Progress in Mathematics, Vol. 143, © 1996 Birkhäuser Verlag Basel/Switzerland

Mixed monomial bases

P. Pedersen[1], B. Sturmfels[2]

1 Introduction

Given a system of n generic Laurent polynomials

$$f_i(x) \quad = \quad \sum_{q \in \mathcal{A}_i} c_{iq}\, x^q; \qquad q = (q_1, \ldots, q_n); \qquad x^q = x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n} \qquad (1.1)$$

with support sets $\mathcal{A}_i \subset \mathbb{Z}^n$, we consider the ring

$$A := K[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]/(f_1, \ldots, f_n),$$

where K is the field $\mathbb{Q}(\{c_{iq}\})$. The K-dimension of A equals the number of toric roots $\{\, x \in (\mathbb{C}^*)^n \,:\, f_i(x) = 0, \quad 1 \leq i \leq n \,\}$. By Bernstein's theorem [Ber], this number equals the mixed volume $\mathcal{MV}(P_1, \ldots, P_n)$ of the Newton polytopes $P_i := conv(\mathcal{A}_i)$. The objective of this note is to construct explicit K-bases for A, using the combinatorial technique of mixed subdivisions of the Minkowski sum $P := P_1 + \cdots + P_n$.

The mixed volume estimate for the number of toric roots of a system of polynomials is frequently much better than the Bezout bound, because it takes into account finer information about the combinatorics of the supporting set of monomials. It is easy to construct examples of families of polynomials of fixed support for which the ratio of the Bezout bound to the mixed volume bound is exponentially large (in n). In many applications, e.g., inverse kinematics, one would like sharp estimates for the number of solutions of a system of polynomial equations, and one knows a priori that the solutions are toric.

Following [GKZ1],[PS],[CE],[HS],[St1],[St2], we consider toric deformations of (1.1),

$$f_i(x,t) \quad := \quad \sum_{q \in \mathcal{A}_i} c_{iq}\, x^q\, t^{\omega_i(q)}, \qquad i = 1, \ldots, n. \qquad (1.2)$$

The weights $\omega_i(q)$ determine polytopes $\widehat{P}_i := \{\, (q, \omega_i(q)) \,:\, q \in P_i \}$ in $\mathbb{R}^{n+1}$. Let $\widehat{P} = \widehat{P}_1 + \cdots + \widehat{P}_n$ denote the Minkowski sum of the lifted polytopes. The *lower convex hull* $\widehat{\Delta}$ of $\widehat{P}$ is the collection of facets $\widehat{F}$ of $\widehat{P}$ whose inner normal

(1) Partially supported by NSF grant CCR-9258533 (NYI).

(2) Partially supported by a David and Lucile Packard Fellowship and NSF grants DMS-9201453 and DMS-9258547 (NYI).

has positive last component. Each such facet has the form $\widehat{F} = \widehat{F}_1 + \cdots + \widehat{F}_n$, where $\widehat{F}_i$ is a face of $\widehat{P}_i$. We say that $\widehat{F}$ is a *mixed facet* if $dim(\widehat{F}_i) = 1$ for $i = 1, \ldots, n$. Suppose $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ projects to the first n coordinates, then $\Delta := \{\pi(\widehat{F}) : \widehat{F} \in \widehat{\Delta}\}$ is a subdivision of P. Each cell C of Δ has the form $\pi(\widehat{F}) = F_1 + \cdots + F_n$ where $\pi(\widehat{F}_i) = F_i$ is a subpolytope of P_i. We assume that the weights ω are *sufficiently generic*, meaning that $\dim(C) = \dim(F_1) + \cdots + \dim(F_n)$, for every $C \in \Delta$. In this case Δ is called a *mixed subdivision* of P. The projection of a mixed facet of $\widehat{\Delta}$ is a *mixed cell* of Δ. The sum of the volumes of the mixed cells of any mixed subdivision Δ equals the *mixed volume* $\mathcal{MV}(P_1, \ldots, P_n)$. This is an integer which does not depend on the choice of subdivision. For details on mixed subdivisions, mixed volumes, and sparse polynomial systems, see [Bet],[HS],[St2] and the references given there.

Mixed cells are parallelotopes. If one considers them "half-open", then their volume equals their number of lattice points. In order to view the mixed cells half-open in a consistent manner, we follow the method of Canny and Emiris [CE] by displacing Δ to $\Delta + \delta$, with $\delta \in \mathbb{R}^n$ generic, and then counting all (now strictly interior) lattice points.

