Complexity of Null- and Positivstellensatz Proofs

Dima Grigoriev*
Nicolai Vorobjov †

Dedicated to N. A. Shanin on the occasion of his 80th anniversary

Abstract

We introduce two versions of proof systems dealing with systems of inequalities: Positivstellensatz refutations and Positivstellensatz calculus. For both systems we prove the lower bounds on degrees and lengths of derivations for the example due to Lazard, Mora and Philippon. These bounds are sharp, as well as they are for the Nullstellensatz refutations and for the polynomial calculus. The bounds demonstrate a gap between the Null- and Positivstellensatz refutations on one hand, and the polynomial calculus and Positivstellensatz calculus on the other.

Introduction

In recent years there was an intensive activity in the research of algebraic proof systems ([BIK 96], [BuGI 99], [BuIK 96], [CEI 96], [G 98], [IPS 97]). The approach relies on the Hilbert's Nullstellensatz and treats the problem of feasibility of a system of polynomial equations

$$f_1 = \cdots = f_k = 0,$$

*IRMAR, Université de Rennes, Campus de Beaulieu, 35042 Rennes, cedex France
†Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, England, supported in part by EPSRC grant GR/L77928

with $f_1, \dots, f_k \in F[X_1, \dots, X_n]$, over the algebraic closure $\overline{F}$. Note that this problem is, in general, NP-hard.

The Nullstellensatz proof system (NS) was first considered in [BIK 96]. The aim of the system is to find the polynomials $g_1, \dots, g_k \in F[X_1, \dots, X_n]$ such that $1 = g_1 f_1 + \cdots + g_k f_k$. The latter representation is sometimes called a *Nullstellensatz refutation.* The number $\max_{1 \le i \le k}\{\deg(g_i f_i)\}$ is called the *Nullstellensatz degree.*

The first example of an exponential lower bound $d^{\Omega(n)}$, where $\deg(f_i) \le d$, for the Nullstellensatz degree is due to Lazard, Mora and Philippon (see [Br 87] and Section 1 below). In the proof system theory ([BIK 96], [BuIK 96], [IPS 97], [P 98]) a similar question of obtaining lower bounds on the Nullstellensatz degree was studied mostly for Boolean systems, in which the input polynomials $f_1, \dots, f_k$ include $X_i^2 - X_i$ for all $1 \le i \le n$. In Boolean case a linear upper bound $O(n)$ is evident, in [BIK 96] a non-constant lower bound was proved, while in [G 98] a *linear* (and thus sharp) lower bound was proved.

In [CEI 96] a stronger proof system — polynomial calculus (PC) was introduced. Starting from axioms $f_1, \dots, f_k$, PC allows to derive from the already obtained polynomials $a, b \in F[X_1, \dots, X_n]$ more polynomials, according to the following two rules:

1. (additive) $\quad a, b \vdash \alpha a + \beta b$, where $\alpha,\ \beta \in F$;

2. (multiplicative) $\quad a \vdash X_i a$ for $1 \le i \le n$.

The aim of a derivation is to reach 1.

The *degree* of a PC derivation is defined as the maximum of the degrees of all intermediately derived polynomials. The first lower bound on the degrees of PC derivations was obtained in [R 96] (see also [IPS 97] and [BuIK 96]). A linear lower bound for PC was proved in [BuGI 99]. Note that the latter bound is sharp.

The aim of the present paper is to involve inequalities along with equations into proof systems, in particular we assume that the input polynomials $f_1, \dots, f_k$ belong to $\mathbf{R}[X_1, \dots, X_n]$. The case of *linear* inequalities with added conditions $X_i^2 = X_i$ (Boolean programming) was widely studied by means of cutting planes proofs, for which an exponential lower bound on the length was obtained (a survey and references can be found in [P 98]). Another approach to systems of *linear* inequalities was undertaken in [LS 91], [L 94], [ST 98], where a derivation system was introduced which allows from *any* linear polynomial e, *already derived* linear inequalities $a_1 \geq 0, a_2 \geq 0$ and quadratic inequalities $p_1 \geq 0, p_2 \geq 0$, to derive quadratic inequalities $e^2 \geq 0$, $a_1 + a_2 \geq 0$, $a_1 a_2 \geq 0$, $p_1 + p_2 \geq 0$. In [P 98] one can find some remarks on the complexity of this Lovász-Schrijver procedure, in particular, an upper bound for the Pigeon Hole Principle which demonstrates an exponential gap between the complexity of cutting planes proofs and the Lovász-Schrijver procedure.

