Technical Report

Bounds for Absolute Positiveness of Multivariate Polynomials

Hoon Hong

Research Institute for Symbolic Computation

Johannes Kepler University

A-4040 Linz, Austria

e-mail: hhong@risc.uni-linz.ac.at

www: http://www.risc.uni-linz/people/hhong

March 11, 1997

A multivariate polynomial $P(x_1, \ldots, x_n)$ with real coefficients is said to be *absolutely positive* from a real number B iff it and all of its non-zero partial derivatives of every order are positive for $x_1, \ldots, x_n \geq B$. We call such B a *bound* for the absolute positiveness of P. This paper provides a simple formula for computing such bounds. We also prove that the resulting bounds are guaranteed to be close to the optimal ones.

Key words: multivariate polynomials, absolute positiveness, roots, bounds.

1. Introduction

A multivariate polynomial $P(x_1, \ldots, x_n)$ with real coefficients is said to be *absolutely positive* from a real number B iff it and all of its non-zero partial derivatives of every order are positive for $x_1, \ldots, x_n \geq B$. We also call such B a *bound* for the absolute positiveness of P. The main goal of this paper is to devise a "nice" formula for computing a bound of a given polynomial.

The initial motivation arose while studying several *partial* methods for testing *positiveness* of multivariate polynomials (Ben-Cherifa and Lescanne, 1987; Dershowitz, 1987; Steinbach, 1992; Steinbach, 1994; Giesl, 1995). We found that these partial methods are in fact *complete* methods for testing *absolute* positiveness (Hong and Jakus, 1996). Since then, we also realized that most previously known formulas for univariate root bounds (Cauchy, 1829; Carmichael and Mason, 1914; Birkhoff, 1914; Fujiwara, 1915; Kelleher, 1916; Kuniyeda, 1916; Cohn, 1922; Berwald, 1934; Montel, 1932; Tôya, 1933; Marden, 1949; Johnson, 1991) are in fact bounds for absolute positiveness (thus, a bound not only for the polynomial, but also for all its non-zero derivatives). Indeed, from Lucas' theorem (Lucas, 1874) one can conclude, in the univariate case, that any complex root bound, when used as a bound for real roots, is also a bound for absolute positiveness (See Section 5). Thus, we believe that the notion of absoluteness positiveness deserves to be investigated.

Not all multivariate polynomials have bounds for absolute positiveness. Thus, first we need to have a method for checking the existence of bounds. An efficient method is given in (Hong and Jakus, 1996), and we use it in this paper.

The main contributions of this paper are: (1) We give a simple formula for computing a bound for a given multivariate polynomial, when exists. (2) We prove that the resulting bound is always "good", in that it is guaranteed to be close the optimal bound, *unlike* previously known bounds.

The structure of the paper is as follows. In Section 2, we give precise statements of the main results of this paper (a bound and its quality). In sections 3 and 4, we prove those main results. Finally in section 5, we compare the bound with known bounds in the univariate case.

For modern treatment of related topics, see the recent books (Milanovic *et al.*, 1994; Borwein and Erdelyi, 1995).

2. Main Results

In this section, we give precise statements of the main results of this paper. The proofs will be given in the later sections (Sections 3 and 4). We begin by fixing some notations/conventions that will be used throughout the paper.

NOTATIONS 2.1.

$$
\begin{array}{lll}
\mu & := & (\mu_1, \ldots, \mu_n) \in \mathbf{N}^n \\
|\mu| & := & \mu_1 + \cdots + \mu_n \\
\mu! & := & \mu_1! \cdots \mu_n! \\
\nu - \mu & := & (\nu_1 - \mu_1, \ldots, \nu_n - \mu_n) \\
\nu \geq \mu & := & \nu_1 \geq \mu_1 \wedge \cdots \wedge \nu_n \geq \mu_n \\
\nu > \mu & := & \nu \geq \mu \ \ \wedge \ \ \nu \neq \mu \\
x & := & (x_1, \ldots, x_n) \\
x^\mu & := & x_1^{\mu_1} \cdots x_n^{\mu_n} \\
P^{(\mu)} & := & \frac{\partial^{|\mu|} P}{\partial x_1^{\mu_1} \cdots \partial x_n^{\mu_n}} \\
\forall x \geq B & := & \forall x_1 \geq B \cdots \forall x_n \geq B
\end{array}
$$

DEFINITION 2.1. (ABSOLUTE POSITIVENESS) *Let $P \in \mathbf{R}[x]$ and let $B \in \mathbf{R}$. We say that P is* absolutely positive from *B iff the following two conditions hold:*

(a) $\forall x \geq B \;\; P(x) > 0$

(b) $\forall x \geq B \;\; P^{(\lambda)}(x) > 0$, *for every non-zero partial derivative $P^{(\lambda)}$ of P.*[†]

We will also say that B is a bound *for the absolute positiveness of P.* □

[†] In our previous paper (Hong and Jakus, 1996), we used a slightly different (weaker) condition: $\forall x \geq B \;\; P^{(\lambda)}(x) \geq 0$ for every partial derivate $P^{(\lambda)}$ of P. The new definition will be essential for proving certain theorem in this paper (Theorem 2.3).

Example 2.1. The polynomial $x^2 + y^2 - 1$ is absolutely positive from 1. But the polynomial $(x-y)^2 + 1$ is not absolutely positive from any bound, because the derivative $\frac{\partial^2 P}{\partial x \partial y} = -2$ is always negative. □

A question arises immediately: *For which polynomial does there exist a bound for the absolute positiveness?* We have given a complete answer to this question in our previous paper (Hong and Jakus, 1996).† We recall this result because we will need it while stating and proving the main results of this present paper. First we need one more notion: *dominating monomial*, which is a generalization of the notion of *leading* monomial to the multivariate case.

