Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Annals of Pure and Applied Logic 125 (2004) 43–74

ANNALS OF PURE AND APPLIED LOGIC

www.elsevier.com/locate/apal

An effective version of Wilkie's theorem of the complement and some effective o-minimality results

Alessandro Berarducci[a,*], Tamara Servi[b]

[a] *Dipartimento di Matematica, Università di Pisa, Via Buonarroti 2, 56127 Pisa, Italy*
[b] *Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, 56100 Pisa, Italy*

Received 13 July 2003; received in revised form 11 August 2003; accepted 25 August 2003
Communicated by A.J. Wilkie

Abstract

Wilkie (Selecta Math. (N.S.) 5 (1999) 397) proved a "theorem of the complement" which implies that in order to establish the o-minimality of an expansion of $\mathbb{R}$ with C^∞ functions it suffices to obtain uniform (in the parameters) bounds on the number of connected components of quantifier free definable sets. He deduced that any expansion of $\mathbb{R}$ with a family of Pfaffian functions is o-minimal. We prove an effective version of Wilkie's theorem of the complement, so in particular given an expansion of the ordered field $\mathbb{R}$ with finitely many C^∞ functions, if there are uniform and computable upper bounds on the number of connected components of quantifier free definable sets, then there are uniform and computable bounds for all definable sets. In such a case the theory of the structure is effectively o-minimal: there is a recursively axiomatized subtheory such that each of its models is o-minimal. This implies the effective o-minimality of any expansion of $\mathbb{R}$ with Pfaffian functions. We apply our results to the open problem of the decidability of the theory of the real field with the exponential function. We show that the decidability is implied by a positive answer to the following problem (raised by van den Dries (in: Logic: From Foundations to applications, Oxford Science Publ., Oxford University Press, New York, 1996, p. 137)): given a language L expanding the language of ordered rings, if an L-sentence is true in every L-structure expanding the ordered field of real numbers, then it is true in every o-minimal L-structure expanding any real closed field.

MSC: 03C64; 03B25

Keywords: o-Minimality; Pfaffian functions; Real exponentiation

* Corresponding author.
E-mail addresses: berardu@paley.dm.unipi.it (A. Berarducci), servi@sns.it (T. Servi).

0168-0072/$ - see front matter
doi:10.1016/j.apal.2003.08.001

1. Introduction

In [12] Tarski proved that the (complete first-order) theory of the structure $\mathbb{R} = (\mathbb{R}, +, \cdot, 0, 1, <)$ is decidable, and asked whether the same holds for the theory $T_{\exp}$ of $\mathbb{R}_{\exp} = (\mathbb{R}, +, \cdot, 0, 1, <, \exp)$, where $\exp(x) = e^x$ is the real exponential function. This problem is still open and has been a main source of motivation for research in this area.

Some information on this structure comes from the study of the zero-sets of *exponential polynomials*, where an exponential polynomial is an expression of the form $p(x_1, \dots, x_n, e^{x_1}, \dots, e^{x_n})$ with $p(x_1, \dots, x_n, y_1, \dots, y_n) \in \mathbb{R}[\vec{x}, \vec{y}]$. A (real) exponential variety $V = V(F) \subseteq \mathbb{R}^n$ is the zero-set of a system $F : \mathbb{R}^n \to \mathbb{R}^k$ of k exponential polynomials. In [7] Khovanskii proved that there is a finite upper bound on the number of connected components of an exponential variety $V = V(F)$, which is computable in terms of the *complexity* of the system $F : \mathbb{R}^n \to \mathbb{R}^k$, where the complexity is given by n, k and the total degrees of the polynomials involved. More generally he obtained similar computable bounds for systems of "Pfaffian equations". (The exponential function e^x is Pfaffian.)

In [13] Wilkie proved that $T_{\exp}$ is *model complete*, and therefore every definable set in $\mathbb{R}_{\exp}$ is a projection of an exponential variety. It then follows that every definable set $X \subseteq \mathbb{R}^n$ in this structure has a finite number of connected components and therefore $\mathbb{R}_{\exp}$ is *o-minimal* (see [3] for an introduction to the subject).

In [8] Macintyre and Wilkie linked the decidability of $T_{\exp}$ to some unproved conjectures in transcendental number theory. *Schanuel's conjecture* (in the real case) states that, given $\alpha_1, \dots, \alpha_n \in \mathbb{R}^n$ linearly independent over $\mathbb{Q}$, the transcendence degree of the field $\mathbb{Q}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n})$ over $\mathbb{Q}$ is $\geqslant n$. It is proved in [8] that Schanuel's conjecture implies the decidability of $T_{\exp}$. Actually the authors show that the decidability is equivalent to a weaker conjecture, the *Last Root Conjecture*, which states that there is a computable upper bound on the norm of the non-singular solutions of an $n \times n$ system of exponential polynomials (by Khovanskii's results there are computable upper bounds on the cardinality of the set of solutions, but not necessarily on their norms).

In [15] Wilkie proved, using the notion of Charbonnel closure introduced in [1], a general "theorem of the complement" which in particular implies that in order to establish the o-minimality of an expansion of $\mathbb{R}$ with C^∞-functions it suffices to prove uniform (in the parameters) bounds on the number of connected components of quantifier free definable sets. He deduced, using [7], that any expansion of $\mathbb{R}$ with a family of Pfaffian functions is o-minimal. So every definable set in any of these structures has a finite number of connected components. The results in [13,15] leave open the question of the effectiveness of the bounds. (The effectiveness would follow trivially from the decidability of the corresponding theory, although the converse is not a priori true.)

In this paper we prove an effective version of Wilkie's theorem of the complement. In particular we prove that, given an expansion of $\mathbb{R}$ with finitely many C^∞ functions, if there are uniform and computable upper bounds on the number of connected components of quantifier free definable sets, then there are uniform and computable bounds for all definable sets. In such a case the theory of the structure is effectively o-minimal: there is a recursively axiomatized subtheory such that all the models are o-minimal. The hypotheses of our theorem hold in the case of an expansion of $\mathbb{R}$ with Pfaffian functions by [7], so in particular we obtain a proof of the effective o-minimality of any expansion of $\mathbb{R}$ by finitely many Pfaffian functions.

We apply our results to the open problem of the decidability of the theory of the real field with the exponential function. We show that the decidability is implied by a positive answer to the following problem (raised by van den Dries (1993)): given a language L expanding the language of ordered rings, if an L-sentence is true in every L-structure expanding the ordered field of real numbers, then it is true in every o-minimal L-structure expanding any real closed field.

After a preliminary version of this paper was completed, Alex Wilkie pointed out to us the article [5] of Gabrielov. This belongs to a series of papers (see for instance [4,5]) where Gabrielov studies the effectiveness and the complexity of several operations and topological invariants of sets belonging to various categories which can be defined in terms of Pfaffian functions: semi-Pfaffian, restricted sub-Pfaffian, sub-Pfaffian, first-order definable (in a language with Pfaffian functions). The semi-Pfaffian is the smallest category, while the definable category is the largest. A nice survey of the research in this area is the preprint [6]. In particular Gabrielov has introduced a representation of the definable sets in an expansion of $\mathbb{R}$ with Pfaffian functions in terms of "limit sets" and has obtained in [5] efficient computable upper bounds on various operations and topological invariants (among which the number of connected components) in terms of the complexity of the limit-set expressions. The existence of recursive bounds in the Pfaffian case also follows from our main theorem. Our results, being based only on the existence of computable bounds in the quantifier free case, do not use specific properties of the Pfaffian functions such as suitable versions of the Łojasiewicz inequality.

2. Consequences of the main result

Our main theorem (Theorem 9.4) applies to the definable sets in the language associated to an "o-minimal effective W-structure" (Definition 5.1) which is "effectively determined by its smooth functions" (Definition 6.1). This is an effective analogue of the setting of [15]. Since the definitions involved are rather technical, we state in this outline a particular case of the theorem which is easier to formulate. We then derive some corollaries.

Definition 2.1. For $X \subseteq \mathbb{R}^n$ let $cc(X)$ be the number of connected components of X and let $\gamma(X)$ be the least $n \in \mathbb{N}$ such that for every affine set $L \subseteq \mathbb{R}^n$ (i.e. a set defined by a system of linear equations over $\mathbb{R}$) we have $cc(X \cap L) \leqslant n$, with the convention that $\gamma(X) = \infty$ if n does not exist.

Clearly $cc(X) \leqslant \gamma(X)$. It is well known that a first order structure with domain $\mathbb{R}$ is o-minimal if and only if for every definable set in the structure one has $cc(X) < \infty$. If a structure is o-minimal it then follows that one actually has $\gamma(X)<\infty$. We can now state the particular case of Theorem 9.4:

Theorem 2.2. *Let $\mathcal{R}$ be an L-structure which is an expansion of $(\mathbb{R};+,\cdot,0,1)$ by finitely many C^∞ functions. Assume that there is a recursive function Γ_0 which, given a quantifier free L-formula $\phi(\vec{x})$ computes a finite upper bound $\Gamma_0(\phi)\in\mathbb{N}$ on $\gamma(X)$, where $X\subseteq\mathbb{R}^n$ is the set defined by ϕ. Then there is a recursive function Γ which, given an arbitrary L-formula $\theta(\vec{x})$ computes a finite upper bound $\Gamma(\theta)\in\mathbb{N}$ on $\gamma(Y)$, where $Y\subseteq\mathbb{R}^n$ is the set defined by θ.*

The corresponding result, dropping the word "recursive", is due to Wilkie [15]. The formulas involved in the above theorem are without parameters, so it makes sense to speak of recursive functions taking such formulas as inputs (an L-formula is just a string of symbols from some finite alphabet). We have not attempted a complexity analysis, but it should be clear by the analysis of the proof that if Γ_0 is primitive recursive, then Γ can also be found primitive recursive. For technical reasons we did not include the order relation in the language. However the order can be defined as usual from $+,\cdot$ using existential quantifiers.

Remark 2.3. Theorem 2.2 refers to formulas without parameters. The following easy observation allows us to obtain bounds on γ also in the presence of parameters: if $X\subseteq\mathbb{R}^n$ is defined by a formula $\phi(\vec{x},\vec{b})$ with n free variables $\vec{x}$ and k parameters $\vec{b}\in\mathbb{R}$, then $\gamma(X)\leqslant\gamma(Y)$, where $Y\subseteq\mathbb{R}^{n+k}$ is defined by the formula without parameters $\phi(\vec{x},\vec{y})$.

Let $\mathbb{R}_{\mathrm{Pfaff}}$ be an expansion of $(\mathbb{R},+,\cdot,0,1)$ by finitely many Pfaffian functions (for instance by the exponential function e^x). Then the hypothesis of Theorem 2.2 are verified by [7] (reasoning as in [15, Theorem 1.9]). We thus obtain

Corollary 2.4. *If $X\subseteq\mathbb{R}^n$ is defined by a formula $\phi=\phi(x_1,\dots,x_n)$ in the language of $\mathbb{R}_{\mathrm{Pfaff}}$, then $\gamma(X)<\Gamma(\phi)$, where $\Gamma:Formulas\to\mathbb{N}$ is a computable (even primitive recursive) function.*

Corollary 2.5. *Let T_{Pfaff} be the complete theory of $\mathbb{R}_{\mathrm{Pfaff}}$. There is a recursively axiomatized subtheory T_{omin} of T_{Pfaff} such that all the models of T_{omin} are o-minimal.*

Proof. A structure M is o-minimal if and only if every definable subset of M, possibly with parameters, is a finite union of open intervals and points. So it suffices to define T_{omin} as the theory which contains, for each formula $\phi(x_1,x_2,\dots,x_n)$ in the language of T_{Pfaff}, an axiom stating that $\forall x_2,\dots,x_n$ the set $\{x_1 \mid \phi(x_1,x_2,\dots,x_n)\}$ is the union of at most $\Gamma(\phi)$ open intervals and points. □

Now let $T_{\exp}$ be the complete theory of $\mathbb{R}_{\exp}$. The corollary allows us to prove the recursive axiomatizability (and therefore the decidability) of $T_{\exp}$ under the assumption that the following open problem, raised by van den Dries in [2], has a positive answer:

Problem 2.6 (Transfer conjecture). Let L be a language expanding the language of ordered rings. If an L-sentence is true in all L-structures expanding the ordered field of real numbers, then it is true in all o-minimal L-structures expanding any real closed field.

Theorem 2.7. *The Transfer conjecture implies the decidability of* T_{exp}.

Proof. By Corollary 2.5 there is a recursive subtheory T_{omin} of T_{exp}, such that all the models of T are o-minimal. Let RCF be the axioms for real closed fields and let EXP be the sentence stating that exp is a C^1 function which satisfies $\exp(0)=1$ and $\exp(x)' = \exp(x)$. It suffices to prove that $RCF + T_{\mathrm{omin}} + EXP$ is a complete axiomatization of T_{exp}. So let $\mathbb{R}_{\mathrm{exp}} \models \varphi$. By the classical uniqueness result for the solution of the differential equation for exp, $\mathbb{R}_{\mathrm{exp}}$ is the only L_{exp}-expansion of the real field which satisfies EXP. So by the transfer conjecture $EXP \to \varphi$ is true in every o-minimal L_{exp}-structure expanding any real closed field. Hence φ is deducible from $RCF + T_{\mathrm{omin}} + EXP$. □

Note that Theorem 2.7 only uses the fact that exp is Pfaffian (to be able to apply the results of Khovanskii) and is the unique solution of its differential equation.

We thus obtain another candidate for a recursive axiomatization of T_{exp}, after the one considered in [8]. In any model M of the proposed axioms, any other definable function satisfying EXP must coincide with the interpretation of e^x in M. This follows from a uniqueness result for differential equations in o-minimal expansions of a field proved in [10].

Independently of the Transfer conjecture, the above argument shows:

Theorem 2.8. *In order to establish the decidability of* T_{exp} *it suffices to show that any two o-minimal models of* $RCF + EXP$ *are elementary equivalent.*

In fact in this case $RCF + T_{\mathrm{omin}} + EXP$ would be a complete recursive axiomatization.

3. Outline of the proof of the main theorem

The proof of Theorem 9.4 and its consequences, as stated in the previous section, is based on suitable effective versions of the results in [15], and in particular of the cell decomposition theorem contained in that paper. We cannot go as far as to claim that there is an algorithm to perform the cell decomposition theorem, since, in the case of $\mathbb{R}_{\mathrm{exp}}$, this would be equivalent to the decidability of T_{exp}: in fact a sentence φ in the language of $\mathbb{R}_{\mathrm{exp}}$ is true in $\mathbb{R}_{\mathrm{exp}}$ if and only if the subset of $\mathbb{R}$ defined by $(\varphi \wedge x = x)$ is non-empty (and one would expect that a reasonable notion of algorithmic cell decomposition should be able to tell if a set is empty). However we will see that, despite this obstacle, we can extract from [15] some "non-deterministic" or "multivalued" algorithms which are good enough for Theorem 9.4 and its corollaries.

In order to carry out this program we begin by presenting the results in [15] in a form that suits our purposes. The main idea in [15] is to give a new characterization of the definable sets, under suitable assumptions. The new characterization is based on the notion of *Charbonnel closure*, introduced by Charbonnel in [1], which we now describe.

Let $\mathscr{S}_n$ be a collection of subsets of $\mathbb{R}^n$ and let $\mathscr{S} = \langle \mathscr{S}_n \mid n \in \mathbb{N} \rangle$. The definable sets in the structure $\mathscr{S}$ form the smallest collection of sets stable under the boolean operations (inside each $\mathbb{R}^n$) and the operation of taking the image of a set under a linear projection $\Pi_n^{n+k} : \mathbb{R}^{n+k} \to \mathbb{R}^n$ (projection onto the first n coordinates). Let $\mathrm{Def}_n(\mathscr{S}) \supseteq \mathscr{S}_n$ be the collection of all definable subsets of $\mathbb{R}^n$ in the sense just described. We call $\mathrm{Def}(\mathscr{S}) = \langle \mathrm{Def}_n(\mathscr{S}) \mid n \in \mathbb{N} \rangle$ the definable closure of $\mathscr{S}$.