In the present paper we introduce two proof systems for dealing with systems of *polynomial* inequalities. While NS and PC are based on Hilbert's Nullstellensatz, our systems involve Positivstellensatz (see [BCR 98], [S 74] and Section 2 below). The first (weaker) proof system (Section 2), called *Positivstellensatz refutation*, is similar to NS and, as NS, can be viewed as a *static* proof system in the sense that we require to present a refutation as a formula containing all the information about the refutation, and the complexity is the size of this formula.

The second (stronger) system (Section 2) is similar to PC and, as PC, can be viewed as a *dynamic* system in the sense that a refutation is derived step by step. The complexity is measured by the size of the intermediate polynomials and by the number of steps, i.e., by the length of the derivation.

The main results (Section 3) provide lower bounds for the Lazard-Mora-Philippon example for the Positivstellensatz refutations (Theorem 2), and for Positivstellensatz calculus (Theorem 3). Both bounds are sharp and coincide with the corresponding upper bounds for NS (see [Br 87] and Section 1), and for PC (Theorem 1, Section 1). In particular, the results show that for the example considered in this paper, there is a gap between the degree of derivations for the Positivstellensatz refutations and NS on one hand, and for the Positivstellensatz calculus and PC on the other. Note that in [CEI 96], [Bu 99] there were constructed examples of Boolean systems of equations for which PC has smaller degrees of derivations than the ones of NS.

We also mention that in [LMR 96] a deductive system based on Positivstellensatz was exhibited, for describing an equivalence of two semialgebraic sets defined by systems of polynomial inequalities. In [CLR 99] a system is described for automatic extracting of some constructive consequences from non-constructive proofs of Null- and Positivstellensatz type. However, both papers do not consider complexity issues.

1 Complexity of polynomial calculus for a telescopic system

Let F be a field. A well-known example of a lower bound for the degrees in NS is due to Lazard, Mora and Philippon [Br 87]. It is called "telescopic system" and is defined by the following polynomials:

$$1 - YX_1, X_1^2 - X_2, X_2^2 - X_3, \ldots, X_{n-1}^2 - X_n, X_n \in F[Y, X_1, \ldots, X_n] \quad (1)$$

Polynomials (1) obviously don't have common roots, and its NS degree is bounded from below by 2^{n-1}. Observe that 2^{n-1} is close to the known upper bound for NS degree [Br 87]. The proof of the lower bound [Br 87] involves the following substitution of Laurent polynomials:

$$Y = X_1^{-1}, X_2 = X_1^2, X_3 = X_1^4, \ldots, X_n = X_1^{2^{n-1}}. \quad (2)$$

For any polynomial $f \in F[Y, X_1, \ldots, X_n]$ let $\overline{f}$ be the the Laurent polynomial being the result of the substitution (2) into f; by its *order* $\mathrm{ord}(f) \in \mathbf{Z}$ we mean the lowest (possibly negative) degree of the Laurent monomials occurring in $\overline{f}$.

Let

$$1 = (1 - YX_1)g_1 + (X_1^2 - X_2)g_2 + \cdots + (X_{n-1}^2 - X_n)g_n + X_n g$$

be a NS refutation for (1), where $g_1, \ldots, g_n, g \in F[Y, X_1, \ldots, X_n]$. Then $1 = X_1^{2^{n-1}}\overline{g}$, hence $\overline{g} = X_1^{-2^{n-1}}$, thus $\mathrm{ord}(g) = -2^{n-1}$. On the other hand, $\mathrm{ord}(g) \geq -\deg_Y(g)$ (see (2)). Therefore, $\deg_Y(g) \geq 2^{n-1}$. This proves the mentioned lower bound on the NS refutation degree for the system (1). Below, in Theorem 2 we'll prove a similar lower bound for a stronger proof system.

