

THE HEIGHT OF THE MIXED SPARSE RESULTANT

MARTÍN SOMBRA

ABSTRACT. We present an upper bound for the height of the mixed sparse resultant, defined as the logarithm of the maximum modulus of its coefficients. We obtain a similar estimate for its Mahler measure.

Let $\mathcal{A}_0, \dots, \mathcal{A}_n \subset \mathbb{Z}^n$ be finite sets of integer vectors and let $\text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n} \in \mathbb{Z}[U_0, \dots, U_n]$ be the associated mixed sparse resultant — or $(\mathcal{A}_0, \dots, \mathcal{A}_n)$ -resultant — which is a polynomial in $n + 1$ groups $U_i := \{U_{ia}; a \in \mathcal{A}_i\}$ of $m_i := \#\mathcal{A}_i$ variables each. We refer to [Stu94] and [CLO98, Chapter 7] for the definitions and basic facts.

This resultant is widely used as a tool for polynomial equation solving, a fact that has sparked a lot of interest in its computation, see e.g. [CLO98, Sec. 7.6], [EM99], [D’An02], [JKSS04], while it is also studied from a more theoretical point of view because of its connections with toric varieties and hypergeometric functions, see e.g. [GKZ94], [CDS98].

We assume for the sequel that the family of supports $\mathcal{A}_0, \dots, \mathcal{A}_n$ is essential (see [Stu94, Sec. 1]) which does not represent any loss of generality, by [Stu94, Cor. 1.1].

Set $\mathcal{A} := (\mathcal{A}_0, \dots, \mathcal{A}_n)$, and let $L_{\mathcal{A}} \subset \mathbb{Z}^n$ denote the \mathbb{Z} -module affinely spanned by the pointwise sum $\sum_{i=0}^n \mathcal{A}_i$. This is a subgroup of \mathbb{Z}^n of finite index

$$[\mathbb{Z}^n : L_{\mathcal{A}}] := \#(\mathbb{Z}^n / L_{\mathcal{A}})$$

because we assumed that the family \mathcal{A} is essential. Also set $Q_i := \text{Conv}(\mathcal{A}_i) \subset \mathbb{R}^n$ for the convex hull of \mathcal{A}_i for $i = 0, \dots, n$.

We note by MV the mixed volume function as defined in e.g. [CLO98, Sec. 7.4]: this is normalized so that for a polytope $P \subset \mathbb{R}^n$, the mixed volume $\text{MV}(P, \dots, P)$ equals $n!$ times its Euclidean volume $\text{Vol}_{\mathbb{R}^n}(P)$. We also set $\text{Vol}(P) := \text{MV}(P, \dots, P) = n! \text{Vol}_{\mathbb{R}^n}(P)$.

Under this notation and assumption, the resultant is a multihomogeneous polynomial of degree

$$\deg_{U_i}(\text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n}) = \frac{1}{[\mathbb{Z}^n : L_{\mathcal{A}}]} \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n) > 0$$

with respect to each group of variables U_i , see [PS93, Cor. 2.4].

The *absolute height* of a polynomial $g = \sum_a c_a x^a \in \mathbb{C}[x_1, \dots, x_n]$ is defined as $H(g) := \max\{|c_a|; a \in \mathbb{N}^n\}$. Hereby we will be mainly concerned with its (*logarithmic*) *height*:

$$h(g) := \log H(g) = \log \max\{|c_a|; a \in \mathbb{N}^n\}.$$

Date: April 15, 2004.

2000 Mathematics Subject Classification. Primary 11G50; Secondary 13P99.

Key words and phrases. Mixed sparse resultant, height of polynomials, Mahler measure.

Supported by a Marie Curie post-doctoral fellowship of the European Community Program *Improving Human Research Potential and the Socio-economic Knowledge Base*, contract n° HPMFCT-2000-00709.