The Charbonnel closure $\mathrm{Ch}(\mathscr{S}) = \langle \mathrm{Ch}_n(\mathscr{S}) \mid n \in \mathbb{N} \rangle$ of $\mathscr{S}$ is defined similarly, but instead of the boolean operation of taking the complement one has the operation of taking the topological closure. More precisely one considers the operations of binary unions, projections, and the operation sending a sequence $A, B_1, \dots, B_k$ of subsets of $\mathbb{R}^n$ into $A \cap \overline{B_1} \cap \cdots \cap \overline{B_k}$, where $\overline{B_i}$ is the topological closure of B_i. We will work in this paper with an equivalent definition, where we replace the latter with the simpler operation of taking the topological closure $B \mapsto \bar{B}$ and we add a rather limited form of intersection with linear sets (see Definition 4.4 below).

Clearly $\mathrm{Def}_n(\mathscr{S}) \supseteq \mathrm{Ch}_n(\mathscr{S})$ (since the topological closure is a definable operation). In [15] Wilkie proves the following two results under suitable assumptions on $\mathscr{S}$. First, for every $X \in \mathrm{Ch}_n(\mathscr{S})$ we have $\gamma(X) < \infty$ (see Definition 2.1 for the definition of γ). This essentially amounts to proving that the operations in the definition of $\mathrm{Ch}(\mathscr{S})$ preserve the finiteness of γ, at least if they are applied to sets already in $\mathrm{Ch}(\mathscr{S})$. Second, and more difficult, under some additional "smoothness assumptions" on $\mathscr{S}$ it is shown that the complement of a set in $\mathrm{Ch}_n(\mathscr{S})$ is also in $\mathrm{Ch}_n(\mathscr{S})$. From this it clearly follows that the equality $\mathrm{Def}(\mathscr{S}) = \mathrm{Ch}(\mathscr{S})$ holds. The needed assumptions on $\mathscr{S}$ are verified if, for instance, $\mathscr{S}_n$ is the collection of all the exponential varieties included in $\mathbb{R}^n$. In this case the sets in $\bigcup_n \mathrm{Def}_n(\mathscr{S})$ coincide with the definable sets in the structure $\mathbb{R}_{\mathrm{exp}}$ and the o-minimality of $\mathbb{R}_{\mathrm{exp}}$ follows.

Our goal is to prove effective versions of these results. In order to do so it is technically convenient to weaken the assumptions on $\mathscr{S}$ (with respect to [15]), so as to allow for instance the possibility that $\mathscr{S}_n$ is the collection of all those exponential varieties included in $\mathbb{R}^n$ which are defined as the zero-sets of exponential polynomials with coefficients in $\mathbb{Z}$ (so in particular we do not require that all the semialgebraic subsets of $\mathbb{R}^n$ are in $\mathscr{S}_n$: actually we do not even require that all the singletons $\{a\}$ with $a \in \mathbb{R}$ are in $\mathscr{S}_1$). Assuming that $\mathscr{S}$ is an "effective W-structure" (Definition 5.1), the sets in $\mathrm{Ch}_n(\mathscr{S})$ can be naturally coded by "Ch-formulas" (Definition 5.3), which correspond to a subset of the first order formulas of the language associated to $\mathscr{S}$ (Definition 9.2). Roughly our Ch-formulas correspond to the "Charbonnel formulas" in [9], although our definition is different (recall that we work with different assumptions on $\mathscr{S}$ and with a different definition of $\mathrm{Ch}(\mathscr{S})$). If $\mathscr{S}_n$ consists of the exponential varieties included in $\mathbb{R}^n$ which are defined as the zero-sets of exponential polynomials with coefficients in $\mathbb{Z}$, then the Ch-formulas correspond to a subset of the first order formulas of T_{exp}.

Our first result (Lemma 4.10, Theorem 5.4) is that if $A \subseteq \mathbb{R}^n$ is defined by a Ch-formula A, then $\gamma(A) < \Gamma(\mathsf{A})$, where Γ : Ch-Formulas $\to \mathbb{N}$ is a computable function.

Using this fact we prove a "non-deterministic" effective version of Wilkie's theorem of the complement. More precisely we show (Theorem 9.1) that there is a recursive function which, given a Ch-formula for a set $A \in \mathrm{Ch}_n(\mathscr{S})$, returns a finite set of Ch-formulas, one of which defines the complement of A in $\mathbb{R}^n$, although we are not able to tell which one.

Granted this the results stated in the introduction follow easily. First we deduce that there is a recursive function which, given a first order formula ϕ (in the language associated to $\mathscr{S}$), returns a finite set of Ch-formulas, one of which defines the subset of $\mathbb{R}^n$ defined by ϕ. In other words we have an effective non-deterministic translation from first order formulas to Ch-formulas. This is the effective version of the result $\mathrm{Def}(\mathscr{S}) = \mathrm{Ch}(\mathscr{S})$. Finally we deduce Theorem 9.4, namely we obtain a recursive function which, given a first order formula $\phi(x_1, \ldots, x_n)$ in the language associated to $\mathscr{S}$ returns an upper bound on $\gamma(A)$, where $A \subseteq \mathbb{R}^n$ is the set defined by ϕ (it suffices to take the maximum of the bounds for the Ch-formulas associated to ϕ).

So it remains to prove Theorem 9.1, the non-deterministic effective version of Wilkie's theorem of the complement. The latter depends on a preliminary (and difficult) "boundary theorem" asserting that for every closed set $X \in \mathrm{Ch}(\mathscr{S})$ there is a closed set with empty interior $Y \in \mathrm{Ch}(\mathscr{S})$ such that Y contains the boundary ∂X of X (a posteriori it will follow that ∂X itself is in $\mathrm{Ch}(\mathscr{S})$). By an analysis of Wilkie's proof it is not difficult to obtain an effective version of this result (Theorem 6.12), namely it can be shown that from a Ch-formula for X we can effectively find a Ch-formula for Y. Note that we are not claiming that Wilkie's *proof* is constructive. What we claim is only that Wilkie's *definition* of Y (implicit in the proof) is constructive, although the proof that Y has the desired properties may not be so.

Granted the boundary theorem, the complement theorem in [15] follows from a cell-decomposition argument. The latter is non-constructive because it is a proof by cases and the task of distinguishing the cases by a computable function seems hopeless (we have already seen that even telling if a set is empty or not is connected with the decidability of $T_{\exp}$). To prove an effective version we do not distinguish the cases. We simply try them all and in at least one case we will obtain the correct result.

To give an idea of how this works, let us prove the non-deterministic effective version of the complement theorem (Theorem 9.1) in the basic case of subsets of $\mathbb{R}$. Note that in [15] this case is obvious and does not even requires the boundary theorem. The effective version is instead nontrivial even in the basic case so it is worth sketching a proof (in the official proof we will give a more complicated argument which is more suitable for the generalization to $\mathbb{R}^n$). So let $A \subseteq \mathbb{R}$ be a set in $\mathrm{Ch}_1(\mathscr{S})$ which we assume to be closed for simplicity. Then $\gamma(A) < \infty$ and therefore A is a finite union of closed intervals. The complement A^c of A is then a finite union of open intervals. We want to prove first of all that A^c is in $\mathrm{Ch}_1(\mathscr{S})$ (this is trivial in [15] since in that paper all the semialgebraic sets are in $\mathscr{S}$) and also that, given a Ch-formula for A, we can effectively find a finite set of Ch-formulas one of which defines A^c. The algorithm is the following. First, using the boundary theorem we find a Ch-formula for a set $B \in \mathrm{Ch}_1(\mathscr{S})$ with empty interior containing the boundary ∂A of A. Note that B is then a finite set of cardinality $\gamma(B)$. Since Ch-formulas do not have negations, it is not clear a priori whether one can effectively find, knowing the Ch-formula for B, a Ch-definition of the least element of the finite set B, or of the other elements of B. This however would be possible if we knew the cardinality of B. So we proceed as follows. First, using the Ch-formula for B, we compute an upper bound N on $\gamma(B)$. Then we choose non-deterministically a number $k \leqslant N$. At least one choice will give us the cardinality of B. Now given k we consider, for each $i \leqslant k \leqslant N$, the set $P_i^k = \{x \in \mathbb{R} \mid \exists x_1 \dots x_k \in B(x_1 < \cdots < x_k \wedge x = x_i)\}$, which can be defined by a Ch-formula as a projections on $\mathbb{R}$ of a Ch-definable set lying in $\mathbb{R}^{k+1}$. If k was the cardinality of B, as we temporarily assume, then these sets are singletons and B is the union of these singletons. Moreover the boundary of A is the union of a subset of these singletons. We now guess non-deterministically which of the above singletons and which of the open intervals determined by the such singletons are disjoint from A, and we take their union. This is the complement of A. If we were unlucky and k was not the cardinality of B, we can still make sense of the rest of the algorithm (e.g. if $P, Q \subseteq \mathbb{R}$ are not singletons we can still define the pseudo-interval $(P, Q) := \{x \mid \exists y \in P \exists z \in Q(y < x < z)\}$). At least one of the non-deterministic choices will lead to a Ch-formula for A^c.

In the general case (in $\mathbb{R}^n$) the proof of Theorem 9.1 will require a rather complex effective non-deterministic version of Wilkie's cell decomposition theorem. At a crucial point of the cell decomposition theorem we must define a certain number of functions (the functions which bound the cells) where the ith functions picks the ith point of a certain finite set $A \in \mathrm{Ch}_1(\mathscr{S})$. The problem is that we do not know the cardinality of A, but we can compute an upper bound on it (since we can compute an upper bound on $\gamma(A)$ given a Ch-formula for A). We have thus only finitely many possibilities and we can non-deterministically guess the exact cardinality and proceed with the construction.

We remark on a difference between our approach and Wilkie's in the structure of the induction. We perform induction on the notion of "rank" introduced in Definition 4.5. Moreover we make the assumption (not present in [15]) that all sets in $\mathscr{S}$ are closed (other differences on $\mathscr{S}$ and $\mathrm{Ch}(\mathscr{S})$ have already been explained). This assumption is inessential and can be dropped (assuming our EDSF condition, Definition 6.1), however since this would not produce any essential gain of generality in the main result, we decided to keep the assumption to simplify some arguments.

4. W-structures and Charbonnel closure

The following definition of W-structure is a modification of the notion of weak-structure in [15]. The difference is that we do not require that all the semi-algebraic sets are in the structure.

Definition 4.1. Let $\mathscr{S} = \langle \mathscr{S}_n \colon n \in \mathbb{N}^+ \rangle$, where $\mathscr{S}_n$ is a collection of subsets of $\mathbb{R}^n$. We say that $\mathscr{S}$ is a *W-structure* if for all $n \in \mathbb{N}$,

W(pol): $\mathscr{S}_n$ contains every subset of $\mathbb{R}^n$ defined as the zero-set of a system of finitely many polynomials with coefficients in $\mathbb{Z}$;

W(perm): if $A \in \mathscr{S}_n$, then $\Sigma[A] \in \mathscr{S}_n$, where $\Sigma : \mathbb{R}^n \to \mathbb{R}^n$ is a linear bijection induced by a permutation of the variables;
W($\cap$): if $A, B \in \mathscr{S}_n$, then $A \cap B \in \mathscr{S}_n$;
W($\times$): if $A \in \mathscr{S}_n$ and $B \in \mathscr{S}_m$, then $A \times B \in \mathscr{S}_{n+m}$.

Definition 4.2. We say that a W-structure $\mathscr{S}$ is *closed* if for every n and $A \in \mathscr{S}_n$, A is a closed subset of $\mathbb{R}^n$; $\mathscr{S}$ is *semi-closed* if for every n and $A \in \mathscr{S}_n$, A can be obtained as the projection onto the first n coordinates of some closed set $B \in \mathscr{S}_{n+k}$, for some suitable k. We say that a W-structure is *o-minimal* if for every n and $A \in \mathscr{S}_n$ we have $\gamma(A) < \infty$.

Example 4.3. Let $\mathscr{S}_n$ be the collection of all zero-sets $X \subseteq \mathbb{R}^n$ of polynomials with coefficients in $\mathbb{Z}$, then $\mathscr{S} = \langle \mathscr{S}_n \mid n \in \mathbb{N} \rangle$ is a W-structure, indeed the minimal one.

The following definition is different from the corresponding one in [15] but equivalent to it.

Definition 4.4. Let $\mathscr{S}$ be an o-minimal W-structure. The *Charbonnel closure* $\tilde{\mathscr{S}} = \mathrm{Ch}(\mathscr{S}) = \langle \tilde{\mathscr{S}}_n \colon n \in \mathbb{N}^+ \rangle$ of $\mathscr{S}$ is defined as follows:

Ch(base): $\tilde{\mathscr{S}}_n$ is a collection of subsets of $\mathbb{R}^n$ and $\mathscr{S}_n \subseteq \tilde{\mathscr{S}}_n$.
Ch($\cup$): If $A, B \in \tilde{\mathscr{S}}_n$, then $A \cup B \in \tilde{\mathscr{S}}_n$.
Ch($\cap_\ell$): If $A \in \tilde{\mathscr{S}}_n$ and $L \subseteq \mathbb{R}^n$ is the zero-set of a system of linear polynomials with coefficients in $\mathbb{Z}$, then $A \cap L \in \tilde{\mathscr{S}}_n$ (the "ℓ" in the label stands for "linear"). We call such an L a *$\mathbb{Z}$-affine set*.
Ch(π): if $A \in \tilde{\mathscr{S}}_{n+k}$ and $\Pi_n^{n+k} : \mathbb{R}^{n+k} \to \mathbb{R}^n$ is the projection onto the first n coordinates, then $\Pi_n^{n+k}[A] \in \tilde{\mathscr{S}}_n$;
Ch($\bar{x}$): if $A \in \tilde{\mathscr{S}}_n$, then $\bar{A} \in \tilde{\mathscr{S}}_n$, where $\bar{A}$ is the topological closure of A.
$\bigcup_n \tilde{\mathscr{S}}_n$ is minimal with these properties.

Our aim is to prove that if $\mathscr{S}$ is a closed o-minimal W-structure, then $\tilde{\mathscr{S}}$ is a semi-closed o-minimal W-structure. The same conclusion would be valid if $\mathscr{S}$ were only assumed to be semi-closed, but we do not need this fact.

Sometimes we write $A \in \tilde{\mathscr{S}}$ instead of $A \in \tilde{\mathscr{S}}_n$ if n is implicit or irrelevant. Similar conventions apply to $\mathscr{S}$.

Definition 4.5. A *Ch-description* of $A \in \tilde{\mathscr{S}}$ is an expression which illustrates one of the possible ways to obtain A from sets in $\mathscr{S}$ using the Ch-operations Ch($\cup$), Ch($\cap_\ell$), Ch(π) and Ch($\bar{x}$). More precisely, we fix a set Σ of symbols (called *labels*) of the same cardinality as $\bigcup_n \mathscr{S}_n$ and a surjection from Σ to $\bigcup_n \mathscr{S}_n$, so that every set $A \in \mathscr{S}$ has a label $\mathsf{A} \in \Sigma$ (possibly not unique). If A is a label for the set $A \in \mathscr{S}$, then A is a Ch-description of A. Inductively, if B, C are Ch-descriptions of the sets $B, C \in \tilde{\mathscr{S}}$, and if L is a label for a $\mathbb{Z}$-affine set L, then the strings of symbols $(\mathsf{B} \cup \mathsf{C}), (\mathsf{B} \cap \mathsf{L}), \bar{\mathsf{B}}$, and $\Pi_n^{n+k} \mathsf{B}$ (the last one includes the strings needed to define the integers n, k) are Ch-descriptions of the sets $(B \cup C), (B \cap L), \bar{B}$ and $\Pi_n^{n+k} B$ respectively. So for instance the expressions $(\mathsf{B} \cup \mathsf{C}) \cap \mathsf{L}$ and $(\mathsf{B} \cap \mathsf{L}) \cup (\mathsf{C} \cap \mathsf{L})$ (where we have omitted the external parenthesis), are two different Ch-descriptions for the same set of $\tilde{\mathscr{S}}$.