Definition 2.2. (Dominating Monomial) *We say that a monomial $a_\nu x^\nu$* dominates *a monomial $a_\mu x^\mu$ iff $\nu > \mu$. We say that p is a* dominating monomial *of P iff no monomial in P dominates p.* □

Example 2.2. Let us consider polynomials $P = x^2 - 2xy + y^2 + 1$ and $Q = x^2y - xy + y^2$. There are three dominating monomials in P, namely $x^2, -2xy$ and y^2. There are two dominating monomials in Q, namely x^2y and y^2. For univariate polynomials there is only one dominating monomial — the leading monomial.□

Theorem 2.1. (Existence (Hong and Jakus, 1996)) *Let $P \in \mathbf{R}[x]$ be a non-zero polynomial*‡*. Then the following two properties are equivalent.*

(A) *There exists a bound for the absolute positiveness of P.*

(B) *Every dominating monomial of P has positive coefficient.*

The above theorem only tells about the existence, and we naturally would like to find a "witness" when there exists a bound. The next theorem (Theorem 2.2) provides a formula for finding a witness. In order to simplify the presentation of this and the subsequent theorems/proofs, we will make the following global assumption on the polynomial P.

Assumption 2.1. *We assume, from here to the end of this paper, that*

(a) *every dominating monomial of P has positive coefficient and*

(b) *at least one monomial of P has negative coefficient.* □

The assumption (a) ensures that there exists a bound for the absolute positiveness (Theorem 2.1). The assumption (b) filters out a trivial degenerate case. If all the monomials of P have positive coefficients, we see immediately that P is absolutely positive from any $B > 0$.

† The question can be easily formulated as a sentence in the first order theory of real closed fields. Thus, in principle, we can use any decision procedure for the theory (Tarski, 1951; Collins, 1975; Arnon, 1981; McCallum, 1984; Grigor'ev, 1988; Canny, 1988; Weispfenning, 1988; Heintz *et al.*, 1989; Hong, 1990; Collins and Hong, 1991; Renegar, 1992) to check the existence of bounds. However, since this is a very structured and special question, one can naturally find a special method which is more efficient than the general methods. In (Lankford, 1976), a special method is given (using partial differentiation and evaluation). But we will use the method in (Hong and Jakus, 1996), because it is simpler and more efficient.

‡ When $P = 0$, we trivially see that P is not absolute positive from any bound.

Further, the following expression will appear frequently throughout the paper, thus, we will introduce a short-hand for it.

NOTATIONS 2.2. $\Omega_n = 1 - \sqrt[n]{\frac{1}{2}}$.

THEOREM 2.2. (BOUND) *Let* $P = \sum_{\mu \in I} a_\mu x^\mu \in \mathbf{R}[x]$ *and let*

$$B_P = \frac{1}{\Omega_n} \max_{a_\mu < 0} \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu - \mu|}}.$$

Then P *is absolutely positive from* B_P*.*

PROOF. Given in Section 3. □

The above expression for B_P is well-defined due to Assumption 2.1, that is, the index sets of max and min are non-empty.

Another question arises: *How good (tight) is the bound given above?* To answer this, one needs a notion such as "optimal bound". However, in general, the set of all bounds for the absolute positiveness of a given polynomial is an *open* set, without a minimum. Thus, we introduce instead another similar notion: *threshold.*

DEFINITION 2.3. (THRESHOLD) *The* threshold *of absolute positiveness of a polynomial* P*, written as* A_P*, is the infimum of all the bounds for the absolute positiveness of* P*.* □

Due to Assumption 2.1, we have that $A_P > 0$. Obviously, we also have $\frac{B_P}{A_P} > 1$. Naturally, we want that this ratio cannot be arbitrarily large. The following theorem tells us that the ratio is indeed bounded from above (when the degree and the number of variables are fixed).

THEOREM 2.3. (QUALITY) *We have*

$$\frac{B_P}{A_P} \leq \frac{1}{\Omega_n} \frac{d_1 + \cdots + d_n}{\ln(2)}$$

where $d_i = \deg_{x_i} P$*.*

PROOF. Given in Section 4. □

Thus, the ratio is bounded by an expression which is linear in the sum of the degrees. How does it depend on the number of variables? For this, we need to understand the behavior of the factor $\frac{1}{\Omega_n}$. The following proposition tells us that it is almost linear in n.

PROPOSITION 2.1. (ALMOST LINEAR BEHAVIOR)

$$\frac{n}{\ln(2)} + \frac{1}{2} \leq \frac{1}{\Omega_n} \leq \frac{n}{\ln(2)} + \frac{1}{2} + \frac{\ln(2)}{12\, n}.$$

PROOF. Immediate from Laurent expansion of $\frac{1}{\Omega_n}$ around $\frac{1}{n} = 0$. □

3. Proof of Bound Theorem

In this section, we will prove Theorem 2.2. The proof is divided into a few lemmas for easier reading and also for separating out main insights. We begin by finding a bound for the positiveness of polynomials of a certain nice type:

Lemma 3.1. *Let P be of the type:*

$$P = a_\nu x^\nu + \sum_{\mu \in I} a_\mu x^\mu$$

where $a_\nu > 0$, $a_\mu < 0$, and $\nu > \mu$ for $\mu \in I$. Let

$$B_P^* = \frac{1}{\Omega_n} \max_{\mu \in I} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu - \mu|}}.$$

Then P is positive from B_P^*, that is,

$$\forall x \geq B_P^* \;\; P(x) > 0.$$

Proof. Let $x \geq B_P^*$ be arbitrary but fixed. We need to show that $P(x) > 0$. We show it by the following repeated rewriting, which in fact also shows how the formula for B_P^* was originally discovered.

