

# On the Curvature of the Central Path of Linear Programming Theory. §

Jean-Pierre Dedieu¶ Gregorio Malajovich|| Mike Shub\*\*

August 29, 2004

**We dedicate this paper with great admiration  
and affection to our friend and teacher  
Steve Smale on his seventy-fifth birthday.**

## Abstract

We prove a linear bound on the average total curvature of the central path of linear programming theory in terms on the number of independent variables of the primal problem, and independent of the number of constraints.

---

§Mathematics Subject Classification (MSC2000): 90C51 (Primary), 90C60, 14Q05, 68Q25, 65H10 (Secondary).

¶MIP. Département de Mathématique, Université Paul Sabatier, 31062 Toulouse cedex 04, France ([dedieu@mip.ups-tlse.fr](mailto:dedieu@mip.ups-tlse.fr)).

||Departamento de Matemática Aplicada, Universidade Federal de Rio de Janeiro, Caixa Postal 68530, CEP 21945-970, Rio de Janeiro, RJ, Brazil ([gregorio@ufrj.br](mailto:gregorio@ufrj.br)) Gregorio Malajovich was sponsored by CNPq and by FAPERJ (Brazil). His visits to Université Paul Sabatier were supported by the International Cooperation Agreement Brazil-France in Mathematics.

\*\*Department of Mathematical Sciences, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA, and: Department of Mathematics, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada ([shub@math.toronto.edu](mailto:shub@math.toronto.edu)).

# 1 Introduction.

Consider a linear programming problem in the following primal/dual form:

$$\begin{array}{ll} \min & \langle c, x \rangle \\ Ax - s = b & \\ s \geq 0 & \end{array} \quad \text{and} \quad \begin{array}{ll} \max & \langle b, y \rangle \\ A^T y = c & \\ y \geq 0 & \end{array} .$$

Here  $m > n \geq 1$  and  $A$  is an  $m \times n$  real matrix assumed to have rank  $n$ ,  $b \in \mathbb{R}^m$  and  $c \in \mathbb{R}^n$  are given vectors,  $b$  is not in the range of  $A$  and  $c$  is non-zero,  $y, s \in \mathbb{R}^m$  and  $x \in \mathbb{R}^n$  are unknown vectors ( $s$  is the vector of slack variables).

Our principal result bounds the total curvature of the union of all the central paths associated with all the feasible regions obtained by considering all the  $2^m$  possible sign conditions

$$s_i \epsilon_i 0, \quad i = 1, \dots, m,$$

where  $\epsilon_i$  is either  $\geq$  or  $\leq$ .

Formal definitions will be given in subsequent sections. The rest of the results in the introduction follow from the next theorem which requires the rest of the paper.

**Theorem 1.1.** *Let  $m > n \geq 1$ . Let  $A$  be an  $m \times n$  matrix of rank  $n$ , and let  $b \in \mathbb{R}^m$  and  $c \in \mathbb{R}^n$ ,  $b$  not in the range of  $A$  and  $c$  non-zero. The sum over all  $2^m$  sign conditions of the total curvature of the primal/dual central paths (resp. primal central paths, dual central paths) is less than or equal to  $2\pi n \binom{m-1}{n}$  (resp.  $2\pi(n-1) \binom{m-1}{n}$ ,  $2\pi n \binom{m-1}{n}$ .)*

Theorem 1.1 allows us to conclude various results on the average curvature of the central paths corresponding to various probability measures on the space of problems. We begin with our main motivating example.

Central paths are numerically followed to the optimal solution of linear programming problems by interior-point methods. For relevant background material on interior-point methods see Renegar [21]. Our point in studying the total curvature is that curves with small total curvature may be easy to approximate with straight lines. So, small total curvature may contribute to the understanding of why long-step interior-point methods are seen to be efficient in practice. In Dedieu-Shub [9] we studied the central paths of

linear programming problems defined on strictly feasible compact polyhedra (polytopes)<sup>1</sup> from a dynamical systems perspective. In this paper we optimistically conjectured that the worst case total curvature of a central path is  $O(n)$ . Our first average result and main theorem lend some credence to this conjecture, proving it on the average.

If we assume that the primal polyhedron  $\{x|Ax - b = s \geq 0\}$  is compact and strictly feasible (i.e., has non-empty interior), then the primal and dual problems have central paths which are each the projection of a primal/dual central path and all these central paths lead to optimal solutions. So for our purposes we will get a meaningful number if we divide the total curvature of the central paths of the all the strictly feasible polytopes arising from all possible sign conditions by the number of distinct strictly feasible polytopes associated with the  $2^m$  sign conditions:

$$Ax - s = b \quad s_i \in \{0, 1\} \quad i = 1, \dots, m,$$

where  $\epsilon_i$  is either  $\geq$  or  $\leq$ . The cardinality of the set of these polytopes is  $\leq \binom{m-1}{n}$  and equality holds for almost all  $(A, b)$ , see section 6. When equality holds we say  $(A, b)$  is in **general position**.

We use Theorem 1.1 to give an upper bound on the sum of the curvatures.

We obtain the following average result.

**Main Theorem.** *Let  $m > n \geq 1$ . Let  $A$  be an  $m \times n$  matrix of rank  $n$ , and let  $b \in \mathbb{R}^m$  and  $c \in \mathbb{R}^n$ ,  $c$  non-zero such that  $(A, b)$  is in general position. Then the average total curvature of the primal/dual central paths (resp. primal central paths, dual central paths) of the strictly feasible polytopes defined by  $(A, b)$  is less than or equal to  $2\pi n$  (resp.  $2\pi(n-1)$ ,  $2\pi n$ .)*

We may also average over more general probability measures on the data  $A, b, c$  defining the problem. First we more precisely define the space of problems  $P$  and measures  $\mu, \nu$  we consider.  $P = I \times \mathbb{R}^m \times \mathbb{R}^n$ . Here  $I$  is the open set of  $rank(n)$ ,  $m$  by  $n$  real matrices, and we assume for convenience that no row of any element of  $I$  is identically zero. Let  $D$  be the group with  $2^m$  elements consisting of those  $m$  by  $m$  diagonal matrices whose diagonal entries are all either 1 or  $-1$ . So for  $D \in D$ ,  $D$  acts on  $P$  by  $D((A, b, c)) = (DA, Db, c)$ . The set of problems defined by the orbit of  $(A, b, c)$  under the action of  $D$

---

<sup>1</sup>The feasible region for a linear programming problem is a polyhedron, a compact polyhedron is a polytope.

is the same as considering  $(A, b, c)$  with all possible sign conditions, so each orbit has  $2^m$  distinct elements. We say that a probability measure  $\mu$  is **sign invariant** if it is invariant under the action of  $D$ , i.e.  $D_*\mu = \mu$  for all  $D \in D$ .

We now generalize Theorem 1 once again averaging over problems with a strictly feasible primal polytope.

Let  $\mu$  be a sign-invariant probability measure on the data  $A, b, c$ . If the set of  $(A, b, c)$  in  $P$  such that  $(A, b)$  are in general position has full measure we will say that  $\mu$  is **full (for general position)**. This is the case for example if  $\mu$  is supported on a finite union of orbits of  $D$  through elements in general position or if  $\mu$  is absolutely continuous with respect to the Lebesgue measure, see section 6. For instance, an independent Gaussian probability distribution with zero mean and arbitrary variance for each coefficient of the data is sign invariant and full.

**Corollary 1.2.** *Let  $m > n$  and let  $\mu$  be a sign-invariant and full (for general position) probability measure on  $P$ . Let Feas be the set of data  $A, b, c$  with a strictly feasible primal polytope. Let  $\nu$  be the conditional probability measure (with respect to Feas) defined for any measurable  $V$  by*

$$\nu(V) = \frac{\mu(V \cap \text{Feas})}{\mu(\text{Feas})}.$$

*Then, the average (w.r.t.  $\nu$ ) total curvature of the primal/dual central path (resp. primal central path, dual central path) is less than or equal to  $2\pi n$  (resp.  $2\pi(n-1)$ ,  $2\pi n$ .)*

This corollary while almost immediate requires a little proof which we carry out in section 6. There is another version of Corollary 1.2 which is perhaps a little more natural from the point of view of regions which have central paths defined for all positive parameter values. We state it below but don't prove it as the proof is the same as for Corollary 1.2. For a primal/dual central path to exist for all positive parameter values a necessary and sufficient condition is that both primal and dual problems are strictly feasible: see [29],[32]. If this is the case we say that the primal/dual polyhedra are **jointly strictly feasible**. Every strictly feasible primal polytope gives rise to primal/dual jointly strictly feasible polyhedra, but there are more of the later generally among the polyhedra arising from the  $2^m$  possible sign conditions in a linear programming problem. Generally the number of jointly strictly feasible primal/dual polyhedra is  $\binom{m}{n}$ . We may see this simply since

there are generally  $\binom{m}{n}$  vertices to the primal polyhedra and at each vertex almost all non-zero  $c$  select a unique primal polyhedron for which that vertex minimizes the optimization problem, see[1]. When the number is  $\binom{m}{n}$  we say that  $(A, b, c)$  is in **joint general position**. If we consider a sign invariant probability measure which is full (for joint general position) i.e. the set of problems  $(A, b, c)$  which are in joint general position has full measure, we get a slight improvement of Corollary 1.2.

