A GENERIC WORST-CASE BOUND ON THE CONDITION NUMBER OF A HOMOTOPY PATH

GREGORIO MALAJOVICH

ABSTRACT. The number of steps of homotopy algorithms for solving systems of polynomials is usually bounded by the condition number of the homotopy path. A generic bound on the condition number of homotopy path between systems with integer coefficients will be given.

1. INTRODUCTION

In [6], it was proven that there is a Zariski closed set Σ' in the space of all systems of homogeneous polynomial equations of degree $\mathbf{d} = (\mathbf{d_1}, \dots, \mathbf{d_n})$ in $n+1$ variables, with the following property : For any f not in Σ', f with integer (resp. Gaussian integer) coefficients, the Shub and Smale condition number $\mu(f)$ of f satisfies :

$$\mu(f) \leq \mu(\Sigma') \ \mathbf{H}\left(\mathbf{f}\right)^{d(\Sigma')}$$

The numbers $\mu(\Sigma')$ and $d(\Sigma')$ depend only on n and $\mathbf{d}$, and :

$$\mathbf{H}\left(\mathbf{f}\right) = \max\left(\mathrm{Re}|f_{iJ}| + \mathrm{Im}|f_{iJ}|\right)$$

where f_{iJ} ranges over all the coefficients of f. For more details, see [6] and [8].

In this paper, a similar theorem is proven for the condition number of a linear homotopy path $\{f^{(t)}\} = \{(1-t)f^{(0)} + tf^{(1)}\}$. Here, t is a real parameter in $[0,1]$. The same homotopy path will be represented by the pair $(f^{(0)}, f^{(1)})$.

This bound will provide a *generic* worst case bound for the number of steps of a homotopy algorithm. See [4, 5, 8, 9, 10, 11, 12].

Let $\mathcal{H}_\mathbf{d}$ be the complex vector space of all systems of n homogeneous polynomial equations of degree $\mathbf{d}$ in $n+1$ variables. The notation $\mathbb{P}(\mathcal{H}_\mathbf{d})$

Date: November 29, 1995.
1991 *Mathematics Subject Classification.* 65H10, 65H20, 68Q25.
Key words and phrases. homotopy, conditioning, gap theorem.
Sponsored by CNPq (Brasil).
Typeset using $\mathcal{AMS}$-LaTeX.

will denote the projectivization of the complex vector space $\mathcal{H}_{\mathbf{d}}$. One may consider a path as a subset of $\mathbb{P}(\mathcal{H}_{\mathbf{d}})$. Its Zariski-closure is always a complex line (provided $f^{(0)} \neq f^{(1)}$). Generically speaking, it meets the discriminant variety $\Sigma \subset \mathbb{P}(\mathcal{H}_{\mathbf{d}})$. This is still true if one fixes one of the systems $f^{(0)}$ and $f^{(1)}$.

We may also represent the path $\{f^{(t)}\}$ by an element $(f^{(0)}, f^{(1)})$ of the space $\mathcal{H} = \mathcal{H}_{\mathbf{d}} \times \mathcal{H}_{\mathbf{d}}$. Once again, it makes sense to look at the Zariski closure of the set of paths meeting the discriminant variety Σ, as subsets of $\mathbb{P}(\mathcal{H}_{\mathbf{d}})$. Clearly, all non-constant paths are in this closure. Therefore, it makes no sense to look for a closed set in $\mathcal{H}$ to generalize Σ' of [6].

However, a generalization is possible if we consider the *real* vector space $\mathbb{R}(\mathcal{H}) = (\mathrm{Re}(\mathcal{H}), \mathrm{Im}(\mathcal{H}))$. This space is endowed with Zariski topology as a real vector space. Indeed, we will prove :

Main Theorem 1. *Let n and $\mathbf{d} = (\mathbf{d_1}, \ldots, \mathbf{d_n})$ be fixed. Let $\mathcal{H}$ be the complex vector space of all pairs $(f^{(0)}, f^{(1)})$ of polynomial systems of degree $\mathbf{d}$. Then there is a non-trivial Zariski closed set Σ'' in $\mathbb{R}(\mathcal{H})$ such that, for all $(f^{(0)}, f^{(1)})$ not in Σ'' and for all $t \in [0,1]$,*

$$\mu(f^{(t)}) \leq \mu(\Sigma'')\ \mathbf{H}\left((\mathbf{f^{(0)}}, \mathbf{f^{(1)}})\right)^{d(\Sigma'')}$$

where the numbers $\mu(\Sigma'')$ and $d(\Sigma'')$ depend only on d, and :

$$\mathbf{H}\left((\mathbf{f^{(0)}}, \mathbf{f^{(1)}})\right) = \max\left(\mathbf{H}\left(\mathbf{f^{(0)}}\right), \mathbf{H}\left(\mathbf{f^{(1)}}\right)\right)$$

