The arithmetical hierarchy over the reals

Felipe Cucker

Report LSI-91-38

The arithmetical hierarchy over the reals

Felipe Cucker [†]
Dept. Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Barcelona 08028
SPAIN
e-mail: `cucker@lsi.upc.es`

0. Introduction

Alternation of quantifiers is a common measure of problem complexity either in the subrecursive setting or in the undecidable one. Thus, the polynomial, arithmetical and analytical hierarchies are semantically defined in terms of machines that make queries to oracles in a certain class, but are syntactically characterized by quantifier alternation. A formula like

$$\exists z_1 \forall z_2 \ldots Q z_n \ R(z_1 \ldots z_n, y_1 \ldots y_s) \tag{1}$$

where Q is a existential or universal quantifier depending on the parity of n, defines a subset of $\mathbb{N}^s$ belonging to Σ_n^P (in the polynomial hierarchy, see [1] ch. 8 for this hierarchy of complexity classes) if the sizes of the z_i are polynomial in the size of $(y_1 \ldots y_s)$, and R is a recursive relation which is computable in polynomial time. Also, it defines a set in Σ_n (in the arithmetical hierarchy of Kleene, see [8]) if no condition is imposed on the z_i or on R aside from being recursive.

On the other hand, a model of computation and a theory of recursiveness that allowed an ordered ring or field as alphabet for the space of admissible inputs has been recently introduced by L. Blum, M. Shub and S. Smale in [3], which emphasised the case when the ring is the field of real numbers, $\mathbb{R}$. In this case, if we take a formula like (1) where R is a recursive relation decidable in polynomial time, we obtain a subset of $\mathbb{R}^\infty$ which is in Σ_n^P in the polynomial hierachy (which is defined from P and NP in the same manner as the polynomial hierarchy in the boolean case).

This last fact contrasts with what happens over the integers, whose existential theory is undecidable, and is a consequence of the existence of quantifier elimination in the theory of the reals. In particular, it leads to the question of finding a syntactical characterization of the arithmetical hierarchy over the reals.

In section 1 of this paper we provide such a characterization in terms of alternations of countable connectives (conjunctions and disjunctions) for formulæ in the infinitary logic $\mathcal{L}_{\omega_1,1}$, that permits such connectives but no quantifiers. From such characterizations we show the existence of complete problems for each level of the hierarchy, and we exhibit some more natural problems that are complete in the low levels.

In order to classify some undecidable problems, one is led to consider expressions that merge infinitary connectives with quantification over variables denoting real numbers, and these expressions do not fit in any level of the arithmetical hierarchy. In section 2, we show that another arithmetical hierarchy can be defined by using nondeterministic machines and that the whole arithmetical hierarchy is contained in one of the lowest levels of its nondeterministic counterpart. We also give some complete natural problems in the low levels of this hierarchy. In section 3 we relate both hierarchies with classes of subsets of $\mathbb{R}^\infty$ currently studied in descriptive set theory. This allows us to obtain some more results concerning them.

---

1. The arithmetical hierarchy and infinitary logic

In what follows, we assume the reader is aquainted with the theory of computability and complexity over the real numbers introduced in [3]. Thus, concepts like real Turing machine, recursive or recursively enumerable subsets of $\mathbb{R}^\infty$ and alike will be freely used. We just recall that by $\mathbb{R}^\infty$ we denote the direct sum $\bigoplus_{i=1}^\infty \mathbb{R}$ and that we denote by $M(x) \downarrow^t$ the fact that the machine M with input x has halted after t steps, and by $M(x) \uparrow^t$ its negation. Also, in many situations, we shall say simply "machine" instead of real Turing machine expecting that no confusion can arise.

We begin by recalling that in [6], C. Michaux states a characterization of r.e. subsets of $\mathbb{R}^\infty$ in terms of countable unions of semialgebraic sets. However, no proof is provided there and, since this characterization together with some of the central ideas underlying the proof are central for what follows, we now give one possible proof.

Definition 1.1. A *semialgebraic set* S is a subset of some finitely dimensional real affine space, $S \subseteq \mathbb{R}^n$, for $n < \infty$, such that there is a finite number of polynomials $f_{ij} \in \mathbb{R}[X_1, \ldots, X_n]$ and sign conditions $\epsilon_{ij} \in \{> 0, = 0, < 0\}$ such that the following equality holds :

$$S = \bigcup_{j=1}^r \bigcap_{i=1}^t \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid sign(f_{ij}(x_1, \ldots, x_n)) = \epsilon_{ij}\}$$

Thus, the class of semialgebraic subsets of $\mathbb{R}^n$ is the closure under finite Boolean operations of the class of subsets defined by polynomial inequalities in the variables $X_1, \ldots, X_n$. If all the polynomials f_{ij} have coefficients in a subfield $H \subset \mathbb{R}$ we shall say that S is defined over H.

Semialgebraic sets are the main object of study in real algebraic geometry. For a comprehensive introduction to the subject see [2] and [4].

Definition 1.2. A language L is said to be *finitely generated* if there is a finitely generated field extension $\mathbb{Q}(S)$ of $\mathbb{Q}$ contained in $\mathbb{R}$ and a countable family of semialgebraic sets $\{A_n \subseteq \mathbb{R}^{i_n}\}$ defined over $\mathbb{Q}(S)$, such that

$$L = \bigcup_{n \in \mathbb{N}} A_n$$

This notion describes the geometrical structure of r.e. sets, as we shall see in the following theorem. We firstly prove a lemma which is of general interest.