The *rank* ρ of a Ch-description of A is defined as follows:

- If $\mathsf{A} \in \Sigma$ is a label, then $\rho(\mathsf{A}) = 0$;
- $\rho(\mathsf{B} \cup \mathsf{C}) = 1 + \max\{\rho(\mathsf{B}), \rho(\mathsf{C})\}$;
- $\rho(\mathsf{B} \cap \mathsf{L}) = 1 + \rho(\mathsf{B})$;
- $\rho(\prod_n^{n+k} \mathsf{B}) = 1 + \rho(\mathsf{B})$;
- $\rho(\bar{\mathsf{B}}) = 4 + \rho(\mathsf{B})$.

Finally we define the rank $\rho(A)$ of a set $A \in \tilde{\mathscr{S}}$ as the least possible rank of a Ch-description of A.

Thus the sets of rank zero are exactly the sets in $\mathscr{S}$, but there are Ch-descriptions of sets in $\mathscr{S}$ of arbitrarily high rank. Note that the equalities in the definition of the rank of a Ch-description become inequalities if we refer to the sets rather than their descriptions. For instance $\rho(A \cup B) \leqslant 1 + \max\{\rho(A), \rho(B)\}$ and the inequality can be strict since the *set* $A \cup B$ can admit simpler Ch-descriptions besides the one which presents it as the union of A and B. The reason why we need to let the operation $\mathrm{Ch}(\bar{x})$ raise the rank so much will be clear in the proof of Lemma 4.10.

Remark 4.6. Since $\mathscr{S}_n \subseteq \tilde{\mathscr{S}}_n$, $\tilde{\mathscr{S}}$ satisfies W(pol). Notice also that, since linear bijections (induced by a permutation of the variables) commute with union, intersection, projection and closure, $\tilde{\mathscr{S}}$ satisfies W(perm). Moreover, by an application of W(perm) the rank of a Ch-description does not increase.

Given a closed o-minimal W-structure $\mathscr{S}$, to prove that $\tilde{\mathscr{S}}$ is a W-structure it remains to show that it verifies W($\times$) and W($\cap$). This will be done by induction on the rank.

Lemma 4.7. *If $X \in \tilde{\mathscr{S}}_m$ and $Y \in \tilde{\mathscr{S}}_n$, then $X \times Y \in \tilde{\mathscr{S}}_{m+n}$. Moreover $\rho(X \times Y) \leqslant \rho(X) + \rho(Y)$.*

Proof. We prove by induction on $\rho(\mathsf{X}) + \rho(\mathsf{Y})$ the following stronger result: if X, Y are Ch-descriptions of X, Y, then there is a Ch-description $\mathsf{X} \times \mathsf{Y}$ of $X \times Y$ such that $\rho(\mathsf{X} \times \mathsf{Y}) = \rho(\mathsf{X}) + \rho(\mathsf{Y})$. We use the following facts:

- $\mathscr{S}$ is closed under $\times$. This handles the case $\rho(\mathsf{X}) + \rho(\mathsf{Y}) = 0$.
- $$(A \cup B) \times Z = (A \times Z) \cup (B \times Z).$$

  This settles the case when one of the two Ch-descriptions X, Y is obtained from descriptions of smaller rank by the operation Ch($\cup$), say $\mathsf{X} = \mathsf{A} \cup \mathsf{B}$ and $\mathsf{Y} = \mathsf{Z}$ (the symmetric case is handled, both here and below, by permuting the variables). In fact, $\rho(\mathsf{A}) + \rho(\mathsf{Z})$ and $\rho(\mathsf{B}) + \rho(\mathsf{Z})$ are strictly smaller than $\rho(\mathsf{A} \cup \mathsf{B}) + \rho(\mathsf{Z})$, hence by induction $(A \times Z)$ and $(B \times Z)$ are in $\tilde{\mathscr{S}}$ and have Ch-descriptions of the prescribed rank. An application of Ch($\cup$) puts $(A \times Z) \cup (B \times Z)$ in $\tilde{\mathscr{S}}$. The correct evaluation of the ranks follows from an easy computation: $\rho(\mathsf{X} \times \mathsf{Y}) = 1 + \max\{\rho(\mathsf{A}) + \rho(\mathsf{Y}), \rho(\mathsf{B}) + \rho(\mathsf{Y})\} = 1 + \max\{\rho(\mathsf{A}), \rho(\mathsf{B})\} + \rho(\mathsf{Y}) = \rho(\mathsf{X}) + \rho(\mathsf{Y})$.

- $$(A \cap L) \times Z = (A \times Z) \cap (L \times \mathbb{R}^n),$$

  where $Z \subseteq \mathbb{R}^n$. This handles the case when one of the two descriptions is obtained from a description of smaller rank by the operation $\mathrm{Ch}(\cap_\ell)$. It is important to note that, if L is $\mathbb{Z}$-affine, so is $L \times \mathbb{R}^n$.
- $$Z \times \Pi_n^{n+k} A = \Pi_{m+n}^{m+n+k}(Z \times A),$$

  where $Z \subseteq \mathbb{R}^m$. This handles the case when one of the two descriptions is obtained from a description of smaller rank by the operation $\mathrm{Ch}(\pi)$.
- It remains to show that if $A \in \tilde{\mathscr{S}}_m$ and $Z \in \tilde{\mathscr{S}}_n$, then $\bar{A} \times Z \in \tilde{\mathscr{S}}_{m+n}$. If Z has a description Z obtained from descriptions of smaller rank by one of the operations considered above, then we are in one of the preceding cases. In the remaining cases Z is either a description for a set in $\mathscr{S}$, or for a set of the form $\bar{B}$. In any case Z is a closed set, so we can write

  $$\bar{A} \times Z = \overline{A \times Z}.$$

  Note that $\rho(\mathsf{A}) + \rho(\mathsf{Z}) < \rho(\bar{\mathsf{A}}) + \rho(\mathsf{Z})$, hence by induction $A \times Z$ is in $\tilde{\mathscr{S}}$ and we conclude by an application of $\mathrm{Ch}(\bar{x})$. □

Lemma 4.8. *If $A \in \tilde{\mathscr{S}}_n$ and $B \in \tilde{\mathscr{S}}_n$ then $A \cap B \in \tilde{\mathscr{S}}_n$. Moreover $\rho(A \cap B) \leqslant 2 + \rho(A) + \rho(B)$.*

Proof. $A \cap B = \Pi_n^{2n}[(A \times B) \cap \Delta]$ where $\Delta \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is the diagonal $\{(\vec{x}, \vec{x}) \mid \vec{x} \in \mathbb{R}^n\}$. The estimate on the rank follows from Lemma 4.7. □

Lemma 4.9. *$\tilde{\mathscr{S}}$ is semi-closed, i.e. if $A \in \tilde{\mathscr{S}}_n$, then there exist $k \in \mathbb{N}$ and a closed set $B \in \tilde{\mathscr{S}}_{n+k}$ such that $A = \Pi_n^{n+k}[B]$.*

Proof. By induction on the rank of a Ch-description of A, using the following facts:

- If $A \in \mathscr{S}$ or A is obtained by an application of $\mathrm{Ch}(\bar{x})$, there is nothing to prove, since A is already closed.
- If $X = \Pi_n^{n+k}[B]$ and $Y = \Pi_n^{n+h}[C]$, then $X \cup Y = \Pi_n^{n+k+h}[(B \times \mathbb{R}^h) \cup (C \times \mathbb{R}^k)]$. This handles the case when $A = X \cup Y$ is obtained by an application of $\mathrm{Ch}(\cup)$.
- If $X = \Pi_n^{n+k}[B]$ and L is $\mathbb{Z}$-affine, then $X \cap L = \Pi_n^{n+k}[B \cap (L \times \mathbb{R}^k)]$. This handles the case when $A = X \cap L$ is obtained by an application of $\mathrm{Ch}(\cap_\ell)$.
- $\Pi_n^{n+h} \circ \Pi_{n+h}^{n+h+k}[B] = \Pi_n^{n+h+k}[B]$. This handles the case when A is obtained by an application of $\mathrm{Ch}(\pi)$. □

The fact that $\tilde{\mathscr{S}}$ is semi-closed will be useful in Section 9 to prove Theorem 9.1. Let us prove that $\tilde{\mathscr{S}}$ is o-minimal.

Lemma 4.10. *If* $A \in \tilde{\mathscr{S}}_n$, *then* $\gamma(A) < \infty$.

Proof. By induction on the rank of a Ch-description of A, using the following facts:

- $\gamma(B \cup C) \leqslant \gamma(B) + \gamma(C)$.
- If L is $\mathbb{Z}$-affine, $\gamma(B \cap L) \leqslant \gamma(B)$.
- $\gamma(\Pi_n^{n+k} B) \leqslant \gamma(B)$.
- $\gamma(\bar{B}) \leqslant \gamma((B \times \mathbb{R}^{m+2}) \cap E)$, where $m = n^2 + n$, E is the semi-algebraic set $\{(\vec{x}, \vec{y}, R, \varepsilon) \in \mathbb{R}^{n+m+2} : |p(\vec{x}, \vec{y})| < \varepsilon^2 \wedge \sum_{i=1}^n x_i^2 < R^2\}$, and p is a polynomial with coefficients in $\mathbb{Z}$ with the property that every subset of $\mathbb{R}^n$ defined by a system of linear polynomials over $\mathbb{R}$ is of the form $\{\vec{x} \mid p(\vec{x}, \vec{y}) = 0\}$ for a suitable $\vec{y}$.

The existence of p and the proof that $\gamma(\bar{B}) \leqslant \gamma((B \times \mathbb{R}^{m+2}) \cap E)$ is in [9, Claim 1.9]. Since E is semi-algebraic, it is the projection of an algebraic set, and moreover in our case it is the projection of the zero-set of a polynomial with coefficients in $\mathbb{Z}$. It thus follows that E is a set in $\tilde{\mathscr{S}}$ of rank at most 1. So by Lemmas 4.7 and 4.8, the rank of the Ch-description $(\mathsf{B} \times \mathsf{R}^{m+2}) \cap \mathsf{E}$ of the set $(B \times \mathbb{R}^{m+2}) \cap E$ is strictly smaller than $\rho(\bar{\mathsf{B}})$. It is now clear how to complete the proof by induction. □

We have thus proved:

Theorem 4.11. *If* $\mathscr{S}$ *is a closed o-minimal W-structure, then its Charbonnel closure* $\tilde{\mathscr{S}}$ *is a semi-closed o-minimal W-structure.*

5. Effective W-structures

Let $\tilde{\mathscr{S}}$ be the Charbonnel closure of a closed o-minimal W-structure $\mathscr{S}$. We have seen that each set $A \in \tilde{\mathscr{S}}_n$ admits a Ch-description which shows how to obtain it from sets in $\mathscr{S}$. If we now assume that each set in $\mathscr{S}$ admits a description as a string of symbols from a finite alphabet (this implies in particular that each $\mathscr{S}_n$ is countable), then a Ch-description of a set $A \in \tilde{\mathscr{S}}$ becomes itself a string of symbols from a finite alphabet, and it makes sense to ask whether an upper bound on $\gamma(A)$ can be effectively found from the description of A. We will see that the answer is positive if we make some natural assumptions on how the sets in the $\mathscr{S}$ are described.

Definition 5.1. Let $\mathscr{S}$ be a W-structure such that each $\mathscr{S}_n$ is countable. A *coding* of $\mathscr{S}$ is a surjective map $\mathscr{I} : \mathsf{Expr} \to \bigcup_n \mathscr{S}_n$, where Expr is a recursive set of strings of symbols from some finite alphabet. If $\mathscr{I}(\mathsf{A}) = A \in \mathscr{S}_n$, we say that A is a *W-formula* for A. We say that $(\mathscr{S}, \mathscr{I})$ is an *effective W-structure* if the following conditions hold:

EW(sort): There is a recursive function which, given a W-formula for $A \in \bigcup_n \mathscr{S}_n$, computes the unique integer n (called the *sort* of A) such that $A \in \mathscr{S}_n$.

EW(pol): There is a recursive function which, given the coefficients of a system of polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$, compute a W-formula for the zero-set of the system.

EW(perm): There is a recursive function which, given a W-formula for $A \in \mathscr{S}_n$ and a permutation σ of $\{1, \ldots, n\}$, computes a W-formula for the set $\Sigma[A]$, where $\Sigma : \mathbb{R}^n \to \mathbb{R}^n$ is the linear bijection induced by the permutation σ on the coordinates.

EW($\cap$): There is a recursive function which, given W-formulas for the sets $A, B \in \mathscr{S}_n$, computes a W-formula for the set $A \cap B$.

EW($\times$): There is a recursive function which, given W-formulas for $A \in \mathscr{S}_n$ and $B \in \mathscr{S}_m$ computes a W-formula for the set $A \times B$.

An effective *o-minimal* W-structure satisfies furthermore:

EW(o-min): There is a recursive function such that, given a W-formula for $A \in \mathscr{S}_n$, computes an upper bound for $\gamma(A)$.

Example 5.2. An example of an effective o-minimal W-structure $\mathscr{S}$ is the following: let $\mathscr{S}_n$ be the collection of all the subsets X of $\mathbb{R}^n$ such that X is the zero-set of a system of exponential polynomials with coefficients in $\mathbb{Z}$. A W-formula for X is any of the systems defining it. By the results of [7] this coding turns $\mathscr{S}$ into an effective o-minimal W-structure.

Our aim is to show that the Charbonnel closure of a closed effective o-minimal W-structure is an effective o-minimal W-structure with respect to an induced coding which we now describe.

Definition 5.3. Let $(\mathscr{S}, \mathscr{I})$ be an effective W-structure. The notion of *Ch-formula* for a set in $\tilde{\mathscr{S}}$ is defined exactly as the notion of Ch-description (Definition 4.5) with the further requirement that the sets of $\mathscr{S}$ are labeled by their W-formulas (so a Ch-formula is a string of symbols from a finite alphabet). So for instance if B is a W-formula for $B \in \mathscr{S}_n$ and C is a W-formula for $C \in \mathscr{S}_n$, then B and C are Ch-formulas for B and C respectively, and the expression $(\mathsf{B} \cup \mathsf{C})$ is a Ch-formula for the set $B \cup C \in \tilde{\mathscr{S}}_n$.

We define a surjective map $\tilde{\mathscr{I}}$ from the set of all Ch-formulas to $\bigcup_n \tilde{\mathscr{S}}_n$ as follows: $\tilde{\mathscr{I}}(\mathsf{A}) = A$ if A is a Ch-formula for the set A. We call $\tilde{\mathscr{I}}$ the coding induced by $\mathscr{I}$.

We define the notion of rank of a Ch-formula exactly as the rank of a Ch-description and we use the same notation.

Theorem 5.4. *If $(\mathscr{S}, \mathscr{I})$ is a closed effective o-minimal W-structure, then $(\tilde{\mathscr{S}}, \tilde{\mathscr{I}})$ is an effective o-minimal W-structure which is semi-closed.*

Proof. It suffices to follow the proof of Theorem 4.11 and notice that one can extract from it the additional information required: for instance the proof of Lemma 4.7 actually shows that there is a recursive function which, given Ch-formulas for $X \in \tilde{\mathscr{S}}_n$ and $Y \in \tilde{\mathscr{S}}_m$ yields a Ch-formula for $X \times Y$; the proof of Lemma 4.10 gives a recursive function which, given a Ch-formula for $A \in \tilde{\mathscr{S}}_n$, computes an upper bound for $\gamma(A)$. □

Remark 5.5. Following the proof of Lemma 4.9 it can also be shown that $\tilde{\mathscr{S}}$ is *effectively semi-closed*, i.e. given a Ch-formula for $A \in \tilde{\mathscr{S}}_n$, we can effectively find $k \in \mathbb{N}$ and the Ch-formula for a closed set $B \in \tilde{\mathscr{S}}_{n+k}$ such that $A = \Pi_n^{n+k}[B]$.

