Isomorphism Theorem for BSS Recursively Enumerable Sets over Real Closed Fields

C. Michaux
Christian.Michaux@umh.ac.be

C. Troestler [1]
Christophe.Troestler@umh.ac.be
Université de Mons-Hainaut[2]

NeuroCOLT2 Technical Report Series

NC2-TR-1998-013

May, 1998[3]

Produced as part of the ESPRIT Working Group
in Neural and Computational Learning II,
NeuroCOLT2 27150

NeuroCOLT2 Coordinating Partner

Royal Holloway
University of London

Department of Computer Science
Egham, Surrey TW20 0EX, England

For more information contact John Shawe-Taylor at the above address
or email neurocolt@neurocolt.com

[1]Chargé de recherches au Fonds National Belge de la Recherche Scientifique.
[2]Institut de Mathématique et d'Informatique, Université de Mons-Hainaut,
Avenue Maistriau, 15 , B-7000 Mons (Belgium)
[3]Received 29-MAR-1998

Abstract

The main result of this paper lies in the framework of BSS computability : it shows roughly that any recursively enumerable set S in R^N, $N \leqslant \infty$, where R is a real closed field, is isomorphic to $R^{\dim S}$ by a bijection φ which is decidable over S. Moreover the map $S \mapsto \varphi$ is computable. Some related matters are also considered like characterization of the real closed fields with a r.e. set of infinitesimals, and the dimension of r.e. sets.

Introduction

In the classical theory of Turing machines — which are the same as BSS machines over the field $\mathbb{F}_2$, see [Po] — it is a basic result that recursively enumerable subsets of $\mathbb{F}_2^\infty$ are either finite or isomorphic to $\mathbb{F}_2^\infty$ (see for example [Sh]). Consequently there is no notion of dimension[1] for Turing recursively enumerable sets. In this paper, we investigate the isomorphism problem for recursively enumerable sets over a real closed field in the sense of the BSS model of computation. When, in the sequel, we will use words as effective, machine, computable,... one has to understand effective in the sense of the BSS model of computation,..., except if it is explicitly stated otherwise. Apart the pioneering paper on the subject [BSS] good references for the BSS model of computability are the books [BCSS] and [Po], see also [MM] for a survey.

Structure of recursively enumerable sets over the field of real numbers (or a general real closed field) is well known [BSS] : they are countable Turing decidable union of semialgebraic sets defined over a finite set of parameters. So in some sense recursively enumerable sets over a real closed field mixes features of Turing decidable subsets of $\mathbb{N}$ and of semialgebraic sets. There exists a natural notion of dimension for semialgebraic sets (see [BCR]) which is invariant under semialgebraic isomorphism but, two semialgebraic sets of same dimension are not necessarily semialgebraically isomorphic. If it was clear that this notion of dimension will play a role in the search of an isomorphism theorem for recursively enumerable sets over a general real closed field,[2] it was not obvious to predict how this feature of semialgebraic sets interplays with the fact that recursively enumerable sets are in general countably infinite union of semialgebraic sets, with a priori no bound on the semialgebraic dimension of the sets involved in this union.
Another problem in the BSS model of computation (which is fully discussed in the extended version of this paper [MT, section 6]) is that maps with a recursively enumerable graph over a real closed fields are in general not computable (contrarily to the classical case). Such maps are called decidable maps. Typical examples are semialgebraic maps (as $x \mapsto \sqrt{x}$ on $\mathbb{R}$). Nevertheless it turns out that the behavior (for this question) of recursively enumerable sets over a real closed field is the best possible:[3] roughly we have

[1] In the usual sense of dimension, see [HW].
[2] Semialgebraic isomorphisms have recursively enumerable graph.
[3] At least when $\mathbb{Z}$ is BSS decidable over the field.

either

> the transcendence degree (over $\mathbb{Q}$) of the field is finite and there is no notion of dimension: all infinite recursively enumerable sets are isomorphic,

or

> the transcendence degree of the field is infinite and there exists a notion of recursively enumerable dimension (with values in $\mathbb{N} \cup \{\infty\}$) which coincides with the semialgebraic dimension for semialgebraic sets, and two recursively enumerable sets are isomorphic if and only if they have same dimension.

In particular all semialgebraic sets of same dimension are isomorphic as recursively enumerable sets in both cases.

The paper is organized as follows. In section 1, we recall various definitions and theorems and set the terminology that shall be used throughout the rest of the paper. Section 2 shows that semialgebraic sets can be effectively split into several parts, each of which being isomorphic to an open cube.[4] These cubes are glued together in section 3 to prove the isomorphism theorem for semialgebraic sets. This theorem is extended from semialgebraic sets to recursively enumerable sets in section 4. The rest of the paper is then dedicated to related questions. Section 5 deals with the characterization of real closed field in which either the infinitesimals or $\mathbb{Z}$ is decidable. Finally, in section 6, we discuss the notion of dimension for recursively enumerable sets.

1 Preliminaries

This section is simply a reminder of some facts about semialgebraic sets and BSS-machines. Probably the reader will be familiar with most (if not all) of them, but the intention is to make clear the terminology we shall use and to make this paper easily accessible both to people working in algebraic geometry and computability.

1.1 Semialgebraic sets and real closed fields

Our basic reference for this section is the book [BCR], of which will closely follow the notations. Let $(R, <)$ be an ordered ring, that is R is a commutative ring with identity and '$<$' is an ordering on R which is compatible with addition and multiplication. We will write (a, b) (resp. $[a, b]$) for the open (resp. closed) interval with endpoints a and b. A number $x \in R$ is infinitesimal if it satisfies $0 < |x| < 1/n$ for all $n \in \mathbb{N}^{\times}$. The equivalence relation '$x \approx y$' is defined by '$x - y$ is infinitesimal or 0.' An element $x \in R$ is said to be finite whenever $|x| \leqslant n$ for some $n \in \mathbb{N}$. Otherwise, it is called infinite. Let us denote $R_{\diamond}$ the

[4]This result is only a natural effective version of the so-called *cell decomposition theorem (or splitting theorem) for semialgebraic sets* (in French: *théorème de démontage*).

set of elements of R which are finite but not infinitesimal. When $N < M$, R^N will be embedded in R^M by setting the last $M - N$ components to zero. Let $f_1, \ldots, f_{p+q}$ be polynomials in $R[x_1, \ldots, x_N]$. A *basic semialgebraic set* is a set of the form:

$$\{x \in R^N : f_1(x) = 0, \ldots, f_p(x) = 0,\ f_{p+1}(x) > 0, \ldots, f_{p+q}(x) > 0\}.$$

Semialgebraic sets are all sets that can be build by finite union of basic semialgebraic sets. Note that finite union, finite intersection, and the complementary of semialgebraic sets are semialgebraic sets. A function $f : R^N \hookrightarrow R^M$ is called a *semialgebraic function* iff its graph, $\operatorname{Graph}(f) \subseteq R^N \times R^M$, is a semialgebraic set. The notation $f : A \hookrightarrow B$ is used to emphasize that the domain of f, $\operatorname{Dom} f$, may not be the whole set A. Writing $f : A \to B$ will mean $f : A \hookrightarrow B$ and $\operatorname{Dom} f = A$. Remark straight away that if $(S_i)_{i=1}^k$ are disjoint semialgebraic sets of R^N and $(f_i : S_i \hookrightarrow R^M)_{i=1}^k$ are semialgebraic functions, the map $\sum f_i : \bigcup S_i \hookrightarrow R^M$ is semialgebraic. A semialgebraic set $S \subseteq \mathbb{R}^N$ is called a *rectangle* (resp. an *integer cube*) iff $S = \prod_{i=1}^N S_i$ with $S_i = (a_i, b_i)$ (resp. $S_i = (a_i, a_i + 1)$) or $S_i = \{a_i\}$ for some $a_i, b_i \in R$ (resp. $a_i \in \mathbb{Z}$). A notion of dimension, $\dim S$, can be defined for any semialgebraic set S (cf. [BCR]). As usual, $\dim \varnothing = -1$. It is invariant under semialgebraic isomorphisms. Over the field of real numbers this dimension is equal to the euclidean dimension.