**Corollary 1.3.** *Let  $m > n$  and let  $\mu$  be a sign-invariant and full (for joint general position) probability measure on  $P$ . Let Feas be the set of data  $A, b, c$  with joint strictly feasible primal/dual polyhedra. Let  $\nu$  be the conditional probability measure (with respect to Feas) defined for any measurable  $V$  by*

$$\nu(V) = \frac{\mu(V \cap \text{Feas})}{\mu(\text{Feas})}.$$

*Then, the average (w.r.t.  $\nu$ ) total curvature of the primal/dual central path (resp. primal central path, dual central path) is less than or equal to  $2\pi n \frac{m-n}{m}$  (resp.  $2\pi(n-1) \frac{m-n}{m}$ ,  $2\pi n \frac{m-n}{m}$ .)*

## 2 Description of the central path.

When the optimal value is attained, the primal and dual problems have the same value and the optimality conditions may be written as

$$\begin{cases} Ax - s = b \\ A^T y = c \\ sy = 0 \\ y \geq 0, s \geq 0 \end{cases}$$

where  $sy$  denotes the componentwise product of these two vectors. The *primal/dual central path* of this problem is the curve  $(x(\mu), s(\mu), y(\mu))$ ,  $0 < \mu < \infty$ , given by

$$\begin{cases} Ax - s = b \\ A^T y = c \\ sy = \mu e \\ y > 0, s > 0 \end{cases} \quad (2.1)$$

where  $\mathbf{e}$  denotes the vector in  $\mathbb{R}^m$  of all 1's.

The *primal central path* is the curve  $(x(\mu), s(\mu))$ ,  $0 < \mu < \infty$ , defined as the curve of minimizers of the function  $-\mu \sum_1^m \ln(s_i) + c^T x$  restricted to the primal polyhedron. By the use of Lagrange multipliers one sees that this is the curve defined by the existence of a vector  $y(\mu)$  satisfying the equations (2.1). Thus the primal central path is the projection of the primal/dual central path into the  $(x, s)$  subspace.

Similarly, the *dual central path* is the curve  $y(\mu)$ ,  $0 < \mu < \infty$ , defined as the curve of maximizers of the function  $\mu \sum_1^m \ln(y_i) + b^T y$  restricted to the dual polyhedron. By use of Lagrange multipliers one sees that this curve is defined by the existence of vectors  $x(\mu), s(\mu)$  satisfying (2.1). So the dual central path is the projection of the primal/dual central path on the  $y$  subspace.

Note, as we have alluded to in the introduction, that when the primal polyhedron is compact and strictly feasible the primal central path is defined for all  $0 < \mu < \infty$  and then so are the primal/dual and dual central paths.

### 3 Curvature.

Let  $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^n$  be a  $C^2$  map with non-zero derivative:  $\dot{\mathbf{c}}(t) \neq 0$  for any  $t \in [a, b]$ . We denote by  $l$  the arc length:

$$l(t) = \int_a^t \|\dot{\mathbf{c}}(\tau)\| d\tau.$$

To the curve  $\mathbf{c}$  is associated another curve on the unit sphere, called the Gauss curve, defined by

$$t \in [a, b] \rightarrow \gamma(t) = \frac{\dot{\mathbf{c}}(t)}{\|\dot{\mathbf{c}}(t)\|} \in S^{n-1}$$

which may also be parameterized by the arc length  $l$  of  $\mathbf{c}$ :

$$l \in [0, L] \rightarrow \dot{\mathbf{c}}(l) \in S^{n-1}$$

with  $L$  the length of the curve  $\mathbf{c}$ . The curvature is

$$\kappa(l) = \frac{d}{dl} \dot{\mathbf{c}}(l)$$

see Spivak [28, chapter 1]. In terms of the original parameter we have

$$\kappa(t) = \frac{1}{\|\dot{\mathbf{c}}(t)\|} \frac{d}{dt} \left( \frac{\dot{\mathbf{c}}(t)}{\|\dot{\mathbf{c}}(t)\|} \right) = \frac{\ddot{\mathbf{c}}(t) \|\dot{\mathbf{c}}(t)\|^2 - \dot{\mathbf{c}}(t) \langle \dot{\mathbf{c}}(t), \ddot{\mathbf{c}}(t) \rangle}{\|\dot{\mathbf{c}}(t)\|^4}. \quad (3.2)$$

The total curvature  $K$  is the integral of the norm of the curvature vector:

$$K = \int_0^L \|\kappa(l)\| dl .$$

Thus,  $K$  is equal to the length of the Gauss curve on the unit sphere  $S^{n-1} \subset \mathbb{R}^n$ . To compute  $K$  we use integral geometry; the next section is devoted to that.

## 4 An integral geometry formula.

Let  $\gamma(t)$ ,  $a \leq t \leq b$ , be a  $C^1$  parametric curve contained in the unit sphere  $S^{n-1}$  with at most a countable number of singularities (i.e.  $\dot{\gamma}(t) = 0$ ). The parameter interval is not necessarily finite:  $-\infty \leq a \leq b \leq \infty$ . Let us denote by  $G_{n,n-1}$  the Grassmannian manifold of hyperplanes through the origin contained in  $\mathbb{R}^n$ . We also denote by  $dG(\mathcal{H})$  the unique probability measure on  $G_{n,n-1}$  invariant under the action of the orthogonal group.

**Theorem 4.1.** *The length of  $\gamma$  is equal to*

$$L(\gamma) = \int_a^b \left\| \frac{d}{dt} \gamma(t) \right\| dt = \pi \int_{\mathcal{H} \in G_{n,n-1}} \#(\mathcal{H} \cap \gamma) dG(\mathcal{H})$$

where  $\#(\mathcal{H} \cap \gamma)$  denotes the number of parameters  $a \leq t \leq b$  such that  $\gamma(t) \in \mathcal{H}$ :  $\#(\mathcal{H} \cap \gamma)$  is the number of intersections counted with multiplicity.

*Proof.* If  $\gamma$  is an embedding then Theorem 4.1 follows from Santaló [23, chapter 18, section 6], or also see Shub and Smale [25, section 4], where a similar theorem is proved for projective spaces or Edelman and Kostlan [11]. Now the set of  $t$  such that  $\frac{d}{dt} \gamma(t) \neq 0$  may be written as a countable union of intervals on each of which  $\gamma$  is an embedding.  $\square$

**Definition 4.2.** The parametric curve  $\gamma$  is transversal to  $\mathcal{H} \in G_{n,n-1}$  (we also say  $\mathcal{H}$  is transversal to  $\gamma$ ) when  $\dot{\gamma}(t) \notin \mathcal{H}$  at the intersection points.

**Corollary 4.3.** *If the number of intersections counted with multiplicity satisfies  $\#(\mathcal{H} \cap \gamma) \leq \mathcal{B}$  for all transversal  $\mathcal{H} \in G_{n,n-1}$  then*

$$L(\gamma) = \int_a^b \left\| \frac{d}{dt} \gamma(t) \right\| dt \leq \pi \mathcal{B}.$$

*Proof.* By a usual application of Sard's Theorem, see Golubitsky-Guillemin [12], non-transversality is a zero measure event. Thus, the integral giving  $L(\gamma)$  only needs to be evaluated on the set  $\mathcal{T}$  of  $\mathcal{H} \in G_{n,n-1}$  such that  $\mathcal{H}$  is transversal to  $\gamma$ . Since  $dG(\mathcal{H})$  is a probability measure we get

$$L(\gamma) = \pi \int_{\mathcal{H} \in \mathcal{T}} \#(\mathcal{H} \cap \gamma) dG(\mathcal{H}) \leq \pi \mathcal{B} \int_{\mathcal{H} \in \mathcal{T}} dG(\mathcal{H}) = \pi \mathcal{B}.$$

□

In order to bound the number of transversal intersections of the Gauss curve with a hyperplane  $\mathcal{H}$ , we will need the following fact: let

$$\begin{aligned} F : \mathbb{R} \times \mathbb{R}^r &\rightarrow \mathbb{R}^r \\ (\mu, z) &\mapsto F_\mu(z) = F(\mu, z) \end{aligned}$$

be of class  $\mathcal{C}^2$ , and assume that we are in the conditions of the Implicit Function Theorem, namely  $F_{\mu_0}(\mathbf{c}_0) \equiv 0$  and  $DF_{\mu_0}(\mathbf{c}_0)$  (the derivative of  $F$  with respect to the  $z$  variables) has full rank. Let  $\mathbf{c}(\mu) : [\mu_0 - \epsilon, \mu_0 + \epsilon] \rightarrow \mathbb{R}^r$  be the associated implicit function,  $\mathbf{c}(\mu_0) = \mathbf{c}_0$  and  $F_\mu(\mathbf{c}(\mu)) = 0$ , and let  $\dot{\mathbf{c}}(\mu)$  (resp.  $\dot{F}_\mu$ ) denote the derivative of  $\mathbf{c}$  (resp.  $F$ ) with respect to  $\mu$ .

Let  $\mathcal{H}$  denote a hyperplane, with normal vector  $h$ :

$$\mathcal{H} = \{z \in \mathbb{R}^r : \langle h, z \rangle = 0\}.$$

**Lemma 4.4.** *In the conditions above, if the Gauss curve  $\gamma(\mu) = \dot{\mathbf{c}}(\mu)/\|\dot{\mathbf{c}}(\mu)\|$  intersects  $\mathcal{H}$  transversally for  $\mu = \mu_0$ , then  $(\mathbf{c}(\mu_0), \dot{\mathbf{c}}(\mu_0), \mu_0)$  is a zero of the function*

$$\Phi(\mathbf{c}, \dot{\mathbf{c}}, \mu) = \begin{bmatrix} F_\mu(\mathbf{c}) \\ DF_\mu(\mathbf{c}) \dot{\mathbf{c}} + \dot{F}_\mu(\mathbf{c}) \\ \langle h, \dot{\mathbf{c}} \rangle \end{bmatrix} \quad (4.3)$$

Moreover,  $D\Phi$  has full rank at that point.

