In this paper, we built on the methods of [4] to further refine their result in Theorem 3.12. It is weaker than the result of [1] for subanalytic sets but holds for the larger class of power-subanalytic sets.

Theorem. *If $X \subset [-1, 1]^n$ is power-subanalytic of dimension $m \leq n$, then for any positive integer r there exists a C^r-parameterization of X consisting of cr^{m^3} maps whose C^r-norm is bounded by 1. Moreover, if X belongs a power-subanalytic family of such sets, the constant c holds for all members of the family.*

The theorem above and the results of [4] actually hold for structures $\mathbb{R}^K_{\mathcal{F}}$, which were studied by Miller in [10]. These structures expand the semi-algebraic sets with restricted analytic functions in a Weierstrass system $\mathcal{F}$ and power maps $x \mapsto x^\mu$ for any $\mu \in K$, where K is a subfield of the field of exponents of $\mathcal{F}$ (see the introduction of § 3 for precise definitions). The structures $\mathbb{R}$, $\mathbb{R}_{\text{an}}$ and $\mathbb{R}^{\text{pow}}_{\text{an}}$, corresponding to the semi-algebraic, subanalytic and power-subanalytic sets respectively, are all examples of this class of structures. The main result of [10] is a preparation theorem for functions definable in these structures, which is one of the main ingredients for the proof of the main result.

In 2011, Jones, Miller and Thomas have shown in [8] that any set definable in (any reduct of) $\mathbb{R}_{\text{an}}$ expanding the real field has an $(A, 0)$-mild parameterization for some $A > 0$, i.e. a parameterization consisting of maps that are $(A, 0)$-mild. The precise definition of a mild function is given in § 2. Informally, for $C \geq 0$, an (A, C)-mild function is a smooth function with good bounds on the derivatives depending on A and C. For any integer $r > 0$, these bounds allow us to bound the C^r-norm by 1 after a suitable substitution. Since we only control the derivatives up to order r, their result is stronger. However, by an elementary example of Yomdin [18, Proposition 3.3], their result cannot be made uniform. Indeed, the family of hyperbolas $\{xy = t^2 \mid (x, y) \in (0, 1)^2, t \in (0, 1)\}$ does not have a uniform $(A, 0)$-mild parameterization, i.e. the amount of maps will depend on the parameter t. Recently, it has been shown in [15] that it has a uniform (A, C)-mild parameterization for all $C > 0$.

The paper is organized as follows. In § 2, we prove various properties of mild functions using only standard techniques of real analysis. A key result on the composition of these functions, Theorem 2.10, permits us to make all of the results of [4] on mild functions explicit and thus also the constant d. This key result is in fact an old result on Gevrey functions [6]. We will give the original proof of this result, but in full generality, and reformulate it in terms of mild functions. In § 3, we provide the necessary background in model theory to state the main theorem and we prove the main theorem. Finally, we

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

explain that when $m \geq 2$, we can in fact obtain a parameterization consisting of $cr^{m^2(m-1)}$ maps.

2. Mild functions

Mild functions are a class of C^∞-maps introduced by Pila in [11]. The upper bounds on the derivatives of a mild function are very suitable to use the determinant method. This method, which Pila developed with Bombieri [2], is a useful tool in diophantine geometry. For example, the determinant method is used in the proof of the counting theorem in [13].

Cluckers, Pila and Wilkie introduced several variants to mild functions in [4]. Since in their result (and our main theorem), they only cared about derivatives up to order r, they defined functions that are mild up to order r. These functions are C^r-maps that satisfy the same bounds on the derivatives as mild functions, but up to order r. In fact, the maps are actually smooth in their work.

Since we will encounter many compositions of mild maps, we need a result on compositions of these maps. This result on compositions will allow us to find the amount of maps in § 3. One can think of mild functions as functions of some class indexed by a real number $C \geq 0$. When $C = 0$, these maps are real analytic, when $C = +\infty$, they are real smooth. Using the theory of Gevrey functions [6], we show that compositions of mild functions of class C are again of class C.

We start with some definitions of multidimensional calculus. Throughout this section, we will work with maps $f = (f_1, \ldots, f_n) : U \subset \mathbb{R}^d \to \mathbb{R}^n$, where U is always assumed to be open in $\mathbb{R}^d$. We say that $f : U \subset \mathbb{R}^d \to \mathbb{R}$ is C^r for a natural number r (or $+\infty$) if f is r times continuously differentiable on U and $f = (f_1, \ldots, f_n) : U \subset \mathbb{R}^d \to \mathbb{R}^n$ is C^r if $f_1, \ldots, f_n$ are all C^r. For any $i \in \{1, \ldots, d\}$, denote $(\partial f/\partial x_i) = (\partial f_1/\partial x_i, \ldots, \partial f_n/\partial x_i)$. For any multi-index $\nu = (\nu_1, \ldots, \nu_d) \in \mathbb{N}^d$, map $f : U \to \mathbb{R}^n$ and $x = (x_1, \ldots, x_d) \in U$ we set:

$$|\nu| = \nu_1 + \ldots + \nu_d$$

$$\nu! = \nu_1! \cdots \nu_d!$$

$$x^\nu = \prod_{i=1}^{d} x_i^{\nu_i}$$

$$f^{(\nu)} = \frac{\partial^{|\nu|}}{\partial x_1^{\nu_1} \cdots \partial x_d^{\nu_d}} f$$

where by definition $0! = 1$, and $0^0 = 1$.

Definition 2.1 (C^r-norm). Suppose that $f : U \subset \mathbb{R}^d \to \mathbb{R}$ is C^r. Then we define the C^r-norm $|\cdot|_r$ of f as follows:

$$|f|_r = \sup_{x \in U} \sup_{\substack{|\nu| \leq r \\ \nu \in \mathbb{N}^d}} \frac{|f^{(\nu)}(x)|}{|\nu|!}.$$

Note that $|f|_r$ can be $+\infty$. We define the C^r-norm of a map $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ to be the maximum of the C^r-norms of the component functions $f_1, \ldots, f_n$.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

This is the norm used in [1]. In [4], they did not divide by $|\nu|!$. This yields equivalent norms but has an impact on the exponent of r in the main theorem (see Lemma 2.5).

Definition 2.2 (Mild functions). Suppose that $A, B > 0$, $C \geq 0$ are real numbers and $f : U \subset \mathbb{R}^d \to \mathbb{R}$. Then f is called (A, B, C)-mild if it is C^∞ and if for all $\nu \in \mathbb{N}^d$ and $x \in U$:

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

A map $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ is (A, B, C)-mild if all component functions $f_1, \ldots, f_n$ are (A, B, C)-mild. If $B \leq 1$, we simply say that f is (A, C)-mild and if we say that f is mild, then we mean that f is (A, C)-mild for some $A > 0$ and $C \geq 0$.

This definition is slightly different than the original one by Pila in [11] ([12] for the multivariate case), where the bounds on the derivatives for an (A, C)-mild function are given by $\nu!(A|\nu|^C)^{|\nu|}$. Now, due to the following inequalities:

$$\nu! \leq |\nu|! \leq d^{|\nu|}\nu!$$

$$|\nu|! \leq |\nu|^{|\nu|} \leq e^{|\nu|}|\nu|!$$

(where the number e appears as a result of Stirling's formula), one sees that the definitions of (A, C)-mild coincide in the following way: f is (A_1, C)-mild for some $A_1 > 0$ as in our definition if and only if f is (A_2, C)-mild for some $A_2 > 0$ as in [12]. The reason to adjust the definition is to keep the proof of Theorem 2.10 as simple as possible.

Following [4], we will use 'up to order r' versions of many definitions and theorems, for example, the following definition.

Definition 2.3 (Mild up to order r). Suppose that $A, B > 0$, $C \geq 0$ are real numbers, $r > 0$ an integer or $+\infty$ and $f : U \subset \mathbb{R}^d \to \mathbb{R}$. Then f is called (A, B, C)-mild up to order r if it is C^r and if for all $\nu \in \mathbb{N}^d$ with $|\nu| \leq r$ and $x \in U$:

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

A map $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ is (A, B, C)-mild up to order r if all component functions $f_1, \ldots, f_n$ are (A, B, C)-mild up to order r. If $B \leq 1$, we simply say that f is (A, C)-mild up to order r and if we say that f is mild up to order r, then we mean that f is (A, C)-mild up to order r for some $A > 0$ and $C \geq 0$.

Note (A, B, C)-mild up to order $+\infty$ agrees with (A, B, C)-mild. The following corollary is an immediate consequence of the definitions.

Corollary 2.4. *Suppose $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ is C^r. Then f has C^r-norm less than or equal to B if and only if f is $(1, B, 0)$-mild up to order r.*

For (A, C)-mild functions up to order r, one can bound the C^r-norm by 1 with an easy substitution. It will be used to count the numbers of maps in the proof of the main theorem. More precisely, we will construct a parameterization consisting of $(A, 0)$-mild maps up to order r.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

Lemma 2.5. *Suppose that $f : U \subset (0, 1)^d \to [-1, 1]$ is (A, C)-mild up to order r. Let $P = (P_1, \ldots, P_d)$ be any point in $(0, 1)^d$ and consider the map*

$$\psi : (x_1, \ldots, x_d) \mapsto \left(\frac{x_1}{Ar^C} + P_1, \ldots, \frac{x_d}{Ar^C} + P_d \right).$$

If $V = \psi^{-1}(U)$ then $(f \circ \psi) : V \to [-1, 1]$ has C^r-norm bounded by 1.