An *atomic formula* in the language of ordered rings is a formula of one of the following forms:

$$f(x) > 0 \quad \text{or} \quad f(x) = 0 \quad \text{or} \quad 0 > f(x)$$

where f is a polynomial in $R[x_1, \ldots, x_N]$ for some $N \in \mathbb{N}$. *Open formulae* in the language of ordered rings are (well formed) expressions made of conjunctions ($\wedge$), disjunctions ($\vee$), and negations ($\neg$) of atomic formulae. All the formulae we will speak about are in the language of ordered rings, so, from now on, we drop the precision. To stress that a formula P depends on the variable(s) x, we will write $P[x]$. From the very definition of semialgebraic sets, it is easy to see that semialgebraic sets in R^N are precisely those that can be written as

$$\{x \in R^N : P[x]\}$$

for an open formula P. We will say that a formula $P[x_1, \ldots, x_N]$ *represents a rectangle* (resp. *an integer cube*) if $P[x] = \bigwedge_{i=1}^N P_i[x_i]$ with each $P_i[x_i]$ being either $a_i < x_i < b_i$ (resp. $a_i < x_i < a_i + 1$) or $x_i = a_i$ for some $a_i, b_i \in R$ (resp. $a_i \in \mathbb{Z}$).

First order formulae are those that can be constructed from atomic formulae using the '$\wedge$', '$\vee$', '$\neg$' connectors and the quantifiers '$\exists$' and '$\forall$'. In general, first order formulae have greater expressive power than open ones. It turns out however that, in some ordered fields, they are in fact equivalent. These fields are known as the *real closed fields.* They are characterized by the following property: R is real closed if and only if R can be endowed with a unique ordering whose positive elements are the squares and such that every polynomial with odd degree has a root in R. For the sequel, the essential result about real closed fields is the following. It is known as 'elimination of quantifiers' or 'Tarski-Seidenberg'.

THEOREM 1. (TARSKI-SEIDENBERG) *In a real closed field, any first order formula is equivalent to an open formula.*

Proofs of this theorem can be found in [Ta] or [BCR], see also [vdD] for the history of this theorem. Actually Tarski-Seidenberg's theorem is another characterization of real closed fields because ordered rings that admit quantifier elimination are necessarily real closed (see e.g. [MMV]).

Tarski-Seidenberg's theorem has important consequences on the class of semialgebraic sets and semialgebraic functions. Indeed, it implies that the closure and the interior of semialgebraic sets are semialgebraic and that the set of semialgebraic functions from R^N to R^M is a vector space. Moreover, the domain and the range of semialgebraic functions are semialgebraic sets, and the composition of two semialgebraic functions is again a semialgebraic function.

1.2 Universal machines over rings

For the definition and basic properties of BSS machines over an ordered ring R, the reader is referred to the original article by Blum, Shub and Smale [BSS]. Since throughout this paper R will be a field, the computation nodes may be rational functions. Following [BSS], the set R^∞ is made of all sequences $(x_n)_{n\geqslant 1}$ with each $x_n \in R$ and such that all but a finite number of x_n's are null. To any $x \in R^\infty$ is associated $\mathrm{length}(x) := \max\{n : x_n \neq 0\}$ if $x \neq 0$ and $\mathrm{length}(x) := 0$ if $x = 0$. From now on, we shall identify R^N with $\{x \in R^\infty : \mathrm{length}(x) \leqslant N\}$. We want to stress that the machines are allowed to use the length of the state space—and therefore the length of any variable $x \in R^\infty$—as the content of any ordinary register.[5]

We will speak throughout this paper of machines inputting, outputting, and acting on other machines. This will mean that the machines will input, output, and act on a *coding* in R^∞ representing the other machines. Description of such a coding can be found in [BSS]. Associated with it is the *universal machine* that simulates the machine described by the coding. As a byproduct, we get the *universal polynomial evaluator* that takes the coding of a polynomial and a value for each of its variables and outputs the evaluation of the polynomial.

We will also consider machines manipulating first order (e.g., open) formulae. As above that means that the formulae are coded in some way in R^∞ (see [BSS,

[5] This is indeed possible because the length of an input x is stored in register $I(x)_4$ of input space (see [BSS]) and, consequently, the length of the data in the state space can be tracked along the computations (an upper bound is readily obtained and then a simple procedure can compute its actual value).

Po]). There exists a *universal open formula evaluator* that inputs (a coding of) an open formula[6] and a value for each of its variables and says whether or not the formula is satisfied by the values. There is an analogous for first order formulae, at least when R is a real closed field. Indeed, Tarski-Seidenberg is effective; that is, there exists a machine that transforms any first order formula into an equivalent open one (see [vdD]). As a result, the truth or falseness of any first order formula can be calculated by a BSS-machine. These facts shall be repeatedly used without necessarily explicit reference.

In what follows, we will write formulae with the language of symbolic logic and will usually omit to say that we are in fact talking about the codings of such formulae. Various operations will be performed on machines and formulae (composing machines, extracting the polynomials of a formula, constructing inductively formulae from other ones, ...). These operations will have to be carried out by machines but we will leave to the (patient) reader the task of designing the specific subroutines to achieve them on the corresponding codings.

2 Splitting semialgebraic sets is effective

The splitting theorem for semialgebraic sets says roughly that any semialgebraic set is semialgebraically isomorphic to a disjoint union of cubes $(0,1)^d$ with $d \in \mathbb{N}$ (see [BCR][7]). This section is driven by the slogan: *there exists an universal splitting machine for semialgebraic sets.* Every worker in the field of real algebraic geometry shall certainly be convinced that the previous statement holds.[8] For this reason we will not give here complete proofs of the main statements of this section. Proofs are included in a longer paper.

DEFINITION 2. A *coding of a semialgebraic set* $S \subseteq R^N$ is the coding of an open formula P such that $S = \{x \in R^N : P[x]\}$. A *coding of a semialgebraic function* $f : R^N \hookrightarrow R^M$ is a coding of its graph $\mathrm{Graph}(f) \subseteq R^N \times R^M$.

Notice that, in general, several formulae can describe a given semialgebraic set, that is codings are not unique. Given a coding of a semialgebraic set $S \subseteq R^N$, the universal formula evaluator can easily tell whether some $x \in R^N$ belongs or not to S. In particular, the universal formula evaluator can take a semialgebraic function $f : R^N \hookrightarrow R^M$ and a point $(x, y) \in R^N \times R^M$ as input and answer the question 'does y equal $f(x)$?'. Remark that this does not mean that $f(x)$ is computable. We will come back later to this question. Now let us state the main theorem of this section.

[6] Note that the set of open formulae is decidable in the set of first order formulae, i.e., a machine can decide whether a first order formula is actually an open one.

[7] A proof of this theorem already appeared in [Lo].

[8] It should be noted that, because we are not interested in the piecewise continuity of the isomorphism, we shall not take into consideration the derivatives of the polynomials defining the semialgebraic set.

THEOREM 3. (UNIVERSAL SPLITTING MACHINE) *Let R be a real closed field and $N \in \mathbb{N}$. There exists a machine $\mathcal{S}_{R^N}$ with the following properties:*

(i) *input: a coding of a semialgebraic set S of R^N;*

(ii) *output: a coding of a semialgebraic set T of R^N and a coding of a semialgebraic function $\varphi : R^N \hookrightarrow R^N$;*

(iii) *the formula coding T has the form $\bigvee_{\alpha\in A} P_\alpha[x]$ where the $P_\alpha[x]$'s are mutually exclusive formulae representing integer cubes;*

(iv) *the function $\varphi : S \to T$ is a bijection.*

Let us mention the following interesting corollary.

COROLLARY 4. *Let R be a real closed field and $N \in \mathbb{N}$. There exists a machine $\mathcal{D}$ that inputs (a coding of) a semialgebraic set of R^N and computes its dimension.*

PROOF. First recall that, if $S \subseteq R^N$ is a semialgebraic set and $\varphi : S \to R^N$ is a semialgebraic injective map, $\varphi(S)$ has the same dimension as S (see [BCR]). Therefore, using the above theorem, we may assume that the semialgebraic set S inputed to $\mathcal{D}$ is coded by a formula $\bigvee_\alpha P_\alpha[x]$ satisfying point (iii) above. The dimension of S is the maximum of the dimensions of the sets $\{x \in R^N : P_\alpha[x]\}$ (see [BCR] proposition 2.8.5), each of which being read on the formula P_α. $\square$

Let us mention that explicit bounds on the complexity of computing the dimension of a semialgebraic set have been recently given in [Ko] and [Vo].

Theorem 3 will follow from a recursion argument on the following lemma.