*Proof.* Equations (4.3-1) and (4.3-2) vanish because of the Implicit Function Theorem. Equation (4.3-3) is zero because of the intersection hypothesis. We write  $D\Phi(\mathbf{c}, \dot{\mathbf{c}}, \mu)$  as the block matrix:

$$D\Phi = \begin{bmatrix} DF_\mu(\mathbf{c}) & 0 & \dot{F}_\mu(\mathbf{c}) \\ D^2F_\mu(\mathbf{c}) \otimes \dot{\mathbf{c}} + D\dot{F}_\mu(\mathbf{c}) & DF_\mu(\mathbf{c}) & D\dot{F}_\mu(\mathbf{c}) \dot{\mathbf{c}} + \ddot{F}_\mu(\mathbf{c}) \\ 0 & h^T & 0 \end{bmatrix} \quad (4.4)$$

where  $D^2F_\mu(\mathbf{c}) \otimes \dot{\mathbf{c}}$  is the linear map  $y \mapsto D^2F_\mu(\mathbf{c})(\dot{\mathbf{c}}, y)$ . By hypothesis,  $DF_\mu(\mathbf{c})$  is invertible. Hence, the block LU factorization of the matrix in (4.4) is:

$$D\Phi(\mathbf{c}, \dot{\mathbf{c}}, \mu) = \begin{bmatrix} I & & & \\ L_{21} & I & & \\ 0 & h^T DF_\mu(\mathbf{c})^{-1} & 1 & \end{bmatrix} \begin{bmatrix} DF_\mu(\mathbf{c}) & 0 & \dot{F}_\mu(\mathbf{c}) \\ & DF_\mu(\mathbf{c}) & U_{23} \\ & & U_{33} \end{bmatrix}$$

where using (4.3):

$$\begin{aligned} L_{21} &= (D^2F_\mu(\mathbf{c}) \otimes \dot{\mathbf{c}}) DF_\mu(\mathbf{c})^{-1} + D\dot{F}_\mu(\mathbf{c}) DF_\mu(\mathbf{c})^{-1} \\ U_{23} &= 2D\dot{F}_\mu(\mathbf{c}) \dot{\mathbf{c}} + \ddot{F}_\mu(\mathbf{c}) + D^2F_\mu(\mathbf{c})(\dot{\mathbf{c}}, \dot{\mathbf{c}}) \\ U_{33} &= -h^T \left( 2DF_\mu(\mathbf{c})^{-1} D\dot{F}_\mu(\mathbf{c}) \dot{\mathbf{c}} + DF_\mu(\mathbf{c})^{-1} \ddot{F}_\mu(\mathbf{c}) + DF_\mu(\mathbf{c})^{-1} D^2F_\mu(\mathbf{c})(\dot{\mathbf{c}}, \dot{\mathbf{c}}) \right). \end{aligned}$$

Note that by construction,  $F_\mu(\mathbf{c}(\mu)) \equiv 0$ . Differentiating once with respect to  $\mu$ , we obtain (4.3-2). Differentiating once again,

$$DF_\mu(\mathbf{c}(\mu)) \ddot{\mathbf{c}} + D^2F_\mu(\mathbf{c}(\mu)) (\dot{\mathbf{c}}, \dot{\mathbf{c}}) + 2D\dot{F}_\mu(\mathbf{c}(\mu)) \dot{\mathbf{c}} + \ddot{F}_\mu(\mathbf{c}(\mu)) = 0$$

Solving for  $\ddot{\mathbf{c}}$  and substituting into  $U_{33}$ , we obtain:

$$U_{33} = h^T \ddot{\mathbf{c}}.$$

We need to show that  $U_{33} \neq 0$ . Our hypothesis was that  $\dot{\gamma}(t) \notin \mathcal{H}$ . Multiplying equation (3.2) by  $h^T$  on the left, we obtain:

$$h^T \dot{\gamma}(t) = \frac{h^T \ddot{\mathbf{c}}(t)}{\|\dot{\mathbf{c}}(t)\|^2}.$$

Hence,  $U_{33}$  does not vanish and  $D\Phi$  is non-singular at  $(\mathbf{c}_0, \dot{\mathbf{c}}_0, \mu_0)$ .  $\square$

## 5 A Bézout bound for multi-homogeneous systems.

According to Theorem 4.1 to estimate the length of a curve we have to count the number of points in a certain set. To give such an estimate we use the multi-homogeneous Bézout Theorem. While this theorem is well-known to algebraic geometers, topologists and homotopy method theorists, the computation of the Bézout number is usually only carried out in the bi-homogeneous case in textbooks. Morgan and Sommese [19] prove the theorem and give a simple description of how to compute the number, which we repeat here.

Let  $f = (f_i)_{1 \leq i \leq n}$  be a system of  $n$  complex polynomial equations in  $n+m$  complex variables. These variables are partitioned into  $m$  groups  $X_1, \dots, X_m$  with  $k_j + 1$  variables in the  $j$ -th group.  $f_i$  is said multi-homogeneous if for any index  $j$  there exists a degree  $d_{ij}$  such that, for any scalar  $\lambda \in \mathbb{C}$ ,

$$f_i(X_1, \dots, \lambda X_j, \dots, X_m) = \lambda^{d_{ij}} f_i(X_1, \dots, X_j, \dots, X_m).$$

In this case the system  $f$  is called multi-homogeneous. The Bézout number  $\mathcal{B}$  associated with this system and this structure is defined as the coefficient of  $\prod_{j=1}^m \zeta_j^{k_j}$  in the product  $\prod_{i=1}^n \sum_{j=1}^m d_{ij} \zeta_j$ .

We say that  $(X_1, \dots, X_m) \in \mathbb{C}^{n+m}$  is a zero for  $f$  when  $f(X_1, \dots, X_m) = 0$ . In that case,  $f(\lambda_1 X_1, \dots, \lambda_m X_m) = 0$  for any  $m$ -tuple of complex scalars  $(\lambda_1, \dots, \lambda_m)$ . For this reason it is convenient to associate a zero to a point in the product of projective spaces  $\mathbb{P}^{k_1}(\mathbb{C}) \times \dots \times \mathbb{P}^{k_m}(\mathbb{C})$ . We use the same notation for a point in  $\mathbb{P}^{k_1}(\mathbb{C}) \times \dots \times \mathbb{P}^{k_m}(\mathbb{C})$  and for any representative  $(X_1, \dots, X_m) \in \mathbb{C}^{n+m}$ .

We say that a zero  $(X_1, \dots, X_m) \in \mathbb{P}^{k_1}(\mathbb{C}) \times \dots \times \mathbb{P}^{k_m}(\mathbb{C})$  is non-singular when the derivative

$$Df(X_1, \dots, X_m) : \mathbb{C}^{n+m} \rightarrow \mathbb{C}^n$$

is surjective. Notice that this definition is independent of the representative  $(X_1, \dots, X_m) \in \mathbb{C}^{n+m}$ . We have

**Theorem 5.1.** (*Multi-homogeneous Bézout Theorem*) *Let  $f$  be a multi-homogeneous system. Then the number of isolated zeros of  $f$  in  $\mathbb{P}^{k_1}(\mathbb{C}) \times \dots \times \mathbb{P}^{k_m}(\mathbb{C})$  is less than or equal to  $\mathcal{B}$ . If all the zeros are non-singular then  $f$  has exactly  $\mathcal{B}$  zeros.*

## 6 The total curvature of the central path on the average.

To the matrix  $A$  and the vector  $b$  not in the range of  $A$ , we associate the set of admissible points of the primal problem via the set of equalities-inequalities

$$Ax - s = b, \quad s \geq 0.$$

We may also consider the other polyhedra contained in the subspace  $Ax - s = b$  and defined by the inequalities

$$s_i \leq \epsilon_i, \quad 1 \leq i \leq m,$$

where  $\epsilon = (\epsilon_1, \dots, \epsilon_m)$  is one of the  $2^m$  vectors of sign conditions.

Let  $\mathcal{F}(A, b)$  denote the set of such primal strictly feasible polyhedra contained in the subspace  $Ax - s = b$  and  $\mathcal{Q}(A, b)$  the set of those which are compact.

**Lemma 6.1.** *For almost all  $A$  and  $b$ ,*

$$\#\mathcal{Q}(A, b) = R_K(m, n) = \binom{m-1}{n}.$$

*Proof.* This statement was proved by Buck [8] for  $A$  and  $b$  in *general position*. In particular, since  $A$  and  $b$  are in general position except in a set of measure zero, Lemma 6.1 holds for almost all  $A$  and  $b$ .  $\square$

**Proposition 6.2.** *A probability measure on  $P$  which is absolutely continuous with respect to Lebesgue measure is full.*

*Proof.* The set of  $(A, b, c)$  in  $P$  where  $(A, b)$  is not in general position has zero Lebesgue measure by the above lemma and Fubini's theorem, thus it has zero measure for any measure absolutely continuous with respect to Lebesgue.  $\square$

Now we prove Corollary 1.2 of the introduction assuming the Main Theorem.

*Proof.* The group  $D$  acts freely on  $P$ , so let  $\frac{P}{\sim}$  denote the orbit space. Then we may decompose the measure  $\mu$  on the orbits of  $D$ . Since  $\mu$  is sign invariant each point in the orbit gets equal measure and the same is true for the conditional measure  $\nu$ , i.e., each strictly feasible polytope in the orbit of  $D$

gets equal measure when the measure  $\nu$  is decomposed on orbits. Now we average over the orbits of points in general position, apply the Main Theorem and then average over  $\frac{P}{\approx}$ .  $\square$

It remains to prove Theorem 1.1.

The proof of this theorem requires Lemmas 6.3, 6.4 and Proposition 6.5 below.

**Lemma 6.3.** *For each  $\mathcal{F} \in \mathcal{F}(A, b)$ , the Gauss curves associated with the central paths  $\mathbf{c}_{\mathbf{PD}}(\mathcal{F})$ ,  $\mathbf{c}_{\mathbf{P}}(\mathcal{F})$  and  $\mathbf{c}_{\mathbf{D}}(\mathcal{F})$  are well-defined.*

*Proof.* The primal/dual (resp. primal, resp. dual) central path associated with a polyhedron  $\mathcal{F} \in \mathcal{F}(A, b)$  satisfies the system of polynomial equations

$$F_\mu(x, s, y) = \begin{bmatrix} Ax - s - b \\ A^T y - c \\ sy - \mu e \end{bmatrix} = 0 \quad (6.5)$$

with  $\mu > 0$ , and this system is the same for all those polyhedra.

