A Numerical Algorithm for Zero Counting.
II: Distance to Ill-posedness and Smoothed Analysis

Felipe Cucker *
Dept. of Mathematics
City University of Hong Kong
HONG KONG
e-mail: macucker@cityu.edu.hk

Teresa Krick †
Departamento de Matemática
Univ. de Buenos Aires & CONICET
ARGENTINA
e-mail: krick@dm.uba.ar

Gregorio Malajovich‡
Depto. de Matemática Aplicada
Univ. Federal do Rio de Janeiro
BRASIL
e-mail: gregorio@ufrj.br

Mario Wschebor
Centro de Matemática
Universidad de la República
URUGUAY
e-mail: wschebor@cmat.edu.uy

Abstract. We show a Condition Number Theorem for the condition number of zero counting for real polynomial systems. That is, we show that this condition number equals the inverse of the normalized distance to the set of ill-posed systems (i.e., those having multiple real zeros). As a consequence, a smoothed analysis of this condition number follows.

1 Introduction

This paper continues the work in [8], where we described a numerical algorithm to count the number of zeros in n-dimensional real projective space of a system of n real homogeneous polynomials. The algorithm works with finite precision and both its complexity and the precision required to ensure correctness are bounded in terms of n, the maximum $\mathbf{D}$ of the polynomials' degrees, and a condition number $\kappa(f)$.

In this paper we replace $\kappa(f)$ —which was originally defined using the computationally friendly infinity norm— for a version $\widetilde{\kappa}(f)$ (defined in Section 2 below) which uses instead Euclidean norms. This difference is of little consequence in complexity estimates since one

*Partially supported by City University SRG grant 7002106.
†Partially supported by grants ANPCyT 33671/05 and UBACyT X113/08-10.
‡Partially supported by CNPq grants 470031/2007-7, 303565/2007-1 and by FAPERJ grant E26/110.849/2009.

has (cf. Proposition 3.3 below) that

$$\frac{\widetilde{\kappa}(f)}{\sqrt{n}} \leq \kappa(f) \leq \sqrt{2n}\widetilde{\kappa}(f). \tag{1}$$

It allows one, however, to prove a result following a classical theme in conditioning —the relation between condition and distance to ill-posedness— and to deduce from this result a probabilistic analysis for $\widetilde{\kappa}(f)$ and, a fortiori, for the complexity of the algorithm in [8]. This analysis is of interest since, in contrast with n and $\mathbf{D}$, the value of $\widetilde{\kappa}(f)$ (or of $\kappa(f)$) is not apparent in f and therefore complexity or accuracy bounds depending on this condition number are not of immediate use. A solution pioneered by John von Neumann and collaborators (see [14, §2.1] and references therein) and reintroduced by Steve Smale [23, 24] is to assume a probability measure on the space of data and to study the condition number at data f as a random variable. This approach yields bounds on accuracy or complexity for random data and has been pursued in several contexts: systems of linear equations [12, 13], polyhedral conic systems [3, 5, 9, 15, 25], linear programs [4], complex polynomial systems [21], [18], etc. In our case it allows us to trade the presence of $\widetilde{\kappa}(f)$ in deterministic bounds for probabilistic bounds in n and $\mathbf{D}$ only.

1.1 Distance to Ill-posedness

It has since long been observed [11] that the condition number for several problems (in its original aception, as a measure of the worst possible magnification of small input errors in the output [19]) either coincides with the relativized inverse of the distance from the input to the set of ill-posed data or is bounded by a small multiple of this inverse. A data is *ill-posed* when the magnification above is unbounded. In our case, a polynomial system f is ill-posed when arbitrary small perturbations of f may change its number of projective real zeros. Systems having this property are exactly those having multiple (real projective) zeros. Let us denote the set of such systems by $\Sigma_{\mathbb{R}}$. Also, let $\mathcal{H}_{\mathbf{d}}$ denote our set of input systems, i.e. the set of $f = (f_1, \ldots, f_n)$, n real homogeneous polynomials in $n+1$ variables of degrees $\mathbf{d} := (d_1, \ldots, d_n)$ respectively, endowed with the Bombieri-Weyl norm $\| \ \|_W$ (defined in Section 2). Finally, let dist be the distance on $\mathcal{H}_{\mathbf{d}}$ induced by this norm.