LEMMA 5. *Let R be a real closed field and $N \in \mathbb{N}$. There exists a machine, denoted S_N, which takes as input a semialgebraic set S of R^N and outputs a semialgebraic set T of R^N and a semialgebraic map $\varphi_N : R^N \hookrightarrow R^N$ such that:*

(i) *T is coded by a formula of the form $\bigvee_{\beta\in B} t_\beta[x]$ where the $t_\beta[x]$'s are mutually exclusive formulae which are equal either to $\tau_\beta[x_1, \ldots, x_{N-1}] \wedge (a_\beta < x_N < a_\beta + 1)$ or $\tau_\beta[x_1, \ldots, x_{N-1}] \wedge (x_N = a_\beta)$ for some satisfiable (in R) open formula τ_β and some $a_\beta \in \mathbb{N}$;*

(ii) *the map φ_N is a bijection from S onto T.*

Moreover, if the formula coding S has the form $\bigvee_{\alpha\in A} \sigma^1_\alpha[x_1, \ldots, x_\ell] \wedge \sigma^2_\alpha[x_{\ell+1}, \ldots, x_N]$ with the σ^1_α's representing integer cubes, then every τ_β can be assumed to have the form $\tau^1_\beta[x_1, \ldots, x_\ell] \wedge \tau^2_\beta[x_{\ell+1}, \ldots, x_{N-1}]$ with the τ^1_β's representing integer cubes.

REMARK 6. The fact that the previous formula coding S has the special form of the last statement of the preceding lemma *cannot*, in general, be checked by a BSS-machine. In fact, this can be done iff $\mathbb{Z}$ is decidable in R. $\square$

Proof of the lemma will not be given here : it is reminiscent of the proof of the main lemma which leads to the proof of the cell decomposition theorem for semialgebraic sets (see [BCR]). A proof is given in [MT].

PROOF OF THEOREM 3. Let $\mathcal{S}_n$ $(1 \leqslant n < N)$ be the machine that does the same thing as $\mathcal{S}_N$ but on the nth component instead of the Nth one. The machine $\mathcal{S}_{R^N}$ is the composition of the $\mathcal{S}_n$'s. More precisely

$$\begin{aligned} (T_N, \varphi_N) &:= \mathcal{S}_N(S) \\ (T_n, \varphi_n) &:= \mathcal{S}_n(T_{n+1}) \qquad\qquad (1 \leqslant n < N) \\ \mathcal{S}_{R^N}(S) &:= (T_1, \varphi_1 \circ \cdots \circ \varphi_N) \end{aligned}$$

The formula coding the set T_1 has the required form. Indeed, if the formula coding T_{n+1} has the form $\bigvee_{\alpha \in A_{n+1}} \sigma^1_{\alpha,n+1}[x_1, \ldots, x_n] \wedge \sigma^2_{\alpha,n+1}[x_{n+1}, \ldots, x_N]$ with the $(\sigma^2_{\alpha,n+1} : \alpha \in A_{n+1})$ representing integer cubes, then, by virtue of lemma 5 (because, of course, we can specify other variables than $x_1, \ldots, x_\ell$), the set T_n is coded by an analog formula for n. A recursion argument completes the proof. □

REMARK 7. A close look at the complete proofs shows that $\mathcal{S}_n$ depends computably on n and therefore $\mathcal{S}_{R^N}$ depends also computably on N. In other words, the map $N \mapsto \mathcal{S}_{R^N}$ is computable. □

3 Decidable isomorphisms

The aim of this section is to show how the disjoint cubes in which semialgebraic sets are split can be glued together. The trouble in doing so is that, for example, $(0,1) \cup (1,2)$ and $(0,1)$ are not semialgebraically isomorphic when $R = \mathbb{R}$. We will need the following more general kind of isomorphism.

DEFINITION 8. A function $\varphi : R^N \circ\!\!\rightarrow R^M$ is said to be *decidable* iff its graph is decidable. Two sets A and B are said to be *decidably isomorphic* whenever there exists a bijective map $\varphi : A \to B$ which is decidable.

For that larger class of isomorphisms, any semialgebraic set is either a finite number of points or isomorphic to some cube $(0,1)^d$ for some $d \in \mathbb{N}$. We start with the following basic gluing result.

PROPOSITION 9. *Let R be a real closed field and $N \in \mathbb{N}$. There exists a machine $\mathcal{G}_N$ satisfying the following:*

(i) *input: a pair (S_1, S_2) of disjoint semialgebraic sets of R^N which are coded by formulae representing rectangles;*

(ii) *output: a semialgebraic set T of R^N and a map $\varphi : R^N \circ\!\!\rightarrow R^N$;*

(iii) $\dim T = \max\{\dim S_1, \dim S_2\}$ *and, if* $\dim T > 0$, *the formula coding* T *represents a rectangle;*

(iv) $\varphi : S_1 \cup S_2 \to T$ *is a bijection;*

(v) *if* $\mathbb{Z}$ *is decidable in* R, $\varphi : S_1 \cup S_2 \to T$ *and* $\varphi^{-1} : T \to S_1 \cup S_2$ *are computable.*

Moreover, if $\dim S_1 < \dim S_2$, *one can assume* $\varphi\lceil S_1 = \mathrm{id}$. *On the other hand, if* $\dim S_1 \geqslant \dim S_2$ *and if* L *is a semialgebraic set coded by a formula representing a rectangle such that* $\dim(S_1 \cap L) \geqslant \max\{1, \dim S_2\}$, *then one may assume* $T = S_1$ *and* $\varphi\lceil(S_1 \setminus L) = \mathrm{id}$, $\varphi\lceil(S_1 \cap L) \subseteq L$ *(this last property shows the local character of* φ *and will play a large role in the proof of theorem 14).*

Remark 10. As the proof will show, the map $N \mapsto \mathcal{G}_N$ is computable. □

Proof. Let $d_i := \dim S_i$ $(i = 1, 2)$. The dimension being computable, the machine will be able to take the appropriate action in each case below. We ought only to consider the case $d_1 > 0$ or $d_2 > 0$—otherwise nothing has to be done. It is no lack of generality[9] to suppose

$$S_1 = (-\tfrac{1}{2}, \tfrac{1}{2})^{d_1} \times \{0\}^{N-d_1} \qquad \text{and} \qquad S_2 = (\tfrac{1}{2}, \tfrac{3}{2})^{d_2} \times \{1\}^{N-d_2}.$$

If $d_1 < d_2$, let us write $x \in R^N$ as $(\overline{x}^1, x', \tilde{x})$ with $\overline{x}^1 \in R^{d_1}$, $x' = x_{d_1+1} \in R$, $\tilde{x} \in R^{N-d_1-1}$, and define $\varphi : S_1 \cup S_2 \to (-\frac{1}{2}, \frac{1}{2})^{d_2} \times \{0\}^{N-d_2}$ by

$$\varphi(x) := \begin{cases} x & \text{if } x \in S_1; \\ (\overline{x}^1 - e^1, -1/2 + 1/2^{n+1}, 0) & \text{if } x \in S_2,\ x' = \frac{1}{2} + \frac{1}{2^n}\ (n \geqslant 1),\ \tilde{x} = \tilde{e}; \\ (\overline{x}^1 - e^1, x' - 1, \tilde{x} - \tilde{e}) & \text{otherwise;} \end{cases}$$

where $e^1 := (1, \ldots, 1) \in R^{d_1}$ and $\tilde{e} := (1, \ldots, 1) \in R^{N-d_1-1}$. If $\mathbb{Z}$ is decidable in R, so is $\{1/2 + 1/2^n : n \geqslant 1\}$ and therefore φ is computable (note that if $x' = 1/2 + 1/2^n$, $-1/2 + 1/2^{n+1} = (2x' - 3)/4$).

If $d_1 > d_2 = 0$, we set $x = (x_1, \tilde{x}) \in R \times R^{N-1}$ and define $\varphi : S_1 \cup S_2 \to S_1$ as

$$\varphi(x) := \begin{cases} (0, 0) & \text{if } x \in S_2; \\ (-1/2 + 1/2^{n+1}, 0) & \text{if } x \in S_1,\ x_1 = -\frac{1}{2} + \frac{1}{2^n}\ (n \geqslant 1), \text{ and } \tilde{x} = 0; \\ x & \text{otherwise.} \end{cases}$$

If $d_1 \geqslant d_2 \geqslant 1$, let us write $x \in R^N$ as $(\overline{x}^2, x', \tilde{x})$ with $\overline{x}^2 \in R^{d_2-1}$, $x' = x_{d_2} \in R$, $\tilde{x} \in R^{N-d_2}$. We first apply to S_2 the computable isomorphism $\varphi_2 : S_2 \to S_2'$ defined by

$$\varphi_2(x) := \begin{cases} (\overline{x}^2 - e^2, 1/2, 0) & \text{if } x' = 1; \\ (\overline{x}^2 - e^2, 1/2 + 1/2^{n-1}, 0) & \text{if } x' = 1/2 + 1/2^n \text{ with } n \geqslant 2; \\ (\overline{x}^2 - e^2, x', 0) & \text{otherwise;} \end{cases}$$

[9] Translations and Rotations are semialgebraic functions.

where $e^2 := (1, \ldots, 1) \in R^{d_2-1}$ and $S_2' := (-\frac{1}{2}, \frac{1}{2})^{d_2-1} \times [\frac{1}{2}, \frac{3}{2}) \times \{0\}^{N-d_2}$. Next let us consider the piecewise polynomial isomorphism $\varphi_1 : S_1 \cup S_2' \to S_1$ whose definition is

$$\varphi_1(x) := \begin{cases} \left(\overline{x}^2, (2x'-1)/4, 0\right) & \text{if } x \in S_1 \text{ and } \tilde{x} = 0, \text{ or } x \in S_2'; \\ x & \text{otherwise.} \end{cases}$$

The required map φ is $\varphi_1 \circ \varphi_2$.