Let  $D_s$  denote the diagonal matrix with diagonal entries  $s_i$ . Since  $sy = \mu e$  (equation 6.5-3),  $D_s$  is invertible. The derivative of  $F_\mu$  is equal to

$$DF_\mu(x, s, y) = \begin{bmatrix} A & -I & 0 \\ 0 & 0 & A^T \\ 0 & D_y & D_s \end{bmatrix}$$

and it factors as:

$$\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} I & & & \\ -\mu D_s^{-1} & I & & \\ 0 & A^T D_s^{-1} & I & \end{bmatrix} \begin{bmatrix} -I & 0 & A \\ & D_s & \mu D_s^{-1} A \\ & & -\mu A^T D_s^{-2} A \end{bmatrix} \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \\ I & 0 & 0 \end{bmatrix}.$$

Therefore, since  $A$  has full column rank,  $\mu > 0$  and  $s_j \neq 0$ , this derivative is nonsingular and we are in the conditions of the Implicit Function Theorem. The speed vector

$$\dot{\mathbf{c}} = (\dot{x}, \dot{s}, \dot{y}) = -DF_\mu(x(\mu), y(\mu), z(\mu))^{-1} \dot{F}_\mu((x(\mu), y(\mu), z(\mu)))$$

is the unique solution of the implicit equations:

$$\begin{cases} A\dot{x} - \dot{s} = 0 \\ A^T \dot{y} = 0 \\ \dot{s}y + s\dot{y} = e. \end{cases} \quad (6.6)$$

The Gauss curve for the primal-dual problem is  $(\dot{x}, \dot{s}, \dot{y})/\|(\dot{x}, \dot{s}, \dot{y})\|$ . Notice that because of (6.6-3),  $\dot{s}$ , and  $\dot{y}$  cannot be together equal to 0 so that this curve is well-defined.

The Gauss curve associated to the primal (resp. dual) central path is  $(\dot{x}, \dot{s})/\|(\dot{x}, \dot{s})\|$  (resp.  $\dot{y}/\|\dot{y}\|$ ). Those curves are well defined, for suppose that  $\dot{s} = 0$ . Then equations (6.5-3) and (6.6-3) combined give:

$$sy = \mu s \dot{y}.$$

Hence, dividing componentwise by  $s$  and then multiplying by  $A^T$ , one obtains:

$$c = A^T y = \mu A^T \dot{y} = 0$$

which contradict the hypothesis  $c \neq 0$ . Suppose now  $\dot{y} = 0$ . Then, by the same reasoning one obtains  $s = \mu \dot{s}$ . Hence, by (6.6-1),  $s$  is in the range of  $A$ . Then by (6.5-1),  $b$  is in the range of  $A$ , contradiction. Thus, we showed that the Gauss curves for the primal-dual, primal and dual central paths are well-defined.  $\square$

A point of the curve  $\gamma_{\mathbf{PD}}$  is the image under the map

$$(x, s, y, \dot{x}, \dot{s}, \dot{y}) \mapsto \frac{(\dot{x}, \dot{s}, \dot{y})}{\|(\dot{x}, \dot{s}, \dot{y})\|}$$

of a point  $(x, s, y, \dot{x}, \dot{s}, \dot{y})$  satisfying the systems (6.5) and (6.6) for some  $\mu > 0$ . Similarly, a point of the curve  $\gamma_{\mathbf{P}}$  (resp.  $\gamma_{\mathbf{D}}$ ) is the image of such a point under the map

$$(x, s, y, \dot{x}, \dot{s}, \dot{y}) \mapsto \frac{(\dot{x}, \dot{s})}{\|(\dot{x}, \dot{s})\|}$$

$$\text{(resp. } (x, s, y, \dot{x}, \dot{s}, \dot{y}) \mapsto \frac{\dot{y}}{\|\dot{y}\|}\text{)}.$$

The symbol  $*$  stands for  $\mathbf{PD}$ ,  $\mathbf{P}$  or  $\mathbf{D}$ . These cases will be known as the *primal-dual*, the *primal* and the *dual* case, respectively.

**Lemma 6.4.** *It is assumed as above that  $\mathcal{F} \in \mathcal{F}(A, b)$  and that  $\mathbf{c}_*$  and  $\gamma_*$  are defined as above. Let  $u \in \mathbb{R}^n$ ,  $v \in \mathbb{R}^m$ ,  $w \in \mathbb{R}^m$  be not all zero.*

1. Each transversal intersection of the Gauss curve  $\gamma_{\mathbf{PD}}$  with the hyperplane

$$\mathcal{H}_{\mathbf{PD}} = \{(\dot{x}, \dot{s}, \dot{y}) : u^T \dot{x} + v^T \dot{s} + w^T \dot{y} = 0\}$$

is the image of a nonsingular solution of the polynomial system

$$\Phi^{A,b,c,u,v,w}(x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu) = \begin{bmatrix} Ax - s - b \\ A^T y - c \\ sy - \mu e \\ A\dot{x} - \dot{s} \\ A^T \dot{y} \\ \dot{s}y + s\dot{y} - e \\ u^T \dot{x} + v^T \dot{s} + w^T \dot{y} \end{bmatrix} = 0 \quad (6.7)$$

such that  $\mu > 0$ .

2. Let  $w = 0$ . Each transversal intersection of the Gauss curve  $\gamma_{\mathbf{P}}$  with the hyperplane

$$\mathcal{H}_{\mathbf{P}} = \{(\dot{x}, \dot{s}) : u^T \dot{x} + v^T \dot{s} = 0\}$$

is the image of a nonsingular solution of the polynomial system (6.7).

3. Let  $u = 0$  and  $v = 0$ . Each transversal intersection of the Gauss curve  $\gamma_{\mathbf{D}}$  with the hyperplane

$$\mathcal{H}_{\mathbf{D}} = \{\dot{y} : w^T \dot{y} = 0\}$$

is the image of a nonsingular solution of the polynomial system (6.7).

*Proof.* Part 1 is Lemma 4.4, where  $F_\mu$  and  $DF_\mu$  are computed in (6.5) and (6.6).

Part 2 follows from the fact that any transversal intersection of  $\gamma_{\mathbf{P}}$  with the hyperplane  $u^T \dot{x} + v^T \dot{s} = 0$  corresponds to a transversal intersection of  $\gamma_{\mathbf{PD}}$  with the hyperplane  $u^T \dot{x} + v^T \dot{s} + 0^T \dot{y} = 0$ . Indeed, if  $\gamma_{\mathbf{P}}(\mu) = \frac{(\dot{x}, \dot{s})}{\|(\dot{x}, \dot{s})\|}$  we set  $\gamma_{\mathbf{PD}}(\mu) = \frac{(\dot{x}, \dot{s}, \dot{y})}{\|(\dot{x}, \dot{s}, \dot{y})\|}$ . Then  $(u, v)^T \gamma_{\mathbf{P}}(\mu) = 0$  if and only if  $(u, v, 0)^T \gamma_{\mathbf{PD}}(\mu) = 0$ .

Now, assume that the intersection of  $\gamma_{\mathbf{P}}$  with  $u^T \dot{x} + v^T \dot{s} = 0$  is transversal. Then,

$$\begin{aligned}
\frac{\partial}{\partial \mu} (u, v, 0)^T \gamma_{\mathbf{PD}}(\mu) &= \frac{1}{\|\dot{x}, \dot{s}, \dot{y}\|} (u^T \ddot{x} + v^T \ddot{s} + 0^T \ddot{y}) + (u^T \dot{x} + v^T \dot{s} + 0^T \dot{y}) \frac{\partial}{\partial \mu} \frac{1}{\|\dot{x}, \dot{s}, \dot{y}\|} \\
&= \frac{1}{\|\dot{x}, \dot{s}, \dot{y}\|} (u^T \ddot{x} + v^T \ddot{s} + 0^T \ddot{y}) \\
&= \frac{\|\dot{x}, \dot{s}\|}{\|\dot{x}, \dot{s}, \dot{y}\|} \frac{\partial}{\partial \mu} (u, v)^T \gamma_{\mathbf{P}}(\mu) \neq 0
\end{aligned}$$

and therefore the intersection of  $\gamma_{\mathbf{PD}}$  with  $u^T \dot{x} + v^T \dot{s} + 0^T \dot{y} = 0$  is also transversal.

The proof of Part 3 is similar.  $\square$

**Proposition 6.5.** *Let  $m > n \geq 1$ . Let  $A$  be an  $m \times n$  matrix of rank  $n$ , and let  $b \in \mathbb{R}^m$  and  $c \in \mathbb{R}^n$ ,  $c$  non-zero. Then, for any transversal hyperplane  $\mathcal{H}_*$ , the polynomial system (6.7) has at most*

$$\mathcal{B}_{\mathbf{PD}} \leq 2n \binom{m-1}{n}$$

nonsingular solutions  $(x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$  with  $\mu > 0$ .

If furthermore we have  $w = 0$ , the number of nonsingular solutions is bounded above by

$$\mathcal{B}_{\mathbf{P}} \leq 2(n-1) \binom{m-1}{n}.$$

If instead we have  $u = 0$  and  $v = 0$ , the number of nonsingular solutions is still bounded above by

$$\mathcal{B}_{\mathbf{D}} \leq 2n \binom{m-1}{n}.$$

The proof of Proposition 6.5 is long, and is postponed to Section 7.

