RC22772 (W0304-100) April 10, 2003
Mathematics

IBM Research Report

On the curvature of the central path of linear programming theory

Jean-Pierre Dedieu
MIP. D´epartement de Mathematique,
Universite Paul Sabatier, 31062
Toulouse cedex 04, France

Michael Shub
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

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

Jean-Pierre Dedieu* Mike Shub†

DRAFT of April 8, 2003

1 Introduction

In this paper we study the curvature of the central path of linear programming theory. Our ultimate goal is to show that the total curvature of the path is polynomial in m and n the numbers of linear constraints and variables respectively. In fact we think that the total curvature may be $O(n)$ independent of any other data in the problem, a family of examples where it is $\Omega(n)$ may be found in Meggido and Shub [8]. Here we do not prove anything so strong, but we establish that for $m > n$ and bounded polytopes the total curvature is less than or equal to

$$\frac{\pi m^2(m+1)}{(m-n+1)(m-n)}$$

at least in a generic average sense which we describe below. 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.

2 Description of the central path.

Let $\mathcal{P}$ be a compact polytope in $\mathbb{R}^n$ defined by m affine inequalities

$$A_i x \geq b_i, \quad 1 \leq i \leq m.$$

*MIP. Département de Mathématique, Université Paul Sabatier, 31062 Toulouse cedex 04, France (dedieu@mip.ups-tlse.fr).

†Department of Mathematical Sciences, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA (mshub@us.ibm.com).

Here $A_i x$ denotes the matrix product of the row vector $A_i = (a_{i1}, \ldots, a_{in})$ by the column vector $x = (x_1, \ldots, x_n)^T$, A is the $m \times n$ matrix with rows A_i and we assume Rank $A = n$. Given $c \in \mathbb{R}^n$, we consider the linear programming problem (LP)

$$\min_{\substack{A_i x \geq b_i \\ 1 \leq i \leq m}} \langle c, x \rangle .$$

Let us denote by

$$f(x) = \sum_{i=1}^{m} \ln(A_i x - b_i)$$

($\ln(s) = -\infty$ when $s \leq 0$) the logarithmic barrier function associated with the description $Ax \geq b$ of $\mathcal{P}$. The barrier technique considers the family of nonlinear convex optimization problems $(LP(t))$

$$\min_{x \in \mathbb{R}^n} \langle c, x \rangle - t f(x)$$

with $t > 0$. The objective function

$$f_t(x) = \langle c, x \rangle - t f(x)$$

is strictly convex, smooth, and satisfies

$$\lim_{\substack{x \to \partial \mathcal{P} \\ x \in \text{Int } \mathcal{P}}} f_t(x) = \infty.$$

Thus, there exists a unique optimal solution $\mathbf{c}(t)$ to $(LP(t))$ for any $t > 0$. This curve is called the central path of our problem. Let us denote D_x the $m \times m$ diagonal matrix $D_x = \text{Diag}(A_i x - b_i)$. This matrix is nonsingular for any $x \in \text{Int } \mathcal{P}$. We also let $e = (1, \ldots, 1)^T \in \mathbb{R}^m$,

$$g(x) = \text{grad } f(x) = \sum_{i=1}^{m} \frac{A_i^T}{A_i x - b_i} = A^T D_x^{-1} e$$

and

$$h(x) = \text{hess } f(x) = -A^T D_x^{-2} A$$

so that

$$Df(x)u = \langle u, g(x) \rangle$$

and

$$D^2 f(x)(u, v) = \langle u, h(x)v \rangle$$

for any $u, v \in \mathbb{R}^n$. Since f_t is smooth and strictly convex the central path is given by the equation grad $f_t(\mathbf{c}(t)) = 0$ i.e.

$$g(\mathbf{c}(t)) = \frac{c}{t}, \; t > 0.$$

Lemma 2.1 *$g : Int\ \mathcal{P} \to \mathbb{R}^n$ is real analytic and invertible. Its inverse is also real analytic.*

Proof. For any $c \in \mathbb{R}^n$ the optimization problem

$$\min_{x\in\mathbb{R}^n} \langle c, x\rangle - f(x)$$

has a unique solution in Int $\mathcal{P}$ because the objective function is smooth, strictly convex and $\mathcal{P}$ is compact. Thus $g(x) = c$ has a unique solution that is g is bijective. We also notice that, for any x, $Dg(x)$ is nonsingular. Thus g^{-1} is real analytic by the inverse function theorem. ■

When c varies in $\mathbb{R}^n$ we obtain a family of curves. Our aim in this paper is to investigate the curvature of these curves.

