DIMACS Technical Report 96-27
July 1996

Hilbert's Nullstellensatz is in the Polynomial Hierarchy

by

Pascal Koiran[1]
LIP, Ecole Normale Supérieure de Lyon – CNRS
46 allée d'Italie, 69364 Lyon Cedex 07
France
email: koiran@lip.ens-lyon.fr

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Research, Bellcore, and Bell Laboratories.
DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

Abstract

We show that if the Generalized Riemann Hypothesis is true, the problem of deciding whether a system of polynomial equations in several complex variables has a solution is in the second level of the polynomial hierarchy. In fact, this problem is in AM, the "Arthur-Merlin" class (recall that $\mathrm{NP} \subseteq \mathrm{AM} \subseteq \mathrm{RP}^{\mathrm{NP}} \subseteq \Pi_2$). The best previous bound was PSPACE.

An earlier version of this paper was distributed as NeuroCOLT Technical Report 96-44. The present paper includes in particular a new lower bound for unsatisfiable systems, and remarks on the Arthur-Merlin class.

[1]A part of this work was done when the author was visiting DIMACS at Rutgers University.

1 Introduction

In its weak form, Hilbert's Nullstellensatz states that a system

$$f_1(x) = 0, \ldots, f_s(x) = 0 \tag{1}$$

of polynomial equations in n unknowns has no solution over $\mathbb{C}$ if and only if there are polynomials $g_1, \ldots, g_s \in \mathbb{C}[X_1, \ldots, X_n]$ such that $\sum_{i=1}^{s} f_i g_i = 1$. For this reason, the problem of deciding whether (1) is satisfiable has also been called *Hilbert's Nullstellensatz* (HN in the sequel). This problem (and similar problems over the reals) has generated a lot of interest due to its importance in algebraic geometry and its potential applications. For instance, it is the basic step in algorithms for the decision and quantifier elimination problems in the first-order theory of $\mathbb{C}$. This opens up applications in, e.g., geometric theorem proving and robot motion planning [6]. Note also that if the f_i's can have arbitrary complex coefficients, HN is the canonical NP-complete problem in the Blum-Shub-Smale model of computation [4].

In this paper we consider systems of equations with integral coefficients only, i.e., we assume that $f_1, \ldots, f_s \in \mathbb{Z}[X_1, \ldots, X_n]$. Our model of computation is the Turing machine model. It is a folklore result that HN is NP-hard (there is a simple reduction from the satisfiability of boolean formulas in conjunctive normal form to HN). In terms of general complexity classes, the best upper bound previously known was PSPACE. Here we show that HN is in the second level of the polynomial hierarchy (in Π_2). In fact, we show that HN is in $\mathrm{RP}^{\mathrm{NP}}$, where RP is the class of problems solved by polynomial-time probabilistic Turing machines with one-sided error (i.e., the algorithm always gives a correct answer for a positive instance of the problem, and for negative instances the probability of answering incorrectly is bounded by some constant $\epsilon < 1$). This is a stronger result since $\mathrm{RP}^{\mathrm{NP}} \subseteq \Pi_2$. Some background on these complexity classes can be found in, e.g., [3]. The above-mentioned algorithm uses randomization and non-determinism *in this order only.* This means that an even stronger statement is true: Hilbert's Nullstellensatz is in AM, the "Arthur-Merlin" class [2] (by definition of AM). This result is almost as good as possible since AM is considered to be "just above" NP (see for instance the sequence of inclusions in the abstract). Note also that from the point of view of non-uniform complexity, the problem of characterizing the difficulty of HN is completely settled since HN is NP-hard on the one hand, and on the other hand $\mathrm{HN} \in \mathrm{AM} \subseteq \mathrm{NP/poly}$.

Previous algorithms were based on a direct application of effective Nullstellensätze: if one knows an upper bound on the degree of the g_i's then the satisfiability of (1) is equivalent to the satisfiability of a certain linear system (of exponential size). Our algorithm is fundamentally different. The Nullstellensatz is used only indirectly, to give a correctness proof. The basic idea is to study the satisfiability of (1) in $\mathbb{Z}/p\mathbb{Z}$. If the system is satisfiable modulo p for "many" primes p then it is satisfiable in $\mathbb{C}$; otherwise it is not satisfiable. This is established in section 4. We need a certain result on the roots of univariate polynomials which is established in section 5. Its proof requires the Generalized Riemman Hypothesis (GRH). Assuming these properties, we show in section 2 that HN is in Π_2. This requires some complexity-theoretic techniques (approximation of #P functions).

Of course now the tantalizing question is whether HN $\in$ NP. (to the author's knowledge, this is an open problem even for sparse univariate polynomials.) The modular techniques used in this paper suggest a natural approach to this question. Unfortunately, as shown in section 6 this approach cannot establish that HN $\in$ NP, even for sparse univariate polynomials. In that section we also give some evidence that HN really is unlikely to be in NP.

