On the complexity of some problems for the Blum, Shub & Smale model

Felipe Cucker
Francesc Roselló

Report LSI–90–42

On the complexity of some problems for the Blum, Shub & Smale model.

Felipe Cucker [1]
Dept. Llenguatges i Sistemes Informàtics
Univ. Politècnica de Catalunya
Barcelona 08028
SPAIN
e-mail: cucker@lsi.upc.es

Francesc Rosselló
Dept. Matemàtiques i Informàtica
Univ. de les Illes Balears
Palma de Mallorca 07071
SPAIN
e-mail: dmifrl0@ps.uib.es

Abstract. We show some problems coming from real algebra and semi-algebraic geometry to be *NP*-complete or co*NP*-complete for the Blum, Shub and Smale model of computation. We also introduce a class of languages R lying between P and NP that uses probabilistic machines, and several problems from the same area are classified as "probably non–complete" by showing their membership to R.

Very recently L. Blum, M. Shub and S. Smale ([2]) introduced a new model of computation, the real Turing machine (the BSS model in the sequel), that modelizes the kind of computations done in numerical analysis or computational geometry, where operations are in principle performed over real numbers. To this new model, a notion of time–complexity is attached that is shown to correspond to what is usually called algebraic complexity, i.e. an estimation of the number of arithmetic operations we must do to solve a problem.

In [2], analogous of the classes P and NP are introduced and the problem of deciding whether a degree 4 polynomial in several variables has a real root is shown to be *NP*-complete (under polynomial time reductions). The problem of deciding whether a semi–algebraic set given by polynomials of degree 2 is non–empty is also shown to be *NP*-complete. A first consequence of these results is that any algorithm solving the above quoted problems in polynomial time would give us, for any other problem in *NP*, an algorithm that solves it in polynomial time, that is, we should have $P{=}NP$ for the BSS model. On the standard model, this is a generally believed fact, even though no proof is known that $P{\neq}NP$. The main evidence supporting this guess is the number of *NP*-complete problems found till now for which no polynomial time algorithm is known. Maybe for that reason Blum, Shub, and Smale ask for other *NP*-complete problems. In this paper we give some examples of such problems. Let us recall that a semi–algebraic subset S of $\mathbb{R}^n$ is the set of points $x \in \mathbb{R}^n$ satisfying a system of the form

$$\bigvee_{i=1}^{s} \bigwedge_{j=1}^{r_i} f_{ij}(x) \left\{ \begin{matrix} = \\ > \\ \geq \end{matrix} \right\} 0 \qquad (*)$$

where the f_{ij} are polynomials. The system is said to be satisfiable when the semi–algebraic set that it defines is non–empty. We shall denote by $\mathcal{L}(d)$ the set of all systems like $(*)$ such that all the f_{ij} have degree smaller or equal than d. Analogously, the space of polynomials in any number of variables having degree smaller than d will be denoted by $\mathcal{P}(d)$.

1 Partially supported by DGICyT PB 860062 and the ESPRIT Basic Research Action Program of the EC under contract no. 3075, project ALCOM.

In section 6 of [2] it is shown that the problem

$$SAS_d = \{\varphi \in \mathcal{L}(d) \mid \varphi \text{ is a satisfiable system } \}.$$

is NP-complete for every $d \geq 2$. Also, for a single polynomial, and for every $d \geq 4$ the following feasibility problem

$$FEAS_d = \{f \in \mathcal{P}(d) \mid f \text{ has a real zero}\}$$

is shown to be NP-complete. In section 1.1. below we show that the problems

$$CONVEX_d = \{\varphi \in \mathcal{L}(d) \mid \text{the semi-algebraic set defined by } \varphi \text{ is convex}\},$$

$$kFINITE_d = \{\varphi \in \mathcal{L}(d) \mid \varphi \text{ has at most } k \text{ solutions}\},$$

$$POSITIVE_d = \{f \in \mathcal{P}(d) \mid \text{for every } x \in \mathbb{R}^n \ f(x) > 0\}$$

$$REGULAR_d = \{f \mid f \text{ is a regular polynomial}\}.$$

are coNP-complete for every $d \geq 2$ in the first two cases, and for every $d \geq 4$ in the other two.