3 Curvature.

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

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

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

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

which may also be parameterized by the arc length s:

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

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

$$\kappa(s) = \frac{d}{ds}\dot{\mathbf{c}}(s)$$

see Spivak [17] chap. 1. In terms of the original parameter we have

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

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

$$K = \int_0^L \|\kappa(s)\| ds$$

that is K is equal to the length of the Gauss curve on the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. 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 into the unit sphere $\mathbb{S}^{n-1}$. The parameter interval is not necessarily finite: $-\infty \leq a \leq b \leq \infty$. Let us denote by $\mathbb{G}_{n,n-1}$ the Grassmannian manifold of hyperplanes through the origin contained in $\mathbb{R}^n$. We also denote by $d\mathbb{G}(H)$ the unique probability measure on $\mathbb{G}_{n,n-1}$ invariant under the action of the orthogonal group.

Theorem 4.1 *The length of γ is equal to*

$$L(\gamma) = \int_a^b \left\| \frac{d}{dt}\gamma(t) \right\| dt = \pi \int_{H \in \mathbb{G}_{n,n-1}} \sharp(H \cap \gamma) \quad d\mathbb{G}(H)$$

where $\sharp(H \cap \gamma)$ denotes the number of parameters $a \leq t \leq b$ such that $\gamma(t) \in H$.

Proof. If γ is an embedding then Theorem 4.1 follows from Santalo [12], chapter 18, section 6 or also see Shub and Smale [14], section 4, where a similar theorem is proved for projective spaces and Edelman and Kostlan [6]. 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 γ is an embedding. ∎

Note that by a usual application of Sard's Theorem the integral only needs to be evaluated on the set of H such that γ is transversal to H.

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 text-books. Morgan and Sommese [9] 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 $X_1, \ldots, X_m$ with $k_j + 1$ variables into 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 $\lambda \in \mathbb{C}$,

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

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 $\Pi_{j=1}^m \alpha_j^{k_j}$ in the product $\Pi_{i=1}^n \sum_{j=1}^m d_{ij}\alpha_j$.

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

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

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

is surjective. Notice that this definition is independent of the representative $(X_1, \ldots, X_m) \in \mathbb{C}^{n+m}$. 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 $P^{k_1}(\mathbb{C}) \times \ldots \times \mathbb{P}^{k_m}(\mathbb{C})$ is less than or equal to $\mathcal{B}$. If all the zeros are non-singular then f has exactly $\mathcal{B}$ zeros.*

6 An estimate for the total curvature of the central path

To the matrix A and the vector b which define the set of admissible points of the linear programming problem we may associate 2^m different systems of inequalities

$$A_i x \ \epsilon_i \ b_i, 1 \leq i \leq m,$$

with $\epsilon_i \in \{\leq, \geq\}$. Some of these systems are consistent and define a compact and non-void polytope. We denote by $\mathcal{P}(A, b)$ the set of these polytopes. To each $\mathcal{P} \in \mathcal{P}(A, b)$ we associate the linear programming problem

$$\min_{x \in \mathcal{P}} \langle c, x \rangle$$

with a central path $\mathbf{c}(\mathcal{P}, c)$ and a total curvature $K(\mathcal{P}, c)$.

In this section we give the following estimate for the total curvature of the union of these central path:

Theorem 6.1 *The sum of the total curvatures of the central path satisfies*

$$\sum_{\mathcal{P} \in \mathcal{P}(A,b)} K(\mathcal{P}, c) < \pi m \begin{pmatrix} m+1 \\ n \end{pmatrix}.$$

Remark 6.1 *We notice that this number depends only on the number of variables and the number of variables in (LP).*

G. Malajovich, using BKK bounds instead of multi-homogeneous Bézout numbers, conjectured the better bound

$$K \leq 2\pi(n-1)\begin{pmatrix} m-1 \\ n \end{pmatrix}.$$

Our main conjecture is a bound of $O(n)$ independent on m for one central path.

Proof. For two vectors u_1 and u_2 with the same dimension we often write $u_1.u_2 = \langle u_1, u_2 \rangle = u_1^T u_2$ for their dot product and $u_1 u_2$ for the componentwise product also equal to $u_1 u_2 = \mathrm{Diag}(u_{1i})u_2 = \mathrm{Diag}(u_{2i})u_1$.