Proof. This is a direct calculation using the chain rule. □

This lemma is nearly identical to [4, Lemma 4.1.3]. The difference is the exponent of r, which is $C + 1$ in [4], because they use the following norm:

$$\sup_{x \in U} \sup_{\substack{|\nu| \leq r \\ \nu \in \mathbb{N}^d}} |f^{(\nu)}(x)|.$$

Since this norm is stronger, the number of maps in the C^r-parameterization is larger.

We conclude this section with the following class of examples which will show up in § 3. For a proof, we refer to [9, Proposition 2.2.10].

Lemma 2.6. *Suppose that $f : U \subset \mathbb{R}^d \to \mathbb{R}$ is analytic on an open neighbourhood of the topological closure $\overline{U}$ of U. Then f is $(A, B, 0)$-mild for some $A, B > 0$.*

2.1. Compositions of mild functions

In this section, we show that if f and g are (A_f, B_f, C)- and (A_g, B_g, C)-mild respectively, then the composition $f \circ g$ is (A, B, C)-mild for some explicit A and B. Of course, we will need a multivariate version of a formula for arbitrary derivatives of a composite function. This is known as the Faà di Bruno formula, which has first been proved in [5] and we reformulate it here. A proof can also be found in [9, Theorem 1.3.2].

Theorem 2.7. *Suppose that n is a positive integer, $V \subset \mathbb{R}^d$ and $U \subset \mathbb{R}^e$ are open, $f : V \to \mathbb{R}$, $g : U \to V$ and that f and g are C^n. For any $x \in U$ and $\nu \in \mathbb{N}^e$ with $|\nu| = n$ we have that*

$$(f \circ g)^{(\nu)}(x) = \sum_{1 \leq |\lambda| \leq n} f^{(\lambda)}(g(x)) \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{(g^{(l_j)}(x))^{k_j}}{k_j!(l_j!)^{|k_j|}}$$

where $p_s(\nu, \lambda)$ is the set consisting of all $k_1, \ldots, k_s \in \mathbb{N}^d$ with $|k_i| > 0$ and $l_1, \ldots, l_s \in \mathbb{N}^e$ with $0 \prec l_1 \prec \ldots \prec l_s$ such that:

$$\sum_{i=1}^{s} k_i = \lambda$$

and

$$\sum_{i=1}^{s} |k_i| l_i = \nu.$$

Here $l_i \prec l_{i+1}$ means that $|l_i| < |l_{i+1}|$ or, if $|l_i| = |l_{i+1}|$, then l_i comes lexicographically before l_{i+1}.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

This formula is the heart of all of the proofs of results on compositions of mild functions that will follow. In [4], Cluckers, Pila and Wilkie use this formula to deduce that the composition of mild functions is mild. However, they show this by roughly estimating the sum and hence they obtain no explicit formulas for A, B and C. In his paper [6], Gevrey introduces 'functions of class α' (see below), later known as Gevrey functions, and showed that the functions of class α with $\alpha \geq 1$ are closed under composition. In fact, Gevrey proved his result only where one of the functions involved has only one variable. The proof uses the Faà di Bruno formula in an essential way. Even though we refer to a more recent paper for the proof of this formula, the formula was already known for a long time. Because the result on Gevrey functions will immediately yield our result on mild functions and the technique is important for all other proofs, we will give a full proof of Gevrey's theorem for functions in arbitrary many variables, that is, the general version of what he proved.

Definition 2.8 (Gevrey functions). Suppose that $\alpha \geq 0$ and $f : U \subset \mathbb{R}^d \to \mathbb{R}$. We say that f is a Gevrey function of class α if it is C^∞ and for all $\nu \in \mathbb{N}^d$ and $x \in U$ we have:

$$\left|f^{(\nu)}(x)\right| \leq \left(M \frac{|\nu|!}{R^{|\nu|}}\right)^\alpha$$

for some M, $R > 0$. A map $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ is a Gevrey function of class α if all of its component functions are Gevrey functions of class α.

The following proposition, together with Theorem 2.10, motivates why we adjusted the definition of a mild function, as we explained below Definition 2.2.

Proposition 2.9. *Suppose that $f : U \subset \mathbb{R}^d \to \mathbb{R}$ is C^∞, then f is (A, B, C)-mild for some A, $B > 0$ if and only if it is a Gevrey function of class $C + 1$.*

Proof. If one sets $R = A^{-1/(C+1)}$ and $M = B$, then it follows immediately that the bounds on the derivatives are the same. □

Of course, one could define a Gevrey function of class α up to order r and the result above has an up to order r version and thus, so do the theorem and the subsequent corollary below.

Theorem 2.10 (Gevrey [6]). *Suppose that $f : V \subset \mathbb{R}^d \to \mathbb{R}$ and $g : U \subset \mathbb{R}^e \to V$ are Gevrey functions of class $\alpha \geq 1$, then $f \circ g$ is a Gevrey function of class α.*

Proof. The idea of this ingenious proof of Gevrey is as follows. For some specific F and G, we can explicitly compute both sides of the Faà di Bruno formula at some well chosen point since all derivatives of $F \circ G$ will be easy to compute. In turns out that the values of these derivatives at this point are exactly the bounds we have for the derivatives of our given maps f and g. Now let $x \in U$ be arbitrary. Using the Faà di Bruno formula and the triangle inequality to bound $|(f \circ g)^{(\nu)}(x)|$, we obtain the same sum of positive terms where the bounds on the derivatives of f and g occur. Moreover, this sum does not depend on x, since the bounds of a Gevrey function are uniform.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

To define the maps F and G, we need the constants that bound the derivatives of f and g. Since they're Gevrey functions of class α, we have:

$$\left|f^{(\nu)}(x)\right| \leq \left(M_f \frac{|\nu|!}{R_f^{|\nu|}}\right)^{\alpha}$$

and

$$\left|g_i^{(\nu)}(x)\right| \leq \left(M_g \frac{|\nu|!}{R_g^{|\nu|}}\right)^{\alpha}$$

for $i = 1, \ldots, d$. The maps that do the trick are the following F and $G = (G_1, \ldots, G_d)$:

$$F = \frac{M_f R_f}{(R_f + dM_g) - (x_1 + \ldots + x_d)}$$

and for $i = 1, \ldots, d$:

$$G_i = \frac{M_g R_g}{R_g - (x_1 + \ldots + x_e)}.$$

We first compute an arbitrary derivative of G_i for some $i \in \{1, \ldots, d\}$. Set $S(x) = x_1 + \ldots + x_e$. Then we have for all $j \in \{1, \ldots, e\}$ that $(\partial/\partial x_j)(S(x)) = 1$. Writing $G_i(x) = M_g R_g (R_g - S(x))^{-1}$ we easily obtain that for any $\lambda \in \mathbb{N}^e$:

$$G_i^{(\lambda)}(x) = (M_g R_g)|\lambda|!(R_g - S(x))^{-(1+|\lambda|)}.$$

Hence we get:

$$G_i^{(\lambda)}(0) = M_g \frac{|\lambda|!}{R_g^{|\lambda|}}. \tag{2.1}$$

Noticing that $G(0) = (M_g, \ldots, M_g)$, we see that for all $\lambda \in \mathbb{N}^d$ we obtain similarly:

$$F^{(\lambda)}(G(0)) = M_f \frac{|\lambda|!}{R_f^{|\lambda|}}. \tag{2.2}$$

Next, one checks that:

$$(F \circ G)(x) = \frac{M_f R_f (R_g - S(x))}{R_f R_g - (R_f + dM_g) S(x)}$$

and that for any $\lambda \in \mathbb{N}^e$ with $|\lambda| \geq 1$:

$$(F \circ G)^{(\lambda)}(x) = \frac{dM_f M_g R_f R_g}{R_f + dM_g} \frac{(R_f + dM_g)^{|\lambda|}|\lambda|!}{(R_f R_g - (R_f + dM_g)S(x))^{|\lambda|+1}}.$$

Finally, we get that:

$$(F \circ G)^{(\lambda)}(0) = M \frac{|\lambda|!}{R^{|\lambda|}} \tag{2.3}$$

with $M = \frac{dM_f M_g}{R_f + dM_g}$ and $R = \frac{R_f R_g}{R_f + dM_g}$. Hence if we plug in values as in (2.2) and (2.1) in the sum of Theorem 2.7, then we know that it is equal to (2.3). More precisely, we get

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

the following equalities:

$$M\frac{|\nu|!}{R^{|\nu|}} = (F\circ G)^{(\nu)}(0) = \sum_{1\le|\lambda|\le n} M_f \frac{|\lambda|!}{R_f^{|\lambda|}} \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{\left(M_g \frac{|l_j|!}{R_g^{|l_j|}}\right)^{|k_j|}}{k_j!(l_j!)^{|k_j|}}.$$

Now first suppose $\alpha = 1$. To compute an upper bound on $|(f\circ g)^{(\nu)}(x)|$, use the Faà di Bruno formula, the triangle inequality and finally the bounds on the derivatives of f and g to obtain the sum on the right-hand side of these equations. We now know it equals the desired form on the left-hand side. This finishes the proof if $\alpha = 1$. If $\alpha > 1$, you additionally use that $r^\alpha + s^\alpha \le (r+s)^\alpha$ $(r, s \ge 0)$, which yields the result. □

Combining this with Proposition 2.9, one deduces from this proof the following result on mild functions.

