Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon
Unité de recherche associée au CNRS n° 1398

Definability of Geometric Properties in Algebraically Closed Fields

Olivier Chapuis and Pascal Koiran

July 1998

Research Report N° 98-32

École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.00
Télécopieur : +33(0)4.72.72.80.80
Adresse électronique : `lip@ens-lyon.fr`

Definability of Geometric Properties in Algebraically Closed Fields

Olivier Chapuis and Pascal Koiran

July 1998

Abstract

We prove that there exists no sentence F of the language of rings with an extra binary predicat I_2 satisfying the following property: for every definable set $X \subseteq \mathbb{C}^2$, X is connected if and only if $(\mathbb{C}, X) \models F$ where I_2 is interpreted by X. We conjecture that the same result holds for the closed subsets of $\mathbb{C}^2$. We prove some results motivated by this conjecture.

Keywords: definability, constraint databases, model theory, algebraically closed fields.

Résumé

On montre qu'il n'existe pas d'énoncé F dans le langage des anneaux muni d'un prédicat binaire supplémentaire I_2 satisfaisant la propriété suivante: pour tout ensemble définissable $X \subseteq \mathbb{C}^2$, X est connexe si et seulement si $(\mathbb{C}, X) \models F$ (I_2 est interprété par X dans l'énoncé F). Nous conjecturons que le même résultat est vrai pour les fermées de $\mathbb{C}^2$. Nous démontrons également quelques résultats motivés par cette conjecture.

Mots-clés: définissabilité, bases de données contraintes, théorie des modèles, corps algébriquement clos.

Definability of Geometric Properties in Algebraically Closed Fields

Olivier Chapuis
Institut Girard Desargues – CNRS
Bâtiment des Mathématiques, Université Lyon I
43 Bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex
chapuis@desargues.univ-lyon1.fr
http://www.desargues.univ-lyon1.fr

Pascal Koiran
Laboratoire de l'Informatique du Parallélisme – CNRS
Ecole Normale Supérieure de Lyon
46 allée d'Italie, F-69364 Lyon Cedex 07
Pascal.Koiran@ens-lyon.fr
http://www.ens-lyon.fr/~koiran

17th July 1998

Abstract

We prove that there exists no sentence F of the language of rings with an extra binary predicat I_2 satisfying the following property: for every definable set $X \subseteq \mathbb{C}^2$, X is connected if and only if $(\mathbb{C}, X) \models F$ where I_2 is interpreted by X. We conjecture that the same result holds for closed subset of $\mathbb{C}^2$. We prove some results motivated by this conjecture.

Keywords: definability, constraint databases, model theory, algebraically closed fields.

1 Introduction

There is a recent and fairly large body of work on the definability of "geometric" properties in first-order logic, originally motivated by database research (geographic databases in particular). We refer the reader to [4] and [12] for an introduction to this subject and a guide to the literature. The structures which have been most studied from this point of view are the integers and the reals with various sets of operations. In this paper, we begin a study of definability over the complex numbers and algebraically closed fields. Since these questions are mathematically interesting in their own right, we have chosen to use a language which may be more appealing to readers who do not specialize in geographic databases.

For us, a property is just a family of definable sets of K^n, where K is a field and n is some fixed constant. Here definable means definable by a first-order formula (with parameters) of the language of rings :

$$\mathcal{L}_{rings} = \{=, +, -, \times, 0, 1\}.$$

By elimination of quantifiers, these sets are the constructible sets of algebraic geometry if K is algebraically closed; they are the semi-algebraic sets if K is real-closed. Here are two examples of properties:

1. The family of definable sets of $\mathbb{C}^4$ which are of dimension 2.

2. The family of connected definable subsets of $\mathbb{C}^2$.

It turns out that Property 1 is definable but Property 2 isn't. Formally, in order to define properties in K^n we work in the language $\mathcal{L}^K_{rings}$ enriched with a n-ary predicate I_n (if $\mathcal{L}$ is a language and if M is an $\mathcal{L}$-structure, $\mathcal{L}^M$ is $\mathcal{L}$ with constants naming the elements of M). The property defined by a first-order sentence F in this language is the family of definable sets $X \subseteq K^n$ such that F is true when I_n is interpreted by membership to X (in this case we write: $(K, X) \models F$).

We shall give examples of definable properties in section 2. The main result of this paper is that Connectivity in $\mathbb{C}^2$ is not definable. That is:

Theorem 1.1 *There exists no sentence F of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_2\}$ satisfying the following property: for every definable set $X \subseteq \mathbb{C}^2$, X is connected if and only if $(\mathbb{C}, X) \models F$.*

It is easy to see that the above theorem implies that connectivity is not definable for the definable subsets of $\mathbb{C}^n$ whenever $n \geq 2$. In fact we shall prove that there exists no sentence F of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_2\}$ satisfying the following property: for every definable set $X \subseteq \mathbb{C}^2$ which can be written as a finite boolean combination of points and lines, X is connected if and only if $(\mathbb{C}, X) \models F$ (the same result is known to be true in $\mathbb{R}^2$ for finite unions of line segments [12]). However, it is shown in the next section that there exists a sentence $F(I_2)$ such that if X is a finite union of points and lines in $\mathbb{C}^2$, then X is connected iff $(\mathbb{C}, X) \models F(I_2)$. A variation on the proof of Theorem 1.1 shows that no such sentence exists in dimension higher than 2.

We present in section 3 a basic geometric construction (reduction from Parity to Connectivity) which is used in the three proofs of Theorem 1.1 that we shall give. A proof of Theorem 1.1 follows by reduction to the real case and by the fact that Parity is not definable over the reals [4]. We give a second proof in section 4. As a byproduct, we obtain a strengthening of the result that Parity is not definable over the reals: Parity remains undefinable even if we restrict our attention only to those subsets X of $\mathbb{R}$ which are made only of integers, with distance at most 2 between two consecutive elements of X. The proof of this result uses the equivalence between active and natural semantics over the reals [5] and the fact that Parity is not in AC^0. See [1, 11, 22, 14] for original proofs of this important theorem of complexity theory, and [18] (chapters 6.12 and 6.13) for an elementary proof and further references. A self-contained model-theoretic proof of Theorem 1.1 is given in section 5. This last proof works also for algebraically closed fields of positive characteristic.

At this stage, it is perhaps useful to make a few remarks of a topological nature. We can view $\mathbb{C}^n$ with the strong (euclidean) topology or the Zariski topology. Note first, that for definable sets to be closed in the strong topology is the same thing as to be closed in the Zariski topology. Clearly, a definable set which is connected in the strong topology has to be connected in the Zariski topology. Moreover, one can use the fact that an irreducible Zariski-closed set is connected for the strong topology (see [21, Chapter VII]) to show that the converse is also true. These two notions of connectivity are therefore equivalent.

We propose the following problems.

Conjecture 1.2 *Let K be an algebraically closed field.*
(a) The family of closed definable subsets of K^2 is not definable.
(b) The family of closed irreducible definable subsets of K^2 is not definable.

Note that in a real-closed field R, the family of closed (for the order topology) definable sets of R^n is obviously definable.

The above conjecture makes precise an intuition that some logicians have regarding the Cherlin-Zil'ber conjecture. We recall that the Cherlin-Zil'ber conjecture states that a simple group of finite Morley rank is an algebraic group over an algebraically closed field (see [7]). This conjecture essentially says that if G is a simple group of finite Morley rank, then we can recover the Zariski topology of G from its definable subsets. This is surely not an easy task and Conjecture 1.2 claims that this is not possible using a first-order sentence in the case of an algebraically closed field.

The last two sections of this paper were motivated by Conjecture 1.2. In Section 6, we show that when a property of an algebraically closed field K is definable with a formula with parameters in K, if there is any hope of eliminating these parameters then this can be done (we leave it as an open problem whether the same result holds in real-closed fields). It follows that Conjecture 1.2 depends only on the characteristic and not on a specific algebraically closed field. In the last section, we show that the method used to prove that Connectivity is not a definable property cannot solve Conjecture 1.2. Namely, we show that for certain families of definable sets of K^n closedness is definable. These families are, roughly speaking, the families of definable sets which can be defined by a formula with "parameters in a class of finite structures". The main tools in this section is a result of quantifier elimination where the degree of the polynomials in the quantifier-free formula depends only on the number of quantifiers and the degree of the polynomials in the quantified formula. Surprisingly, such a result does not seem to appear in the literature.

Finally, we would like to point out that many undefinability results in first-order logic hinge on the fact that the property under consideration (e.g., parity or connectivity) is not "local." However, closedness is a local property. This explains perhaps why it seems difficult to tackle Conjecture 1.2 with standard techniques.

2 Examples of Definable Properties

In this section we give some examples of definable properties. We fix an algebraically closed field K.