Let us now prove the "local" property. Since $\dim(S_1 \cap L) \geqslant \max\{1, \dim S_2\}$, there exists a semialgebraic set L' coded by a formula representing a rectangle which has at least the dimension $\max\{1, \dim S_2\}$ and is included in $S_1 \cap L$. The rectangle L' can obviously be computed by a BSS-machine for all there is to do is to solve one-variable affine inequalities. Let $\varphi' : L' \cup S_2 \to L'$ be the computable isomorphism constructed in the appropriate case above—using L' in place of S_1. Then $\varphi : S_1 \cup S_2 \to S_1$ defined by

$$\varphi(x) := \begin{cases} \varphi'(x) & \text{if } x \in L' \cup S_2; \\ x & \text{otherwise;} \end{cases}$$

possesses the desired properties. □

THEOREM 11. (ISOMORPHISM THEOREM) *Let R be a real closed field in which $\mathbb{Z}$ is decidable, and $N \in \mathbb{N}$. There exists a machine $\mathcal{G}_{R^N}$ that inputs a semialgebraic set S of R^N and outputs a semialgebraic set T of the same dimension as S and a decidable isomorphism $\varphi : S \to T$ such that T is either a finite number of points or is coded by a formula representing an integer cube. Moreover $N \mapsto \mathcal{G}_{R^N}$ is computable.*

PROOF. First use theorem 3 to split S as a disjoint union of integer cubes coded by $\bigvee_{\alpha \in A} P_\alpha[x]$; and then apply proposition 9 $\mathrm{card}(A) - 1$ times to glue them together. Of course, $(-\frac{1}{2}, \frac{1}{2})^d \times \{0\}^{N-d} \cong (0,1)^d \times \{0\}^{N-d}$.

The fact that $N \mapsto \mathcal{G}_{R^N}$ is computable results from the computability of $N \mapsto \mathcal{S}_{R^N}$ and $N \mapsto \mathcal{G}_N$. □

REMARK 12. Note that the map $\varphi = \varphi^{\mathrm{glue}} \circ \varphi^{\mathrm{split}}$ where φ^{split} is a semialgebraic map (given by theorem 3) and φ^{glue} is a computable isomorphism (given by the successive applications of proposition 9). □

4 Recursively enumerable sets

We now turn our attention to recursively enumerable sets in R^N with $N \in \mathbb{N}$ or $N = +\infty$. To input a recursively enumerable set S to a machine will mean to feed it with a coding of a machine whose halting set is S. The link with semialgebraic sets is the following:

PROPOSITION 13. *Let R be a ring, $N \in \mathbb{N} \cup \{\infty\}$, and S be a recursively enumerable set in R^N. Then there exists a decidable (in the classical sense) subset Γ of $\mathbb{N}$ such that*

$$S = \bigcup_{\gamma \in \Gamma} S_\gamma$$

for some disjoint finite dimensional semialgebraic sets S_γ of R^N.

Moreover, there exists a machine that, given any such S, can compute the corresponding Γ and the function $\gamma \mapsto S_\gamma$.

PROOF. See [BSS], §4, proposition 2. Let Γ' be the set of all computation paths of a machine M whose halting set is S, and $S'_{\gamma'}$ the set of inputs that reach an output node by following the path γ'. When $N < \infty$, let us consider $\Gamma \subseteq \mathbb{N}$ computably isomorphic to Γ' and $S_\gamma := S'_\gamma$. Therefore, it is clear that Γ and $\gamma \mapsto S_\gamma$ are computable once the machine M is known. When $N = \infty$, $x \in S'_{\gamma'}$ iff a computable boolean condition on expressions of the form $f\big(\mathrm{length}(x), x\big) \gtreqless 0$ holds, where $f : R^\infty \to R^\infty$ are polynomials. We set $\Gamma :\cong \Gamma' \times \mathbb{N}$ and $S_{(\gamma',N)} := \{x \in S'_{\gamma'} : \mathrm{length}(x) = N\} = S'_{\gamma'} \cap (R^N \setminus R^{N-1}) = S'_{\gamma'} \cap \{x \in R^\infty : \mathrm{length}(x) \leqslant N \ \wedge\ x_N \neq 0\}$. The sets $S_{(\gamma',N)}$ are clearly disjoint and are finite dimensional semialgebraic subsets of R^∞ because they are subsets of R^N which are defined by the same inequations as $S'_{\gamma'}$ (where the length is now fixed) plus $x_N \neq 0$. $\square$

We are now in position to prove the main theorem of this paper.

THEOREM 14. *Let R be a real closed field and $N \in \mathbb{N} \cup \{\infty\}$. There exists a machine with the following properties:*

(i) *input: a recursively enumerable set S of R^N for some $N \in \mathbb{N} \cup \{\infty\}$;*

(ii) *output: a recursively enumerable set T of R^N and a function $\varphi : R^N \circ\!\!\rightarrow R^N$;*

(iii) *T is either an at most countable set of points or R^d for some $0 < d \leqslant N$;*

(iv) *the map $\varphi : S \to T$ is a bijection;*

(v) *if $x \approx 0$ is recursively enumerable in R, φ is decidable over S.*

REMARK 15. 1) Recall that 'φ is decidable over S' is equivalent to 'Graph φ is decidable over $S \times R^N$' which means that there exists a machine that inputs $(x, y) \in R^N \times R^N$ and, at least when $(x, y) \in S \times R^N$, halts and says whether or not $(x, y) \in \mathrm{Graph}\,\varphi$.

2) If $x \approx 0$ is recursively enumerable, $x \approx 0$ is decidable (because $x \not\approx 0$ means $\exists k,\ |x| > 1/k$ and so is recursively enumerable) and then $\mathbb{Z}$ is decidable in R. We shall show in section 5 that in fact the converse is true as well. $\square$

PROOF. Because of proposition 13, we may assume that we have at our disposal $\Gamma \subseteq \mathbb{N}$ and $\gamma \mapsto S_\gamma$ such that $S = \bigcup_{\gamma \in \Gamma} S_\gamma$. Without lack of generality, we may also assume that $\Gamma = \mathbb{N}$, and $x \approx 0$ is recursively enumerable. We will argue by induction over γ. The cases $N < \infty$ and $N = \infty$ are treated together.

We will construct by induction a sequence of computable semialgebraic subsets T_γ of R^N and maps $\varphi_\gamma : R^N \hookrightarrow R^N$ such that

(i) $T_\gamma = (-\frac{1}{2}, \frac{1}{2})^{d(\gamma)}$ provided $d(\gamma) > 0$, where $d(\gamma) := \max\{\dim S_\mu : \mu \leqslant \gamma\}$;

(ii) if $\gamma = 0$, $\varphi_0 : S_0 \to T_0$ is a decidable bijection and, if $\gamma \geqslant 1$, we may assume S_γ is disjoint of $T_{\gamma-1}$ and $\varphi_\gamma : T_{\gamma-1} \cup S_\gamma \to T_\gamma$ is a decidable bijection;

(iii) when $\gamma \geqslant 1$, two things can happen: either $\dim S_\gamma > \dim T_{\gamma-1}$, and $\varphi_\gamma \restriction T_{\gamma-1} = \mathrm{id}$; or $\dim S_\gamma \leqslant \dim T_{\gamma-1}$, and then, provided $\dim T_{\gamma-1} > 0$, $\varphi_\gamma \restriction (T_{\gamma-1} \setminus L_\gamma) = \mathrm{id}$ and $\varphi_\gamma(T_{\gamma-1} \cap L_\gamma) \subseteq L_\gamma$ with $L_\gamma := \big(1/(\gamma+2), 1/(\gamma+1)\big)^{d(\gamma-1)}$.