Corollary 2.11. *Suppose that $f : V \subset \mathbb{R}^d \to \mathbb{R}$ is (A_f, B_f, C)-mild and that $g : U \subset \mathbb{R}^e \to V$ is (A_g, B_g, C)-mild. Then $f\circ g$ is (A, B, C)-mild where:*

$$A = A_f A_g (A_f^{-1/(C+1)} + dB_g)^{C+1},$$
$$B = \frac{dB_f B_g}{A_f^{-1/(C+1)} + dB_g} < B_f.$$

In particular, if f is $(A_f, 0)$-mild and g is $(A_g, 0)$-mild then $f\circ g$ is $(A, 0)$-mild with:

$$A = A_g(dA_f + 1).$$

2.2. Weakly mild functions and power substitutions

Many nice functions are not mild. For instance, if f is analytic on U, this still is not sufficient: consider, for example, the function $x \mapsto x^{1/2}$ on $(0, 1)$. So one can not weaken the conditions of Lemma 2.6. Maps of this form will be crucial later on. They are also examples of the next definition, a variant on mild functions that has been introduced by Cluckers, Pila and Wilkie in [4]. By a reparameterization, which we will call a power substitution, we can make these functions mild up to any order if they satisfy an additional condition.

Definition 2.12 (Weakly mild functions). Suppose that $A, B > 0$, $C \ge 0$ are real numbers, $r > 0$ an integer or $+\infty$ and $f : U \subset (0, 1)^d \to \mathbb{R}$. Then f is called weakly (A, B, C)-mild up to order r if it is C^r and if for all $\nu \in \mathbb{N}^d$ with $|\nu| \le r$ and $x \in U$:

$$\left|f^{(\nu)}(x)\right| \le \frac{B^{C+1}A^{|\nu|}|\nu|!^{C+1}}{x^\nu}.$$

A map $f = (f_1, \ldots, f_n) : U \subset (0, 1)^d \to \mathbb{R}^n$ is weakly (A, B, C)-mild up to order r if all component functions $f_1, \ldots, f_n$ are weakly (A, B, C)-mild up to order r. If $B \le 1$, we simply say that f is weakly (A, C)-mild up to order r and if we say that f is weakly mild

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

up to order r, then we mean that f is weakly (A, C)-mild up to order r for some $A > 0$ and $C \geq 0$. Finally, if $r = +\infty$, we might just say weakly (A, B, C)-mild or weakly mild.

Using the technique of the proof of Theorem 2.10, we can make the result of [4, Proposition 4.1.5] more explicit in the proposition below. In particular, we have an explicit formula for $\tilde{A}$ and the constant C is preserved.

Proposition 2.13 (Power substitution). *Let $A \geq 1$, $B > 0$ and $C \geq 0$ be real numbers and suppose that $f : U \subset (0, 1)^d \to \mathbb{R}$ is a map such that f and all first-order derivatives of f are weakly (A, B, C)-mild. Let $r > 0$ be an integer and define $\phi : (0, 1)^d \to (0, 1)^d$ by:*

$$(x_1, \ldots, x_d) \mapsto (x_1^{n_1}, \ldots, x_d^{n_d})$$

with $n_i \geq r$ for all $i \in \{1, \ldots, d\}$, and denote $V = \phi^{-1}(U)$. Then the map $f \circ \phi : V \to \mathbb{R}$ is $(\tilde{A}, B, C)$-mild up to order r with $\tilde{A} = NA(d+1)^{C+1}$, with $N = \max(n_1, \ldots, n_d)$.

Remark. If f satisfies the conditions above, but with $0 < A < 1$ instead, one can apply the proposition after enlarging A to 1.

Proof. We first take a look at the derivatives of ϕ. Take any component function ϕ_i of ϕ. Because it only depends on x_i, we just have to bound the following derivatives:

$$\left(\frac{\partial}{\partial x_i}\right)^k \phi_i(x) = n_i \cdots (n_i - (k+1)) x_i^{n_i - k} \leq N^k x_i^{n_i - k} \leq N^k.$$

Thus, ϕ is $(N, 0)$-mild. However, we will use the sharper bound (involving the power $n_i - k$ of x_i) to deal with the derivatives of f. By Theorem 2.7 and the triangle inequality, we have for every $x \in V$ that

$$|(f \circ \phi)^{(\nu)}(x)| \leq \sum_{1 \leq |\lambda| \leq n} |f^{(\lambda)}(\phi(x))| \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{|(\phi^{(l_j)}(x))^{k_j}|}{k_j!(l_j!)^{|k_j|}}.$$

Consider one term in this sum, so we also have a fixed $\lambda \in \mathbb{N}^d$ with $1 \leq |\lambda| \leq |\nu|$, $s \in \{1, \ldots, n\}$ and $(k_1, \ldots, k_s, l_1, \ldots, l_s) \in p_s(\nu, \lambda)$. Then, for any $i \in \{1, \ldots, d\}$ and $j \in \{1, \ldots, s\}$, we have:

$$\left|\left(\phi_i^{(l_j)}(x)\right)^{k_{j,i}}\right| \leq N^{l_{j,i} k_{j,i}} x_i^{(n_i - l_{j,i}) k_{j,i}}$$

and thus the power of x_i coming from the product over j is $\sum_{j=1}^{s} (n_i - l_{j,i}) k_{j,i}$.

We now compute the 'negative' contribution to the powers of x_i in this term coming from $f^{(\lambda)}(\phi(x))$. Write $\lambda = \lambda' + \beta$ for some $\beta \in \mathbb{N}^d$ with $|\beta| = 1$ and let $k \in \{1, \ldots, d\}$ be the unique index such that $\beta_k = 1$. For any choice of β, we have the following (thus, we may pick some particular β, equivalently, make a choice for k, later):

$$\left|f^{(\lambda)}(\phi(x))\right| = \left|(f^{(\beta)})^{(\lambda')}(\phi(x))\right| \leq B^{C+1} A^{|\lambda'|} |\lambda'|!^{C+1} \frac{1}{\phi(x)^{\lambda'}} \leq B^{C+1} A^{|\lambda|} |\lambda|!^{C+1} \frac{1}{\phi(x)^{\lambda'}}.$$

Hence the power of x_i coming from $f^{(\lambda)}(\phi(x))$ in the term is equal to $-n_i \lambda_i'$. We can now compute the total power of x_i occurring in the term. The total power of x_i when $i \neq k$

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

is (then $\lambda_i' = \lambda_i$):

$$\sum_{j=1}^{s} \big((n_i - l_{j,i})k_{j,i}\big) - n_i\lambda_i = \sum_{j=1}^{s} \big((n_i - l_{j,i})k_{j,i}\big) - n_i \sum_{j=1}^{s} k_{j,i} = -\sum_{j=1}^{s} k_{j,i}l_{j,i}.$$

When $i = k$ (then $\lambda_k' = \lambda_k - 1$) one computes in the same way that the power of x_k is $n_k - \sum_{j=1}^{s} k_{j,k}l_{j,k}$. We now pick k such that

$$x_k = \min_{i:\lambda_i \neq 0} x_i.$$

Then we have for all $i \in \{1, \ldots, d\}$:

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

and we can bound the product of all x_i and their powers by:

$$x_k^{n_k - \sum_{i=1}^{d}\sum_{j=1}^{s} |k_j|l_{j,i}} = x_k^{n_k - \sum_{i=1}^{d} \nu_i} = x_k^{n_k - |\nu|} \leq 1$$

since $n_k \geq r$ and $|\nu| \leq r$. So we can bound this term independent of x.

Putting everything together we obtain that:

$$\begin{aligned} |(f \circ \phi)^{(\nu)}(x)| &\leq \sum_{1 \leq |\lambda| \leq n} B^{C+1}A^{|\lambda|}|\lambda|!^{C+1} \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{N^{|l_j||k_j|}}{k_j!(l_j!)^{|k_j|}} \\ &\leq \sum_{1 \leq |\lambda| \leq n} B^{C+1}A^{|\lambda|}|\lambda|!^{C+1} \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{(N^{|l_j|}|l_j|!^{C+1})^{|k_j|}}{k_j!(l_j!)^{|k_j|}}. \end{aligned}$$

This is the same upper bound as the one that one would find for the composition of an (A, B, C)-mild function up to order r and an $(N, 1, C)$-mild function up to order r using Proposition 2.9 and then the proof of Theorem 2.10. Therefore, one can use the formula of Corollary 2.11 to find the expression for $\tilde{A}$. □

The extra condition that all first-order derivatives are weakly mild is crucial and cannot be omitted. For instance, consider again the map $x \mapsto x^{1/2}$. Then we see that by composing with the power map $x \mapsto x^3$, we do not obtain a map that is mild up to order 3. Of course, using the power map $x \mapsto x^2$, it becomes mild (up to order $+\infty$). This observation will be the key to slightly improve the main theorem in § 3.

Even though we can say something about the derivatives of a composition of weakly mild functions, it is not necessarily weakly mild. We do have the following result.

Proposition 2.14. *Suppose that $f : V \subset \mathbb{R}^d \to \mathbb{R}$ is (A_f, B_f, C)-mild up to order r and that $g : U \subset (0, 1)^e \to V$ is weakly (A_g, B_g, C)-mild up to order r. Then $f \circ g$ is weakly (A, B, C)-mild up to order r, where A and B are as in Corollary 2.11.*

Proof. The strategy of the proof is completely the same as in Proposition 2.13: one just checks that in each term one obtains exactly a factor $1/x^{\nu}$ when bounding $(f \circ g)^{(\nu)}(x)$ using Theorem 2.7. □

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

We conclude this section with a result on the product of (weakly) mild functions. In this proposition, one can replace $(0, 1)^d$ by $\mathbb{R}^d$ in the case of mild functions.