Proposition 2.1 *The family* $\mathrm{DIM}_{n,d}$ *of d-dimensional definable subsets of K^n is definable without parameters.*

Proof. We consider first the case $d = n$. A definable subset of $X \subseteq K^n$ has dimension n (i.e., is dense in K^n) if and only if K^n can be covered by $n+1$ translates of X (see [16], Theorem 4.8 for a proof; this is probably well known from model theory). Hence $\mathrm{DIM}_{n,n}$ is defined by the following formula:

$$\exists t_1, \ldots, t_{n+1} \in K^n \ \forall v \in K^n \ \bigvee_{i=1}^{n+1} I_n(v - t_i). \tag{1}$$

For $d < n$, we use the fact that X has dimension at least d if and only if it has a dense projection on some d-dimensional coordinate subspace. Hence $\dim X \geq d$ can be expressed by a disjunction of $\binom{n}{d}$ formulas of the form (1) (projecting X amounts to adding an existential quantifier in front of I_n). A formula for $\dim X = d$ follows immediately. $\square$

This implies that connectivity is definable in K since a definable set $X \subseteq K$ is connected if it is one-dimensional or has a single element. This also implies that closedness is definable in K since a definable set $X \subseteq K$ is closed if it is zero-dimensional or equal to K.

Proposition 2.2 *The following properties of definable subsets of K^2 are definable.*

1. *X is a finite union of points.*
2. *X is a finite union of lines.*
3. *X is a finite union of points and lines.*
4. *X is connected, and is a finite union of points and lines.*

Proof. We will only give informal descriptions of the required formulas. Supplying the details should be straightforward.

1. Follows immediately from Theorem 2.1.
2. X is a union of lines iff for every $x \in X$ there exists a line $\Delta \subseteq X$ which goes through x. The "finite union" condition can be enforced by requiring X to be of dimension ≤ 1.
3. This is equivalent to a conjunction of two conditions:
   (a) X has dimension ≤ 1.
   (b) If we remove from X all points x such that there exists a line $\Delta \subseteq X$ going through x, the remaining set is 0-dimensional.
   A formula for condition (b) can thus be obtained from $\mathrm{DIM}_{2,0}$ by replacing each occurrence of $I_2(x)$ in this formula by $I_2(x) \wedge \neg\mathrm{line}(x)$, where $\mathrm{line}(x)$ expresses that there exists a line $\Delta \subseteq X$ going through x.
4. A finite union of points and lines is connected iff one of the following conditions holds:
   (a) X is reduced to a single point.
   (b) X is reduced to a single line.
   (c) X is a finite union of at least 2 lines, and there exists no line D such that all lines $\Delta \subseteq X$ are parallel to D.

□

These observations lead to many interesting questions. For instance, one can ask whether connectivity is definable for closed definable subsets of K^2. Or, restricting our attention to a special class of closed sets, we can make the following definition: let us say that a closed subset of K^n has pseudo-degree $\leq d$ if it is a finite union of closed sets defined by systems of polynomial equations of degree at most d. Then, given $d \geq 2$, one can ask whether connectivity is definable for closed subsets of K^2 of pseudo-degree $\leq d$ (for $n \geq 3$ and $d \geq 1$, we can answer this question by the negative using a variation of the construction in section 3).

We shall need the following bound. If X is a closed set of K^n defined by polynomial equations of degree $\leq d$, then the irreducible components of X can be defined by polynomial equations of degree $\leq D$ where D depends only on n and d (see [20, n^o 65] and [9, 2.10.v] for more general results which imply this bound). The first consequence of this bound is that if X is a closed set of K^n of pseudo-degree $\leq d$, then X is a finite union of irreductible closed sets defined by polynomial equations of degree $\leq D$. The second one is that we can say in a first-order formula "there exists an irreductible closed set W of K^n defined by polynomial equations of degree $\leq d$". Indeed, note first that if a closed set is defined by polynomial equations of degree $\leq d$, then it can be defined by $\leq c$ polynomial equations of degree $\leq d$ where $c = \binom{d+n}{d}$ (since c is the dimension of the K-subspace of polynomials of degree $\leq d$ in $K[x_1, \ldots x_n]$). Then we considere the following sentence : "There exist c polynomials f_i of degree $\leq d$ such that for any familly of $\binom{D+n}{D}$ polynomials g_i of degree $\leq D$, if the closed set W defined by the f_i contains the closed set Y defined by the g_i and if these sets have the same dimension, then $Y = W$". This can be expressed with a first-order formula by quantifing the coeficients of polynomials and using Proposition 2.1.

The following result will be useful in Section 6 and 7.

Proposition 2.3 *Let d and n be integers ≥ 1. There exists a formula F of $\mathcal{L}_{rings} \cup \{I_n\}$ such that for any definable set $X \subseteq K^n$ the two following properties hold:*

(i) If X is closed of pseudo-degree $\leq d$ then $(K, X) \models F$.

(ii) If $(K, X) \models F$ then X is closed.

Proof. We consider the following "algorithm".
Step 0: check whether $\dim X \leq n-1$. If not, then accept X if $X = K^n$, reject X if $X \neq K^n$.
Step i ($1 \leq i \leq n-1$): Let X_i be the set obtained from X by removing every point $x \in X$ such that there exists a closed irreducible set $W \subseteq X$ of dimension at least $n-i$ defined by polynomials of degree $\leq D$ such that $x \in W$ (here D depends only on n and d as explained before the Proposition).

If the dimension of X_i is $\geq n-i$ reject. If not, then if $i = n-1$ accept X and if $i < n-1$ goto step $i+1$.

We claim that this algorithm accepts all definable subsets of K^n which are closed of pseudo-degree $\leq d$, and that conversely any definable subsets of K^n accepted by the algorithm must be closed. This will imply the Proposition since we can then use Proposition 2.1, the remark before the statement of the proposition and the algorithm to construct a formula F satisfying (i) and (ii).

Let us now prove the claim. Assume first that X is closed of pseudo-degree $\leq d$, and write $X = \cup_{j \in J} B_j$ where B_j is a closed irreducible set defined by polynomial equations of degree $\leq D$. Let $x \in X$ be a point lying on a component B_j with $\dim B_j \geq n-i$. Then $x \notin X_i$ since we can take $W = B_j$ at step i of the algorithm. This shows that $\dim X_i \leq n-i-1$, and therefore X is not rejected at step i and is eventually accepted after step $n-1$.

Assume now that X is a definable (i.e., constructible) set accepted by the algorithm. In order to prove that X is closed, we will show by induction on i that $Y_i = \{x \in X;\ \dim_x X \geq n-i\}$ is closed for any $0 \leq i \leq n-1$. This is clear for $i=0$ since either $Y_0 = \emptyset$ or $Y_0 = K^n$ (by step 0 of the algorithm). Induction step: assume that the result is true for $i-1$. There is nothing to prove if $Y_i = Y_{i-1}$. Note that since X is constructible, Y_i is also constructible. Let us now examine the case $\Delta_i = Y_i \setminus Y_{i-1} \neq \emptyset$ (note that $\dim \Delta_i = n-i$ since $\Delta_i = \{x \in X;\ \dim_x X = n-i\}$). Let $\mathcal{F}$ be the family of closed irreducible sets $W \subseteq X$ of dimension $\geq n-i$ defined polynomials of degree $\leq D$ such that $W \cap \Delta_i \neq \emptyset$. In fact $W \in \mathcal{F}$ must be of dimension exactly $n-i$ since $\dim_x X = n-i$ for any $x \in \Delta_i$ (and in particular for $x \in W \cap \Delta_i$). We shall see that $\mathcal{F}$ is finite and $Y_i = Y_{i-1} \cup \bigcup_{W \in \mathcal{F}} W$ (implying that Y_i is closed as claimed). The inclusion from right to left is clear since $W \subseteq Y_i$ for any $W \in \mathcal{F}$ (this follows from: $\dim W = n-i$, W irreducible, and $W \subseteq X$). This inclusion implies that $\bigcup_{W \in \mathcal{F}} (W \setminus Y_{i-1}) \subseteq \Delta_i$. Each term $W \setminus Y_{i-1}$ in the left-hand side has dimension $n-i$. This follows from $\dim W = n-i$, W irreducible and $W \not\subseteq Y_{i-1}$ (which follows in turn from $W \cap \Delta_i \neq \emptyset$). Since Δ_i has dimension $n-i$ too and is constructible, we conclude that $\mathcal{F}$ must be finite. In order to establish the inclusion $Y_i \subseteq \bigcup_{W \in \mathcal{F}} W$, we need to show that $\Delta_i \subseteq \bigcup_{W \in \mathcal{F}} W$. Assume to the contrary that there exists $x \in \Delta_i$ such that $x \notin \bigcup_{W \in \mathcal{F}} W$. Since $\mathcal{F}$ is finite, there exists a Zariski open set O (containing x) such that $(\Delta_i \cap O) \cap \bigcup_{W \in \mathcal{F}} W = \emptyset$. By definition of $\mathcal{F}$, this implies that $\Delta_i \cap O \subseteq X_i$. This is a contradiction since $\dim(\Delta_i \cap O) = n-i$ and $\dim X_i < n-i$ by hypothesis. $\square$

Lemma 2.4 *Let $\phi(x_1, \ldots, x_n)$ be a quantifier-free formula which is a boolean combination of formulas of the form $f(\bar{x}) = 0$ where f is a polynomial of $K[\bar{x}]$ of degree $\leq d$. If the set defined by $\phi(\bar{x})$ is closed, then its pseudo-degree is $\leq D$ where D depends only on n and d.*

Proof. Assume that ϕ is of the form

$$\bigvee_{j} (\bigwedge_{i}^{s_j} f_{i,j}(\bar{x}) = 0 \wedge g_{i,j}(\bar{x}) \neq 0).$$

Denote by X the set defined by $\phi(\bar{x})$. Let F_j be the closed set defined by the polynomials $f_{i,j}(\bar{x})$, $i = 1, \ldots s_j$. Then, X is a union of sets of the form $V \cap O$ where V is an irreducible component of one of the F_j and where O is a nonempty open subset of K^n. Note that the closure of such a set is equal to V whenever $V \cap O \neq \emptyset$. Assume that X is closed. Then, X is a union of the closure of certain set $V \cap O$ where V is an irreducible component of one of the F_j and where O is a nonempty open subset of K^n. Thus, X is a union of certain irreducible components of the F_j. Since the F_j are defined by polynomial equations of degree $\leq d$, X is of pseudo-degree $\leq D$ where D depends only on n and d. $\Box$

3 Parity from Connectivity

A family $(G_n)_{n \geq 1}$ of undirected graphs on the set of vertices $\{1, \ldots, n\}$ will play an important role. There is an edge between vertices i and j in G_n if $|i - j| = 2$ or $i = 1$ and $j = n$. One checks easily that G_n is connected only when n is even. Given a finite set $X = \{a_1, \ldots, a_n\} \subseteq K$, we now construct a “geometric realization” $S_X \subseteq K^2$ of G_n as a boolean combination of points and lines in K^2. This set is a geometric realization of G_n in the following sense:

1. The points $A_i = (a_i, 0)$ belong to S_X (A_i represents vertex i of G_n).

2. There exists a path in S_X between A_i and A_j which does not go through any other A_k if and only if $(i, j) \in G_n$.