Moreover, one can choose $d(\Sigma'') = 2n \prod \mathbf{d_j}\ \sum \mathbf{d_j}$

We will first construct the set Σ'' containing all the singular paths. Then, using a result in [6], we will bound the 'distance' between a path $\{f^{(t)}\} \not\in \Sigma''$ and Σ'', in terms of $\mathbf{H}\left((\mathbf{f^{(0)}}, \mathbf{f^{(1)}})\right)$. Finally, we will bound the condition number $\mu(\{f^{(t)}\})$ in terms of the inverse of the distance to Σ'' . A suitable distance may be introduced in the 'real projectivization' of $\mathbb{R}(\mathcal{H})$ by :

$$d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), (g^{(0)}, g^{(1)}))^2 = \frac{1}{2}\left(d_{\mathbb{RP}}(f^{(0)}, g^{(0)})^2 + d_{\mathbb{RP}}(f^{(1)}, g^{(1)})^2\right)$$

On the right hand side, $d_{\mathbb{RP}}(.,.)$ is the projective 2-distance :

$$d_{\mathbb{RP}}(f,g) = \min_{\lambda \in \mathbb{R}_*} \frac{\|f - \lambda g\|_{\mathrm{k}}}{\|f\|_{\mathrm{k}}}$$

This distance can also be interpreted as the sine of the (real) angle between f and g. The norm $\|.\|_{\mathrm{k}}$ denotes the $SU(n+1)$ invariant norm in $\mathcal{H}_{\mathbf{d}}$ (See [2, 8]).

This is similar to the usual projective distance :

$$d_{\mathbb{P}}(f,g) = \min_{\lambda \in \mathbb{C}_*} \frac{\|f - \lambda g\|_{\mathrm{k}}}{\|f\|_{\mathrm{k}}}$$

Clearly, $d_{\mathbb{P}}(f,g) \le d_{\mathbb{RP}}(f,g)$.

2. Breaking the algebraic structure

In order to construct the set Σ'', we will need somehow to 'break' the algebraic structure of the problem. The crucial step for this is the following, elementary fact :

Lemma 1. *Let $g \in \mathbb{C}[x]$. Let R denote the resultant of two degree $\deg g$ polynomials. Then g has a real factor of degree ≥ 1 if and only if $R(g, \bar{g}) = 0$.*

Proof. Suppose g has a real factor r. Then r has a real zero ζ, or a pair of conjugate zeros ζ and $\bar{\zeta}$. In both cases, ζ is a common zero of g and $\bar{g}$. Therefore the resultant $R(g, \bar{g})$ vanishes.

Conversely, suppose that $R(g, \bar{g}) = 0$. Then g and $\bar{g}$ have a common zero ζ. Furthermore, $g(\bar{\zeta}) = \overline{\bar{g}(\zeta)} = 0$, so $\bar{\zeta}$ is a zero of g, and the polynomial $(x-\zeta)(x-\bar{\zeta}) = x^2 - 2x\ \mathrm{Re}(\zeta) + |\zeta|^2$ divides g. □

We may now construct the polynomial $h(t) = R(f^{(t)} \det D' f^{(t)})$ where D' denotes the derivative with respect to $x_1, \ldots, x_n$, and where R denotes Macaulay's resultant [1, 3] of $n+1$ homogeneous polynomials in $n+1$ variables. R is a polynomial of degree $(\prod_{j \ne i} \mathbf{d_j})(\sum \mathbf{d_j} - \mathbf{n}) + \prod_{\mathbf{j}} \mathbf{d_j}$ in each set of 'variables' $f_j^{(t)}$. As a polynomial in t, it has degree bounded by $n \prod \mathbf{d_j} \sum \mathbf{d_j}$.