Also, let us consider the following problems,

$$FINITE_d = \{\varphi \in \mathcal{L}(d) \mid \text{the semi-algebraic set defined by } \varphi \text{ is finite}\},$$

$$BOUNDED_d = \{\varphi \in \mathcal{L}(d) \mid \text{the semi-algebraic set defined by } \varphi \text{ is bounded}\},$$

$$ADH_d = \{\varphi \in \mathcal{L}(d) \mid 0 \text{ belongs to the adherence of the semi-algebraic set defined by } \varphi\}.$$

It is not difficult to show that these problems are NP-hard, while it is far from obvious the membership to NP. However, NP algorithms for them can be given if we restrict the inputs to have integer coefficients. The arguments used take advantage of some fine bounds on certain distances; they are exposed in section 1.2. From the logical point of view, the restriction to formulæ with integer coefficients corresponds to the decision of sentences in the basic theory of real closed fields where no symbols of constants are provided for elements in $\mathbb{R}$.

Little is known about the problems and classes between P and NP-complete problems for this model. One good candidate to this position seems to be, for any $d \in \mathbb{N}$,

$$SPOS_d = \{f \in \mathcal{P}(d) \mid \text{there is an } x \text{ such that } f(x) > 0\}$$

In section 2, a new probabilistic complexity class R lying between P and NP containing $SPOS_d$ is defined and it is shown there that the problems

$$NEI_d = \{\varphi \in \mathcal{L}(d) \mid \text{the semi-algebraic set defined by } \varphi \text{ has non-empty interior}\}$$

$$17^{\text{th}}\,HILBERT_d = \{f \in \mathcal{P}(d) \mid \text{for every } x_1, \ldots, x_n,\ f(x_1, \ldots, x_n) \geq 0\}$$

$$HYPERSURFACE_d = \{f \in \mathcal{P}(d) \mid \text{the zero set of } f \text{ has codimension } 1\}$$

also belong to R. Probability plays a role here since the machines randomly generate real numbers and the definition of acceptance involves the probability of getting an accepting output as a function of these random numbers. In this sense, this class is a first step towards [2] 11.4. where the authors ask "to develop a theory of probabilistic algorithms" for this computational model.

1. Around NP–completeness.

1.1. Some NP-complete problems.

Let us recall that a semi–algebraic subset S of $\mathbb{R}^n$ is the set of points $x \in \mathbb{R}^n$ satisfying a system of the form

$$\bigvee_{i=1}^{s} \bigwedge_{j=1}^{r_i} f_{ij}(x)\ \sigma_{ij}\ 0 \qquad (*)$$

where the f_{ij} are polynomials and the σ_{ij} are sign conditions taken from $\{=, >, \geq\}$.

We shall denote by $\mathcal{L}(d)$ the set of all systems like $(*)$ such that all the f_{ij} have degree smaller than or equal to d. Also, for every $\varphi \in \mathcal{L}(d)$ we shall denote by $\mathcal{S}(\varphi)$ the semi–algebraic set of the points $x \in \mathbb{R}^n$ satisfying φ and, if $\mathcal{S}(\varphi)$ is non-empty, we shall say that φ is satisfiable. In the same way, the space of polynomials in any number of variables having degree smaller than d will be denoted by $\mathcal{P}(d)$.

One of the main results exposed in [2] is the existence of *NP*-complete problems. In fact, let us consider

$$FEAS_4 = \{f \in \mathcal{P}(4) \mid f \text{ has a real zero}\}$$

and

$$SAS_2 = \{\varphi \in \mathcal{L}(2) \mid \varphi \text{ is a satisfiable system }\}.$$

It is shown there that both problems are *NP*-complete. We shall use them to show some new problems having this property.

Let us consider the following sets

$$CONVEX_d = \{\varphi \in \mathcal{L}(d) \mid \mathcal{S}(\varphi) \text{ is convex}\},$$

$$kFINITE_d = \{\varphi \in \mathcal{L}(d) \mid \mathcal{S}(\varphi) \text{ has less than } k \text{ points}\},$$

$$POSITIVE_d = \{f \in \mathcal{P}(d) \mid \text{for every } x \in \mathbb{R}^n\ f(x) > 0\}$$

and

$$REGULAR_d = \{f \in \mathcal{P}(d) \mid f \text{ is regular}\}.$$

To determine membership in those sets is a common problem in real algebraic geometry, always with a clear geometrical meaning.