Our main result in this note is the following.

Theorem 1.1 *For all* $f \in \mathcal{H}_{\mathbf{d}}$, $\widetilde{\kappa}(f) = \frac{\|f\|_W}{\mathsf{dist}(f, \Sigma_{\mathbb{R}})}$.

Remark 1.2 It is worth noting that, although $\widetilde{\kappa}(f)$ is somehow related to the condition number $\mu_{\rm norm}(f)$ for complex polynomial systems (cf. [1, 20]) a result like Theorem 1.1 does not hold for the latter. Actually, such a result holds on the fibers of the zeros (see [1, Ch. 12, Theorem 3]) but not globally.

We also note that, as a consequence of Theorem 1.1, we have $\widetilde{\kappa}(f) \geq 1$ for all $f \in \mathcal{H}_{\mathbf{d}}$. This feature, although not immediate from the definition of $\widetilde{\kappa}(f)$, follows immediately from the fact that $0 \in \Sigma_{\mathbb{R}}$.

1.2 Smoothed Analysis

Theorem 1.1 carries out meaningful consequences in the probabilistic analysis of $\widetilde{\kappa}(f)$.

The usual probabilistic analysis for polynomial systems assume that random f are drawn from the unit sphere $S(\mathcal{H}_{\mathbf{d}})$ with the uniform distribution (or from such a distribution on the real projective space induced by this sphere). A different approach to the randomization of input data has been recently proposed under the name of *smoothed analysis* [22]. The idea consists on replacing 'random data' by 'random perturbations of given data'. A recent result in [2] derives smoothed analysis bounds for condition numbers which can be written as inverses to distances to ill-posedness. Because of Theorem 1.1, these bounds can be straightforwardly applied to $\widetilde{\kappa}(f)$.

To state this result we need to introduce some notation. Let $\mathbb{P}^p(\mathbb{R})$ denote the real projective space of dimension p and $d_{\mathbb{P}}$ be the projective distance on $\mathbb{P}^p(\mathbb{R})$ (i.e. the sinus of the Riemannian distance). For a point $a \in \mathbb{P}^p(\mathbb{R})$ and $\sigma \in (0,1]$ we denote by $B(a,\sigma)$ the ball (w.r.t. $d_{\mathbb{P}}$) centered at a and of radius σ. That is,

$$B(a,\sigma) := \{x \in \mathbb{P}^p(\mathbb{R}) \mid d_{\mathbb{P}}(x,a) \leq \sigma\}.$$

In what follows we assume $B(a,\sigma)$ endowed with the uniform probability measure. Note that in the particular case $\sigma = 1$ we obtain $B(a,\sigma) = \mathbb{P}^p(\mathbb{R})$ for each $a \in \mathbb{P}^p$ and hence, the usual probabilsitic analysis referred to above.

Theorem 1.3 [2] *Let $S \subset \mathbb{P}^p$ be contained in a projective hypersurface H of degree at most d and $\mathscr{C} : \mathbb{P}^p \to [1,\infty]$ be given by*

$$\mathscr{C}(z) = \frac{1}{d_{\mathbb{P}}(z,S)}.$$

Then, for all $\sigma \in (0,1]$ and all $t \geq (2d+1)\frac{p}{\sigma}$,

$$\sup_{a\in\mathbb{P}^p} \operatorname*{Prob}_{z\in B(a,\sigma)} \{\mathscr{C}(z) \geq t\} \leq 13dp\frac{1}{\sigma t}$$

and

$$\sup_{a\in\mathbb{P}^p} \operatorname*{\mathbf{E}}_{z\in B(a,\sigma)} [\ln \mathscr{C}(z)] \leq 2\ln p + 2\ln d + \ln\left(\frac{1}{\sigma}\right) + 5. \qquad \square$$

As a consequence of this result we obtain the following corollary (which we prove in Section 3). Here $\mathbb{P}(\mathcal{H}_{\mathbf{d}})$ denotes the projective space associated to $\mathcal{H}_{\mathbf{d}}$, N denotes its dimension and $\mathcal{D} = d_1 \cdots d_n$ its associated Bézout number.