The main result of this paper is the following upper bound for the height of the resultant:

Theorem 1.1.

$$h(\text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n}) \leq \frac{1}{[\mathbb{Z}^n : L_{\mathcal{A}}]} \sum_{i=0}^n \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n) \log(\#\mathcal{A}_i).$$

We write for short $\text{Res}_{\mathcal{A}} := \text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n}$ and $\text{MV}_i(\mathcal{A}) := \frac{1}{[\mathbb{Z}^n : L_{\mathcal{A}}]} \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n)$ for $i = 0, \dots, n$. The previous result can thus be rephrased as

$$H(\text{Res}_{\mathcal{A}}) \leq \prod_{i=0}^n (\#\mathcal{A}_i)^{\text{MV}_i(\mathcal{A})}.$$

This improves our previous bound for the unmixed case [Som02, Cor. 2.5] and extends it to the general case. We remark that the obtained upper bound is *polynomial* in the size of the input family of supports \mathcal{A} and in the mixed volumes $\text{MV}_i(\mathcal{A})$, and hence it represents a truly substantial improvement over all previous general estimates. These are the ones which follow either from the Canny-Emiris type formulas (Inequality (4) in the appendix, see also [KPS01, Prop. 1.7] or [Roj00, Thm. 23]) or from direct application of the unmixed case (see the inequality (3) below for $k = 1$).

We also consider the Mahler measure, which is another usual notion for the size of a n -variate polynomial. The *Mahler measure* of $g \in \mathbb{C}[x_1, \dots, x_n]$ is defined as

$$m(g) := \int_{S_1^n} \log |g| d\mu^n,$$

where $S_1 \subset \mathbb{C}$ is the unit circle and $d\mu$ is the Haar measure over S_1 of total mass 1. This can be compared with the height: in our case

$$(1) \quad - \sum_{i=0}^n \text{MV}_i(\mathcal{A}) \log(m_i) \leq m(\text{Res}_{\mathcal{A}}) - h(\text{Res}_{\mathcal{A}}) \leq \sum_{i=0}^n \text{MV}_i(\mathcal{A}) \log(m_i)$$

by [KPS01, Lem. 1.1]. We refer to [KPS01, Sec. 1.1.1] for an account on some of the notions of height of complex polynomials: just note that the height $h(g)$ here coincides with $\log |g|_{\infty}$ in that reference.

We obtain the same estimate as before for the Mahler measure of the resultant.

Theorem 1.2.

$$m(\text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n}) \leq \frac{1}{[\mathbb{Z}^n : L_{\mathcal{A}}]} \sum_{i=0}^n \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n) \log(\#\mathcal{A}_i).$$

Note that this improves by a factor of 2 the estimate which would derive from direct application of Theorem 1.1 and the inequalities (1) above.

Both estimates are a consequence of the following:

Lemma 1.3. *Let $f_0 \in \mathbb{C}^{\mathcal{A}_0}, \dots, f_n \in \mathbb{C}^{\mathcal{A}_n}$. Then*

$$\log |\text{Res}_{\mathcal{A}_0, \dots, \mathcal{A}_n}(f_0, \dots, f_n)| \leq \frac{1}{[\mathbb{Z}^n : L_{\mathcal{A}}]} \sum_{i=0}^n \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n) \log \|f_i\|_1,$$

where $\|f_i\|_1 := \sum_{a \in \mathcal{A}_i} |f_{ia}|$ denotes the ℓ^1 -norm of the vector $f_i = (f_{ia}; a \in \mathcal{A}_i)$.

Let $g = \sum_a c_a x^a \in \mathbb{C}[x_1, \dots, x_n]$. Then for $a \in \mathbb{N}^n$ we have that

$$c_a = \int_{S_1^n} \frac{g(z_1, \dots, z_n)}{z_1^{a_1+1} \dots z_n^{a_n+1}} d\mu^n$$

by Cauchy's formula and so $h(g) \leq \sup \{ \log |g(\xi)|; \xi \in S_1^n \}$. Thus Theorem 1.1 is a consequence of this inequality applied to $g := \text{Res}_{\mathcal{A}}$, together with Lemma 1.3.

On the other hand, Theorem 1.2 follows from Lemma 1.3 by a straightforward estimation of the integral in the definition of the Mahler measure.