6. Smooth approximation of the boundary

The results of this section correspond to the ones in [15, Section 3]. We include the proofs both to keep the paper self-contained and because we work with a slightly different definition of W-structure (see Remark 6.8). Moreover we find it convenient to give a definition of approximation (our Definition 6.4, corresponding to [15, Definition 3.2]) and a proof of the approximation theorem (Theorem 6.11) which does not make explicit use of Wilkie's notion of "moduli". We replace the moduli by a systematic use of the quantifier "for all sufficiently small" (the moduli are essentially the Skolem functions associated to the quantifiers).

Let $\mathscr{S}$ be a closed o-minimal W-structure.

Definition 6.1.

- We say that $\mathscr{S}$ is *determined by its smooth functions* (DSF) if, given a set $A \in \mathscr{S}_n$, there exist $k \in \mathbb{N}$ and a C^∞-function $f_A : \mathbb{R}^{n+k} \to \mathbb{R}$ whose graph lies in $\mathscr{S}$, such that A is the projection onto the first n coordinates of the zero-set of f_A.
- Moreover, $\mathscr{S}$ is *effectively determined by its smooth functions* (EDSF) if $\mathscr{S}$ is an effective W-structure and there is an algorithm that, given a Ch-formula A for A, yields k and a Ch-formula for the graph of f_A.

Example 6.2. Let $\mathscr{S}_n$ be the collection of all zero-sets $X \subseteq \mathbb{R}^n$ of exponential polynomials with coefficients in $\mathbb{Z}$. Then $\mathscr{S} = \langle \mathscr{S}_n \mid n \in \mathbb{N} \rangle$ is an effective W-structure, which is EDSF. More generally, let $\mathscr{S}_n$ be the collection of all quantifier free definable sets of the structure $\mathscr{R}$, where $\mathscr{R}$ is an expansion of $(\mathbb{R}, +, \cdot)$ with finitely many C^∞ functions satisfying a Khovanskii-type result, namely for which there are recursive bounds on the number of connected components (actually on γ) of quantifier free definable sets. Then $\mathscr{S} = \langle \mathscr{S}_n \mid n \in \mathbb{N} \rangle$ is an effective W-structure, which is EDSF (we argue as in [15, Theorem 1.9] eliminating negations and compositions by introducing existential quantifiers).

Recall that the aim is to prove that the Charbonnel closure of $\mathscr{S}$ coincides with the definable closure of $\mathscr{S}$. Since $\tilde{\mathscr{S}}$ is closed under finite unions and projections, it remains to show that $\tilde{\mathscr{S}}$ is closed under complementation (Theorem 7.11). We still have not proved that if $A \in \tilde{\mathscr{S}}$ then the boundary $\partial A = \bar{A} \setminus \mathrm{int}(A)$ is in $\tilde{\mathscr{S}}$. Anyway, by assuming the DSF condition we are able to confine the boundary of a closed set $A \in \tilde{\mathscr{S}}_n$, into a closed set $B \in \tilde{\mathscr{S}}_n$ with empty interior, and this will suffice to prove the stability of $\tilde{\mathscr{S}}$ under complementation (this will be clear in Section 6). The set B will be obtained as the projection of a sort of "limit of smooth manifolds", by a procedure described in [15]. Moreover, if $\mathscr{S}$ is EDSF, the Ch-description of B can be effectively found from a Ch-description of A. The main difficulty lies in the attempt to confine the boundary of the projection of a set (and this is the reason why we need a *smooth* approximation). For this we make use of Lemma 10.6, which is a variant of the fact that, for a smooth function f having zero as a regular value with compact preimage, the boundary in $\mathbb{R}^n$ of a set of the form $\{\vec{x} \mid \mathbb{R}^n : \exists x_{n+1}(f(\vec{x}, x_{n+1}) = 0)\}$ is contained in the set with empty interior $\{\vec{x} \mid \mathbb{R}^n : \exists x_{n+1}(f(\vec{x}, x_{n+1}) = 0 \wedge (\partial f / \partial x_{n+1})(\vec{x}, x_{n+1}) = 0)\}$.

To give the precise notion of limit, we need some definitions and lemmas.

Definition 6.3. Let $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$.

- Given $A \subseteq \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}_+$, define the *$\varepsilon$-neighborhood* A^ε of A as the set $\{x \in \mathbb{R}^n \mid \exists y \in A |x - y| < \varepsilon\}$.
- The *Hausdorff distance* $d(A, B)$ between two subsets A, B of $\mathbb{R}^n$ is the infimum of all the $\varepsilon \in \mathbb{R}_+$ such that the ε-neighborhood of each set contains the other.
- (The quantifier "for all sufficiently small") We write $\forall^s \varepsilon \phi$ as a shorthand for $\exists \mu \forall \varepsilon < \mu \phi$, where μ, ε are always assumed to range in $\mathbb{R}_+$. These quantifiers can be iterated: so $\forall^s \varepsilon_1 \forall^s \varepsilon_2 \phi$ abbreviates $\exists \mu_1 (\forall \varepsilon_1 < \mu_1) \exists \mu_2 (\forall \varepsilon_2 < \mu_2) \phi$, which is not the same as $\forall^s \varepsilon_2 \forall^s \varepsilon_1 \phi$. The expression $\forall^s \varepsilon_1, \ldots, \forall^s \varepsilon_k \phi$ can be read as: ϕ holds for all sufficiently small $\varepsilon_1, \ldots, \varepsilon_k$ provided each ε_i with $i > 1$ is also sufficiently small with respect to the preceding ones.
- (Sections) Given $S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$ and given $\varepsilon_1, \ldots, \varepsilon_k \in \mathbb{R}^+$, we define $S_{\varepsilon_1, \ldots, \varepsilon_k}$ as the set $\{x \in \mathbb{R}^n \mid (x, \varepsilon_1, \ldots, \varepsilon_k) \in S\}$.

The Hausdorff distance is a metric if we restrict to compact subsets of $\mathbb{R}^n$. In this case $\lim_{t \to 0} A_t = B$ if $\forall^s \varepsilon \forall^s t (B \subseteq A_t^\varepsilon \wedge A_t \subseteq B^\varepsilon)$. This is equivalent to $\forall^s \varepsilon \forall^s t (B \subseteq A_t^\varepsilon) \wedge \forall^s \varepsilon \forall^s t (A_t \subseteq B^\varepsilon)$.

Definition 6.4 (Wilkie [15, Definition 3.2]). Let $A \subseteq \mathbb{R}^n, S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$.

(1) S approximates A from below if

$$\forall^s \varepsilon_0 \forall^s \varepsilon_1 \ldots \forall^s \varepsilon_k (S_{\varepsilon_1, \ldots, \varepsilon_k} \subseteq A^{\varepsilon_0}).$$

(2) S approximates A from above on bounded sets if

$$\forall^s \varepsilon_0 \forall^s \varepsilon_1 \ldots \forall^s \varepsilon_k (A \cap B(0, 1/\varepsilon_0) \subseteq S_{\varepsilon_1, \ldots, \varepsilon_k}^{\varepsilon_0})$$

where $B(0, 1/\varepsilon_0) \subseteq \mathbb{R}^n$ is the compact ball of radius $1/\varepsilon_0$ centered at the origin.

Note that if A is bounded, we can omit in the above definition the intersection with the compact ball, and we recover in the special case $k = 1$ the limit in the Hausdorff distance.

Definition 6.5. Let $M(\mathscr{S}) = \bigcup_n M_n(\mathscr{S})$, where $M_n(\mathscr{S})$ is the smallest ring of functions from $\mathbb{R}^n$ to $\mathbb{R}$ closed under partial differentiation and containing:

- all polynomials $p \in \mathbb{Z}[x_1, \ldots, x_n]$;
- all functions f_A, for $A \in \mathscr{S}$, which provide the DSF condition for $\mathscr{S}$ (see Definition 6.1);
- the functions $(x_1, \ldots, x_n) \mapsto (1 + x_{i_1}^2 + \ldots + x_{i_s}^2)^{-1}$, with $s \leqslant n$ and $1 \leqslant i_1 \ldots i_s \leqslant n$.

Note that every function in $M(\mathscr{S})$ is C^∞ and we have $M(\mathscr{S}) \subseteq \tilde{\mathscr{S}}$, in the sense that if $f \in M_n(\mathscr{S})$, then the graph of f is in $\tilde{\mathscr{S}}_{n+1}$. In fact in [9, Lemma 4.11] it is proved that if $f \in \tilde{\mathscr{S}}_{n+1}$ is a C^1 function, then all partial derivatives $\partial f / \partial x_i$ belong to $\tilde{\mathscr{S}}_{n+1}$.

The idea is to simulate the limit of the differential quotient using sections and the topological closure:

$$\mathrm{Graph}(\partial f/\partial x_i) = (\overline{\{(\vec{x}, y, \varepsilon) \mid y\varepsilon = f(\vec{x} + \vec{\varepsilon_i}) - f(\vec{x})\}})_0,$$

where $\vec{\varepsilon_i} = (0, \ldots, 0, \varepsilon, 0, \ldots, 0)$, with ε in the ith coordinate and we have used the notation $X_0 = \{\vec{u} \mid (\vec{u}, 0) \in X\}$. This also shows that, given a Ch-description for f, we can effectively find a Ch-description for $\partial f/\partial x_i$.

Definition 6.6. An $M(\mathscr{S})$-*constituent* is a set of the form

$$\{(\vec{x}, \vec{\varepsilon}) \in \mathbb{R}^n \times \mathbb{R}^k_+ \mid \exists \vec{y} \in \mathbb{R}^{k-1} F(\vec{x}, \vec{y}) = \vec{\varepsilon}\},$$

where $F : \mathbb{R}^{n+k-1} \to \mathbb{R}^k$ belongs to $M(\mathscr{S})^k$. An $M(\mathscr{S})$-*set* $S \subseteq \mathbb{R}^n \times \mathbb{R}^k_+$ is a finite union of $M(\mathscr{S})$-constituents (with the same k).

Given a set $A \in \tilde{\mathscr{S}}_n$ and an $M(\mathscr{S})$-set $S \subseteq \mathbb{R}^{n+k}$, we say that S is an $M(\mathscr{S})$-*approximant* for A if S both approximates $\partial \bar{A}$ from above on bounded sets and approximates $\bar{A}$ from below.

Lemma 6.7. *Every* $M(\mathscr{S})$*-set* $S \subseteq \mathbb{R}^{n+k}$ *is in* $\tilde{\mathscr{S}}_{n+k}$ *and has empty interior.*

Proof. The fact that $S \in \tilde{\mathscr{S}}_{n+k}$ depends on the inclusion $M(\mathscr{S}) \subseteq \tilde{\mathscr{S}}$ and the closure properties of $\tilde{\mathscr{S}}$. To show that every such set S has empty interior, first recall that, as a consequence of Sard's Theorem, the image of a C^∞-function $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m > n$ has empty interior. Now, let $T = \{(\vec{x}, \vec{\varepsilon}) \mid \exists \vec{y} \in \mathbb{R}^{k-1} F(\vec{x}, \vec{y}) = \vec{\varepsilon}\}$ be an $M(\mathscr{S})$-constituent of S and, for each fixed $\vec{x}$, consider the fiber $T_{\vec{x}} = \{\vec{\varepsilon} \mid (\vec{x}, \vec{\varepsilon}) \in T\}$ over $\vec{x}$. Note that $T_{\vec{x}}$ is the (positive part of the) image of the C^∞-function $h : \mathbb{R}^{k-1} \to \mathbb{R}^k$ which sends $\vec{y}$ to $F(\vec{x}, \vec{y})$, hence for every $\vec{x}$, $T_{\vec{x}}$ has empty interior by the remark above. It follows that T (and hence S) has empty interior. □

Remark 6.8. Our definition of W-structure is different from the corresponding definitions given in [1,9,15], where it is required that $\mathscr{S}$ contains all real semi-algebraic sets. Nevertheless we can apply to our $\tilde{\mathscr{S}} = \mathrm{Ch}(\mathscr{S})$ all the results of these authors concerning the regularity properties of the sets in (their) $\tilde{\mathscr{S}}$ (e.g. the fact that if a set $A \in \tilde{\mathscr{S}}$ has empty interior, then so does its closure $\bar{A}$). The reason is the following. Let $\mathscr{S}^*_n$ be the collection of all sets of the form $A \cap L$ where $A \in \mathscr{S}_n$ and $L \subseteq \mathbb{R}^n$ is defined by a system of linear equations with coefficients in $\mathbb{R}$. We call $\mathscr{S}^* = \langle \mathscr{S}^*_n \mid n \in \mathbb{N}\rangle$ the enlargement of $\mathscr{S}$ with parameters from $\mathbb{R}$. Next, define $\widehat{\mathscr{S}^*}$ as the closure of $\mathscr{S}^*$ under the Ch-operation $\mathrm{Ch}(\pi)$. It can be readily verified that, if $\mathscr{S}$ is a closed o-minimal W-structure with DSF, then $\langle \widehat{\mathscr{S}^*_n} \mid n \in \mathbb{N}\rangle$ is a semi-closed o-minimal W-structure with DSF, and since $\mathscr{S}^*$ contains all real semi-algebraic sets, $\widehat{\mathscr{S}^*}$ is also a weak structure in Wilkie's sense. Moreover $\tilde{\mathscr{S}} \subseteq \mathrm{Ch}(\widehat{\mathscr{S}^*})$, so we can apply to our $\tilde{\mathscr{S}}$ the regularity results of $\mathrm{Ch}(\widehat{\mathscr{S}^*})$. To prove the DSF condition for $\widehat{\mathscr{S}^*}$ note that a generic set in $\widehat{\mathscr{S}^*}$ is of the form $\Pi^{n+k}_n[A \cap L]$, where $A \in \mathscr{S}$ and L is the zero-set of a system of linear polynomial $p_1, \ldots, p_r$ over $\mathbb{R}$; the DSF condition for $\mathscr{S}$ provides a C^∞-function f_A with graph in $\mathscr{S}$ such that $A = \Pi^{n+k+h}_{n+k}[V(f_A)]$ (we recall that $V(f)$ is the zero-set of

f); then the function $g = f_A^2 + \sum_i p_i^2$ is C^∞ with graph in $\widehat{\mathscr{S}^*}$ (note that the graph of the square of a function is existentially definable) and $\Pi_n^{n+k}[A \cap L] = \Pi_n^{n+k+h}[V(g)]$.

We need the following result of Charbonnel [1].

Lemma 6.9.
- *If $A \in \tilde{\mathscr{S}}_n$ has empty interior, then so does $\bar{A}$.*
- *If $A \in \tilde{\mathscr{S}}_{n+1}$ and $A \subseteq \mathbb{R}^n \times \mathbb{R}_+$ then $\bar{A}_0 = \{\vec{x} \in \mathbb{R}^n \mid (\vec{x}, 0) \in A\} \in \tilde{\mathscr{S}}_n$, and if A has no interior points nor does $\bar{A}_0$.*

Proof. See [1] or also [9, Lemma 2.7] for the first statement and [15, Lemma 2.2] for the second. The proof depends on the o-minimality condition for $\tilde{\mathscr{S}}$. □

From Lemmas 6.7 and 6.9 we obtain

Lemma 6.10. *Suppose $S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$ is an $M(\mathscr{S})$-set. Then the section $\bar{S}_{\vec{0}} = \{\vec{x} \in \mathbb{R}^n \mid (\vec{x}, \vec{0}) \in \bar{S}\}$ is closed, lies in $\tilde{\mathscr{S}}_n$, and has empty interior.*

Theorem 6.11.
- *Suppose $\mathscr{S}$ is DSF; then, every set $A \in \tilde{\mathscr{S}}_n$, has an $M(\mathscr{S})$-approximant $S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$ for some $k \geqslant 0$.*
- *Moreover, if $\mathscr{S}$ is EDSF, then there is an algorithm which, given a Ch-description for A, produces a Ch-description for S.*

The first part is in [15, Theorem 3.13], except that we are working with a slightly different definition of $\mathscr{S}$ and $\tilde{\mathscr{S}}$. From the analysis of the proof it is easy to obtain the second part. We include a proof in the last section.