Let V_i be the vertical line of equation $x = a_i$, and D_j the line $x + y = a_j$. Let us remove from V_i and D_j the intersection point $V_i \cap D_j$ whenever both $i \neq j$ and $(i, j) \notin G_n$. This yields one-dimensional sets $V'_1, D'_1, \ldots, V'_n, D'_n$. We take S_X to be the union of these $2n$ sets.

Proposition 3.1 *S_X is connected if and only $|X|$ is even.*

Proof. For $|X|$ even, consider the arrangement of these $2n$ (connected) sets in the following order:

$$D'_2 V'_2 D'_4 V'_4 \cdots V'_{2i} D'_{2i+2} V'_{2i+2} \cdots D'_n V'_n D'_1 V'_1 D'_3 V'_3 \cdots V'_{2i-1} D'_{2i+1} V'_{2i+1} \cdots D'_{n-1} V'_{n-1}.$$

By construction, two consecutive sets in this sequence have a nonempty intersection. Their union is thus connected. This can be proved as follows.

Let U and V be two Zariski-closed subsets of K^2 such that $S_X \subseteq U \cup V$ and $S_X \cap U \cap V = \emptyset$. We need to show that either $S_X \subseteq U$ or $S_X \subseteq V$. Since $D'_2 \subseteq S_X \subseteq U \cup V$, $D'_2 \cap U \neq \emptyset$ or $D'_2 \cap V \neq \emptyset$. Assume for instance that $D'_2 \cap U \neq \emptyset$. Since D'_2 is connected, this implies that $D'_2 \subseteq U$. Hence $V'_2 \cap U \neq \emptyset$ since $D'_2 \cap V'_2 \neq \emptyset$. V'_2 being connected too, this implies that $V'_2 \subseteq U$ as well. This process can be continued until we have shown that the $2n$ sets in the above sequence are all included in U.

For n odd, let C_1 be the union of the $n+1$ lines

$$D_1, V_1, D_3, V_3, \ldots, V_{2i-1}, D_{2i+1}, V_{2i+1}, \ldots, D_n, V_n$$

and C_2 the union of the $n-1$ lines

$$D_2, V_2, D_4, V_4, \ldots, V_{2i}, D_{2i+2}, V_{2i+2}, \ldots, D_{n-1}, V_{n-1}.$$

S_X is included in the union of these two nonempty closed sets, and the intersection $S_X \cap C_1 \cap C_2$ is empty. This implies that S_X is not connected. □

Assume now that we have a total order on X, and $a_1 < a_2 \cdots < a_n$. We can then construct first-order formulas in the language $\mathcal{L}_{rings} \cup \{I_1, <\}$ (where I_1 is a predicate for X and $<$ is a predicate for an order on X), $\min(x)$, $\max(x)$ and $\mathrm{succ}(x, y)$ expressing respectively that $x = a_1$, $x = a_n$, $x = a_i$ and $y = a_{i+1}$ for some $i \in \{1, \ldots, n-1\}$:

$$\min(x) \equiv I_1(x) \wedge \forall y \in I_1 \ (x = y \vee x < y),$$

$$\max(x) \equiv I_1(x) \wedge \forall y \in I_1 \ (x = y \vee y < x),$$

$$\mathrm{succ}(x, y) \equiv I_1(x) \wedge I_1(y) \wedge \forall z \in I_1 \ \neg(x < z < y).$$

Membership of a point $(x, y) \in K^2$ to the union of the V'_i's is expressed by a formula $\psi_V(x, y)$ of the form:

$$\exists z \in I_1 \ [x = z \wedge \forall t \in I_1 (z = t \vee \mathrm{edge}(z, t) \vee \mathrm{edge}(t, z) \vee x + y \neq t)]$$

where $\mathrm{edge}(z, t)$ stands for:

$$(\min(z) \wedge (\max(t)) \vee \exists u \in I_1 \ (\mathrm{succ}(z, u) \wedge \mathrm{succ}(u, t)).$$

The construction of a similar formula $\psi_D(x, y)$ for membership to the union of the D'_i's is left to the reader. Membership to S_X is then defined by $\psi = \psi_V \vee \psi_D$.

One problem with the above construction is that if we work with an algebraically closed field in the language $\mathcal{L}_{rings} \cup \{I_1\}$, there is no way to construct a total order on an arbitrary finite X (however, a related construction over the reals can be used to show that Connectivity is not definable in $\mathbb{R}^3$).

In the case where $K = \mathbb{C}$ it is possible to circumvent this difficulty by performing a reduction to the real case.

Proposition 3.2 *There exists a formula* $\phi(x_1, x_2, x_3, x_4)$ *of* $\mathcal{L}_{rings} \cup \{I_1\}$ *which satisfies the following property. For any finite set* $X \subseteq \mathbb{R}$, *let* ϕ_X *be the subset of* $\mathbb{C}^2$ *(identified to* $\mathbb{R}^4$*) defined by* ϕ *when* I_1 *is interpreted by membership to* X. *Then* ϕ_X *is a definable subset of* $\mathbb{C}^2$, *and* ϕ_X *is connected if and only if* $|X|$ *is even.*

Proof. We use the formula ψ constructed above, with the order on X induced by the real order which is definable by : $x < y$ iff $\exists z\ y - x = z^2 \wedge z \neq 0$. If X is fixed, then one easily see that ϕ_X defines a definable subset of $\mathbb{C}^2$ using a formula with parameters in X. □

Lemma 3.3 *If Connectivity in* $\mathbb{C}^2$ *is definable, then Parity over* $\mathbb{R}$ *is definable, i.e., there exists a formula* G *of over* $\mathcal{L}^{\mathbb{R}}_{rings} \cup \{I_1\}$ *which satisfies the following property: for any finite set* $X \subseteq \mathbb{R}$, $(\mathbb{R}, X) \models G$ *if and only if* $|X|$ *is even.*

Proof. Let F be the formula of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_2\}$ which defines Connectivity. By separating real and imaginary parts of variables and parameters in F, we obtain a formula F' of $\mathcal{L}^{\mathbb{R}}_{rings} \cup \{I_4\}$ which satisfies the following property: whenever I_4 is interpreted by membership to a definable subset $S \subseteq \mathbb{C}^2$, $(\mathbb{R}, S) \models F'$ if and only if S is connected. Formula G is obtained from F' by replacing each occurrence of I_4 in this formula by formula ϕ from Proposition 3.2. □

This proves Theorem 1.1 since Parity is not definable over the reals [4]. As announced in the introduction, we give a strengthening of this result in the next section. This yields an alternative proof of Theorem 1.1. In section 5 we will give a self-contained proof of this theorem which does not use any reduction to the real case. Note that we perform such a reduction in section 4 only because to this date, the equivalence between natural and active domain semantics has been established only for the reals (these notions are defined in the next section).

4 From Complexity to Logic

This proof of Theorem 1.1 is by a series of reductions, beginning with a reduction from a restricted version of Parity to Connectivity. As mentioned in the introduction, we will be interested in defining Parity only for a very special class of finite subsets X of $\mathbb{C}$: those that are made only of integers with distance either one or two between two consecutive elements of X. Let $\mathcal{X}$ be this class of subsets of $\mathbb{N}$. Along the way, we will give (in Theorem 4.4) a strengthening of the recent result [4] that Parity is not definable over the reals: no first-order formula can correctly "compute" Parity even if we restrict our attention to the input sets X that belong to $\mathcal{X}$. The only property of the reals which will be used for this result is the equivalence between active domain and natural domain semantics (see [5] for a nonconstructive proof, [6] for a constructive proof, and [2] for an efficient translation algorithm). Let K be a field, F a closed formula of $\mathcal{L}^K_{rings} \cup \{I_1\}$ and $X \subseteq K$. F is said to be true under the active domain semantics (this is denoted $X \models F$) if this formula is true when I_1 is interpreted by membership to X and *the range of every quantified variable in F is taken to be X instead of the "natural domain" K*. We refer to [5, 6, 2] for more details. The natural domain semantics $(K, X) \models F$ has already been defined (for $X \subseteq K^n$) in the introduction and is the only semantics used outside this section. Note that the predicate I_1 is no longer needed under the active domain semantics.