Corollary 1.4 *For all $\sigma \in (0,1]$ and all $t \geq (4n\mathcal{D}^2+1)\frac{N}{\sigma}$,*

$$\sup_{f\in\mathbb{P}(\mathcal{H}_{\mathbf{d}})} \operatorname*{Prob}_{g\in B(f,\sigma)} \{\widetilde{\kappa}(g) \geq t\} \leq 13n^2\mathbf{D}\mathcal{D}N\frac{1}{\sigma t}$$

and

$$\sup_{f\in\mathbb{P}(\mathcal{H}_{\mathbf{d}})} \mathop{\mathbf{E}}_{g\in B(f,\sigma)} [\ln \widetilde{\kappa}(g)] \le 2\ln N + 4\ln n + 2\ln \mathcal{D} + \ln \mathbf{D} + \ln\left(\frac{1}{\sigma}\right) + 6.$$

In particular, taking $\sigma = 1$*, we obtain average analysis: for all* $t \ge N(4n\mathcal{D}^2 + 1)$*,*

$$\mathop{\mathrm{Prob}}_{g\in\mathbb{P}^p}\{\widetilde{\kappa}(g) \ge t\} \le 13n^2\mathbf{D}\mathcal{D}N\frac{1}{t}$$

and

$$\mathop{\mathbf{E}}_{g\in\mathbb{P}^p}[\ln \widetilde{\kappa}(g)] \le 2\ln N + 4\ln n + 2\ln \mathcal{D} + \ln(\mathbf{D}) + 6.$$

A recent result [7] extends Theorem 1.3 to absolutely continuous measures on $B(a,\sigma)$ whose densities are radially symmetric around a and may have a pole at a. Applications of this result to $\widetilde{\kappa}(f)$ readily follow.

2 Setting and Notations

For $d \in \mathbb{N}$, $\mathcal{H}_d$ denotes the subspace of $\mathbb{R}[x_0,\dots,x_n]$ of homogeneous polynomials of degree d. We endow $\mathcal{H}_d$ with the Bombieri-Weyl inner product $\langle\ ,\ \rangle_W$, defined for $f = \sum_{|j|=d} a_j \boldsymbol{x}^j$ and $g = \sum_{|j|=d} b_j \boldsymbol{x}^j$ by

$$\langle f, g\rangle_W = \sum_{|j|=d} \frac{a_j b_j}{\binom{d}{j}}$$

where for $j = (j_0,\dots,j_n)$, $\binom{d}{j} := \frac{d!}{j_0!\cdots j_n!}$. A main feature of this inner product is its invariance under the action of the orthogonal group $O(n+1)$. That is, for all $\psi \in O(n+1)$ and all $f, g \in \mathcal{H}_d$, $\langle f\circ\psi, g\circ\psi\rangle_W = \langle f, g\rangle_W$. Next, for $d_1,\dots,d_n \in \mathbb{N}$, we endow $\mathcal{H}_{\mathbf{d}} = \mathcal{H}_{d_1}\times\ldots\times\mathcal{H}_{d_n}$ with the inner product

$$\langle f, g\rangle_W = \sum_{i=1}^{n} \langle f_i, g_i\rangle_W$$

where $f = (f_1,\dots,f_n)$, $g = (g_1,\dots,g_n) \in \mathcal{H}_{\mathbf{d}}$. We write $\|\ \|_W$ and dist to denote the norm and distance on $\mathcal{H}_{\mathbf{d}}$ induced by this inner product.

Projective zeros of polynomial systems $f \in \mathcal{H}_{\mathbf{d}}$ correspond to pairs of zeros $(-\zeta,\zeta)$ of the restriction $f_{|S^n}$ of f to the n-dimensional unit sphere $S^n \subset \mathbb{R}^{n+1}$. We will thus consider a system $f \in \mathcal{H}_{\mathbf{d}}$ as a (centrally symmetric) mapping of S^n into $\mathbb{R}^n$.