A weaker form of Theorem 6.11—i.e. given a set in $A \in \tilde{\mathscr{S}}_n$ we can find an $M(\mathscr{S})$-set $S \subseteq \mathbb{R}^{n+k}$ (for some k) such that S approximates $\partial\bar{A}$ from above on bounded sets—would be enough to our purposes, but we are not able to prove the weaker statement without proving the statement of Theorem 6.11 first.

Let us prove the main theorem of this section (corresponding to [15, Theorem 3.1]).

Theorem 6.12.
- *Let $\mathscr{S}$ be a closed o-minimal W-structure which is DSF. Then, given a closed set $A \subseteq \mathbb{R}^n$ in $\tilde{\mathscr{S}}$, there exists a closed set $B \subseteq \mathbb{R}^n$ in $\tilde{\mathscr{S}}$ such that B has empty interior and $\partial A \subseteq B$.*
- *Furthermore, if $\mathscr{S}$ is EDSF, then there is an effective procedure which, given a Ch-description for A, produces a Ch-description for B.*

Proof. Given a set $A \in \tilde{\mathscr{S}}_n$, we can find an $M(\mathscr{S})$-approximant $S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$ for A as in Theorem 6.11. So in particular S approximates $\partial\bar{A} \in \mathbb{R}^n$ from above on bounded sets. But then so does the section $\bar{S}_{\vec{0}} = \{\vec{x} \in \mathbb{R}^n \mid (\vec{x}, \vec{0}) \in \bar{S}\}$ (see the proof of Lemma 3.3 in [15]). Moreover, the set $\bar{S}_{\vec{0}}$ is closed, lies in $\tilde{\mathscr{S}}_n$ and has empty interior, by Lemma 6.10. Hence we can set $B = \bar{S}_{\vec{0}}$.

As to the effectiveness of this procedure, in case $\mathscr{S}$ is EDSF, notice that given a Ch-description for A, we can effectively find a Ch-description for the set S, from which we can easily compute a Ch-description for the set $\bar{S}_{\vec{0}}$. □

7. Cell decomposition

We give a presentation of Wilkie's cell decomposition omitting some details of the proofs but emphasizing the definitions implicit in the proofs. We will refer to such definitions in the next section, where we will give an effective non-deterministic version of these results.

Fix a W-structure $\mathscr{S}$ which is DSF and let $\tilde{\mathscr{S}}$ be the Ch-closure of $\mathscr{S}$.

Definition 7.1. Given $A \in \tilde{\mathscr{S}}_n$, consider the set with empty interior $B \in \tilde{\mathscr{S}}_n$ with $\partial A \subseteq B$ given by Theorem 6.12 and define $A^* \in \tilde{\mathscr{S}}_n$ as $B \cap \bar{A}$. So $\partial A \subseteq A^* \subseteq \bar{A}$ and A^* has empty interior.

So A^* may contain, besides ∂A, some points in the interior of A.

Definition 7.2. Given $C \in \tilde{\mathscr{S}}_n$ and functions $f : C \to \mathbb{R}$ and $g : C \to \mathbb{R}$, both in $\tilde{\mathscr{S}}_{n+1}$, we denote by $(f)_C$ the graph of f and by $(f, g)_C$ the set $\{(\vec{x}, y) \in C \times \mathbb{R} \mid f(\vec{x}) < y < g(\vec{x})\}$.

In the sequel we identify a function with its graph, so a function $f : \mathbb{R}^n \to \mathbb{R}$ is a subset of $\mathbb{R}^{n+1}$.

Definition 7.3.

(1) A cell in $\mathbb{R}$ is either a singleton $\{a\}$ belonging to $\tilde{\mathscr{S}}_1$ or an interval $(a, b) \in \tilde{\mathscr{S}}_1$.

(2) A cell in $\mathbb{R}^{n+1}$ is either a set of the form $(f)_C$, where $f : C \to \mathbb{R}$ is a continuous function in $\tilde{\mathscr{S}}_{n+1}$ and C is a cell in $\mathbb{R}^n$, or else a set of the form $(f, g)_C$ where C is a cell in $\mathbb{R}^n$ and $f, g : C \to \mathbb{R}$ are continuous bounded functions in $\tilde{\mathscr{S}}_{n+1}$ satisfying $f(\vec{x}) < g(\vec{x})$ for all $\vec{x} \in C$.

The definition of cell depends on $\tilde{\mathscr{S}}$, so our cells are $\tilde{\mathscr{S}}$-cells. According to our definition, which departs from the usual one, a singleton $\{a\} \subseteq \mathbb{R}$ is not necessarily a cell, unless it belongs to $\tilde{\mathscr{S}}_1$ (recall that we did not put in $\mathscr{S}_1$ all the singletons). Similarly an interval (a, b) is a cell only if it belongs to $\tilde{\mathscr{S}}_1$. Note moreover that every cell is bounded, as in [15].

Definition 7.4. Let $D \in \tilde{\mathscr{S}}_n$ be a cell. A cell decomposition $\mathscr{D}$ of D is a partition of D into cells where we require, if $n > 1$, that the projections $\Pi^n_{n-1} E$ of the cells $E \in \mathscr{D}$ form a cell decomposition of $\Pi^n_{n-1} D$ (which is clearly a cell). We say that $\mathscr{D}$ is compatible with a set $A \subseteq \mathbb{R}^n$ if $A \cap D$ is the union of some cells of $\mathscr{D}$. We say that $\mathscr{D}$ is compatible with a finite collection of sets, if it is compatible with each of them.

Remark 7.5. A cell decomposition $\mathscr{D}$ of D which is compatible with A^* is also compatible with $\bar{A}$.

Lemma 7.6. *Let $\mathscr{D}$ and $\mathscr{F}$ be two cell decompositions of the same cell $D \in \tilde{\mathscr{S}}_n$. If $\mathscr{D}$ is compatible with the closure of each cell of $\mathscr{F}$, then $\mathscr{D}$ is compatible with each cell of $\mathscr{F}$.*

Proof. By induction on the definition of cell one shows that given two distinct cells C_0 and C_1 of $\mathscr{F}$, the closure of C_i $(i=0,1)$ does not intersect C_{i-1}. Granted this, if for a contradiction there is a cell E of $\mathscr{D}$ which intersects two distinct cells C_0, C_1 of $\mathscr{F}$, then by the compatibility condition E is included in $\bar{C}_i$ for $i=1,2$. Now $E \cap C_i$ is nonempty and is included in $\overline{C_1}$ and $\overline{C_2}$, so the closure of each C_i intersects C_{i-1}, and we have a contradiction. □

We can now state the cell decomposition theorem:

Theorem 7.7. *Let $n \geqslant 1$ and suppose that D is a cell in $\mathbb{R}^n$. Given a finite collection $\mathscr{A} = \{A_1, \ldots, A_m\}$ of subsets of D which are closed in D and lie in $\tilde{\mathscr{S}}_n$, there exists a cell decomposition $\mathscr{D}$ of D compatible with each set of the collection.*

The proof is by induction on n. The key step in the induction is based on Lemma 7.8 below, which provides the functions needed to define the cells.

Lemma 7.8. *Let $D = (f,g)_C \in \tilde{\mathscr{S}}_{n+1}$ be an open cell (i.e. a cell which is an open subset of $\mathbb{R}^{n+1}$) and let $A \in \tilde{\mathscr{S}}_{n+1}$ be a subset of D which is closed in D. There is a finite collection $\mathscr{H} \subseteq \tilde{\mathscr{S}}_n$ of subsets of C which are closed in C and such that, if $\mathscr{F}$ is a cell decomposition of C compatible with $\mathscr{H}$ and C' is an open cell of $\mathscr{F}$, then:*

(1) *the fibers of A^* over C' (namely the sets $A^*_{\vec{x}} = \{y \in \mathbb{R} \mid (\vec{x}, y) \in A^*\}$ for $\vec{x} \in C'$), have constant finite cardinality $\kappa = \kappa(C') \leqslant \gamma(A^*)$;*

(2) *for $1 \leqslant i \leqslant \kappa$ the function $f_i : C' \to \mathbb{R}$, where $f_i(\vec{x})$ is defined as the ith point in increasing order of $A^*_{\vec{x}}$, is continuous and lies in $\tilde{\mathscr{S}}_{n+1}$ (this is vacuous if $\kappa = 0$).*

The lemma permits us to decompose a "large" subset of $D = (f,g)_C \subseteq \mathbb{R}^{n+1}$ compatibly with A^* provided we can find a decomposition $\mathscr{F}$ of the open cell $C \subseteq \mathbb{R}^n$ as required in the lemma. Indeed, for each open cell $C' \subseteq C$ of $\mathscr{F}$, the functions $f_i : C' \to \mathbb{R}$, together with $f_{|C'}$ and $g_{|C'}$, allow us to define a cell decomposition of $(f,g)_{C'} = (C' \times \mathbb{R}) \cap D$ compatible with A^*. In this way we decompose the union $\bigcup_{C'} (C' \times \mathbb{R}) \cap D$, where C' varies among the open cells of C. This set is large in D in the sense that its relative complement in D has empty interior.

Proof of Lemma 7.8. We must define $\mathscr{H}$ and, for each C', the functions $f_i : C' \to \mathbb{R}$. Let $\{A^* \geqslant i\} \subseteq C$ be the set of points $\vec{x} \in C$ such that the fiber $A^*_{\vec{x}} \subseteq \mathbb{R}$ of A^* over x has cardinality $\geqslant i$. This set is in $\tilde{\mathscr{S}}_n$ since it admits the definition

$$\{A^* \geqslant i\} = \left\{ x \in C \mid \exists y_1, \ldots, y_i \left(y_1 < \cdots < y_i \wedge \bigwedge_{j=1}^{i} (x, y_j) \in A^* \right) \right\}$$

which presents it as a projection of a set in $\tilde{\mathscr{S}}_{n+i}$. Note that if a fiber $A^*_{\vec{x}}$ has cardinality $>\gamma(A^*)$ then it has a nonempty interior. By The Kuratowski–Ulam Theorem (see for example [11, Theorem 15.1]), the set of those points $\vec{x} \in C$ for which this happens has empty interior, as otherwise A^* would have interior. So, by Lemma 6.9, for $N > \gamma(A^*)$ the set $\overline{\{A^* \geqslant N\}}$ has empty interior. Note that $\{A^* \geqslant N\} = \{A^* \geqslant N'\}$ for $N, N' > \gamma(A^*)$. Now consider the following sets:

$$H = \{(\vec{x}, \varepsilon) \in C \times \mathbb{R}_+ \mid \exists y_1, y_2 (y_1 < y_2 \wedge (\vec{x}, y_1) \in A \wedge (\vec{x}, y_2) \in A^* \wedge y_2 - y_1 = \varepsilon)\},$$

$$H_f = \{(\vec{x}, \varepsilon) \in C \times \mathbb{R}_+ \mid \exists y ((\vec{x}, y) \in A^* \wedge y - f(\vec{x}) = \varepsilon)\},$$

$$H_g = \{(\vec{x}, \varepsilon) \in C \times \mathbb{R}_+ \mid \exists y ((\vec{x}, y) \in A^* \wedge g(\vec{x}) - y = \varepsilon)\}.$$

Let $\tilde{H} = \{\vec{x} \in C \mid (\vec{x}, 0) \in \bar{H}\}$, $\tilde{H}_f = \{\vec{x} \in C \mid (\vec{x}, 0) \in \overline{H_f}\}$ and define $\tilde{H}_g$ similarly. Finally define:

$$\mathscr{H} := \{\overline{\{A^* \geqslant 1\}}, \ldots, \overline{\{A^* \geqslant N\}}, \tilde{H}, \tilde{H}_f, \tilde{H}_g\}.$$

Using the fact that A^* is a closed set with empty interior, it is not difficult to prove (see [15, Theorem 4.5]) that given a cell decomposition $\mathscr{F}$ of C compatible with $\mathscr{H}$, and given an open cell C' of $\mathscr{F}$, there is an integer $\kappa = \kappa(C') \leqslant \gamma(A^*)$ such that the set $(C' \times \mathbb{R}) \cap A^*$ is the union of the graphs of κ continuous functions $f_i : C' \to \mathbb{R}$ $(1 \leqslant i \leqslant \kappa)$ with $f_1 < \cdots < f_\kappa$ on C'. Granted this, it remains to prove that $f_i \in \tilde{\mathscr{S}}_{n+1}$. This follows from the following definition

$$f_i = \left\{ (\vec{x}, y) \in C' \times \mathbb{R} \mid \exists y_1, \ldots, y_\kappa \left(y_1 < \cdots < y_\kappa \wedge \bigwedge_{j=1}^{\kappa} (\vec{x}, y_j) \in A^* \wedge y = y_i \right) \right\}$$

which presents f_i as the projection of a set in $\tilde{\mathscr{S}}_{n+1+\kappa}$. □

Remark 7.9. The definition of f_i given above depends on κ and so it is nonconstructive inasmuch as we do not know how to compute $\kappa = \kappa(C')$ given a description of C'. Using negations we could give an alternative definition of f_i which makes no reference to κ: $(\vec{x}, y) \in f_i$ if there are at least i points in the fiber $A^*_{\vec{x}}$ below y, and it is not the case that there are at least $i+1$ points in $A^*_{\vec{x}}$ below y. Unfortunately we cannot give this definition until we prove that $\tilde{\mathscr{S}}$ is stable under complementation.

We are now ready to prove the cell decomposition theorem.

Proof of Theorem 7.7: case $n = 1$. Suppose $n = 1$, namely D is a cell of $\mathbb{R}$. If D is a singleton the result to be proved is obvious, so assume that D is an interval $(a, b) \in \tilde{\mathscr{S}}_1$. Assume first that $m = 1$, namely the collection $\{A_1, \ldots, A_m\}$ contains only one set A. The set $A^* \subseteq D$ is finite since it has empty interior and belongs to $\tilde{\mathscr{S}}_1$. Let κ be the cardinality of A^*. The singleton of the ith point of A^* belongs to $\tilde{\mathscr{S}}_1$ since it admits the definition

$$P_{i,\kappa} = \left\{ y \in \mathbb{R} \mid \exists y_1 \dots y_\kappa \left(y_1 < \cdots < y_\kappa \wedge \bigwedge_{j=1}^{\kappa} y_j \in A^* \wedge y = y_i \right) \right\}$$

which presents $P_{i,\kappa}$ as a projection of a set in $\tilde{\mathscr{S}}_{\kappa+1}$ $(1 \leqslant i \leqslant \kappa)$. Define $P_{0,\kappa} = \{a\}$ and $P_{\kappa+1,\kappa} = \{b\}$. A cell decomposition of D is obtained by considering the singletons $P_{i,\kappa}$ $(1 \leqslant i \leqslant \kappa)$ and the intervals $(P_{i,\kappa}, P_{i,\kappa+1})$ $(0 \leqslant i < i+1 \leqslant \kappa+1)$. This decomposition is compatible with A^*, hence with A. The case $m > 1$ is similar. □

To prove the general case we need:

Lemma 7.10. *For each cell C in $\mathbb{R}^n$ there exists a unique sequence of integers $1 \leqslant i_1 < \cdots < i_d \leqslant n$ such that if we let $\pi : \mathbb{R}^n \to \mathbb{R}^d$ be the projection $\pi(x_1, \dots, x_n) = (x_{i_1}, \dots, x_{i_d})$ we have that the restriction of π to C is an homeomorphism onto an open cell of $\mathbb{R}^d$.*

Proof. The lemma is well known but we give a proof for future reference. A cell of the form $(f)_C$ is homeomorphic to its base C through the projection onto the first coordinates. So if $i_1, \dots, i_d$ is the sequence associated to C, then the same sequence is associated to $(f)_C$, while the sequence associated to a cell of the form $(f,g)_C$ is $i_1, \dots, i_d, i_n$. □

Proof of Theorem 7.7: general case. Assume that the theorem holds in dimension $< n+1$. We prove it for $n+1$, dealing first with the case in which $m = 1$, namely the collection $\{A_1, \dots, A_m\}$ contains only one set A. There are two cases two distinguish.