Lemma 1.3. *For every partial mapping $s : \mathbb{N} \longmapsto \mathbb{N}$ there is a machine M which computes s when restricted to $\mathbb{N}$.*

Proof: Let us recall from [3] section 1, example 6, that given a set $S \subset \mathbb{N}$ we can decide whether a real number x belongs to S. Now, we consider the polynomial

$$f(n, m) = \frac{1 + n + m}{2}(n + m) + (m + 1),$$

the set $S = f(\text{graph}(s))$, and the algorithm

```
input(n)
m := 0
while f(n, m) ∉ S do
        m := m + 1
od
RETURN(m)
```

This algorithm computes $s(n)$. ∎

2

Theorem 1.4. *For any subset $L \subset \mathbb{R}^\infty$, L is r.e. iff L is finitely generated.*

Proof:

($\Rightarrow$) Suppose that L is recognised by a machine M. We know that for each pair $(t, n) \in \mathbb{N}^2$, the set

$$I_{t,n} := \{x \in \mathbb{R}^\infty \mid |x| = n \text{ and } x \text{ is accepted by } M \text{ in time } t\}$$

is a semialgebraic subset of $\mathbb{R}^n$. Moreover there is a finitely generated field extension $\mathbb{Q}(a_1, \ldots, a_r)$ of $\mathbb{Q}$, where $\{a_1, \ldots, a_r\}$ is the collection of all constants appearing in the description of M, such that for all $n \in \mathbb{N}$, $I_{t,n}$ can be described by a finite number of polynomials in $\mathbb{Q}(a_1, \ldots, a_r)[X_1, \ldots, X_n]$. Thus, we have that

$$L = (\bigcup_{t=1}^{\infty} I_{t,n})$$

from where we deduce the statement.

($\Leftarrow$) Let us now suppose that L is finitely presented. Then there is a finite extension $\mathbb{Q}(S) = \mathbb{Q}(a_1, \ldots, a_r)$ of $\mathbb{Q}$ and a countable family of semialgebraic sets defined over $\mathbb{Q}(S)$, $\{A_n \subseteq \mathbb{R}^{d_n}\}$ such that L is the union of the A_n's.

For each n the set A_n can be described as a finite system Φ_n of equalities and inequalities (as in the definition) involving polynomials with coefficients in $\mathbb{Q}(S)$. Thus, for each n, A_n is described by a family

$$\{(f_{ij}^n, \epsilon_{ij}^n)\}$$

where ϵ_{ij}^n is a sign and $f_{ij}^n \in \mathbb{Q}(S)[X_1, \ldots, X_{d_n}]$.

Now, each f_{ij}^n can be expressed as a polynomial in $\mathbb{Q}[a_1, \ldots, a_r, X_1, \ldots, X_{d_n}]$ and can then be coded as an element of $\mathbb{N}^\infty$. Thus, the same happens with the whole family $\{(f_{ij}^n, \epsilon_{ij}^n)\}$. We therefore consider the function

$$\varphi : \mathbb{N} \to \mathbb{N}^\infty$$
$$n \to \text{the encoding of } \{(f_{ij}^n, \epsilon_{ij}^n)\}$$

together with the machine given by

```
input(x)
n := |x|
t := 1
accept := false
while not accept do
        compute {(f_{ij}^t, ε_{ij}^t)} := φ(t)
        if d_t = n then
                for every (i, j) evaluate f_{ij}^n(a_1, ..., a_r, x_1, ..., x_n)
                if the evaluated values satisfy Φ_n then
                        accept := true
                fi
        fi
od
ACCEPT
```

The preceding machine accepts exactly L. ∎

Remarks 1.5. We want to attract the reader's attention to the fact that the preceding theorem provides a particular finite representation for r.e. sets. In fact, if a r.e. set L is given by

$$\bigcup_{n=1}^{\infty} A_n$$

where all the A_n are defined over $\mathbb{Q}(a_1, \ldots, a_r)$, we can represent L by the point $(a_1, \ldots, a_r, s)$ where s is the real number coding the function φ.

We now define the arithmetical hierarchy in the real setting, in the same manner as it is classically defined, i.e. using real Turing machines with oracles.

3

Definition 1.6. Let Σ_0 be the class of recursive sets. We inductively define Σ_{k+1} to be the class of sets accepted by real Turing machines that consult an oracle in Σ_k. Also, we define Π_k to be the class of sets whose complements are in Σ_k, for every $k \geq 0$, and $\Delta_k = \Sigma_k \cap \Pi_k$. The class $\mathbf{AH} = \cup_{k \geq 0} \Sigma_k$ is called the *arithmetical hierarchy*.

The standard relativization arguments *mutatis mutandis* allow one to prove that the inclusions between the different levels of the hierarchy are strict.

Proposition 1.7. *For every $k \geq 1$ the following inclusions are strict:*
  i) $\Sigma_k \subset \Delta_{k+1}$
  ii) $\Pi_k \subset \Delta_{k+1}$
  iii) $\Delta_k \subset \Sigma_k$
  iv) $\Delta_k \subset \Pi_k$ ∎

Our next concern is to characterize syntactically the classes in $\mathbf{AH}$. To do so, we begin by recalling that formulas in the logic $\mathcal{L}_{\omega_1,1}$ are constructed from the atomic ones by using countable conjunctions and disjunctions. The logic $\mathcal{L}_{\omega_1,1}$ is an example of an infinitary logic, and the reader interested in such logics should consult [5].