This sequence is easy to construct. First, if $\gamma = 0$, an application of theorem 11 gives the representation of an integer cube T_0 — that up to translation and rotation may be assumed to be $(-\frac{1}{2}, \frac{1}{2})^{d(0)}$ with $d(0) = \dim T_0$ — and an isomorphism $\varphi_0 : S_0 \to T_0$. Now let us suppose (i)–(iii) hold for $\gamma - 1$ and let us show they still hold for γ. By theorem 11 we may assume that, up to an isomorphism, S_γ is coded by a formula representing an integer cube and, up to a translation, is disjoint of $T_{\gamma-1}$. Then we apply proposition 9 to $(T_{\gamma-1}, S_\gamma)$ with $L := L_\gamma$ and we get T_γ — which has the form $(-\frac{1}{2}, \frac{1}{2})^{d(\gamma)} \times \{0\}^{N-d(\gamma)}$ — and an isomorphism $\varphi_\gamma : T_{\gamma-1} \cup S_\gamma \to T_\gamma$. Clearly properties (i)–(iii) are satisfied.

Define $\varphi : S \to (-\frac{1}{2}, \frac{1}{2})^d$, where $d := \max\{\dim S_\gamma : \gamma \in \mathbb{N}\}$, by

$$\varphi(x) := \cdots \circ \varphi_{\gamma+2} \circ \varphi_{\gamma+1} \circ \varphi_\gamma(x) \quad \text{if } x \in S_\gamma \text{ for some } \gamma.$$

The map φ is well defined. Let $x \in R^N$. If $x \in S$, we can compute the unique γ such that $x \in S_\gamma$. Two cases can happen. First, one component of $\varphi_\gamma(x)$ is infinitesimal. Then, in view of (iii), none of the φ_μ, $\mu > \gamma$, will modify $\varphi_\gamma(x)$. Thus $\varphi(x) = \varphi_\gamma(x)$. Second, all components of $\varphi_\gamma(x)$ are greater than, say, $1/k$ for some computable $k \in \mathbb{N}^\times$—which can be supposed to be greater than γ. Then (iii) implies that $\varphi_k \circ \cdots \circ \varphi_\gamma(x)$ will be left invariant by all φ_μ, $\mu > k$. Indeed, either some S_μ, $\gamma < \mu \leqslant k$, has a greater dimension than T_γ in which case this is clear because $T_\gamma = T_{\mu-1} \cap R^{d(\gamma)}$ shall never be touched from φ_μ on (remark that no L_ν, $\nu > \mu$, intersects T_γ), or all S_μ, $\gamma \leqslant \mu \leqslant k$, have the same dimension and then $\varphi_\gamma(x)$ can only be moved provided that $\varphi_\gamma(x) \in L_\mu$ in which case $\varphi_\mu(\varphi_\gamma(x)) \in L_\mu$ and that precludes any further φ_ν to act on it. Thus, we have shown $\varphi(x) = \varphi_k \circ \cdots \circ \varphi_\gamma(x)$. The relation $y = \varphi(x)$ is decidable because it can be written as $y = \varphi_k^{\mathrm{glue}} \circ \cdots \circ \varphi_\gamma^{\mathrm{glue}} \circ \varphi_\gamma^{\mathrm{split}}(x)$ with $\varphi_\gamma^{\mathrm{split}}$ being semialgebraic and all $\varphi_\mu^{\mathrm{glue}}$ being computable isomorphisms (see remark 12). Consequently, if we can decide whether or not some component of $\varphi_\gamma(x)$ is ≈ 0, the map φ is decidable over S. This concludes the proof—because $(-\frac{1}{2}, \frac{1}{2})^d \cong R^d$ computably. $\square$

5 On real closed fields with a recursively enumerable set of infinitesimals

In view of proposition 9 and theorem 14, two natural questions raise themselves: is it possible to give a characterization of the real closed fields in which $\mathbb{Z}$ (resp. $x \approx 0$) is decidable? It turns out that both questions are equivalent. They are settled in theorem 24. As a consequence of this theorem, we show that the properties that $\mathbb{Z}$ is or isn't decidable are not "stable" through extensions—see proposition 25. The various results leading to theorem 24 may already be known (see a remark in [Wi, page 1087]). As we could not find a suitable reference however, we gave our own proofs. Theorem 24 is new. Throughout this section, we shall say 'infinitely large' for 'positive and infinite'. To start with, note that the following statements are equivalent (see also remark 15):

(i) the set of infinitesimals is recursively enumerable;

(ii) the set of infinitely large numbers is recursively enumerable;

(iii) the set of infinitesimals is decidable;

(iv) the set of infinitely large numbers is decidable.

For the shake of convenience, we shall here consider formulation (ii). We shall also make an extensive use of the following notion.

Definition 16. For any $x, y \in R$ with y infinitely large, the notation $x \ll y$ will stand for $|x|^k < y$ for all positive integer k.

Remark 17. Of course, if R is real closed, this is equivalent to $|x| < y^{1/k}$ which may be more intuitive. $\square$

Proposition 18. *Let $x, y, z, w > 0$. The following properties hold.*

(i) *the relations $\ll$ and $\not\ll$ are reflexive and transitive;*

(ii) $w \leqslant x \ll y \leqslant z$ *implies* $w \ll z$.

If x and y are infinitely large,

(iii) $x \ll y$ *implies* $P(x) \ll y$ *for any polynomial P whose coefficients are finite.*

In particular, if x and z/y are infinitely large, then

(iv) $x \ll z/y$ *implies* $r\,x^k\,y \leqslant z$ *for any finite $r \in R$ and $k \in \mathbb{N}$.*

Proof. These properties readily follow from the definition of '$\ll$'. $\square$

Lemma 19. *Let R be an ordered field, $x = (x_1, \ldots, x_m) \in R^m$, $C_\alpha, D_\alpha \in R$ for all multi-indices $\alpha \in \mathbb{N}^m$ satisfying $0 \leqslant |\alpha| \leqslant d$. Let β be the largest multi-index (for the lexicographic order) such that $D_\beta = \max\{D_\alpha : 0 \leqslant |\alpha| \leqslant d\}$. Assume*

- *$x_1, \dots, x_m$ are infinitely large numbers of R such that $x_1 \gg \cdots \gg x_m$;*
- *$C_\beta \not\approx 0$ and the other C_α's are finite ($\alpha \neq \beta$);*
- *$D_\alpha > 0$ for all α;*
- *$x_1 \ll D_\beta / D_\alpha$ (and the right hand side is infinitely large) whenever $D_\alpha < D_\beta$.*

Then $\sum_{0 \leqslant |\alpha| \leqslant d} C_\alpha \, D_\alpha \, x^\alpha \not\approx 0$ (e.g., $\neq 0$).

Proof. Without lack of generality, we may assume $C_\beta > 0$. The polynomial splits in two parts:

$$\sum_{D_\alpha = D_\beta} C_\alpha \, D_\alpha \, x^\alpha + \sum_{D_\alpha < D_\beta} C_\alpha \, D_\alpha \, x^\alpha .$$

When $D_\alpha = D_\beta$, necessarily $\alpha \leqslant \beta$, and so, if $\alpha \neq \beta$, $x^{\alpha - \beta} \approx 0$. Indeed, let $k \in \{1, \dots, m\}$ be the first value such that $\alpha_k \neq \beta_k$. Then $\alpha_k < \beta_k$ and $x^{\alpha-\beta} = x_k^{\alpha_k - \beta_k} \prod_{i>k} x_i^{\alpha_i - \beta_i} \leqslant x_k^{-1} x_{k+1}^{\ell}$ where $\ell := \sum_{i>k} \max\{\alpha_i - \beta_i, 0\}$. If $k = m$, $\ell = 0$ and we are done. If not, since $x_{k+1} \ll x_k$, one deduces $x_{k+1}^{\ell+1} \leqslant x_k$ and therefore $x^{\alpha - \beta} \leqslant x_{k+1}^{-1} \approx 0$ which is the claim.