$$\begin{aligned}
P(x) &= a_\nu x^\nu + \sum_{\mu \in I} a_\mu x^\mu \\
&= a_\nu x^\nu \left[1 + \sum_{\mu \in I} \frac{a_\mu}{a_\nu} \frac{1}{x^{\nu-\mu}} \right] \\
&= a_\nu x^\nu \left[1 - \sum_{\mu \in I} \left| \frac{a_\mu}{a_\nu} \right| \frac{1}{x^{\nu-\mu}} \right] && \text{since } a_\nu > 0 \text{ and } a_\mu < 0. \\
&\geq a_\nu x^\nu \left[1 - \sum_{\mu \in I} \left| \frac{a_\mu}{a_\nu} \right| \frac{1}{{B_P^*}^{|\nu-\mu|}} \right] && \text{since } x \geq B_P^* > 0. \\
&= a_\nu x^\nu \left[1 - \sum_{\mu \in I} \left(\left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu-\mu|}} \frac{1}{B_P^*} \right)^{|\nu-\mu|} \right] \\
&\geq a_\nu x^\nu \left[1 - \sum_{\mu \in I} \Omega_n^{|\nu-\mu|} \right] \\
&\geq a_\nu x^\nu \left[1 - \sum_{\mu < \nu} \Omega_n^{|\nu-\mu|} \right] && \text{since possibly more is subtracted.} \\
&= a_\nu x^\nu \left[2 - \sum_{\mu \leq \nu} \Omega_n^{|\nu-\mu|} \right] && \text{since the summand is 1 when } \mu = \nu.
\end{aligned}$$

$$\begin{aligned}
&= a_\nu x^\nu \left[2 - \sum_{0\le\mu_1\le\nu_1} \cdots \sum_{0\le\mu_n\le\nu_n} \Omega_n^{\nu_1-\mu_1} \cdots \Omega_n^{\nu_n-\mu_n} \right] \\
&= a_\nu x^\nu \left[2 - \sum_{0\le\mu_1\le\nu_1} \Omega_n^{\nu_1-\mu_1} \cdots \sum_{0\le\mu_n\le\nu_n} \Omega_n^{\nu_n-\mu_n} \right] \\
&= a_\nu x^\nu \left[2 - \sum_{0\le\mu_1\le\nu_1} \Omega_n^{\mu_1} \cdots \sum_{0\le\mu_n\le\nu_n} \Omega_n^{\mu_n} \right] \\
&> a_\nu x^\nu \left[2 - \sum_{0\le\mu_1} \Omega_n^{\mu_1} \cdots \sum_{0\le\mu_n} \Omega_n^{\mu_n} \right] \\
&= a_\nu x^\nu \left[2 - \frac{1}{1-\Omega_n} \cdots \frac{1}{1-\Omega_n} \right] \qquad \text{since } \Omega_n < 1. \\
&= a_\nu x^\nu \left[2 - \left(\frac{1}{1-\Omega_n} \right)^n \right] \\
&= a_\nu x^\nu \left[2 - \left(\frac{1}{1-\left(1-\sqrt[n]{\frac{1}{2}}\right)} \right)^n \right] \\
&= 0.
\end{aligned}$$

Thus, we have shown that $P(x) > 0$. □

REMARK 3.1. Note that only at the very last step of the rewriting (the second line from the bottom) we have used the definition of Ω_n. In fact, we originally discovered the definition of Ω_n by examining the expression on the third line from the bottom. We simply looked for the value of Ω_n that will make the expression to become 0, which is obviously $1 - \sqrt[n]{1/2}$. □

REMARK 3.2. A referee suggested that a slight improvement could be obtained by replacing the 5-th line from the bottom with the following:

$$a_\nu x^\nu \left[2 - \sum_{0\le\mu_1\le d} \Omega_n^{\mu_1} \cdots \sum_{0\le\mu_n\le d} \Omega_n^{\mu_n} \right]$$

where $d = \max_i \nu_i$. Though this observation is correct in itself, the approach eventually requires solving the polynomial equation:

$$\Omega_n^d + \Omega_n^{d-1} + \cdots + \Omega_n + 1 - \sqrt[n]{2} = 0$$

which in general does not have a "closed" form solution. □

Next we generalize this result to find a bound for the positiveness of arbitrary polynomials.

LEMMA 3.2. *Let* $P = \sum_{\mu \in I} a_\mu x^\mu \in \mathbf{R}[x]$ *and let*

$$B_P = \frac{1}{\Omega_n} \max_{a_\mu < 0} \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu - \mu|}}.$$

Then P *is positive from* B_P, *that is,*

$$\forall x \geq B_P \;\; P(x) > 0.$$

PROOF. Consider a partition of the monomials of P

$$P = P_1 + \cdots + P_\ell + R$$

such that

(a) Each P_k is of the type studied in the previous lemma, that is,

$$P_k = a_{\nu^{(k)}} x^{\nu^{(k)}} + \sum_{\mu \in I^{(k)}} a_\mu x^\mu$$

where $a_{\nu^{(k)}} > 0$, $a_\mu < 0$, and $\nu^{(k)} > \mu$ for all $\mu \in I^{(k)}$.

(b) R is either 0 or a polynomial consisting of only positive monomials.