Vanishing of the resultant is a necessary and sufficient condition for $f^{(t)}$ and $\det D' f^{(t)}$ to have a common root in $\mathbb{P}^{\Bbbk}$. This common root may be a degenerate root of $f^{(t)}$ or a root of $f^{(t)}$ at 'infinity' $x_0 = 0$. Indeed, if $f^{(t)}(x) = 0$ and $D' f^{(t)}(x)$ is not surjective, we obtain :

$$0 = Df^{(t)}(x).x = x_0 \frac{\partial f^{(t)}}{\partial x_0} + D' f^{(t)}(x). \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

However, if $Df^{(t)}(x)$ is surjective, the columns of $D' f^{(t)}(x)$ cannot spawn $\frac{\partial f^{(t)}}{\partial x_0}$, hence $x_0 = 0$.

Clearly, if $f_t \in \Sigma$ for some real t, then h has a real factor. We now define the mapping :

$$\begin{array}{cccc} p: & \mathcal{H} & \rightarrow & \mathbb{C} \\ & (f^{(0)}, f^{(1)}) & \mapsto & R(h, \bar{h}) \end{array}$$

The mapping p defines a polynomial from $\mathbb{R}(\mathcal{H})$ into $\mathbb{R}^{\not k}$. If some $f^{(t)} \in \Sigma$, then $p(f^{(0)}, f^{(1)})$ vanishes. Let $\Sigma'' = Z(p)$.

Lemma 2. *The set Σ'' is a non-trivial closed set.*

We mean that p does not vanish uniformly on $\mathbb{R}(\mathcal{H})$.

Proof. Let $(f^{(0)}, f^{(1)})$ be generic, in the following sense : We require $f^{(0)}$ and $f^{(1)}$ to be non-degenerate, and to have no root at 'infinity' $x_0 = 0$. We also want $f^{(0)}$ and $f^{(1)}$ not colinear.

We will prove that for a 'generic' complex number λ (in a sense we will precise later), the path $(f^{(0)}, \lambda f^{(1)})$ is not in Σ''. Compare with Theorem 1 in [7].

Indeed, let $h^\lambda(t) = R(f^{(t)}, \det D' f^{(t)})$ where $f^{(t)} = (1-t) f^{(0)} + t\lambda f^{(1)}$.

The polynomial h^λ does not vanish uniformly in t, since $f^{(0)}$ has no degenerate solution, and no solution at infinity. Let D be the (maximal) degree of h^λ, as a polynomial in t.

Let $t_1, \dots, t_D$ be the roots of h^1. We will see that a 'generic' choice of λ will put $t_1, \dots, t_D$ in position $s_1, \dots, s_D$ such that $s_i \neq \bar{s}_j$ for all i, j, possibly $i = j$. Therefore, h^λ has no real factor in general, and $(f^{(0)}, \lambda f^{(1)})$ is not in Σ'' .

Indeed, for almost all λ, we may choose s_i such that :

$$(1 - t_i) f^{(0)} + t_i f^{(1)} = c_i \left((1 - s_i) f^{(0)} + s_i \lambda f^{(1)} \right)$$

where c_i is some complex number. If we do that, $h^\lambda(s_i) = h(t_i) {c_i}^D = 0$. We have to solve :

$$c_i = \frac{1 - s_i}{1 - t_i} = \frac{\lambda s_i}{t_i}$$

Recall that the genericity hypothesis in $(f^{(0)}, f^{(1)})$ prevents $t_i = 0$ or $t_i = 1$. We obtain :

$$s_i \lambda - s_i \lambda t_i = t_i - s_i t_i$$

Solutions are :

$$s_i = \frac{t_i}{\lambda - \lambda t_i + t_i}$$

Those s_i are finite for all $\lambda \neq \frac{-t_i}{1-t_i}$, all i. We still need to prove that for 'generic' λ, there are no i, j (possibly $i = j$) such that $s_i = \bar{s}_j$, or again : $\mathrm{Im}({s_i}^{-1} + {s_j}^{-1}) = 0$. (Recall that $s_i \neq 0$).