Proposition 4.1 *There exists a formula $\phi(x, y)$ of $\mathcal{L}_{rings} \cup \{I_1\}$ which satisfies the following property.*

For any $X \in \mathcal{X}$, let ϕ_X be the subset of $\mathbb{C}^2$ defined by ϕ when I_1 is interpreted by membership to X. Then ϕ_X is connected if and only if $|X|$ is even.

Proof. We use the formula ψ constructed in section 3 (hence $\phi_X = S_X$), but here we define the predicates min, max and succ as follows:

$$\min(x) \equiv I_1(x) \wedge \neg I_1(x-1) \wedge \neg I_1(x-2),$$

$$\max(x) \equiv I_1(x) \wedge \neg I_1(x+1) \wedge \neg I_1(x+2),$$

and $\mathrm{succ}(x, y) \equiv I_1(x) \wedge I_1(y) \wedge [y = x + 1 \vee (y = x + 2 \wedge \neg I_1(x+1))]$. □

The following result is then clear.

Lemma 4.2 *If Connectivity in $\mathbb{C}^2$ is definable, there exists a formula G of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_1\}$ which satisfies the following property:*

(*) for any $X \in \mathcal{X}$, $(\mathbb{C}, X) \models F$ if and only if $|X|$ is even.

Proof. Let F be the formula of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_2\}$ which defines Connectivity: G is obtained from F by replacing each occurrence of I_2 by formula ϕ from Proposition 4.1. □

If (*) holds, we say by abuse of language that Restricted Parity is definable.

In a second reduction, we show that if Restricted Parity is definable, it is also definable over the reals. This follows immediately from the next proposition.

Proposition 4.3 *Let F be a formula of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_1\}$. There exists a formula G over $\mathcal{L}^{\mathbb{R}}_{rings} \cup \{I_1\}$ such that for any finite set $X \subseteq \mathbb{R}$, $(\mathbb{C}, X) \models F$ if and only if $(\mathbb{R}, X) \models G$.*

Proof. Separate real and imaginary parts of variables in F. $\square$

The fact that Restricted Parity is not definable, and Theorem 1.1, will then follow from the next result.

Theorem 4.4 *There exists no formula F of $\mathcal{L}^{\mathbb{R}}_{rings} \cup \{I_1\}$ satisfying the following property: for every $X \in \mathcal{X}$, $(\mathbb{R}, X) \models F$ if and only if $|X|$ is even.*

Corollary 4.5 *There exists no formula F of $\mathcal{L}^{\mathbb{C}}_{rings} \cup \{I_2\}$ satisfying the following property: for every $X \in \mathcal{X}$, $(\mathbb{C}, X) \models F$ if and only if $|X|$ is even.*

These two results are in a sense optimal since Parity becomes definable if we restrict our attention further, by considering only sets X made of consecutive integers.

The remainder of this section is devoted to the proof of Theorem 4.4. By the equivalence between natural and active domain semantics over the reals it is sufficient to prove the following result.

Proposition 4.6 *Fix a first-order structure*

$$M = (\mathbb{N}, \mathcal{R}_1, \ldots, \mathcal{R}_m, f_1, \ldots, f_p)$$

where $\mathcal{R}_i \subseteq \mathbb{N}^{n_i}$ is an arbitrary predicate, and $f_i : \mathbb{N}^{q_i} \to \mathbb{N}$ an arbitrary function. There exists no formula F over M satisfying the following property: for every $X \in \mathcal{X}$, $X \models F$ if and only if $|X|$ is even.

The proof is by a reduction from the familiar Parity problem of complexity theory to Restricted Parity: we will see that if Restricted Parity was definable then Parity would be in AC^0. For this we need to know how fast query in natural semantics can be evaluated. We shall work with the following encoding of finite sets of integers: a vector $u \in \{0,1\}^n$ represents the set $X_u = \{i;\ x_i = 1\}$ (of course there are many different encodings for a given X). It is not hard to see that under this encoding, queries in an arbitrary first-order language can be evaluated in AC^0.

Proposition 4.7 *Fix as in Proposition 4.6 an arbitrary first-order structure over $\mathbb{N}$, and a first-order formula F. Then $\mathrm{Eval}_F \in \mathrm{AC}^0$, where Eval_F denotes the following problem: given $u \in \{0,1\}^*$, decide whether $X_u \models F$.*

Proof. We may assume that F is in prenex form: $F \equiv Q_1 x_1 \cdots Q_k x_k\ G(x_1, \ldots, x_k)$ where G is quantifier-free and $Q_i \in \{\exists, \forall\}$. We now describe a polynomial-size, $O(k)$ depth circuit $C_n(a, u)$ which solves Eval_F for inputs in $u \in \{0,1\}^n$. Here a is a vector of n^k "hardwired" boolean constants corresponding to the n^k elements of $\{1, \ldots, n\}^k$. The component a_x of a associated to $x \in \{1, \ldots, n\}^k$ is 1 if and only if $G(x)$

is true. It is clear that F can be evaluated from a by replacing each existential quantifier by a disjunction, and each universal quantifier by a conjunction. To be completely precise, one can define inductively the formulas $F_k(x_1, \ldots, x_k) \equiv G(x_1, \ldots, x_k)$ and $F_{i-1}(x_1, \ldots, x_{i-1}) \equiv Q_i x_i F_i(x_1, \ldots, x_i)$ (note that $F_0 = F$). The 2^{i-1} formulas $F_{i-1}(x_1, \ldots, x_{i-1})$ are evaluated in parallel as follows. If Q_i is existential,

$$F_{i-1}(x_1, \ldots, x_{i-1}) = \bigvee_{x_i \in X_u} F_i(x_1, \ldots, x_i) = \bigvee_{j=1}^{n} [F_i(x_1, \ldots, x_{i-1}, j) \wedge u_j = 1].$$

If Q_i is universal,

$$F_{i-1}(x_1, \ldots, x_{i-1}) = \bigwedge_{x_i \in X_u} F_i(x_1, \ldots, x_i) = \bigwedge_{j=1}^{n} [F_i(x_1, \ldots, x_{i-1}, j) \vee u_j = 0].$$

□

The next and final lemma completes the proof of Proposition 4.6, Theorem 4.4 and Theorem 1.1.

Lemma 4.8 *If Restricted Parity is definable then* Parity $\in$ AC^0.

Proof. By Proposition 4.7, Restricted Parity can be evaluated in AC^0 if it is definable. The result then follows from a straightforward AC^0 reduction from Parity to Restricted Parity: map $x \in \{0,1\}^n$ to the code $u \in \{0,1\}^{3n}$ satisfying $u_{3i-2} = u_{3i-1} = 1$ and $u_{3i} = x_i$ for $i = 1, \ldots, n$. □

The fact that Parity$\notin$AC^0 was used in [13] to show that Parity is not definable with linear and order constraints (see also [12]).

5 A Logical Proof of Theorem 1.1

In this section we present a self-contained (and direct) proof of Theorem 1.1 modulo some basic model theory and field theory (we refer the reader to [19] and [15] for the basic facts and notions from model theory that we shall used freely). Moreover, the proof works for an arbitrary algebraically closed field. In this general case, we define connectivity using the Zariski topology.

Let K be an algebraically closed field of characteristic p (prime or zero). We denote by ACF_p the theory of algebraically closed fields of characteristic p in $\mathcal{L}_{rings}$. ACF_p is a complete theory. We denote by $\mathcal{L}^*$ the language $\mathcal{L}_{rings} \cup \{I_1, <\}$ where $<$ is a binary predicate. We denote by (M, X) the $\mathcal{L}^*$-structures where M is the base set and X is the interpretation of I_1. We do not stress the interpretation of $<$ in the notation because we shall only consider $\mathcal{L}^*$-structure where the interpretation of $<$ is on I_1.

Let T^* be the theory of $\mathcal{L}^*$ constituted by the following axioms:
(i) the axioms of ACF_p;
(ii) $\forall xy\ x < y \to I_1(x) \wedge I_1(y)$;
(iii) $<$ is a linear order on I_1 and this order is discrete with a smallest and a largest element;
(iv) I_1 is infinite: for every n we consider the axiom

$$\exists x_1 \dots x_n \bigwedge_{i \neq j} x_i \neq x_j \wedge \bigwedge_{i=1}^{n} I_1(x_i)$$

(v) the elements of I_1 are algebraically independent: for every non-zero polynomial $f(x_1, \dots x_n)$ with coefficients in $\mathbb{Z}/p\mathbb{Z}$ we consider the axiom

$$\forall x_1 \dots x_n \bigwedge_{i \neq j} x_i \neq x_j \wedge \bigwedge_{i=1}^{n} I_1(x_i) \to f(x_1, \dots x_n) \neq 0.$$

Note first that T^* is a consistent theory. Indeed, let L be an algebraically closed field of characteristic p with infinite transcendence degree and let X be a transcendence basis of L. Fix on X a discrete linear order $<$ with a smallest and a largest element. Then, clearly, (L, X) is a model of T^*.

The technical result of this section is the following proposition, which one may consider as folklore.