1.1 Notations

Let S be a system of the form (1), where the f_i's have degree $d_i \leq d$ and coefficients of size at most L. (the size of an integer a is $\log|a|$;[1] by convention the size of 0 is 0.) By definition, the total degree of S is $\sigma = 2 + \sum_{i=1}^{s} d_i$. The size of this system is the bit size of a representation of S in a suitable binary encoding scheme. In this paper we use a sparse representation. This means that we do not charge for monomials with a coefficient equal to 0. Sometimes the opposite choice is made (see e.g. [9, 11]). With that dense representation, the decision algorithms mentioned above are downsized from PSPACE to LOGSPACE. For univariate systems, there is an equally important representational issue, namely, whether exponents are coded in unary or binary. If binary (or *sparse*) encoding is used then polynomials can have exponential degree in the system's size (for instance, it takes about n bits to code the monomial x^{2^n}). This is not an important issue for us because one can always represent polynomials of exponential degree by introducing intermediate variables and using "repeated squaring". In fact, one could assume that the f_i's are of degree $d_i \leq 2$ and have all their coefficients in $\{-2, \ldots, 2\}$ without loss of generality (i.e., the general case of HN is polynomial-time

[1] Throughout the paper log stands for $\log_2$.

many-one reducible to this special case).

As usual we denote by $\pi(x)$ the number of primes in $\{2, \ldots, x\}$. Let R_S be the set of prime numbers such that S is satisfiable in $F_p = \mathbb{Z}/p\mathbb{Z}$. $\pi_S(x)$ denotes the cardinality of the set $R_S(x) = R_S \cap \{1, 2, \ldots, x\}$. Given a polynomial $f \in \mathbb{Z}[X]$, we use the abbreviations R_f, $R_f(x)$ and $\pi_f(x)$ for $R_{\{f=0\}}$, $R_{\{f=0\}}(x)$ and $\pi_{\{f=0\}}(x)$.

2 Position in the Polynomial Hierarchy

Theorem 1 is the crucial fact which makes it possible to locate HN in the polynomial hierarchy.

Theorem 1 *There exist constants $c_1, c_2, c_3 \in \mathbb{N}$ such that if $A = d^{c_1 n} s(\lceil \log s \rceil + L)$ and $x_0 \geq L^{c_2} 2^{(n \log \sigma)^{c_3}}$ the following two properties hold:*

- *If (1) is not satisfiable in $\mathbb{C}$ then $\pi_S(x_0) \leq A$.*
- *If (1) is satisfiable in $\mathbb{C}$ then $\pi_S(x_0) \geq B = 8A(\log A + 3)$.*

Proof. In Theorem 5 we show that if S is not satisfiable then there are at most $d^{c_1 n} s(\log s + L)$ primes p such that S is satisfiable in F_p. Hence the bound $\pi_s(x_0) \leq A$ holds for every x_0.

In Theorem 8, we show that there are absolute constants c_4, c_5, c_6 such that

$$\pi_S(x) \geq \frac{\pi(x) - c_4 n x^{1/2} \log x}{L.2^{(n \log \sigma)^{c_5}}} - L.2^{(n \log \sigma)^{c_6}} x^{1/2}$$

if S is satisfiable. The result follows from the prime number theorem: $\pi(x) \sim x/\ln x$. □

The rationale for the setting of B in this theorem will become clear in the proof of Theorem 2. It is already clear that HN is in $\mathrm{P}^{\#\mathrm{P}^{\mathrm{NP}}}$: in order to decide whether S is satisfiable we just have to compute $\pi_S(x_0)$ (note that Theorem 1 provides a bound on the size of x_0 which is polynomial in the system's size). If $\pi_S(x_0) \leq A$ then S is not satisfiable, otherwise it is satisfiable. This counting problem can be solved in #P with the help of an oracle in NP. This oracle decides, given an integer p, whether p is prime and S is satisfiable modulo p. The first task is feasible since the set of prime numbers is known to be in NP [17], and for the second one we just have to guess a solution.

In fact, we do not need exact counting since there is a large gap between A and B. Stockmeyer [21] has shown that approximate counting can be performed in the polynomial hierarchy (in Δ_3), and this result relativizes to an arbitrary oracle. Hence HN is in the polynomial hierarchy. By taking a closer look at Stockmeyer's argument, one can prove the following result.

Theorem 2 *Hilbert's Nullstellensatz is in* AM *(and therefore in* Π_2*).*

Proof. As mentioned before, the second part of the claim follows from AM $\subseteq$ Π_2. However, for the clarity of exposition, we will show first that HN $\in \Pi_2$.

Stockmeyer's result is based on a lemma of Sipser [20] on universal hashing: there exists a Σ_2 predicate $\mathrm{Hash}(E, m)$ which has the following property. If a set $E \subseteq \{0,1\}^k$ has at most 2^{m-2} elements then $\mathrm{Hash}(E, m)$ is true; but if $|E| \geq m2^m$, $\mathrm{Hash}(E, m)$ is false. The Π_2 predicate $\neg\mathrm{Hash}(E, m)$, which expresses that E is *not* hashable into $\{0,1\}^m$, has the form:

$$\forall f_1 \dots f_m \; C(f_1, \dots, f_m) \tag{2}$$

where $C(f_1, \dots, f_m)$ is the predicate:

$$\exists x, x_1, \dots, x_m \in E \; \bigwedge_{i=1}^{m} [f_i(x) = f_i(x_i) \wedge x \neq x_i].$$

Here $f_i : \{0,1\}^k \to \{0,1\}^m$ is a hash function, i.e., a function of the form $f_i(x) = A_i x$ where A_i is a binary matrix and arithmetic is performed modulo 2 (actually the exact form of f_i does not really matter for our purposes).