Proof of Lemma 1.3.— Let $k \in \mathbb{N}$. Then let $k\mathcal{A}_i \subset \mathbb{Z}^n$ denote the pointwise sum of k copies of \mathcal{A}_i , and set $k\mathcal{A} := (k\mathcal{A}_0, \dots, k\mathcal{A}_n)$. It is easy to verify that $k\mathcal{A}$ is also essential, $L_{k\mathcal{A}} = L_{\mathcal{A}}$ and $\text{Conv}(k\mathcal{A}_i) = kQ_i$.

We identify each $f_i \in \mathbb{C}^{\mathcal{A}_i}$ with the corresponding Laurent polynomial $f_i = \sum_{a \in \mathcal{A}_i} f_{ia} x^a$, and we set $f_i^k \in \mathbb{C}^{k\mathcal{A}_i}$ for the vector which corresponds to the k -th power of f_i . By the factorization formula for resultants [PS93, Prop. 7.1] we get that

$$\text{Res}_{k\mathcal{A}}(f_0^k, \dots, f_n^k) = \text{Res}_{\mathcal{A}}(f_0, \dots, f_n)^{k^{n+1}}$$

and so

$$\begin{aligned} k^{n+1} \log |\text{Res}_{\mathcal{A}}(f_0, \dots, f_n)| &\leq h(\text{Res}_{k\mathcal{A}}) + \sum_{i=0}^n \text{MV}_i(k\mathcal{A}) \log \|f_i^k\|_1 \\ (2) \qquad \qquad \qquad &\leq h(\text{Res}_{k\mathcal{A}}) + k^{n+1} \sum_{i=0}^n \text{MV}_i(\mathcal{A}) \log \|f_i\|_1. \end{aligned}$$

The first inequality follows from the straightforward estimate $|G(u_0, \dots, u_n)| \leq H(G) \prod_{i=0}^n \|u_i\|_1^{d_i}$ for a multihomogeneous polynomial G of degree d_i in each group of variables, applied to $G := \text{Res}_{k\mathcal{A}}$ and $u_i := f_i^k$. The second one follows from the linearity of the mixed volume, and the sub-additivity of the ℓ^1 -norm with respect to polynomial multiplication (which implies that $\log \|f_i^k\|_1 \leq k \log \|f_i\|_1$).

Now let $\mathcal{B} \subset \mathbb{Z}^n$ be any finite set such that $L_{\mathcal{B}} = \mathbb{Z}^n$ and such that $\mathcal{A}_0, \dots, \mathcal{A}_n \subset \mathcal{B}$. Set $n(k) := \#k\mathcal{B}$ and $P := \text{Conv}(\mathcal{B}) \subset \mathbb{R}^n$. Then the (unmixed) resultant $\text{Res}_{k\mathcal{B}}$ is a polynomial in $(n+1)n(k)$ variables and total degree $(n+1)\text{Vol}(kP) = (n+1)k^n \text{Vol}(P)$. We have also that $L_{k\mathcal{B}} = \mathbb{Z}^n$ and so we are in the hypothesis of [Som02, Cor. 2.5], which gives the height estimate

$$h(\text{Res}_{k\mathcal{B}}) \leq 2(n+1) \log(n(k)) \text{Vol}(kP) = 2(n+1) \log(n(k)) k^n \text{Vol}(P).$$

We have that $k\mathcal{A}_i \subset k\mathcal{B}$ for $i = 0, \dots, n$ and so by [Stu94, Cor. 4.2] there exists a monomial order \prec such that $\text{Res}_{k\mathcal{A}}$ divides the initial form $\text{init}_{\prec}(\text{Res}_{k\mathcal{B}})$. This is a polynomial in $(n+1)n(k)$ variables of degree and height bounded by those of $\text{Res}_{k\mathcal{B}}$, and so

$$\begin{aligned} h(\text{Res}_{k\mathcal{A}}) &\leq h(\text{Res}_{k\mathcal{B}}) + 2 \log((n+1)n(k) + 1) (n+1) k^n \text{Vol}(P) \\ (3) \qquad \qquad &\leq 4(n+1) \log((n+1)n(k) + 1) k^n \text{Vol}(P) \end{aligned}$$

by the inequality $h(f) \leq h(g) + 2 \deg(g) \log(n+1)$, which holds for $f, g \in \mathbb{Z}[x_1, \dots, x_n]$ such that $f|g$ (see [KPS01, Lem. 1.2(1.d)]) applied to $f := \text{Res}_{k\mathcal{A}}$ and $g := \text{init}_{\prec}(\text{Res}_{k\mathcal{B}})$.