Theorem 1.1. *Let $f_1, \ldots, f_n$ be generic Laurent polynomials with Newton polytopes $P_1, \ldots, P_n$, respectively. The monomials corresponding to the lattice points lying in the mixed cells of any generically displaced mixed subdivision $\Delta + \delta$ of $P = P_1 + \cdots + P_n$ form a vector space basis for the quotient ring $A = K[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}] / (f_1, \ldots, f_n)$.*

We remark that an alternative proof of theorem 1.1 has been provided by Emiris and Rege [ER].

Consider a mixed cell $C = E_1 + \cdots + E_n$, where E_j is the one-dimensional subpolytope $[q_{j1},\, q_{j2}]$ with $q_{jk} \in \mathcal{A}_j$. Then C supports the *binomial system*

$$c_{11}x^{q_{11}} + c_{12}x^{q_{12}} \quad = \quad \cdots \quad = \quad c_{n1}x^{q_{n1}} + c_{n2}x^{q_{n2}} \quad = \quad 0. \tag{1.3}$$

We use the following notations consistently from here on:

$$b_C \quad := \quad (\,-c_{12}/c_{11}, \ldots, -c_{n2}/c_{n1}\,), \tag{1.4}$$

$$U_C \quad := \quad \text{the } (n \times n)\text{-matrix with column vectors} \quad q_{j1} - q_{j2}, \tag{1.5}$$

$$a_C \quad := \quad \sum \{\, q \,:\, q \in (C + \delta) \cap \mathbb{Z}^n \,\}. \tag{1.6}$$

We shall refer to the case when $b_C = \underline{1} := (1, 1, \ldots, 1)$ as the *unit* case.

There is a natural representation of the coordinate algebra A in $\mathrm{Hom}_K(A, A)$ which maps $f \mapsto$ (multiplication by f). This defines a trace $Tr(\,\cdot\,)$ on A, and a bilinear form $B : A \times A \to K$, $(g, h) \mapsto \mathrm{Tr}(g \cdot h)$. We call B the *trace form* and represent it by a symmetric matrix. The rank of B equals the number of distinct roots of (1.1). If the field K contains the reals $\mathbb{R}$, then the signature of B equals the number of distinct real roots (see [BW],[PRS]). Our second result concerns the asymptotic behavior of the trace form.

Theorem 1.2. *With respect to the basis arising from a displaced mixed subdivision as in theorem 1.1, the trace form of the deformed system (1.2) is a matrix polynomial in t,*

$$B(t) \quad = \quad B_0\, t^d\, (1 + o(1)),$$

where $d = \sum_C \gamma_C \cdot a_C$, $\det(B_0) = \prod_C \mathrm{Vol}(C)^{\mathrm{Vol}(C)} \cdot b_C^{2U_C^{-1} a_C}$. *and* $(\gamma_C, 1)$ *is the vector supporting the mixed facet above C. (The sum and product are over all mixed cells of Δ.)*

The expression $\mathrm{Vol}(C)^{\mathrm{Vol}(C)}$ plays a prominent role in the fundamental work of Gel'fand-Kapranov-Zelevinsky on $\mathcal{A}$-discriminants [GKZ1],[GKZ2]. (See Section 5 for further information regarding the relationship to their work.)

Throughout this paper we employ the following generalized multi-index notation: For a row vector $x = (x_j)$ and a matrix $U = (u_{ji})$, we let x^U denote the row vector whose i-th entry equals the monomial $x_1^{u_{1i}} \cdots x_n^{u_{ni}}$. This definition reduces to the usual multi-index notation when U is a column vector, and it reduces to a vector of powers of t when $x = t$ is a scalar and U is a row vector. One has the relation $(x^U)^V = x^{(UV)}$.

2 Binomial systems

Suppose $A = K[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]\,/\,I$ is a finite-dimensional reduced K-algebra of dimension d, and X the variety in $(\mathbb{C}^*)^n$ defined by an ideal I of generic polynomials as above. (We are using the fact that X is reduced.) Then a set of d monomials $\{\, x^q \,:\, q \in M \,\}$ is a basis of A if and only if the $d \times d$-matrix $S := (\xi^q)_{\xi \in X,\, q \in M}$ is non-singular. To establish our results we may also use the matrix $S^T S$, whose entries lie in the ground field K.