Proposition 2.15. *Suppose that $f_1, \ldots, f_l : U \subset (0, 1)^d \to \mathbb{R}$ are (weakly) (A, B, C)-mild up to order r for some $A, B > 0$ and $C \geq 0$. Then the product $f_1 \cdots f_l$ is (weakly) (lA, B^l, C)-mild up to order r.*

Proof. As in [12, Proposition 2.6], we have that:

$$|(f_1 \cdots f_l)^{(\nu)}(x)| \leq \sum_{\nu_1+\ldots+\nu_l=\nu} \mathrm{Ch}(\nu_1, \ldots, \nu_l) \prod_{i=1}^{l} |f_i^{(\nu_i)}(x)|,$$

where the sum runs over all $\nu_1, \ldots, \nu_l \in \mathbb{N}^d$ such that $\nu_1 + \ldots + \nu_l = \nu$ and $\mathrm{Ch}(\nu_1, \ldots, \nu_l)$ is a constant depending on $\nu_1, \ldots, \nu_l$. If $f_1, \ldots, f_l$ are (A, B, C)-mild, we find:

$$\prod_{i=1}^{l} |f_i^{(\nu_i)}(x)| \leq \prod_{i=1}^{l} B^{C+1} A^{|\nu_i|} |\nu_i|!^{C+1} \leq B^{l(C+1)} A^{|\nu|} |\nu|!^{C+1}.$$

If $f_1, \ldots, f_l$ are weakly (A, B, C)-mild, one obtains an extra factor $1/x^\nu$ as desired. Finally, in § 3.2 of [4], it is shown that $\sum_{\nu_1+\ldots+\nu_l=\nu} \mathrm{Ch}(\nu_1, \ldots, \nu_l) \leq l^{|\nu|}$. □

3. The C^r-parameterization theorem

In this section, we give a precise definition of the structure wherein the family $X_T, T \subset \mathbb{R}^k$, of our main theorem should be definable. This structure is o-minimal, a powerful tool of model theory, and thus enables us to use the cell decomposition theorem, see [14]. Furthermore, in [10], Miller has shown a preparation theorem for definable functions in this structure. Combining these results, we obtain a very strong parameterization theorem from which we will easily deduce the main theorem with the results of the last section.

We start with the necessary definitions of model theory and the result of Miller. In his paper definability (in $\mathbb{R}^n$) is with respect to the following language. Let $\mathcal{L}_r = \{+, -, \cdot, <, 0, 1\}$ be the language of ordered rings and expand it with a symbol for each of the following functions:

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

where $f : U \to \mathbb{R}$ is a real analytic function on an open neighbourhood U of $[-1, 1]^n$. Denote this language $\mathcal{L}_{\mathrm{an}}$. If X is $\mathcal{L}_{\mathrm{an}}$-definable in $\mathbb{R}^n$, then it is called (globally) subanalytic. Remember that a function is definable if its graph is a definable set. Next, if we expand $\mathcal{L}_{\mathrm{an}}$ with a symbol for all functions

$$x \mapsto \begin{cases} x^r & \text{if } x > 0 \\ 0 & \text{if } x \leq 0, \end{cases}$$

for $r \in \mathbb{R}$, we obtain the language $\mathcal{L}_{\mathrm{an}}^{\mathrm{pow}}$ and the corresponding structure $\mathbb{R}_{\mathrm{an}}^{\mathrm{pow}}$. If X is $\mathcal{L}_{\mathrm{an}}^{\mathrm{pow}}$-definable in $\mathbb{R}^n$, we say that X is power-subanalytic. In [10], Miller considers the

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

following reducts of $\mathbb{R}_{\text{an}}^{\text{pow}}$. Let $\mathcal{F}$ be a Weierstrass system: a collection $\mathcal{F}_n$ of real analytic functions $\mathbb{R}^n \to \mathbb{R}$ for each $n \in \mathbb{N}$, containing the polynomials in n variables, for all n which is closed under several operations such as the ring operations, composition and Weierstrass preparation (for a precise definition, see [10]). If one adds to the language $\mathcal{L}_r$ symbols for each $\tilde{f}$, where $f \in \mathcal{F}$, one obtains the language $\mathcal{L}_{\mathcal{F}}$. For any Weierstrass system $\mathcal{F}$, one can consider its field of exponents:

$$K = \{r \in \mathbb{R} \mid x \mapsto (1+x)^r \in \mathcal{F}\}.$$

(Miller shows in his paper that this is indeed a field.) For the smallest Weierstrass system, this field is $\mathbb{Q}$ and thus any field of exponents K contains $\mathbb{Q}$. For the largest Weierstrass system, i.e. consisting of all analytic functions, this field is $\mathbb{R}$. Denote $\mathcal{L}_{\mathcal{F}}^K$ the language obtained by adding to $\mathcal{L}_{\mathcal{F}}$ a symbol for all power maps $x \mapsto x^r$ for $r \in K$ (as above), where K can be any subfield of $\mathbb{R}$. From now on, definability is with respect to the language $\mathcal{L}_{\mathcal{F}}^K$, where K is a subfield of the field of exponents of $\mathcal{F}$. The corresponding structure $\mathbb{R}_{\mathcal{F}}^K$ is o-minimal.

Definition 3.1 (Cell). A cell in $\mathbb{R}^m$ is a set of the following form:

$$\{(x_1, \ldots, x_m) \in \mathbb{R}^m \mid \alpha_i(x_{<i}) \,\square_{i1}\, x_i \,\square_{i2}\, \beta_i(x_{<i}), i = 1, \ldots, m\}$$

where $x_{<i} = (x_1, \ldots, x_{i-1})$, α_i and β_i are continuous definable functions, with $\alpha_i < \beta_i$, and $\square$ is 'no condition' or the conditions $<$ or $=$, where $\square_{i2}$ is $<$ or no condition if $\square_{i1}$ is equality. Obviously, a cell is open in $\mathbb{R}^m$ if and only if $\square_{i1}$ is not equality for all i. In that case we call α_i and β_i the walls of x_i.

Suppose that C is a cell in $\mathbb{R}^n$. Then, up to reordering the variables if necessary, we may suppose that for $i = 1, \ldots, m$ $(m \leq n)$ the condition $\square_{i1}$ is inequality and for the last $n - m$ variables $\square_{i1}$ is equality. In this way, we see that any cell in $\mathbb{R}^n$ corresponds to the graph of a definable function $f : U \subset \mathbb{R}^m \to \mathbb{R}^{n-m}$. Combining this with the fact that an open cell in $\mathbb{R}^n$ is definably homeomorphic to $(0, 1)^n$, we get that a cell in $\mathbb{R}^n$ is the same as the graph of a definable function $(0, 1)^m \to \mathbb{R}^{n-m}$. The number m is the dimension of the cell and thus is nothing more than counting how many times $\square_{i1}$ is inequality. We will do many manipulations with these functions and thus end up in general with definable maps $U \subset (0, 1)^m \to \mathbb{R}^{n-m}$, where U is an open cell in $(0, 1)^m$. To conclude, because any definable set in an o-minimal structure is a finite union of cells, it suffices to prove the main theorem, where we assume X_T to be a definable family of m-dimensional subsets of $[-1, 1]^n$, with $T \subset \mathbb{R}^k$ the set of parameters, in the case that X_T is the graph of a definable map $f : T \times (0, 1)^m \to [-1, 1]^{n-m}$. (If X does not depend on parameters, $T = \mathbb{R}^0$, which is by definition a set with only one point.) Later on, our map will be of the form $f : C \subset T \times (0, 1)^m \to [-1, 1]^{n-m}$, where the fibres C_t of C are open cells in $(0, 1)^m$ for any $t \in T$ and we will denote f_t for the map $C_t \to [-1, 1]^{n-m}$ given by $x \mapsto f(t, x)$.

In this section, T is a definable subset of $\mathbb{R}^k$ and we will often use the following notation. Suppose $U \subset \mathbb{R}^m$. Then we denote by $\overline{U}$ the topological closure of U in $\mathbb{R}^m$ endowed with the standard topology. Furthermore, for any $i \in \{2, \ldots, m\}$, $\pi_{<i} : T \times \mathbb{R}^m \to T \times \mathbb{R}^{i-1}$

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

denotes the map

$$(t, x_1, \ldots, x_m) \mapsto (t, x_1, \ldots, x_{i-1})$$

and $\pi_{<1}(t, x) = t$. The following two definitions coincide with [4, Definition 4.4.1].

Definition 3.2 (Centre of a cell). Suppose that C is a cell in $\mathbb{R}^m$. A definable continuous map $\theta : \pi_{<m}(C) \to \mathbb{R}$ is called a centre for C if its graph and $\overline{C}$ are disjoint or if its graph is contained in $\overline{C} \setminus C$, and θ is identically zero or $\theta(x_{<m}) \sim x_m$ for all $x \in C$, that is, there exists some $\epsilon \in (0, 1)$ such that for all $x \in C$:

$$\epsilon x_m \leq \theta(x_{<m}) \leq \epsilon^{-1} x_m.$$

The next definition motivates why θ is called the centre of a cell.

Definition 3.3 (Prepared with centre). A bounded definable function $f : C \to \mathbb{R}$, where C is a cell in $\mathbb{R}^m$, is called prepared with centre θ if θ is a centre for C and f can be written in the following way:

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

where $b : C \to \mathbb{R}^N$ (for some $N \in \mathbb{N}$) is bounded, b_j is a component function of b and the component functions b_i, $i \in \{1, \ldots, N\}$, of b are of the following form:

$$a_i(x_{<m}) |x_m - \theta(x_{<m})|^{r_i}$$

with $r_i \in K$ and $a_i : \pi_{<m}(C) \to \mathbb{R}$ definable and, finally, F is a non-vanishing analytic function on an open neighbourhood of $\overline{b(C)}$. We call b the associated bounded range map of f. A map $f : C \to \mathbb{R}^n$ is prepared with centre θ if all of its component functions are prepared with centre θ and moreover have the same bounded range map b.

We now state the preparation theorem as in [4, Proposition 4.4.2], which follows from the main theorem of [10, Lemma 3.7 and Lemma 4.4].