Finally we set $\mathcal{B} := b + d[0, 1]^n \subset \mathbb{R}^n$ where $[0, 1]$ denotes the unit interval of \mathbb{R} , for some $b \in \mathbb{Z}^n$ and $d \in \mathbb{N}$ such that $\mathcal{A}_0, \dots, \mathcal{A}_n \subset b + d[0, 1]^n$. Then $n(k) =$

$\log(\#(kb + kd[0, 1]^n \cap \mathbb{Z}^n)) = \log(kd + 1)^n = O_k(\log k)$ (here the notation O_k refers to the dependence on k) and so

$$h(\text{Res}_{k\mathcal{A}}) = O_k(k^n \log k).$$

Note that alternatively, we could have obtained this from the inequality (4) in the appendix.

Togther with the inequality (2) this implies that

$$\log |\text{Res}_{\mathcal{A}}(f_0, \dots, f_n)| \leq \sum_{i=0}^n \text{MV}_i(\mathcal{A}) \log \|f_i\|_1 + O_k\left(\frac{\log k}{k}\right),$$

from where we conclude by letting $k \rightarrow \infty$. \square

Let us consider some examples. For short we set $H(\mathcal{A}) := H(\text{Res}_{\mathcal{A}})$ and $E(\mathcal{A}) := \prod_{i=0}^n (\#\mathcal{A}_i)^{\text{MV}_i(\mathcal{A})}$; we also set

$$q(\mathcal{A}) := \frac{\log E(\mathcal{A})}{\log H(\mathcal{A})}$$

for the quotient between the height of the resultant and the estimate from Theorem 1.1.

Example 1.1. *Sylvester resultants.* For $d \in \mathbb{N}$ we let

$$\mathcal{A}_0(d) = \mathcal{A}_1(d) := \{0, 1, 2, \dots, d\} \subset \mathbb{Z}.$$

The corresponding resultant coincides with the Sylvester resultant of two univariate polynomials of the same degree d . In this case $\text{MV}_0(d) = \text{MV}_1(d) = d$ and $\#\mathcal{A}_0(d) = \#\mathcal{A}_1(d) = d + 1$, and so $E(d) := E(\mathcal{A}_0(d), \mathcal{A}_1(d)) = (d + 1)^{2d}$.

We compute the height $H(d) := H(\mathcal{A}_0(d), \mathcal{A}_1(d))$ for $2 \leq d \leq 7$ with the aid of Maple and we collect the results in the following comparative table:

d	2	3	4	5	6	7
$H(d)$	2	3	10	23	78	274
$E(d)$	81	4,096	390,625	60,466,176	13,841,287,201	4,398,046,511,104
$q(d)$	6.33	7.57	5.59	5.71	5.35	5.18

Example 1.2. We take this example from [EM99, Example 3.5]. Let

$$\mathcal{A}_0 := \{(0, 0), (1, 1), (2, 1), (1, 0)\},$$

$$\mathcal{A}_1 := \{(0, 1), (2, 2), (2, 1), (1, 0)\},$$

$$\mathcal{A}_2 := \{(0, 0), (0, 1), (1, 1), (1, 0)\}.$$

Then $\text{MV}_0 = 4$, $\text{MV}_1 = 3$ and $\text{MV}_2 = 4$, so that $E(\mathcal{A}) = 4^4 4^3 4^4$. On the other hand, we can compute the resultant using its expression in [EM99, Example 3.19] as a quotient of determinants and we obtain that $H(\mathcal{A}) = 8$. Hence

$$H(\mathcal{A}) = 8 \quad , \quad E(\mathcal{A}) = 4,194,304 \quad , \quad q(\mathcal{A}) = 7.33 .$$