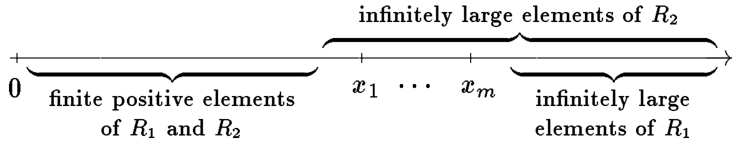
Consequently, the left hand side can be written $D_\beta x^\beta (C_\beta + \text{infinitesimals})$. This is greater or equal to $D_\beta x^\beta (C_\beta - \varepsilon) =: C D_\beta x^\beta$ for some small $\varepsilon \in \mathbb{Q}$—so that $C := C_\beta - \varepsilon \not\approx 0$. On the other hand, since $x_1 \ll D_\beta / D_\alpha$ when $D_\alpha < D_\beta$, proposition 18 (iv) implies that $|C_\alpha \, D_\alpha \, x^\alpha| \leqslant |C_\alpha| \, D_\alpha \, x_1^{|\alpha|} \leqslant x_1^{-1} C D_\beta$. Thus

$$\sum C_\alpha D_\alpha x^\alpha \geqslant C D_\beta \, x^\beta - (N/x_1) C D_\beta = (x^\beta - N/x_1) C D_\beta \not\approx 0$$

where N is the number of terms in the right hand side. The fact $N/x_1 \approx 0$ concludes the proof. □

Theorem 20. *Let R_1, R_2 be two ordered fields with $R_1 \hookrightarrow R_2$ (ordered field morphism). If $x_1 \ll \cdots \ll x_m$ are infinitely large elements of R_2 such that $x_m \ll r$ for any infinitely large $r \in R_1$, then $x_1, \dots, x_m$ are algebraically independent on R_1.*

The situation can be pictured as follows:



Proof. Let us suppose on the contrary that $x_1, \dots, x_m$ are algebraically dependent on R_1; that is there exists a polynomial P whose coefficients belong to R_1 such that

$$0 = P(x_m, \dots, x_1) =: \sum_{0 \leqslant |\alpha| \leqslant d} p_\alpha \, x^\alpha \tag{1}$$

where $x := (x_m, \ldots, x_1)$ and $p_\alpha \in R_1$. We can assume none of the p_α's is infinitesimal—otherwise multiply (1) by p_α^{-1}. Moreover, by virtue of lemma 19, not all p_α's can be finite:—consider $C_\alpha := p_\alpha$, $D_\alpha := 1$. So, if γ is a multi-index such that $|p_\gamma| = \max\{|p_\alpha| : 0 \leqslant |\alpha| \leqslant d\}$, at least p_γ is infinite and may be assumed to be positive. For each α, let us set $C_\alpha := p_\alpha$, $D_\alpha := 1$ if p_α is finite; $C_\alpha := \mathrm{sign}(p_\alpha)$, $D_\alpha := |p_\alpha|$ if both p_α and p_γ/p_α are infinite; $C_\alpha := p_\alpha/p_\gamma$, $D_\alpha := p_\gamma$ if p_α is infinite but p_γ/p_α is finite. Noting that $D_\alpha \in R_1$ for all α, it is easy to check that the assumptions of lemma 19 are satisfied—with β possibly greater than γ but nonetheless verifying $D_\beta = p_\gamma$. Then, lemma 19 asserts that (1) cannot hold and the proof is complete. □

Let us now draw some interesting corollaries.

Corollary 21. (Algebraic extensions) *Let R be an ordered field and $\tilde{R}$ its real closure. For any infinitely large element $b \in \tilde{R}$, there exists an infinitely large $a \in R$ and $k \in \mathbb{N}$ such that $b^k \geqslant a$.*

Proof. If the claim does not hold, one has $b \ll a$ for all infinitely large $a \in R$. This is absurd in view of theorem 20 (take $R_1 := R \hookrightarrow R_2 := \tilde{R}$, $m := 1$, and $x_1 := b$). □

Corollary 22. (Extensions of archimedean fields) *Let R_2 be an ordered extension of an archimedean field R_1 (e.g., $R_1 = \mathbb{Q}$). Every sequence $x_1 \ll \cdots \ll x_m$ of infinitely large elements of R_2 is algebraically independent on R_1.*

Proof. Obvious. □

Corollary 23. (Extensions with finite transcendence degree) *Let R be an ordered extension with finite transcendence degree of $\mathbb{Q}$. Then either R is archimedean or there exists an infinitely large element $a \in R$ such that, for any infinitely large $b \in R$, there is some $k \in \mathbb{N}$ such that $b^k \geqslant a$.*

Proof. If the conclusion of the corollary does not hold, R is not archimedean and one can find a sequence of infinitely large numbers $(x_i)_{i \geqslant 1}$ such that $\cdots \ll x_3 \ll x_2 \ll x_1$. According to corollary 22, $x_1, \ldots, x_m$ must be algebraically independent no matter what $m \in \mathbb{N}$ is. This contradicts the fact that the transcendence degree is finite. □

After these algebraic preliminaries, let us turn to the main theorem of this section.

Theorem 24. *Let R be a real closed field. The following three statements are equivalent.*

(i) *The set of infinitely large numbers is recursively enumerable.*

(ii) $\mathbb{Z}$ *is decidable in* R.

(iii) *Either* R *is archimedean or there exists an infinitely large number* $a \in R$ *such that any infinitely large* b *is greater or equal to* $a^{1/k}$ *for some* $k \in \mathbb{N}^{\times}$.

PROOF. We may of course assume R is non-archimedean.

(i) $\Rightarrow$ (ii). It is clear that $\mathbb{Z}$ is recursively enumerable. The complement of $\mathbb{Z}$ is also recursively enumerable because $x \in R \setminus \mathbb{Z}$ is equivalent to: either $|x|$ is infinitely large or $0 < |x| - n < 1$ for some $n \in \mathbb{N}$.

(iii) $\Rightarrow$ (i) is easy. Indeed it suffices to run the machine that compares an input x successively to a, $a^{1/2}$, $a^{1/3}, \dots$ (i.e., that checks whether $x - a \geqslant 0$, $x^2 - a \geqslant 0$, $x^3 - a \geqslant 0, \dots$). By hypothesis, this machine will stop iff x is infinitely large.

(ii) $\Rightarrow$ (iii). Since $R_+ \setminus \mathbb{N}$ is recursively enumerable, proposition 13 says that it can be written as an at most countable disjoint union of semialgebraic sets of R_+:

$$R_+ \setminus \mathbb{N} = \bigcup\nolimits_{\gamma \in \Gamma} S_\gamma.$$

Let $R_{<\infty}$ (resp. $R_{\geqslant\infty}$) denote the set of finite (resp. infinitely large) numbers of R and Γ' the set of $\gamma \in \Gamma$ such that $S_\gamma \cap R_{\geqslant\infty} \neq \varnothing$. Each S_γ is a finite disjoint union of intervals of R_+ (see [BCR]): $S_\gamma = \bigcup_i S_{\gamma i}$. Since $S_{\gamma i} \cap \mathbb{N} = \varnothing$, one necessarily has that either $S_{\gamma i} \subseteq R_{<\infty}$ or $S_{\gamma i} \subseteq R_{\geqslant\infty}$. Therefore $S_\gamma \cap R_{\geqslant\infty}$ consists of finitely many intervals and so we can speak of the infinitely large number $b_\gamma := \inf(S_\gamma \cap R_{\geqslant\infty})$ whenever $\gamma \in \Gamma'$. It is readily checked that

$$R_{\geqslant\infty} = \bigcup\nolimits_{\gamma \in \Gamma'} [b_\gamma, \rightarrow[. \tag{2}$$

Let $a_1, \dots, a_n$ be the constants of the machine whose halting set is $R_+ \setminus \mathbb{N}$. Each S_γ is described by an open formula with these parameters. Since b_γ is the lower endpoint of some interval $S_{\gamma i}$, it must solve a polynomial equation with parameters $a_1, \dots, a_n$; that is b_γ is algebraic on $\mathbb{Q}(a_1, \dots, a_n)$. Let a be given by corollary 23. This corollary implies there exists some $k \in \mathbb{N}$ such that $b_\gamma^k \geqslant a$—because the real closure of $\mathbb{Q}(a_1, \dots, a_n)$ has finite transcendence degree over $\mathbb{Q}$. But then (2) shows that, for any infinitely large number, such a k also exists. This concludes the proof. $\square$

PROPOSITION 25. *Let* R *be an ordered ring. There exist* R_1, R_2 *two real closed fields such that* $R \hookrightarrow R_1 \hookrightarrow R_2$ *and* R_1 *satisfies the equivalent properties of theorem 24 whereas* R_2 *does not.*

PROOF. Let us order the field $R(a)$ by imposing that a is infinitely large and $a \ll b$ for any infinitely large $b \in R$. This defines an ordered field by the compactness theorem for first-order languages because this extension is characterized by adding to the theory of R the axioms '$a > n$' for all $n \in \mathbb{N}$ and

'$a^k < b$' for all $k \in \mathbb{N}$ and all infinitely large $b \in R$. Let R_1 be the real closure of $R(a)$. Corollary 21 shows that R_1 will satisfy property (iii) of theorem 24 iff $R(a)$ does. We claim that is for any infinitely large $b \in R(a)$, one can find some $k \in \mathbb{N}$ such that $b^k \geqslant a$. If not, there must be an infinitely large $b \in R(a)$ with $b \ll a$. Theorem 20 implies that a, b are independent on R. That contradicts the fact that the transcendence degree of $R(a)$ on R is less or equal to 1.