Any of the considered central paths $\mathbf{c}(t)$ satisfies the equations

$$\mathbf{c}(t) = x, \quad A^T D_x^{-1} e = \frac{c}{t}, \quad t > 0.$$

By differentiating we get

$$A^T D_x^{-2} A \dot{\mathbf{c}}(t) = -\frac{c}{t^2}.$$

This last formula proves that $\dot{\mathbf{c}}(t) \neq 0$ so that the curvature is well defined. The total curvature is the length of the curve $\dot{\mathbf{c}}(t)/\|\dot{\mathbf{c}}(t)\|$ which is itself given by the number of parameters corresponding to the intersections of this curve with a generic hyperplane

$$\mathcal{H}(d) = \{x \in \mathbb{R}^n \ : \ \langle x, d \rangle = 0\}$$

see Theorem 4.1. Such an intersection point satisfies

$$\begin{aligned} &A^T D_x^{-1} e = \frac{c}{t}, \\ &\langle (A^T D_x^{-2} A)^{-1} c, d \rangle = 0. \end{aligned} \tag{6.1}$$

Notice that the number of parameters corresponding to the intersections is equal to the number of solutions of system 6.1 because, by Lemma 2.1, the map $x \rightarrow A^T D_x^{-1} e$ is invertible. Let us now prove that each solution of this system is isolated. We only consider the case of a hyperplane $\mathcal{H}(d)$ that cuts the image of the Gauss map transversaly. That means, see section 3,

$$\langle \ddot{\mathbf{c}}(t) \|\dot{\mathbf{c}}(t)\|^2 - \dot{\mathbf{c}}(t) \langle \dot{\mathbf{c}}(t), \ddot{\mathbf{c}}(t) \rangle, d \rangle \neq 0.$$

Since

$$\langle (A^T D_x^{-2} A)^{-1} c, d \rangle = 0$$

that is $\langle \dot{\mathbf{c}}, d \rangle = 0$ our transversality hypothesis is equivalent to $\langle \ddot{\mathbf{c}}, d \rangle \neq 0$. We already noticed that, with $x = \mathbf{c}(t)$,

$$A^T D_x^{-2} A \dot{\mathbf{c}} = -\frac{c}{t^2}.$$

Differentiating this equation gives

$$-2A^T D_x^{-3}(A\dot{\mathbf{c}})^2 + A^T D_x^{-2} A\ddot{\mathbf{c}} = 2\frac{c}{t^3}$$

(the product $(A\dot{\mathbf{c}})^2$ is taken componentwise) so that

$$\ddot{\mathbf{c}} = -2\frac{\dot{\mathbf{c}}}{t} + 2(A^T D_x^{-2} A)^{-1} A^T D_x^{-3}(A\dot{\mathbf{c}})^2.$$

Since $\langle \dot{\mathbf{c}}, d\rangle = 0$ our transversality hypothesis becomes

$$d^T (A^T D_x^{-2} A)^{-1} A^T D_x^{-3}(A\dot{\mathbf{c}})^2 \neq 0.$$

Let us denote

$$F(x,t) = \begin{pmatrix} A^T D_x^{-1} e - \frac{c}{t} \\ d^T (A^T D_x^{-2} A)^{-1} c \end{pmatrix}.$$

Our aim is to show that, under the previous transversality hypothesis, the derivative $DF(x,t)$ is non-singular. An easy computation shows that

$$DF(x,t)(\dot{x},\dot{t}) = \begin{pmatrix} A^T D_x^{-2} A\dot{x} + \dot{t}\frac{c}{t^2} \\ 2d^T (A^T D_x^{-2} A)^{-1} A^T D_x^{-3} \mathrm{Diag}(A_i \dot{x}) A (A^T D_x^{-2} A)^{-1} c \end{pmatrix}.$$

Let us suppose that $DF(x,t)(\dot{x},\dot{t}) = 0$. Since $A^T D_x^{-2} A$ is injective and $A^T D_x^{-2} A\dot{\mathbf{c}} = -c/t^2$ the first equation shows that $\dot{x}$ is proportional to $\dot{\mathbf{c}}$: $\dot{x} = \dot{t}\dot{\mathbf{c}}$. If we insert this fact into the second equation we obtain $\dot{t} = 0$ by the transversality hypothesis. This gives $\dot{x} = 0$ thus $DF(x,t)$ is invertible and all the zeros of system 6.1 are isolated and have multiplicity 1.