*Proof of Theorem 1.1.* The total curvature is the sum of the lengths of the Gauss curves corresponding to strictly feasible regions. According to Corollary 4.3, a bound  $\mathcal{B}_*$  for the number of intersections (counted with multiplicity) of the associated Gauss curves with a transversal hyperplane gives the bound  $\pi \mathcal{B}_*$  for the length. Finally, by Lemma 6.4 and Proposition 6.5  $\mathcal{B}_*$  may be taken as in Proposition 6.5.  $\square$

## 7 Proof of Proposition 6.5

The proof of Proposition 6.5 is quite long, and occupies all of this section. There are actually three cases, that are quite similar and will be treated in parallel.

We proceed as follows:

### 7.1 Complexification of the equations

The first step is to complexify the equations, i.e. to keep the coefficients fixed and to consider the variables as complex instead of real.

**Lemma 7.1.** *The number of nonsingular solutions of (6.7) in  $\mathbb{R}^{4m+2m+1}$  with  $\mu > 0$  is bounded above by the number of nonsingular solutions of (6.7) in  $\mathbb{C}^{4m+2m+1}$  with  $\mu \neq 0$ .*

*Proof.* A real root is, in particular, a complex root. It is nonsingular if and only if the determinant of the Jacobian matrix of the derivative does not vanish. The non-vanishing of this determinant does not depend on whether the matrix is considered as real or complex.  $\square$

Note that when we complexify the equations, the terms  $u^T \dot{x} + v^T \dot{s} + w^T \dot{y}$  stand for the usual transpose.

A standard application of Bézout's Theorem implies that:

**Lemma 7.2.** *The number of nonsingular solutions of (6.7) in  $\mathbb{C}^{4m+2m+1}$  with  $\mu \neq 0$  is bounded above by  $2^{2m}$ .*

This estimate, while ensuring finiteness, is not sharp enough for our theorem.

### 7.2 Continuation of nonsingular roots

More formally, we denote by  $\mathcal{A}_{\mathbf{PD}}$  the set of all complex  $A, b, c, u, v, w$  where  $A$  has rank  $n$ ,  $c \neq 0$  and  $u, v, w$  are not simultaneously zero. We also denote by  $\mathcal{A}_{\mathbf{P}}$  (resp.  $\mathcal{A}_{\mathbf{D}}$ ) the intersection of  $\mathcal{A}_{\mathbf{PD}}$  with the linear space  $w = 0$  (resp.  $u = 0$  and  $v = 0$ ).

Then,  $\mathcal{B}_*$  will denote the maximal number of nonsingular complex roots of (6.7) with  $\mu \neq 0$ , where  $*$  is one of  $\mathbf{PD}, \mathbf{P}, \mathbf{D}$  and the maximum is taken over all parameters in  $\mathcal{A}_*$ . As in Remark 7.2,  $\mathcal{B}_*$  is finite. Hence this maximal

number is attained, and at that point all the nonsingular complex roots may be continued in a certain neighborhood. Thus,

**Lemma 7.3.** *The maximal number  $\mathcal{B}_*$  of nonsingular complex roots is attained in a certain open set of  $\mathcal{A}_*$ .*

*Proof.* As in Lemma 7.2,  $\mathcal{B}_*$  is attained for some parameter  $(A, b, c, u, v, w)$ .

By the Implicit Function Theorem, the nonsingular complex roots of (6.7) with  $\mu \neq 0$  can be continued to nonsingular complex roots with  $\mu \neq 0$ , in a certain neighborhood of the parameter  $(A, b, c, u, v, w)$ .  $\square$

### 7.3 Non-degeneracy at the maximum

The following fact will be needed in the sequel:

**Proposition 7.4.** *The complex roots of (6.7) with  $\mu \neq 0$  are all nonsingular, almost everywhere in  $\mathcal{A}_*$ .*

*Proof.* Let  $\mathcal{X} = \{x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu \in \mathbb{C}^{4m+2n+1} : \mu \neq 0\}$ . We consider the evaluation function

$$\begin{aligned} \text{ev} : \mathcal{A}_* \times \mathcal{X} &\rightarrow \mathbb{C}^{4m+2n+1} \\ (A, b, c, u, v, w; x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu) &\mapsto \Phi^{A,b,c,u,v,w}(x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu) \end{aligned}$$

where  $\Phi$  was defined in (6.7).

0 is a regular value of  $\text{ev}$  if and only if  $\text{Dev}(A, b, c, \dots, \mu)$  is onto when  $\text{ev}(A, b, c, \dots, \mu) = 0$  (see [12, Ch II S1]).

**Lemma 7.5.** *0 is a regular value for  $\text{ev}$ .*

This lemma guarantees that  $V = \text{ev}^{-1}(0)$  is a smooth manifold and  $\dim V = \dim \mathcal{A}_*$ .

Now we consider the natural projection  $\pi_1 : V \rightarrow \mathcal{A}_*$ . By Sard's theorem, the regular values of  $\pi_1$  have full measure in  $\mathcal{A}_*$ . Since  $\dim V = \dim \mathcal{A}_*$ ,  $(A, b, \dots, w)$  is a regular value if and only if  $D\pi_1$  is an isomorphism at every point  $(A, b, \dots, \mu)$  such that  $\pi_1(A, b, \dots, \mu) = (A, b, \dots, w)$ . For such systems, all the roots with  $\mu \neq 0$  are nonsingular.  $\square$

*Proof of Lemma 7.5.* We first reorder the equations and the variables of (6.7) as follows:



## 7.4 Genericity

In this section we show that it is sufficient to bound the number of nonsingular roots of systems satisfying conditions 1 through 4 of Proposition 7.6.

Let  $K \subset \{1, \dots, m\}$ . We define  $S_K$  as the linear space of all  $\{s \in \mathbb{C}^m : s_k = 0 \forall k \in K\}$ .

**Proposition 7.6.** *Let  $m > n$ . There is a point  $(A, b, c, u, v, w) \in \mathcal{A}_*$  such that:*

1. *The maximal number  $\mathcal{B}_*$  of nonsingular complex solutions of (6.7) with  $\mu \neq 0$  is attained at this point.*
2. *All the solutions at that point are nonsingular.*
3. *For any  $K \subset \{1, \dots, m\}$ , the linear space  $S_K$  and the affine space  $(\ker A^H)^\perp - b$  intersect if and only if  $n - \#K \geq 0$ . In that case, the intersection has dimension  $n - \#K$ .*
4. *For any  $K \subset \{1, \dots, m\}$ , the linear space  $S_{\{1, \dots, m\} \setminus K}$  and the affine space  $\{y : A^T y = c\}$  intersect if and only if  $\#K - n \geq 0$ . In that case, the intersection has dimension  $\#K - n$ .*

*Proof.* By Lemma 7.3, item 1 holds on an open set  $U \subset \mathcal{A}^*$ . Item 2 will fail only on zero measure set (Proposition 7.4). For items 3 and 4, notice that with probability one,  $\dim(\ker A^H)^\perp - b = n$  and  $\dim\{y : A^T y = c\} = m - n$ . On the other hand,  $\text{codim } S_K = \#K$  and  $\text{codim } S_{\{1, \dots, m\} \setminus K} = m - \#K$ . Thus it is easy to see that points where item 3 or 4 fails lie in a finite union of zero measure sets.

Hence, items 2 to 4 will hold on a subset of  $\mathcal{A}^*$  of full measure which has a non-empty intersection with the open set of Lemma 7.3.  $\square$

This result has the following consequence: to give a bound for the number of nonsingular solutions of the system (6.7) with  $\mu \neq 0$ , we can replace the initial data by the data  $(A, b, c, u, v, w)$  of Proposition 7.6.

Also, for convenience, we will count the number of *isolated* roots of the corresponding system, which is the same.

## 7.5 Simplification of the equations

**Lemma 7.7.** *Set  $\hat{u} = u + A^T v$ . The polynomial systems (6.7) and (7.9) below have the same solutions with  $\mu \neq 0$  so that, the isolated solutions of (6.7) with  $\mu \neq 0$  are identical to the isolated solutions of (7.9) with  $\mu \neq 0$ .*

$$\Psi_\mu(x, s, y, \dot{x}, \dot{s}, \dot{y}) = \begin{bmatrix} Ax - s - b \\ A^T y - c \\ sy - \mu e \\ A\dot{x} - \dot{s} \\ A^T \dot{y} \\ \mu \dot{y} + y^2 \dot{s} - y \\ \hat{u}^T \dot{x} + w^T \dot{y} \end{bmatrix} = 0. \quad (7.9)$$

*Proof.* This last system is obtained from (6.7) by the transformation

$$\begin{cases} \Psi_k = \Phi_k, & 1 \leq k \leq 5, \\ \Psi_6 = y\Phi_6 - \dot{y}\Phi_3, \\ \Psi_7 = \Phi_7 + v^T \Phi_4. \end{cases} \quad (7.10)$$

When  $\mu \neq 0$ , then no component of  $y$  is zero so the solutions of (6.7) and the solutions of (7.9) coincide.  $\square$

Let  $g_1, \dots, g_{m-n}$  and  $f \in \mathbb{C}^m$ ,  $\gamma, \delta$  and  $\hat{w} \in \mathbb{C}^{m-n}$  be such that

- (a)  $(g_1, \dots, g_{m-n})$  is a basis of  $\ker A^T$ ,
- (b)  $A^T f = c$ ,
- (c)  $y = f + \sum_{j=1}^{m-n} \gamma_j g_j$ ,
- (d)  $\dot{y} = \sum_{j=1}^{m-n} \delta_j g_j$ ,
- (e)  $\hat{w}_j = w^T g_j, \quad j = 1 \dots m - n$ .