Now we prove an upper bound on the degree of PC refutation for (1). Recall that the *length* of a PC refutation is the number of applications of the rules of PC.

Theorem 1 *For the system (1) one can produce a PC refutation with the degree bounded by $O(n)$ and the length not exceeding $O(n2^n)$.*

Remark 1 *This bound is close to sharp. This will follow from the lower bound in Theorem 3 below.*

Proof. We start with the axiom X_n and derive on each step of the proof a certain monomial of one of the two types:

(i) $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_{l_k}$ or

(ii) $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_{l_k}^2$,

where $n > l_1 > l_2 > \cdots > l_k \geq 1$.

If $l_k > 1$ (we call it a *regular step*), then in case (i) we derive the monomial $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_{l_k-1}^2$ invoking the axiom $X_{l_k} - X_{l_k-1}^2$, and applying $k-1$ times the multiplication rule of PC and, subsequently, once the addition rule. In case (ii) we derive the monomial $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_{l_k}X_{l_k-1}^2$ in a similar way using the same axiom.

If $l_k = 1$ (we call it a *singular step*), then in case (i) we derive first $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_1Y$ and then $X_{l_1}X_{l_2}\cdots X_{l_{k-1}}$ (as above). In case (ii) we similarly derive

$$X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_1^2 \vdash X_{l_1}X_{l_2}\cdots X_{l_{k-1}}X_1 \vdash X_{l_1}X_{l_2}\cdots X_{l_{k-1}}.$$

Observe that after each step of the described derivation the *order* of the current monomial does not increase in all cases. More precisely, at a regular step the order remains the same, while on a singular step the order decreases by either 1 or 2. Thereby, the number of singular steps does not exceed $O(2^n)$. Moreover, if as a result of a singular step we derive a monomial $X_{l_1}\cdots X_{l_k}$, then after that we perform $l_k - 1$ regular steps, followed by a singular one. Since $\mathrm{ord}(X_{l_1}\cdots X_{l_k}) = 2^{l_1-1} + \cdots + 2^{l_k-1}$, the total number of regular steps in the derivation can be bounded from above by $O(2^n)$. Because each (regular or singular) step involves at most $O(n)$ derivation rules of PC, we obtain the bound $O(n2^n)$ on the length of the derivation.

While the order of a current monomial is positive we are able to perform either a regular or a singular step. The derivation terminates when we attain the order 0, thereby the monomial 1. Obviously, the degree of any intermediately derived polynomial does not exceed $O(n)$.$\square$

2 Positivstellensatz proofs

Similar to NS and PC which rely on Hilbert's Nullstellensatz, we introduce here Positivstellensatz proof system and Positivstellensatz calculus respectively. Both are based on the Positivstellensatz [BCR 87], [S 74].

Definition 1 *The* cone $c(h_1,\dots,h_m)$ *generated by polynomials* $h_1,\dots,h_m \in \mathbf{R}[X_1,\dots,X_n]$ *is the smallest family of polynomials containing* $h_1,\dots,h_m$ *and satisfying the following rules:*

(a) $e^2 \in c(h_1,\dots,h_m)$ *for any* $e \in \mathbf{R}[X_1,\dots,X_n]$*;*

if $a, b \in c(h_1,\dots,h_m)$*, then*
(b) $a+b \in c(h_1,\dots,h_m)$*;*
(c) $ab \in c(h_1,\dots,h_m)$*.*

Remark 2 *The minimal cone* $c(\emptyset)$ *consists of all sums of squares of polynomials.*