Such a partition exists due to Assumption 2.1. Let $B^*_{P_k}$ be the bound for the positiveness of P_k given in the previous lemma. Then obviously P is positive from $\max_k B^*_{P_k}$, which is

$$\begin{aligned}
& \max_k \left[\frac{1}{\Omega_n} \max_{\mu \in I^{(k)}} \left| \frac{a_\mu}{a_{\nu^{(k)}}} \right|^{\frac{1}{|\nu^{(k)} - \mu|}} \right] \\
= & \frac{1}{\Omega_n} \max_k \max_{\mu \in I^{(k)}} \left| \frac{a_\mu}{a_{\nu^{(k)}}} \right|^{\frac{1}{|\nu^{(k)} - \mu|}} \\
= & \frac{1}{\Omega_n} \max_{a_\mu < 0} \left| \frac{a_\mu}{a_{\nu^{(\mu)}}} \right|^{\frac{1}{|\nu^{(\mu)} - \mu|}}
\end{aligned}$$

where $\nu^{(\mu)}$ stands for the exponent vector of the positive monomial which belongs to the same partition as the monomial $a_\mu x^\mu$.

Now we only have to fix a partition. Which partition shall we choose? Equivalently put, for each negative monomial $a_\mu x^\mu$, which positive monomial $a_\nu x^\nu$ shall we choose to put in the same partition? The best one is naturally the one that minimizes $\max_k B^*_{P_k}$. One sees immediately that it means to choose ν that minimizes $\left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu - \mu|}}$ under the condition $a_\nu > 0$ and $\nu > \mu$. Thus, we obtain the following bound:

$$B_P = \frac{1}{\Omega_n} \max_{a_\mu < 0} \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu - \mu|}}.$$

□

PROOF OF BOUND THEOREM. In Lemma 3.2, we have already shown that

$$\forall x \geq B_P \;\; P(x) > 0.$$

Thus it only remains to show that

$$\forall x \geq B_P \;\; P^{(\lambda)}(x) > 0$$

for every non-zero partial derivative $P^{(\lambda)}$ of P. If $P^{(\lambda)}$ consists of only positive monomials, then it is obviously true. Thus from now on assume that $P^{(\lambda)}$ has at least one negative monomial.

The idea for the proof is to apply Lemma 3.2 to $P^{(\lambda)}$, obtaining a bound $B_{P^{(\lambda)}}$, and to show that $B_{P^{(\lambda)}} \leq B_P$. But before doing so, we need to ensure that $P^{(\lambda)}$ satisfies the conditions in Assumption 2.1.

Note that the condition (b) is already satisfied since we assumed it in the above. In order to see whether the condition (a) is also satisfied, we first recall that all the dominating monomial of P have positive coefficient (from Assumption 2.1). During differentiation, a dominating monomial of P either disappears or stays as a dominating monomial (multiplied with some positive integer). Further every dominating monomial of $P^{(\lambda)}$ originates from a dominating monomial of P. So all the dominating monomials of $P^{(\lambda)}$ have positive coefficients. Thus, $P^{(\lambda)}$ satisfies the condition (a) also. Hence, we can safely apply Lemma 3.2.

Now, by Lemma 3.2, we know that

$$\forall x \geq B_{P^{(\lambda)}} \quad P^{(\lambda)}(x) > 0.$$

Next, we will show that $B_{P^{(\lambda)}} \leq B_P$. For this, let us recall the elementary fact from calculus:

$$P^{(\lambda)}(x) = \sum_{\mu \in I, \mu \geq \lambda} \frac{\mu!}{(\mu-\lambda)!} a_\mu x^{\mu-\lambda}.$$

Thus, we have

$$\begin{aligned}
B_{P^{(\lambda)}} &= \frac{1}{\Omega_n} \max_{\frac{\mu!}{(\mu-\lambda)!} a_\mu < 0} \quad \min_{\frac{\nu!}{(\nu-\lambda)!} a_\nu > 0, \nu-\lambda > \mu-\lambda} \left| \frac{\frac{\mu!}{(\mu-\lambda)!} a_\mu}{\frac{\nu!}{(\nu-\lambda)!} a_\nu} \right|^{\frac{1}{|(\nu-\lambda)-(\mu-\lambda)|}} \\
&= \frac{1}{\Omega_n} \max_{a_\mu < 0, \mu \geq \lambda} \quad \min_{a_\nu > 0, \nu \geq \lambda, \nu > \mu} \left| \frac{\frac{\mu!}{(\mu-\lambda)!} a_\mu}{\frac{\nu!}{(\nu-\lambda)!} a_\nu} \right|^{\frac{1}{|\nu-\mu|}} \\
&= \frac{1}{\Omega_n} \max_{a_\mu < 0, \mu \geq \lambda} \quad \min_{a_\nu > 0, \nu > \mu} \left| \frac{\frac{\mu!}{(\mu-\lambda)!} a_\mu}{\frac{\nu!}{(\nu-\lambda)!} a_\nu} \right|^{\frac{1}{|\nu-\mu|}} \\
&\qquad \text{since } \mu \geq \lambda \text{ and } \nu > \mu \text{ implies } \nu \geq \lambda. \\
&\leq \frac{1}{\Omega_n} \max_{a_\mu < 0, \mu \geq \lambda} \quad \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu-\mu|}} \\
&\qquad \text{since } \frac{\frac{\mu!}{(\mu-\lambda)!}}{\frac{\nu!}{(\nu-\lambda)!}} < 1. \\
&\leq \frac{1}{\Omega_n} \max_{a_\mu < 0} \quad \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{|\nu-\mu|}} \\
&= B_P.
\end{aligned}$$

Thus we have shown that $B_{P^{(\lambda)}} \leq B_P$. Hence obviously we have

$$\forall x \geq B_P \quad P^{(\lambda)}(x) > 0.$$

$\square$

4. Proof of Quality Theorem

In this section, we will prove Theorem 2.3.