In this section we prove theorem 1.1 in the "local case" where I is generated by the binomial system (1.3). This system can be rewritten in the form

$$-(c_{11}/c_{12}) \cdot x^{q_{11} - q_{12}} \quad = \quad \cdots \quad = \quad -(c_{n1}/c_{n2}) \cdot x^{q_{n1} - q_{n2}} \quad = \quad 1, \qquad (2.1)$$

or more compactly $x^{U_C} = b_C$ according to the notations (1.4), (1.5). Here Δ consists of a single mixed cell C, and our basis-to-be is $M := (C + \delta) \cap \mathbb{Z}^n$.

Lemma 2.1. *Let Λ denote the subgroup of $\mathbb{Z}^n$ generated by the edges $v_1, \ldots, v_n$ of C. Then the points in M are coset representatives for the finite abelian group $G := \mathbb{Z}^n/\Lambda$.*

Proof: Let $C = \{\, v_0 + \sum \lambda_i v_i \;:\; 0 \le \lambda_i \le 1 \}$, where $v_0 \in \mathbb{Z}^n$. By the genericity of δ $C + \delta$ has no lattice points on its boundary. Given $q, q' \in M$, we have $q - q' = \sum \mu_i v_i$, where $0 \le \mu_i < 1$ are unique rational numbers. Therefore $q \not\equiv q' \pmod{\Lambda}$. On the other hand, any $x \in \mathbb{Z}^n$ may be expressed as $v_0 + \delta + \sum \mu_i v_i$, and this expression is normalized by extracting the integer parts of the μ_i to put it in the form $v' + v$, where $v' \in \Lambda$, and $v \in M$. $\square$

Lemma 2.2. *With the notations of lemma 2.1, the group algebra $\mathbb{C}[G]$ is isomorphic to the Laurent polynomial ring $\mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}] \,/\, (x^{v_1}-1, \dots, x^{v_n}-1)$.*

Proof: Define a homomorphism

$$\begin{array}{rclcl} \phi & : & \mathbb{Z}^n & \longrightarrow & \mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}] \,/\, (x^{v_1}-1, \dots, x^{v_n}-1) \\ (a_1, \dots, a_n) & & & \mapsto & x_1^{a_1} \cdots x_n^{a_n}. \end{array}$$

Then the kernel of ϕ has a $\mathbb{Z}$-basis $v_1, \dots, v_n$, and these generate Λ. □

Lemma 2.3. *In the unit case $b_C = \underline{1}$, the set X of roots of (2.1) may be identified with the group $\mathrm{Hom}(G, \mathbb{C}^*)$ of characters of $G = \mathbb{Z}^n/\Lambda$.*

Proof: We have the identifications

$$\begin{array}{rcl} \mathrm{Hom}_{\mathrm{grp.}}(G, \mathbb{C}^*) & \cong & \mathrm{Hom}_{\mathrm{alg.}}(\mathbb{C}[G], \mathbb{C}) \\ & \cong & \mathrm{spec}_m(\mathbb{C}[G]) \\ & \cong & \mathrm{spec}_m(\mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}] \,/\, (x^{v_1}-1, \dots, x^{v_n}-1)). \end{array}$$

The first line follows from general principles of character theory (see [CR], §10). The second line follows from the identification of maximal ideals with kernels of homomorphisms into fields. The last line follows by lemma 2.2. □

Proposition 2.4. *In the unit case, the matrix $S = (\xi^q)_{\xi \in X,\, q \in M}$ is non-singular, and $\det(S^T S) = \pm \mathrm{Vol}(C)^{\mathrm{Vol}(C)}$.*

Proof: By lemma 2.1 and lemma 2.3, the matrix S may be viewed as the matrix of characters $\xi \in X = \mathrm{Hom}(G, \mathbb{C}^*)$ evaluated at coset representatives $q \in M$ of $\mathbb{Z}^n$ modulo Λ. Applying the second orthogonality relation for group characters (which sums across characters; see [CR], §31) in the abelian case, the matrix $\overline{S}^T S$ is a diagonal matrix with entries $|G|$ on the diagonal. On the other hand, the matrix S differs from its complex conjugate $\overline{S}$ only by some permutation of the columns corresponding to the automorphism $q \mapsto -q$ of G. Therefore, $\det(S^T S) = \pm |G|^{|G|}$. On the other hand, $|G|$ equals the number of lattice points in $C + \delta$, which equals the volume of C. □