For reference, the straightforward estimation of $H(\mathcal{A})$ via the Canny-Emiris formula (see the appendix below) gives:

$$H(\mathcal{A}) \leq 2^{82} = 4,835,703,278,458,516,698,824,704 .$$

Example 1.3. We take this example from [Stu94, Example 2.1]. Let

$$\begin{aligned}\mathcal{A}_0 &:= \{(0, 0), (2, 2), (1, 3)\}, \\ \mathcal{A}_1 &:= \{(0, 1), (2, 0), (1, 2)\}, \\ \mathcal{A}_2 &:= \{(3, 0), (1, 1)\}.\end{aligned}$$

Then $MV_0 = 5$, $MV_1 = 7$ and $MV_2 = 7$, so that $E(\mathcal{A}) = 3^5 3^7 2^7$. From the explicit monomial expansion of the resultant (see [Stu94, Example 2.1]) we find that $H(\mathcal{A}) = 14$ and so

$$H(\mathcal{A}) = 14 \quad , \quad E(\mathcal{A}) = 68,024,448 \quad , \quad q(\mathcal{A}) = 6.83 .$$

These examples show that there is still some room for improvement over Theorem 1.1. It is however possible that our estimate is quite sharp anyway: in spite of the large difference between $H(\mathcal{A})$ and $E(\mathcal{A})$ in the computed examples, the quotient $q(\mathcal{A})$ is quite small, and moreover it does not seem to grow when $E(\mathcal{A}) \rightarrow \infty$.

In any case, it would be very interesting to have an *exact* expression for $h(\text{Res}_{\mathcal{A}})$ — as was remarked to me by B. Sturmfels — or at least a non trivial lower bound. Note that the only information that we dispose about the exact value of the coefficients of $\text{Res}_{\mathcal{A}}$ is for the extremal ones, which are equal to ± 1 [Stu94, Cor. 3.1].

Remark 1.4. *After a first version of this paper was circulating, C. D’Andrea (personal communication) obtained a non trivial lower bound for the height of the Sylvester resultant, and an improvement of the upper bound for this case: in the notation of Example 1.1 above, he obtains that $H(d) \leq d!$.*

APPENDIX: ESTIMATION OF THE HEIGHT VIA THE CANNY-EMIRIS FORMULA

For purpose of easy reference, we establish herein the estimate for $h(\text{Res}_{\mathcal{A}})$ which follows from the Canny-Emiris formula and the standard estimates for the behavior of the height of polynomials under addition, multiplication and division.

Assume that $L_{\mathcal{A}} = \mathbb{Z}^n$ and set $Q := \sum_{i=0}^n Q_i \subset \mathbb{R}^n$. Let $\mathcal{M}_0, \dots, \mathcal{M}_n$ be a family of Canny-Emiris (square, non singular) matrices for \mathcal{A} ; we refer to [CLO98, Sec. 7.6] for their precise definition. In the sequel we just describe the aspects needed for the height estimate.

A family of Canny-Emiris matrices is not unique, as their construction depends on a choice of a coherent mixed subdivision of Q and of a sufficiently small and generic vector $\delta \in \mathbb{Q}^n$. Set

$$\mathcal{E} := (Q + \delta) \cap \mathbb{Z}^n.$$

The procedure then uses the given subdivision of Q to split this set into a disjoint union $\mathcal{E} = \mathcal{E}_0(j) \cup \dots \cup \mathcal{E}_n(j)$, for each $0 \leq j \leq n$. The elements in \mathcal{E} are in bijection with the rows of \mathcal{M}_j , and to each $p \in \mathcal{E}_i(j)$ corresponds a row of \mathcal{M}_j with exactly $m_i = \#\mathcal{A}_i$ non zero entries, which consist of the variables in $U_i := \{U_{ia}; a \in \mathcal{A}_i\}$.

Set $D_j := \det(\mathcal{M}_j) \in \mathbb{Z}[U_0, \dots, U_n] \setminus \{0\}$. The *Canny-Emiris formula* [CLO98, Ch. 7, Thm. 6.12] states that $\text{Res}_{\mathcal{A}} = \gcd(D_0, \dots, D_n)$.