For a point $x \in S^n$ and a system $f \in \mathcal{H}_{\mathbf{d}}$ one may define both ill-posedness and condition relative to this point. For the first, one defines

$$\Sigma_{\mathbb{R}}(x) = \{f \in \mathcal{H}_{\mathbf{d}} \mid x \text{ is a multiple zero of } f\},$$

the set of systems which are ill-posed at x. Note that $\Sigma_{\mathbb{R}}(x) \ne \emptyset$ for all $x \in S^n$ and that

$$\Sigma_{\mathbb{R}} = \{f \in \mathcal{H}_{\mathbf{d}} \mid f \text{ has a multiple zero in } S^n\} = \bigcup_{x\in S^n} \Sigma_{\mathbb{R}}(x).$$

Towards the definition of $\widetilde{\kappa}(f)$, for $f \in \mathcal{H}_{\mathbf{d}}$ and $x \in S^n$, we define

$$\widetilde{\mu}_{\rm norm}(f,x) = \|f\|_W \left\| Df(x)^{-1}_{|T_xS^n} \begin{bmatrix} \sqrt{d_1} & & & \\ & \sqrt{d_2} & & \\ & & \ddots & \\ & & & \sqrt{d_n} \end{bmatrix} \right\| \tag{2}$$

where $Df(x)_{|T_xS^n}$ is the restriction to the tangent space of x at S^n of the derivative of f at x and the norm is the spectral norm, i.e. the operator norm with respect to $\| \, \|_2$.

Next, we define the *condition of f relative to x* to be

$$\widetilde{\kappa}(f,x) = \frac{\|f\|_W}{\left(\|f\|_W^2 \widetilde{\mu}_{\rm norm}(f,x)^{-2} + \|f(x)\|_2^2\right)^{1/2}}.$$

Finally, we take the *condition number* $\widetilde{\kappa}(f)$ of $f \in \mathcal{H}_{\mathbf{d}}$ to be its condition relative to its worst conditioned point,

$$\widetilde{\kappa}(f) = \max_{x \in S^n} \widetilde{\kappa}(f,x)$$

Note that for all $\lambda \neq 0$, $\widetilde{\kappa}(\lambda f) = \widetilde{\kappa}(f)$ and $\mathsf{dist}(\lambda f, \Sigma_{\mathbb{R}}) = |\lambda| \mathsf{dist}(f, \Sigma_{\mathbb{R}})$. The same is true relative to a point $x \in S^n$. We will therefore assume, without loss of generality, that $\|f\|_W = 1$, and denote by $S(\mathcal{H}_{\mathbf{d}})$ the unit sphere in $\mathcal{H}_{\mathbf{d}}$.

3 The proofs

3.1 The main results

Proposition 3.1 *For all $x \in S^n$ and $f \in S(\mathcal{H}_{\mathbf{d}})$,*

$$\widetilde{\kappa}(f,x) = \frac{1}{\mathsf{dist}(f, \Sigma_{\mathbb{R}}(x))}.$$

PROOF. For $0 \le i \le n$, let $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ denote the ith coordinate vector. The group $O(n+1)$ acts on $\mathcal{H}_{\mathbf{d}} \times S^n$ and leaves $\mu_{\rm norm}, \widetilde{\kappa}$ and distance to $\Sigma_{\mathbb{R}}(\,)$ invariant. Therefore, we may assume without loss of generality that $x = e_0$. This implies that $T_{e_0}S^n \simeq \langle e_1, \ldots, e_n \rangle$ and we may write the singular value decomposition

$$\mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Df(e_0)_{|T_{e_0}S^n} = \underbrace{\begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}}_{U} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} V^{\rm t}$$

with U and V orthogonal and $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge 0$. Since the subgroup of $O(n+1)$ leaving e_0 invariant is isomorphic to $O(n)$ acting on $T_{e_0}S^n$ we may as well assume that $V = \mathrm{Id}$. Note that $\widetilde{\mu}_{\rm norm}(f, e_0) = \sigma_n^{-1}$, and therefore $\widetilde{\kappa}(f, e_0) = (\sigma_n^2 + \|f(e_0)\|_2^2)^{-1/2}$.