Corollary 2.5. *For all choices of non-zero coefficients in (2.1), the matrix $S = (\xi^q)_{\xi \in X,\, q \in M}$ is non-singular. With b_C, U_C, and a_C defined as in (1.4)–(1.6), we have*

$$\det(S^T S) \;=\; \pm \mathrm{Vol}(C)^{\mathrm{Vol}(C)} \cdot b_C^{2\, U_C^{-1} a_C}. \tag{2.2}$$

Proof: Suppress the subscript C. Let $\mathrm{diag}(y_1, \dots, y_n)$ denote the diagonal matrix with y_i's on the diagonal, and make the change of variables $x = \mathrm{diag}(y_1, \dots, y_n) \cdot b^{U^{-1}}$. Then

$$b \;=\; (\mathrm{diag}(y_1, \dots, y_n) \cdot b^{U^{-1}})^U \;=\; y^U \cdot b^{U^{-1} U} \;=\; y^U \cdot b,$$

or, equivalently, $y^U = \underline{1}$. Thus after the change of variables we are in the unit case. The general matrix entry ξ^q of S differs from the same entry for the unit case as follows:

$$(\mathrm{diag}(\xi_1, \ldots, \xi_n) \cdot b^{U^{-1}})^q \;=\; \xi^q \cdot b^{(U^{-1} q)}. \tag{2.3}$$

We can factor the constant $b^{(U^{-1} q)}$ out of the column labelled q. The aggregate factor is

$$b^{(U^{-1} \sum_{q \in M} q)} \;=\; b^{U^{-1} a}. \qquad \square$$

3 General sparse systems

The roots of (1.2) are algebraic functions $x(t)$ of the parameter t. Their Puiseux series for small t may be determined from the ansatz

$$x(t) \;=\; y \cdot t^{\gamma}\,(1 + o(1)) \;:=\; (y_1 t^{\gamma_1}, \ldots, y_n t^{\gamma_n})\,(1 + o(1)), \tag{3.1}$$

where y is a vector of complex variables and $\gamma = (\gamma_1, \ldots, \gamma_n)$ is a vector of rational numbers. Substituting (3.1) into the i-th equation of (1.2), one obtains terms

$$c_{iq}\, y^q\, t^{\langle\, \gamma,\, q\, \rangle + \omega_i(q)}\;(1 + o(1)). \tag{3.2}$$

The exponent of t in (3.2) may be expressed as $\langle\ (\gamma, 1),\ (q, \omega_i(q))\ \rangle$, that is to say, the value of the linear functional $(\gamma, 1)$ on the lifting of some point $q \in \mathcal{A}_i$. Lowest-order terms are determined by faces $\widehat{F}_{i,\gamma}$ of $\widehat{P}_i$ on which $(\gamma, 1)$ is minimized. Let $F_{i,\gamma} := \pi(\widehat{F}_{i,\gamma})$. We refer to each

$$f_{i,\gamma}(y) \;:=\; \sum_{q \in \pi(\pi^{-1}(\mathcal{A}_i) \cap \widehat{F}_{i,\gamma})} c_{iq}\, y^q, \quad i = 1, \ldots, n \tag{3.3}$$

as the *degeneration* of f_i with respect to the linear functional γ under the lifting ω_i. (We include only monomials which lift to the face $\widehat{F}_{i,\gamma}$.) The following result is proved in lemmas 3.1 and 3.2 of [HS].

Lemma 3.1. *The system $f_{1,\gamma}(y) = \cdots = f_{n,\gamma}(y) = 0$ has a solution $y \in (\mathbb{C}^*)^n$ if and only if $(\gamma, 1)$ supports a mixed facet $\widehat{F}$ of the lower envelope $\widehat{\Delta}$ of $\widehat{P}$. For γ fixed, the number of solutions y equals the volume of the mixed cell $C = \pi(\widehat{F})$.*

This lemma shows that all relevant degenerations (3.3) are binomial systems of the form shown in (1.3) and (2.1). Every pair $\gamma \in \mathbb{Q}^n$ and $y \in (\mathbb{C}^*)^n$ as in lemma 3.1 contributes a branch (3.1) to the vector-valued algebraic function $x(t)$ defined by (1.2). Taking the sum over all mixed facets in $\widehat{\Delta}$, one finds that the total number of branches equals the mixed volume $\mathcal{MV}(P_1, \ldots, P_n)$. This technique was used in [HS] to give a new proof and algorithm for Bernstein's theorem. We are now prepared to prove our main result.