Proposition 1.1.

i) $CONVEX_d$ is co*NP*-complete for every $d \geq 2$,

ii) $kFINITE_d$ is co*NP*-complete for every $d \geq 2$ and for every k,

iii) $POSITIVE_d$ is co*NP*-complete for every $d \geq 4$, and

iv) $REGULAR_d$ is co*NP*-complete for every $d \geq 4$.

Proof. i) The problem is clearly in co*NP* because, given a system φ, the following *NP* algorithm can prove the non–convexity of $\mathcal{S}(\varphi)$

input(φ)
begin
guess $x, y \in \mathbb{R}^n$ and $t \in (0,1)$
if $\varphi(x) \wedge \varphi(y) \wedge \neg\varphi(x + t(y - x))$ **then** ACCEPT
  else REJECT
fi
end

On the other hand, given a system φ, we can consider the system

$$\tilde{\varphi} := \varphi \wedge (z^2 - 1 = 0)$$

where z is a new variable. It is clear that φ is satisfiable if and only if $\tilde{\varphi}$ is so, and in this case the semi–algebraic set given by $\tilde{\varphi}$ is not convex. Therefore, $\varphi \in SAS_2 \iff \tilde{\varphi} \notin CONVEX_2$.

ii) Since a semi–algebraic set S defined by a formula φ is finite with its cardinal smaller than k if and only if

$$\forall x_1 \ldots \forall x_{k+1} \left(\left(\bigwedge_{i=1}^{k+1} \varphi(x_i) \right) \Rightarrow \left(\bigvee_{1 \leq i < j \leq k+1} x_i = x_j \right) \right),$$

it is not difficult to write up an NP algorithm which accepts when S has cardinal strictly greater than k. It implies that $kFINITE_d$ is in coNP.

On the other hand, given a system $\varphi \in \mathcal{L}(2)$, we can consider the new one

$$\tilde{\varphi}^{(k)} = \varphi \wedge (z_1^2 - 1 = 0) \wedge \ldots \wedge (z_m^2 - 1 = 0)$$

where $m = \lceil \log(k) \rceil + 1$ and $z_1, \ldots, z_m$ are new variables. It is again clear that φ is satisfied if and only if $\tilde{\varphi}^{(k)}$ is satisfied by strictly more than k different points. So $\varphi \in SAS_2 \iff \tilde{\varphi}^{(k)} \notin kFINITE_2$, which shows the corresponding hardness.

iii) Since the membership to coNP is obvious, we shall only show the hardness. To do so, we recall that the degree 4 polynomial obtained in the reduction given in [2] to show that $FEAS_4$ is NP-complete is a sum of squares. Thus, it is strictly positive if and only if it has no real root.

iv) Since an n–variated polynomial f is regular if and only if

$$\forall x \left(f(x) = 0 \Rightarrow \bigwedge_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \neq 0 \right)$$

it is clear than $REGULAR_d$ is in coNP.

On the other side, again since the degree 4 polynomial obtained in the reduction of [2] is a sum of squares, it has no real root if and only if it is regular. ∎

1.2. On the complexity of some problems with integer coefficients.

In this section we show that some problems concerning geometrical properties of semi–algebraic sets are in NP provided the defining polynomials have integer coefficients. Without this hypothesis it is not difficult to see that they are NP-hard but no argument seems to be available to prove the belonging to NP. From a logical point of view, this turns out to be equivalent to deciding satisfiability for a certain class of formulæ in the basic theory of real closed fields, when no symbols of constants are introduced in the language for elements in $\mathbb{R}$, and thus, the polynomials appearing in such formulæ have integer coefficients. So, in this section all the polynomials are supposed to have integer coefficients.

Let us consider for every $d \in \mathbb{N}$ the following problems

$$FINITE_d = \{\varphi \in \mathcal{L}(d) \mid \mathcal{S}(\varphi) \text{ is discrete}\},$$

$$BOUNDED_d = \{\varphi \in \mathcal{L}(d) \mid \mathcal{S}(\varphi) \text{ is bounded}\},$$

and

$$ADH_d = \{\varphi \in \mathcal{L}(d) \mid 0 \text{ belongs to the aherence of } \mathcal{S}(\varphi)\},$$

where all polynomials appearing in the systems φ have integer coefficients.