Theorem 3.4. *Suppose that $f : X \subset \mathbb{R}^m \to \mathbb{R}^n$ is a bounded definable map. Then there exists a finite partition of X into cells C_i with centre θ_i such that the restriction $f|_{C_i}$ of f to C_i is prepared with centre θ_i.*

We follow [4, § 4], but in our proofs, we avoid transforming parameters, i.e. we prove everything uniformly. The main reason we do this is to clearly point out that the power substitution in the proof of [4, Theorem 2.1.3] does not require a power substitution in the parameter variables. The idea here is that we stop the inductive proof of [4, Theorem 4.3.1] when we reach the parameter variables (see Theorem 3.11) and then apply a (uniform) power substitution (Theorem 3.10). To achieve this, we slightly adjusted some definitions, namely the notion of an a–b–m function in [4], which is here replaced by 'prepared in x' (see below). The main theorem will be deduced from Theorem 3.11. In this theorem, we parameterize X_T with maps that satisfy the conditions of Proposition 2.13. Finally, we conclude by Lemma 2.5.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

Definition 3.5. Suppose that C is a cell in $\mathbb{R}^k \times (0,1)^m$ and denote an element of C as a tuple (t, x). A bounded definable function $f : C \to \mathbb{R}$ is prepared in x if it can be written as:

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

where $b : C \to \mathbb{R}^N$ (for some $N \in \mathbb{N}$) is bounded, b_j is a component function of b, any component function b_i, $i \in \{1, \ldots, N\}$, of b is of the form:

$$b_i(t,x) = a_i(t)x^{\mu_i} = a_i(t)\prod_{l=1}^{m} x_l^{\mu_{i,l}}$$

for some definable function $a_i : \pi_{<k+1}(C) \to \mathbb{R}$, $\mu_i \in K^m$ and F is analytic and non-vanishing on an open neighbourhood of $\overline{b(C)}$. We call b the associated bounded monomial map of f. A bounded definable map $f : C \to \mathbb{R}^n$ is prepared in x if all of its component functions are prepared in x and have the same associated bounded monomial map b.

Note that k can be zero. In that case, up to a scalar a, the associated bounded monomial map b is a monomial and the definition of prepared in x coincides with the definition of an analytic-bounded-monomial map of [4]. We now show some properties of these maps that relate to the previous section.

Lemma 3.6. *Let $b : U \subset (0,1)^m \to \mathbb{R}$, where U is open, be given by $b(x) = ax^\mu$ for some $a \in \mathbb{R}$ and some $\mu \in \mathbb{R}^m$. If b is bounded, then b is weakly $(M, B, 0)$-mild, where $M = \max(1, |\mu_1|, \ldots, |\mu_m|)$ and $B = \sup_U b(x)$.*

Proof. Clearly, for any $\nu \in \mathbb{N}^m$, by the form of b, we have:

$$b^{(\nu)}(x) = c(\nu,\mu)\frac{b(x)}{x^\nu},$$

for some constant $c(\nu, \mu)$ depending on ν and μ, which remains to be bounded as desired. It is easy to see that computing the next order derivative would multiply $c(\nu, \mu)$ with a factor that is at most $M + |\nu|$ in absolute value. Therefore, we find that:

$$|c(\nu,\mu)| \leq M(M+1)\cdots(M+|\nu|-1) \leq M^{|\nu|}|\nu|!. \qquad \square$$

Lemma 3.7. *Suppose that $f : U \subset T \times (0,1)^m \to \mathbb{R}$ is prepared in x with associated bounded monomial map b, where for any $t \in T$ the fibre U_t is open. Then there exist $A, B > 0$ such that for any $t \in T$ the map f_t is weakly $(A, B, 0)$-mild. Moreover, if the C^1-norm of the associated bounded monomial map b_t of f_t is bounded independently of t, then there exist $A, B > 0$ such that for any $\beta \in \mathbb{N}^m$ with $|\beta| \leq 1$ and $t \in T$ the map $f_t^{(\beta)}$ is weakly $(A, B, 0)$-mild.*

Proof. Since f is prepared in x, we have that $f(t, x) = b_j(t, x)F(b(t, x))$ as in Definition 3.5. By Lemma 3.6, there exist $A, B > 0$, such that for any $t \in T$ the map b_t is weakly $(A, B, 0)$-mild. The constant A does not depend on t and, since b is bounded, we may suppose B does not depend on t. Finally, by Proposition 2.14 and Lemma 2.6, for any $t \in T$ we have that $F \circ b_t$ is weakly $(A, B, 0)$-mild for some $A, B > 0$ that do not

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

depend on t and we can conclude by Proposition 2.15 that for any $t \in T$, f_t is weakly $(A, B, 0)$-mild for some $A, B > 0$ that do not depend on t.

Now fix any $t \in T$ and $1 \leq i \leq m$. By the product rule, we have

$$\frac{\partial f_t}{\partial x_i} = \frac{\partial b_{j,t}}{\partial x_i}(F \circ b_t) + \sum_{l=1}^{N} b_{j,t}\left(\frac{\partial F}{\partial b_{l,t}} \circ b_t\right)\frac{\partial b_{l,t}}{\partial x_i}.$$

Since we can bound the derivatives of b_t independently of t, this is for any t a sum of weakly $(A, B, 0)$-mild functions for some $A, B > 0$ independent of t and thus is also weakly $(A, B, 0)$-mild (for possibly larger $A > 0$ and $B > 0$). □

This lemma motivates the following definition.

Definition 3.8. Suppose that $f : U \subset \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n$ is a family of functions. If there exists a B such that for all $t \in \pi_{<1}(U)$ and $x \in U_t$ we have that $|f_t(x)| < B$, we say that f is bounded in x. More generally, if r is a natural number, we say f is C^r-bounded in x if there exists a B such that for any $t \in \pi_{<1}(U)$ the map f_t is C^r and for any $x \in U_t$ the C^r-norm of f_t is bounded by B.

As an example, let $T = (0, 1)$ and consider the family of maps $f_T : (t, x) \mapsto t^2/x$ on the cell given by

$$0 < t < 1,$$
$$t < x < 1.$$

Then for a fixed $t \in T$, the map f_t is in fact mild, but the upper bound depends on t. However, one sees that f is C^1-bounded in x.

Lemma 3.7 shows that if f is prepared in x such that its associated bounded monomial map b is C^1-bounded in x, then we can apply Proposition 2.13 to f. The following result, which we need in the proof of Proposition 3.10, is a more refined version of Proposition 2.13 for the bounded monomial map b. The technique of the proof is similar, but will use the particular form of b.

Proposition 3.9. *Let $b : U \subset (0, 1)^m \to \mathbb{R}$, where U is open, be given by $b(x) = ax^\mu$ for some $a \in \mathbb{R}$ and $\mu \in \mathbb{R}^m$. Suppose that b is C^1-bounded. Let $r > 0$ be an integer and $\phi : (0, 1)^m \to (0, 1)^m$ be the map given by:*

$$(x_1, \ldots, x_m) \mapsto (x_1^{n_1}, \ldots, x_m^{n_m}),$$

where for all $i \in \{1, \ldots, m\}$, we have that $n_i \geq r^l$ for some integer $l \geq 1$ and set $V = \phi^{-1}(U)$. Then $(b \circ \phi)^{1/r^{l-1}} : V \to \mathbb{R}$ is $(\max(n_1, \ldots, n_m)A, B, 0)$-mild up to order r, for some $A, B > 0$ depending on b only.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

Proof. For simplicity, we assume that $a = 1$. We have that $(b \circ \phi)(x) = x^{N\mu}$, where $N\mu = (n_i\mu_i)_i$. We apply Theorem 2.7 to the composition $(b \circ \phi)^{1/r^{l-1}} = b^{1/r^{l-1}} \circ \phi$ to get:

$$\left|(b^{1/r^{l-1}} \circ \phi)^{(\nu)}(x)\right| \leq \sum_{1 \leq |\lambda| \leq n} \left|(b^{1/r^{l-1}})^{(\lambda)}(\phi(x))\right| \sum_{s=1}^{n} \sum_{p_s(\nu,\lambda)} \nu! \prod_{j=1}^{s} \frac{|(\phi^{(l_j)}(x))^{k_j}|}{k_j!(l_j!)^{|k_j|}}$$

for any $x \in V$. We again compute the power of x_i in a fixed term of this sum as in the proof of Proposition 2.13. The product of the $|(\phi^{(l_j)}(x))^{k_j}|$ gives a power $\sum_{j=1}^{s}(n_i - l_{j,i})k_{j,i}$. Using the notation of Lemma 3.6, we have that

$$(b^{1/r^{l-1}})^{(\lambda)}(\phi(x)) = c(\lambda, r^{-(l-1)}\mu)\frac{x^{(1/r^{(l-1)})N\mu}}{x^{N\lambda}},$$

where $N\lambda = (n_i\lambda_i)_i$ and $c(\lambda, r^{-(l-1)}\mu)$ is a constant depending on λ, r, l and μ. We have that $x_i^{\sum_{j=1}^{s}((n_i - l_{j,i})k_{j,i}) - n_i\lambda_i} = x_i^{-\sum_{j=1}^{s} l_{j,i}k_{j,i}}$. Let $k \in \{1, \ldots, m\}$ be such that $x_k = \min_i\{x_i \mid \lambda_i \neq 0\}$. We may then assume that $\mu_k \neq 0$ for the computations below, since if $\mu_k = 0$, the partial derivative of $b^{1/r^{l-1}}$ with respect to λ is zero since $\lambda_k \neq 0$. Then we bound the product over all x_i in terms of x_k:

$$\prod_{i=1}^{m} x_i^{-\sum_{j=1}^{s} l_{j,i}k_{j,i}} \leq x_k^{-|\nu|}.$$

Finally, assume $|\nu| \leq r$. Using that $n_k \geq r^l$, we further bound:

$$\frac{x^{(1/r^{(l-1)})N\mu}}{x_k^{|\nu|}} \leq \frac{x^{(1/r^{(l-1)})N\mu}}{x_k^{r}} = \left(\frac{x^{N\mu}}{x_k^{r^l}}\right)^{1/r^{l-1}} \leq \left(\frac{x^{N\mu}}{x_k^{n_k}}\right)^{1/r^{l-1}}.$$

Finally, observe that

$$\frac{x^{N\mu}}{x_k^{n_k}} = \frac{1}{\mu_k}\left(\frac{\partial}{\partial x_k}b\right)(\phi(x)),$$

which is bounded since $\phi(x) \in U$ and b is C^1-bounded on U. This bound only depends on b and the r^{l-1}-th root of this bound can be bounded by some constant that only depends on b, not on r or l. As in the proof of Lemma 3.6, we can bound $|c(\lambda, r^{-(l-1)}\mu)|$ by $M^{|\lambda|}|\lambda|!$, where $M > 0$ can be further bounded to only depend on b. One now finishes the proof similarly as the (end of) the proof of Proposition 2.13. □

The next proposition is one of the two main ingredients for the proof of the main theorem.