Remark 3 *Any element of* $c(h_1,\dots,h_m)$ *can be represented in a form*

$$\sum_{I\subset\{1,\dots,m\}} \left(\prod_{i\in I} h_i\right)\left(\sum_j e_{I,j}^2\right)$$

for some polynomials $e_{I,j} \in \mathbf{R}[X_1,\dots,X_n]$*.*

Positivstellensatz *A system of equations* $f_1 = \cdots = f_k = 0$ *and inequalities* $h_1 \geq 0,\dots,h_m \geq 0$*, where* $f_1,\dots,f_k,h_1,\dots,h_m \in \mathbf{R}[X_1,\dots,X_n]$ *has no common solutions in* $\mathbf{R}^n$ *if and only if for a suitable polynomial* $f \in \mathbf{R}[X_1,\dots,X_n]$ *from the ideal* $(f_1,\dots,f_k)$ *and a polynomial* $h \in c(h_1,\dots,h_m)$ *we have:* $f+h=-1$*.*

Consider a system of equations and inequalities

$$f_1 = \cdots f_k = 0,\ h_1 \geq 0,\dots,h_m \geq 0. \tag{3}$$

The following proof system is stronger than NS refutations and could be viewed as its Positivstellensatz analogue.

Definition 2 *A pair of polynomials*

$$(f,h) = \left(\sum_{1\leq s\leq k} f_s g_s,\quad \sum_{I\subset\{1,\dots,m\}} \left(\prod_{i\in I} h_i\right)\left(\sum_j e_{I,j}^2\right)\right)$$

with $f+h=-1$ *where* $g_i, e_{I,j} \in \mathbf{R}[X_1,\dots,X_n]$ *we call a* Positivstellensatz refutation *for (3). The* degree *of the refutation is*

$$\max_{s,I,j}\{\deg(f_s g_s),\ \deg(e_{I,j}^2 \prod_{i\in I} h_i)\}.$$

The following proof system is stronger than PC and could be viewed as its Positivstellensatz analogue.

Definition 3 *Let a polynomial $f \in (f_1, \dots, f_k)$ be derived in PC from the axioms $f_1, \dots, f_k$, and a polynomial $h \in c(h_1, \dots, h_m)$ be derived, applying the rules (a), (b), (c) (see Definition 1), from the axioms $h_1, \dots, h_m$. Suppose that $f + h = -1$. This pair of derivations we call a* Positivstellensatz calculus refutation *for (3). By its* degree *we mean the maximum of the degrees of intermediate polynomials from both derivations. The* length *of the refutation we define as the total number of steps in both derivations.*

Note that the system of Lovász-Schrijver [LS 91] could be viewed as the degree 2 fragment of the Positivstellensatz calculus.

3 Lower bounds for Positivstellensatz proof systems

Generalizing the lower bound on NS degree for the telescopic system (1), we prove the following theorem.

Theorem 2 *The degree of any Positivstellensatz refutation of the system (1) is greater or equal to 2^{n-1}.*

Proof. Let

$$h + 1 = -f, \tag{4}$$

where

$$h = \sum_j e_j^2$$

and

$$-f = (1 - YX_1)g_1 + (X_1^2 - X_2)g_2 + \cdots + (X_{n-1}^2 - X_n)g_n + X_n g.$$

Make the substitution (2) into (4) and assume contrary to the statement of the theorem (in fact we need only to assume that $\deg_Y(g) < 2^{n-1}$). We conclude that the order of the right-hand side of (4),

$$\operatorname{ord}(-f) = \operatorname{ord}(g) + 2^{n-1} \geq 1.$$

On the other hand consider the result of the substitution (2) into the left-hand side of (4). Let q be the minimal power of X_1 occurring in all $\overline{e}_j$. If $q \leq 0$, then X_1^{2q} will occur in the left-hand side $\overline{h}+1$ with a positive coefficient, which contradicts to the proved above inequality $\mathrm{ord}(-f) \geq 1$. Otherwise, if $q > 0$, then $\mathrm{ord}(h+1) = 0$, and again we get a contradiction.$\Box$

The next theorem provides a lower bound for the system (1) for Positivstellensatz calculus, which together with Theorem 1 (Section 1) shows that the obtained complexity bounds are sharp for the Positivstellensatz calculus (and *a fortiori*, they are sharp for PC).