We need the following fact:

Proposition 1.2.
a) There exists a $p \in \mathbb{Z}^+$ such that, for any closed and bounded semi–algebraic set $S \subset \mathbb{R}^n$ given by a system $\varphi \in \mathcal{L}(d)$ and for any $x \in S$, the set $B(x, 1/L^{D^{n^p}}) \cap S$ is contained in only one semi–algebraically connected component of S, where L and D are upper bounds for the absolute value of the coefficients of the polynomials appearing in φ and the total degree of φ.
b) There exists a $q \in \mathbb{Z}^+$ such that, for any bounded semi–algebraic set $S \subset \mathbb{R}^n$ given by a system $\varphi \in \mathcal{L}(d)$ and for any $x \in S$, S is contained in $B(x, L^{D^{n^q}})$, with L and D as before.

Proof. Point (a) is [5] Prop. 14. Here we shall sketch the proof of (b), which is very similar to that one.

Let S be a bounded semi–algebraic set given by $\varphi \in \mathcal{L}(d)$. Then the adherence $\overline{S}$ can be defined by a formula $\overline{\varphi}$ whose total degree is $\overline{D} = D^{n^{O(1)}}$ and whose coefficients are bounded in absolute value by $\overline{L} = L^{D^{n^{O(1)}}}$ ([6]). Since S is bounded, $\overline{S}$ is also bounded.

Let us consider the map

$$\begin{array}{rccc} f: & \overline{S} \times \overline{S} & \longrightarrow & \mathbb{R} \\ & (x, y) & \longmapsto & \|x - y\|^2 \end{array}$$

The set $\mathcal{E}$ of images of local extrema with respect to f is a finite semi–algebraic set ([5] Lemma 13), which is included in the set of zeroes of a univariated integer polynomial F of degree $\overline{D}^{n^{O(1)}}$ whose coefficients are bounded in absolute value by $\overline{L}^{\overline{D}^{n^{O(1)}}} = L^{D^{n^{O(1)}}}$ (cf. [5] Prop. 14).

Thus, if R is the maximum of the roots of F and if $\xi > R$ then for any $x \in S$ the ball $B(x, \xi)$ contains S (since it contains its adherence $\overline{S}$).

Using the usual upper bound for the roots of a univariated polynomial (see [7]) we have that we can take $\xi = L^{D^{n^q}}$ for some $q \in \mathbb{Z}^+$. ∎

Notice that, for any $c \in \mathbb{Z}^+$, $L^{D^{n^c}}$ can be computed in polynomial time with respect to n.

Theorem 1.3. For every d we have that
i) $FINITE_d$ is in coNP,
ii) $BOUNDED_d$ is in coNP, and
iii) ADH_d is in NP.
Proof.
(i) Because of proposition 1.2. a), a semialgebraic set S defined by a formula $\varphi \in \mathcal{L}(d)$ has dimension ≤ 0 if and only if

$$\forall x \forall y (\varphi(x) \wedge \varphi(y) \wedge \|x - y\|^2 < 1/L^{D^{n^q}} \Rightarrow x = y).$$

So, φ does not belong to $FINITE_d$ if and only if

$$\exists x \exists y (\varphi(x) \wedge \varphi(y) \wedge \|x - y\|^2 < 1/L^{D^{n^q}} \wedge x \neq y),$$

which shows that $FINITE_d$ is in coNP.
(ii) Just notice that, because of proposition 1.2. b), a semi–algebraic set S defined by a formula $\varphi \in \mathcal{L}(d)$ is unbounded if and only if

$$\exists x \exists y (\varphi(x) \wedge \varphi(y) \wedge \|x - y\|^2 > L^{D^{n^q}}).$$

(iii) Is an easy consequence of the following
FACT: There exists a $c \in \mathbb{Z}^+$ such that, for any semi–algebraic set $S \subset \mathbb{R}^n$ defined by a formula $\varphi \in \mathcal{L}(d)$, $0 \in \overline{S}$ if and only if

$$\exists x(0 < \|x\|^2 < 1/L^{D^{n^c}} \wedge \varphi(x))$$

where L and D are as in proposition 1.2.
In order to prove it, let us consider the involutive transformation $T : \mathbb{R}^n - \{0\} \longrightarrow \mathbb{R}^n - \{0\}$ defined by