The situation to avoid is :

$$\mathrm{Im} \left(\frac{\lambda - \lambda t_i + t_i}{t_i} + \frac{\lambda - \lambda t_j + t_j}{t_j} \right) = 0$$

This is :

$$\operatorname{Im}\left(\frac{t_j - 2t_i t_j + t_i}{t_i t_j}\lambda + 2\right) = 0$$

Therefore, it suffices that λ avoids a finite set of points and real lines in complex plane. □

3. End of the proof

We are now under the hypotheses of Theorem 1 in [6] :

Theorem 1. *Let p be a multi-homogeneous polynomial of degree $r_1, \dots, r_n$ in sets of variables $f_1 \in \mathbb{C}^{m_1}, \dots, f_n \in \mathbb{C}^{m_n}$, with integer coefficients. Assume also that groups of variables f_i range over Gaussian integers. Then either $p(f) = 0$, or :*

$$d_{\mathbb{P}}(f, Z(p)) \geq \frac{1}{\frac{\pi}{2} \max \sqrt{m_i} \sum r_i \mathbf{B}(\mathbf{p})} \left(\frac{1}{\mathbf{H}(\mathbf{f})}\right)^{\sum r_i}$$

where $Z(p)$ is the zero-set of p and d is the complex projective 2-distance.

Here, the number $\mathbf{B}(\mathbf{p})$ depends only on p. We set $d(\Sigma'') = \sum r_i \leq 2n(\prod \mathbf{d_j})(\sum \mathbf{d_j} - \mathbf{n})$. We define $\mu(\Sigma'')$ as $\frac{\pi}{2} \max \sqrt{m_i} \sum r_i \mathbf{B}(\mathbf{p})$. Then, using $d_{\mathbb{RP}} \leq d_{\mathbb{RP}}$, we obtain a weaker version of the Main Theorem :

Theorem 2. *Let n and $\mathbf{d} = (\mathbf{d_1}, \dots, \mathbf{d_n})$ be fixed. Let $\mathcal{H}$ be the space of all pairs $(f^{(0)}, f^{(1)})$ of polynomial systems of degree $\mathbf{d}$. Then there is a non-trivial Zariski closed set Σ'' in $\mathbb{R}(\mathcal{H})$ such that, for all $(f^{(0)}, f^{(1)})$ not in Σ'' and for all $t \in [0, 1]$,*

$$\frac{1}{d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), \Sigma'')} \leq \mu(\Sigma'')\ \mathbf{H}\left((\mathbf{f^{(0)}}, \mathbf{f^{(1)}})\right)^{d(\Sigma'')}$$

where the numbers $\mu(\Sigma'')$ and (Σ'') depend only on d, and :

$$\mathbf{H}\left((\mathbf{f^{(0)}}, \mathbf{f^{(1)}})\right) = \max(\mathbf{H}\left(\mathbf{f^{(0)}}\right), \mathbf{H}\left(\mathbf{f^{(1)}}\right))$$

In order to conclude the proof of the Main Theorem, we will need the

Lemma 3.

$$\max_{t \in [0,1]} \mu(f^{(t)}) \leq \frac{1}{d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), \Sigma'')}$$

Since μ is real-scaling invariant, we may assume without loss of generality that $\left\|f^{(t)}\right\|_{\mathrm{k}} = 1$ always.

It was proven in [8] that for a given system f,

$$\mu(f) \leq \frac{1}{d_{\mathbb{P}}(f, \Sigma)}$$

The condition number of a homotopy path was defined by :

$$\mu(\{f^{(t)}\}) = \max_{t \in [0,1]} \mu(f^{(t)})$$

Hence :

$$\mu(\{f^{(t)}\}) \leq \max_{\substack{t \in [0,1] \\ g \in \Sigma}} \frac{1}{d_{\mathbb{P}}(f^{(t)}, g)} = \frac{1}{\min_{\substack{t \in [0,1] \\ g \in \mathbb{R}(\Sigma)}} d_{\mathbb{P}}(f^{(t)}, g)}$$

Suppose that this minimum was attained at some $t \in [0,1]$ and some $g \in \Sigma$:

$$d_{\mathbb{P}}(f^{(t)}, g) = \frac{\min_{\lambda \in \mathbb{C}_*} \left\| f^{(t)} - \lambda g \right\|_{\mathrm{k}}}{\| f^{(t)} \|_{\mathrm{k}}}$$

Since λg also belongs to Σ, we may scale g by λ so that :

$$d_{\mathbb{P}}(f^{(t)}, g) = \frac{\left\| f^{(t)} - g \right\|_{\mathrm{k}}}{\| f^{(t)} \|_{\mathrm{k}}} = d_{\mathbb{RP}}(f^{(t)}, g)$$