Proof of theorem 1.1: Suppose $\{ C_j \,:\, j = 1,\ldots,r \}$ are the mixed cells of the subdivision Δ, and let $M_j := (C_j + \delta) \cap \mathbb{Z}^n$ and $M := M_1 \cup \cdots \cup M_r$. We need to show that the matrix $S := (\xi^q)_{\xi \in X,\, q \in M}$ is non-singular, where $X := \{ \xi \in (\mathbb{C}^*)^n \,:\, f_1(\xi) = \cdots = f_n(\xi) = 0 \}$. If the determinant of S were zero, then this identity would continue to hold for the branches $x(t)$ in (3.1). Writing Ω for the set of all branches, the corresponding matrix is $S(t) = (x(t)^q)_{x(t) \in \Omega,\, q \in M}$. We will prove $\det(S(t)) \neq 0$ by showing that the lowest-order term in t is non-zero. We do this by showing that to lowest order the matrix $S(t)$ is block diagonal, and each block is of the type considered in the previous section.

Let $(\gamma_i, 1)$ be the inner normal vector of a mixed facet $\widehat{F}_i$ lying over the mixed cell C_i. Let Ω_i denote the set of branches $x_i(t) = y_i\, t^{\gamma_i}\,(1 + o(1))$ arising from C_i as in lemma 3.1. The vector y_i is a solution of the binomial system determined by the edges of C_i. We consider the matrix $\underline{S}(t)$ obtained from $S(t)$ by multiplying the *column* labelled $q \in M_j$ by $t^{\omega_j(q)}$. This has the effect of multiplying $\det(S(t))$ by a factor t^d, where $d = \sum_{j=1}^{r} \sum_{q \in M_j} \omega_j(q)$. We have $\det(\underline{S}(t)) \;=$

$$
\begin{array}{c|cccc|}
 & q_1 \in M_1 & q_2 \in M_2 & \ldots & q_r \in M_r \\
x_1 \in \Omega_1 & (t^{w_1(q_1)} x_1(t)^{q_1}) & (t^{w_2(q_2)} x_1(t)^{q_2}) & \ldots & (t^{w_r(q_r)} x_1(t)^{q_r}) \\
x_2 \in \Omega_2 & (t^{w_1(q_1)} x_2(t)^{q_1}) & (t^{w_2(q_2)} x_2(t)^{q_2}) & \ldots & (t^{w_r(q_r)} x_2(t)^{q_r}) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_r \in \Omega_r & (t^{w_1(q_1)} x_r(t)^{q_1}) & (t^{w_2(q_2)} x_r(t)^{q_2}) & \ldots & (t^{w_r(q_r)} x_r(t)^{q_r})
\end{array}
\tag{3.4}
$$

Let S_{ij} denote the block corresponding to Ω_i and M_j. The typical entry of S_{ij} looks like

$$
t^{w_j(q_j)} x_i(t)^{q_j} \;=\; t^{w_j(q_j)} (y_i\, t^{\gamma_i}\,(1 + o(1)))^{q_j} \;=\; y_i^{q_j}\, t^{\langle\, (\gamma_i,\, 1),\;\, (q_j, \omega_j(q_j))\,\rangle}\,(1 + o(1)).
$$

The order in t of this expression is abbreviated

$$
a_{ij}(q) \quad := \quad \langle\; (\gamma_i,\, 1),\;\; (q,\, \omega_j(q))\; \rangle; \quad q \in M_j.
$$

We note that $a_{ii}(q)$ has a constant value a_{ii} for $q \in M_i$. Any mixed facet $\widehat{F}_j \neq \widehat{F}_i$ of the lower envelope $\widehat{\Delta}$ of $\widehat{P}$ lies above $\widehat{F}_i$ in the direction $(\gamma_i,\, 1)$. There can be no "ties" in the values of $(\gamma_i,\, 1)$ on vertices appearing on faces which $\widehat{F}_i$ shares with $\widehat{F}_j$, since all the vertices we consider are *strictly interior* to the mixed cells of the displaced subdivision. In more precise terms, we have $a_{ii} < a_{ij}(q),\ \forall q \in M_j,\ j \neq i$. Therefore the components of the diagonal block S_{ii} have strictly smaller order in t than any other entries in their rows. We may extract from the *rows* corresponding to Ω_i a factor $t^{a_{ii}}$. This leaves terms in the S_{ii} block of order $(1 + o(1))$, and terms in the off-diagonal blocks S_{ij} with positive exponents $a_{ij}(q) - a_{ii} > 0$. The complete power of t extracted from (3.4) is