We are going to apply this result to $E = R_S(x_0)$ (given x_0, one can easily code the elements of $R_S(x_0)$ by binary strings of equal length k). The membership of $x, x_1, \dots, x_m$ in E can be expressed by Σ_1 predicates. When these predicates are substituted in (2) the blocks of existential quantifiers can be merged, and we still have a Π_2 predicate. This predicate will be of polynomial size by Theorem 1. The satisfaction of $\neg\mathrm{Hash}(E, m)$ will be equivalent to the satisfiability of S if $A \leq 2^{m-2} \leq m2^m \leq B$. Let m be the unique integer such that $A \leq 2^{m-2} < 2A$: since $m < \log A + 3$, the condition $m2^m \leq B$ is satisfied with the choice of B made in Theorem 1.

To see that HN is in AM, consider a variation of this Π_2 algorithm where the matrices $A_1, \dots, A_m$ are randomly chosen. It follows from the analysis above that this randomized algorithm always gives a correct answer for satisfiable systems. On the other hand, Sipser's lemma [20] states that, if $|E| \leq 2^{m-2}$ and $A_1, \dots, A_m$ are chosen at random, the probability of collision is at most $1/2$ ("collision" meaning that $C(f_1, \ldots, f_m)$ is true). Hence this randomized algorithm fails with probability at most $1/2$ for unsatisfiable systems. Since randomization and non-determinism are used in this order only, this is in fact an AM algorithm. $\Box$

At first glance one might think that a slightly stronger statement than just HN $\in$ AM holds since our algorithm has one-sided error (it always gives the correct answer for satisfiable systems), and the definition of AM allows for two-sided error. However, it turns out that the one-sided and two-sided versions of AM are identical ([23], Theorem 2).

One can ask whether there is a similar algorithm which always gives the correct answer for *unsatisfiable* systems, i.e., in the terminology of [10], whether HN has a *perfectly sound* Arthur-Martin proof system (in the same terminology, the algorithm of Theorem 2 is a *perfectly complete* Arthur-Martin proof system). This is possible if and only if HN $\in$ NP since perfectly sound Arthur-Martin proof systems exist only for languages in NP [10].

Another question is whether HN $\in$ coAM. This is quite unlikely since by the NP-hardness of HN, coNP would then be included in AM. It is known [5, 2] that coNP $\subseteq$ AM implies the collapse of the polynomial hierarchy to its second level ($\Sigma_2 = \Pi_2 =$ AM).

There is a more direct proof that HN $\in$ AM. The corresponding algorithm does not use hash functions, and is based instead on the fact that for a satisfiable system S and any integer x, $R_S \cap [x, x+L[$ must be non-empty if L is large enough. (In Theorem 2, we have used intervals starting at $x = 2$ only.) More precisely, given a system S, let $(x_n)_{n \in \mathbb{N}}$ be a sequence of integers which grows fast enough to insure that if S is satisfiable, $R_S \cap [x_n, x_{n+1}[\neq \emptyset$ for any $n \in \mathbb{N}$. It can be shown using Theorem 9 that x_n needs to grow only polynomially. The algorithm picks an integer $k < 2A$ at random (where A is the constant of Theorem 1), and accepts if $R_S \cap [x_k, x_{k+1}[\neq \emptyset$ (this condition can be checked non-deterministically). By definition of the sequence $(x_n)_{n \in \mathbb{N}}$ this algorithm never fails if S is satisfiable. If S is not satisfiable, $|R_S| \leq A$ by Theorem 1. Hence the probability of failure is at most $1/2$.

3 The Complexity of Primitive Elements

This section is of a technical nature. We establish some bounds on the complexity of primitive elements for use in section 4.

The following result is an effective version of the primitive element theorem. Recall that the norm of a polynomial $P = \sum_{k=0}^{d} a_k X^k \in \mathbb{Z}[X]$ is

$N(P) = (\sum_{k=0}^{d} a_k^2)^{1/2}$.

Theorem 3 *Let α_1 and α_2 be roots of two squarefree polynomials $P_1, P_2 \in \mathbb{Z}[X]$ of degree d_1 and d_2, and maximum norm N. Let $d = \max\{d_1, d_2\}$. There exists a squarefree polynomial $R \in \mathbb{Z}[X]$ of degree at most $n_1 n_2$ and a root β of R (called a* primitive element*) such that $\alpha_i = Q_i(\beta)/a_i$ $(i = 1, 2)$ where $Q_i \in \mathbb{Z}[X]$ and $a_i \in \mathbb{Z}$, $|a_i| \leq c.N^{2d}$. Here c is a universal constant,* $\deg(Q_i) < n_1 n_2$ *and $N(R) \leq c.N^{2d}$.*

In fact R depends only on P_1 and P_2. See [7, 15] for proofs of this result. We need a generalization to several polynomials. This is tedious but straightforward.