Now, let us order $R_1(a_i : i \in \mathbb{N})$ by asking that all a_i are infinitely large and $\cdots \ll a_2 \ll a_1 \ll a_0 \ll b$ for any infinitely large $b \in R_1$. As before, these constraints define an ordered field. Let R_2 be the real closure of $R_1(a_i : i \in \mathbb{N})$. Again, by virtue of corollary 21, R_2 will satisfy property (iii) of theorem 24 iff $R_1(a_i : i \in \mathbb{N})$ does. But, if the latter holds, one can find some $a_* \in R_1(a_i : i \in \mathbb{N})$ such that any infinitely large b satisfies $b^k \geqslant a_*$ for some k. Thus $a_* \ll a_m \ll \cdots \ll a_0$ for all m—because $a_* \leqslant a_{m+1}^k \ll a_m$—and so theorem 20 implies that $a_*, a_m, \ldots, a_0$ are algebraically independent on R_1 for all m. On the other hand, a_* is the quotient of two polynomials depending only on finitely many a_i's; that is there is a m such that $a_* \in R_1(a_i : 0 \leqslant i \leqslant m)$. But then $a_*, a_m, \ldots, a_0$ cannot be independent on R_1. $\square$

REMARK 26. The above proposition shows that a field with the property that the set of its infinitely large numbers is recursively enumerable can loose it through an extension and that, conversely, a field that does not have that property can gain it by passing to a suitable extension. $\square$

6 Dimension of r.e. sets

In real closed fields, there is a well-known notion of dimension for semialgebraic sets. This notion is invariant under semialgebraic isomorphisms or, more precisely, if S is a semialgebraic set and φ is a semialgebraic map which is one-to-one on S, then $\dim \varphi(S) = \dim S$. The aim of this section is to construct a notion of dimension for recursively enumerable sets. The maps under which it should be invariant are the one-to-one decidable maps over the recursively enumerable set in question. Due to the fact that this class is larger than the one of semialgebraic maps, this notion of dimension shall be nontrivial only on real closed fields with an infinite transcendence degree over $\mathbb{Q}$. For these fields, the notions of dimension for recursively enumerable and semialgebraic sets will coincide. Let us start with drawing some consequences of $\mathrm{tr}_{\mathbb{Q}} R < \infty$.

THEOREM 27. *Let R be a real closed field. The following statements are equivalent.*

(i) *The transcendence degree of R over $\mathbb{Q}$, $\mathrm{tr}_{\mathbb{Q}} R$, is finite.*

(ii) *There exists a computable one-to-one map $\varepsilon : R \to R$ whose range $\mathrm{Im}\, \varepsilon = \mathbb{N}^\times$.*

(iii) *For any semialgebraic set $S \subseteq R^N$, $1 \leqslant \dim S \leqslant N < \infty$, one can find* *a sequence* $(S_i)_{i\in\mathbb{N}}$ *of disjoint semialgebraic sets of dimension less than* $\dim S$ *such that* $S = \bigcup_{i\in\mathbb{N}} S_i$ *and* $i \mapsto S_i$ *is computable.*

(iv) *There exists some semialgebraic set* $S \subseteq R^N$, $1 \leqslant \dim S \leqslant N < \infty$, *which can be covered by a countable sequence* $(S_i)_{i\in\mathbb{N}}$ *of semialgebraic sets of dimension less than* $\dim S$ *and such that* $i \mapsto S_i$ *is computable.*

Before going into the proof, let us introduce some notation. If $P \in R[X]$, $\mathcal{P}$ is a finite subset of $R[X]$, and $x \in R$, we denote $Z(P;\mathcal{P})$ (resp. $Z(P,x;\mathcal{P})$) the number of roots of P (resp. less or equal to x) which are not root of any of the polynomials in $\mathcal{P}$. We claim that these two quantities are computable. Indeed, since $Z(P,x;\mathcal{P}) \leqslant \deg P$, we only have to check whether $Z(P,x;\mathcal{P}) = \ell$ for a finite number of ℓ's. But $Z(P,x;\mathcal{P}) = \ell$ is equivalent to

$$\exists r_1 < \cdots < r_\ell \left[\begin{array}{l} \bigwedge_{i=1}^{\ell} P(r_i) = 0 \quad \wedge \quad r_\ell \leqslant x \\ \wedge \\ \bigwedge_{i=0}^{\ell} \forall \xi, \quad r_i < \xi < r_{i+1} \Rightarrow P(\xi) \neq 0 \\ \wedge \\ \bigwedge_{i=1}^{\ell} \bigwedge_{Q\in\mathcal{P}} Q(r_i) \neq 0 \end{array}\right.$$

where we have set $r_0 := -\infty$ and $r_{\ell+1} := +\infty$ for writing convenience. These formulae are constructible and, thanks to Tarski-Seidenberg, a machine can determine which one is true. A similar argument shows that $Z(P;\mathcal{P})$ is also computable.

Proof of theorem 27. (i) $\Rightarrow$ (ii). Since $\operatorname{tr}_{\mathbb{Q}} R < \infty$ and R is real closed, it is well known that R is equal to the real closure of $\mathbb{Q}(a_1,\dots,a_n)$ for some $a_1,\dots,a_n \in R$. The number of polynomials on $\mathbb{Q}(a_1,\dots,a_n)$ is countable. In fact, using a diagonal procedure to enumerate $\mathbb{Q} \cup \{a_1,\dots,a_n\}$ and the polynomials, one can suppose that $\mathbb{Q}(a_1,\dots,a_n)[X] \setminus \{0\} = \{P_0, P_1, \dots\}$ with the map $i \mapsto P_i$ being computable. For any $x \in R$, let us define $\varepsilon(x)$ by

$$\varepsilon(x) := \sum\nolimits_{0\leqslant i<k} Z\big(P_i;\{P_j : j < i\}\big) + Z\big(P_k, x; \{P_j : j < k\}\big)$$

where $k := \min\{i : P_i(x) = 0\}$. First note that there is always some P_i of which x is a root and so k is easily computed by successively looking whether $P_0(x) = 0$, $P_1(x) = 0, \dots$ It is no difficulty to check that the map ε is one-to-one and reaches all positive integers.

(ii) $\Rightarrow$ (iii). Let us first deal with the case $S = R^N$. Since the map ε is computable, it is well known (see [BSS] §4, proposition 2) that one can write $R = \bigcup_{\gamma\in\Gamma} T_\gamma$ for some decidable subset Γ of $\mathbb{N}$ and some semialgebraic sets T_γ such that $\gamma \mapsto T_\gamma$ is computable and $\varepsilon\restriction T_\gamma$ is a rational map. The fact that ε is one-to-one and can only take integer values implies that none of the T_γ's can contain an open interval; that is, every T_γ is a finite set of points–(remember that semialgebraic subsets of R are finite unions of intervals). Let us consider a computable bijection $\nu : \mathbb{N} \to \Gamma$. One has

$$R^N = \bigcup\nolimits_{i\in\mathbb{N}} S_i \qquad \text{with } S_i := R^{N-1} \times T_{\nu(i)}.$$

Since $T_{\nu(i)}$ contains finitely many points, it is clear that $\dim S_i \leqslant N-1$.

Now, let us consider an arbitrary semialgebraic set S of R^N. According to theorem 3, one can write

$$S = \bigcup_{\alpha \in A} \varphi^{-1}(T_\alpha) \qquad \text{where } T_\alpha := \{x \in S : P_\alpha[x]\}. \tag{3}$$

The T_α's are disjoint integer cubes and φ is a semialgebraic bijection on S. Recall also that

$$\dim S = \max_{\alpha \in A} \dim T_\alpha. \tag{4}$$

Like R^N above, each T_α can be written as $T_\alpha = \bigcup_{i \in \mathbb{N}} T'_{(\alpha,i)}$ for some semialgebraic sets $T'_{(\alpha,i)}$ such that $\dim T'_{(\alpha,i)} < \dim T_\alpha$ and $i \mapsto T'_{(\alpha,i)}$ is computable. Let $\nu : \mathbb{N} \to A \times \mathbb{N}$ be a computable map. Then

$$S = \bigcup_{i \in \mathbb{N}} S_i \qquad \text{with } S_i := \varphi^{-1}(T'_{\nu(i)}).$$

As $\dim S_i = \dim T'_{\nu(i)} < \dim S$ and since $x \in S_i$ is equivalent to $\exists y,\ (x,y) \in \operatorname{Graph} \varphi \wedge y \in T'_{\nu(i)}$ and then to an open formula by Tarski-Seidenberg, the sets S_i have the desired properties.

(iii) $\Rightarrow$ (iv). Obvious.