Proposition 3.10. *Suppose that $f : C \subset T \times (0, 1)^m \to [-1, 1]$ is prepared in x such that the associated bounded monomial b is C^1-bounded in x. Suppose moreover that for any $t \in T$, the cell C_t is open and that the walls of C are also prepared in x with*

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

associated bounded monomial maps that are C^1-bounded in x. Let $r > 0$ be any integer and consider the map $\phi_r : T \times (0, 1)^m \to T \times (0, 1)^m$ given by

$$\phi_r(t, x) = (t, x_1^{r^m}, x_2^{r^{m-1}}, \ldots, x_m^r).$$

Then $C_r = \phi_r^{-1}(C)$ is a cell and there exists an $A > 0$ such that for any $t \in T$ the walls of the open cell $C_{r,t}$ of x_1 are $(A, 0)$-mild, for $i = 2, \ldots, m$, the walls of x_i are $(r^m A, 0)$-mild up to order r and $(f \circ \phi_r)_t : C_{r,t} \to [-1, 1]$ is $(r^m A, 0)$-mild up to order r.

Proof. Fix some $t \in T$ and consider a wall of the cell C_t, say α_i of x_i. If $i = 1$, the walls are constant and there is nothing to show, so suppose $i \geq 2$. Then $\alpha_i(x_{<i}) = b_{j,t}(x_{<i})F(b_t(x_{<i}))$ for some b and F as in Definition 3.5, since α_i is prepared in x. It follows that the corresponding wall of x_i in the cell $C_{r,t}$ is given by:

$$\alpha_{r,t}(x_1, \ldots, x_{i-1}) = \sqrt[r^{m-i+1}]{\alpha_i(x_1^{r^m}, \ldots, x_{i-1}^{r^{m-i+2}})}.$$

By a uniform version of Proposition 3.9, we have that the function $\sqrt[r^{m-i+1}]{b_{j,t}(x_1^{r^m}, \ldots, x_{i-1}^{r^{m-i+2}})}$ is $(r^m A_1, B_1, 0)$-mild up to order r for some A_1, $B_1 > 0$ independent of t and r. Indeed, A_1 does not depend on t since it only depends on the exponent of x in b_j, which does not depend on t by Definition 3.5. To show that B_1 does not depend on t, one should dismiss the assumption that $a = 1$ in the proof of Proposition 3.9, but can later use that b_j is C^1-bounded in x to find B_1 independent of t (and also of r and l).

Since F is a non-vanishing analytic function on some open neighbourhood of $\mathrm{Im}(b_t)$, we may suppose that there is some $S > 1$ such that $\mathrm{Im}(F) \subset (1/S, S)$. On this domain, the function $\sqrt[r^{m-i+1}]{}$ is $(S, S, 0)$-mild. Hence, $\sqrt[r^{m-i+1}]{} \circ F$ is an $(A_F, B_F, 0)$-mild map for some A_F, $B_F > 0$ by Lemma 2.6 and Corollary 2.11, where A_F and B_F only depend on F, not on t or r. Next, $b_t(x_1^{r^m}, \ldots, x_{i-1}^{r^{m-i+2}})$ is $(r^m A', B', 0)$-mild up to order r by Proposition 2.13 for some $A', B' > 0$ depending only on b, not on t or r. Applying the up to order r version of Corollary 2.11, we see that $\sqrt[r^{m-i+1}]{F(b_t(x_1^{r^m}, \ldots, x_{i-1}^{r^{m-i+2}}))}$ is $(r^m A_2, B_2, 0)$-mild up to order r for some A_2, $B_2 > 0$ depending only on α_i, not on t or r.

By Proposition 2.15, we see that $\alpha_{r,t}$ is $(r^m A, B, 0)$-mild up to order r for some A, $B > 0$ independent of t and r. Since $\mathrm{Im}(\alpha_{r,t}) \subset (0, 1)$, after possibly enlarging A, we can conclude that $\alpha_{r,t}$ is $(r^m A, 0)$-mild up to order r for some $A > 0$ independent of t and r. Finally, by Proposition 2.13 and Lemma 3.7, the function $(f \circ \phi_r)_t$ is $(r^m A'', B'', 0)$-mild up to order r for some A'', $B'' > 0$ independent of t and r. Since $|f| \leq 1$, after possibly enlarging A'', we have that $(f \circ \phi_r)_t$ is $(r^m A'', 0)$-mild up to order r. We conclude after redefining A to be the maximum over all A corresponding to the mildness of the walls and A'' of the mildness of $(f \circ \phi_r)_t$. □

We will now show one can obtain the conditions of this proposition. It is a parameterization result, a slight modification of [4, Theorem 4.3.1], that will allow us to easily deduce the main theorem.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

Theorem 3.11. *Suppose that X_T is the graph of a definable function $\varphi : C \subset T \times (0, 1)^m \to [-1, 1]^{n-m}$, where C is open in $T \times (0, 1)^m$. Then there exist finitely many definable maps*

$$f_l : C_l \to X_T$$

such that:

1. $\bigcup_l Im(f_l) = X_T$,
2. *for each l, C_l is an open cell in $T_l \times (0, 1)^m$, where T_l is a cell contained in T,*
3. *for each l and for any $(t, x) \in C_l$, $f_l(t, x) \in X_t$, thus f_l is a family of maps $C_{l,t} \to X_t$ with $C_{l,t}$ open in $(0, 1)^m$,*
4. *each f_l is prepared in x and the associated bounded monomial map of f_l is C^1-bounded in x,*
5. *the walls of all C_l are prepared in x and their associated bounded monomial maps are C^1-bounded in x.*

Proof. The proof uses induction on m. The case $m = 0$ is trivial. By o-minimality, we may assume that there exist finitely many $f_l : C_l \subset T \times (0, 1)^m \to X_T$, satisfying properties 1, 2 and 3. For the rest of the proof, we focus on one such map f_l. Next, we apply Theorem 3.4 to f_l such that we may suppose that f_l is prepared on C_l with centre θ_l. Note that some of the cells obtained by applying (3.4) might not be open, but then they have lower dimension, therefore can be considered separately and can be ignored by induction. For this reason, whenever we further partition C_l, we always assume the cells are open.

Next, we will show that we may assume $\theta_l = 0$ on C_l. Since $x_m \sim \theta_l(x_{<m})$ on C_l, there exists some $\epsilon \in (0, 1)$ such that $\epsilon x_m < \theta_l(t, x_{<m}) < (1/\epsilon)x_m$ for all $(t, x) \in C_l$. Up to finite partitioning, if necessary, we may assume that either $\theta_l(t, x_{<m}) > x_m$ for all $(t, x) \in C_l$ or $\theta_l(t, x_{<m}) < x_m$ for all $(t, x) \in C_l$. In the first case, the other is similar, consider the *cell*

$$\tilde{C}_l = \{(t, x) \in T \times (0, 1)^m \mid \left(t, x_{<m}, \frac{-1}{\epsilon} x_m + \theta_l(t, x_{<m})\right) \in C_l\}$$

and define $\tilde{f}_l : \tilde{C}_l \to X_T$ by $\tilde{f}_l(t, x) = f_l(t, x_{<m}, ((-1)/\epsilon)x_m + \theta_l(t, x_{<m}))$. Then $\tilde{f}_l$ is prepared on $\tilde{C}_l$ with centre zero. Moreover, properties 1, 2 and 3 are still satisfied. Thus, from now on, we will assume additionally that f_l is prepared on C_l with centre zero.