PROOF OF QUALITY THEOREM. Let P be an arbitrary but fixed polynomial such that $\deg_{x_i} P = d_i$. We need to show that

$$\frac{B_P}{A_P} \leq \frac{1}{\Omega_n}\frac{d_1+\cdots+d_n}{\ln(2)}.$$

Recall the definition of B_P:

$$B_P = \frac{1}{\Omega_n} \max_{a_\mu<0} \min_{a_\nu>0,\nu>\mu} \left|\frac{a_\mu}{a_\nu}\right|^{\frac{1}{|\nu-\mu|}}.$$

Suppose that $\max_{a_\mu<0}$ is achieved at μ^*, so that

$$B_P = \frac{1}{\Omega_n} \min_{a_\nu>0,\nu>\mu^*} \left|\frac{a_{\mu^*}}{a_\nu}\right|^{\frac{1}{|\nu-\mu^*|}}.$$

In order to simplify the notation, we will write μ instead of μ^* from now on. For every $\nu>\mu$ such that $a_\nu>0$, we have

$$\begin{aligned}
\left|\frac{a_\mu}{a_\nu}\right|^{\frac{1}{|\nu-\mu|}} &\geq B_P\Omega_n \\
\left|\frac{a_\mu}{a_\nu}\right| &\geq (B_P\Omega_n)^{|\nu-\mu|} \\
\frac{|a_\mu|}{a_\nu} &\geq (B_P\Omega_n)^{|\nu-\mu|} \\
a_\nu &\leq |a_\mu|\frac{1}{(B_P\Omega_n)^{|\nu-\mu|}}.
\end{aligned}$$

When $a_\nu<0$, the above inequality trivially holds, thus for every $\nu\in I$ such that $\nu>\mu$, we have

$$a_\nu \leq |a_\mu|\frac{1}{(B_P\Omega_n)^{|\nu-\mu|}}. \tag{4.1}$$

From the elementary calculus, we have

$$P^{(\mu)}(x) = \sum_{\nu\in I,\nu\geq\mu} \frac{\nu!}{(\nu-\mu)!}a_\nu x^{\nu-\mu}.$$

This is a non-constant polynomial since there exists $\nu\in I$ such that $\nu>\mu$, due to Assumption 2.1. Thus we can rewrite this as

$$P^{(\mu)}(x) = -\mu!\,|a_\mu| \;+\; \sum_{\nu\in I,\nu>\mu} \frac{\nu!}{(\nu-\mu)!}a_\nu x^{\nu-\mu}.$$

Using the inequality (4.1), we see that, for every $x>0$,

$$P^{(\mu)}(x) \leq -\mu!\,|a_\mu| \;+\; \sum_{\nu\in I,\nu>\mu} \frac{\nu!}{(\nu-\mu)!}|a_\mu|\frac{x^{\nu-\mu}}{(B_P\Omega_n)^{|\nu-\mu|}}.$$

Let $Q(x)$ denote the polynomial at the right hand side of the above inequality. Then we have just shown that

$$\forall x > 0 \;\; P^{(\mu)}(x) \le Q(x).$$

Let $\hat{P}^{(\mu)}(t) = P^{(\mu)}(t, \ldots, t)$ and let $\hat{Q}(t) = Q(t, \ldots, t)$. Then we immediately have

$$\forall t > 0 \;\; \hat{P}^{(\mu)}(t) \le \hat{Q}(t).$$

Note that the signs of the coefficients of $\hat{Q}$ alternate only once. Thus, by Descartes' rule of sign, there exists a unique positive root of $\hat{Q}$. Let's call the positive root α. Then we have $\hat{P}^{(\mu)}(\alpha) \le \hat{Q}(\alpha) = 0$. Hence, we have

$$\hat{P}^{(\mu)}(\alpha) \le 0. \tag{4.2}$$

We see that

$$\alpha \le A_P \tag{4.3}$$

because if $\alpha > A_P$, then P would be absolutely positive from α, contradicting the inequality (4.2).

Now it remains to estimate α. Note

$$\begin{aligned}
\hat{Q}(t) &= -\mu!\,|a_\mu| + \sum_{\nu \in I, \nu > \mu} \frac{\nu!}{(\nu-\mu)!} |a_\mu| \frac{t^{|\nu-\mu|}}{(B_P \Omega_n)^{|\nu-\mu|}} \\
&= \mu!\,|a_\mu| \left[-1 + \sum_{\nu \in I, \nu > \mu} \binom{\nu}{\mu} \left(\frac{t}{B_P \Omega_n} \right)^{|\nu-\mu|} \right].
\end{aligned}$$

Let

$$S_{I,\mu}(t) = -1 + \sum_{\nu \in I, \nu > \mu} \binom{\nu}{\mu} t^{|\nu-\mu|}.$$

Then obviously, we have

$$\hat{Q}(t) = \mu!\,|a_\mu|\, S_{I,\mu}\left(\frac{t}{B_P \Omega_n} \right).$$

Note that the sign of the coefficients of $S_{I,\mu}$ alternate only once. Thus, by Descartes' rule of sign, there exists a unique positive root, say β, of $S_{I,\mu}$. Obviously we have

$$\beta = \frac{\alpha}{B_P \Omega_n}. \tag{4.4}$$

Let's estimate β. Follow the repeated rewriting:

$$\begin{aligned}
0 &= S_{I,\mu}(\beta) \\
&= -1 + \sum_{\nu \in I, \nu > \mu} \binom{\nu}{\mu} \beta^{|\nu-\mu|} \\
&= -2 + \sum_{\nu \in I, \nu \ge \mu} \binom{\nu}{\mu} \beta^{|\nu-\mu|} \\
&\le -2 + \sum_{d_1 \ge \nu_1 \ge \mu_1} \cdots \sum_{d_n \ge \nu_n \ge \mu_n} \binom{\nu_1}{\mu_1} \beta^{\nu_1-\mu_1} \cdots \binom{\nu_n}{\mu_n} \beta^{\nu_n-\mu_n}
\end{aligned}$$