Formulæ in $\mathcal{L}_{\omega_1,1}$ can have a quite complicated structure of nested infinitary connectives, corresponding to the countable ordinals. Since we shall only deal with a class of formulæ possesing a simpler structure, we shall give a name to this class.

Definition 1.8. For any subfield F of $\mathbb{R}$ we define *the arithmetical formulæ* of $\mathcal{L}_{\omega_1,1}$ in the theory of ordered fields with constants in F to be those of the form

$$\bigvee_{n_1=1}^{\infty} \bigwedge_{n_2=1}^{\infty} \cdots \bigsqcup_{n_k=1}^{\infty} \varphi_{n_1\ldots n_k}(x_1 \ldots x_n) \tag{2}$$

or of the form

$$\bigwedge_{n_1=1}^{\infty} \bigvee_{n_2=1}^{\infty} \cdots \bigsqcup_{n_k=1}^{\infty} \varphi_{n_1\ldots n_k}(x_1 \ldots x_n) \tag{3}$$

where $\varphi_{n_1\ldots n_k}(x_1 \ldots x_n)$ is a quantifier-free formula with constants in F whose number of free variables n depends on the tuple $(n_1, \ldots, n_k)$, and $\bigsqcup$ is a conjunction or disjunction depending on the parity of k, the number of connective alternations.

Remarks 1.9. i) The requirement that $\varphi_{n_1\ldots n_k}(x_1 \ldots x_n)$ be quantifier-free is not strictly necessary since any first order formula in the theory of the real closed fields has a quantifier-free equivalent one.

ii) In the same manner as in remark 1.5, we observe now that any formula like (2) or (3) with constants in a field $F = \mathbb{Q}(a_1, \ldots, a_r)$ can be coded by a point $(a_1, \ldots, a_r, \alpha) \in \mathbb{R}^{r+1}$. This is the main point for our syntactical characterization of the sets in the arithmetical hierarchy.

We finally recall that we say that a point $(a_1, \ldots, a_n)$ satisfies a system of equations and inequalities $\Phi(x_1, \ldots, x_m)$ when $n = m$ and the closed formula $\Phi(a_1, \ldots, a_n)$ is a tautology. It is important to require the first condition since we want that finite formulæ define semialgebraic sets in $\mathbb{R}^{\infty}$ with finite dimension and not infinite cylinders over a semialgebraic basis. We extend this notion to formulæ in $\mathcal{L}_{\omega_1,1}$ by defining satisfaction of a countable disjunction as satisfaction of at least one of its terms, and satisfaction of a countable conjunction as satisfaction of all of its terms. The fact that a point $a = (a_1, \ldots, a_r)$ in $\mathbb{R}^{\infty}$ satisfies a formula $\Phi(x_1, x_2, \ldots)$ in $\mathcal{L}_{\omega_1,1}$ will be denoted by $\mathbb{R} \models \Phi(a)$.

4

Proposition 1.10.

   i) *For every $k \geq 1$ and for every set $S \in \Sigma_k$, there is a finitely generated extension F of $\mathbb{Q}$ contained in $\mathbb{R}$ and a formula Φ_S like (2) with constants in F such that $S = \{x \in \mathbb{R}^\infty \mid \mathbb{R} \models \Phi(x)\}$.*

   ii) *For every $k \geq 1$ and for every set $S \in \Pi_k$, there is a finitely generated extension F of $\mathbb{Q}$ contained in $\mathbb{R}$ and a formula Φ_S like (3) with constants in F such that $S = \{x \in \mathbb{R}^\infty \mid \mathbb{R} \models \Phi(x)\}$.*

Proof: We just prove the first statement. The second follows straight.

$(k = 1)$ Directly follows from theorem 1.4.

$(k > 1)$ Let $S \in \Sigma_k$. Then, there is a set $A \in \Sigma_{k-1}$ and a machine M which accepts S making queries to A. Now, if we add a ω-ary symbol relation $Orac(z)$ to our language $(+, -, *, /, <, =)$, the same arguments used to prove the "if" part of theorem 1.4 allow us to describe S as the set of the $x \in \mathbb{R}^\infty$ satisfying a formula

$$\bigvee_{n=1}^{\infty} \psi_n(x_1 \ldots x_{d_n}) \tag{4}$$

where now the atomic subformulæ of ψ_n are of the kind $t > 0$, $t = 0$ or $Orac(z)$ for any term t and any point $z \in \mathbb{R}^\infty$. By the induction hypothesis, $Orac(z)$ is equivalent to a formula

$$\bigvee_{n_1=1}^{\infty} \bigwedge_{n_2=1}^{\infty} \cdots \bigsqcup_{n_k=1}^{\infty} \varphi_{n_1 \ldots n_k}(x_1 \ldots x_n) \tag{5}$$

We now substitute (5) in (4) and since infinitary connectives behave like quantifiers with respect to the finitary connectives, by standard manipulations we can now "push them to the outside of the formula" getting a new one of the stated form.   ∎

Proposition 1.11.   *For every $k \geq 1$ the problem*

$$S_k = \{(\Phi, x) \mid \Phi \text{ is a formula like (2)}, x \in \mathbb{R}^\infty \text{ and } \mathbb{R} \models \Phi(x)\}$$

belongs to Σ_k, and the problem

$$P_k = \{(\Phi, x) \mid \Phi \text{ is a formula like (3)}, x \in \mathbb{R}^\infty \text{ and } \mathbb{R} \models \Phi(x)\}$$

belongs to Π_k.