Lemma 1 *Let $\alpha_1, \ldots, \alpha_n$ be roots of n squarefree polynomials $P_1, \ldots, P_n \in \mathbb{Z}[X]$ of degree $2 \leq d_i \leq d$, and norm $N(P_i) \leq N$. There exists a squarefree polynomial $R_n \in \mathbb{Z}[X]$ of degree at most d^n and a root β_n of R_n such that $\alpha_i = Q_{in}(\beta_n)/a_{in}$, where $Q_{in} \in \mathbb{Z}[X]$ and $a_{in} \in \mathbb{Z}$, $|a_{in}| = N^{d^{O(n^2)}}$. Here c' is a universal constant,* $\deg(Q_{in}) < \sum_{i=2}^{n} d^i$ *and*

$$N(R_n) \leq c'.N^{2^{n-1}.d^{n(n-1)/2}} \tag{3}$$

Proof. By induction. Let $R_{n-1} \in \mathbb{Z}[X]$ be a polynomial of degree at most d^{n-1} such that $\alpha_1, \ldots, \alpha_{n-1}$ can be represented as $\alpha_i = Q_{i,n-1}(\beta_{n-1})/a_{i,n-1}$ where $R_{n-1}(\beta_{n-1}) = 0$, $Q_{i,n-1} \in \mathbb{Z}[X]$ and $a_{i,n-1} \in \mathbb{Z}$. In order to obtain β_n, we can apply Theorem 3 to β_{n-1} and α_n: there exists $R_n \in \mathbb{Z}[X]$ and a root β_n of R_n such that each element $\gamma \in \{\beta_{n-1}, \alpha_n\}$ can be expressed as $\gamma = Q(\beta_n)/a$ where $a \in \mathbb{Z}$ and $Q \in \mathbb{Z}[X]$. Each α_i $(i = 1, \ldots, n-1)$ can now be expressed as

$$\alpha_i = Q_{i,n-1}(Q(\beta_n)/a)/a_{i,n-1}. \tag{4}$$

Let $N_{n-1} \geq N$ be (an upper bound on) the norm of R_{n-1} and N_n the norm of R_n. By Theorem 3, $N_n \leq c.N_{n-1}^{2d^{n-1}}$. Equation (3) follows from this inductive relation.

By induction hypothesis, $\deg(Q_{i,n-1}) < \sum_{i=2}^{n-1} d^i$. Since $\deg(Q) < d^n$, using (4) one can write $\alpha_i = Q_{in}(\beta_n)/a_{in}$ with $\deg(Q_{in}) < \sum_{i=2}^{n} d^i$ and

$$a_{in} = a^{\deg(Q_{i,n-1})} a_{i,n-1}.$$

The bound on $|a_{in}|$ in the Lemma's statement follows from this inductive relation and the bound $|a| \leq c.N_{n-1}^{2d^{n-1}}$ (Theorem 2). $\square$

The degree of Q_{in} can be reduced to at most $d^n - 1$ by computing $\mathrm{rem}(Q_{in}, R_n)$ (this will slightly increase a_{in}). One can give somewhat better bounds if instead of treating the list of α_i's iteratively, the primitive element is computed by "divide-and-conquer" (i.e., if the first and second half of the list are processed separately, and the 2 primitive elements are put together at the end).

Theorem 4 *Let $x_1, \ldots, x_n$ be n algebraic numbers which are roots of polynomials $A_i \in \mathbb{Z}[X]$ of degree at most d with coefficients of size at most L. There exists a primitive element r for $x_1, \ldots, x_n$ which is a root of an irreducible polynomial $B \in \mathbb{Z}[X]$ of degree at most d^n. The coefficients of B are of size at most $L.d^{n^{O(1)}}$. Moreover, each x_i can be represented as $x_i = Q_i(r)/a_i$ where $Q_i \in \mathbb{Z}[X]$ and $\log|a_i| = L.d^{n^{O(1)}}$.*

Proof. The primitive element can be obtained in three steps:

1. Let A_i' be the derivative of A_i. Make the A_i's squarefree by computing $P_i = A_i/\gcd(A_i, A_i')$.

2. Apply Lemma 1 to $P_1, \ldots, P_n$; this gives a_i, Q_i, and a polynomial $R \in \mathbb{Z}[X]$.

3. The primitive element r is a root of R. So B is an irreducible factor of R.

The stated bounds follow from Lemma 3. Indeed, steps 1 and 3 are very cheap compared to step 2. This follows from the bound on polynomial factors in [16]. □

4 Solutions modulo p

One can take care easily of unsatisfiable systems with the effective Hilbert Nullstellensatz. The case of satisfiable systems is more involved, and requires the Generalized Riemman Hypothesis.

4.1 Unsatisfiable Systems

Since the f_i's have integer coefficients, if (1) has no solution in $\mathbb{C}$ then by Hilbert's Nullstellensatz there exist $a \in \mathbb{Z}$, $a \neq 0$ and polynomials $g_1, \ldots, g_s \in \mathbb{Z}[X_1, \ldots, X_n]$ such that

$$a = g_1 f_1 + \cdots + g_s f_s. \tag{5}$$

A bound on the size of a can be easily obtained if one has an a-priori bound on the degrees D_i of the g_i's. It is shown in [12] that one can take $D_i = \max\{3, d\}^n$. This leads to a $s^{O(1)}d^{O(n^2)}L$ bound on the size a. This bound was improved to $d^{O(n)}s(\log s + L)$ in [13]. Theorem 5 follows easily (one could also use the simpler $d^{O(n^2)}$ bound).

Theorem 5 *If (1) has no solution in $\mathbb{C}$ then R_S is finite and $|R_S| \leq d^{O(n)}s(\log s + L)$.*

Recall that R_S is the set of prime numbers p such that (1) has a solution in $\mathbb{Z}/p\mathbb{Z}$.