Let $g_i(x) := f_i(x) - f_i(e_0)x_0^{d_i} - \sqrt{d_i}\sigma_n u_{in} x_0^{d_i-1}x_n$, where $u_n = (u_{1n}, \dots, u_{nn})^{\mathrm{t}}$. Clearly, $g_i(e_0) = 0$ and $Dg_i(e_0)e_n = 0$ (here and in the sequel we denote $Dg_i(e_0)_{|T_{e_0}S^n}$ simply by $Dg_i(e_0)$) since $\partial g_i/\partial x_n(e_0) = \partial f_i/\partial x_n(e_0) - \sqrt{d_i}u_{in}\sigma_n = 0$. Thus, $g = (g_1, \dots, g_n) \in \Sigma_{\mathbb{R}}(e_0)$. Moreover

$$\|f_i - g_i\|_W^2 = \binom{d_i}{d_i}^{-1} f_i(e_0)^2 + \binom{d_i}{d_i - 1, 1}^{-1} (\sqrt{d_i}\sigma_n u_{in})^2 = |f_i(e_0)|^2 + \sigma_n^2 u_{in}^2$$

and hence, using $\|u_n\| = 1$,

$$\|f - g\|_W^2 = \|f(e_0)\|_2^2 + \sigma_n^2 = \widetilde{\kappa}(f, e_0)^{-2}.$$

It follows that

$$\mathsf{dist}(f, \Sigma_{\mathbb{R}}(e_0)) \leq \|f - g\|_W = \widetilde{\kappa}(f, e_0)^{-1}.$$

For the reciprocal, let $g \in \Sigma_{\mathbb{R}}(e_0)$. Then, $g(e_0) = 0$ and $Dg(e_0)$ is singular. We want to show that $\|f - g\|_W \geq \widetilde{\kappa}(f, e_0)^{-1}$. To this end, we write

$$f_i(x) = f_i(e_0)x_0^{d_i} + \frac{\partial f_i}{\partial x_1}(e_0)x_0^{d_i-1}x_1 + \cdots + \frac{\partial f_i}{\partial x_n}(e_0)x_0^{d_i-1}x_n + Q_i(x)$$

with $\deg_{x_0} Q_i \leq d_i - 2$ and, similarly,

$$g_i(x) = \frac{\partial g_i}{\partial x_1}(e_0)x_0^{d_i-1}x_1 + \cdots + \frac{\partial g_i}{\partial x_n}(e_0)x_0^{d_i-1}x_n + R_i(x).$$

Then

$$\|f_i - g_i\|_W^2 \geq |f_i(e_0)|^2 + \frac{1}{d_i}\|Df_i(e_0) - Dg_i(e_0)\|_2^2$$

and

$$\|f - g\|_W^2 \geq \|f(e_0)\|_2^2 + \left\|\mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Df(e_0) - \mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Dg(e_0)\right\|_F^2.$$

We know that $\mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Dg(e_0)$ is singular. Hence, denoting by Sing_n the set of singular $n \times n$ matrices and by dist_F the Frobenius distance on this set of matrices,

$$\mathsf{dist}_F\left(\mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Df(e_0), \mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Dg(e_0)\right) \geq \mathsf{dist}_F\left(\mathsf{diag}\left(\frac{1}{\sqrt{d_i}}\right) Df(e_0), \mathsf{Sing}_n\right) = \sigma_n,$$

the last by the Eckart-Young Theorem [1, §11.1]. It follows that

$$\|f - g\|_W^2 \geq \|f(e_0)\|_2^2 + \sigma_n^2 = \widetilde{\kappa}(f, e_0)^{-2}. \qquad \square$$

Proof of Theorem 1.1. Again we can assume $f \in S(\mathcal{H}_{\mathbf{d}})$. Note that

$$\mathsf{dist}(f, \Sigma_{\mathbb{R}}) = \min_{g \in \Sigma_{\mathbb{R}}} \mathsf{dist}(f, g) = \min_{x \in S^n} \mathsf{dist}(f, \Sigma_{\mathbb{R}}(x))$$

since $\Sigma_{\mathbb{R}} = \bigcup_{x\in S^n} \Sigma_{\mathbb{R}}(x)$. Then,

$$\widetilde{\kappa}(f) = \max_{x\in S^n} \widetilde{\kappa}(f,x) = \max_{x\in S^n} \frac{1}{\mathsf{dist}(f,\Sigma_{\mathbb{R}}(x))} = \frac{1}{\min\limits_{x\in S^n} \mathsf{dist}(f,\Sigma_{\mathbb{R}}(x))} = \frac{1}{\mathsf{dist}(f,\Sigma_{\mathbb{R}})}. \qquad \square$$

Before proving Corollary 1.4 we recall some useful facts in algebraic geometry.