Proof: As before we shall only prove that $S_k \in \Sigma_k$. For $k = 1$ the part ($\Leftarrow$) of theorem 1.4 proves our statement.

    For $k > 1$, for every Φ like (1) and for every $n \in \mathbb{N}$ we shall denote by Φ_n the subformula

$$\bigwedge_{n_2=1}^{\infty} \cdots \bigsqcup_{n_k=1}^{\infty} \varphi_{n, n_2 \ldots n_k}(x_1 \ldots x_n)$$

We have that $\Phi = \bigvee_{n=1}^{\infty} \Phi_n$, and from this equality and the fact that $S_{k-1} \in \Sigma_{k-1}$ by the induction hypothesis, we design the following machine

```
input(Φ, x)
n := 1
accept := false
while not accept do
        compute Φn
        if (Φn, x) ∈ Sk-1 then
                accept := true
        fi
od
ACCEPT
```

that accepts S_k by querying a set in Σ_{k-1}.   ∎

In the rest of this section we exhibit some problems that are complete in the first levels of the arithmetical hierarchy.

Let INJ be the set $\{x \in \mathbb{R}^\infty \mid$ the function computed by M_x is injective$\}$.

Theorem 1.12.   *The set* INJ *is* Π_1-*complete.*

Proof: We first note that the set belongs to Π_1 since $x \in$ INJ if and only if it satisfies

$$\bigwedge_{n=1}^{\infty} \bigwedge_{m=1}^{\infty} \bigwedge_{t=1}^{\infty} \forall z_1 \ldots \forall z_n \forall y_1 \ldots \forall y_m \quad M_x(z) \uparrow^t \vee M_x(y) \uparrow^t \vee M_x(z) = M_x(y)$$

On the other hand, we can map any formula

$$\Phi(v) = \bigwedge_{t=1}^{\infty} \varphi_t(v)$$

and any $z \in \mathbb{R}^\infty$ to the coding $f(\Phi, z)$ of the following machine

```
input(y)
n := |y|
for i = 1 to n do
        if ¬φᵢ(z) then
                RETURN(1) and HALT
        fi
od
RETURN(y)
```

Clearly, if $\Phi(z)$ is true, then the function computed by $M_{f(\Phi,z)}$ is injective, since it is the identity. On the other hand, if $\Phi(z)$ is false, then there is a $t \in \mathbb{N}$ such that $\varphi_t(z)$ is false, and so, the function computed by $M_{f(\Phi,z)}$ returns 1 for every y such that $|y| \geq t$. That shows that INJ is Π_1-hard.   ∎

A classical undecidable problem in the Boolean setting is to decide wheter the language recognized by a given machine is finite. Of course, this is equivalent to saying that such a language is contained in $\{0,1\}^k$ for some $k \in \mathbb{N}$, and then that the language can be recognized by a "finite" machine (a circuit for instance). This last problem is also interesting in the real case since, in particular, Blum, Shub and Smale introduced a class of finite machines (see [3] §2). So we define the problem FIN to be the set

$$\{x \in \mathbb{R}^\infty \mid \text{ the set acepted by the machine } M_x \text{ is contained in } \mathbb{R}^k \text{ for some } k \in \mathbb{N}\}$$

As in the Boolean case (cf. [8] §14.8) we have the following result

Theorem 1.13.   *The set* FIN *is* Σ_2-*complete.*

Proof: We first note that FIN is indeed in Σ_2 since it can be written in the following way

$$\bigvee_{n=1}^{\infty} \bigwedge_{m=1}^{\infty} \bigwedge_{t=1}^{\infty} \forall x_1 \ldots \forall x_m \ (m < n \vee M_x \text{ has not accepted } (x_1 \ldots x_m) \text{ after } t \text{ steps})$$

Now, we consider the reduction which associates with any formula of the form

$$\Phi(v) = \bigvee_{i=1}^{\infty} \bigwedge_{t=1}^{\infty} \varphi_{i,t}(v)$$

and any $z \in \mathbb{R}^\infty$ the coding $f(\Phi, z)$ of the following machine

```
input(y)
n := |y|
```

6

```
              for i = 1 to n do
                      t := 1
                      while not accept do
                              if ¬φ_{i,t}(z) then
                                      accept := true
                              fi
                      od
              od
              ACCEPT
```

In case that z satisfies Φ there is a $j \in \mathbb{N}$ such that for every $t \in \mathbb{N}$ we have that $\varphi_{i,t}(z)$ is true, and then, for all inputs with size greater than t, the machine $M_{f(\Phi,z)}$ will loop forever, showing that $f(\Phi, z) \in$ FIN. On the other hand, if z does not satisfy that formula, it is immediate that $M_{f(\Phi,z)}$ accepts all $\mathbb{R}^\infty$ and that finishes our proof. ∎

We recall that a subset S of $\mathbb{R}^\infty$ is bounded when there is a constant $K \in \mathbb{N}$ such that for every $x \in S$ $\|x\|^2 = x_1^2 + \ldots + x_{|x|}^2 < K$. Thus we define BOU to be the set

$$\{x \in \mathbb{R}^\infty \mid \text{the language accepted by } M_x \text{ is bounded}\}$$

for which we can prove the following result.

