Zero-patterns of polynomials and Newton polytopes *

Alan G.B. Lauder[†]

May 10, 2002

1 Introduction

We prove a simple upper bound on the number of regions into which the torus over an arbitrary field may be paritioned by a finite collection of algebraic hypersurfaces. This is a polytope-refinement to an existing result, after the fashion of many other recent improvements to classical bounds.

Let $\mathbf{f} = (f_1, \ldots, f_m)$ be a sequence of polynomials in $F[X_1, \ldots, X_n]$ where F is a field and $X_1, \ldots, X_n$ commuting indeterminates. Given any point u in affine space F^n of dimension n over F the *zero-pattern* $\delta(\mathbf{f}, u)$ is the vector $(\delta_1, \ldots, \delta_m)$, where each $\delta_i \in \{0, *\}$ with $\delta_i = 0$ if and only

*The author gratefully acknowledges the support of the EPSRC and St John's College, Oxford, and thanks Professors Richard Brent, Shuhong Gao and Lajos Rónyai for helpful comments on the paper. *Mathematics Subject Classification 2000*: Primary 05A99, 12E05, 52B20 *Key words and phrases*: multivariate polynomial, zero-pattern, Newton polytope.

†Computing Laboratory, Oxford University, Oxford OX1 3QD, U.K. E-mail: alan.lauder@comlab.ox.ac.uk.

if $f_i(u) = 0$. Let $Z_F(\mathbf{f})$ denote the number of distinct zero-patterns as u ranges over F^n. Let F^* denote the set of non-zero elements, and $Z_F^*(\mathbf{f})$ the number of distinct zero patterns as u ranges over the torus $(F^*)^n$.

In [9] upper bounds are proved for $Z_F(\mathbf{f})$ in terms of the degrees of the polynomials in the sequence, and a number of combinatorial applications of these bounds are presented. The essence of these results is that if all the polynomials have degree bound by d, say, then there exists a polynomial function $E_d(m)$ of degree n which bounds the number of zero-patterns. Note that this improves the trivial exponential bound of 2^m. We present some refinements of these bounds, and also new bounds for $Z_F^*(\mathbf{f})$, which take into account the non-zero terms which actually occur in the polynomials, rather than just their degrees. More precisely, we replace $E_d(m)$ by a more refined polynomial function $E_\Delta(m)$ related to the Newton polytopes of the polynomials. Our refinement is of the same spirit as many other results in which a quantity related to a single polynomial or family of polynomials is bound in terms of their Newton polytopes, rather than their degrees. The most famous example is perhaps Bernstein's theorem, the polytope version of Bezout's theorem which gives a bound on the number of common zeros in the torus of n polynomials in n unknowns (see [4, 10]). Similar polytope-bounds also occur in polynomial factorisation [6]. At a deeper level, Adolphson and Sperber's [1] total degree bound on the zeta function of a hypersurface in terms of the volume of the Newton polytope of the defining polynomial improves earlier work of Bombieri. Indeed, this latter bound is used to refine the complexity estimates in the author and Wan's proof of the polynomial-time computability of these zeta functions in small characteristic [8].

Over the real field, one may also obtain bounds on the number of "sign-patterns", see [9]. Both zero-patterns and sign-patterns have been considered by a number of authors, and the bounds obtained have variety of combinatorial applications [2, 9]. We would be very interested in learning of any new combinatorial applications of our bounds. In [9] explicit constructions are given to show that their bounds are optimal to within a factor $(7.25)^n$, assuming m and the field F are suitably large. An easy modification of their approach shows that our bounds are optimal within a factor of 2^n, for suitably large m and field F.