Theorem 3 *Any Positivstellensatz calculus refutation for (1) has a degree greater than $n - \lceil \log_2 n \rceil - 2$, and the length at least 2^{n-1}.*

Proof. Consider (4). As it was shown in the proof of Theorem 2, $\mathrm{ord}(-f) \leq 0$. Let us consider a PC derivation from (1) of f provided by Positivstellensatz calculus refutation (see Definition 3).

At the beginning of the derivation the only non-vanishing (after the substitution (2)) axiom is $\overline{X}_n = X_1^{2^{n-1}}$. The minimum of the orders of the polynomials, intermediate in the course of the derivation, could decrease on one step by at most 1 (only due to the multiplication by Y). Thus, this minimum attains successively the values $2^{n-1}, 2^{n-1}-1, \dots, 1, 0$ because $\mathrm{ord}(-f) \leq 0$. It follows that the length of the derivation is at least 2^{n-1}. In particular, there is a polynomial p, intermediate in the course of the derivation, with the order $\mathrm{ord}(p) = 2^{n-1} - 2^{\lceil \log_2 n \rceil + 1}$. Take a monomial $X_{l_1}^{i_1} \cdots X_{l_k}^{i_k} Y^s$ with $l_1 > \cdots > l_k \geq 1$ occurring in p with the order exactly

$$\mathrm{ord}(X_{l_1}^{i_1} \cdots X_{l_k}^{i_k} Y^s) = \mathrm{ord}(p).$$

Since $s \leq n$ (otherwise the theorem is proved),

$$2^{n-1} - 2^{\lceil \log_2 n \rceil + 1} \leq \mathrm{ord}(X_{l_1}^{i_1} \cdots X_{l_k}^{i_k}) = \mathrm{ord}(p) + s \leq 2^{n-1} - 2^{\lceil \log_2 n \rceil + 1} + 2^{\lceil \log_2 n \rceil}.$$

Thus, the binary representation of the integer $\mathrm{ord}(X_{l_1}^{i_1} \cdots X_{l_k}^{i_k})$ contains at least $n - \lceil \log_2 n \rceil - 2$ digits 1 at the positions from $\lceil \log_2 n \rceil + 2$ to $n-1$.

We prove that $i_1 + \cdots + i_k \geq n - \lceil \log_2 n \rceil - 2$. Indeed, replacing in the monomial $X_{l_1}^{i_1} \cdots X_{l_k}^{i_k}$ any $X_{l_j}^{i_j}$ with $i_j \geq 2$ by $X_{l_j+1} X_{l_j}^{i_j-2}$, and keeping doing this as long as possible, we do not change the order whereas decreasing the degree of the monomial. At the end we arrive to a monomial of the form

$$X_{n-1} X_{n-2} \cdots X_{\lceil \log_2 n \rceil + 2} X_{t_1} \cdots X_{t_r},$$

$$\lceil \log_2 n \rceil + 2 > t_1 > \cdots > t_r \geq 1$$

with the degree at least $n - \lceil \log_2 n \rceil - 2$.$\Box$

4 Further research

One could further extend the Positivstellensatz calculus by admitting a following derivation of two polynomials f, H. The polynomial $f \in (f_1, \ldots, f_k)$ is derived by means of PC from the axioms $f_1, \ldots, f_k$. The polynomial $H \in (f_1, \ldots, f_k) + c(h_1, \ldots, h_m)$ is derived from the axioms $f_1, \ldots, f_k, h_1, \ldots h_m$ using the rules (a), (b), (c), and allowing to take f as a, b (see Definition 1). Thus, we derive separately the elements from $(f_1, \ldots, f_k)$ and from $(f_1, \ldots, f_k) + c(h_1, \ldots, h_1)$. The aim is to derive $H = -1$, i.e., a refutation for the initial system of equations and inequalities. It would be interesting to obtain lower bounds for this extended calculus.