Proof. If (1) has no solution in $\mathbb{C}$ then (5) holds in $\mathbb{C}$ and thus also in $\mathbb{Z}/p\mathbb{Z}$. This implies that (1) has no solution in $\mathbb{Z}/p\mathbb{Z}$ if $a \bmod p \neq 0$. The result follows since a cannot have more than $\log a$ prime factors. $\square$

4.2 Satisfiable Systems

The first step toward the analysis of satisfiable systems will be to show (in Theorem 7) that these systems have algebraic solutions which have "short" descriptions. For this we need the following theorem. It follows from the quantifier elimination results in [8].

Theorem 6 *Let Φ a prenex formula in the first-order theory of $\mathbb{C}$. Let r be the number of quantifier blocks, n the total number of variables, and $\sigma(\Phi)$ the total degree of Φ, defined as:*

$$\sigma(\Phi) = 2 + \sum_{i=1}^{s} \deg\ F_i$$

where $F_1, \ldots, F_s$ are the polynomials occurring in Φ. Φ is equivalent to a quantifier-free formula Ψ in which all polynomials have degree at most

$$2^{n^{O(r)}(\log \sigma(\Phi))^{O(1)}}.$$

The number of polynomials occurring in Ψ is $O(\sigma(\Phi)^{n^{O(r)}})$.

Moreover, when the constants in Φ are integers of bit-size at most L, the constants in Ψ are integers of bit size at most $L.2^{n^{O(r)}(\log \sigma(\Phi))^{O(1)}}$.

Theorem 7 *There are absolute constants c_1 and c_2 such that if (1) has a solution over $\mathbb{C}$ then there exists a solution $x = (x_1, \ldots, x_n)$ such that each x_i is a root of a polynomial of degree at most $2^{(n \log \sigma)^{c_1}}$ with coefficients of bit size at most $L.2^{(n \log \sigma)^{c_2}}$.*

We shall first prove this result in a special case.

Lemma 2 *Theorem 7 holds for systems that have a finite number of solutions.*

Proof. Let S be the solution set of (1) and $S_i \subseteq \mathbb{C}$ the projection of S on the i-th coordinate axis. By Theorem 6, S_i can be defined by a quantifier-free formula in which polynomials $P_{i1}, \ldots, P_{im_i}$ of degree at most $2^{(n \log \sigma)^{c'_1}}$ and coefficients of size at most $L.2^{(n \log \sigma)^{c'_2}}$ appear (moreover, $m_i = \sigma^{n^{O(1)}}$; c'_1 and c'_2 are absolute constants). If S is finite then each S_i is finite. Hence each element of S_i is a root of some P_{ij}. The result follows since by definition the components of any solution $x \in S$ must be in $S_1, \ldots, S_n$. $\Box$

Proof of Theorem 7. By induction on n; the constants will satisfy $c_1 = c'_1$ and $c_2 \geq c'_2$. If (1) has finitely many solutions the result holds by Lemma 2.

Assume now that (1) has infinitely many solutions. Then at least one S_i must be infinite. This implies that $\mathbb{C} \backslash S_i$ is finite, and that its elements are chosen among the roots of $m_i = \sigma^{n^{O(1)}}$ polynomials of degree at most $2^{(n \log \sigma)^{c_1}}$. Hence $|\mathbb{C} \backslash S_i| \leq 2^{(n \log \sigma)^{c_3}}$ for some absolute constant c_3. This guarantees the existence of an integer $\alpha \in S_i$ such that $0 \leq \alpha \leq 2^{(n \log \sigma)^{c_3}}$. Since α is of polynomial size, this integer can be substituted to x_i in (1) without blowing up the system's size too much. More precisely, we obtain a new satisfiable system in $n-1$ variables where the polynomials are of degree at most d and have coefficients of size at most $L + d \log \alpha \leq L + d(n \log \sigma)^{c_3}$. By induction hypothesis this system has a solution whose components are roots of polynomials of degree at most $2^{((n-1) \log \sigma)^{c_1}}$. They have coefficients of size bounded by

$$B = [L + d(n \log \sigma)^{c_3}] 2^{((n-1) \log \sigma)^{c_2}}.$$

For the induction hypothesis to hold in dimension n, we need to have $B \leq L.2^{(n \log \sigma)^{c_2}}$. Assuming without loss of generality that L and $d(n \log \sigma)^{c_3}$ are both larger than 2, we have:

$$B \leq L.2^{((n-1) \log \sigma)^{c_2} + \log[d(n \log \sigma)^{c_3}]}.$$

Recalling that $d \leq \sigma$, it suffices to have:

$$(n \log \sigma)^{c_2} - ((n-1) \log \sigma)^{c_2} \geq \log \sigma + c_3 \log(n \log \sigma).$$

It is not hard to see that this constraint is satisfied if the absolute constant c_2 is large enough. $\Box$

Lemma 3 *Let $x = (x_1, \ldots, x_n)$ be a vector of algebraic numbers solution of (1). Let r be a primitive element for $x_1, \ldots, x_n$: there exist polynomials $Q_1, \ldots, Q_n \in \mathbb{Z}[x]$ and $a \in \mathbb{N}$ such that $x_i = Q_i(r)/a$. Let $R \in \mathbb{Z}[X]$ be an irreducible polynomial such that $R(r) = 0$. If R has a root in F_p and $a \bmod p \neq 0$, (1) is satisfiable in F_p.*