We will denote by  $g_{ij}$  the  $i$ -th coordinate of  $g_j$ , by  $G$  the  $m \times (m - n)$  matrix with entries  $g_{ij}$ , by  $A_i$  the  $i$ -th row of the matrix  $A$ ,  $i = 1, \dots, m$ , and by  $A^\dagger = (A^H A)^{-1} A^H$  the Moore-Penrose inverse of  $A$  (injective);  $AA^\dagger$  is the orthogonal projection onto  $\text{range } A$  while  $A^\dagger A = I_n$ .

**Lemma 7.8.** *The system 7.9 has the same solutions as*

$$\Omega_\mu(x, s, y, \dot{x}, \dot{s}, \dot{y}) = \begin{bmatrix} x - A^\dagger(s + b) \\ G^T(b + s) \\ A^T y - c \\ sy - \mu e \\ A\dot{x} - \dot{s} \\ A^T \dot{y} \\ \mu \dot{y} + y^2 \dot{s} - y \\ \hat{u}^T \dot{x} + w^T \dot{y} \end{bmatrix} = 0 \quad (7.11)$$

*Proof.* The equations 7.11-2 are equivalent to  $b + s \in (\ker A^T)^\perp$  that is  $b + s \in \text{range } A$ . Under this assumption,  $x = A^\dagger(s + b)$  gives  $Ax = b + s$ . Thus 7.11-1 and 7.11-2 imply 7.9-1. Conversely, if  $Ax = b + s$  we get  $b + s \in \text{range } A$  and  $x = A^\dagger(s + b)$  that is 7.11-1 and 7.11-2.  $\square$

## 7.6 Elimination of variables

Let us now introduce a new polynomial system with the same number of zeros:

$$\Xi_\mu(s, \dot{x}, \gamma, \delta) = \begin{bmatrix} G^T(b + s) \\ s(f + G\gamma) - \mu \\ \mu G\delta + (f + G\gamma)^2 \\ A\dot{x} - (f + G\gamma) \\ \hat{u}^T \dot{x} + \hat{w}^T \delta \end{bmatrix} = 0, \quad (7.12)$$

We also define the following maps:

$$\Lambda : \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^{m-n} \times \mathbb{C}^{m-n} \times \mathbb{C} \rightarrow \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C}^m \times \mathbb{C}$$

$$\Lambda(s, \dot{x}, \gamma, \delta, \mu) = (A^\dagger(s + b), s, f + G\gamma, \dot{x}, A\dot{x}, G\delta, \mu),$$

and the projection

$$\Pi_{2478} : \mathbb{C}^n \times \mathbb{C}^{m-n} \times \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C} \rightarrow \mathbb{C}^{m-n} \times \mathbb{C}^m \times \mathbb{C}^m \times \mathbb{C}$$

$$\Pi_{2478}(z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8) = (z_2, z_4, z_7, z_8).$$

**Lemma 7.9.** *The construction of the system (7.12) is such that the diagram*

$$\begin{array}{ccc}
 & \Omega & \\
 \mathbb{C}^{4m+2n+1} & \longrightarrow & \mathbb{C}^{4m+2n+1} \\
 \Lambda \uparrow & & \downarrow \Pi_{2478} \\
 \mathbb{C}^{3m-n+1} & \longrightarrow & \mathbb{C}^{3m-n+1} \\
 & \Xi & 
 \end{array}$$

is commutative. Moreover

1. *If  $(s, \dot{x}, \gamma, \delta, \mu)$  is a solution of (7.12) then  $(x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu) = \Lambda(s, \dot{x}, \gamma, \delta, \mu)$  is a solution of (7.11).*
2. *Any solution  $(x, s, y, \dot{x}, \dot{s}, \dot{y}, \mu)$  of (7.11) is equal to  $\Lambda(s, \dot{x}, \gamma, \delta, \mu)$  for a unique solution of (7.12).*
3. *Any isolated solution of (7.11) with  $\mu \neq 0$  corresponds to an isolated solution of (7.12) with  $\mu \neq 0$ .*

*Proof.* The proof is easy and left to the reader. □

Now we look at equations 7.12-3, 7.12-4 as a linear system of  $m + 1$  equations in the  $m$  unknowns  $(\delta, \dot{x})$ , with coefficients depending on  $\gamma$  and  $\mu$ . When  $(s, \dot{x}, \gamma, \delta, \mu)$  is a solution of (7.12) with  $\mu \neq 0$ , those equations have a solution if and only if the determinant of the corresponding augmented matrix vanishes. We can write the augmented matrix as:

$$M(\gamma, \mu) = \begin{bmatrix}
 \mu g_{11} & \cdots & \mu g_{1,m-n} & a_{11}Q_1(\gamma) & \cdots & a_{1n}Q_1(\gamma) & f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j} \\
 \vdots & & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & & \vdots & \vdots & & \vdots & \vdots \\
 \mu g_{m1} & \cdots & \mu g_{m,m-n} & a_{m1}Q_m(\gamma) & \cdots & a_{mn}Q_m(\gamma) & f_m + \sum_{j=1}^{m-n} \gamma_j g_{mj} \\
 \hat{w}_1 & \cdots & \hat{w}_{m-n} & \hat{u}_1 & \cdots & \hat{u}_n & 0
 \end{bmatrix} \tag{7.13}$$

where  $Q_i(\gamma) = \left( f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij} \right)^2$ . By dividing the first  $m - n$  columns by  $\mu$  and then multiplying the last row by  $\mu$ , we obtain a new matrix  $M'$  whose determinant vanishes if and only if the determinant of  $M$  vanishes:

$$M'(\gamma, \mu) = \begin{bmatrix} g_{11} & \cdots & g_{1,m-n} & a_{11}Q_1(\gamma) & \cdots & a_{1n}Q_1(\gamma) & f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ g_{m1} & \cdots & g_{m,m-n} & a_{m1}Q_m(\gamma) & \cdots & a_{mn}Q_m(\gamma) & f_m + \sum_{j=1}^{m-n} \gamma_j g_{mj} \\ \hat{w}_1 & \cdots & \hat{w}_{m-n} & \mu \hat{u}_1 & \cdots & \mu \hat{u}_n & 0 \end{bmatrix} \quad (7.14)$$

Its determinant is the same as the determinant of:

$$M''(\gamma, \mu) = \begin{bmatrix} g_{11} & \cdots & g_{1,m-n} & a_{11}Q_1(\gamma) & \cdots & a_{1n}Q_1(\gamma) & f_1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ g_{m1} & \cdots & g_{m,m-n} & a_{m1}Q_m(\gamma) & \cdots & a_{mn}Q_m(\gamma) & f_m \\ \hat{w}_1 & \cdots & \hat{w}_{m-n} & \mu \hat{u}_1 & \cdots & \mu \hat{u}_n & -\sum \gamma_i \hat{w}_i \end{bmatrix} \quad (7.15)$$

We now have to distinguish the three cases of Proposition 6.5. In the primal-dual case, we define the eliminating polynomial  $h_{\mathbf{PD}}(\gamma, \mu) = \det M''(\gamma, \mu)$ . In the dual case, we also define  $h_{\mathbf{D}}(\gamma) = \det M''(\gamma, \mu)$  but now, since  $\hat{u} = 0$ , the eliminating polynomial is independent of  $\mu$ . In the primal case,  $\hat{w} = 0$  and we notice that the last row of  $M''$  is divisible by  $\mu$ . Hence, we set  $h_{\mathbf{P}}(\gamma) = \mu^{-1} \det M''(\gamma, \mu)$ .

**Lemma 7.10.** *With the notations above,  $(s, \dot{x}, \gamma, \delta, \mu)$  is a solution of (7.12) with  $\mu \neq 0$  if and only if*

$$\Upsilon(s, \dot{x}, \gamma, \delta, \mu) = \begin{bmatrix} s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) - \mu \\ M(\gamma, \mu) \begin{bmatrix} \delta \\ \dot{x} \\ -1 \end{bmatrix} \\ G^T(b + s) \\ s_i(f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij}) - s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \\ h_*(\gamma, \mu) \end{bmatrix} = 0, \quad (7.16)$$

and  $\mu \neq 0$ .

*Proof.* The system  $\Upsilon = 0$  is the same as  $\Xi = 0$  plus the equation  $h(\gamma, \mu) = 0$  which is a consequence of  $\Xi = 0$  as has been explained previously.  $\square$

**Lemma 7.11.** *The number of isolated solutions of the system (7.16) with  $\mu \neq 0$  is less than or equal to the number of isolated solutions with  $s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \neq 0$  of*

$$\Theta(s, \gamma) = \begin{bmatrix} G^T(b + s) \\ s_i(f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij}) - s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \\ h_*(\gamma, f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \end{bmatrix} = 0 \quad (7.17)$$

where the range for  $i$  in the second equation is  $2 \leq i \leq m$ .

*Proof.* This lemma is a consequence of the following facts:

- An isolated solution  $(s, \dot{x}, \gamma, \delta, \mu)$  of (7.16) gives an isolated solution  $(s, \gamma)$  of (7.17),
- Two distinct solutions  $(s, \dot{x}, \gamma, \delta, \mu)$  and  $(s', \dot{x}', \gamma', \delta', \mu')$  of (7.16) with  $\mu \neq 0$  and  $\mu' \neq 0$  give two distinct solutions  $(s, \gamma)$  and  $(s', \gamma')$  of (7.17) with  $s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \neq 0$  and  $s'_1(f_1 + \sum_{j=1}^{m-n} \gamma'_j g_{1j}) \neq 0$ .

The first fact is true because (7.17) is a sub-system of (7.16).