For $1 \le i \le n$, let $f_i = \sum_{|j|=d_i} u_{ij}\boldsymbol{x}^j$ be a generic (i.e. with indeterminate coefficients) homogeneous polynomial of degree d_i in the variables $\boldsymbol{x} = (x_0,\dots,x_n)$ and $f = (f_1,\dots,f_n)$. Set $N := \sum_{i=1}^n \binom{d_i+n}{n} - 1$, the dimension of the projective coefficients space. The **d**-*discriminant variety* $\Sigma_{\mathbb{C}} \subset \mathbb{P}^N(\mathbb{C})$ is the locus of such polynomial systems $f = (f_1,\dots,f_n)$ with multiple zeros, i.e. such that there exists $z \in \mathbb{C}^{n+1}$, $z \ne 0$, with $f_1(z) = \cdots = f_n(z) = 0$ and $Df(z)$ has rank $< n$. It is well-known that $\Sigma_{\mathbb{C}}$ is a hypersurface in $\mathbb{P}^N(\mathbb{C})$ defined by an irreducible polynomial $\mathrm{Disc}(f) \in \mathbb{Z}[u_{ij}]$ (see [17] or [16, Ch. 10]). For lack of a precise reference we prove the following result.

Lemma 3.2

$$\deg(\Sigma_{\mathbb{C}}) = n\mathcal{D} + (d_1 + \cdots + d_n - n - 1)\mathcal{D}\sum_{j=1}^n \frac{1}{d_j}.$$

Proof. We know that $\deg(\Sigma_{\mathbb{C}}) = \deg(\mathrm{Disc}(f))$. We apply Identity (13) of [10]:

$$\mathrm{Res}_{\rho,d_1,\dots,d_n}(J_f, f_1,\dots,f_n) = \mathrm{Res}_{d_1,\dots,d_n}(f_1^0,\dots,f_n^0)\mathrm{Disc}(f),$$

where the standard notation $\mathrm{Res}_{d_1,\dots,d_n}$ is for the multihomogeneous projective resultant of n generic homogeneous polynomials of respective degrees $d_1,\dots,d_n$ in n variables, $\rho := d_1 + \cdots + d_n - n$, J_f is the determinant of the matrix $(\partial f_i/\partial x_j)_{1\le i,j\le n}$ and f_i^0 denotes the homogeneous component (of degree d_i) of $f_i(1,x_1,\dots,x_n)$.

We note that $\deg(J_f)$ is a polynomial of degree ρ in $\boldsymbol{x}$ whose coefficients are polynomials in u_{ij} of degree n. On the other hand $\mathrm{Res}_{d_1,\dots,d_n}$ is a multihomogeneous polynomial of degree $\prod_{k\ne i} d_k$ in the group of variables u_{ij} ([6]). Therefore, since J_f has degree n in the u_{ij}, we derive

$$n\mathcal{D} + \rho \sum_{1\le j\le n} \frac{\mathcal{D}}{d_j} = \sum_{1\le j\le n} \frac{\mathcal{D}}{d_j} + \deg\big(\mathrm{Disc}(f)\big).$$

The statement easily follows. $\square$

Proof of Corollary 1.4. We have $\Sigma_{\mathbb{R}} \subset \Sigma_{\mathbb{C}}$ and for all $f \in \mathcal{H}_{\mathbf{d}}$, $f \ne 0$,

$$\widetilde{\kappa}(f) = \frac{\|f\|_W}{\mathsf{dist}(f,\Sigma_{\mathbb{R}})} = \frac{1}{d_{\mathbb{P}}(f,\Sigma_{\mathbb{R}} \cap \mathbb{P}^N(\mathbb{R}))}.$$