Proof. For $i \in \{1, \ldots, s\}$, let

$$g_i(X) = a^{d_i} f_i(Q_1(X)/a, \ldots, Q_n(X)/a) \tag{6}$$

Note that $g_i \in \mathbb{Z}[X]$. These polynomials must be multiple of R since R is irreducible and $g_i(r) = 0$. Hence there are polynomials $A_1, \ldots, A_s \in \mathbb{Z}[X]$ such that

$$g_i(X) = R(X)A_i(X). \tag{7}$$

If $a \bmod p \neq 0$, (6) and (7) must also hold in F_p. This implies that if x_0 is a root of R in F_p, $(Q_1(x_0)/a, \ldots, Q_n(x_0)/a)$ is a solution of (1) in F_p. $\square$

Theorem 8 *There are absolute constants c_4, c_5, c_6 such that if (1) is satisfiable,*

$$\pi_S(x) \geq \frac{\pi(x) - c_6 n x^{1/2} \log x}{L.2^{(n \log \sigma)^{c_4}}} - L.2^{(n \log \sigma)^{c_5}} x^{1/2}.$$

Proof. We just need explicit estimates on a and the Q_i's in order to apply Lemma 3. The algebraic numbers $x_1, \ldots, x_n$ are roots of polynomials $P_1, \ldots, P_n$ whose degree and coefficient size can be bounded by Theorem 7. The complexity of the primitive element r can then be estimated by Theorem 4. The number of primes p in $\pi_R(x)$ can be estimated by Corollary 1 (here we use the fact that the bit size of the discriminant of a polynomial is polynomially bounded in its degree and the bit size of its coefficients). ¿From this estimate one has to subtract the prime factors of a; there are at most $\log a$ such primes. $\square$

5 Roots of univariate polynomials

Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree n; let Δ be the discriminant of f. For a prime p,

$$W(p) = |\{k;\ 0 \leq k \leq p-1, f(k) \equiv 0 \pmod{p}\}|$$

denotes the number of roots of f in F_p. Let $S(x) = \sum'_{p \leq x}(1 - W(p))$ where Σ' denotes summation on those primes p which do not divide Δ. The following bound is due to Adleman and Odlyzko ([1], proof of Lemma 3); see also [22].

Theorem 9 $|S(x)| = O(x^{1/2}\log(\Delta x^n))$.

This result relies on GRH through an effective version of the Chebotarev density theorem [14] (in fact, only the effective prime ideal theorem is used). A bound on $\pi_f(x)$ follows easily.

Corollary 1 *There exists an absolute constant c such that:*

$$\pi_f(x) \geq \frac{1}{n}[\pi(x) - \log\Delta - c.x^{1/2}\log(\Delta x^n)].$$

Proof. By Theorem 9,

$$\sum_{p\leq x}{}' W(p) \geq \sum_{p\leq x}{}' 1 - c.x^{1/2}\log(\Delta x^n) \tag{8}$$

for some universal constant c. Since Δ has at most $\log\Delta$ prime factors, $\sum'_{p\leq x} 1 \geq \pi(x) - \log\Delta$. Let $r_f(p) = 1$ if f has a root in F_p, and $r_f(p) = 0$ otherwise. Since f is irreducible in $\mathbb{Z}[X]$ this polynomial cannot be identically 0 in F_p; hence $W(p) \leq nr(p)$. The result now follows from (8) since $\pi_f(x) = \sum_{p\leq x} r_f(p) \geq \sum_{p\leq x} W(p)/n$. $\square$

Remark 1 Corollary 1 provides a lower bound of $1/n$ on the density of R_f in the set of prime numbers. In fact, by the Chebotarev density theorem, the exact value of the density is $|C|/|G|$, where G is the Galois group of f and C is the set of permutations $g \in G$ which have at least one fixed point. It can be shown that $|C| = 1$ when $|G| = n$.[2] Hence the $1/n$ lower bound cannot be improved in the worst case. However, "generic" polynomials with large coefficients have the full symmetric group as Galois group. In this case, $|C|/|G| \geq 1/2$ (and $|C|/|G| \simeq 1 - 1/e \simeq 0.63$ for polynomials of high degree which have S_n as Galois group).

Theorem 8 provides a lower bound on the density of R_S which can be exponentially small in the size of a satisfiable system. The following example shows that in some cases the density of R_S can really be exponentially small (however, by Remark 1 one can expect its density to be at least 1/2 for most satisfiable systems).

[2]For any transitive subgroup G of the symmetric group (and in particular for the Galois group of an irreducible polynomial) and any $i \in \{1,\ldots,n\}$, there are exactly $|G|/n$ permutations $g \in G$ such that $g(i) = i$. Hence only the identity can have fixed points when $|G| = n$.

Example 1 *Let S_p be the system*

$$\begin{cases} x^p = 1 \\ y(x-1) = 1 \end{cases}$$

where p is a prime number. The density of R_{S_p} is $1/(p-1)$.

The $p-1$ solutions of S_p over $\mathbb{C}$ are of the form $(x, 1/(x-1))$ where x is a p-th root of unity different from 1. One can write $X^p - 1 = (X-1)C_p(X)$ where $C_p(X) = \sum_{k=0}^{p-1} X^k$ is the cyclotomic polynomial of order p. Given a prime q, S_p is satisfiable in F_q if and only if C_p has a root different from 1 in F_q. There is a single value of q ($q = p$) for which $C_p(1) = 0$ in F_q. It is well-known that C_p is irreducible over $\mathbb{Z}$, and that its Galois group is (isomorphic to) $\mathbb{Z}/(p-1)\mathbb{Z}$. Hence by Remark 1, the density of R_{S_p} is indeed $1/(p-1)$.