Case 1: suppose D is a cell of $\mathbb{R}^{n+1}$ which is not open. Then there is $d < n+1$ and integers $1 \leqslant i_1 < \cdots < i_d \leqslant n$ such that the projection $\pi : \mathbb{R}^n \to \mathbb{R}^d$, $(x_1, \dots, x_n) \mapsto (x_{i_1}, \dots, x_{i_d})$, maps D homeomorphically onto an open cell of $\mathbb{R}^d$. By induction we can decompose the image of D under π compatibly with the image of A. The preimages give us the desired decomposition of D.

Case 2: suppose $D = (f,g)_C$ is an open cell of $\mathbb{R}^{n+1}$. Consider the finite collection $\mathscr{H} \subseteq \tilde{\mathscr{S}}_n$ of Lemma 7.8. By induction there is a cell decomposition $\mathscr{F}$ of C compatible with each set in $\mathscr{H}$. If C' is an open cell of $\mathscr{F}$, then we decompose $(f_{|C'}, g_{|C'})_{C'} = (C' \times \mathbb{R}) \cap D$ into cells bounded by the functions $f_{|C'}, g_{|C'}$ and the functions $f_i : C' \to \mathbb{R}$ of Lemma 7.8 $(1 \leqslant i \leqslant \kappa(C'))$. On the other hand if C' is a cell of $\mathscr{F}$ which is not open in $\mathbb{R}^n$, then the cell $(f_{|C'}, g_{|C'})_{C'}$ is not open in $\mathbb{R}^{n+1}$ and we argue as in case 1.

It remains to consider the case in which $m > 1$, namely we want a decomposition of a cell $D \in \tilde{\mathscr{S}}_{n+1}$ compatible with a finite collection $\mathscr{A} \subseteq \tilde{\mathscr{S}}_{n+1}$ of subsets of D which are closed in D. To begin with we apply the construction of case 1 and 2 to each $A \in \mathscr{A}$ separately. We obtain in this way, for each $A \in \mathscr{A}$, a cell decomposition $\mathscr{D}_A$ of D compatible with A. Projecting from $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$ we obtain, for each $A \in \mathscr{A}$, a decomposition $\mathscr{D}_{A,C}$ of $C = \Pi_n^{n+1} D$ compatible with $\Pi_n^{n+1} A$. By induction there is a cell decomposition $\mathscr{D}_C$ of C compatible with the following closed sets: (i) the closures of the cells of the various decompositions $\mathscr{D}_{A,C}$ of C described above; (ii) the sets of the form $\{x \in \mathbb{R}^n \mid h_i(x) = h_j(x)\}$, where h_i, h_j are functions bounding some cell in some of the given decompositions $\mathscr{D}_A$ of D (the nontrivial case is when h_i, h_j belong to different decompositions relative to distinct choices of $A \in \mathscr{A}$). By Lemma 7.6, $\mathscr{D}_C$ is then compatible with the cells of $\mathscr{D}_{A,C}$, not only with their closures. It is now obvious how to lift $\mathscr{D}_C$ to a cell decomposition of D. It suffices to consider cells which are bounded by the same functions $h_i, h_j, \ldots$ as before, but with domain restricted to the appropriate cells of $\mathscr{D}_C$ (this guarantees that the graphs of two different functions do not intersect). □

The cell decomposition theorem is only proved for closed sets. However using the fact that $\tilde{\mathscr{S}}$ is semi-closed we can reduce to closed sets and conclude as in [15, Theorem 1.8]:

Theorem 7.11. *$\tilde{\mathscr{S}}$ is closed under complementation.*

A nonconstructive aspect of the above proof is that it makes use of Lemma 7.8 and therefore it requires the knowledge of the number $\kappa = \kappa(C')$. We observe that $\kappa \leqslant \gamma(A^*)$, so if $\mathscr{S}$ is an effective W-structure we can compute an upper bound on κ. This will suffice to turn the above proof into a "multivalued" algorithm, namely an algorithm which tries systematically all the possible values of κ (inductively) so as to yield a finite list of "objects" among which there is the description of a cell-decomposition compatible with A. To make this precise it turns out that the main difficulty is to give the correct definitions of what kind of objects our algorithm is going to manipulate. The problem here is that the notion of cell itself is not very constructive: from the Ch-description of a set we do not know how to recognize if the set is empty, or a singleton, a function, a continuous function, a cell of the form $(f, g)_C$, etc. Moreover, until we prove that $\tilde{\mathscr{S}}$ coincides with the family of sets which are first order definable from sets in $\mathscr{S}$, from the Ch-description of a cell of the form $(f, g)_C$ it is not clear how to obtain a Ch-description of f and g. To handle these problems we will define in the sequel the notion of good representation of a cell.

8. Effective non-deterministic cell decomposition

Fix a closed o-minimal effective W-structure $\mathscr{S}$ which is EDSF.

Definition 8.1. A *good representation of a cell* is given by the Ch-formulas for all the functions which are needed to define the cell. More precisely:

- A good representation of a singleton $P \in \tilde{\mathscr{S}}_1$ is the sequence of length one whose only element is a Ch-formula P for P. Such a sequence is denoted by (P).
- A good representation of a cell of the form $(a, b) \in \tilde{\mathscr{S}}_1$ is a pair whose first element is a Ch-formula P for the singleton $P = \{a\}$ and whose second element is a Ch-formula Q for the singleton $Q = \{b\}$. Such a pair is denoted by (P, Q).

- A good representation of a cell of the form $(f, g)_C$ is a triple whose first two elements are Ch-formulas f and g for f and g respectively and whose third element is a good representation C of C. Such a triple is denoted by $(\mathsf{f}, \mathsf{g})_{\mathsf{C}}$.
- A good representation of a cell of the form $(f)_C$ is a pair whose first element is a Ch-formula f for f and whose second element is a good representation C of C. Such a pair is denoted by $(\mathsf{f})_{\mathsf{C}}$.

We denote by Cell_n the set of all good representations of cells in $\tilde{\mathscr{S}}_n$.

So Cell_n is a hereditary finite sequence of Ch-formulas, namely a finite sequence whose elements are Ch-formulas or other hereditary finite sequences. A hereditary sequence of Ch-formulas can be considered as a syntactic expression, namely a finite sequence of symbols from some finite alphabet. So it makes sense to ask whether Cell_n is a recursive set of syntactic expressions. A priori there is no reason to believe so, since we are not able to determine if a Ch-formula is the Ch-formula for a function. This is the reason to consider a larger recursive set $\mathrm{PCell}_n \supseteq \mathrm{Cell}_n$ which is defined exactly as Cell_n but without the requirement that the various Ch-formulas involved are Ch-formulas for functions. The precise definition follows.

Definition 8.2. The set PCell_n is defined by induction on n as follows:

- If P is the Ch-formula for a set $P \in \tilde{\mathscr{S}}_1$ (not necessarily a singleton), then the sequence of length 1 whose only element is P belongs to PCell_1. Such a sequence is denoted by (P) and represents the set P.
- If P and Q are Ch-formulas for two sets P and Q in $\tilde{\mathscr{S}}_1$ (not necessarily singletons), then the pair whose first element is P and whose second element is Q belongs to PCell_1. Such a pair is denoted by (P, Q) and represents the set

$$(P, Q) := \{y \in \mathbb{R} \mid \exists u \in P \exists v \in Q. u < y < v\}.$$

- If f and g are Ch-formulas for two sets $f \in \tilde{\mathscr{S}}_{n+1}$ and $g \in \tilde{\mathscr{S}}_{n+1}$ (not necessarily functions), and $\mathsf{C} \in \mathrm{PCell}_n$, then the triple whose first element is f, whose second element is g, and whose third element is C, belongs to PCell_{n+1}. Such a triple is denoted $(\mathsf{f}, \mathsf{g})_{\mathsf{C}}$ and represents the set

$$(f, g)_C := \{(\vec{x}, y) \in C \times \mathbb{R} \mid \exists u, v \in \mathbb{R}. (\vec{x}, u) \in f \wedge (\vec{x}, v) \in g \wedge u < y < v\},$$

  where $C \in \tilde{\mathscr{S}}_n$ is the set represented by C.
- If f is a Ch-formula for a set $f \in \tilde{\mathscr{S}}_{n+1}$, and $\mathsf{C} \in \mathrm{PCell}_n$, then the pair whose first element is f and whose second element is C belongs to PCell_{n+1}. Such a triple is denoted $(\mathsf{f})_{\mathsf{C}}$ and represents the set

$$(f)_C := \{(\vec{x}, y) \in C \times \mathbb{R} \mid (\vec{x}, y) \in f\},$$

  where $C \in \tilde{\mathscr{S}}_n$ is the set represented by C.

Unlike Cell_n, the set PCell_n can be considered as a recursive set of syntactic expressions. The expressions in PCell_n will be called *representations of pseudo-cells*, and the sets they represent will be called *pseudo-cells*. Clearly $\mathrm{Cell}_n \subseteq \mathrm{PCell}_n$.

A pseudo-cell $D \in \tilde{\mathscr{S}}_n$ admits two kinds of representations. We can represent D by a Ch-formula, or we can represent it as a pseudo-cell, namely by a syntactic expression in PCell_n. The advantage of this second representation is that it allows us to compute a representation for the boundary of the cells. For instance from the representation of a cell of the form $(f,g)_C$ as a pseudo-cell, we can extract the representations of f and g (which are part of its boundary). The next lemma shows that there is an algorithm to pass from the pseudo-cell representation to the Ch-formula. We denote by Ch_n the set of Ch-formulas for sets in $\tilde{\mathscr{S}}_n$.

Lemma 8.3. *For each $n>0$ there is a recursive function $\psi_n : \mathrm{PCell}_n \to \mathrm{Ch}_n$ (uniform in n) such that if $\mathsf{D} \in \mathrm{PCell}_n$ represents the pseudo-cell $D \in \tilde{\mathscr{S}}_n$ (according to Definition 8.2), then $\psi_n(\mathsf{D})$ is a Ch-formula for D. So in particular from a good representation of a cell we can compute its Ch-formula.*

Here and below, "uniform in n" means that the function is recursive even as a function of n.

Proof. Given $\mathsf{D} \in \mathrm{PCell}_n$, let D be the set represented by D according to Definition 8.2. That definition is inductive and can be naturally turned into an algorithm to compute the Ch-formula for D. $\square$

Definition 8.4. We represent a cell decomposition D by the set of the good representations of its cells according to Definition 8.2. So the representation of D belongs to the set $\mathrm{Dec}_n := \wp_{<\omega}(\mathrm{Cell}_n)$ (the family of all finite subsets of Cell_n). Anyway it is convenient to work with $\mathrm{PDec}_n := \wp_{<\omega}(\mathrm{PCell}_n) \supseteq \mathrm{Dec}_n$, since PDec_n can be naturally identified with a recursive set of syntactic expressions.

Definition 8.5. A non-deterministic function f from A to B is a function $f : A \to \wp_{<\omega}(B)$, namely a function from A to the finite subsets of B. We write $f : A \Rightarrow B$ as a shorthand for $f : A \to \wp_{<\omega}(B)$. So if $f : A \Rightarrow B$ and $b \in f(a)$ $(a \in A, b \in B)$ we can consider b as one of the possible non-deterministic outputs of $f(a)$. We say that a non-deterministic function from A to B is recursive if it is recursive as a function from A to $\wp_{<\omega}(B)$ (this makes sense if A, B are recursive sets of strings of symbols).

We can now give our effective version of Wilkie's cell decomposition theorem.

Theorem 8.6. *For each $n>0$ there is a recursive non-deterministic function*

$$F_n:\ \mathrm{PCell}_n \times \wp_{<\omega}(\mathrm{Ch}_n) \Rightarrow \mathrm{PDec}_n$$

(uniform in n) such that if $\mathsf{D} \in \mathrm{PCell}_n$ is a good representation of a cell $D \in \tilde{\mathscr{S}}_n$ and $\mathsf{A}_1, \dots, \mathsf{A}_m \in \mathrm{Ch}_n$ are Ch-formulas for subsets $A_1, \dots, A_m$ of D which are closed in D, then there is a decomposition $\mathscr{D}$ of D compatible with $A_1, \dots, A_m$ which admits a representation $\{\mathsf{D}_1, \dots, \mathsf{D}_s\} \in F_n(\mathsf{D}, \{\mathsf{A}_1, \dots, \mathsf{A}_m\})$.

The theorem says that from the expressions representing $D, A_1, \ldots, A_m$ we can effectively find a finite set $Y = F_n(\mathsf{D}, \{\mathsf{A}_1, \ldots, \mathsf{A}_m\})$ of candidates, one of which is a representation of a cell decomposition of D compatible with each A_i.

Proof of Theorem 8.6: case $n = 1$. We describe in the sequel the algorithm to compute $F_1(\mathsf{D}, \{\mathsf{A}_1, \ldots, \mathsf{A}_m\})$ where D is a representation of a pseudo-cell D of $\mathbb{R}$ and A_i is a Ch-formula for $A_i \in \tilde{\mathscr{S}}_n$. Assume first that $m = 1$, namely $\{\mathsf{A}_1, \ldots, \mathsf{A}_m\} = \{\mathsf{A}\}$. If D has the form (P) where $\mathsf{P} \in \mathrm{Ch}_1$, then F_1 outputs the finite set Y whose only element is (P). This does what is required, since the unique possible decomposition of a singleton is the singleton itself. Consider now the case when D has the form (P, Q) and represents the pseudo-cell (P, Q) (with P, Q not necessarily singletons). The set $A^* \subseteq \mathbb{R}$ (see Definition 7.1) is finite since it has empty interior and belongs to $\tilde{\mathscr{S}}_1$. Given A we can compute a Ch-formula A^* for A^* and an upper bound $N = \Gamma(\mathsf{A}^*)$ on $\gamma(A^*)$. Choose non-deterministically a non-negative integer $\kappa \leqslant N$ (this means that we try all the possible values of κ and we proceed with the construction for each possible choice, putting in the final output all the outcomes of the various computation paths). For $1 \leqslant i \leqslant \kappa$ compute the Ch-formula $\mathsf{P}_{i,\kappa}$ for the set

$$P_{i,\kappa} = \{y \in \mathbb{R} \mid \exists y_1 \ldots y_\kappa (y_1 < \cdots < y_\kappa \wedge y_i \in A^* \cap (P, Q) \wedge y = y_i)\}.$$

(If $\kappa = 0$ we skip this step.) Define $P_{0,\kappa} = P$ and $P_{\kappa+1,\kappa} = Q$. The output corresponding to these non-deterministic choices is the element of PDec_1 consisting of the following set of expressions: $\mathsf{P}_{i,\kappa}$ (for $1 \leqslant i \leqslant \kappa$) and $(\mathsf{P}_{i,\kappa}, \mathsf{P}_{i+1,\kappa})$ (for $0 \leqslant i < i + 1 \leqslant \kappa + 1$). To verify that the algorithm does what is required to do, note that, if the input $\mathsf{D} = (\mathsf{P}, \mathsf{Q}) \in \mathrm{PCell}_n$ were a good representation of a cell (i.e. if P, Q are singletons) and if κ was non-deterministically chosen as the cardinality of $A^* \cap (P, Q)$, then all the $\mathsf{P}_{i,\kappa}$ represent singletons and the output represents a cell decomposition of (P, Q) compatible with A^*, hence with A (if A was closed in D: see Remark 7.5). The case $m > 1$ is similar. □

Definition 8.7. In the proof of Lemma 7.10 we have defined, for each cell C in $\mathbb{R}^n$, a sequence of integers $1 \leqslant i_1 < \cdots < i_d \leqslant n$ and the corresponding projection $\pi : \mathbb{R}^n \to \mathbb{R}^d$ (if $d = 0$ the sequence is empty and we project onto $\mathbb{R}^0$). The sequence can be computed by an algorithm which takes as input the representation C of C as a pseudo-cell, in the sense of Definition 8.2. If C represents a pseudo-cell C which is not an actual cell, the algorithm will still return a projection $\pi : \mathbb{R}^n \to \mathbb{R}^d$, but in this case $\pi_{|C}$ may not be an homeomorphism onto its image (for instance $(f)_E$ need not be homeomorphic to E if f is not a function). In any case we call d the *pseudo-dimension* of C and π the *associated projection* (it may depend on the representation C, not just on the set C, in case C is only a pseudo-cell).