$$T(x) = \frac{x}{\|x\|^2}.$$

Notice that $\|T(x)\| = 1/\|x\|$.
Given a polynomial $f \in \mathcal{P}(d)$, let $T^*(f) = \|x\|^{2d} f(T(x))$. Given a formula $\varphi \in \mathcal{L}(d)$, we denote by $T^*\varphi$ the formula obtained after replacing any polynomial f in it by $T^*(f)$, and given the semi–algebraic set S defined by φ, we denote by T^*S the semi–algebraic set defined by $T^*\varphi$. Notice that $T^*\varphi \in \mathcal{L}(2d)$, and thus its total degree is $2D$, and that if $\tilde{L}$ denotes an upper bound for the absolute values of the coefficients of the polynomials in $T^*\varphi$ then $\tilde{L}$ is a polynomial in d and L.
By the elementary properties of T one has that $0 \in \overline{S - \{0\}}$ if and only if T^*S is unbounded. Thus, by proposition 1.2. b),

$$\begin{aligned} S \in ADH_d &\iff \exists x(\|x\|^2 > \tilde{L}^{D^{n^q}} \wedge T^*\varphi(x)) \\ &\iff \exists x(0 < \|x\|^2 < 1/\tilde{L}^{D^{n^q}} \wedge \varphi(x)) \end{aligned}$$

whence we easily obtain the statement. ∎

2. A probabilistic complexity class.

Given a real polynomial f of degree 4, we know that deciding whether there exists x such that $f(x) = 0$ is an NP-complete problem. We also know that the same happens for the problem of deciding whether there exists an x such that $f(x) \geq 0$ by theorem 1.1. iii). However, no such result seems to be true for the existence of an x satisfying $f(x) > 0$. While it is clear that this problem ($SPOS_4$ in the sequel) is in NP, nothing leads us to think that it is hard for such a class. Indeed let us consider the following non–deterministic algorithm for $SPOS_4$

```
input(f)
begin
if f = 0 then REJECT
else
        guess x_1, ..., x_n
        if f(x_1, ..., x_n) > 0 then ACCEPT
        else REJECT
        fi
fi
end
```

One remarkable feature of this algorithm, is that if it exists an accepting guess then there exists a set of accepting guesses with non–empty interior in $\mathbb{R}^n$. In other words the accepted polynomials are accepted with probability strictly greater than zero for all continuous distributions given to $\mathbb{R}^n$. This is certainly a property that is not shared for any NP algorithm accepting $FEAS_4$.

We shall see that such property defines a complexity class lying between P and NP.

We briefly recall from [2] that a *real Turing machine* consists of an input space $\bar{I}_M = \mathbb{R}^*$, an output space $\mathbb{R}^*$ and a state space $S = \mathbb{Z}^+ \times \mathbb{Z}^+ \times \mathbb{R}^*$, together with a connected directed graph whose nodes labelled $1 \ldots N$ (the set of differents instructions) are of certain types and with associated functions. The internal content of S at time t is $(i, j, x_1, x_2, x_3, \ldots)$ where for $t = 1$ the input is in the x_s with s odd (thus we reserve the even coordinates to leave work space), and x_2 can denote the length of the input. The five types of nodes are as follows:

1) *Exactly one input node*: node 1. Associated with this node is a next node $\beta(1)$.
2) *Exactly one output node*: node N. Once it is reached the computation halts, the contents of the real part of S being considered as the output.
3) *Computation nodes.* Associated with a node m of this type there is a next node $\beta(m)$ and a map $g_m : S \to S$. The g_n is of the form $g_m(i, j, x) = (i'(i), j'(j), x'(x))$, with $i'(i) = i + 1$ or 1, $j'(i) = j + 1$ or 1, and x' is a polynomial or rational map.
4) *Branch nodes.* There are two nodes associated with this node: $\beta^+(m)$ and $\beta^-(m)$. The next node is $\beta^+(m)$ if $x_1 \geq 0$ and $\beta^-(m)$ otherwise.
5) *Move nodes.* It has a unique next node $\beta(m)$. If the current element of S is $(i, j, x_1, \ldots)$ it operates replacing x_j by x_i in the j^{th} place of the vector $\mathbb{R}^*$ in S.

An instantaneous description of any moment of the computation can be given by providing an element in S and the current node. The first one changes according to the function associated with the current node and the node itself according to the function β.

We also recall from [2] that a machine M is said *to work in polynomial time* when there are constants $c, q \in \mathbb{Z}^+$ such that for every input $y \in \mathbb{R}^*$, M reaches its output node after at most $c(size(y))^q$ steps. The class P is then defined as the set of all subsets of $\mathbb{R}^*$ that can be accepted by a machine working in polynomial time.