$$\begin{aligned} &= -2+\sum_{d_1\ge\nu_1\ge\mu_1}\binom{\nu_1}{\mu_1}\beta^{\nu_1-\mu_1}\cdots\sum_{d_n\ge\nu_n\ge\mu_n}\binom{\nu_n}{\mu_n}\beta^{\nu_n-\mu_n}\\ &= -2+R_{d_1,\mu_1}\cdots R_{d_n,\mu_n}\end{aligned}$$

where

$$R_{p,q}=\sum_{p\ge i\ge q}\binom{i}{q}\beta^{i-q}.$$

Now follow the repeated rewriting again:

$$R_{p,q} \;\le\; \sum_{p\ge i\ge q}\frac{p^{i-q}}{(i-q)!}\beta^{i-q} \;=\; \sum_{p-q\ge i\ge 0}\frac{(p\beta)^i}{i!} \;\le\; \sum_{i\ge 0}\frac{(p\beta)^i}{i!} \;=\; e^{p\beta}. \tag{4.5}$$

Thus, we have

$$0 \;\le\; -2+R_{d_1,\mu_1}\cdots R_{d_n,\mu_n} \;\le\; -2+e^{d_1\beta}\cdots e^{d_n\beta}.$$

Thus,

$$2 \;\le\; e^{(d_1+\cdots+d_n)\beta}.$$

Solving for β, we get

$$\beta \;\ge\; \frac{\ln(2)}{d_1+\cdots+d_n}. \tag{4.6}$$

Now recall the (in)equalities (4.3), (4.4), and (4.6):

$$\begin{aligned}\alpha &\le A_P\\ \beta &= \frac{\alpha}{B_P\Omega_n}\\ \beta &\ge \frac{\ln(2)}{d_1+\cdots+d_n}.\end{aligned}$$

From these, we obtain immediately

$$\frac{\ln(2)}{d_1+\cdots+d_n}B_P\Omega_n \;\le\; A_P.$$

Hence, we finally have

$$\frac{B_P}{A_P} \;\le\; \frac{1}{\Omega_n}\frac{d_1+\cdots+d_n}{\ln(2)}.$$

□

Remark 4.1. A referee pointed out that the estimation of $R_{p,q}$ given in (4.5) can be sharpened as follows:

$$R_{p,q} \;=\; \sum_{j=0}^{p-q}\binom{j+q}{q}\beta^j \;=\; \sum_{j=0}^{p-q}\binom{j+q}{j}\beta^j \;\le\; \sum_{j=0}^{p-q}\binom{p}{j}\beta^j \;\le\; \sum_{j=0}^{p}\binom{p}{j}\beta^j \;\le\; (1+\beta)^p.$$

Note that

$$(1+\beta)^p \;<\; (1+\beta+\frac{\beta^2}{2!}+\cdots)^p \;=\; \left(e^\beta\right)^p \;=\; e^{p\beta}.$$

Continuing with this sharper bound for $R_{p,q}$, one can obtain

$$\frac{B_P}{A_P} \leq \frac{1}{\Omega_n} \frac{1}{2^{\frac{1}{d_1+\cdots+d_n}} - 1}.$$

How much sharper is it than the one given in Theorem 2.3? By carrying out Laurent expansion of the expression $\frac{1}{2^x-1}$, one can straightforwardly observe that

$$\frac{d_1+\cdots+d_n}{\ln(2)} - \frac{1}{2^{\frac{1}{d_1+\cdots+d_n}} - 1} < \frac{1}{2}$$

Thus, the reduction is at most $\frac{1}{2\Omega_n}$. □

5. Comparison with Known Bounds in Univariate Case

Naturally, one is interested to know how the proposed bound compares with any previously known bounds. However, we are not aware of any bounds for multivariate case. Thus, from now on, we will consider only univariate case.

Let P be a univariate polynomial (satisfying Assumption 2.1). Obviously any bound B for the absolute positiveness of P is also a bound for the real roots of P, that is, every real root of P is less than B. Now a question arises: *How does B_P given in Theorem 2.2 compare with the known bounds for the real roots of univariate polynomials?* As far as we are aware, the known bounds can be classified into two types:

(a) A bound B for the modulus of complex roots (Cauchy, 1829; Carmichael and Mason, 1914; Birkhoff, 1914; Fujiwara, 1915; Kelleher, 1916; Kuniyeda, 1916; Cohn, 1922; Berwald, 1934; Montel, 1932; Tôya, 1933; Marden, 1949), that is,

$$\forall z \in \mathbf{C} \;\; P(z) = 0 \implies |z| < B.$$

(b) A bound B for positive real roots (Johnson, 1991), that is,

$$\forall z \in \mathbf{R} \;\; P(z) = 0 \implies z < B.$$

Admittedly, the bounds of type (a) were not originally intended for real roots. However bounds used in most real roots isolation/approximation algorithms (Heindel, 1970; Collins and Loos, 1976; Akritas and Collins, 1976; Collins *et al.*, 1992; Collins and Krandick, 1992; Johnson, 1991; Johnson, 1992; Collins and Krandick, 1993) are in fact bounds for complex roots. Thus we find that the comparison is still worthwhile to do. As for previously known bounds for the positive real roots, we are aware of only the bound given by Johnson (Johnson, 1991), which is a slight modification of Knuth's bound for the modulus of complex roots.