Proof of Theorem 8.6: general case. We assume that $F_1, \ldots, F_n$ have already been defined with the desired properties and we describe the algorithm to compute $F_{n+1}(\mathsf{D}, \{\mathsf{A}_1, \ldots, \mathsf{A}_m\})$, dealing first with the case in which $\{\mathsf{A}_1, \ldots, \mathsf{A}_m\} = \{\mathsf{A}\}$. First we compute the pseudo-dimension d of D and the associated projection $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^d$. Let D be the pseudo-cell represented by D. We distinguish two cases.

Case $d<n+1$. Note that the image $D'=\pi D$ of the pseudo-cell D is a pseudo-cell, and moreover we can compute its representation $\mathsf{D}'\in \mathrm{PCell}_d$. Now we compute the Ch-formula A' for the image of A under π and we choose non-deterministically an element $\mathsf{E}'\in F_d(\mathsf{D}',\{\mathsf{A}'\})$. Then $\mathsf{E}'\in \mathrm{PDec}_d$ and we can find an element $\mathsf{E}\in \mathrm{PDec}_n$ which represents the set of those pseudo-cells which are the preimages under $\pi_{|D}:D\to D'$ of the pseudo-cells of E'. The output corresponding to this non-deterministic computation is $\mathsf{E}\in F_{n+1}(\mathsf{D},\{\mathsf{A}\})$. We must verify that this does the required job in case D was an actual cell and A was closed in D. Indeed in this case, by our inductive assumption on F_d, at least one of the non-deterministic choices of E' is a cell decomposition of πD compatible with $A'=\pi A$. Corresponding to this choice, E represents a cell decomposition of D compatible with A.

Case $d=n+1$. In this case D has necessarily the form $(\mathsf{f},\mathsf{g})_{\mathsf{C}}$ where $\mathsf{C}\in \mathrm{PCell}_n$ has pseudo-dimension n, and represents a pseudo-cell $D=(f,g)_C\in \tilde{\mathscr{S}}_{n+1}$ (f,g are not necessarily functions). We can consider the sets $\tilde{H},\tilde{H}_f$ and $\tilde{H}_g$ defined exactly as in Lemma 7.8, except that in the definition of H_f we replace "$y-f(x)=\varepsilon$" with "$(x,y-\varepsilon)\in f$" (which makes sense even if f is not a function) and similarly with g in the role of f. From the available data we can compute Ch-formulas for these sets. Now from A we can compute a Ch-formula A^* for the set A^* and an upper bound $N=\Gamma(\mathsf{A}^*)+1$ on $\gamma(A^*)$. Define

$$\mathscr{H}:=\{\overline{\{A^*\geqslant 1\}},\dots,\overline{\{A^*\geqslant N\}},\tilde{H},\tilde{H}_f,\tilde{H}_g\}$$

as in Lemma 7.8 and compute Ch-formulas for all the elements of $\mathscr{H}$. Let $\mathsf{H}\in \wp_{<\omega}(\mathrm{Ch}_n)$ be the set of these Ch-formulas. Choose non-deterministically an element $\mathsf{F}\in F_n(\mathsf{C},\mathsf{H})$ and let $\mathscr{F}\subseteq \tilde{\mathscr{S}}_n$ be the corresponding family of sets. So $\mathscr{F}$ is a candidate for a cell decomposition of C. From A and D we can compute an upper bound N on $\gamma(A^*\cap D)$. For each $\mathsf{C}'\in \mathsf{F}$ of pseudo-dimension n, choose non-deterministically a non-negative integer $\kappa(C')\leqslant N$ and define

$$f_i=\Bigg\{(\vec{x},y)\in C'\times\mathbb{R}\mid \exists y_1,\dots,y_\kappa \left(y_1<\cdots<y_\kappa\wedge\bigwedge_{j=1}^{\kappa}(\vec{x},y_j)\in A^*\cap D\wedge y=y_i\right)\Bigg\}.$$

Although f_i may not be a function, we can proceed as in the proof of Theorem 7.7 to define some pseudo-cells over C' "bounded" by the various f_i, together with $(f)_{C'}=\{(\vec{x},y)\in C'\times\mathbb{R}\mid(\vec{x},y)\in f\}$ and $g_{|C'}$ (defined similarly). Clearly we can compute representations for all these pseudo-cells. For $\mathsf{C}'\in\mathsf{F}$ of pseudo-dimension $<n$ we can proceed as in Theorem 7.7 and compute non-deterministically, with the help of the functions $F_1,\dots,F_{n-1}$, the appropriate pseudo-cell decompositions. Putting everything together we have obtained, non-deterministically, an element of PDec_n. If D was an actual cell, and A was a subset of D closed in D, at least one of these non-deterministic computations gives the correct result, namely a representation of a cell decomposition of D compatible with A.

We leave to the reader the definition of $F_{n+1}(\mathsf{D},\{\mathsf{A}_1,\dots,\mathsf{A}_m\})$ in the case in which $m>1$. □

9. Non-deterministic computation of the complement

Fix as above a closed o-minimal effective W-structure $\mathscr{S}$ which is EDSF and let $\tilde{\mathscr{S}}$ be its Charbonnel closure.

Theorem 9.1. *For each $n>0$ there is a recursive non-deterministic function $G_n:\mathrm{Ch}_n \Rightarrow \mathrm{Ch}_n$ (uniform in n) which, given a Ch-formula for a set $A\in\tilde{\mathscr{S}}_n$, returns a finite set of Ch-formulas, one of which defines the complement of A in $\mathbb{R}^n$.*

Proof. Since $\tilde{\mathscr{S}}$ is effectively semi-closed (Remark 5.5), given a Ch-formula for $A\in\tilde{\mathscr{S}}_n$, we can compute a Ch-formula which defines a closed set $B\in\tilde{\mathscr{S}}_{n+k}$ such that $A=\Pi_n^{n+k}[B]$. We can easily find a semi-algebraic homeomorphism $f:\mathbb{R}^{n+k}\to D$ in $\tilde{\mathscr{S}}$, where $D\in\tilde{\mathscr{S}}_{n+k}$ is a cell, such that f commutes with the projection Π_n^{n+k}. Compute the Ch-formula B' of $B'=f(B)$. Apply Theorem 8.6 to compute a finite set of candidates for a good representation of a cell decomposition of D compatible with B'. Choose non-deterministically a candidate $\mathscr{D}$. Take the preimages under f of the sets of $\mathscr{D}$ (and find their Ch-formulas). If $\mathscr{D}$ was the correct candidate, we obtain a partition of $\mathbb{R}^{n+k}$ (technically it is not a cell decomposition since cells must be bounded) such that B is a finite union of classes of the partition. Project all these preimages down to $\mathbb{R}^n$ using Π_n^{n+k}: we obtain a finite collection of subsets of $\mathbb{R}^n$ (together with their Ch-formulas) which is a candidate for a partition of $\mathbb{R}^n$ such that A is the union of some classes of the partition. Select non-deterministically a sub-collection (a candidate for the sets of the collection which do not meet A), consider their union, and return its Ch-formula as the output corresponding to these non-deterministic choices. At least one of the possible outputs is a Ch-formula of the complement of A. □

Definition 9.2. The first order language associated to $\mathscr{S}$ consists of an n-ary predicate symbol P_{A} for every Ch-formula A, which is interpreted as the set $A\subseteq\mathbb{R}^n$ associated to A (we identify a predicate with the set of elements which satisfy it).

Corollary 9.3. *There is a recursive function which, given a first order formula $\phi(x_1,\dots,x_n)$ in the language associated to $\mathscr{S}$, returns a finite set of Ch-formulas, one of which denotes the subset of $\mathbb{R}^n$ defined by ϕ.*

Proof. We can assume that the only logical connectives of ϕ are existential quantifiers, disjunctions, and negations. The first two can be simulated by the Ch-operations of projections and unions, while negations can be non-deterministically simulated with complements using Theorem 9.1. □

Finally we can prove our main result:

Theorem 9.4. *There is a recursive function which, given a first order formula $\phi(x_1,\dots,x_n)$ in the language associated to $\mathscr{S}$, returns an upper bound on $\gamma(A)$, where $A\subseteq\mathbb{R}^n$ is the set defined by ϕ.*

Proof. Given ϕ we compute a finite set of Ch-formulas, one of which denotes the set A defined by ϕ. Using the effective o-minimality of $\tilde{\mathscr{S}}$ (see Theorem 5.4), for each such Ch-formula, we compute an upper bound on γ of the corresponding set. Taking the greatest of these upper bounds, we also get an upper bound on $\gamma(A)$. □

10. Proof of Theorem 6.11

In this section we will go through the proof of Theorem 6.11. Most of the proofs of the following Lemmas can be found in [15], we simply give a presentation suitable to our purposes. Lemmas 10.3, 10.6, 10.8 below are of the form: "given certain sets in $\tilde{\mathscr{S}}$, we can find other sets in $\tilde{\mathscr{S}}$ with some required properties". The proofs show that, if $\mathscr{S}$ is EDSF, then the procedure is effective.

Remark 10.1. $|(1,x_1,\dots,x_n)|^2\leqslant 1/\varepsilon$ iff $\exists y.(1+x_1^2+\cdots+x_n^2+y^2)^{-1}=\varepsilon$. Recall that the function $(x_1,\dots,x_n,y)\mapsto(1+x_1^2+\cdots+x_n^2+y^2)^{-1}$ belongs to $M(\mathscr{S})$.

The first task is to find an $M(\mathscr{S})$-approximant for the zero-set of a smooth function.

Lemma 10.2. *Let $g:\mathbb{R}^n\to\mathbb{R}$ be a continuous function with $g\geqslant 0$ and let $S=\{(\vec{x},t)\in\mathbb{R}^n\times\mathbb{R}_+ \mid g(\vec{x})=t\}$. Then S approximates $\partial[g^{-1}(0)]$ from above on bounded sets, i.e. $\forall^s\varepsilon\forall^s t(g^{-1}(t)^\varepsilon\supseteq\partial[g^{-1}(0)]\cap B(0,1/\varepsilon))$.*

Proof. Fix $\varepsilon>0$ and suppose for a contradiction that there is a sequence of positive real numbers t_n converging to zero such that the inclusion fails for t_n, so that we can choose $\vec{x}_n\in\partial[g^{-1}(0)]\cap B(0,1/\varepsilon)$ with $\vec{x}_n\notin g^{-1}(t_n)^\varepsilon$. By compactness of $B(0,1/\varepsilon)$, choosing a subsequence we can assume that $\vec{x}_n\to\vec{x}\in\partial[g^{-1}(0)]\cap B(0,1/\varepsilon)$. Let O be the $\varepsilon/2$-neighborhood of $\vec{x}$. Since $\vec{x}\in\partial[g^{-1}(0)]$, g assumes some positive value γ on O, and since O is connected g assumes all values in the interval $[0,\gamma]$ on O. Now choose n so big that $\vec{x}_n\in O$ and $t_n<\gamma$. Then O intersects $g^{-1}(t_n)$ and therefore it is contained in $g^{-1}(t_n)^\varepsilon$. So $\vec{x}_n\in g^{-1}(t_n)^\varepsilon$, contrary to the choice of $\vec{x}_n$. □

Lemma 10.3 (See Wilkie [15, Lemma 3.8]). *If $f:\mathbb{R}^n\to\mathbb{R}$ is in $M(\mathscr{S})$, then its zero-set $V(f)$ has an $M(\mathscr{S})$-approximant $S\in\tilde{\mathscr{S}}_{n+2}$.*

Proof. Let $S=\{(x_1,\dots,x_n,\varepsilon_1,\varepsilon_2)\in\mathbb{R}^n\times\mathbb{R}^2_+ \mid |(1,x_1,\dots,x_n)|^2\leqslant 1/\varepsilon_1\wedge f^2(\vec{x})=\varepsilon_2\}$. By Remark 10.1, S is an $M(\mathscr{S})$-set. We prove that S approximates $V(f)$ from below, namely $\forall^s\varepsilon_0\forall^s\varepsilon_1\forall^s\varepsilon_2 S_{\varepsilon_1,\varepsilon_2}\subseteq V(f)^{\varepsilon_0}$. To see this note first that for fixed $\varepsilon_0,\varepsilon_1$, $S_{\varepsilon_0,\varepsilon_1}$ is contained in the compact set $K=\{(x_1,\dots,x_n)\in\mathbb{R}^n \mid |(1,x_1,\dots,x_n)|^2\leqslant 1/\varepsilon_1\}$. If ε_2 is smaller than the minimum of f on $K - V(f)^{\varepsilon_0}$ (or ε_2 is arbitrary if this set is empty), then $S_{\varepsilon_1,\varepsilon_2} \subseteq V(f)^{\varepsilon_0}$.

We prove that S approximates $\partial V(f)$ from above on bounded sets, namely $\forall^s \varepsilon_0 \forall^s \varepsilon_1 \forall^s \varepsilon_2 . \partial V(f) \cap B(0, 1/\varepsilon_0) \subseteq S^{\varepsilon_0}_{\varepsilon_1,\varepsilon_2}$. Fix ε_0 and choose ε_1 so that the set K considered above contains $B(0, 1/\varepsilon_0)$. By Lemma 10.2 for all sufficiently small ε_2, setting $g = f^2$, we have $g^{-1}(\varepsilon_2)^{\varepsilon_0} \supseteq \partial V(f) \cap B(0, 1/\varepsilon_0)$. Thus $\partial V(f) \cap B(0, 1/\varepsilon_0) \subseteq S^{\varepsilon_0}_{\varepsilon_1,\varepsilon_2}$. □

The smoothness assumptions are used in the following key lemma, which give us some information on the boundary of the projection of a set.

Lemma 10.4 (See Wilkie [15, Lemma 2.9]). *Let $F : \mathbb{R}^{m+k} \to \mathbb{R}^k$ be a C^1-function in $\tilde{\mathscr{S}}$, and consider the manifold $V = F^{-1}(a) \subseteq \mathbb{R}^{m+k}$, where $a \in \mathbb{R}^k$ is a regular value of F. Consider the projection $\pi : \mathbb{R}^{m+k} \to \mathbb{R}^m$ on the first m coordinates. Let O be an open ball in $\mathbb{R}^m$ intersecting $\partial \pi V$. Then for every sufficiently small $\varepsilon > 0$, O intersects $\pi V[\varepsilon]$, where $V[\varepsilon] \subseteq V$ is defined as the set of points $(x_1, \ldots, x_{m+k}) \in V$ such that one of the following conditions is satisfied for some $1 \leqslant i_1 < \ldots < i_k \leqslant m + k$:*
- $|(1, x_{m+1}, \ldots, x_{m+k})|^2 = 1/\varepsilon$;
- $\det \left(\frac{\partial^k F}{\partial(x_{i_1} \ldots x_{i_k})} \right)^2 = \varepsilon$.

Definition 10.5. We call $V[\varepsilon]$ the *ε-critical part* of V. This of course depends on the representation $V = F^{-1}(a)$.