Another challenging problem is to obtain lower bounds for the Boolean problems, e.g., for the ones studied in [BuIK 96], with respect to the Positivstellensatz calculi introduced in this paper.

It would be also interesting to construct examples of a system of equations for which Positivstellensatz derivations have the degrees or lengths less than their Nullstellensatz analogues.

Observe that the lower bounds in Theorems 2 and 3 remain true for the sum of squares of polynomials from (1):

$$f_0 = (1 - YX_1)^2 + (X_1^2 - X_2)^2 + \cdots + (X_{n-1}^2 - X_n)^2 + X_n^2.$$

The proofs for f_0 go through almost literally. What are the upper bounds for the complexities of Positivstellensatz refutation and calculus for the equation $f_0 = 0$?

Acknowledgement

The first author would like to acknowledge the support of EPSRC visiting fellowship grant GR/M12308 and the Department of Mathematical Sciences of the University of Bath which he had been visiting during the work on the paper.

We are grateful to Paul Beame, Jan Krajíček, Toni Pitassi, Pavel Pudlák and Marie-Françoise Roy for useful discussions.

References

[BIK 96] P. Beame, R. Impagliazzo, J. Krajíček, T. Pitassi, and P. Pudlák *"Lower bounds on Hilbert's Nullstellensatz and propositional proofs,"* Proc. London Math. Soc., V. 73, 1996, 1–26.

[BCR 87] J. Bochnak, M. Coste, and M.-F.Roy *"Géometrie Algébrique Réelle,"* Springer-Verlag, Berlin, 1987.

[Br 87] D. Brownawell *"Bounds for the degrees in the Nullstellensatz,"* Ann. Math., 1987, V. 126, 577–591.

[BuGI 99] S. Buss, D. Grigoriev, R. Impagliazzo, and T. Pitassi *"Linear gaps between degrees for polynomial calculus modulo distinct primes,"* to appear in: Proc. 31st Ann. ACM Symp. on Theory of Computing, 1999.

[BuIK 96] S. R. Buss, R. Impagliazzo, J. Krajíček, and P. Pudlák, A. Razborov, and J. Sgall *"Proof complexity in algebraic systems and bounded depth Frege systems with modular counting,"* Computational Complexity, 6, 1996/1997, 256–298.

[CEI 96] M. Clegg, J. Edmonds, and R. Impagliazzo *"Using the Groebner basis algorithm to find proofs of unsatisfiability,"* Proc. 28th Ann. ACM Symp. on Theory of Computing, 1996, 174–183.

[CLR 99] M. Coste, H. Lombardi, and M.-F. Roy *Dynamical method in algebra: effective Nullstellensätze,"* Preprint, 1999.

[G 98] D. Grigoriev *"Nullstellensatz lower bounds for Tseitin tautologies,"* Proc. 39th Ann. IEEE Symp. on Foundations of Computer Science, 1998, 648–652.

[IPS 97] R. Impagliazzo, P. Pudlák, and J. Sgall *"Lower bounds for polynomial calculus and the Groebner basis algorithm,"* 1997, to appear in Computational Complexity.

[LMR 96] H. Lombardi, N. Mnev, and M.-F. Roy *"The Positivstellensatz and small deduction rules for systems of inequalities,"* Math. Nachr., V. 181, 1996, 245–259.

[L 94] L. Lovász *"Stable sets and polynomials,"* Discrete Mathematics, V. 124, 1994, 137–153.

[LS 91] L. Lovász and A. Schrijver *"Cones of matrices and set-functions and 0–1 optimization,"* SIAM J. Optimization, V. 1, 1991, 166–190.

[P 98] P. Pudlák *"On the complexity of the propositional calculus,"* Preprint, 1998.

[R 96] A. Razborov *"Lower bounds for the polynomial calculus,"* 1996, to appear in Computational Complexity.

[S 74] G. Stengle *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry,"* Math. Ann., V. 207, 1974, 87–97.

[ST 98] T. Stephen and L. Tunçel *"On representation of the matching polytope via semidefinite liftings,"* Preprint, 1998.