Then D_0 is a multihomogeneous polynomial of degree $N_i := \#\mathcal{E}_i(0)$ in each set of variables U_i and of height bounded by $\sum_{i=0}^n N_i \log(m_i)$. We have that $\text{Res}_{\mathcal{A}} | D_0$ and

so $m(\text{Res}_{\mathcal{A}}) \leq m(D_0)$, which combined with [KPS01, Lem. 1.1] gives

$$(4) \quad \begin{aligned} h(\text{Res}_{\mathcal{A}}) &\leq h(D_0) + \sum_{i=0}^n (N_i + MV_i(\mathcal{A})) \log(\#\mathcal{A}_i) \\ &\leq \sum_{i=0}^n (2N_i + MV_i(\mathcal{A})) \log(\#\mathcal{A}_i). \end{aligned}$$

Applied to Example 1.2, this gives the stated estimate: $N_0 = N_1 = 4$ and $N_2 = 7$ (see [EM99, Example 3.5]) and so the previous estimate gives $H(\mathcal{A}) \leq 4^{2 \cdot 15 + 11} = 2^{82}$.

In general, the estimate so obtained is *much* worse than that of Theorem 1.1, especially for $n \gg 0$. Consider e.g. $\mathcal{A}_i := \{0, \dots, d\}^n \subset \mathbb{Z}^n$ for $i = 0, \dots, n$. Then it is easy to show that Inequality (4) gives

$$h(\text{Res}_{\mathcal{A}}) \leq (2((n+1)d)^n + (n+1)!d^n) \log(d+1)^n = n(2(n+1)^n + (n+1)!)d^n \log(d+1)$$

while Theorem 1.1 gives $h(\text{Res}_{\mathcal{A}}) \leq n(n+1)!d^n \log(d+1)$.

Acknowledgements. I thank Carlos D’Andrea for several helpful discussions, and for having communicated me his work on the height of the Sylvester resultant. I also thank Matilde Lalín and Carlos D’Andrea for disproving a conjecture I made in a previous version of this paper, on the value of $m(\text{Res}_{\mathcal{A}})$,

REFERENCES

- [CDS98] E. Cattani, A. Dickenstein, B. Sturmfels, *Residues and resultants*, J. Math. Sci. Univ. Tokyo **5** (1998) 119-148.
- [CLO98] D. Cox, J. Little, D. O’Shea, *Using algebraic geometry*, Graduate Texts in Mathematics **185**, Springer, 1998.
- [D’An02] C. D’Andrea, *Macaulay style formulas for sparse resultants*, Trans. Amer. Math. Soc. **354** (2002) 2595-2629.
- [EM99] I. Emiris, B. Mourrain, *Matrices in elimination theory*, J. Symbolic Comput. **28** (1999) 3-44.
- [GKZ94] I.M. Gelfand, M.M. Kapranov, A.V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Birkhäuser, 1994.
- [JKSS04] G. Jeronimo, T. Krick, J. Sabia, M. Sombra, *The computational complexity of the Chow form*, Found. Comput. Math. **4** (2004) 41-117.
- [KPS01] T. Krick, L.M. Pardo, M. Sombra, *Sharp estimates for the arithmetic Nullstellensatz*, Duke Math. J. **109** (2001) 521-598.
- [PS93] P. Pedersen, B. Sturmfels, *Product formulas for resultants and Chow forms*, Math. Z. **214** (1993) 377-396.
- [Roj00] J.M. Rojas, *Algebraic geometry over four rings and the frontier to tractability*, Contemp. Math. **270**, Amer. Math. Soc. 2000, 275-321.
- [Som02] M. Sombra, *Minimums successifs des variétés toriques projectives*, to appear in the J. Reine Angew. Math., 25 pp.. E-print: math.NT/0209195.
- [Stu94] B. Sturmfels, *On the Newton polytope of the resultant*, J. Algebraic Combin. **3** (1994) 207-236.

UNIVERSITÉ DE PARIS 7, UFR DE MATHÉMATIQUES, ÉQUIPE DE GÉOMÉTRIE ET DYNAMIQUE, 2 PLACE JUSSIEU, 75251 PARIS CEDEX 05, FRANCE.

E-mail address: sombra@math.jussieu.fr