2 Theorems

We introduce further notation necessary to explain our results. Let $\mathbb{R}^n$ denote real n-dimensional affine space, and Δ any bounded subset of $\mathbb{R}^n$. A *lattice point* in $\mathbb{R}^n$ is simply a point whose coordinates lie in $\mathbb{Z}$. We denote by $\#(\Delta)$ the number of lattice points which lie in Δ. For any $f \in F[X_1, \ldots, X_n]$ the *support set* is the set of all integer vectors $(e_1, \ldots, e_n) \in \mathbb{R}^n$ which occur as exponents of non-zero terms $aX_1^{e_1} \ldots X_n^{e_n}$ in f. The *Newton polytope* of f, written $N^*(f)$, is defined as the convex hull of the support set. It is a convex polytope in $\mathbb{R}^n$. We also introduce one further polytope associated to the polynomial. Let $\mathbb{R}_{\leq 0}$ and $\mathbb{R}_{\geq 0}$ denote the set of non-positive and non-negative real numbers, respectively. Define $N(f)$ as

$$N(f) = (N^*(f) + (\mathbb{R}_{\leq 0})^n) \cap (\mathbb{R}_{\geq 0})^n.$$

Here the summation is the Minkowski sum, defined for $A, B \subseteq \mathbb{R}^n$ by $A + B := \{a + b \,|\, a \in A, b \in B\}$. Notice that $N^*(0), N(0) = \emptyset$. For this reason it is convenient in Theorem 1 to assume that none of the polynomials f_i in the sequence are zero; of course, this is no real restriction.

Theorem 1 *Let* $\mathbf{f} = (f_1, \ldots, f_m)$ *be a sequence of non-zero polynomials in* $F[X_1, \ldots, X_n]$ *and* $Z_F(\mathbf{f})$, $Z_F^*(\mathbf{f})$ *the number of zero-patterns of* $\mathbf{f}$ *over affine and toric space, respectively. Write* $f = f_1 f_2 \ldots f_m$*, so* $f \neq 0$*. Then*

$$\begin{aligned} Z_F(\mathbf{f}) &\leq \#(N(f)) \\ Z_F^*(\mathbf{f}) &\leq \#(N^*(f)) \end{aligned}$$

where the righthand sides are the number of lattice points in certain polytopes associated to f*, as defined above.*

Note that Theorem 1.1 in [9] can be recovered by taking the bound $\#(N(f)) \leq \binom{n+d}{n}$ where $d(= \sum_{i=1}^m d_i\,,\, d_i = \deg f_i)$ is the degree of f. A similar refinement to Theorem 4.1 in [9] can also be obtained.

If all of the polynomials f_i in the sequence have the same Newton polytope Δ then we can get a simple bound in terms of the Erhart polynomial of Δ [7, page 780].

Corollary 2 *Let* $\mathbf{f} = (f_1, \ldots, f_m)$ *be a sequence of non-zero polynomials in* $F[X_1, \ldots, X_n]$ *and* $Z_F^*(\mathbf{f})$ *be the number of zero-patterns of* $\mathbf{f}$ *in the torus. Suppose that each* f_i *has Newton polytope* Δ*. Then* $Z_F^*(\mathbf{f}) \leq E_\Delta(m)$*, where* E_Δ *is the Erhart polynomial of* Δ*.*

An alternative bound for $Z_F(\mathbf{f})$ in terms of $N^*(f)$ can be obtained by using the bound for $Z_F^*(\mathbf{f})$ from Theorem 1 and a torus decomposition of affine space. For $T \subseteq \{1, 2, \ldots, n\}$ and any $g \in F[X_1, \ldots, X_n]$ define $g|_T$ to be the polynomial obtained by setting to zero all X_i for $i \in T$ which occur in g. Let

$$f_T := \prod_{f_i|_T \neq 0} f_i|_T.$$

We use the convention $f_T = 1$ when $\{i \mid f_i|_T \neq 0\} = \emptyset$.

$$N_T := \#(N^*(f_T)).$$

(Note that $N_T = 1$ in the case that $\{i \mid f_i|_T \neq 0\} = \emptyset$; also, when this is not the case N_T is just the number of lattice points in the intersection of $N^*(f)$ with the appropriate coordinate axes.)