THEOREM 5.1. (KNUTH-JOHNSON BOUND (JOHNSON, 1991)) *Let $P = \sum_{\mu=0}^{d} a_\mu x^\mu$ be a univariate polynomial and let*

$$B_P^K = 2 \max_{a_\mu < 0} \left| \frac{a_\mu}{a_d} \right|^{\frac{1}{d-\mu}}.$$

Then every real root of P is smaller than B_P^K. □

First, we would like to know whether the previously known bounds are also bounds for absolute positiveness. The following theorem answers affirmatively.

THEOREM 5.2. (THEY ARE BOUNDS FOR ABSOLUTE POSITIVENESS)

(a) *Every bound for the modulus of the complex roots of a polynomial P is also a bound for the absolute positiveness of P.*

(b) *The Knuth-Johnson bound B_P^K is also a bound for the absolute positiveness of P.*

PROOF. To show (a), we recall Lucas' theorem (Lucas, 1874) which states that all the complex roots of the derivative of a non-constant polynomial P lie in the convex hull of the set of the complex zeros of the polynomial P. From this, we see immediately that the real roots of the derivative of P are smaller than B. We can apply the same reasoning repeatedly on the derivatives to see that the real roots of all the (non-zero) derivative of P are smaller than B. Hence B is a bound for the absolute positiveness of P.

To show (b), let us recall Theorem 2.2. After putting $n = 1$, we see that B_P, given by

$$B_P = 2 \max_{a_\mu < 0} \min_{a_\nu > 0, \nu > \mu} \left| \frac{a_\mu}{a_\nu} \right|^{\frac{1}{\nu - \mu}}$$

is a bound for the absolute positiveness of P. By comparing the expressions for B_P^K and B_P, one immediately observes that

$$B_P^K \geq B_P.$$

Hence B_P^K is a bound for the absolute positiveness of P. □

Since all the previously known bounds are also bounds for absolute positiveness, now we would like to know how they compare with the bound given in this paper.

THEOREM 5.3. (COMPARISON) *Let B_P^C be any bound for the modulus of the complex roots of P, let B_P^K be the Knuth-Johnson bound for the positive real root of P, and let B_P be the bound given in this paper. Further let A_P be the threshold of the absolute positiveness of P. Then we have*

(a) *The ratio $\frac{B_P^C}{A_P}$ can be arbitrarily large (even when the degree of P is fixed).*

(b) *The ratio $\frac{B_P^K}{A_P}$ can be arbitrarily large (even when the degree of P is fixed).*

(c) *The ratio $\frac{B_P}{A_P}$ is bounded above by $\frac{2d}{\ln(2)}$, where d is the degree of P.*

PROOF.

To show (a), consider the quadratic polynomials † of the form $P = x^2 + ax - a$, where $a \geq 1$. By elementary calculation, one sees that

$$A_P = \frac{-a + \sqrt{a^2 + 4a}}{2}.$$

Note also that P has two real roots (and no other roots):

$$\begin{aligned} \alpha &= \frac{-a + \sqrt{a^2 + 4a}}{2} \\ \beta &= \frac{-a - \sqrt{a^2 + 4a}}{2} \end{aligned}$$

† This example was formulated by Dalibor Jakuš and communicated to the author.

Thus, B_P^C is greater than $|\alpha|$ and $|\beta|$. In particular $B_P^C > |\beta|$. Now observe

$$\frac{B_P^C}{A_P} > \frac{|\beta|}{A_P} = \frac{\left|\frac{-a-\sqrt{a^2+4a}}{2}\right|}{\frac{-a+\sqrt{a^2+4a}}{2}} = \frac{a+\sqrt{a^2+4a}}{-a+\sqrt{a^2+4a}} = \frac{a+2+\sqrt{a^2+4a}}{2}.$$

Thus $\frac{B_P^C}{A_P}$ can be arbitrarily large since we can choose arbitrarily large a.

To show (b), consider again the polynomials of the form $P = x^2+ax-a$, where $a \geq 1$. From the definition of B_P^K, we immediately see

$$B_P^K = 2\sqrt{a}.$$

Now observe

$$\frac{B_P^K}{A_P} = \frac{2\sqrt{a}}{\frac{-a+\sqrt{a^2+4a}}{2}} = \sqrt{a}\left(1+\sqrt{1+\frac{4}{a}}\right) > 2\sqrt{a}.$$

Thus $\frac{B_P^K}{A_P}$ can be arbitrarily large since we can choose arbitrarily large a.

To show (c), one only needs to recall Theorem 2.3. The claim follows immediately from the Theorem by setting $n = 1$. The proof is finished.

But to satisfy curiosity, we continue to check the quality of B_P for the particular form of polynomials used for proving the claims (a) and (b). We immediately see that

$$B_P = 2$$

no matter what a is. Thus we have

$$\frac{B_P}{A_P} = \frac{2}{\frac{-a+\sqrt{a^2+4a}}{2}} = \left(1+\sqrt{1+\frac{4}{a}}\right) \leq 1+\sqrt{5}.$$

As expected, we also have

$$1+\sqrt{5} \approx 3.236067978 < 5.70780164 \approx \frac{2\cdot 2}{\ln(2)}.$$

□

I would like to thank Dalibor Jakuš for inspiring discussions. His works on univariate case and Cauchy-like multivariate root bounds provided the initial motivation for starting this work. I also would like to thank the anonymous referees for their interesting suggestions for refinements, for instance Remark 4.1. The term "threshold" of absolute positiveness was kindly suggested by a referee. The proposition 2.1 was suggested by Michael Moeller.

References

Akritas, A., Collins, G. E. (1976). Polynomial real root isolation using Descartes' rule of signs. In *Proceedings of SYMSAC 76*, pages 272–275. ACM, New York.

Arnon, D. S. (1981). *Algorithms for the Geometry of Semi-Algebraic Sets.* PhD thesis, Comp. Sci. Dept., Univ. of Wisconsin-Madison. Tech. Report No. 436.