The statement now follows from Theorem 1.3 with $p = N$, $S = \Sigma_{\mathbb{R}} \cap \mathbb{P}^N(\mathbb{R})$, $H = \Sigma_{\mathbb{C}} \cap \mathbb{P}^N(\mathbb{R})$, and $d = n^2\mathbf{D}\mathcal{D}$, the last by Lemma 3.2 since $n\mathcal{D} + (\rho - 1)\mathcal{D}\sum_{j=1}^n \frac{1}{d_j} \le n^2\mathbf{D}\mathcal{D}$. $\square$

3.2 The equivalence of condition numbers

In [8] we considered $\mathcal{H}_{\mathbf{d}}$ endowed with the norm given by

$$\|f\| := \max_{1\le i\le n} \|f_i\|_W$$

(we note $\|f\| \le \|f\|_W \le \sqrt{n}\|f\|$) and defined

$$\kappa(f) := \max_{x\in S^n} \min\left\{\mu_{\text{norm}}(f,x), \frac{\|f\|}{\|f(x)\|_\infty}\right\}$$

with

$$\mu_{\text{norm}}(f,x) := \sqrt{n}\,\|f\|\,\left\|Df(x)^{-1}_{|T_xS^n}\mathsf{diag}(\sqrt{d_i})\right\| = \sqrt{n}\frac{\|f\|}{\|f\|_W}\widetilde{\mu}_{\text{norm}}(f,x)$$

(we note $\widetilde{\mu}_{\text{norm}}(f,x) \le \mu_{\text{norm}}(f,x) \le \sqrt{n}\,\widetilde{\mu}_{\text{norm}}(f,x)$). The next result shows that the $\kappa(f)$ thus defined is closely related to $\widetilde{\kappa}(f)$.

Proposition 3.3

$$\frac{\widetilde{\kappa}(f)}{\sqrt{n}} \le \kappa(f) \le \sqrt{2n}\;\widetilde{\kappa}(f).$$

Proof. Let $x \in S^n$. We observe that since $\|f\|_W^2\widetilde{\mu}_{\text{norm}}(f,x)^{-2} + \|f(x)\|_2^2 \ge \|f\|_W^2\widetilde{\mu}_{\text{norm}}(f,x)^{-2}$, we have $\widetilde{\kappa}(f,x) \le \widetilde{\mu}_{\text{norm}}(f,x) \le \mu_{\text{norm}}(f,x)$. Similarly, using $\|f\|_W \le \sqrt{n}\|f\|$ and $\|f(x)\|_2 \ge \|f(x)\|_\infty$, we have $\widetilde{\kappa}(f,x) \le \sqrt{n}\|f\|/\|f(x)\|_\infty$. Therefore

$$\widetilde{\kappa}(f,x) \le \sqrt{n}\min\left\{\mu_{\text{norm}}(f,x), \frac{\|f\|}{\|f(x)\|_\infty}\right\},$$

which implies

$$\widetilde{\kappa}(f) = \max_{x\in S^n}\widetilde{\kappa}(f,x) \le \sqrt{n}\max_{x\in S^n}\min\left\{\mu_{\text{norm}}(f,x), \frac{\|f\|}{\|f(x)\|_\infty}\right\} = \sqrt{n}\,\kappa(f).$$

To prove the other inequality note that, for any $x \in S^n$,

$$\begin{aligned}
\min\left\{\|f\|^2\left\|Df(x)^{-1}_{|T_xS^n}\mathsf{diag}(\sqrt{d_i})\right\|^2, \frac{\|f\|^2}{\|f(x)\|_2^2}\right\} &\le \frac{2\|f\|^2}{\left\|Df(x)^{-1}_{|T_xS^n}\mathsf{diag}(\sqrt{d_i})\right\|^{-2} + \|f(x)\|_2^2}\\
&\le \frac{2\|f\|_W^2}{\|f\|_W^2\widetilde{\mu}_{\text{norm}}(f,x)^{-2} + \|f(x)\|_2^2}\\
&= 2\widetilde{\kappa}(f,x)^2,
\end{aligned}$$

and $\|f(x)\|_2 \le \sqrt{n}\,\|f(x)\|_\infty$. Therefore,

$$\min\left\{\mu_{\text{norm}}(f,x), \frac{\|f\|}{\|f(x)\|_\infty}\right\} \le \sqrt{2n}\,\widetilde{\kappa}(f,x).$$

This implies $\kappa(f) \le \sqrt{2n}\,\widetilde{\kappa}(f)$. □

Acknowledgment. In June 2008 we discussed condition and distance to ill-posedness with Mike Shub who suggested to us that a result in the spirit of Theorem 1.1 should be true. We are grateful to him since that conversation is at the origin of our paper. We also thank Carlos D'Andrea for the discussion on discriminants.