This shows that :

$$\mu(\{f^{(t)}\}) \leq \frac{1}{d_{\mathbb{RP}}(f^{(t)}, g)}$$

We may now define a new homotopy path $g^{(s)}$ that is, in some sense, the translation of $f^{(t)}$:

$$g^{(s)} = f^{(s)} + (g - f^{(t)})$$

With that definition :

$$d_{\mathbb{RP}}(\{(f^{(0)}, f^{(1)})\}, \{(g^{(0)}, g^{(1)})\})^2 = \frac{1}{2} \left(d_{\mathbb{RP}}(f^{(0)}, g^{(0)})^2 + d_{\mathbb{RP}}(f^{(1)}, g^{(1)})^2 \right)$$

But $d_{\mathbb{RP}}(f^{(0)}, g^{(0)}) \leq \frac{\|g - f^{(t)}\|_{\mathrm{k}}}{\|f^{(0)}\|_{\mathrm{k}}} = \left\| g - f^{(t)} \right\|_{\mathrm{k}}$, and similarly for $d_{\mathbb{RP}}(f^{(0)}, g^{(0)})$. Therefore :

$$d_{\mathbb{RP}}(\{(f^{(0)}, f^{(1)})\}, \{(g^{(0)}, g^{(1)})\})^2 \leq \left\| g - f^{(t)} \right\|_{\mathrm{k}}^2 = d_{\mathbb{RP}}(f^{(t)}, g)^2$$

Therefore,

$$\mu(\{f^{(t)}\}) \leq \frac{1}{d_{\mathbb{RP}}(f^{(t)}, g)} \leq \frac{1}{d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), (g^{(0)}, g^{(1)}))}$$

Moreover, since $(g^{(0)}, g^{(1)}) \in \Sigma''$,

$$d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), (g^{(0)}, g^{(1)})) \geq d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), \Sigma'')$$

Thus, we obtained :

$$\mu(\{f^{(t)}\}) \leq \frac{1}{d_{\mathbb{RP}}((f^{(0)}, f^{(1)}), \Sigma'')}$$

This proves the Lemma, and hence the Main Theorem.

References

[1] John Canny, *The complexity of robot motion planning*, MIT Press, Cambridge, Mass., 1988.

[2] Eric Kostlan, *Random polynomials and the statistical fundamental theorem of algebra*, Preprint, Univ. of Hawaii, 1987.

[3] F. S. Macaulay, Some formulæ in elimination, *Proceedings of the London Mathematical Society*, Vol XXXV, 1903.

[4] Gregorio Malajovich *On the complexity of path-following Newton algorithms for solving systems of polynomial equations with integer coefficients.* PhD Thesis, Berkeley, 1993. United Microfilms Inc, 300 North Zeeb Road, Ann Arbor, MI, 48106-1346 USA, Phone 800-521-0600.

[5] Gregorio Malajovich On generalized Newton algorithms : quadratic convergence, path-following and error analysis. *Theoretical Computer Science* **133** 65-84, 1994.

[6] Gregorio Malajovich *Worst possible condition number of polynomial systems.* Preprint, Rio de Janeiro, 1995.

[7] Alexander Morgan and Andrew Sommese, Computing all solutions to polynomial systems using homotopy continuation. *Applied Mathematics and Computation* **24**, 115-138, 1987.

[8] Michael Shub and Steve Smale, On the Complexity of Bezout's Theorem I - Geometric aspects. *Journal of the AMS*, **6**, 2, Apr 1993.

[9] Michael Shub and Steve Smale, On the complexity of Bezout's Theorem II - Volumes and Probabilities. in: F. Eysette and A. Galligo, eds : *Computational Algebraic geometry.* Progress in Mathematics **109**, Birkhauser, 267-285, 1993.

[10] Michael Shub and Steve Smale, Complexity of Bezout's Theorem III ; Condition number and packing. *Journal of Complexity* **9**, 4-14, 1993.

[11] Michael Shub and Steve Smale, Complexity of Bezout's Theorem IV ; Probability of success ; Extensions Preprint, Berkeley, 1993.

[12] Michael Shub and Steve Smale, Complexity of Bezout's Theorem V : Polynomial time ; Preprint, Barcelona, 1993.

Departamento de Matematica Aplicada da UFRJ, Caixa Postal 68530, CEP 21945, Rio, RJ, Brasil

E-mail address: gregorio@lyric.labma.ufrj.br