$$\sum_{i=1}^{r} |M_i| \cdot a_{ii} \quad = \quad \sum_{i=1}^{r} \sum_{q \in M_i} \gamma_i \cdot q + \omega_i(q) \quad = \quad \sum_{i=1}^{r} \sum_{q \in M_i} \gamma_i \cdot q + \sum_{i=1}^{r} \sum_{q \in M_i} \omega_i(q).$$

The last term cancels the extra factor t^d we had introduced when passing from $S(t)$ to $\underline{S}(t)$. We therefore have, to lowest order in t,

$$\begin{aligned} \det(S(t)) &= \left\{ \prod_{i=1}^{r} \det(S_{ii}) \right\} t^{\sum_{i=1}^{r} \sum_{q \in M_i} \gamma_i \cdot q} + \cdots = \\ &= \left\{ \prod_{i=1}^{r} \det(S_{ii}) \right\} t^{\sum_C \gamma_C \cdot a_C} + \cdots . \end{aligned}$$

Here γ_C denotes the vector which selects the mixed cell C. By Proposition 2.5, each of the factors $\det(S_{ii})$ is non-zero. This concludes the proof of theorem 1.1. □

4 Trace forms and real roots

Let $B(\cdot, \cdot)$ denote the trace form (cf. [PRS],[BW]) of the finite-dimensional reduced K-algebra A, and let $X = Spec_m(A)$ as before. It is known that the trace of a polynomial is the sum of its values at all roots. Therefore the trace form can be written as

$$B(f, g) \quad = \quad Tr(f \cdot g) \quad = \quad \sum_{\xi \in X} f(\xi) g(\xi), \qquad \text{for } f, g \in A. \tag{4.1}$$

Proof of theorem 1.2: Let M be the basis of A which was established in theorem 1.1, and let $S = (\xi^q)_{\xi \in X, q \in M}$ as before. Given any element $f = \sum_{q \in M} \lambda_q x^q$ in A, with column vector of coefficients λ, then $S \cdot \lambda$ equals the column vector $(f(\xi) : \xi \in X)$. Therefore,

$$\mu^T \cdot (S^T S) \cdot \lambda \quad = \quad (S\mu)^T \cdot (S\lambda) \quad = \quad B(f, g), \quad \text{where} \quad g = \sum_{q \in M} \mu_q x^q.$$

In other words, the matrix $S^T S$ represents the trace form B with respect to the basis M.

By the results of the previous sections, the matrix polynomial $B(t)$ is asymptotically

$$B(t) = S(t)^T S(t) = \begin{bmatrix} S_{11}^T S_{11} & 0 & \cdots & 0 \\ 0 & S_{22}^T S_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_{rr}^T S_{rr} \end{bmatrix} \cdot t^{2 \cdot \sum_i \gamma_i \cdot a_i} + \cdots \tag{4.2}$$

From the second orthogonality relation of character theory we saw that, up to a constant factor, each block $S_{ii}^T S_{ii}$ is a permutation of the diagonal matrix $\mathrm{diag}(\mathrm{Vol}(C_i), \ldots, \mathrm{Vol}(C_i))$. By Proposition 2.5, the leading term of the matrix polynomial $B(t)$ has determinant $\pm \prod_C \mathrm{Vol}(C)^{\mathrm{Vol}(C)}\, b_C^{2U_C^{-1}a_C}$. □

For the remainder of this section we suppose that K contains $\mathbb{R}$, and that the coefficients of the input system (1.1) are real. An important theorem of computational algebra (cf. [Pe],[PRS],[BW]) states that the number of real roots $\# X(\mathbb{R})$ counted without multiplicity equals the signature of the trace form B.