Definition. We now introduce the class R as the set of decision problems Y for which there are constants $c, q \in \mathbb{N}$ and a machine M over $\mathbb{R}$ with $\bar{I}_M = \bar{I} \times \bar{I}'$, where $\bar{I} = \bar{I}' = \mathbb{R}^*$ such that

(a) the outputs of M are 1 (*yes*) and 0 (*no*).
(b) for every y, y' the computation time of M for the input (y, y') is smaller than $c(\text{size}(y))^q$.
(c) $y \in Y$ iff the set $\{y' \in \bar{I}' \mid \phi_M(y, y') = 1\}$ has non–zero measure under the assumption that each y' is a random variable with normal standard distribution.
(d) for every node of type 4 we have that, if there exists (y, y') such that some queried value at this node during the computation for the input (y, y') involves y', then for every $y \in \bar{I}$ the set $\{y' \in \bar{I}' \mid$ that queried value is $0\}$ has a measure of zero. Nodes of type 4 satisfying this condition will be called *proper tests.*

Remarks.

i) The y' in the definition plays the role of the random choices made by the machine, as with the Boolean probabilistic machines (see [1] ch.6). It is clear that, as for languages in NP the size of y' can be considered to be $c(\text{size}(y))^q$. So, we can assume that we take the measure of the choices leading to an accepting output in a real space with that dimension.

Moreover, it is clear that condition (c) is equivalent to restricting the set of choices leading to an accepting output to have non–empty interior.

ii) With this definition, the algorithm given at the beginning of this section shows that $SPOS_4 \in R$.

iii) Condition (d) in the definition does not seem, to us, to be a condition testable in polynomial time for a given machine, even for a given input. So it is a responsibility of the programmer to check that the program is correct in that sense (something similar happens with real RAM's with

indirect addressing, where the programmer must show that no addressing to non–integer addresses is made).

iv) One feature over which we want to attract the reader's attention is that the inclusion $R \subseteq NP$ is not staightforward, as in the Boolean case, due to the fact that in our continuous distribution, probability zero does not entail emptyness.

v) We finally remark that since we can not give an *a priori* strictly positive lower bound for the probability of reaching an accepting output, no probability amplification lemma is now available.

The following property is a trivial consequence of condition (d).

Lemma 2.1. For every problem L in R we can devise a machine that accepts L whose proper tests have the form

if $E > 0$ **then** go to β^+
elif $E < 0$ **then** go to β^-
else REJECT and halt
fi

Proof. Let us consider a machine M accepting L. We define a new machine M' by replacing all proper tests in M of the form

if $E \geq 0$ **then** go to β^+
else go to β^-
fi

by tests with the desired form.

Clearly, M' accepts L because it behaves like M except for a set of choices that has a measure of zero since it is defined by *non–constant* equalities. ∎

Theorem 2.2. We have that $P \subseteq R \subseteq NP$.

Proof. The first inclusion is trivial. For the second one, let us consider a problem L in R and a machine M accepting L that we shall suppose has proper tests in the form given by the lemma above. We claim that the same machine accepts L considered as a non–deterministic one. In fact, if $x \in L$ then there are (many) accepting guesses and then we have the non–deterministic acceptance of x. On the other hand, if M accepts x as a non–deterministic machine, then we have a sequence of guesses $y_1, \ldots, y_m$ that lead to an accepting output. During these computation we pass through a finite number of proper tests. For each one of them let $y_{i_1}, \ldots, y_{i_s}$ be the guesses involved in the test. Since the equality sign leads to a rejection, the queried sign must have been > 0 or < 0, and then there exist $\epsilon_{i_1}, \ldots, \epsilon_{i_s}$ such that the same sign holds if we replace y_{i_j} by any point in the open ball of center y_{i_j} and radius ϵ_{i_j}. For every $i = 1, \ldots, m$ we consider ϵ_i defined as the minimum of the ϵ_{i_j} for all the proper tests where y_i is involved. Clearly the elements in the open set

$$\{z_1, \ldots, z_m \mid z_i \in (y_i - \epsilon_i, y_i + \epsilon_i) \quad i = 1, \ldots, m\}$$

are all sequences of accepting guesses. ∎

A natural question concerning this class is whether it has complete problems. We do not have an answer to it, but we recall that in the Boolean model and for the probabilistic complexity classes contained in NP (R and ZPP) no complete problems are known.