Let us prove the second assertion. Let  $(s, \dot{x}, \gamma, \delta, \mu)$  and  $(s, \dot{x}', \gamma, \delta', \mu')$  be two solutions of (7.16) with  $s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \neq 0$ . Our aim is to prove that  $(\dot{x}, \delta, \mu) = (\dot{x}', \delta', \mu')$ . We have clearly  $\mu = \mu' \neq 0$  and  $s_i(f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij}) \neq 0$  for each  $i$ . Moreover  $(\dot{x}, \delta)$  is given by the system

$$M(\gamma, \mu) \begin{bmatrix} \delta \\ \dot{x} \\ -1 \end{bmatrix} = \begin{bmatrix} \mu G & \text{Diag}(Q_i(\gamma))A & f + G\gamma \\ \hat{w} & \hat{u} & 0 \end{bmatrix} \begin{bmatrix} \delta \\ \dot{x} \\ -1 \end{bmatrix} = 0.$$

This system has a unique solution if and only if  $\text{Rank } M(\gamma, \mu) = m$ . Let us prove that the  $m \times m$  submatrix

$$\begin{bmatrix} \mu G & \text{Diag}(Q_i(\gamma))A \end{bmatrix}$$

is nonsingular. Let  $\alpha \in \mathbb{C}^{m-n}$  and  $\beta \in \mathbb{C}^n$  be given. If

$$\begin{bmatrix} \mu G & \text{Diag}(Q_i(\gamma))A \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 0$$

then

$$\mu G\alpha + \text{Diag} (Q_i(\gamma))A\beta = 0$$

so that, multiplying by  $A^T$ , we get

$$A^T \text{Diag} (Q_i(\gamma))A\beta = 0.$$

Since  $(f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij})^2 > 0$  and  $\text{Rank } A = n$  the matrix  $A^T \text{Diag} (Q_i(\gamma))A$  is positive definite and consequently  $\beta = 0$ . Thus  $\mu G\alpha = 0$  so that  $\alpha = 0$  because  $\mu \neq 0$  and  $\text{Rank } G = m - n$ . This completes the proof.  $\square$

## 7.7 The Bézout bound

To count the number of isolated roots of this last system we use the multi-homogeneous Bézout bound given in Theorem 5.1. For this purpose we bi-homogenize the system (7.17): we introduce the homogenizing variables  $s_0$  in the first group of variables and  $\gamma_0$  in the second group. We obtain the system:

$$\Theta(s_0, s, \gamma_0, \gamma) = \begin{bmatrix} G^T(s_0 b + s) \\ s_i(\gamma_0 f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij}) - s_1(\gamma_0 f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \\ h_*(\gamma_0, \gamma, s_0, s) = 0 \end{bmatrix} = 0. \quad (7.18)$$

Here  $h_*(\gamma_0, \gamma, s_0, s)$  is just the homogenization of  $h_*(\gamma, s)$ .

### Lemma 7.12.

1. The system 7.18 is bi-homogeneous in the groups of variables  $(s_0, s) \in \mathbb{C}^{m+1}$  and  $(\gamma_0, \gamma) \in \mathbb{C}^{m-n+1}$ ,
2. There are  $m - n$  equations of multi-degree  $(1, 0)$  and  $m - 1$  equations of multi-degree  $(1, 1)$ . The multi-degree of the last equation depends on whether  $*$  stands for **PD**, **P** or **D**:

The multi-degree of  $h_{\mathbf{PD}}$  is  $(1, 2n + 1)$ .

The multi-degree of  $h_{\mathbf{P}}$  is  $(0, 2n - 2)$ .

The multi-degree of  $h_{\mathbf{D}}$  is  $(0, 2n + 1)$ .

3.  $(s, \gamma)$  is an isolated solution of (7.17) if and only if  $(1, s, 1, \gamma)$  is an isolated solution of (7.18),

4. The number of isolated solutions of (7.18) is bounded as follows:

When  $*$  is equal to **PD**, by  $2n \binom{m-1}{n} + \binom{m}{n}$ ,

when  $*$  is equal to **P**, by  $(2n-2) \binom{m-1}{n}$ , and

when  $*$  is equal to **D**, by  $(2n+1) \binom{m-1}{n}$ .

*Proof.* Assertion 1 holds by construction.

For the second assertion, we write down the Laplace expansion of the determinant of  $M''$  in equation (7.15) in terms of the last row:

$$\det M''(\gamma, \mu) = \sum_{j=1}^{m-n} (-1)^{m+j-1} \hat{w}_j C_j + \sum_{i=1}^n (-1)^{n+i-1} \mu \hat{u}_i C'_i + \left( \sum_{j=1}^{m-n} \gamma_j \hat{w}_j \right) C''$$

The cofactors  $C_j$ ,  $C'_i$  and  $C''$  have multi-degree  $(0, 2n)$ ,  $(0, 2n-2)$  and  $(0, 2n)$ , respectively. They get multiplied to factors of multi-degree  $(0, 0)$ ,  $(1, 1)$  and  $(0, 1)$ . Hence  $\det M''$  has multi-degree at most  $(1, 2n+1)$ .

The cofactors  $C_j$  and  $C''$  are irrelevant to the primal case, since  $\hat{w} = 0$ . In this case,

$$\det M''(\gamma, \mu) = \mu \sum_{i=1}^n (-1)^{n+i-1} \hat{u}_i C'_i$$

Moreover,  $h_{\mathbf{P}}$  was defined as  $\mu^{-1} \det M''$  and hence the multi-degree of  $h_{\mathbf{P}}$  is  $(0, 2n-2)$ .

Similarly, the cofactors  $C'_i$  are irrelevant to the dual case, and hence  $\deg h_{\mathbf{D}} = (0, 2n+1)$ .

The third assertion comes from a classical fact: when both product spaces are equipped with the usual topology, the canonical injection

$$i : \mathbb{C}^m \times \mathbb{C}^{m-n} \rightarrow \mathbb{P}^m(\mathbb{C}) \times \mathbb{P}^{m-n}(\mathbb{C})$$

is open and continuous.

By the Multi-homogeneous Bézout Theorem (Theorem 5.1), the number of isolated solutions in the primal-dual case ( $*$  = **PD**) is no more than the coefficient of  $\zeta_1^m \zeta_2^{m-n}$  in the expression

$$\zeta_1^{m-n} (\zeta_1 + \zeta_2)^{m-1} (\zeta_1 + (2n+1) \zeta_2).$$

In the primal-dual case (item 6), this coefficient is precisely

$$(2n + 1) \binom{m-1}{n} + \binom{m-1}{n-1} = 2n \binom{m-1}{n} + \binom{m}{n}.$$

The primal and dual cases are similar.  $\square$

## 7.8 The spurious roots

Except for the primal case, the number of roots we have obtained is still too big to give the bound announced in Theorem 1.1. We have to eliminate some of them to obtain the right result.

There are three classes of bi-homogeneous solutions  $(s_0, s, \gamma_0, \gamma)$  of the system (7.18).

1. **Roots at infinity** are roots for which  $\gamma_0 = 0$  or  $s_0 = 0$ . We will not worry about roots at infinity here.
2. **Spurious roots.** These are the finite roots for which  $\mu = 0$  that is  $s_1 = 0$  or  $\gamma_0 f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j} = 0$ .
3. **Legitimate roots** are all the other solutions.

Notice that solutions  $(s, \gamma, \mu)$  of the system (7.17) with  $s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) \neq 0$  correspond always to legitimate roots of (7.18).

**Lemma 7.13.** *The number of spurious roots of (6.7) for \* one of PD or D is  $\binom{m}{n}$ .*

*Proof.* We will only deal with the primal-dual case, the dual case being exactly the same.

The idea of the proof is to produce a bijection from the spurious roots to the class of subsets  $K \subset \{1, \dots, m\}$  of cardinality  $n$ .

Since  $\mu = s_1(f_1 + \sum_{j=1}^{m-n} \gamma_j g_{1j}) = 0$ , spurious roots are the zeros of the system

$$\begin{cases} G^T s = -s_0 G^T b, \\ s_i \left( \gamma_0 f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij} \right) = 0, \quad i = 1, \dots, m, \\ h_{\mathbf{PD}}(\gamma_0, \gamma, s_0, s) = 0 \end{cases} \quad (7.19)$$

Let  $(\gamma_0, \gamma, s_0, s)$  be a spurious root. Since spurious roots are finite we assume that  $\gamma_0 = 1$  and  $s_0 = 1$ . We set  $K = K(s) = \{k \in \{1, \dots, m\} : s_k = 0\}$ . Then, the system (7.19) breaks into:

$$\begin{cases} G^T(s + b) = 0, \text{ with } s \in S_K \\ f_i + \sum_{j=1}^{m-n} \gamma_j g_{ij} = 0, \text{ } i \notin K \\ h_{\mathbf{PD}}(\gamma, s) = 0 \end{cases} \quad (7.20)$$

The first equation is equivalent to say that  $s \in S_K$  and  $s + b \in (\ker A^H)^\perp$ . Again, this is the same as saying that  $s \in S_K \cap ((\ker A^H)^\perp - b)$ .

If the cardinality of  $K$  is strictly larger than  $n$ , then Proposition 7.6 item 3 guarantees that the intersection of  $S_K$  with the affine space  $(\ker A^H)^\perp + b$  is empty. Therefore the cardinality of  $K$  is at most  $n$ .

The second equation of (7.20) is satisfied if and only if the preimage of  $c$  by  $A^T$  and  $S_{\{1, \dots, m\} \setminus K}$  have a non-empty intersection. In that case,  $\gamma_i$  is associated to the  $i$ -th non-zero coordinate of a point  $y$  at the intersection.

If the cardinality of  $K$  is strictly smaller than  $n$ , then the intersection is empty because of Proposition 7.6 item 4.

Therefore  $K$  has cardinality  $n$ . Conversely, let  $K$  be a cardinality  $n$  subset of  $\{1, \dots, m\}$ . Then because of Proposition 7.6 item 3,  $S_K$  and  $(\ker A^H)^\perp - b$  do intersect. Because of item 4, the spaces  $(A^T)^{-1}(c)$  and  $S_{\{1, \dots, m\} \setminus K}$  also have a non-empty intersection.