Ben-Cherifa, A., Lescanne, P. (1987). Termination of rewriting systems by polynomial interpretations and its implementation. *Science of Computer Programming*, **9**(2):137–159.

Berwald, L. (1934). Elementare Sätze über die Abgrenzung der Wurzeln einer algebraischen Gleichung. *Acta. Sci. Math. Litt. Sci. Szeged*, **6**:209–221.

Birkhoff, G. D. (1914). An elementary double inequality for the roots of an algebraic equation having greatest absolute value. *Bull. Amer. Math. Soc.*, **21**:494–495.

Borwein, P., Erdelyi, T. (1995). *Polynomials and Polynomial Inequalities.* Springer Verlag.

Canny, J. (1988). Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th annual ACM symposium on the theory of computing*, pages 460–467.

Carmichael, R. D., Mason, T. E. (1914). Note on the roots of an algebraic equations. *Bull. Amer. Math. Soc.*, **21**:14–22.

Cauchy, A. L. (1829). Exercises de mathematique. *Oeuvres*, **9**(2):122.

Cohn, A. (1922). Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise. *Math. Z.*, **14**:110–148.

Collins, G. E. (1975). Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. In *Lecture Notes In Computer Science*, pages 134–183. Springer-Verlag, Berlin. Vol. 33.

Collins, G. E., Hong, H. (1991). Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, **12**(3):299–328.

Collins, G. E., Johnson, J. R., Kuechlin, W. (1992). Parallel real root isolation using the coefficient sign variation method. In Zippel, R. E., editor, *Proc. Computer Algebra and Parallelism*, pages 71–88. Springer-Verlag.

Collins, G. E., Krandick, W. (1992). An efficient algorithm for infallible polynomial complex root isolation. In Wang, P. S., editor, *ISSAC-92*, pages 189–194. Association for Computing Machinery.

Collins, G. E., Krandick, W. (1993). A hybrid method for high precision calculation of polynomial real roots. In Bronstein, M., editor, *ISSAC-93*, pages 47–52. Association for Computing Machinery.

Collins, G. E., Loos, R. (1976). Polynomial real root isolation by differentiation. In *Proceedings of SYMSAC 76*, pages 15–25, New York. ACM.

Dershowitz, N. (1987). Termination of rewriting. *Journal of Symbolic Computation*, **3**:69–116.

Fujiwara, M. (1915). Über die Wurzeln der algebraischen Gleichungen. *Tôhoku Math. J.*, **8**:78–85.

Giesl, J. (1995). Generating polynomial orderings for termination proofs. In *Proceedings of the 6th International Conference on Rewriting Techniques and Applications.*

Grigor'ev, D. Y. (1988). The complexity of deciding Tarski algebra. *Journal of Symbolic Computation*, **5**(1,2):65–108.

Heindel, L. E. (1970). *Algorithms for Exact Polynomial Root Calculation.* PhD thesis, University of Wisconsin.

Heintz, J., Roy, M.-F., Solernó, P. (1989). On the complexity of semialgebraic sets. In Ritter, G. X., editor, *Proc. IFIP*, pages 293–298. North-Holland.

Hong, H. (1990). *Improvements in CAD-based Quantifier Elimination.* PhD thesis, The Ohio State University.

Hong, H., Jakus, D. (1996). Testing Positiveness of Polynomials. Technical Report 96-02, RISC-Linz, Johannes Kepler University, Linz, Austria.

Johnson, J. R. (1991). *Algorithms for Polynomial Real Root Isolation.* PhD thesis, The Ohio State University.

Johnson, J. R. (1992). Real algebraic number computation using interval arithmetic. In *Proceedings of International Symposium on Symbolic and Algebraic Computation*, pages 195–205. ACM Press.

Kelleher, S. B. (1916). Des limites des zéros dùn polynome. *J. Math. Pres Appl.*, **2**:169–171.

Kuniyeda, M. (1916). Note on the roots of algebraic equations. *Tôhoku Math. J.*, **9**.

Lankford, D. (1976). A finite termination algorithm. Technical report, Southwestern University, Georgetown, Texas.

Lucas, F. (1874). Propriétés géométriques des fractions rationnelles. *C. R. Acad. Sci. Paris*, **77**:431–433.

Marden, M. (1949). *The Geometry of Zeros of a Polynomial in a Complex Variable*, volume 3 of *Mathematics Surveys.* American Mathematics Society, Providence Rhode Island.

McCallum, S. (1984). *An Improved Projection Operator for Cylindrical Algebraic Decomposition.* PhD thesis, University of Wisconsin-Madison.

Milanovic, Mitrinovic, Rassias (1994). *Topics in polynomials: Extremal problems, inequalities, zeros.* World Scientific Publishing.

Montel, P. (1932). Sur La limite supérieure du module des racines dùne équation algébrique. *C. R. Soc. Sci. Varsovie*, **24**:317–326.

Renegar, J. (1992). On the computational complexity and geometry of the first-order theory of the reals. *Journal of Symbolic Computation*, **13**(3):255–300.

Steinbach, J. (1992). Proving Polynomials Positive. In *Proc. 12th FST&TCS, LNCS 652, New Delhi, India.*

Steinbach, J. (1994). *Termination of rewriting.* PhD thesis.

Tarski, A. (1951). *A Decision Method for Elementary Algebra and Geometry.* Univ. of California Press, Berkeley, second edition.

Tôya, T. (1933). Some remarks on Montel's paper concerning upper limit of absolute values of roots of algebraic equations. *Sci. Rep. Tokyo Bunrika Daigaku A*, **1**:275–282.

Weispfenning, V. (1988). The complexity of linear problems in fields. *Journal of Symbolic Computation*, **5**.