Let us now introduce five new vectors and scalars $s, q, v \in \mathbb{R}^m$, $w \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}$ with

$$\begin{aligned} &\sigma t = 1, \\ &s = Ax - b, \\ &q = D_x^{-1} e, \\ &w = (A^T D_x^{-2} A)^{-1} c, \\ &v = D_x^{-2} A w. \end{aligned} \tag{6.2}$$

System 6.1 becomes

$$\begin{aligned} &s > 0, \\ &Ax - s = b, \\ &sq = e, \\ &A^T q = \sigma c, \\ &w.d = 0, \\ &A^T v = c, \\ &sv = q(Aw). \end{aligned} \tag{6.3}$$

This system has $3m + 2n + 1$ equations and $3m + 2n + 1$ unknowns: $s, q, v \in \mathbb{R}^m$, $x, w \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}$. Notice that it has the same number of solutions than

6.1 and only isolated ones. Since $A^T q = \sigma c$ with $\sigma \in \mathbb{R}$ the vector q lies in the reciprocal image under A^T of the vector line through c. For this reason q depends only on $m - n + 1$ parameters: we can write $q = Bu$ for a linear and injective operator $B : \mathbb{R}^{m-n+1} \to \mathbb{R}^m$. This gives, with $q = Bu$ and $s = Ax - b$ a new system of equations

$$\begin{array}{l} (Ax-b)(Bu) = e, \\ w.c = 0, \\ A^T v = c, \\ (Ax-b)v = (Bu)(Aw), \end{array} \tag{6.4}$$

which contains $2m+n+1$ equations and $2m+n+1$ unknowns, $v \in \mathbb{R}^m$, $x, w \in \mathbb{R}^n$ and $u \in \mathbb{R}^{m-n+1}$. This system also has the same number of solutions than 6.1 and only isolated ones.

To bound this number we compute the Bézout number associated with a multi-homogenization of this system. We introduce three new variables $\nu_1, \nu_2, \nu_3 \in \mathbb{R}$. System 6.4 becomes

$$\begin{array}{l} (Ax-\nu_1 b)(Bu) = \nu_1\nu_2 e, \\ w.c = 0, \\ A^T v = \nu_3 c, \\ \nu_2(Ax-\nu_1 b)v = \nu_1(Bu)(Aw). \end{array} \tag{6.5}$$

It has $2m + n + 1$ equations and $2m + n + 4$ unknowns. This system is trilinear according to the structure

$$X_1 = \{x, \nu_1\}, \quad X_2 = \{u, \nu_2\}, \quad X_3 = \{v, w, \nu_3\}.$$

The degrees of the different equations are

- $(1, 1, 0)$ for $1 \le i \le m$,
- $(0, 0, 1)$ for $i = m + 1$,
- $(0, 0, 1)$ for $m + 2 \le i \le m + n + 1$,
- $(1, 1, 1)$ for $m + n + 2 \le i \le 2m + n + 1$.

The Bézout number is equal to the coefficient of

$$\alpha_1^n \alpha_2^{m-n+1} \alpha_3^{m+n}$$

in the product

$$(\alpha_1 + \alpha_2)^m \alpha_3^1 \alpha_3^n (\alpha_1 + \alpha_2 + \alpha_3)^m.$$

This coefficient is equal to $m \begin{pmatrix} m+1 \\ n \end{pmatrix}$, this achieves the proof. ■

7 The total curvature of the central path on the average.

Before stating our main theorem we have to explain the meaning we give to the word *average*.

We already noticed in the previous section that, to the data (A, b, c) which defines the linear programming problem, we may associate 2^m different systems of inequalities

$$A_i x \ \epsilon_i \ b_i, 1 \leq i \leq m,$$

with $\epsilon_i \in \{\leq, \geq\}$. All these systems are not necessarily consistent. In general the number of non-void polytopes is given by the recurrence

$$R(m, n) = R(m-1, n) + R(m-1, n-1), \quad R(1, n) = 2, \quad R(m, 1) = m + 1.$$

So

$$R(m, n) = \sum_{i=0}^{n} \binom{m}{i}.$$

The number of non-compact polytopes, in a generic sense, satisfy a similar relation

$$R_\infty(m, n) = R_\infty(m-1, n) + R_\infty(m-1, n-1), \quad R_\infty(1, n) = 2, \quad R_\infty(m, 1) = 2,$$

so that the number of compact polytopes is such that

$$R_K(m, n) = R_K(m-1, n) + R_K(m-1, n-1), \quad R_K(1, n) = 0, \quad R_K(m, 1) = m-1,$$

thus

$$R_K(m, n) = \binom{m-1}{n}.$$

Let $\mathcal{P}(A, b)$ denotes the set of compact polytopes defined by the inequalities related with A and b. In general

$$\sharp\mathcal{P}(A, b) = \binom{m-1}{n}.$$

To each $\mathcal{P} \in \mathcal{P}(A, b)$ we associate a linear programming problem

$$\min_{x \in \mathcal{P}} \langle c, x \rangle$$

with a central path $\mathbf{c}(\mathcal{P}, c)$ and a total curvature $K(\mathcal{P}, c)$.