Theorem 1.14.   *The set* BOU *is* Σ_2-*complete.*

Proof: The membership in Σ_2 comes from the fact that BOU can be expressed by the formula

$$\bigvee_{K=1}^{\infty} \bigwedge_{n=1}^{\infty} \bigwedge_{t=1}^{\infty} \forall x_1 \ldots \forall x_n \ (\|(x_1, \ldots, x_n)\| < K \vee M_x \text{ accepts } (x_1 \ldots x_n) \text{ in less than } t \text{ steps})$$

To see the hardness in the class, just consider the reduction which associates with any formula

$$\Phi(v) = \bigvee_{i=1}^{\infty} \bigwedge_{t=1}^{\infty} \varphi_{i,t}(v)$$

and any $z \in \mathbb{R}^\infty$ the coding $f(\Phi, z)$ of the following machine

```
              input(y)
              K := ||y||^2
              for i = 1 to n do
                      t := 1
                      while not accept do
                              if ¬φ_{i,t}(z) then
                                      accept := true
                              fi
                      od
              od
              ACCEPT
```

∎

7

2. Nondeterminism strikes again

If we try to classify the problem TOT of deciding wheter a given machine computes a total function, we find that the problem is expressed by a formula like

$$\bigwedge_{n=1}^{\infty} \forall v_1 \ldots \forall v_n \bigvee_{t=1}^{\infty} (M_x \text{ halts on input } (v_1 \ldots v_n) \text{ after less than } t \text{ steps})$$

but no way seems to be available for moving the quantifiers inside the scope of all infinitary connectives (to further eliminate them) and thus, no membership in any class in **AH** clearly follows. However, the membership of a given machine in TOT can be disproved by a real Turing machine with an oracle in Σ_1 if we allow this machine to make nondeterministic guesses. This property motivates the following definition.

Definition 2.1. Let Σ_0^N be the class of recursive sets. We inductively define Σ_{k+1}^N to be the class of sets accepted by nondeterministic real Turing machines that consult an oracle in Σ_k^N. Also, we define Π_k^N to be the class of sets whose complements are in Σ_k^N, for every $k \geq 0$, and $\Delta_k^N = \Sigma_k^N \cap \Pi_k^N$. The class $\mathbf{NAH} = \cup_{k \geq 0} \Sigma_k^N$ is called the *nondeterministic arithmetical hierarchy*.

Again, we have the very first relation between classes in **NAH**.

Proposition 2.2. *For every $k \geq 1$ the following inclusions are strict:*
i) $\Sigma_k^N \subset \Delta_{k+1}^N$
ii) $\Pi_k^N \subset \Delta_{k+1}^N$
iii) $\Delta_k^N \subset \Sigma_k^N$
iv) $\Delta_k^N \subset \Pi_k^N$ ∎

A syntactical characterization of **NAH** can also be done, this time by means of classes of formulæ in the logic $\mathcal{L}_{\omega_1,\omega}$, an extension of $\mathcal{L}_{\omega_1,1}$ obtained by allowing also first order quantification.

Definition 2.3. For any subfield F of $\mathbb{R}$ we define *the arithmetical formulæ* of $\mathcal{L}_{\omega_1,\omega}$ in the theory of ordered fields with constants in F to be those of the form

$$\bigvee_{n_1=1}^{\infty} \exists x_{1,1} \ldots \exists x_{1,r_1} \bigwedge_{n_2=1}^{\infty} \forall x_{2,1} \ldots \forall x_{2,r_2} \ldots \bigsqcup_{n_k=1}^{\infty} \varphi_{n_1 \ldots n_k}(y_1 \ldots y_n) \tag{6}$$

or the form

$$\bigwedge_{n_1=1}^{\infty} \forall x_{1,1} \ldots \forall x_{1,r_1} \bigvee_{n_2=1}^{\infty} \exists x_{2,1} \ldots \exists x_{2,r_2} \ldots \bigsqcup_{n_k=1}^{\infty} \varphi_{n_1 \ldots n_k}(y_1 \ldots y_n) \tag{7}$$

where $\varphi_{n_1 \ldots n_k}(y_1 \ldots y_n)$ is a quantifier-free formula with constants in F whose number of free variables n depends on the tuple $(n_1, \ldots, n_k)$, and $\bigsqcup$ is a conjunction or disjunction depending on the parity of k, the number of connective alternations.

Lemma 2.4. *For any formula like (6) or (7) there is an equivalent one with the same form and such that for any $1 \leq j \leq k-1$, $r_j = n_j$.*

Proof: Let us consider a formula like (6)

$$\varphi = \bigvee_{n=1}^{\infty} \exists x_1 \ldots \exists x_{r_n} \psi_n(x_1, \ldots, x_{r_n}, \underline{y})$$

To every pair (n, r_n) we associate $m_n = 2^n 3^{r_n}$. Clearly, $m_n \geq max\{n, r_n\}$. Now, if $S = \{m_n \mid n \in \mathbb{N}\}$, we define

$$\bar{\varphi} = \bigvee_{m=1}^{\infty} \exists x_1 \ldots \exists x_m \bar{\psi}_r(x_1, \ldots, x_m, \underline{y})$$

where

$$\bar{\psi}_m = \begin{cases} \psi_n(x_1, \ldots, x_{r-n}, \underline{y}) \vee (\bigvee_{j=r_n+1}^{m} x_j^2 < 0), & \text{if } m = m_n \in S \\ (\bigvee_{j=1}^{m} x_j^2 < 0), & \text{if } m \notin \S \end{cases}$$

Clearly, $\bar{\varphi}$ is equivalent to φ and repeating this procedure with the Ψ_n we eventually get a formula as stated. ∎

8

As in the deterministic case, we have the following result.