References

[1] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation.* Springer-Verlag, 1998.

[2] P. Bürgisser, F. Cucker, and M. Lotz. The probability that a slightly perturbed numerical analysis problem is difficult. *Mathematics of Computation*, 77:1559–1583, 2008.

[3] D. Cheung and F. Cucker. Probabilistic analysis of condition numbers for linear programming. *Journal of Optimization Theory and Applications*, 114:55–67, 2002.

[4] D. Cheung and F. Cucker. Solving linear programs with finite precision: I. Condition numbers and random programs. *Math. Program.*, 99:175–196, 2004.

[5] D. Cheung, F. Cucker, and R. Hauser. Tail decay and moment estimates of a condition number for random linear conic systems. *SIAM J. Optim.*, 15:1237–1261, 2005.

[6] D. Cox, J. Little, and D.O'Shea. *Using algebraic geometry.* Graduate Texts in Mathematics, 185. Springer-Verlag, New York, 1998.

[7] F. Cucker, R. Hauser, and M. Lotz. Adversarial smoothed analysis. Preprint, 2008.

[8] F. Cucker, T. Krick, G. Malajovich, and M. Wschebor. A numerical algorithm for zero counting. I: Complexity and accuracy. *J. Compl.*, 24:582–605, 2008.

[9] F. Cucker and M. Wschebor. On the expected condition number of linear programming problems. *Numer. Math.*, 94:419–478, 2003.

[10] C. D'Andrea, and G. Jeronimo. Rational Formulas for Traces in zero-dimensional Algebras. Preprint, 2008.

[11] J. Demmel. On condition numbers and the distance to the nearest ill-posed problem. *Numer. Math.*, 51:251–289, 1987.

[12] J. Demmel. The probability that a numerical analysis problem is difficult. *Math. Comp.*, 50:449–480, 1988.

[13] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. of Matrix Anal. and Applic.*, 9:543–556, 1988.

[14] A. Edelman. *Eigenvalues and Condition Numbers of Random Matrices.* Ph. D. Thesis, M.I.T., 1989.

[15] R. Hauser and T. Müller. Conditioning of random conic systems under a general family of input distributions. *Found. Comput. Math.*, 9:335–358, 2009.

[16] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants.* Mathematics: Theory & Applications. Birkhauser Boston, Inc., Boston, MA, 1994. x+523 pp.

[17] J.P. Jouanolou. Cours DEA, University of Strasbourg.

[18] G. Malajovich and J. M. Rojas. High probability analysis of the condition number of sparse polynomial systems. *Theoret. Comp. Sci.*, 315:524–555, 2004.

[19] J.R. Rice. A theory of condition. *SIAM J. Numer. Anal.*, 3:217–232, 1966.

[20] M. Shub and S. Smale. Complexity of Bézout's theorem I: geometric aspects. *Journal of the Amer. Math. Soc.*, 6:459–501, 1993.

[21] M. Shub and S. Smale. Complexity of Bézout's theorem II: volumes and probabilities. In F. Eyssette and A. Galligo, editors, *Computational Algebraic Geometry*, volume 109 of *Progress in Mathematics*, pages 267–285. Birkhäuser, 1993.

[22] D.A. Spielman and S.-H. Teng. Smoothed analysis of algorithms. In *Proceedings of the International Congress of Mathematicians*, volume I, pages 597–606, 2002.

[23] S. Smale. On the efficiency of algorithms of analysis. *Bulletin of the Amer. Math. Soc.*, 13:87–121, 1985.

[24] S. Smale. Complexity theory and numerical analysis. In A. Iserles, editor, *Acta Numerica*, pages 523–551. Cambridge University Press, 1997.

[25] M.J. Todd, L. Tunçel, and Y. Ye. Characterizations, bounds and probabilistic analysis of two complexity measures for linear programming problems. *Math. Program.*, 90:59–69, 2001.