Proposition 5.1 *T^* is a complete theory. Moreover, if Σ is a finite subset of T^*, there exists an integer n such that for all integer $m \geq n$ there exists a subset X of K of cardinality m such that $(K, X) \models \Sigma$ for an arbitrary linear order on X.*

Proof. First we show the second part of the proposition. It is easy to see that it suffices to prove that for every integer m and every finite family of non-zero polynomials $f_j(x_1, \dots x_{n_j})$, $j = 1, \dots s$, with coefficients in $\mathbb{Z}/p\mathbb{Z}$ and with $n_j \leq m$ indeterminates, K satisfies the sentence

$$\exists x_1 \dots x_m \bigwedge f_j(x_{i_1}, \dots, x_{i_{n_j}}) \neq 0.$$

where the conjunction is taken over the $j = 1, \dots s$ and the sequences $(i_1, \dots i_{n_j})$ of distinct elements of $\{1, \dots, m\}$. Since, K has an elementary extension with infinite transcendence degree such a sentence is always satisfied in K.

Now we shall show that T^* is complete. We denote by $\mathcal{L}_0$ the sublanguage $\{=, <\}$ of $\mathcal{L}^*$. Let $(\mathcal{M}, \mathcal{X})$ and $(\mathcal{N}, \mathcal{Y})$ be two $\aleph_0$-saturated models of T^*. Since any completion of T^* has an $\aleph_0$-saturated model, it suffices to prove that $(\mathcal{M}, \mathcal{X})$ and $(\mathcal{N}, \mathcal{Y})$ are elementarily equivalent in $\mathcal{L}^*$. Clearly, the $\mathcal{L}_0$-structures $\mathcal{X}$ and $\mathcal{Y}$ are $\aleph_0$-saturated models of the theory of discrete linear orders with endpoints. Since this theory is complete, by $\aleph_0$-saturation, there exists a set Λ_0 of $\mathcal{L}_0$-isomorphism $\sigma_0 : X \longrightarrow Y$ where X is a finite $\mathcal{L}_0$-substructure of $\mathcal{X}$ and Y is a finite $\mathcal{L}_0$-substructure of $\mathcal{Y}$, with the back-and-forth property.

We consider the set Λ of $\mathcal{L}^*$-isomorphisms $\sigma : (M, X) \longrightarrow (N, Y)$ where (M, X) is an $\mathcal{L}^*$-substructure of $(\mathcal{M}, \mathcal{X})$ and (N, Y) is an $\mathcal{L}^*$-substructure of $(\mathcal{N}, \mathcal{Y})$ such that:
(i) the restriction of σ to X is an element of Λ_0;
(ii) there exists a tuple $\alpha_1, \dots \alpha_n$ (possibly empty) in M algebraically independent over $\mathcal{X}$ and a tuple $\beta_1, \dots \beta_n$ in N algebraically independent over $\mathcal{Y}$ such that M is the algebraic closure of $\{\alpha_1, \dots \alpha_n\} \cup X$ and N is the algebraic closure of $\{\beta_1, \dots \beta_n\} \cup Y$.
To prove that $(\mathcal{M}, \mathcal{X})$ and $(\mathcal{N}, \mathcal{Y})$ are elementarily equivalent for $\mathcal{L}^*$ it suffices to prove that Λ is nonempty and has the back-and-forth property.

Let us show that Λ is nonempty. Let $\sigma_0 : X \longrightarrow Y$ be an element of Λ_0. Since X and Y are constituted of algebraically independent elements of $\mathcal{M}$ and $\mathcal{N}$ respectively, σ_0 extends to a fields isomorphism σ from the algebraic closure M_X of X in $\mathcal{M}$ into the algebraic closure N_Y of Y in $\mathcal{N}$. Moreover, it is easy to see that $M_X \cap \mathcal{X} = X$ and that $N_Y \cap \mathcal{Y} = Y$, thus (M_X, X) is an $\mathcal{L}^*$-substructure of $(\mathcal{M}, \mathcal{X})$, (N, Y) is an $\mathcal{L}^*$-substructure of $(\mathcal{N}, \mathcal{Y})$ and σ is in fact an $\mathcal{L}^*$-isomorphism. Then, clearly, σ is an element of Λ.

Let us show that Λ has the back-and-forth property. By symmetry it is enough to show that Λ has the forth property. So, let $\sigma : (M, X) \longrightarrow (N, Y)$ be an element of Λ and assume that M is the algebraic closure of $\{\alpha_1, \dots \alpha_n\} \cup X$ where the α_i are algebraically independent over $\mathcal{X}$ and N is the algebraic closure of $\{\beta_1, \dots \beta_n\} \cup Y$ where the β_i are algebraically independent over $\mathcal{Y}$. Let α be an element of $\mathcal{M}$. Let us denote by σ_0 the restriction of σ to X. Of course we may assume that α is not in M and thus algebraically independent over M.

Firstly, assume that $\alpha \in \mathcal{X}$. Since Λ_0 has the forth property there exists a $\beta \in \mathcal{Y}$ such that the map $\hat{\sigma}_0 : X \cup \{\alpha\} \longrightarrow Y \cup \{\beta\}$ which extend σ_0 and which send α on β is in Λ_0. Note that β is algebraically independent over N ($N \cap \mathcal{Y} = Y$ since (N, Y) is an $\mathcal{L}^*$-substructure of $(\mathcal{N}, \mathcal{Y})$). Now, since α is algebraically independent over M, σ extends to a fields isomorphism $\hat{\sigma}$ from the algebraic closure M' of $M \cup \{\alpha\}$ into the algebraic closure N' of $N \cup \{\beta\}$ which send α on β. Set $X' = X \cup \{\alpha\}$ and $Y' = Y \cup \{\beta\}$. We claim that $M' \cap \mathcal{X} = X'$ and that $N' \cap \mathcal{Y} = Y'$. Indeed, let $a \in M' \cap \mathcal{X}$. Then, a is algebraic over $\{\alpha_1, \dots \alpha_n\} \cup X'$. Since the α_i are algebraically independent over $\mathcal{X}$ and since $a \in \mathcal{X}$ it follows that a is algebraic over X'. Thus, $a \in X'$ because $\mathcal{X}$ is a set of algebraically independent elements. The same proof shows that $N' \cap \mathcal{Y} = Y'$. It follows that $\hat{\sigma}$ is in fact an $\mathcal{L}^*$-isomorphism between $\mathcal{L}^*$-substructures of $(\mathcal{M}, \mathcal{X})$ and $(\mathcal{N}, \mathcal{Y})$. Then, clearly, $\hat{\sigma}$ is in Λ.

Secondly, assume that α is algebraic over $M \cup \mathcal{X}$. Then, there exist elements $a_1, \dots a_m$ of $\mathcal{X}$ such that α is algebraic over $M \cup \{a_1, \dots a_m\}$.

Applying, m times the above case we obtain an element $\hat{\sigma} : M' \longrightarrow N'$ of Λ such that the a_i are in M'. Since M' is algebraically closed $\alpha \in M'$.

Finally, assume that α is algebraically independent over $M \cup \mathcal{X}$. Note first that $\alpha, \alpha_1, \dots \alpha_n$ are algebraically independent over $\mathcal{X}$. Assume that we have found an element β of $\mathcal{N}$ algebraically independent over $N \cup \mathcal{Y}$. Then, $\beta, \beta_1, \dots \beta_n$ are algebraically independent over $\mathcal{Y}$. Moreover, we can extend σ to a fields isomorphism $\hat{\sigma} : M' \longrightarrow N'$ which send α on β where M' is the algebraic closure of $M \cup \{\alpha\}$ and where N' is the algebraic closure of $N \cup \{\beta\}$. Again to show that $\hat{\sigma}$ is in Λ, it suffices to show that $M' \cap \mathcal{X} = X$ and $N' \cap \mathcal{Y} = Y$. So, let $a \in M' \cap \mathcal{X}$. Assume that a is not in X. Then, since $M \cap \mathcal{X} = X$, a is algebraic over $M \cup \{\alpha\}$ but not algebraic over M. It follows (by the exchange law) that α is algebraic over $M \cup \{a\}$. This is absurd by hypothesis on α. In the same way, one shows that $N' \cap \mathcal{Y} = Y$.