We now come to the problem of determining the *asymptotic number of real roots* of the system (1.2), meaning the number of roots $x \in (\mathbb{R}^*)^n$ for any sufficiently small fixed real value of $t > 0$. For each mixed cell C of Δ, we consider the finite abelian group $G_C := \mathbb{Z}^n \,/\, \Lambda$, where Λ is the lattice generated by the edges of C. Consider the decomposition into invariant factors,

$$G_C \;:=\; \mathbb{Z}^n \,/\, \Lambda \;\cong\; \mathbb{Z}/n_1\mathbb{Z} \,\oplus\, \cdots \,\oplus\, \mathbb{Z}/n_k\mathbb{Z} \tag{4.3}$$

with $n_1 \,|\, n_2 \,|\, \cdots \,|\, n_k$. Let $p(C)$ denote the number of even invariant factors n_i of G_C. The following result was proved in [St1] for the special case of unmixed systems.

Proposition 4.1. *The asymptotic number of real roots of (1.2) is at most* $\sum_C 2^{p(C)}$.

Proof: To lowest order in t, the signature of $S^T S$ is the sum of the signatures of the blocks $S_{ii}^T S_{ii}$. Writing C for the i-th mixed cell and $S_C = S_{ii}$, we need to show that the signature of $S_C^T S_C$ is bounded above by $2^{p(C)}$.

We fix C and consider the invariant decomposition (4.3). Let W_j denote the character table of the cyclic group $\mathbb{Z}/n_j\mathbb{Z}$. It is easy to see that

$$W_j^T\, W_j \quad = \quad \begin{bmatrix} n_j & 0 & \ldots & 0 \\ 0 & 0 & \ldots & n_j \\ \vdots & \vdots & .\,\cdot\,\dot{} & \vdots \\ 0 & n_j & \ldots & 0 \end{bmatrix}, \tag{4.4}$$

because the automorphism $x \mapsto -x$ of $\mathbb{Z}/n_j\mathbb{Z}$ leaves 0 invariant, and cycles the remaining elements as shown. In the matrix (4.4) there are either one or two diagonal entries depending on whether n_i is odd or even. The signature of such a matrix equals the number of non-zero diagonal entries (cf. [Sh], p. 12). Hence the signature of $W_j^T W_j$ is 2 if n_j is even, and 1 if n_j is odd. Let W be the character table of G_C. This is the tensor product of the character tables W_j, and therefore

$$W^T W = (W_1 \otimes \cdots \otimes W_r)^T (W_1 \otimes \cdots \otimes W_r) = (W_1^T W_1) \otimes (W_2^T W_2) \otimes \cdots \otimes (W_r^T W_r).$$

Since the signature of real symmetric matrices is multiplicative with respect to tensor products, the signature of W^TW equals $2^{p(C)}$.

Now, the matrix S_C is obtained from W by multiplying each column by a scalar as in (2.3), and it can be written as $S_C = W'_1 \otimes \cdots \otimes W'_r$, where W'_i is obtained from W_i by multiplying each column by a non-zero scalar. Then $(W'_i)^T W'_i$ is obtained from $W_i^T W_i$ by multiplying each entry by a non-zero scalar, preserving symmetry of the matrix and non-negativity of the signature. The new symmetric matrix has signature either 0 or 2 if n_j is even, while it remains 1 if n_j is odd. Therefore the signature of

$$S_C^T S_C = (W'_1 \otimes \cdots \otimes W'_r)^T (W'_1 \otimes \cdots \otimes W'_r) = ((W'_1)^T W'_1) \otimes \cdots \otimes ((W'_r)^T W'_r)$$

is bounded above by the signature of W^TW, which is $2^{p(C)}$. $\square$

5 Two problems

We close with two remarks which suggest directions for future research.

(1) Our theorem 1.1 is non-trivial in the sense that, even for generic coefficients c_{iq}, not all collections of $\mathcal{MV}(P_1, \ldots, P_n)$ monomials form a basis for A. (For instance, the set $\{1, x^4y^2, x^3y^3, x^2y^4\}$ is a non-base modulo the system $\{a_0 + a_1x + a_2xy + a_3y,\ b_0 + b_1x^2y + b_2xy^2\}$.) It would be interesting to find a combinatorial characterization of the non-bases. This problem can be rephrased as follows: The data $\mathcal{A}_1, \ldots, \mathcal{A}_n \subset \mathbb{Z}^n$ define a matroid of rank $\mathcal{MV}(P_1, \ldots, P_n)$ on the (infinite) set $\mathbb{Z}^n$, by independence of monomials in A. The idea is to study this matroid. For instance, when is it uniform?