Theorem 2.5.   *For every $k \geq 1$ the problem*

$$S_k = \{(\Phi, x) \mid \Phi \text{ is a formula like (6)}, x \in \mathbb{R}^\infty \text{ and } \mathbb{R} \models \Phi(x)\}$$

is Σ_k^N-complete, and the problem

$$P_k = \{(\Phi, x) \mid \Phi \text{ is a formula like (7)}, x \in \mathbb{R}^\infty \text{ and } \mathbb{R} \models \Phi(x)\}$$

is Π_k^N-complete.   ∎

Our next goal is to relate the deterministic arithmetical hierarchy to the nondeterministic one.

Remarks 2.6. We have seen in lemma 1.3 that any partial recursive function can be computed by a real Turing machine since it can be coded in a real number. On the other hand, any real number can be considered to code a function from the naturals to the naturals. In the sequel, if $x \in \mathbb{R}$, we shall denote by $[x]$ such function and we just recall here that given $n, m \in \mathbb{N}$ and $x \in \mathbb{R}$, the predicate $[x](n) = m$ is recursive.

Lemma 2.7.   *For every $k \geq 2$ and for every $\varphi \in \Pi_k$ there is a formula $\bar{\varphi}$ equivalent to φ of the form $\exists x \psi$ and such that if $k = 2$ then $\psi \in \Pi_1$, else $\psi \in \Pi_{k-2}$.*

Proof: Let

$$\varphi = \bigwedge_{n_k=1}^{\infty} \bigvee_{n_{k-1}=1}^{\infty} \cdots \bigsqcup_{n_1=1}^{\infty} \varphi_{n_1,\ldots,n_k}(y)$$

$$= \bigwedge_{n_k=1}^{\infty} \bigvee_{n_{k-1}=1}^{\infty} \Phi_{n_k,n_{k-1}}(y)$$

For every $y \in \mathbb{R}^\infty$ we have that

$$\mathbb{R} \models \varphi(y) \iff \forall n_k \in \mathbb{N} \; \exists n_{k-1} \in \mathbb{N} \text{ s.t. } \Phi_{n_k,n_{k-1}}(y) \text{ holds}$$

$$\iff \exists f : \mathbb{N} \to \mathbb{N} \text{ total s.t. } \forall i \in \mathbb{N} \; \Phi_{i,f(i)}(y) \text{ holds}$$

$$\iff \mathbb{R} \models \exists x \bigwedge_{i=1}^{\infty} \bigwedge_{m=1}^{\infty} (m \neq [x](i) \vee \Phi_{i,m}(y))$$

Now, since $m \neq [x](i)$ is a recursive predicate, we can write it as a countably conjunction and by contracting the three conjunctions in one, we get the desired result.   ∎

Proposition 2.8. $\mathbf{AH} \subseteq \Delta_2^N$.

Proof: By repeated applications of the preceeding lemma, it easily follows that $\forall k \, \Pi_k \subseteq \Sigma_2^N$. Then, since $\Sigma_k \subset \Pi_{k+1}$, it follows that $\forall k \Sigma_k \subseteq \Sigma_2^N$. Taking complements we get that $\forall k \Sigma_k \subseteq \Pi_2^N$ and $\Pi_k \subseteq \Pi_2^N$ from where we deduce the desired result.   ∎

We close this section classifying a couple of problems inside **NAH**.

9

Theorem 2.9. *The set* TOT *is* Π_2^N*-complete.*

Proof: We have already seen that it belongs to Π_2^N. For the hardness, just consider for any formula

$$\Phi = \bigwedge_{i=1}^{\infty} \forall v_1 \ldots \forall v_i \bigvee_{t=1}^{\infty} \varphi_{i,t}(z, v)$$

and any z the code $f(\Phi, z)$ of the machine

```
input(y)
n := |y|
t := 1
accept := false
        while not accept do
                if φ_{i,t}(z, y) then
                        accept := true
                fi
        od
od
RETURN(y)
```

which computes a total function if and only if $z \in$ TOT. ∎

Theorem 2.10. *The set* EXH$= \{x \in \mathbb{R}^\infty \mid$ *the function computed by* M_x *is exhaustive*$\}$ *is* Π_2^N*-complete.*

Proof: The set belongs to Π_2^N since it can be written by

$$\bigwedge_{n=1}^{\infty} \forall v_1 \ldots \forall v_n \bigvee_{m=1}^{\infty} \bigvee_{t=1}^{\infty} \exists w_1 \ldots \forall w_m \ (M_x(w_1 \ldots w_m) \downarrow^t \text{ returning } (v_1 \ldots v_n))$$

The reduction given in the preceding theorem shows also the hardness of this problem. ∎

3. Real Turing machines and descriptive set theory

The study of subsets of $\mathbb{R}$, $\mathbb{R}^n$, $\bigoplus \mathbb{R}$ or $\prod \mathbb{R}$ lies at the core of what is called descriptive set theory. Sets of real numbers are classified there according to a certain measure of the complexity of their descriptions. A central rôle is played by the Borelian sets which are those generated from the open ones by performing complements, and countable unions and intersections, and also by the analytical sets which are continuous images of Borelians. In this section we shall prove that the aritmetical subsets of $\mathbb{R}^\infty$ are exactly the Borelian ones of finite order, and that sets in class Σ_2^N are exactly the analytical ones. From those results we shall deduce that the inclusion $\mathbf{AH} \subset \Delta_2^N$ is strict.