Thus, to complete the proof of the proposition we just need to show that there exists a β in $\mathcal{N}$ which is algebraically independent over $N \cup \mathcal{Y}$. Let us consider the set $\Gamma(y)$ of formulas of the form

$$\forall x_1 \dots x_m \bigwedge_{i \neq j} x_i \neq x_j \wedge \bigwedge_{i=1}^{m} I_1(x_i) \rightarrow f(y, x_1, \dots x_m, \beta_1, \dots, \beta_n) \neq 0$$

where f is a nonzero polynomial with coefficients in $\mathbb{Z}/p\mathbb{Z}$. It suffices to show that $\Gamma(y)$ is satisfiable in $\mathcal{N}$. Since $\mathcal{N}$ is $\aleph_0$-saturated we only need to show that $\Gamma(y)$ is finitely satisfiable in $\mathcal{N}$. So, let $\Gamma_0(y)$ be a finite subset of $\Gamma(y)$. There exists an integer d such that every polynomial which "appears in" in $\Gamma_0(y)$ has degree at most d in y. Let $\mathcal{N}_0$ be the subfield of $\mathcal{N}$ generated by $\{\beta_1, \dots \beta_n\} \cup \mathcal{Y}$. $\mathcal{N}_0$ is isomorphic to the field of rational fractions $K_0(\mathcal{T})$ where K_0 is the prime field of characteristic p and where $\mathcal{T}$ is a set of indeterminates of the same cardinality than $\{\beta_1, \dots \beta_n\} \cup \mathcal{Y}$. There exists an irreducible polynomial $g(u)$ of $K_0(\mathcal{T})[u]$ of degree $d+1$ (one may consider the polynomial $u^{d+1} - t$ for a $t \in \mathcal{T}$). Since $\mathcal{N}$ is algebraically closed it follows that there exists an element β in $\mathcal{N}$ of degree $d+1$ over $\mathcal{N}_0$. By definition of d this element satisfy $\Gamma_0(y)$. This completes the proof of the proposition. □

We are now ready to prove Theorem 1.1 for K. First we note that there exists a formula $\psi(x, y)$ of $\mathcal{L}^*$ such that if X is a finite subset of K linearly ordered by $<$, then in the associated $\mathcal{L}^*$-structure (K, X) the subset of K^2 defined by $\psi(x, y)$ is definable in K and is connected iff $|X|$ is even. Such a formula is constructed in Section 3. Assume that there exists a sentence $F(I_2)$ of $\mathcal{L}_{rings} \cup \{I_2\}$ such that if A is a definable subset of K^2, $(K, A) \models F(I_2)$ iff A is connected. We denote by $F(\psi)$ the sentence of $\mathcal{L}^*$ obtained from $F(I_2)$ by replacing each occurrence of I_2 by ψ. We shall obtain a contradiction by showing that $T^* \cup \{F(\psi)\}$ and $T^* \cup \{\neg F(\psi)\}$ are consistent theories. This is absurd because, by Proposition 5.1, T^* is complete. Let Σ be a finite subset of T^*. By Proposition 5.1, there exists a finite subset X of K of even cardinality such that given a linear order on X, the $\mathcal{L}^*$-structure $(K, X) \models \Sigma$. Moreover, since X is finite of even cardinality the set defined by $\psi(x, y)$ in (K, X) is connected and definable in K, thus $(K, X) \models F(\psi)$. By Proposition 5.1, there also exists a finite subset Y of K of odd cardinality such that given a linear order on Y, the $\mathcal{L}^*$-structure $(K, Y) \models \Sigma$. Again, since Y is finite the set defined by $\psi(x, y)$ in (K, Y) is definable in K. But, since $|Y|$ is odd, $(K, Y) \models \neg F(\psi)$. We have shown that for every finite subset Σ of T^*, $\Sigma \cup \{F(\psi)\}$ and $\Sigma \cup \{\neg F(\psi)\}$ are consistent theories. By compactness, $T^* \cup \{F(\psi)\}$ and $T^* \cup \{\neg F(\psi)\}$ are consistent theories.

Note that the above proof works for sentence $F(I_2)$ without parameters from K. In the case where $F(I_2)$ contains a tuple of parameters $\bar{\alpha}$ from K a slight modification of the theory T^* yields a proof in this case. For if, we add a tuple a constants in $\mathcal{L}_{rings}$ for naming the α_i and instead of ACF_p we consider the theory of K in $\mathcal{L}_{rings}$ (or equivalently we add to ACF_p the diagram of $\bar{\alpha}$). Moreover, in the new theory T^* we say that the elements of I_1 are algebraically independent over $\bar{\alpha}$. Then, Proposition 5.1 holds for this theory T^* (with essentially the same proof, we just need to work over $\bar{\alpha}$).

6 Elimination of Parameters

Let K be an algebraically closed field. Let $\mathcal{P}$ be a property of K^n defined by a sentence F in the language $\mathcal{L}_{rings} \cup \{I_n\}$ with parameters in K. One may ask to which extent it is possible to eliminate the parameters, i.e., to define $\mathcal{P}$ by a sentence F in the language $\mathcal{L}_{rings} \cup \{I_n\}$ without parameters. More generally, given a subfield $k \subset K$, one may try to define $\mathcal{P}$ with parameters in k only.

We have mentioned at the end of section 5 that Parity remains undefinable in the presence of parameters. In this section, we show that parameters can be eliminated for a large class of properties (in fact this class is as large as possible). In order to investigate the definability of a property in this class, one is therefore free to focus on parameter-free definability.

Let k be a subfield of K and let $\mathcal{P}$ be a family of definable subsets of K^n. We say that $\mathcal{P}$ is locally definable with parameters in k if for every (parameter-free) formula $\phi(x_1, \ldots x_n, y_1, \ldots, y_l)$ of $\mathcal{L}_{rings}$, there exists a formula $\psi(\bar{y})$ of $\mathcal{L}^k_{rings}$ such that for all $a \in K^l$, $\phi(K, a) \in \mathcal{P}$ iff $K \models \psi(a)$ (here $\phi(K, a)$ is the subset of K^n defined by the formula $\phi(\bar{x}, a)$). Let us give an example (one can also prove this lemma for irreducible closed sets).

Lemma 6.1 *The family of closed sets of an algebraically closed field is locally definable without parameters. More precisely, for every (parameter-free) formula $\phi(x_1, \ldots x_n, y_1, \ldots, y_l)$ of $\mathcal{L}_{rings}$, there exists a parameter-free* *formula $\psi(\bar{y})$ of $\mathcal{L}_{rings}$ such that if L is an algebraically closed field, then for all $a \in L^l$, the set defined by $\phi(\bar{x}, a)$ in L is closed iff $L \models \psi(a)$.*

Proof. Since the theory of algebraically closed fields admits quantifier elimination, we may assume that ϕ is quantifier-free. if $\phi(L, a)$ is closed, then $\phi(L, a)$ is of pseudo-degree $\leq D$ for a D which depends only on $\phi(\bar{x}, \bar{y})$. We may apply Proposition 2.3 (which does not depend on the field under consideration) to complete the proof of the lemma : replace $I_n(\bar{x})$ by $\phi(\bar{x}, \bar{y})$ in the formula of Proposition 2.3. □

Clearly, if $\mathcal{P}$ is definable with parameters in k it is also locally definable with parameters in k. In this section we show that the converse is true.

Theorem 6.2 *Let k be a subfield of K. A definable property of K^n is definable with parameters in k if and only if it is locally definable with parameters in k.*

We first eliminate algebraic parameters.

Lemma 6.3 *Let $\mathcal{P}$ be a property of K^n which is definable with parameters in an algebraic extension $k[\alpha]$ of a field $k \subset K$. If $\mathcal{P}$ is locally definable with parameters in k, it is also definable with parameters from k only.*

Proof. Let m be the minimal polynomial of α over k. Property $\mathcal{P}$ is defined by a formula $F(\alpha)$ where the parameters of $F(z)$ are in k. We claim that this property is also defined by the following formula G:

$$\forall \beta\ [m(\beta) = 0 \Rightarrow F(\beta)].$$

Let X be a definable subset of K^n. If $(K, X) \models G$ it is clear that $(K, X) \models F(\alpha)$ (take $\beta = \alpha$). Conversely, assume that $(K, X) \models F(\alpha)$ and that X is defined by a formula $\phi(\bar{x}, a)$ where $a \in K^l$. Since $\mathcal{P}$ is locally definable with parameters in k, there exists a formula $\psi(\bar{y})$ with parameters in k such that

$$\forall b \in K^l\ [K \models \psi(b) \text{ iff } (K, \phi(K, b)) \models F(\alpha)]. \tag{2}$$

Let Γ be the set of element of K which can "play the role" of α in (2). That is, $\gamma \in \Gamma$ if and only if γ satisfies the following formula $\Gamma(z)$ of $\mathcal{L}_{rings}$

$$\forall \bar{y}\ \ [\psi(\bar{y}) \Leftrightarrow F_{\phi(\bar{x}, \bar{y})}(z)] \tag{3}$$

where $F_{\phi(\bar{x}, \bar{y})}$ is the formula obtained from F by substitution of $\phi(\bar{x}, \bar{y})$ to I_n. Since (3) has parameters in k and is satisfied by α, it is also satisfied by the conjugates of α. Since $\psi(a)$ holds true, this implies in particular that

$$\forall \beta\ [m(\beta) = 0 \Rightarrow F_{\phi(\bar{x}, a)}(\beta)],$$

that is, $(K, X) \models G$. □

We now eliminate algebraically independent parameters.