This density is exponentially small in the size of S_p, which is $O(\log p)$. Recall also that, as pointed out in section 1.1, S_p can be replaced by an equivalent system S'_p of size $O(\log p)$ where all polynomials have degree 2. Since $R_{S'_p} = R_{S_p}$, the density of $R_{S'_p}$ is again exponentially small in the system's size.

6 Is the Nullstellensatz NP-complete ?

In Theorem 5 we have given an upper bound on the size of $|R_S|$ which may be exponential in the size $|S|$ of an unsatisfiable system. Improvements of this bound would be interested. In particular, a polynomial upper bound (say, $|S|^c$) would imply under GRH that HN $\in$ NP. (a certificate for a satisfiable system would simply consist of a list $p_1, \ldots, p_{|S|^c+1}$ of distinct primes of polynomial size such that S is satisfiable modulo p_i, together with a list of corresponding solutions; the existence of such primes follows from Theorem 8.) Unfortunately, there is no such polynomial upper bound. We shall see in Theorem 10 that $|R_S|$ can be exponential in $|S|$. First, we need the following result, which is due to Noam Elkies (personal communication).

Lemma 4 *Let π_n be the product of the first n prime numbers: for $n \geq 2$, $\pi_n^{\pi_n} - 1$ has at least 2^n distinct prime factors.*

Proof. The factorization of the polynomial $x^N - 1$ ($N \geq 2$) over the integers is:

$$x^N - 1 = \prod_{d|N} \Psi_d(x) \tag{9}$$

where

$$\Psi_d(x) = \prod_{1 \le k \le d, \gcd(k,d)=1} (x - e^{2ik\pi/d}) \tag{10}$$

is the cyclotomic polynomial of order d. These two propeties of cyclotomic polynomials will be useful.

1. If a prime number p does not divide N then the polynomials Ψ_d in (9) have pairwise distinct roots in F_p. Indeed, if Ψ_d and $\Psi_{d'}$ had a common root r in F_p, the derivative of $x^N - 1$ would vanish at r. However the derivative (Nx^{N-1}) has no roots in F_p since $N \bmod p \neq 0$.

2. $\Psi_d(x) \ge 2$ for any integer $x \ge 3$. Indeed, each factor $(x - e^{2ik\pi/d})$ in (10) has modulus larger than 1. Hence $|\Psi_d(x)| > 1$, and in fact $\Psi_d(x) \ge 2$ since Ψ_d has integer coefficients.

Let p be a prime factor of $\Psi_d(N)$, where $d|N$. Since p is *a fortiori* a prime factor of $N^N - 1$, p does not divide N (N and $N^N - 1$ are relatively prime). Hence by property 1, p is not a factor of any other $\Psi_{d'}(N)$ where $d'|N$. Setting $N = \pi_n$ gives the desired result since in this case there are 2^n factors in (9). (Each factor $\Psi_d(N)$ gives at least one distinct prime factor since $\Psi_d(N) \ge 2$ by property 2.) □

The motivation for this lemma came from a conjecture of Shub and Smale [19] on the length of computations for $k!$. The "naive" method for computing $k!$ requires $\Theta(k \log k)$ operations (additions and multiplications; subtractions are also allowed). Elkies has pointed out that a version of the elliptic curve factoring method suggests that there should exist computations of length growing slower than any power of k, indeed no faster than $\exp[\log(k)^{1/2+\epsilon}]$.

Theorem 10 *For $n \ge 1$, let S_n be the system:*

$$\begin{cases} x^{\pi_n} - 1 = 0 \\ x - \pi_n = 0 \end{cases}$$

where π_n is the product of the first n prime numbers. This system is unsatisfiable over $\mathbb{C}$, and $|R_{S_n}| \ge 2^n$.

Proof. S_n is obviously unsatisfiable over $\mathbb{C}$. The prime factors of π_n are in R_{S_n}, and by Lemma 4 there are at least 2^n of them. □

By the prime number theorem, the n-th prime is $O(n \log n)$. Hence the bit size of π_n is also $O(n \log n)$. Therefore $|R_{S_n}|$ is exponential in $|S_n|$.

This result shows that a particular method fails to establish that HN $\in$ NP. In the rest of this section we would like to give some evidence that HN is unlikely to be in NP (or at least that proving that HN $\in$ NP may be quite difficult). The argument is based on the observation that HN is at least as hard as ACVP, the Arithmetic Circuit Value Problem. An instance of this problem is a pair (C, x) where $x \in \{0,1\}^n$ and C is a circuit with n inputs made of addition, subtraction, multiplication, and sign gates (a sign gates returns 1 for a non-zero input, and 0 otherwise). An instance is positive if C outputs a non-zero value on input x. This problem occurs in the simulation of Random Access Machines on Turing machines [18]. It is well-known that ACVP is in BPP; however, this problem is not known to be in NP. It is not hard to see that ACVP is polynomial-time many-one reducible to HN. Hence HN $\in$ NP would imply ACVP $\in$ NP.