Other natural questions are whether P=R or R=NP. The only partial answer we can now give is the following, very easy, one.

Proposition 2.3. If some problem in R is NP-complete, then R=NP. ∎

We close this section exhibiting some problems in R. Let us consider the following sets:

$$NEI_d = \{\varphi \in \mathcal{L}(d) \mid \mathcal{S}(\varphi) \text{ has non-empty interior}\}$$

$$17^{\text{th}}\,HILBERT_d = \{f \in \mathcal{P}(d) \mid \text{for every } x_1, \ldots, x_n,\ f(x_1, \ldots, x_n) \geq 0\}$$

and

$$HYPERSURFACE_d = \{f \in \mathcal{P}(d) \mid \text{the zero set of } f \text{ is a hypersurface i.e. has codimension } 1\}$$

The second problem has considerable historical importance, because the polynomials being positive on the whole space are exactly those which can be written as a sum of squares of rational functions in $R(X_1, \ldots, X_n)$. This characterization was asked by Hilbert in one of his famous 23 problems (the 17^{th}) and the positive answer was given by Artin and Schreier in the 20's and motivated the introduction of the real closed fields (for more data concerning Hilbert's seventeenth problem, see chapter 6 of [3]).

Lemma 2.4. For any squarefree $f \in \mathbb{R}[X_1, \ldots, X_n]$, the following conditions are equivalent:

i) f changes sign (i.e. $\exists x, y \in \mathbb{R}^n\ f(x)f(y) < 0$),

ii) the zero set of f has dimension $n-1$.

Proof.

$i) \Rightarrow ii)$ Let $U_1 = \{x \in \mathbb{R}^n \mid f(x) > 0\}$ and $U_2 = \{x \in \mathbb{R}^n \mid f(x) < 0\}$. Since those sets are open, non–empty and disjoint, we have that $\dim(\mathbb{R}^n - (U_1 \cup U_2)) \geq n-1$ (see [3] 4.5.2.) and ii) follows.

$ii) \Rightarrow i)$ Let p be a prime divisor of f whose zero set has dimension $n-1$. Since p is prime, it generates the ideal of polynomials vanishing at its zero set. Now, if $f = p \cdot q$, p and q are coprimes because f is squarefree. We can then find a point $a = a_1, \ldots, a_n$ that is a non–singular zero of p but not a zero of q. Therefore, there exists $i \leq n$ such that $p(a) = 0$, $\frac{\partial p}{\partial X_i}(a) \neq 0$ and $q(a) \neq 0$. This implies that $f(a) = 0$ and $\frac{\partial f}{\partial X_i}(a) \neq 0$. Thus, the function

$$x \;\rightarrow\; f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$$

is strictly monotonic over an interval containing a_i, and f must change sign. ∎

Proposition 2.5. We have that for every d

(i) NEI_d is in R,

(ii) $17^{\text{th}}\text{HILBERT}_d$ is in co$R$, and

(iii) $HYPERSURFACE_d$ is in R.

Proof. For (i), let φ be the system

$$\bigvee_{i=1}^{s} \bigwedge_{j=1}^{r_i} f_{ij}(x)\ \sigma_{ij}\ 0$$

The set $\mathcal{S}(\varphi)$ has non–empty interior if and only if one of the union sets has this property. But these union sets, being intersections of sets defined by equations and inequations, have non–empty interior if and only if no equality appears in the set and there exists a point $x \in \mathbb{R}^n$ satisfying the strict inequalities.

Part (ii) is trivial.

For part (iii) just note that, because of the preceding lemma, the algorithm

input(f)
begin
if $f = 0$ **then** REJECT

else
compute $\tilde{f}$ the squarefree part of f
randomly choose $x_1, \ldots, x_n, y_1, \ldots, y_n$
if $\tilde{f}(x_1, \ldots, x_n)\tilde{f}(y_1, \ldots, y_n) < 0$ **then** ACCEPT
else REJECT
fi
fi
end

accepts f if and only if the zero set of $\tilde{f}$ (which coincides with the one of f) has dimension $n-1$. ∎

3. Some open problems.

1. A first related problem is the complexity of quantifier elimination. In recent years many algorithms have been given that decide whether a quantified sentence in the elementary language of the real closed fields is true or not (see for instance [4], or [8]). These algorithms have all the same time complexity, which is simply exponential in the number of variables and doubly exponential in the number of quantifier alternations. This differs from the Boolean case in which the decision of quantified Boolean formulæ can be done in single exponential time and moreover is known to be *PSPACE*-complete.