Lemma 6.4 *Let $\mathcal{P}$ be a property of K^n which is definable with parameters in an extension $k(\alpha)$ of a field $k \subset K$, where $\alpha = (\alpha_1, \ldots, \alpha_m)$ is a tuple of elements of K which are algebraically independent over k. If $\mathcal{P}$ is locally definable with parameters in k, it is also definable with parameters in k only.*

Proof. Property $\mathcal{P}$ is defined by a formula $F(\alpha)$ where the parameters of $F(\bar{z})$ are in k. We claim that $(K, X) \models F(\alpha)$ if and only if $(K, X) \models F(\beta)$ for a generic β. This will prove the theorem since, as we have seen in section 2 (see the proof of Proposition 2.1), $\mathcal{P}$ can then be defined by

$$\exists t_1, \ldots, t_{m+1} \in K^m \; \forall \beta \in K^m \; \bigvee_{i=1}^{k+1} F(\beta - t_i).$$

The proof of the claim is similar to the proof of Lemma 6.3. Let $\phi(x_1, \ldots, x_n, \bar{y})$ be a parameter-free formula. Since $\mathcal{P}$ is locally definable with parameters in k, there exists a formula $\psi(\bar{y})$ with parameters in k such that

$$\forall b \in K^l \; [K \models \psi(b) \text{ iff } (K, \phi(K, b)) \models F(\alpha)]. \tag{4}$$

Consider the formula $\Gamma(z_1, \ldots, z_k)$ of $\mathcal{L}_{rings}$

$$\forall \bar{y} \;\; [\psi(\bar{y}) \Leftrightarrow F_{\phi(\bar{x}, \bar{y})}(\bar{z})]. \tag{5}$$

Since (5) has parameters in k and is satisfied when $\bar{z} = \alpha$, it is satisfied by a generic $\beta \in K^m$. Hence for a generic β and any $a \in K^l$, $(K, \phi(K, a)) \models F(\alpha)$ if and only if $(K, \phi(K, a)) \models F(\beta)$. $\square$

One can also use the special case $m = 1$ to prove the lemma by induction on m (this does not result in any significant simplification).

Proof of Theorem 6.2. Using Lemma 6.3 repeatedly, we can assume that $\mathcal{P}$ is definable with parameters in $k \cup \{\alpha_1, \ldots, \alpha_m\}$ where $\alpha_1, \ldots, \alpha_m$ are algebraically independent over k. We eliminate $\alpha_1, \ldots, \alpha_m$ with Lemma 6.4. $\square$

Problem 6.5 *Is Theorem 6.2 still true for $\mathbb{R}$?*

We conclude this section with an application of Theorem 6.2 (which holds as well for irreducible closed sets).

Corollary 6.6 *Let p be prime or equal to zero. Assume that there exists an algebraically closed field K of characteristic p such that the family of closed sets of K^n is definable. Then, there exists a parameter-free sentence $F(I_n)$ which defines the family of closed set of L^n for every algebraically closed field of characteristic p.*

Proof. By Theorem 6.2 and Lemma 6.1 if the family of closed set of K^n is definable, then it is definable with a sentence $F(I_n)$ without parameters. Then, using Lemma 6.1 again and the fact that two algebraically closed fields of same characteristic are elementarily equivalent, it is easy to see that $F(I_n)$ defines the family of closed set of L^n for every algebraically closed field of characteristic p. □

7 Definable Sets over Finite Structures

Again in this section K is an algebraically closed field. Let $\mathcal{L}_0 = \{R_1, \ldots, R_u\}$ be a finite set of relational symbols with R_i of arity r_i. Let $\Phi(x_1, \ldots, x_n)$ be a formula of $\mathcal{L}^K_{rings} \cup \mathcal{L}_0$. Given for each R_i a finite subset X_i of K^{r_i}, if we interpret R_i by X_i, the formula $\Phi(\bar{x})$ define a subset of K^n which is definable by a formula of $\mathcal{L}^K_{rings}$. We denote by $\mathcal{X}_\Phi$ the family of all these possible definable subset of K^n.

To show that connectivity is not a definable property of K^2, we have shown that there exists a family of the form $\mathcal{X}_\Phi$ (with $\mathcal{L}_0 = \{I_1, <\}$) such that there exists no sentence $F(I_2)$ that can recognize the connected sets in $\mathcal{X}_\Phi$. The point is that if such a sentence exists, then, roughly speaking, a given class of finite $\mathcal{L}_0$-structures (totally ordered structure of even cardinality) becomes finitely axiomatisable in a given world (which yields a contradiction).

For closed sets the situation is different. We say that a formula $\Phi(x_1, \ldots, x_n)$ of $\mathcal{L}^K_{rings} \cup \mathcal{L}_0$ is of type (m, d) if it is equivalent to a formula in prenex form which contains at most m quantifiers and where every polynomial $f(\bar{x}, \bar{z})$ of $K[\bar{x}, \bar{z}]$ which appears in this formula is of degree $\leq d$.

Theorem 7.1 *Let n, d, m be three integers ≥ 1. There exists a sentence $F_{n,m,d}$ of $\mathcal{L}_{rings} \cup \{I_n\}$ such that if $\mathcal{L}_0$ is a finite set of relational symbols and if $\Phi(x_1, \ldots, x_n)$ is a formula of $\mathcal{L}^K_{rings} \cup \mathcal{L}_0$ of type (m, d) then : if $X \in \mathcal{X}_\Phi$, X is closed if and only if $(K, X) \models F_{n,m,d}$.*

This result gives a kind of strengthening of Lemma 6.1. Of course, Theorem 7.1 no longer holds if we replace "X is closed" by "X is connected." The main point in the proof is the following result.

Proposition 7.2 *There exists a function B of $\mathbb{N} \times \mathbb{N}$ into $\mathbb{N}$ such that the following holds. Let $\phi(x_1, \ldots x_n)$ be a formula of the form*

$$Q_1 z_1 \ldots Q_m z_m \quad \phi^*(x_1, \ldots x_n, z_1, \ldots z_m)$$

where the Q_i are $\exists$ or $\forall$ and where ϕ^* is a boolean combination of atomic formulas of the form $f(\bar{x}, \bar{z}) = 0$ where f is a polynomial of $K[\bar{x}, \bar{z}]$ of degree $\leq d$. Then, $\phi(x_1, \ldots x_n)$ is equivalent to a quantifier-free formula $\psi(x_1, \ldots x_n)$ which is a boolean combination of atomic formulas of the form $f(\bar{x}) = 0$ where f is a polynomial of $K[\bar{x}]$ of degree $\leq B(m, d)$.

In fact for the proof of Theorem 7.1 we only need of a bound which depends only on m, d and n. However it is not more difficult to obtain a bound which depends only on m and d. Note that the usual bound depends on m, n and the *sum* of the degree of the polynomials which appear in the formula. The proof below shows that we may take for B the function $(d+1)^{11^m}$ (using the best available bound for Hilbert's Nullstellensatz [17]). One can prove that there is no simply exponential bound. However, one can hope to prove the existence of a bound of the form $(d+2)^{\prod O(m_i)}$ in the case where, in the formula of the proposition, z_i is a tuple of variable of length m_i (i.e., one can hope to obtain a simply exponential bound if the number of alternation of quantifier is fixed). Note that such a bound is known to be true in the case of real-closed fields in the language of ordered rings (see [3] and in this case the other "complexity parameters" are quite optimal). However, we have not found such a result for algebraically closed fields neither Proposition 7.2 in the literature (see [10] for the "faster" algorithm for algebraically closed fields). One possible reason is that it is perhaps difficult to obtain such results if one want to keep reasonable bounds on the other "complexity parameters".

Proof of Proposition 7.2. It is easy to see that we may assume that $\phi(x_1, \ldots x_n)$ is of the form $\exists z\ \phi^*(\bar{x}, z)$ where ϕ^* is equal to

$$\bigwedge_{i=1}^{s} f_i(\bar{x}, z) = 0 \wedge \bigwedge_{i=1}^{t} g_i(\bar{x}, z) \neq 0$$

where the degree of the f_i and the g_i are $\leq d$. We may also assume that $d \geq 1$. We write:

$$f_i(\bar{x}, z) = \sum_j p_{j,i}(\bar{x}) z^j \quad \text{and} \quad g_i(\bar{x}, z) = \sum_j q_{j,i}(\bar{x}) z^j.$$

We denote by $\mu(\bar{x})$ the formula

$$\bigwedge_{i,j} p_{j,i}(\bar{x}) = 0$$

and we denote by $\theta(\bar{x})$ the formula

$$\bigwedge_i \left(\bigvee_j q_{j,i}(\bar{x}) \neq 0 \right).$$

For a $a \in K^n$, $K \models \mu(a)$ iff $\{b \in K \mid K \models \bigwedge_{i=1}^{s} f_i(a, b) = 0\}$ is equal to K. Moreover, if $K \models \neg\mu(a)$, then the above set is finite of cardinality $\leq d$ since it is the intersection of zero sets of non-zero polynomials in one variable of degree $\leq d$. On the other hand, For a $a \in K^n$, $K \models \theta(a)$

iff $\{b \in K \mid K \models \bigwedge_{i=1}^{s} g_i(a,b) \neq 0\}$ is nonempty (the intersection of two cofinite subsets of K is cofinite).

Let us first assume that $t > d$. For a subset I of $\{1,\dots t\}$ we denote by $\phi_I(\bar{x})$ the formula

$$\exists z \ \bigwedge_{i=1}^{s} f_i(\bar{x},z) = 0 \wedge \bigwedge_{i\in I} g_i(\bar{x},z) \neq 0.$$

Then, we consider the formula $\phi_1(\bar{x})$

$$(\ \mu(\bar{x}) \wedge \theta(\bar{x})\ ) \ \vee \ (\ \neg\mu(\bar{x}) \wedge \bigwedge_{I\subseteq\{1,\dots t\} \text{ and } |I|=d} \phi_I(\bar{x})\ )$$

Let us prove that $\phi(\bar{x})$ is equivalent to $\phi_1(\bar{x})$. Let $a \in K^n$. It is clear that if $K \models \phi(a)$, then $K \models \phi_1(a)$. Assume that $K \models \neg\phi(a)$. If $K \models \mu(a)$, then $K \models \neg\theta(a)$ and clearly $K \models \neg\phi_1(a)$. Thus we assume that $K \models \neg\mu(a)$ and the set $B = \{b \in K \mid K \models \bigwedge_{i=1}^{s} f_i(a,b) = 0\}$ is finite of cardinality $\leq d$. Since $K \models \neg\phi(a)$, for all $b \in B$ there exists a $i_b \in \{1,\dots t\}$ such that $g_{i_b}(a,b) = 0$. Thus there exists a subset I of $\{1,\dots t\}$ of cardinality d such that $K \models \neg\phi_I(a)$. It follows that $K \models \neg\phi_1(a)$.