(iv) $\Rightarrow$ (i). As above, S can be written like in (3). The relation (4) implies there must be some $\alpha \in A$ such that $\dim T_\alpha = \dim S$. The integer cube T_α is equal to the union of $\varphi\big(S_i \cap \varphi^{-1}(T_\alpha)\big)$, $i \in \mathbb{N}$, which are sets of dimension $\leqslant \dim S_i < \dim T_\alpha$. But $T_\alpha \cong R^{\dim T_\alpha}$ by a semialgebraic isomorphism. As a result, we can suppose from now on that $S = R^d$ for some $1 \leqslant d \leqslant N$.

Let us suppose for a moment that $d = 1$ and let $a_1, \dots, a_n$ be the constants of a machine computing $i \mapsto S_i$. Since $\dim S_i = 0$, S_i is a finite set of points. Therefore every $x \in S_i$ is a root of some nontrivial polynomial defining S_i and so $S_i \subseteq R'$ where R' denote the real closure of $\mathbb{Q}(a_1, \dots, a_n)$. But that implies $R \subseteq R'$ and consequently $\operatorname{tr}_{\mathbb{Q}} R \leqslant \operatorname{tr}_{\mathbb{Q}} R' \leqslant n$.

To conclude the proof, it is enough to show that, if (iv) is true for $S = R^d$, $d > 1$, then it is also true for $S = R^{d-1}$. More precisely, we claim that $\operatorname{tr}_{\mathbb{Q}} R < \infty$ or else there is some $x \in R$ such that $S^x := R^{d-1} \times \{x\}$ satisfies (iv), and so we can conlude by induction on d.

Let us prove the claim. Let $a_1, \dots, a_n$ be the constants of a machine that computes the covering $i \mapsto S_i$ of R^d. Set $S_i^x := S^x \cap S_i$. Obviously $S^x = \bigcup_{i \in \mathbb{N}} S_i^x$ and $\dim S_i^x \leqslant \dim S_i \leqslant \dim S^x$. Assume that $\dim S_i^x = \dim S^x$. Then x must be algebraic on the real closure of $\mathbb{Q}(a_1, \dots, a_n)$. Let us denote by f_i the product of the nontrivial polynomial f appearing in the formula defining S_i.[10] Indeed, either

(a) there exists some $\xi \in S_i^x$ such that $f_i(\xi, x) \neq 0$;

[10]There is only a finite number of polynomials in the open formula defining S_i.

or, otherwise, every $\xi \in S_i^x$ must be the root of f_i, and so, taking into account that the cardinal of S_i^x is infinite (because $\dim S_i^x = \dim S^x = d-1 \geqslant 1$) and that , one infers

(b) the polynomial $f_i(\cdot, x)$ is identically zero.

Case (a) cannot occur because, if it does, the continuity of polynomials would allow to find some cubic neighborhood C of $(\xi, x) \in R^d$ such that $C \subseteq S_i$ and this would contradict the fact that $\dim S_i < d$. So (b) is true. But then the coefficients of the polynomial f_i must vanish at x, and since f_i is nontrivial, at least one of these coefficients is a nontrivial polynomial over $\mathbb{Q}(a_1, \ldots, a_n)$. So, to sum up, or $\dim S_i^x < \dim S^x$ for all i i.e., S^x satisfies (iv), or else x belongs to the real closure of $\mathbb{Q}(a_1, \ldots, a_n)$. Therefore, if $\mathrm{tr}_{\mathbb{Q}} R = \infty$, one can find some $x \in R$ such that $S^x \cong R^{d-1}$ can be effectively covered by semialgebraic sets of smaller dimension. $\square$

Now the time has come to state the definition of the dimension for recursively enumerable sets. This definition is natural in view of proposition 2.8.5 of [BCR].

DEFINITION 28. Let R be a real closed field such that $\mathrm{tr}_{\mathbb{Q}} R = \infty$ and $S \subseteq R^N$, $N \leqslant \infty$, be a recursively enumerable set. We define $\dim_{\text{r.e.}} S$ by

$$\dim_{\text{r.e.}} S := \sup_{i \in \mathbb{N}} \dim S_i \qquad \text{where } S = \bigcup\nolimits_{i \in \mathbb{N}} S_i$$

and S_i are finite dimensional semialgebraic sets such that $i \mapsto S_i$ is computable.

By proposition 13, we know that at least one covering of S by some S_i's enjoying the above properties exists. We shall show that $\dim_{\text{r.e.}} S$ is independent of the chosen covering. So let $S = \bigcup_{i \in \mathbb{N}} S_i = \bigcup_{j \in \mathbb{N}} T_j$ where S_i and T_j are two covering of S having the required properties. Let $\nu : \mathbb{N} \to \mathbb{N}^2$ be a BSS-computable bijection. Then $U_k := S_{\nu_1(k)} \cap T_{\nu_2(k)}$ is another suitable covering of S. For the claim to be true, it suffices to show that

$$\dim S_i = \max_{j \in \mathbb{N}} \dim(S_i \cap T_j). \tag{5}$$

But that's exactly what the implication $\neg$(i) $\Rightarrow \neg$(iv) of theorem 27 states—the inequality '$\geqslant$' being obvious. Equality (5) also proves that, if S is a semialgebraic set, $\dim_{\text{r.e.}} S = \dim S$. The following proposition establishes the invariance of that notion of dimension. The proof is given in the extended version of this paper [MT]

PROPOSITION 29. *Let $S \subseteq R^N$ be a recursively enumerable set and $\varphi : R^N \to R^M$ a map which is decidable and one-to-one over* $\mathrm{Dom}\,\varphi \supseteq S$. *Then* $\dim_{\text{r.e.}} \varphi(S) = \dim_{\text{r.e.}} S$.

With this notion of dimension, some other equivalences with the statements of theorem 27 can easily be proved. They are analoguous to well-known results in classical recursion theory (see e.g., [Sh]). Of course, when $\mathrm{tr}_{\mathbb{Q}} = \infty$, the negation of the following satements hold. They in particular say that R^N cannot be decidably isomorphic to R^M unless $N = M$. This expresses the non-triviality of the dimension. This is similar to what happens for the topological dimension.

PROPOSITION 30. *Let R be a real closed field. The following statements are equivalent to those of theorem 27.*

(i) *For any integers $N > M \geqslant 1$, there is a one-to-one decidable map $R^N \to R^M$.*

(ii) *For any integers $N > M \geqslant 1$, there exists a surjective decidable map $R^M \to R^N$.*

(iii) *All R^N, $N \geqslant 1$, are decidably isomorphic.*

PROOF. Left to the reader. □

References

[BSS] L. BLUM, M. SHUB & S. SMALE, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, *Bull. of the AMS*, **21**, N° 1, 1–46 (1989).

[BCSS] L. BLUM, F. CUCKER, M. SHUB & S. SMALE, *Complexity and Real Computation*, Springer Verlag, 1998.

[BB] D. BERTILSSON & V. BLONDEL, Transcendence in simultaneous stabilization, *J.of Math. Systems, Estimation, and Control* **6**, N° 3, 1–22 (1996).

[BCR] J. BOCHNAK, M. COSTE & M.-F. ROY, *Géométrie Algébrique Réelle*, Springer Verlag, Berlin Heidelberg 1987.

[Co] J. P. COHEN, Decision procedures for real and p-adic fields, *Comm. Pure Appl. Math.* **22**, 131–151 (1969).

[vdD] L. VAN DEN DRIES, Alfred Tarski's elimination theory for real closed fields, *J. Symbolic Logic* **53**, 7–19, (1988).

[HW] W. HUREWICZ & H. WALLMAN *Dimension Theory*, Princeton University Press, 1948.

[Ko] P. KOIRAN, Randomized and deterministic algorithms for the dimension of Algebraic Varieties, preprint, 1997.

[Ło] S. ŁOJASEWICZ, *Ensembles semi-analytiques*, I.H.E.S. (prepublication) (1964). § 20 lemma 3

[MMV] A. Macintyre, K. McKenna & L. van den Dries, Elimination of quantifiers in algebraic structures, *Adv. in Math.* **47**, 74–87 (1983).

[MM] K. Meer & C. Michaux, A survey on real structural complexity theory, *Bull. Belg. Math. Soc.* **4**, 113–148 (1997).

[MT] C. Michaux & C. Troestler, Dimension theory for BSS recursively enumerable sets over real closed fields, preprint, 1998.

[Po] B. Poizat, *Les Petits Cailloux*, Aleas, Lyon 1995.

[Sh] J. Schoenfield, *Degrees of Unsolvability*, North-Holland, 1971.

[Ta] A. Tarski, *A Decision Method for Elementary Algebra and Geometry,* 2e ed., Berkeley and Los Angeles, 1951.

[Vo] N. Vorobjov, Computing dimensions of semialgebraic sets, preprint 1997.

[Wi] A. J. Wilkie, Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function, *J. Am. Math. Soc.* **9**, N° 4, 1051–1094 (1996).