Theorem 3 *Let $\mathbf{f} = (f_1, \ldots, f_m)$ be a sequence of polynomials in $F[X_1, \ldots, X_n]$ and $Z_F(\mathbf{f})$ the number of zero-patterns in affine space. Then $Z_F(\mathbf{f}) \leq \sum_T N_T$, where the sum is over all sets $T \subseteq \{1, 2, \ldots, n\}$ and N_T is as defined above.*

The above result suggests a natural generalisation: suppose each of the polynomials f_i has a common Newton polytope Δ. As in [3, Sections 2.1,3.1] one may associate a toric variety "T_Δ" with the integral polytope Δ, and a hypersurface "$\overline{Z}_{f_i,\Delta}$" with each of the polynomials f_i. The number of distinct zero-patterns over all points "$u \in T_\Delta$" for the sequence of hypersurfaces may then be defined, and a bound on this proved in terms of the lattice points in Δ; however, we do not pursue this.

We finish with two examples to illustrate Theorems 1 and 3.

Example 4 Take the case $n = 2$ with polynomials in $F[X, Y]$. Consider the sequence $\mathbf{f} = (X, Y, X + Y)$ so $f = XY(X + Y)$. The original bound from Theorem 1.1 in [9] is 10. We have $\#(N(f)) = 8$, $\#(N^*(f)) = 2$, with $Z_F(\mathbf{f}) = 5$ and $Z_F^*(\mathbf{f}) = 2$. Hence only one of the bounds in Theorem 1 is sharp in this case. However, using the bound from Theorem 3, we find that $N_\emptyset = 2$, $N_{\{1\}} = 1$, $N_{\{2\}} = 1$ and $N_{\{1,2\}} = 1$. Thus the bound from Theorem 3 is tight in this example.

Example 5 (This example is based upon [9, Sections 6,7].) Assume that $r := m/n$ is an integer, and let $d_1, \ldots, d_n$ be positive integers each at most m. Assume $\#(F) > 1 + \lfloor r(\max_i(d_i) + 1)/2 \rfloor$. Let d be one of the integers $d_1, \ldots, d_n$. Let $A(r, d)$ be a collection of subsets $\{S_1, \ldots, S_{M(r,d)}\}$ of $\{1, \ldots, r\}$ such that each element occurs in at most d subsets. By the comment following [9, Proposition 6.2], we may take $M(r, d) = 1 + \lfloor r(d + 1)/2 \rfloor$. Construct a sequence of polynomials $f_1, \ldots, f_r \in F[X]$ each of degree bounded by d as follows: Choose a set of $M(r, d)$ distinct elements $u_1, \ldots, u_{M(r,d)} \in F^*$, and define $f_j := \prod_{k,\, j \in S_k} (X - u_k)$. By construction $\{u_1, \ldots, u_{M(r,d)}\}$ is a complete set of witnesses to the zero patterns of $\mathbf{f} = (f_1, \ldots, f_r)$. Hence $Z_F^*(\mathbf{f}) = M(r, d) = 1 + \lfloor r(d + 1)/2 \rfloor$. Repeating this construction n times, we can find a sequence $\mathbf{f}$ of $nr = m$ polynomials, the first r univariate in X_1 of degree at most d_1, and so on, such that $Z_F^*(\mathbf{f}) = \prod_{i=1}^n (1 + \lfloor r(d_i + 1)/2 \rfloor)$. The upper bound from Theorem 1 in this case is $\prod_{i=1}^n (rd_i + 1)$. Hence we are within a factor of 2^n. (Also here $Z_F(\mathbf{f}) = Z_F^*(\mathbf{f})$ and the bounds from Theorem 1 for affine space are the same as for the torus.) In [9] the upper bound is within a factor of $(7.25)^n$, and so we have a slight improvement.