Next, we show that, up to finite partitioning, if necessary, the associated bounded range map $b_l = (b_{l,1}, \ldots, b_{l,N})$ of f_l is C^1 and $|\partial b_l / \partial x_m| \leq 1$. That we may suppose that b_l is C^1 is a classical consequence of the cell decomposition theorem. Up to further partitioning using o-minimality, we may suppose that there exists some $i \in \{1, \ldots, N\}$ such that for any other $i' \in \{1, \ldots, N\}$ we have: $|\partial b_{l,i} / \partial x_m| \geq |\partial b_{l,i'} / \partial x_m|$ on C_l and that either $|\partial b_{l,i} / \partial x_m| \geq 1$ or $|\partial b_{l,i} / \partial x_m| < 1$ on C_l. The second case is exactly what we want. In the first case, we do a change of variables. To ensure that we recover a map $T \times (0, 1)^m \to [-1, 1]^{m-n}$, we first have to further partition C_l such that $b_{l,i}$ is either identically -1,0 or 1, $b_{l,i} > 0$ or $b_{l,i} < 0$. We only have to consider the last two cases. Suppose we have

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

$b_{l,i} > 0$ (the case $b_{l,i} < 0$ is similar up to changing a sign). Once more using o-minimality, we may assume that for fixed $(t, x_1, \ldots, x_{m-1})$ the map $x_m \mapsto b_{l,i}(t, x_1, \ldots, x_{m-1}, x_m)$ is injective and it follows that the map

$$\phi : C_l \to \phi(C_l) : (t, x) \mapsto (t, x_1, \ldots, x_{m-1}, b_{l,i}(t, x))$$

is invertible. Set $\bar{C}_l = \phi(C_l)$ and $\bar{f}_l = f_l \circ \phi^{-1}$. Then $\bar{C}_l$ is an open cell in $T \times (0, 1)^m$ and $\bar{f}_l$ satisfies the same properties as f_l, in particular, it is prepared in x_m with bounded range map $\bar{b}_l$ but moreover $|\partial \bar{b}_{l,i}/\partial x_m| \leq 1$ for all component functions $\bar{b}_{l,i}$ of $\bar{b}_l$ (by the chain rule and the choice of i).

So up to now, we may suppose we have finitely many maps $f_l : C_l \to X_T$ satisfying the first three properties, that are prepared in x_m and the associated bounded range map b_l of f_l is C^1 and $|\partial b_{l,i}/\partial x_m| \leq 1$. Thus, for any $i \in \{1, \ldots, N\}$, the component function $b_{l,i}$ of b_l is given by:

$$b_{l,i}(t, x) = a_{l,i}(t, x_{<m}) x_m^{r_{l,i}}.$$

Denote α_m for the wall of C_l bounding the variable x_m from below and β_m for the wall of C_l bounding x_m from above. Up to further partitioning, we may suppose $\alpha_m \neq 0$ on C_l and we will also suppose that $r_{l,i} \neq 0$ for every $i \in \{1, \ldots, N\}$. We will come back to this at the end of the proof. Now consider the following maps:

$$h_\alpha : \pi_{<m}(C_l) \to [-1, 1]^N : (t, x_{<m}) \mapsto \lim_{x_m \to \alpha_m(t, x_{<m})} b_l(t, x_{<m}, x_m)$$

$$g_\alpha : \pi_{<m}(C_l) \to [-1, 1]^N : (t, x_{<m}) \mapsto \lim_{x_m \to \alpha_m(t, x_{<m})} \frac{\partial b_l}{\partial x_m}(t, x_{<m}, x_m)$$

and define h_β and g_β analogously. Note that g_α and g_β are well defined by our effort to bound $|\partial b_{l,i}/\partial x_m|$ by 1. For any $i \in \{1, \ldots, N\}$, by the form of $b_{l,i}$, the component functions of $h_{\alpha,i}$ and $g_{\alpha,i}$ (similarly for h_β and g_β) become:

$$h_{\alpha,i}(t, x_{<m}) = a_{l,i}(t, x_{<m}) \alpha_m(t, x_{<m})^{r_{l,i}},$$

$$g_{\alpha,i}(t, x_{<m}) = r_{l,i} a_{l,i}(t, x_{<m}) \alpha_m(t, x_{<m})^{r_{l,i}-1} \quad (r_{l,i} \neq 0).$$

Now define the map $F : \pi_{<m}(C_l) \to \mathrm{Im}(F)$ whose component functions are α_m, β_m, h_α, h_β, g_α and g_β. Next, apply the induction hypothesis to the graph of F. Hence we obtain finitely many maps $\psi_{l,j} : D_{l,j} \to \mathrm{graph}(F)$ satisfying all properties of the theorem. In particular, they are prepared in $x_{<m}$ with associated bounded monomial map $c_{l,j}$ that is C^1-bounded in $x_{<m}$. Therefore, by their form, the maps $\psi_{l,j}$ are C^1-bounded in $x_{<m}$. Note that $\psi_{l,j}(t, x_{<m})_{<m} \in \pi_{<m}(C_l)$ and thus, for instance, $\alpha_m(\psi_{l,j}(t, x_{<m})_{<m})$ is a component function of $\psi_{l,j}$. Now set

$$C_{l,j} = \{(t, x_{<m}, x_m) \in D_{l,j} \times (0, 1) \mid (\psi_{l,j}(t, x_{<m})_{<m}, x_m) \in C_l\}.$$

By construction, these cells have all the properties we want. Finally define $f_{l,j} : C_{l,j} \to X_T$ by:

$$f_{l,j}(t, x_{<m}, x_m) = f_l(\psi_{l,j}(t, x_{<m})_{<m}, x_m).$$

By the form of f_l, it remains to show that for any $i \in \{1, \ldots, N\}$, $b_{l,i}(\psi_{l,j}(t, x_{<m})_{<m}, x_m)$ is prepared in x and is C^1-bounded in x. Thus, let $b_{l,i}$ be a component function of b_l and

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

suppose $r_{l,i} < 0$ (in the case $r_{l,i} > 0$, use β instead of α in the calculations). We have that:

$$\begin{aligned} b_{l,i}(\psi_{l,j}(t, x_{<m})_{<m}, x_m) &= a_{l,i}(\psi_{l,j}(t, x_{<m})_{<m}) x_m^{r_{l,i}} \\ &= h_{\alpha,i}(\psi_{l,j}(t, x_{<m})_{<m}) \left(\frac{x_m}{\alpha_m(\psi_{l,j}(t, x_{<m})_{<m})} \right)^{r_{l,i}}. \end{aligned}$$

By the construction, $h_{\alpha,i}(\psi_{l,j}(t, x_{<m})_{<m})$ and $\alpha_m(\psi_{l,j}(t, x_{<m})_{<m})$ are prepared in $x_{<m}$ with associated bounded monomial map $c_{l,j}$ that is C^1-bounded in $x_{<m}$. Therefore, it follows that $b_{l,i}(\psi_{l,j}(t, x_{<m})_{<m}, x_m)$ is prepared in x. We now look at the C^1-norm. To simplify notation, denote $y = \psi_{l,j}(t, x_{<m})_{<m}$. Let $s \in \{1, \ldots, m-1\}$, then:

$$\begin{aligned} \frac{\partial}{\partial x_s}\big(b_{l,i}(y, x_m)\big) &= \frac{\partial}{\partial x_s}\left(h_{\alpha,i}(y) \left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} \right) \\ &= \frac{\partial}{\partial x_s}(h_{\alpha,i}(y)) \left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} + h_{\alpha,i}(y) \frac{\partial}{\partial x_s}\left(\left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} \right). \end{aligned}$$

The first term is bounded in x since $h_{\alpha,i}(y) = h_{\alpha,i}(\psi_{i,l}(t, x_{<m})_{<m})$ is C^1-bounded in $x_{<m}$ (because it is prepared in $x_{<m}$, with C^1-bounded associated bounded monomial map) and $(x_m/\alpha_m(y))^{r_{l,i}} < 1$ since $x_m > \alpha_m(y)$ and $r_{l,i} < 0$. We further compute the last term:

$$\begin{aligned} h_{\alpha,i}(y) \frac{\partial}{\partial x_s}\left(\left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} \right) &= -a_{l,i}(y) \alpha_m(y)^{r_{l,i}} \frac{r_{l,i}}{\alpha_m(y)} \left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} \frac{\partial}{\partial x_s}\big(\alpha_m(y)\big) \\ &= -g_{\alpha,i}(y) \left(\frac{x_m}{\alpha_m(y)} \right)^{r_{l,i}} \frac{\partial}{\partial x_s}(\alpha_m(y)). \end{aligned}$$

Hence we see that also this term is bounded in x. Since $(\partial b_l / \partial x_m)$ was already bounded, we obtain that $b_{l,i}(\psi_{l,j}(t, x_{<m})_{<m}, x_m)$ is C^1-bounded in x.

To conclude we explain the cases $\alpha_m = 0$ and $r_{l,i} = 0$. If $\alpha_m = 0$, this forces the exponent $r_{l,i}$ to be positive or zero since b_l is bounded. If $r_{l,i} > 0$, we can just use the maps h_β and g_β as above since $\beta_m > 0$. If $r_{l,i} = 0$, one can use the induction hypothesis on $b_{l,i} = a_{l,i}$ immediately. □

We can now easily prove the main theorem using this parameterization result and Lemma 2.5. It refines the C^r-parameterization theorem of [4] by making the constant d more explicit; we have $d = m^3$.

Theorem 3.12. *Suppose that X_T is a definable family of m-dimensional subsets in $[-1, 1]^n$. Then there exists a constant $c > 0$ such that for any integer $r > 0$ and $t \in T$ there is a collection of finitely many analytic maps*

$$\{f_{r,i,t} : (0,1)^m \to X_t \mid i \in \{1, \ldots, cr^{m^3}\}\}$$

whose C^r-norm is bounded by 1 and such that for any $t \in T$ the ranges of $f_{r,i,t}$, for $i = 1, \ldots, cr^{m^3}$, cover X_t. Moreover, for each i and r, $\{f_{r,i,t} \mid t \in T\}$ is a definable family of maps.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

Proof. As we explained below Definition 3.1, we may suppose that X_T is the graph of a definable function $T \times (0,1)^m \to [-1,1]^{n-m}$. By the previous result, we obtain finitely many $f : C \to X_T$, where C is an open cell in $T \times (0,1)^m$. Furthermore, f is prepared in x with associated bounded monomial map b that is C^1-bounded in x and also the walls of C are of this form. By Proposition 3.10 using the map

$$(t, x_1, \ldots, x_m) \mapsto \left(t, x_1^{r^m}, x_2^{r^{m-1}}, \ldots, x_m^r\right),$$

we get finitely many maps $f_r : C_r \to X_T$, where C_r is open in $T \times (0,1)^m$. Moreover, there is an $A > 0$ such that for any $t \in T$, the walls α_1 and β_1 of $C_{r,t}$ of x_1 are $(A, 0)$-mild, for $i = 2, \ldots, m$, the walls α_i and β_i of x_i are $(Ar^m, 0)$-mild up to order r and $f_{r,t} : C_{r,t} \to X_t$ is $(Ar^m, 0)$-mild up to order r.