Definition 7.1 *We call "average total curvature" associated with the data* (A, b, c) *the number*

$$\mathcal{K}(A, b, c) = \frac{\sum_{\mathcal{P} \in \mathcal{P}(A,b)} K(\mathcal{P}, c)}{\sharp\mathcal{P}(A, b)}.$$

Our main result is the following

Theorem 7.1 *In general*

$$\mathcal{K}(A,b,c) \leq \frac{\pi m^2(m+1)}{(m-n)(m-n+1)}.$$

Proof. This number is equal to

$$\frac{\pi m \begin{pmatrix} m+1 \\ n \end{pmatrix}}{\begin{pmatrix} m-1 \\ n \end{pmatrix}}$$

that is the bound obtain in Theorem 6.1 divided by the number of compact polytopes. ■

8 Concluding remarks

1. We have averaged the total curvature over compact regions because each compact region has a central path. Non-compact regions may also have a central path. For a polytope $\mathcal{P}$ defined by the inequalities $A_i x \geq b_i$ a sufficient condition is: for any non-zero $x \in \mathbb{R}^n$, if $A_i x \geq 0$ for any $i = 1 \ldots m$ then $\langle c, x \rangle > 0$ giving an average total curvature much smaller. We have not averaged on the total number of possible systems of inequalities as Beling and Verma in [3].

2. We have estimated the curvature by the number of complex roots of a system of equations including the 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 [13], Edelman and Kostlan [6], McLennan [7] and Rojas [11]. 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 [15] and [16], Stoer and Zhao [18], Zhao [20] and [21]. These papers use a different notion of curvature, closer to $1/\gamma$ where γ is Smale's γ, see also Dedieu and Smale [5]. The integral of these quantities is infinite.

4. The Riemannian geometry of the central path has been studied by quite a few authors, see Nesterov and Todd [10], Bayer and Lagarias [1] and [2].

5. Vavasis and Ye study regions where the central path is straight or crossing over. In straight regions the tangent vectors stay in definite cones. Curvature estimates may be used as a refinement of this information.

6. We have studied central paths for the linear programming problem in the format $\min\langle c, x\rangle$ subject to $Ax \geq b$. The format is not important for our results.

7. We have studied via a similar argument the total curvature of the central path in a primal-dual formulation. Consider the standard-form problem

$$\begin{array}{ll} \min & \langle c, x\rangle \\ Ax = b & \\ x \geq 0 & \end{array}$$

together with

$$\begin{array}{ll} \max & \langle b, y\rangle. \\ A^T y + s = c & \\ s \geq 0 & \end{array}$$

Here A is an $m \times n$ matrix with Rank $A = m$ and the vectors are of appropriate size. The associated central path $t > 0 \to (x(t), y(t), s(t))$ is given by the equations

$$\begin{array}{l} A^T y + s = c \\ Ax = b \\ xs = te. \end{array} \tag{8.6}$$

To compute the associated total curvature we count the number of solutions of the system

$$\begin{array}{l} A^T y + s = c \\ Ax = b \\ xs = te \\ A^T \dot{y} = 0 \\ A\dot{x} = 0 \\ x\dot{s} + \dot{x}s = e \\ c_1\dot{x} + c_2\dot{s} + c_3\dot{y} = 0 \end{array} \tag{8.7}$$

which has $4n + 2m + 1$ equations and unknowns. We multi-homogenize it in partitioning the variables in four groups: $X_1 = \{s, y, \nu_1\}$, $X_2 = \{x, \nu_2\}$, $X_3 = \{t, \nu_3\}$, $X_4 = \{\dot{x}, \dot{s}, \dot{y}, \nu_4\}$. We have obtain, via this computation, the bound

$$\pi \frac{n^3}{m}$$

that is $n^2 \begin{pmatrix} n \\ m \end{pmatrix}$ zeros of the system of equations divided by $\begin{pmatrix} n-1 \\ n-m \end{pmatrix}$ compact regions.