3 Proofs

Proof of Theorem 1 : We follow the proof of Theorem 1.1 in [9], making appropriate modifications. Assume $M = Z_F(\mathbf{f})$ and let $u_1, \ldots, u_M$ be witnesses to the distinct zero-patterns. (That is, the set of zero patterns is precisely $\{\delta(\mathbf{f}, u_i)\}_{1 \leq i \leq M}$.) The *support set* of a zero-pattern $\delta(\mathbf{f}, u_i)$ is just the set of indices $S_i \subseteq \{1, 2, \ldots, m\}$ which mark $*$'s in the zero-pattern. Define

$$g_i = \prod_{k \in S_i} f_k$$

and so

$$g_i(u_j) \neq 0 \text{ if and only if } S_i \subseteq S_j. \tag{1}$$

Now the polynomials $g_1, \ldots, g_M$ are linearly independent over F, exactly as proved in [9, page 721]. Moreover, each is a factor of $f = f_1 f_2 \ldots f_m$, a non-zero polynomial. It follows from Lemma 6 below that each $N^*(g_i)$ lies in the polytope $N(f)$. The dimension over F of the space of all polynomials whose Newton polytopes lie in $N(f)$ is exactly $\#(N(f))$, and the first inequality follows.

For the second inequality, now let $M = Z_F^*(\mathbf{f})$ and $u_1, \ldots, u_M$ be witnesses for the distinct zero patterns in the torus $(F^*)^n$. Define the g_i exactly as before. Each g_i is a factor of $f (\neq 0)$. Thus by Lemma 6 we can find a monomial r_i such that $N^*(r_i g_i) \subseteq N^*(f)$. We claim that $r_1 g_1, \ldots, r_M g_M$ are linearly independent as polynomials over F. To prove this, assume that a nontrivial linear relation $\sum_{i=1}^{M} \lambda_i (r_i g_i) = 0$ exists ($\lambda_i \in F$). Let j be a subscript such that $|S_j|$ is minimal among the S_i with $\lambda_i \neq 0$. Substitute $u_j (\in (F^*)^n)$ in the relation. Now $\lambda_j r_j(u_j) g_j(u_j) \neq 0$, since $g_j(u_j) \neq 0$ from (1) and r_j is a monomial with u_j having no zero coordinates. However,

$\lambda_i r_i(u_j) g_i(u_j) = 0$ for all $i \neq j$ since $g_i(u_j) = 0$, from (1) and the minimality of S_j. This is a contradiction, establishing the linear independence of $r_1 g_1, \ldots, r_M g_M$. Now each polynomial in this sequence lies in the space of all polynomials over F spanned by the monomials whose exponents are lattice points in $N^*(f)$. Hence $M \leq \#(N^*(f))$, as required.

$\square$

In the proof we used the following lemma.

Lemma 6 *Let $f, g \in F[X_1, \ldots, X_n]$ with $f \neq 0$ and g a factor of f. Then $N^*(g) \subseteq N(f)$. Moreover, there exists a monomial r such that $N^*(rg) \subseteq N^*(f)$.*

Proof: Let $f = gh$ and so $N^*(f) = N^*(g) + N^*(h)$. (See, for example, [5, Lemma 2.1].) We first show $N^*(g) \subseteq N(f) (= (N^*(f) + (\mathbb{R}_{\leq 0})^n) \cap (\mathbb{R}_{\geq 0})^n)$. Let $w \in N^*(g)$ be any point. Then there exists some $u \in N^*(h)$ with $v := w + u \in N^*(f)$. Now $u \in (\mathbb{R}_{\geq 0})^n$ and so $-u \in (\mathbb{R}_{\leq 0})^n$. So $w = v - u \in N^*(f) + (\mathbb{R}_{\leq 0})^n$ and certainly $w \in (\mathbb{R}_{\geq 0})^n$. Hence $w \in N(f)$ and so $N^*(g) \subseteq N(f)$, as required.

For the second part, select a vertex v, say, of $N^*(f)$ with supporting hyperplane l, say. Let w be the vertex of $N^*(g)$ which is supported by l, with the same inner normal. Define $z = v - w$, a vertex of $N^*(h)$. Then $N^*(g) + z \subseteq N^*(g) + N^*(h) = N^*(f)$. Write $r = X_1^{z_1} \ldots X_n^{z_n}$ where $z = (z_1, \ldots, z_n)$. It follows that $N^*(rg) (= N^*(g) + z) \subseteq N^*(f)$, as we wished to show.