We thus begin by recalling some basic notions of descriptive set theory (the interested reader can find a good exposition of the topic in [7]).

Definition 3.1. A *Polish space* is any topological space homeomorphic to a separated, complete metric one. A Polish space is said to be *perfect* if it has no isolated points. Thus the real field is a perfect Polish space. On the other hand, the class of Polish spaces is closed under countable sums and products as well as under closed and open subspaces.

Definition 3.2. The class of *Borelian subsets* of a Polish space $\mathcal{X}$ is the smallest class of subsets of $\mathcal{X}$ containing the open sets and closed under complements, countable unions and countable intersections. A Borelian subset of $\mathcal{X}$ is said to be of *finite order* when it can be obtained from the open sets of $\mathcal{X}$ by a finite number of such operations. A subset A of a Polish space $\mathcal{X}$ is said to be *analytical* when there is another Polish space $\mathcal{Y}$ and a continuous function $f : \mathcal{Y} \to \mathcal{X}$ such that $A = f(\mathcal{Y})$. In the case that $\mathcal{X} = \mathbb{R}^\infty$ we shall denote these three classes of sets by $\text{Bor}_\mathbb{R}$, $\text{Bor}_\mathbb{R}^F$ and Analytical$_\mathbb{R}$ respectively.

10

Theorem 3.3.   *The class of Borelian subsets of finite order of $\mathbb{R}^\infty$ coincides with the class of arithmetical sets.*

Proof: Since open sets are in Σ_1 and **AH** is closed under finitely many complements, countable unions and intersections, we get that borelian sets of finite order are arithmetic.

In order to prove the converse, it is enough to show that any semialgebraic subset of $\mathbb{R}^n$ is borelian of finite order. Now, isf S is such a set, it can be described by a Boolean combination of equations of the form

$$f(x_1,\ldots,x_n) = 0 \qquad \text{or} \qquad f(x_1,\ldots,x_n) > 0$$

The first formula defines a closed set of $\mathbb{R}^\infty$ and so, a Borelian one of finite order. The second one can be written as

$$\bigvee_{k=1}^{\infty} f(x_1,\ldots,x_n) \geq \frac{1}{k}$$

and thus as a countable union of closed subsets of $\mathbb{R}^\infty$, again Borelian of finite order. The conclusion now follows straightforwardly.   ∎

Theorem 3.4.   *The sets in Σ_2^N are precisely the analytical sets in $\mathbb{R}^\infty$.*

Proof: We shall rely on a characterization of analytical sets as projections of Borelian sets of finite order. In fact, we recall that if $\mathcal{X}$ is a perfect Polish space, and $\mathcal{Y}$ is any other Polish space, a subset $P \subset \mathcal{Y}$ is analytical if and only if there is a Borelian set of finite order $Q \subset \mathcal{X} \times \mathcal{Y}$ such that P is the projection of Q (see [7] 1G.12). We now take $\mathcal{X}$ to be $\mathbb{R}$ and $\mathcal{Y}$ to be $\mathbb{R}^\infty$. Now, by the preceding theorem, P is analytical if and only if there is an arithmetical formula $\varphi(x,y)$ such that

$$P = \{y \in \mathbb{R}^\infty \mid \mathbb{R} \models \exists x\, \varphi(x,y)\}$$

By repeatedly aplying lemma 2.7. we can transform $\varphi(x,y)$ into a Σ_2^N formula from where we deduce that analytical subsets of $\mathbb{R}^\infty$ are in Σ_2^N.

For the converse, let us consider a set S defined by a formula

$$\bigvee_{n=1}^{\infty} \exists x_1 \ldots \exists x_n \bigwedge_{t=1}^{\infty} \varphi_{n,t}(x_1,\ldots,x_n,y)$$

and the set $B \subset \mathbb{R}^\infty \times \mathbb{R}^\infty$ defined by

$$\bigcup_{n=1}^{\infty} \bigcap_{t=1}^{\infty} \varphi_{n,t}(x_1,\ldots,x_n,y)$$

Since that last formula is arithmetical, B is a borelian set. Now, we consider the projection

$$\pi : \mathbb{R}^\infty \times \mathbb{R}^\infty \to \mathbb{R}^\infty$$
$$(x,y) \to y$$