We now inductively map $T \times (0,1)^m$ on the cell C_r as follows. For $i = 1, \ldots, m$, define the map

$$\Phi_i : \pi_{<i}(C_r) \times (0,1)^{m-i+1} \to \pi_{<i+1}(C_r) \times (0,1)^{m-i}$$

given by:

$$(t, x) \mapsto (t, x_1, \ldots, x_{i-1}, ((\beta_i - \alpha_i)x_i + \alpha_i)(t, x_{<i}), x_{i+1}, \ldots, x_m)$$

and let $\Phi = \Phi_1 \circ \ldots \circ \Phi_m : T \times (0,1)^m \to C_r$. If $i \geq 2$, since for any $t \in T$, α_i and β_i are $(Ar^m, 0)$-mild up to order r, there is some $A' > 0$ such that for any $t \in T$, the map $\Phi_{i,t}$ is $(A'r^m, 0)$-mild up to order r, where $A' > 0$ is possibly larger than A due to the product (2.15) and addition of mild functions. If $i = 1$, $\Phi_{1,t}$ is $(A', 0)$-mild, for possibly larger A'. Thus, we obtain that there is some $A' > 0$, possibly larger than A' by Corollary 2.11, such that for any $t \in T$, the map $(f_r \circ \Phi)_t$ is $(A''r^{m^2}, 0)$-mild up to order r.

Finally, by Lemma 2.5, covering the unit cube with cubes of size $1/(A''r^{m^2})$, we obtain cr^{m^3} maps that parametrize X_T with bounded C^r-norm. □

Remark. The main obstacle to obtain a better result (e.g. cr^m maps as in [1]), is the power substitution using r^m to ensure that the walls of the cell are mild up to order r too. Thus it would be interesting to prove a stronger version of Theorem 3.11 that has better walls.

A small improvement that we show now slightly improves the walls. More precisely, if $m \geq 2$, we will show that we can first improve the walls bounding the variable x_2 and use a slightly better power substitution with powers up to r^{m-1} (instead of r^m).

Suppose we are given a map as a result of Theorem 3.11, call it $f : C \to X_T$, where C is an open cell in $T' \times (0,1)^m$ and T' a cell contained in T. Now, since the walls bounding x_1 and x_2 are prepared in x, the first two inequalities defining C_t are:

$$\alpha_1(t) < x_1 < \beta_1(t)$$
$$a_i(t)x_1^r F(a(t, x_1)) < x_2 < b_j(t)x_1^s G(b(t, x_1)).$$

We will now manipulate f and C to ensure that the maps $a_i(t)x_1^r$ and $b_j(t)x_1^s$ become mild for any $t \in T$, by transforming r and s to natural numbers. If they are already mild, there is no power substitution required for these walls and we can use a better version

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

of Proposition 3.10. Suppose that $r - s \geq 0$, the other case is similar. Since G is analytic and non-vanishing on $\mathrm{Im}(b)$, there exists $S > 0$ such that $G(b(t, x_1)) \in (1/S, S)$ for all t and x. Equivalently: $G(b(t, x_1))/S \in (1/S^2, 1)$. Now consider the map ϕ given by:

$$\phi : T \times (0,1)^m \to \mathrm{Im}(\phi) : \phi(t, x_1, \ldots, x_m) = (t, x_1, b_j(t)x_1^s S x_2, x_3, \ldots, x_m).$$

Since $b_j(t)x_1^s$ is C^1-bounded in x_1 by assumption, the map ϕ is C^1-bounded in x. We see that $\phi^{-1}(C)$ is a cell where the walls bounding x_1 and x_2 are now given by:

$$\alpha_1(t) < \ x_1 < \beta_1(t)$$
$$(a_i/b_j)(t)x_1^{r-s} F(a(t, x_1))/S < x_2 < G(b(t, x_1))/S.$$

By the form of ϕ, the walls bounding $x_3, \ldots, x_m$ are still prepared in x and their associated bounded monomial map is C^1-bounded in x. If $r - s = 0$, the walls are already as desired. Denote by R the smallest integer greater than or equal to $r - s$. We then use the power substitution

$$(t, x_1, x_2, \ldots, x_m) \mapsto (t, x_1^{R/(r-s)}, x_2, \ldots, x_m)$$

to obtain a cell $\tilde{C}$ and a map $\tilde{f} : \tilde{C} \to X_T$ with the same properties as f, i.e. $\tilde{f}$ and all of the walls of $\tilde{C}$ are prepared in x with an associated bounded monomial map that is C^1-bounded in x. The walls of $\tilde{C}$ bounding x_1 and x_2 are of the form:

$$\tilde{\alpha}_1(t) < x_1 < \tilde{\beta}_1(t)$$
$$\tilde{a}_i(t)x_1^N F(\tilde{a}(t, x_1)) < x_2 < G(\tilde{b}(t, x_1))$$

where $N \in \mathbb{N}$. We now have that for any $t \in T$, $\tilde{a}_i(t)x_1^N$ is mild. One now proceeds as in the proof of (3.12) using a slightly modified version of Proposition 3.10, namely with the power substitution

$$(t, x_1, \ldots, x_m) \mapsto \left(t, x_1^{r^{m-1}}, x_2^{r^{m-1}}, x_3^{r^{m-2}}, \ldots, x_m^r\right).$$

In the proof of Theorem 3.12, we would now have that there is some $A > 0$ such that for any $t \in T$, the walls α_i and β_i of x_i are $(Ar^{m-1}, 0)$-mild up to order r and $f_{r,t}$ is $(Ar^{m-1}, 0)$-mild up to order r. It follows that for some possibly larger A and for any $t \in T$, $(f_r \circ \Phi)_t$ is $(Ar^{m(m-1)}, 0)$-mild up to order r and thus obtain $cr^{m^2(m-1)}$ maps as result.

It would be interesting to know if we could also transform the walls of the other variables in this way. However, currently, it is not clear how to achieve this.

Acknowledgements. The author would like to thank the referee for reading the article in great detail and giving numerous helpful comments to improve the presentation. The author is partially supported by KU Leuven grant IF C14/17/083.

References

1. G. Binyamini and D. Novikov, Complex cellular structures, *Ann. Math. (2)* **190**(1) (2019), 145–248.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press

2. E. Bombieri and J. Pila, The number of integral points on arcs and ovals, *Duke Math. J.* **59**(2) (1989), 337–357.
3. D. Burguet, A proof of Yomdin-Gromov's algebraic lemma, *Israel J. Math.* **168** (2008), 291–316.
4. R. Cluckers, J. Pila and A. Wilkie, Uniform parameterization of subanalytic sets and diophantine applications, *Ann. Sci. Ecole Norm. Sup.* **53**(1) (2020), 1–42.
5. G. M. Constantine and T. H. Savits, A multivariate Faà di Bruno formula with applications, *Trans. Amer. Math. Soc.* **348**(2) (1996), 503–520.
6. M. Gevrey, Sur la nature analytique des solutions des équations aux dérivées partielles, *Premier mémoire. Ann. Sci. École Norm. Sup.* **35**(3) (1918), 129–190.
7. M. Gromov, Entropy, homology and semialgebraic geometry. Séminaire Bourbaki, vol. 1985-1986, *Astérisque* **145–146**(5) (1987), 225–240.
8. 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) (2011), 67–74.
9. S. G. Krantz and H. R. Parks, *A primer of real analytic functions*. Birkhäuser Advanced Texts, Second edition (2002). doi:10.1007/978-0-8176-8134-0.
10. D. J. Miller, A preparation theorem for Weierstrass systems, *Trans. Amer. Math. Soc.* **358**(10) (2006), 4395–4439. (electronic).
11. J. Pila, Mild parameterization and the rational points of a Pfaff curve, *Comment. Math. Univ. St. Pauli* **55**(1) (2006), 1–8.
12. J. Pila, Counting rational points on a certain exponential-algebraic surface, *Ann. Inst. Fourier (Grenoble)* **60**(2) (2010), 489–514.
13. J. Pila and A. J. Wilkie, The rational points of a definable set, *Duke Math. J.* **133**(3) (2006), 591–616.
14. L. van den Dries, *Tame topology and O-minimal structures*, Lecture Note Series, Volume **248** (Cambridge University Press, 1998). doi:10.1017/CBO9780511525919.
15. S. Van Hille, On a family of mild functions, *Int. J. Number Theory* (2021), (online ready). https://doi.org/10.1142/S179304212150041X.
16. Y. Yomdin, C^k-resolution of semialgebraic mappings. Addendum to "Volume growth and entropy", *Israel J. Math.* **57**(3) (1987), 301–317.
17. Y. Yomdin, Volume growth and entropy, *Israel J. Math.* **57**(3) (1987), 285–300.
18. Y. Yomdin, Analytic reparametrization of semi-algebraic sets, *J. Complexity* **24**(1) (2008), 54–76.
19. Y. Yomdin, Smooth parametrizations in dynamics, analysis, diophantine and computational geometry, *Jpn. J. Ind. Appl. Math.* **32**(2) (2015), 411–435.

https://doi.org/10.1017/S0013091521000298 Published online by Cambridge University Press