$\square$

Proof of Corollary 2 : In this case $N^*(f) = \sum_{i=1}^{m} N^*(f_i) = m\Delta$. Hence, from the properties [7, page 780] of the Erhart polynomial, $\#(N^*(f)) =$

$\#(m\Delta) = E_\Delta(m)$, and the corollary now follows from the second estimate in Theorem 1.

$\square$

Proof of Theorem 3 : For each $T \subseteq \{1, 2, \ldots, n\}$ let

$$\mathbb{G}_T := \{r \in F^n \mid x_i(r) = 0 \text{ if and only if } i \in T\}$$

where x_i are coordinate functions on F^n. Then the torus decomposition expresses F^n as the disjoint union of these sets over all T. The number of zero-patterns over F^n is certainly bounded by the sum $\sum_T M_T$ of the number of zero-patterns, M_T say, over each $\mathbb{G}_T$. Thus it suffices to show that $M_T \leq N_T$.

Now on the torus $\mathbb{G}_T$ each polynomial in the sequence $f_1, \ldots, f_m$ takes the same value as the corresponding polynomial in the sequence $f_1|_T, \ldots, f_m|_T$. Thus we need to show that the number of zero-patterns of this latter sequence on $\mathbb{G}_T$ is bounded by N_T. If this is the zero sequence, then by definition $N_T = 1$, which is exactly the number of zero-patterns. Otherwise, $f_T \neq 0$ is a polynomial, and the Newton polytope $N^*(f_T)$ is defined.

Now let $x_1, \ldots, x_n$ be coordinate functions for $\mathbb{R}^n$ and for each $T \subseteq \{1, 2, \ldots, n\}$ denote $\mathbb{R}_T = \{r \in \mathbb{R}^n \mid x_i(r) = 0 \text{ for } i \in T\}$. (Note that we do not assume that $x_i(r) \neq 0$ for $r \in \mathbb{R}_T$ and $i \notin T$.) Thus $N^*(f_T)$ is a polytope in $\mathbb{R}_T$. We may then apply the second bound in Theorem 1 with $f, N^*(f)(\subseteq \mathbb{R}^n)$ and $(F^*)^n$ replaced by $f_T, N^*(f_T)(\subseteq \mathbb{R}_T)$ and $\mathbb{G}_T$. We deduce that the number of zero-patterns of this sequence in $\mathbb{G}_T$ is bounded by $\#(N^*(f_T))$. But this last quantity is just N_T by definition, which completes the proof.

$\square$

References

[1] A. Adolphson and S. Sperber, Newton polyhedra and the degree of the L-functions associated to an exponential sum, Invent. Math., 88, (1987), 555-569.

[2] N. Alon, Tools from higher algebra, Handbook of Combinatorics, Elsevier and MIT Press, 1995 (R. Graham, M. Grötschel, L. Lovász eds.), 1749-1783.

[3] V.V. Batyrev, Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties, J. Algebraic Geometry **3** (1994), 493-535.

[4] D.N. Bernstein, The number of roots of a system of equations, *Functional Analysis Appl.* 9 (1975), 1-4.

[5] S. Gao, Absolute irreducibility of polynomials via Newton polytopes, *J. Algebra* **237** (2001), 501-520.

[6] S. Gao and V.M. Rodrigues, Irreducibility of polynomials modulo p via Newton polytopes, preprint 2002.

[7] P.M. Gruber and J.M. Wills (Eds), Handbook of Convex Geometry, Volume B, Elsevier Science, 1993.

[8] A.G.B. Lauder and D. Wan, Counting points on varieties over finite fields of small characteristic, to appear in the proceedings of the 2000 MSRI workshop on Algorithmic Number Theory.

[9] L. Rónyai, L. Babai, M.K. Ganapathy, On the number of zero-patterns of a sequence of polynomials, *J. Amer. Math. Soc.* **14** no. 3, (2001), 717-735.

[10] B. Sturmfels, Polynomial equations and convex polytopes, Amer. Math. Monthly 105 No. 10, (1998), 907-922.