Therefore, equations (7.20–1, 2) admit a finite common solution  $(1, \gamma, 1, s)$ . Here, it is crucial to use the fact that we are in the primal-dual or dual case, and hence  $h_*$  is the determinant of  $M''$  in (7.15) or equivalently the determinant of  $M'$  in (7.14). It will follow that  $h_*(\gamma_0, \gamma, s_0, s) = 0$ . This happens because for  $i \notin K$ ,  $Q_i$  will vanish in (7.13). Also for  $i \notin K$ ,  $M'(\gamma, \mu)_{i, m+1} = 0$ .

Since  $\mu$  was replaced by  $s_1(f_1 + \sum \gamma_j g_{1j})$  in the last row of (7.14), the matrix  $M'(\gamma, \mu)$  has (after reordering) an  $(m - n + 1) \times (n + 1)$  block of zeros. This means that an  $n + 1$ -dimensional subspace is mapped into a  $m + 1 - (m - n + 1) = n$ -dimensional subspace. Hence  $h_*(\gamma, s) = \det M'(\gamma, \mu) = 0$ .  $\square$

Notice that spurious roots must be isolated. Otherwise, there would be a component of non-isolated roots of (6.7) for  $\mu \neq 0$ , and this would contradict Proposition 7.6 item 2. Hence, when  $*$  is one of **PD** or **D**, we can subtract the number of spurious roots from the bounds obtained in Lemma 7.12, and Proposition 6.5 follows. Notice that the bound for the dual case is not sharp.

## 8 Concluding remarks.

1. Beling and Verma [5] is a predecessor to our paper. They prove a similar result to our Proposition 6.5 but only for subspaces defined by the vanishing of one of the coordinates and their estimate is not as strong.

2. We have estimated the curvature by the number of complex roots of a system of equations including possibly roots at infinity. In fact only real and finite roots count. The number of real roots is in general much less and can in some contexts be compared with the square root of the number of complex roots, see Shub and Smale [24], Edelman and Kostlan [11], McLennan [17], Rojas [22] and Malajovich and Rojas [15, 16]. Thus the total curvature at least on average may be very small indeed for large problems. We find a better understanding of the total curvature of the central path in worst and average case analysis an interesting problem.

3. There is a body of literature on the curvature of the central path, relating the curvature to the complexity of Newton type algorithms that approximate the central path and produce approximations to the solutions: see Sonnevend, Stoer and Zhao [26, 27], Stoer and Zhao [31], Zhao [33, 34]. These papers use a different notion of curvature, closer to  $1/\gamma$  where  $\gamma$  is Smale's  $\gamma$ , see also Dedieu and Smale [10]. The integral of these quantities is infinite.

4. The Riemannian geometry of the central path has been studied by quite a few authors, see Karmarkar [13], Bayer and Lagarias [2, 3, 4], Nesterov and Todd [20].

5. Vavasis and Ye [30] study regions where the tangent vectors to the central path stay in definite cones. Curvature estimates may be used as a refinement of this information.

6. Malajovich and Meer [14] showed that the problem of computing (or even approximating up to a fixed constant) the sharpest multi-homogeneous Bézout bound for a system of polynomial equations is actually NP-hard.

### Acknowledgements

We would like to thank the two anonymous referees and the communicating editor, Mike Todd, for their advice and for suggesting that we state some of our results in terms of sign invariant measures.

## References

- [1] I. ADLER , *The expected number of pivots needed to solve parametric linear programs and the efficiency of the self-dual simplex method*, Department of Industrial Engineering and Operations Research, University of California at Berkeley, June, 1983.
- [2] D. A. BAYER and J. C. LAGARIAS, *The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories*, Trans. Amer. Math. Soc. **314** (1989), no. 2, 499–526.
- [3] ———, *The nonlinear geometry of linear programming. II. Legendre transform coordinates and central trajectories*, Trans. Amer. Math. Soc. **314** (1989), no. 2, 527–581.
- [4] ———, *Karmarkar’s linear programming algorithm and Newton’s method*, Math. Programming **50** (1991), no. 3, (Ser. A), 291–330.
- [5] PETER A. BELING and SUSHIL VERMA, *A probabilistic analysis of a measure of combinatorial complexity for the central curve*, Math. Program. **87** (2000), no. 1, Ser. A, 177–187.
- [6] D. N. BERNSTEIN, *The number of roots of a system of equations*, Funkcional. Anal. i Priložen. **9** (1975), no. 3, 1–4.
- [7] L. BLUM, F. CUCKER, M. SHUB AND S. SMALE, *Complexity and real computation*. Springer-Verlag, New York, 1998. xvi+453 pp.
- [8] R. C. BUCK, *Partition of space*, Amer. Math. Monthly, **50** (1943), 541–544.
- [9] JEAN-PIERRE DEDIEU and MICHAEL SHUB, *Newton vector fields and interior-point methods in linear programming*, To appear.
- [10] JEAN-PIERRE DEDIEU and STEVE SMALE, *Some lower bounds for the complexity of continuation methods*, J. Complexity **14** (1998), no. 4, 454–465.
- [11] ALAN EDELMAN and ERIC KOSTLAN, *How many zeros of a random polynomial are real?*, Bull. Amer. Math. Soc. (N.S.) **32** (1995), no. 1, 1–37.

- [12] M. GOLUBITSKY and V. GUILLEMIN, *Stable Mappings and Their Singularities*, Springer-Verlag, New York (1973).
- [13] NARENDRA KARMARKAR, *Riemannian geometry underlying interior-point methods for linear programming*. In: J. C. Lagarias and M. Todd, editors, *Mathematical Developments Arising from linear programming*. Contemporary Mathematics, Amer. Math. Soc. **144** (1990), 51–75.
- [14] GREGORIO MALAJOVICH and KLAUS MEER, *Computing Multi-Homogeneous Bézout Numbers is Hard*. Preprint, <http://www.arxiv.org/abs/cs/0405021>, May 2004.
- [15] GREGORIO MALAJOVICH and J. MAURICE ROJAS, *Polynomial systems and the momentum map*, Foundations of Computational Mathematics, Proceedings of the Smalefest 2000 (F. CUCKER and M. ROJAS, eds.), World Scientific, Singapore, 2002.
- [16] ———, *High probability analysis of the condition number of sparse polynomial systems*, Theoretical Computer Science **315** No 2-3 Pages 525–555, May 2004.
- [17] ANDREW MCLENNAN, *The expected number of real roots of a multihomogeneous system of polynomial equations*, Amer. J. Math. **124** (2002), no. 1, 49–73.
- [18] NIMROD MEGIDDO and MICHAEL SHUB, *Boundary behavior of interior-point algorithms in linear programming*, Math. Oper. Res. **14** (1989), no. 1, 97–146.
- [19] ALEXANDER MORGAN and ANDREW SOMMESE, *A homotopy for solving general polynomial systems that respects  $m$ -homogeneous structures*, Appl. Math. Comput. **24** (1987), no. 2, 101–113.
- [20] YU. E. NESTEROV and M. J. TODD, *On the Riemannian geometry defined by self-concordant barriers and interior-point methods*, Found. Comput. Math. **2** (2002), no. 4, 333–361.
- [21] JAMES RENEGAR, *A mathematical view of interior-point methods in convex optimization*, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

- [22] J. MAURICE ROJAS, *On the average number of real roots of certain random sparse polynomial systems*, The mathematics of numerical analysis (Park City, UT, 1995), Lectures in Appl. Math., vol. 32, Amer. Math. Soc., Providence, RI, 1996, pp. 689–699.
- [23] LUIS A. SANTALÓ, *Integral geometry and geometric probability*, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976, With a foreword by Mark Kac, Encyclopedia of Mathematics and its Applications, Vol. 1.
- [24] M. SHUB and S. SMALE, *Complexity of Bezout's theorem. II. Volumes and probabilities*, Computational algebraic geometry (Nice, 1992), Progr. Math., vol. 109, Birkhäuser Boston, Boston, MA, 1993, pp. 267–285.
- [25] ———, *Complexity of Bezout's theorem. V. Polynomial time*, Theoret. Comput. Sci. **133** (1994), no. 1, 141–164.
- [26] G. SONNEVEND, J. STOER, and G. ZHAO, *On the complexity of following the central path of linear programs by linear extrapolation*, XIV Symposium on Operations Research (Ulm, 1989), Methods Oper. Res., vol. 62, Hain, Frankfurt am Main, 1990, pp. 19–31.
- [27] ———, *On the complexity of following the central path of linear programs by linear extrapolation. II*, Math. Programming **52** (1991), no. 3, Ser. B, 527–553 (1992).
- [28] MICHAEL SPIVAK, *A comprehensive introduction to differential geometry. Vol. II*, second ed., Publish or Perish Inc., Wilmington, Del., 1979.
- [29] M.J. TODD, *Detecting Infeasibility in Infeasible-Interior-Point Methods for Optimization*, Foundations of Computational Mathematics, Minneapolis 2002, London Mathematical Society Lecture Note Series 312, Cambridge University Press, 2004 pp. 157–192.
- [30] STEPHEN A. VAVASIS and YINYU YE, *A primal-dual interior-point method whose running time depends only on the constraint matrix*, Math. Programming **74** (1996), no. 1, Ser. A, 79–120.
- [31] G. ZHAO and J. STOER, *Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals*, Appl. Math. Optim. **27** (1993), no. 1, 85–103.

- [32] S.J. WRIGHT, *Primal-dual Interior-Point Methods*, SIAM, Philadelphia.
- [33] GONGYUN ZHAO, *On the relationship between the curvature integral and the complexity of path-following methods in linear programming*, SIAM J. Optim. **6** (1996), no. 1, 57–73.
- [34] ———, *Interior point algorithms for linear complementarity problems based on large neighborhoods of the central path*, SIAM J. Optim. **8** (1998), no. 2, 397–413 (electronic).