The question now is: is there a complexity class for which the decision of quantified sentences over the reals is a complete problem?

This seems to be a non–trivial question, one of its disturbing features being the fact that no complexity classes defined by bounds on the used space appear to be useful. Indeed, consider for instance the set $\{(x, y) \in \mathbb{R}^2 \mid y \neq e^x\}$. This set can be accepted in constant space but is not even recursive because its complement is the graph of the exponential function which cannot be written as a countable union of semi–algebraic sets (see [2] proposition 2 for this characterization of recursive sets).

2. In the BSS model one can define, as in the Boolean one, a polynomial hierarchy of complexity classes (for the Boolean case, see [1] ch.8), and a syntactical characterization of those classes by the number of alternations of quantifiers easily follows. Many of the problems we have seen in the preceding sections can be "naturally" stated as problems in some higher levels of this hierarchy. For instance, the fact that $\varphi \in NEI_d$ is usually expressed by

$$\exists x \in \mathbb{R}^n \ \exists \epsilon \forall y \in \mathbb{R}^n \ (\varphi(x) \wedge \|x - y\|^2 < \epsilon \Rightarrow \varphi(y))$$

and the fact that $\varphi \in BOUNDED_d$ is expressed by

$$\exists K \forall y \in \mathbb{R}^n \ (\|y\|^2 < K \vee \neg\, \varphi(y))$$

For the first problem, we have seen that a more simple expression can be found since the problem is in *NP* (in fact, in *R*). For the second problem, however, no *NP* algorithm is known. Syntactically, the obstruction comes from the $\exists K$ in the beginning of the formula, and this kind of obstruction (an alternance produced by a single quantifier not adjacent to the quantifier–free part of the formula) also appears in the other problems dealt with in section 1.2. Notice that in the Boolean case, the cost of eliminating this single quantifier is linear in time and duplicates the size of the formula. In fact, a formula like

$$\forall y \, \exists x_1 \cdots \exists x_n \ \varphi(y, x_1, \ldots, x_n)$$

is equivalent to

$$\exists x_1 \cdots \exists x_n \exists z_1 \cdots \exists z_n \ (\varphi(0, x_1, \ldots, x_n) \wedge \varphi(1, z_1, \ldots, z_n)).$$

The above mentioned algorithms for quantifier elimination in the theory of real closed fields do not share this property, the eliminated quantifier being always the innermost one.

The question that arises then is: is there a way of performing this kind of elimination in the real case with polynomial cost?

3. It is easy to define a new complexity class by putting aside condition (d) in the definition of class R, and this class has complete problems. However, what is not obvious now is that this class is contained in NP. An open problem, for us, is whether condition (d) is superfluous in the definition of R.

Acknowledgment. Thanks are due to José Luis Balcázar for several discussions as well as for the careful reading of this manuscript.

References.

[1] J.L. Balcázar, J. Díaz and J. Gabarró; *Structural Complexity.* vol.1, EATCS Monographs of Theoretical Computer Science, Springer Verlag, 1988.

[2] 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. A preliminary version appeared in 29^{th} *Found. of Comp. Sc.* pp.387-397, 1988.

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

[4] D. Grigor′ ev; " Complexity of deciding Tarski algebra". *J. of Symb. Comp.*, **5**, pp.65-108, 1988.

[5] J. Heintz, T. Krick, M.-F. Roy and P. Solerno;"Single exponential time algorithms for basic constructions in elementary geometry", *Proceedings of the* 10^{th} *Int. Conf. of the Chilean Comp. Sc. Soc.*, Santiago de Chile, 1990.

[6] J. Heintz, M.-F. Roy and P. Solerno;"Sur la complexité du principe de Tarski–Seidenberg". *Bull. Soc. Math. France*, **118**, pp.101-126, 1990.

[7] M. Mignote; "Some useful bounds" in *Computer Algebra, Symbolic and Algebraic Computation*, pp.259-263, Springer Verlag, 1982.

[8] J. Renegar; "On the computational complexity and geometry of the first order thery of the reals", parts I, II and III. *Cornell University, Technical Reports 853,854 and 856*, 1989.