The above paragraph shows that to prove the proposition we may assume that $t \leq d$. Then, the formula ϕ is equivalent to the formula:

$$\exists zw \ \bigwedge_{i=1}^{s} f_i(\bar{x},z) = 0 \wedge g(\bar{x},z,w) = 0$$

where $g(\bar{x},z,w) = (\prod_{i=1}^{t} g_i(\bar{x},z)))w - 1$. The point is that since $t \leq d$ the degree of g is $\leq d^2 + 1$.

Now we apply the "effective" Hilbert Nullstellensatz of [17] (see [8] for a similar bound and a more elementary proof; moreover, if one only wants to prove the existence of B one may use [9] or the bound of G. Hermann with the proof of [20]). The negation of $\phi(\bar{x})$ is true iff there exists polynomials $h_1(z,w),\dots h_s(z,w), h_s(z,w)$ with coefficients in the field of fractions of $K[\bar{x}]$ and of degree (in z and w) $\leq (d^2+2)^2 \overset{\text{def}}{=} d_1$ such that $gh + \sum_i f_i h_i = 1$. Thus, the negation of $\phi(\bar{x})$ is true iff a system (*) of $u \leq (d_1+1)(d_1+2)/2 \overset{\text{def}}{=} d_2$ equations in $v \leq (s+1)d_2$ unknowns has a solution in the field of fractions $K[\bar{x}]$. Let $A = (r_{i,j}(\bar{x}))_{1\leq i\leq u,\, 1\leq j\leq v}$ be the matrix associated to the homogeneous system and let B be the matrix of (*) (i.e., the matrix constituted of A plus a column of zeros and one 1). The polynomials $r_{i,j}(\bar{x})$ which appear in A and B come from the coefficients of the f_i and of g (viewed as polynomial in z and w). They are of degree $\leq d^2$. Moreover, (*) has a solution iff $rank(A) = rank(B)$. The condition $rank(A) = r$ is equivalent to (i) there exists a $r \times r$ sub-matrix of A with a nonzero determinant and (ii) every $(r+1)\times(r+1)$ sub-matrix of A has a zero determinant (or $r = u$).

Thus, since r should be $\leq d_2$, the condition $rank(A) = r$ is equivalent to a quantifier free formula $\rho_r^A(\bar{x})$ which is a boolean combination of atomic formulas of the form $f(\bar{x}) = 0$ where f is of degree $\leq d_2 d^2$. We have the same kind of formulas $\rho_r^B(\bar{x})$ for B. Then, $\neg\phi(\bar{x})$ is equivalent to the formula

$$\bigvee_{r=0}^{u} (\ \rho_r^A(\bar{x}) \wedge \rho_r^B(\bar{x})\ ).$$

Thus, $\phi(\bar{x})$ is equivalent to a quantifier-free formula which is a boolean combination of atomic formulas of the form $f(\bar{x}) = 0$ where f is of degree $\leq d_2 d^2$. This completes the proof since $d_2 d^2$ depends only on d. To get a compact bound one may check that $d_2 d^2 \leq (d+1)^{10}$. $\square$

It is possible to use more elementary GCD computations instead of the effective Nullstellensatz in the proof of this proposition.

Corollary 7.3 *Let m, d be two integers ≥ 1. There exists an integer B such that if $\mathcal{L}_0$ is a finite set of relational symbols, if $\Phi(x_1, \ldots, x_n)$ is a formula of $\mathcal{L}_{rings}^K \cup \mathcal{L}_0$ of type (m, d) and if $X \in \mathcal{X}_\Phi$, then X is defined by a quantifier-free formula which is a boolean combination of formulas of the form $f(\bar{x}) = 0$ where f is a polynomial of $K[\bar{x}]$ of degree $\leq B$.*

Proof. We may assume that Φ is in prenex form :

$$Q_1 z_1 \ldots Q_m z_m \quad \Phi^*(x_1, \ldots x_n, z_1, \ldots z_m)$$

with Φ^* quantifier-free and where the polynomials which appear in Φ^* is of degree $\leq d$. Set $B = B(m, d)$ where B is the function of the above proposition. Let $X_1, \ldots X_u$ be finite interpretations of $P_1, \ldots P_u$. Then, the set that $\Phi(\bar{x})$ defines is defined by a formula of $\mathcal{L}_{rings}^K$ $Q_1 z_1 \ldots Q_m z_m \quad \phi^*(x_1, \ldots x_n, z_1, \ldots z_m)$. This formula is obtained from Φ by replacing each occurrence of a subformula of Φ^* of the form $P_i(f_1(\bar{x}, \bar{z}), \ldots f_{r_i}(\bar{x}, \bar{z}))$ by a formula of the form

$$\bigvee_{\bar{a} \in X_i} \bigwedge_{j=1}^{r_i} f_j(\bar{x}, \bar{z}) = a_j.$$

Clearly, the degree of the polynomials which appear in ϕ^* are bounded by d. The above proposition completes the proof. $\square$

Proof of Theorem 7.1. Immediate consequence of the above corollary, Lemma 2.4 and Proposition 2.3. $\square$

Again, one can prove Theorem 7.1 for irreductible closed sets.

Acknowledgments

A construction akin to the construction of section 3 seems to be well-known in the real case. The real construction was pointed out to the authors by Saugata Basu, and inspired some of this work.

References

[1] M. Ajtai, Σ_1^1-formulae on finite structures, *Ann. Pure Appl. Logic* **24** (1983) 1–48.

[2] S. Basu, An improved algorithm for quantifier elimination over real closed fields, In: *38th IEEE Symposium on Foundations of Computer Science*, 56–65, 1997.

[3] S. Basu, R. Pollack and M.-F. Roy, On the combinatorial and algebraic complexity of quantifier-elimination, *J. ACM* **43** (1996) 1002–1045.

[4] M. Benedikt, G. Dong, L. Libkin, and L. Wong, Relational expressive power of constraint query languages, *J. ACM* **45** (1998) 1-34.

[5] M. Benedikt and L. Libkin, On the structure of queries in constraint query languages, In: *11th Annual IEEE Symposium on Logic in Computer Science (New Brunswick, NJ, 1996)*, 25–34, IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

[6] M. Benedikt and L. Libkin, Languages for relational databases over interpreted structures, In: *Proceedings of the 16th ACM Symposium on Principles of Database Systems*, 87–98, 1997.

[7] A. Borovik and A. Nesin, *Groups of finite Morley rank*, Oxford University Press, Oxford, 1994.

[8] L. Caniglia, A. Galligo and J. Heintz, Borne simple exponentielle pour les degrés dans le théorème des zéros sur un corps de caractéristique quelconque, *C. R. Acad. Sci. Paris Serie I* **307** (1988) 255–258.

[9] L. van den Dries and K. Schmidt, Bound in the theory of polynomials rings over fields. A nonstandard approach, *Invent. math.* **76** (1984) 77-91.

[10] N. Fitchas, A. Galligo and J. Morgenstern, Precise sequential and parallel complexity bounds for quantifier elimination over algebraically closed fields, *J. Pure Appl. Algebra* **67** (1990) pp. 1–14.

[11] M. Furst, J. Saxe, and M. Sipser, Parity, circuits and the polynomial-time hierarchy. *Math. Systems Theory* **17** (1984) 13–27.

[12] S. Grumbach and J. Su, Queries with arithmetical constraints, *Theoret. Comput. Sci.* **173** (1997) pp. 151–181.

[13] S. Grumbach, J. Su, and C. Tollu, Linear constraint query languages: Expressive power and complexity, In: *Logic and computational complexity (Indianapolis, IN, 1994)*, 426–446, Lecture Notes in Comput. Sci., 960, Springer, Berlin, 1995.

[14] J. Håstad, *Computational limitations for small depth circuits.* MIT Press, 1987.

[15] W. A. Hodges, *Model theory*, Cambridge University Press, Cambridge, 1993.

[16] P. Koiran, Elimination of parameters in the polynomial hierarchy, LIP research report 98-15, 1998.

[17] J. Kollàr, Sharp effective Nullstellensatz, *J. Amer. Math. Soc.* **1** (1988) pp. 963–975.

[18] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications (second edition)*, Graduate Texts in Computer Science, Springer-Verlag, New-York, 1997.

[19] B. Poizat, *Cours de théorie des modèles*, Nur al'Mantiq wal-Ma'rifah, Villeurbanne, 1985.

[20] A. Seidenberg, Constructions in algebra, *Trans. Amer. Math. Soc.* **197** (1974) pp. 273-313,

[21] I. Shafarevich, *Basic Algebraic Geometry 2*, Springer-Verlag, Berlin, 1994.

[22] A. Yao, Separating the polynomial-time hierarchy by oracles, In *Proc. 26th IEEE Symposium on Foundations of Computer Science*, 1–10, 1985.