Lemma 10.6 (See Wilkie [15, Lemma 3.10]). *If $A \subseteq \mathbb{R}^{n+1}$ has an $M(\mathscr{S})$-approximant $S \subseteq \mathbb{R}^{n+1} \times \mathbb{R}^k_+$, then there is a $M(\mathscr{S})$-approximant $S' \subseteq \mathbb{R}^n \times \mathbb{R}^{k+1}_+$ for $\Pi^{n+1}_n A \subseteq \mathbb{R}^n$.*

Proof. The sections $S_{\varepsilon_1,\ldots,\varepsilon_k} \subseteq \mathbb{R}^{n+1}$ of S have the form:

$$S_{\varepsilon_1,\ldots,\varepsilon_k} = \Pi^{n+1+k-1}_{n+1} \{F_1 = (\varepsilon_1, \ldots, \varepsilon_k)\} \cup \cdots \cup \Pi^{n+1+k-1}_{n+1} \{F_s = (\varepsilon_1, \ldots, \varepsilon_k)\},$$

where $F_i : \mathbb{R}^{n+1+k-1} \to \mathbb{R}^k$ is a C^∞ function in $M(\mathscr{S})$ and $\{F_i = (\varepsilon_1, \ldots, \varepsilon_k)\}$ is the pre-image of $(\varepsilon_1, \ldots, \varepsilon_k) \in \mathbb{R}^k$ under F_i. Define $S_{\varepsilon_1,\ldots,\varepsilon_k}[\varepsilon_k]$ as the set

$$\Pi^{n+1+k-1}_{n+1} (\{F_1 = (\varepsilon_1, \ldots, \varepsilon_k)\}[\varepsilon_{k+1}]) \cup \cdots \cup \Pi^{n+1+k-1}_{n} (\{F_s = (\varepsilon_1, \ldots, \varepsilon_k)\}[\varepsilon_{k+1}]),$$

where $\{F_i = (\varepsilon_1, \ldots, \varepsilon_k)\}[\varepsilon_{k+1}]$ is the ε_{k+1}-critical part of $\{F_i = (\varepsilon_1, \ldots, \varepsilon_k)\}$. Define S' as the set whose sections $S'_{\varepsilon_1,\ldots,\varepsilon_{k+1}} \subseteq \mathbb{R}^n$ are given by:

$$S'_{\varepsilon_1,\ldots,\varepsilon_{k+1}} = \Pi^{n+1}_n S_{\varepsilon_1,\ldots,\varepsilon_k}[\varepsilon_{k+1}].$$

It is easy to see that S' is an $M(\mathscr{S})$-set. Let us verify that S' approximates $\overline{\Pi^{n+1}_n A}$ from below. From the definition of S' it follows that $S'_{\varepsilon_1,\ldots,\varepsilon_k,\varepsilon_{k+1}} \subseteq \Pi^{n+1}_n S_{\varepsilon_1,\ldots,\varepsilon_k}$. On the other hand since S approximates $\bar{A}$ from below, given $\varepsilon_0 > 0$, we have $\forall^s \varepsilon_1 \ldots \forall^s \varepsilon_k$ $S_{\varepsilon_1,\ldots,\varepsilon_k} \subseteq (\bar{A})^{\varepsilon_0}$. It follows that $\forall \varepsilon_0 > 0 \forall^s \varepsilon_1 \ldots \forall^s \varepsilon_k$ we have $S'_{\varepsilon_1,\ldots,\varepsilon_{k+1}} \subseteq \Pi^{n+1}_n S_{\varepsilon_1,\ldots,\varepsilon_k} \subseteq (\overline{\Pi^{n+1}_n A})^{\varepsilon_0}$.

It remains to verify that S' approximates $\partial\overline{\Pi_n^{n+1} A}$ from above on bounded sets. Fix $\varepsilon_0 > 0$. Choose open balls $O_1, \dots, O_m \subseteq \mathbb{R}^n$ of radius ε_0 such that $O_1 \cup \cdots \cup O_m \supseteq \partial\overline{\Pi_n^{n+1}[A]} \cap B(0, 1/\varepsilon_0)$ and each O_i intersects $\partial\overline{\Pi_n^{n+1} A}$. Then O_i intersects $\Pi_n^{n+1} \partial\bar{A}$, and since S approximates $\partial\bar{A}$ from above on bounded sets, it easily follows that $\forall^s \varepsilon_1, \dots, \forall^s \varepsilon_k$ O_i intersects $\Pi_n^{n+1} S_{\varepsilon_1,\dots,\varepsilon_k}$. On the other hand since S approximates $\bar{A}$ from below and $O_i \not\subseteq \overline{\Pi_n^{n+1} A}$, it is easy to see that O_i is not included in $\Pi_n^{n+1} S_{\varepsilon_1,\dots,\varepsilon_k}$, and therefore must intersect its frontier. Thus by Lemma 10.4, $\forall^s \varepsilon_1, \dots, \forall^s \varepsilon_{k+1}$ O_i intersects $\Pi_n^{n+1} S_{\varepsilon_1,\dots,\varepsilon_k}[\varepsilon_{k+1}] = S'_{\varepsilon_1,\dots,\varepsilon_{k+1}}$, so O_i is contained in the ε_0 neighborhood of the latter set. Since $\partial\bar{A} \cap B(0, 1/\varepsilon_0)$ is covered by the balls O_i, it is contained in $S'^{\varepsilon_0}_{\varepsilon_1,\dots,\varepsilon_{k+1}}$. □

We give without proof the following easy lemma.

Lemma 10.7. *Let $A, B \subseteq \mathbb{R}^n$ be closed sets, and let $K \subseteq \mathbb{R}^n$ be compact. Then $\forall^s \varepsilon_1 \forall^s \varepsilon_2$ $A^{\varepsilon_2} \cap B^{\varepsilon_2} \cap K \subseteq (A \cap B)^{\varepsilon_1}$.*

Lemma 10.8 (See Wilkie [15, Lemma 3.12]). *Let $A \in \tilde{\mathcal{S}}$ have an $M(\mathcal{S})$-approximant $S \subseteq \mathbb{R}^n \times \mathbb{R}_+^k$ and suppose Y is an $n-1$ dimensional $\mathbb{Z}$-affine set; suppose further that $\bar{A} \cap Y = \partial\bar{A} \cap Y$. Then there is an $M(\mathcal{S})$-approximant $S' \subseteq \mathbb{R}^n \times \mathbb{R}_+^{k+2}$ for $\bar{A} \cap Y$.*

Proof. The requirement on the frontier is equivalent to asking that Y does not meet the interior of $\bar{A}$, hence we only need to worry about a subset of $\partial\bar{A}$. Suppose Y is the zero-set of a linear polynomial l with coefficients in $\mathbb{Z}$. The sections $S_{\varepsilon_1,\dots,\varepsilon_k} \subseteq \mathbb{R}^n$ of S have the form

$$S_{\varepsilon_1,\dots,\varepsilon_k} = \Pi_n^{n+k-1}\{F_1 = (\varepsilon_1, \dots, \varepsilon_k)\} \cup \cdots \cup \Pi_n^{n+k-1}\{F_s = (\varepsilon_1, \dots, \varepsilon_k)\},$$

where $F_i : \mathbb{R}^{n+k} \to \mathbb{R}^k$ is a C^∞ function in $M(\mathcal{S})$ and $\{F_i = (\varepsilon_1, \dots, \varepsilon_k)\}$ is the pre-image of $(\varepsilon_1, \dots, \varepsilon_k) \in \mathbb{R}^k$ under F_i.

Define $S' \subseteq \mathbb{R}^{n+k+2}$ as the set whose sections $S'_{\varepsilon_1,\dots,\varepsilon_{k+2}} \subseteq \mathbb{R}^n$ have the form

$$S'_{\varepsilon_1,\dots,\varepsilon_{k+2}} = S_{\varepsilon_3,\dots,\varepsilon_{k+2}} \cap Y(\varepsilon_2) \cap K_{\varepsilon_1},$$

where $K_{\varepsilon_1} = \{\vec{x} \in \mathbb{R}^n \mid |(1, x_1, \dots, x_n)|^2 \leqslant 1/\varepsilon_1\} = \{\vec{x} \mid \exists x_{n+k}(1 + \sum_{i=1}^n x_i^2 + x_{n+k}^2)^{-1} = \varepsilon_1\}$ and $Y(\varepsilon_2) = \{\vec{x} \mid \exists x_{n+k+1} l(x_1, \dots, x_n)^2 + x_{n+k+1}^2 = \varepsilon_2\}$, so that S' is an $M(\mathcal{S})$-set.

Let us prove that S' approximates $\bar{A} \cap Y$ from below. By Lemma 10.7

$$\forall^s \varepsilon_1 \forall^s \varepsilon_2 \bar{A}^{\varepsilon_2} \cap B(0, \varepsilon_1^{-1}) \cap Y(\varepsilon_2) \subseteq (\bar{A} \cap Y)^{\varepsilon_1}.$$

Since S approximates $\bar{A}$ from below we have

$$\forall^s \varepsilon_0 \forall^s \varepsilon_1 \dots \forall^s \varepsilon_{k+2} S'_{\varepsilon_1,\dots,\varepsilon_{k+2}} \subseteq S_{\varepsilon_3,\dots,\varepsilon_{k+2}} \subseteq \bar{A}^{\varepsilon_2}.$$

From the definition it follows that $S'_{\varepsilon_1,\dots,\varepsilon_{k+2}} \subseteq K_{\varepsilon_1} \cap Y(\varepsilon_2)$, hence combining all these equations we get

$$\forall^s \varepsilon_0 > 0 \forall^s \varepsilon_1 \dots \forall^s \varepsilon_{k+2} S'_{\varepsilon_1,\dots,\varepsilon_{k+2}} \subseteq (\bar{A} \cap Y)^{\varepsilon_0}.$$

It remains to prove that S' approximates $\bar{A} \cap Y$ from above on bounded sets. Since S approximates $\bar{A}$ from above on bounded sets we have

$$\forall^s \varepsilon_2 \ldots \forall^s \varepsilon_{k+2} \partial\bar{A} \cap B(0, \varepsilon_2^{-1}) \subseteq S^{\varepsilon_2}_{\varepsilon_3 \ldots \varepsilon_{k+2}}.$$

Since $\forall^s \varepsilon_0 \forall^s \varepsilon_1 B(0, \varepsilon_2^{-1}) \subseteq K_{\varepsilon_1}$ and by our hypothesis $\partial A \cap Y = \bar{A} \cap Y$, we obtain, using again Lemma 10.7:

$$\forall^s \varepsilon_0 \forall^s \varepsilon_1 \ldots \forall^s \varepsilon_{k+2} \bar{A} \cap Y \subseteq (S_{\varepsilon_3,\ldots,\varepsilon_{k+2}} \cap Y(\varepsilon_2) \cap K_{\varepsilon_1})^{\varepsilon_0}.$$

This concludes the proof of the Lemma. □

Proof of Theorem 6.11. We prove the first part, the second part follows by the analysis of the proof of the first. We proceed by induction on the rank (see Definition 4.5) of a Ch-description of A of A. Assume that $\mathscr{S}$ has DSF (EDSF for the second part).

If A is described as a set in $\mathscr{S}$, then Lemmas 10.3 and 10.6, combined with the DSF condition, provide the result; this is the only reason why we had to assume DSF.

If A is described as $A_1 \cup A_2$, then an $M(\mathscr{S})$-approximant for A is given by the union of the $M(\mathscr{S})$-approximants for A_1 and A_2, respectively. The reason why this arguments works is that topological closure commutes with union. The same is not true with intersection instead of union, and this is the reason why we will need a more complicated argument for the intersection.

If A is described as $\Pi_n^{n+h}[A_1]$, then an iterated use of Lemma 10.8 tells us what to do.

If A is described as $\bar{B}$, then it is trivial since by definition an $M(\mathscr{S})$-approximant for B is an $M(\mathscr{S})$-approximant for A.

So, the only case which requires more care is the case when A is described as $A_1 \cap L$, where L is $\mathbb{Z}$-affine. We need to analyze all subcases.

If A_1 is described as a set in $\mathscr{S}$, then A too can be described as a set in $\mathscr{S}$ and we already know how to deal with these sets. If A_1 is obtained by an application of Ch($\cup$), then by the distributivity laws for $\cup, \cap$, by inductive hypothesis and by an application of the argument above on how to approximate unions, we know how to approximate A. If $A_1 = \Pi_n^m[U]$, then we use the equation

$$\Pi_n^m[U] \cap L = \Pi_n^m[(U \times L) \cap (\Delta \times \mathbb{R}^{m-n})],$$

where $\Delta \subset \mathbb{R}^{2n}$ is the diagonal, and we conclude again by an application of Lemma 10.6 and by inductive hypothesis (notice that $U \times L$ has a description $\mathsf{U} \times \mathsf{L}$ of the same rank as U, by Lemma 4.7). If A_1 is obtained by an application of Ch($\cap_\ell$), then we conclude by the inductive hypothesis (as the intersection of two $\mathbb{Z}$-affine sets is $\mathbb{Z}$-affine). The only difficult case is when A_1 is described as $\bar{U}$. Let $L = Y_1 \cap \cdots \cap Y_m$, where Y_i is a $\mathbb{Z}$-affine set of codimension 1. Notice that

$$\bar{U} \cap Y_1 = \overline{U \cap Y_1} \cup (\overline{U \cap Y_1^+} \cap Y_1) \cup (\overline{U \cap Y_1^-} \cap Y_1),$$

where, Y_1 is the zero set of a linear polynomial l over $\mathbb{Z}$, $Y_1^+ = \{\vec{x} \in \mathbb{R}^n \mid l(\vec{x}) > 0\}$, and Y_1^- is defined similarly by $l < 0$.

The descriptions of $U \cap Y_1^{\pm}$ have lower rank than $\rho(\bar{U})$, hence the inductive hypothesis can be applied to them. Now, Y_1 does not meet the interior of $\overline{U \cap Y_1^{\pm}}$ (since it does not meet the interior of $\overline{Y_1^{\pm}}$), hence to approximate the sets $\overline{U \cap Y_1^{\pm}} \cap Y_1$ we can use Lemma 10.8; while by inductive hypothesis we can get an approximant for the set $U \cap Y_1$. Now notice that $\bar{U} \cap Y_1$ has empty interior, so that we can make use of Lemma 10.8 for $(\bar{U} \cap Y_1) \cap Y_2$, and continue this way until we complete the proof of the theorem. □

References

[1] J.-Y. Charbonnel, Sur certains sous-ensembles de l'espace euclidien, Ann. Inst. Fourier (Grenoble) 41 (3) (1991) 679–717.

[2] L. van den Dries, o-minimal structures, in: Logic: from foundations to applications (Staffordshire, 1993), Oxford Sci. Publ., Oxford University Press, New York, 1996, pp. 137–185.

[3] L. van den Dries, Tame topology and o-minimal structures, in: London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998.

[4] A. Gabrielov, Frontier and closure of a semi-Pfaffian set, Discrete Comput. Geom. 19 (1998) 605–617.

[5] A. Gabrielov, Relative closure and the complexity of Pfaffian elimination, in: B. Aronov (Ed.), Discrete and Computational Geometry: The Goodman-Pollack Festschrift, Springer, 2003.

[6] A. Gabrielov, N. Vorobjov, Complexity of computations with Pfaffian and Noetherian functions, preprint: February 18, 2003.

[7] A.G. Khovanskii, A class of systems of transcendental equations, Dokl. Akad. Nauk SSSR 255 (1980) 762–765.

[8] A. Macintyre, A. Wilkie, On the decidability of the real exponential field, in: A.K. Peters (Eds.), Kreiseliana, Wellesley, MA, 1996, pp. 441–467.

[9] S. Maxwell, A general model completeness result for expansions of the real ordered field, Ann. Pure Appl. Logic 95 (1998) 185–227.

[10] M. Otero, Y. Peterzil, A. Pillay, Groups and rings definable in o-minimal expansions of real closed fields, Bull. London Math. Soc. 28 (1996) 7–24.

[11] J.C. Oxtoby, Measure and Category, Springer, New York–Heidelberg–Berlin, 1980.

[12] A. Tarski, A Decision Method for Elementary Algebra and Geometry, 2nd Edition, University of California Press, Berkeley, Los Angeles, CA, 1951.

[13] 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 (1996) 1051–1094.

[14] A.J. Wilkie, Model theory of analytic and smooth functions, in: Models and computability (Leeds, 1997), London Mathematical Society Lecture Note Series, vol. 259, Cambridge University Press, Cambridge, 1999, pp. 407–419.

[15] A.J. Wilkie, A theorem of the complement and some new o-minimal structures, Selecta Math. (N.S.) 5 (1999) 397–421.