(2) Given any monomial basis for A, we can express the trace form B by a symmetric matrix, and its determinant $\det(B)$ is well-defined rational function of degree 0 in the coefficients c_{iq}. The most important divisor of the numerator of this rational function is the $\mathcal{A}$-*discriminant* $D_{\mathcal{A}}$ due to Gel'fand, Kapranov, and Zelevinsky [GKZ1],[GKZ2]. Here $\mathcal{A} = \cup_{i=1}^n e_i \times \mathcal{A}_i \subset \mathbb{Z}^{2n}$ is obtained by the *Cayley trick* from (1.1). It would be very interesting to understand all other factors, and to see how they depend on the choice of monomial basis. In light of our theorem 1.2, we conjecture that the leading coefficient of the $\mathcal{A}$-discriminant $D_{\mathcal{A}}$ with respect to ω is precisely $\pm \prod_C \mathrm{Vol}(C)^{\mathrm{Vol}(C)}$, where C runs over all mixed cells of the mixed subdivision Δ. This would provide a refinement of Thm 3D.2 in [GKZ1].

6 References

[BW] Becker, E., Wörmann, T.: "On the trace formula for quadratic forms", Recent Advances in Real Algebraic Geometry and Quadratic Forms, Proceedings of the RAGSQUAD Year, Berkeley 1990-1991, W.B. Jacob, T.-Y. Lam, R.O. Robson (editors), *Contemporary Mathematics*, **155** (1993) pp. 271–291.

[Ber] Bernstein, D. N.: "The number of roots of a system of equations", *Functional Analysis and its Applications* **9** (1975), pp. 1–4.

[Bet] Betke, U.: "Mixed volumes of polytopes": *Archiv d. Mathematik* **58** (1992), pp. 388–391.

[CE] Canny, J., Emiris, I.: "An Efficient Algorithm for the Sparse Mixed Resultant", in Proc. AAECC-10, edited by G. Cohen, T. Mora and O. Moreno", Springer Lecture Notes in Computer Science 263 (1993), pp. 89–104.

[CR] Curtis, C.W., I. Reiner, I.: "Representation Theory of Finite Groups and Associative Algebras", Wiley Interscience Publishers, 1962.

[ER] Emiris, I.Z., Rege, A.: "Monomial bases and polynomial system solving", submitted to ISSAC 1994.

[GKZ1] Gelfand, I. M., Kapranov, M. M., Zelevinsky, A. V.: "Discriminants of polynomials in several variables and triangulations of Newton polytopes", *Algebra i analiz (Leningrad Math. J.)* **2** (1990) pp. 1–62.

[GKZ2] Gelfand, I. M., Kapranov, M. M., Zelevinsky, A. V.: "Discriminants, Resultants and Multidimensional Determinants", Birkhäuser, Boston, 1994.

[HS] Huber, B., Sturmfels, B.: "A polyhedral method for solving sparse polynomial systems", *Mathematics of Computation*, to appear.

[Pe] Pedersen, P.: "Multivariate Sturm Theory", in Proc. AAECC-9, New Orleans 1991, Springer Lect. Notes in Comp. Sci. 537 (1991), pp. 318–332.

[PRS] Pedersen, P., Roy, M. F., Szpirglas, A.: "Counting real zeros in the multivariate case", in *"Computational Algebraic Geometry"* (eds. F. Eyssette, A. Galligo), Proceedings MEGA-92, Birkhäuser, 1992, pp. 203–223.

[PS] Pedersen, P., Sturmfels, B.: "Product formulas for resultants and Chow forms", *Mathematische Zeitschrift* **214** (1993) pp. 377–396.

[Sch] Scharlau, W.: "Quadratic and Hermitian Forms", Springer, Grundlehren der mathematischen Wissenschaften **270**, 1985.

[St1] Sturmfels, B.: "The asymptotic number of real zeros of a sparse polynomial system", to appear in Proceedings of the Workshop on "Hamiltonian and Gradient Flows: Algorithms and Control", (ed. A.Bloch), Fields Institute, Waterloo, Ontario, March 1992.

[St2] Sturmfels, B.: "On the Newton polytope of the resultant", *Journal of Algebraic Combinatorics* **3** (1994) pp. 207–236.

P. Pedersen (`paul@cs.cornell.edu`)
Department of Computer Science, Cornell University, Ithaca, NY 14853 (USA).

B. Sturmfels (`bernd@math.cornell.edu`)
Department of Mathematics, Cornell University, Ithaca NY 14853 (USA).