Here m and n do not play the same role as in the main part of this paper. To compare both numbers we make the substitution $n := m$ and $m := m - n$ and we obtain

$$\pi \frac{m^3}{n - m}$$

which has the same flavor as our main result.

References

[1] D. Bayer and J. Lagarias, *The Non-Linear Geometry of Linear Programming I: Affine and projective scaling trajectories, Trans. Amer. Math. Soc. 314 (1989) no.2, 499-526, II: Legendre Transform Coordinates and Central Trajectories.* Trans. Amer. Math. Soc. 314 (1989) 527-581.

[2] D. Bayer and J. Lagarias, *Karmarkar's Linear Programming Algorithm and Newton's Method.* Math. Progr. A. 50 (1991) 291-330.

[3] P. Beling and S. Verma, *A Probabilistic Analysis of a Measure of Combinatorial Complexity for the Central Curve.* Math. Program. A, 87 (2000) 177-187.

[4] J.-P. Dedieu and M. Shub, *On Random and Mean Exponents for Unitarily Invariant Probability Measures on* $\mathbb{G}L_n(\mathbb{C})$. To appear in: Astérisque, 2003.

[5] J.-P. Dedieu and S. Smale, *Some Lower Bounds for the Complexity of Continuation Methods* J. of Complexity Vol 14 (1998) 454-465.

[6] A. Edelman and E. Kostlan, *How many zeros of a random polynomial are real* Bull. Amer. Math. Soc. 32 (1995) 1-37.

[7] A. McLennan, *The Expected Number of Real Roots of a Multi-homogeneous System of Polynomial Equations,* Amer. J. of Math. 124 (2002) 49-73.

[8] N. Meggido and M. Shub, *Boundary Behaviour of of Interior Point Algorithms in Linear Programming.* Mathematics of Operations Research 14 (1989) 97-146.

[9] A. Morgan, A. Sommese *A Homotopy for Solving General Polynomial Systems that Respects m-Homogeneous Structures.* Appl. Math. Comput. 24 (1987) 101-113.

[10] Y. Nesterov and M. Todd, *On the Riemannian Geometry Defined by Self-Concordant Barriers and Interior-Point Method.* FOCM, 2 (2002) 333-361.

[11] M. Rojas, *On the Average Number of Certain Random Sparse Polynomial Systems,* Lectures in Appl. Math. vol. 32, J. Renegar, M. Shub and S. Smale eds., Amer. Math. Soc. (1996) 689-699.

[12] L. Santalo, *Integral Geometry and Geometric Probability.* Addison-Wesley Publishing Company, 1976.

[13] Shub, M., S. Smale, *Complexity of Bézout's Theorem II: Volumes and Probabilities* in: *Computational Algebraic Geometry,* F. Eyssette and A. Galligo eds., em Progress in Mathematics, vol. 109, Birkhäuser, 1993, 267-285.

[14] M. Shub, S. Smale, *Complexity of Bézout's Theorem V: Polynomial Time*, Theoretical Computer Science, (1994) vol. 133, pp.141-164.

[15] G. Sonnevend, J. Stoer and G. Zhao, *On the Complexity of Following the Central Path of Linear Programs by Linear Extrapolation.* In: Methods of Op. Research, Proceedings of the Symposium on Op. Research, Ulm, 1989. Anton Hain. 19-31.

[16] G. Sonnevend, J. Stoer and G. Zhao, *On the Complexity of Following the Central Path of Linear Programs by Linear Extrapolation: II.* Math. Programming, 52 (1991) 527-553.

[17] M. Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. 2.* 2nd ed. Publish or Perish, Berkeley, 1990.

[18] J. Stoer and G. Zhao, *Estimating the Complexity of a Class of Path-Following Methods for Solving Linear Programs by Curvature Integrals.* Applied Math. Opt. 27 (1993) 85-103.

[19] S. Vavasis and Y. Ye *A Primal-Dual Accelerated Interior Point Method Whose Running Time Depends Only on A.* Math. Progr. 74 (1996) 79-120.

[20] G. Zhao, *On the Relationship Between the Curvature Integral and the Complexity of Path-Following Methods in Linear Programming.* SIAM J. Opt. 6 (1996) 57-73.

[21] G. Zhao, *Interior Point Algorithms for Linear Complementarity Problems Based on Large Neighborhoods on the central Path.* SIAM J. Opt. 8 (1998) 397-413.