7 Final Remarks

Some of the techniques in this paper might be useful in practice. For instance, one can try and solve (1) modulo several randomly drawn primes. If there are "many" positive answers, then one can conclude that (1) is satisfiable over $\mathbb{C}$ with high probability. Precise bounds on the proportion of positive answers that should be obtained can be worked out using only the effective Nullstellensatz (see section 4.1). By Remark 1, one can expect that trying a "small" number of random primes will be enough to establish that a satisfiable system is indeed satisfiable (with high probability), if the defining polynomial R for the primitive element has the full symmetric group as Galois group (in this case, (1) is solvable modulo p for at least one half of all primes p). For other satisfiable systems, one may have to try an exponential number of primes. Another related shortcoming is that one cannot establish (at least if we try only a polynomial number of primes) that a system is *not* satisfiable, even in a probabilistic sense. The reason is again that the density of "good" primes for a satisfiable system may be exponentially small (see section 5). It would be interesting to have a rigorous average-case analysis of this method.

Acknowledgments

I would like to thank Andrew Odlyzko and Marc Perret for their help with the Chebotarev density theorem. Thanks also to Mike Shub and Noam Elkies for useful discussions on the number of prime factors. Stockmeyer's paper was pointed out by Stéphane Boucheron. Marcos Kiwi suggested several improvements in the presentation of this paper.

References

[1] L. M. Adleman and A. M. Odlyzko. Irreducibility testing and factorization of polynomials. *Mathematics of Computation*, 41(164):699–709, 1983.

[2] L. Babai and S. Moran. Arthur-Merlin games: A randomized proof system, and a hierarchy of complexity classes. *Journal of Computer and System Sciences*, 36:254–276, 1988.

[3] J.L. Balcázar, J. Diáz, and J. Gabarró. *Structural Complexity I.* EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1988.

[4] 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 American Mathematical Society*, 21(1):1–46, July 1989.

[5] R. Boppana, J. Hastad, and S. Zachos. Does co-NP have short interactive proofs? *Information Processing Letters*, 25:127–132, 1987.

[6] J. Canny. *The Complexity of Robot Motion Planning.* MIT Press, Cambridge, Massachusetts, 1988.

[7] J. Canny. Some algebraic and geometric computations in PSPACE. In *Proc. 20th ACM Symposium on Theory of Computing*, pages 460–467, 1988.

[8] N. Fichtas, A. Galligo, and J. Morgenstern. Precise sequential and parallel complexity bounds for quantifier elimination over algebraically closed fields. *Journal of Pure and Applied Algebra*, 67:1–14, 1990.

[9] M. Giusti and J. Heintz. La détermination des points isolés et la dimension d'une variété algébrique peut se faire en temps polynomial. In *Computational Algebraic Geometry and Commutative Algebra (Cortona, 1991)*, pages 216–256. Sympos. Math. XXXIV, Cambridge University Press, 1993.

[10] O. Goldreich, Y. Mansour, and M. Sipser. Interactive proof systems: Provers that never fail and random selection. In *Proc. 28th IEEE Symposium on Foundations of Computer Science*, pages 449–461, 1987.

[11] J. Heintz and J. Morgenstern. On the intrinsic complexity of elimination theory. *Journal of Complexity*, 9:471–498, 1993. preprint: Technical Report 93-17, Laboratoire I3S, Université de Nice – Sophia Antipolis.

[12] J. Kollár. Sharp effective Nullstellensatz. *Journal of the AMS*, 1:963–975, 1988.

[13] T. Krick and L. M. Pardo. Une approche informatique pour l'approximation diophantienne (A computational approach to diophantine approximation). *Compte-Rendus de l'Académie des Sciences, Série I*, 318(5):407–412, 1994.

[14] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. In A. Frölich, editor, *Algebraic Number Fields*, pages 409–464. Academic Press, 1977.

[15] R. Loos. Computing in algebraic extensions. In Buchberger, Collins, and Loos, editors, *Computer Algebra, Symbolic and Algebraic Computation*, pages 173–187. Springer-Verlag, 1982.

[16] M. Mignotte. Some useful bounds. In Buchberger, Collins, Loos, and Albrecht, editors, *Computer Algebra – Symbolic and Algebraic Computation*, pages 259–263. Springer-Verlag, 1982.

[17] V. Pratt. Every prime has a succint certificate. *SIAM Journal on Computing*, 4(3):214–220, 1975.

[18] A. Schönhage. On the power of random access machines. In *Proc. 6th ICALP*, volume 71 of *Lecture Notes in Computer Science*. Springer-Verlag, 1979.

[19] M. Shub and S. Smale. On the intractability of Hilbert's Nullstellensatz and an algebraic version of "P=NP". To appear in *Duke Journal of Mathematics*.

[20] M. Sipser. A complexity theoretic approach to randomness. In *Proc. 15th ACM Symposium on Theory of Computing*, pages 330–335, 1983.

[21] Larry Stockmeyer. On approximation algorithms for #P. *SIAM Journal on Computing*, 14(4):849–861, 1985.

[22] P. J. Weinberger. Finding the number of factors of a polynomial. *Journal of Algorithms*, 5:180–186, 1984.

[23] S. Zachos. Probabilistic quantifiers, adversaries, and complexity classes: an overview. In *Proc. 1st Structure in Complexity Theory Conference*, volume 223 of *Lecture Notes in Computer Science*. Springer-Verlag, 1986.