and it is trivial that $S = \pi(B)$, which proves that S is analytical.   ∎
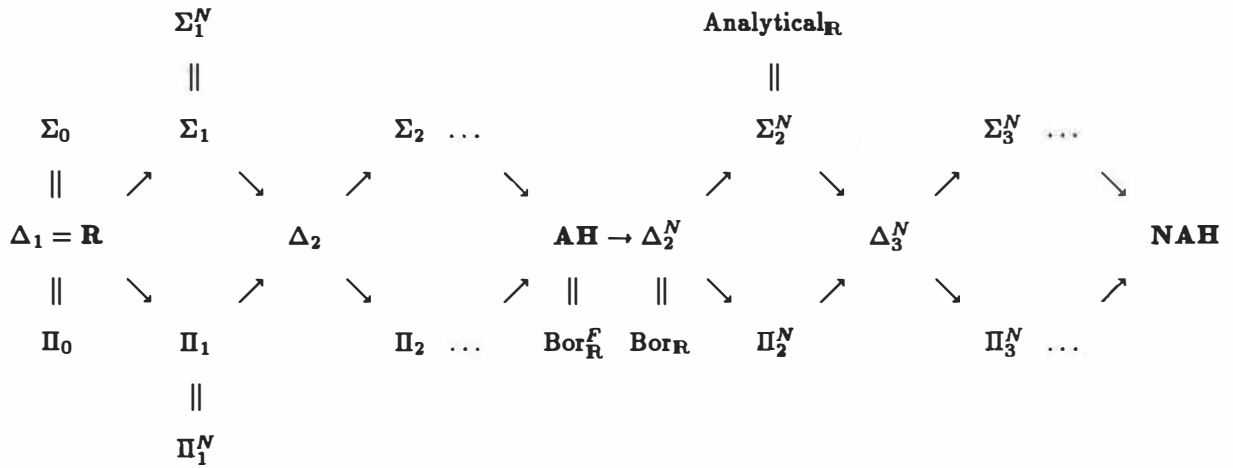
Corollary 3.5.   *The sets in Δ_2^N are exactly the Borelian ones.*

Proof: It is well known that a subset of a Polish space is analytical and co-analytical if and only if it is Borelian ([7] 2E.2).   ∎

Corollary 3.6.   *The inclusion* **AH** $\subset \Delta_2^N$ *is strict.*

Proof: Let us consider the set $B = \{(k,x) \mid x \in S_k\}$. Trivially, B is Borelian, since it is the union of a collection of sets isomorphic to the S_k's. On the other hand, let us suppose that $B \in$ **AH**. Then, there is a k such that $B \in \Sigma_k$. But using B as an oracle, we can decide S_{k+2}, and thus this set should belong to Σ_{k+1} contradicting proposition 1.7.   ∎

11

We can then summarize the relations between both arithmetical hierarchies in the following diagram where the arrows denote strict inclusions.

$$\Sigma_1^N \qquad\qquad\qquad\qquad \text{Analytical}_{\mathbf{R}}$$
$$\|\qquad\qquad\qquad\qquad\qquad\qquad \|$$
$$\Sigma_0 \quad \Sigma_1 \qquad \Sigma_2 \;\ldots \qquad\qquad \Sigma_2^N \qquad\qquad \Sigma_3^N \;\ldots$$
$$\| \quad \nearrow \quad \searrow \quad \nearrow \quad \searrow \quad \nearrow \quad \searrow \quad \nearrow \quad \searrow$$
$$\Delta_1 = \mathbf{R} \qquad\quad \Delta_2 \qquad\qquad \mathbf{AH} \to \Delta_2^N \qquad\qquad \Delta_3^N \qquad\qquad\qquad \mathbf{NAH}$$
$$\| \quad \searrow \quad \nearrow \quad \searrow \quad \nearrow \; \| \quad \| \; \searrow \quad \nearrow \quad \searrow \quad \nearrow$$
$$\Pi_0 \quad \Pi_1 \qquad \Pi_2 \;\ldots \quad \text{Bor}_{\mathbf{R}}^F \; \text{Bor}_{\mathbf{R}} \quad \Pi_2^N \qquad\qquad \Pi_3^N \;\ldots$$
$$\|$$
$$\Pi_1^N$$

Remarks 3.7. More relations can be pursued between the classes defined here and classes of real sets studied elsewhere. For instance, as a consequence of theorem 3.4 we have that the part of the nondeterministic arithmetical hierarchy constructed from Σ_2^N coincides with Suslin's projective hierarchy ([7] 1E). Also, subsets of $\mathbf{R}$ (or $\mathbf{R}^\infty$) can be defined using classical recursion theory: an open set is semirecursive when the basic open sets (open balls with rational center and radius) that are contained in it is a recursive set. From this class a whole hierarchy is obtained, an it can be seen that its classes are strictly contained in their corresponding classes within **AH**.

References.

[1] J.L. Balcásar, J. Días and J. Gabarró; *Structural Complexity I.* EATCS Monographs of Theoretical Computer Science, n 11, Springer Verlag, 1988.

[2] R. Benedetti and J.J. Risler; *Real algebraic and semialgebraic sets.* Hermann, Paris, 1990.

[3] L. Blum, M. Shub and S. Smale; "On a theory of computation and complexity over the real numbers: *NP*-completeness, recursive functions and universal machines". *Bulletin of the Amer. Math. Soc.*, vol.21, n.1, pp.1-46, 1989.

[4] J. Bochnak, M. Coste and M.-F. Roy; *Géométrie algébrique réelle.* Ergebnisse der Math., 3.Folge, Band 12, Springer Verlag, 1987.

[5] M. Dickmann; *Large Infinitary Languages*, Studies in Logic and the Foundations of Mathematics, vol. 83, North Holland, 1975.

[6] C. Michaux; "Machines sur les réels et problèmes NP-complets", *Séminaire de Structures Algébriques Ordonnées.* Prépublications de l'equipe de logique mathématique de Paris 7, 1990.

[7] Y. N. Moschovakis; *Descriptive set theory*, Studies in Logic and the Foundations of Mathematics, vol. 100, North Holland, 1980.

[8] H. Rogers; *Theory of recursive functions and effective computability*, McGraw-Hill, 1967.