EFFICIENT POLYNOMIAL SYSTEM SOLVING BY NUMERICAL METHODS

C. BELTRÁN AND L.M. PARDO

Dedicated to Steve Smale on his 80th Birthday.

Abstract. These pages contain a short overview on the state of the art of efficient Numerical Analysis methods that solve systems of multi–variate polynomial equations. We focus on the work of Steve Smale who initiated this research framework, and on the collaboration between Steve Smale and Mike Shub, which set the foundations of this approach to polynomial system–solving.

1. Introduction

In 1981, a manuscript by Steve Smale initiated a research framework: the design and analysis of efficient polynomial equation solvers by numerical methods, a cornerstone of the Foundations of Numerical Analysis. In the mid–eighties and early nineties, a close collaboration between Steve Smale and Mike Shub established the foundations of this new framework. Key to their work was the development of a new Model of Computation (with algebraic alphabets, allowing infinite alphabet cases) that emerged from a collaboration with Lenore Blum. It was known as the BSS machine model.

The influence of Smale's paper [Sma81] has been especially remarkable in the modern treatment of polynomial root–finding and polynomial system–solving. Smale pointed to Numerical Analysis tractable algorithms, i.e. algorithms whose "running time" (in terms of their BSS machine model) is polynomial in the input length. The 17^{th} problem in his list of problems [Sma00] is

Problem. *Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?*

This problem belongs to a long tradition of mathematical questions relating efficiency to solving, and whose origins may go back to the *"Rhind Mathematical Papyrus"* or the *"Sulba Sutras"* on tractable methods to compute square roots. In more recent times, the problem goes back to I. Newton and his *"De analysi per aequationes numero terminorum infinitas,"* and to the comment in E. Galois' last manuscript where this young mathematician claimed *"En un mot, les calculs sont impracticables."* Galois was probably the first mathematician to recognize the modern concept of Computational

Date: April 30, 2009.

Key words and phrases. Complexity, polynomial equation–solving, Newton's method, homotopy methods.

Research was Partially Supported by MTM2007–62799.

Complexity as the amount of effort required to solve mathematical problems: after describing a method to compute symbolic descriptions of the solutions of univariate polynomial equations, he noticed that it was feasible, but required too much effort in practice.

The problem of solving multivariate polynomial equations underlies most of the developments of 19^{th} and early 20^{th} centuries on foundations of Algebraic Geometry, also called Elimination Theory. See for instance, the works by L. Kronecker [Kr1882], D. Hilbert [Hi1893] or the version of G. Hermann of Hilbert's Nullstellensatz [Her26]. Traces of this combination of efficiency and multi–variate polynomial equation–solving can also be found in classical texts like [Mac16] or [Kö1903]. However, during the first half of the past century, the problem seemed to fall into oblivion. Since then, it has regained its previous importance by being one of the central questions of Symbolic Computation. A description of this symbolic approach is beyond the scope of these pages; the reader may follow wide bibliographic collections in [BeWe93], [BePa06], [GiHe01], [Kri04], [DuLe08], [Mor05], [Par95], [MiSt05], for instance.

We want to make a few comments on the term efficiency (sometimes referred to as tractability). Computational Complexity studies the design and analysis of algorithms that are tractable in practice. Even after the appearance of the first modern computers and powerful processors, certain problems have resisted a computational treatment. Problems are classified as unsolvable when there is no algorithm that solves them. This is the case of Gödel's Undecidablity Theorem, Turing's Halting Problem or Robinson–Matiyasevich's negative answer to Hilbert's Tenth Problem. Some other problems are intractable in practice: they are algorithmically solvable, but they require too much computational resources (either in terms of running time or memory/space). Tractable or efficient problems are those algorithmically solvable problems whose resource requirements grow polynomially with the size of the input. In these cases, we say that the associated algorithm runs on polynomial time, or is a polynomial time algorithm. Not only are polynomial time algorithms considered efficient, but also their probabilistic versions, including average polynomial time algorithms or bounded error probability polynomial time algorithms. In order to analyze these Theoretical Computation problems in the context of Numerical Analysis – rather than in a Turing Machine discrete framework – Steve Smale and his collaborators suggested a "continuous" model of computation in [BSS89]. See [BCSS98] or the more recent [Blm04] for a more detailed reference list. Note that the model is a uniform version of the classical models of Algebraic Complexity Theory.

This manuscript focuses only on the influence played by Steve Smale's kick-off on the Foundations of Numerical Analysis through the study of efficient polynomial equation–solving.

2. Approximate Zero Theory and the Univariate Case

In [Sma81], Steve motivated the problem as an interdisciplinary subject. In his own words:

"A second goal [for studying the tractability of Newton's method] *is to give the background of the various areas of mathematics, pure and applied, which motivate and give the environment for our problem. These areas are parts of (a) Algebra, the "Fundamental theorem of algebra", (b) Numerical Analysis, (c) Economic equilibrium theory and (d) Complexity Theory of computer science"*

Many of the future developments around the topic of equation–solving are already pointed out in [Sma81]: approximate zero theory, probabilistic behavior, liaisons with Integral Geometry... Steve recalls Gauss' first"proof" of the Fundamental Theorem of Algebra as one of his inspirations.

One of the main outcomes of that work was an algorithm (tractable with high probability) for solving univariate complex polynomials f using a "modified Newton operator" $T_h(z) := z - hf(z)/f'(z)$. One innovation of his approach is, for example, that tractability (i.e. running time) is analyzed in probabilistic terms, thus yielding an "almost always very fast" algorithm that solves "almost every" instance problem.

Theorem 1 ([Sma81])**.** *There is a universal polynomial $S(d, 1/\mu)$ and a function $h = h(d, \mu)$ such that for degree d and $0 < \mu < 1$, the following is true with probability $1 - \mu$. Let $x_0 = 0$. Then $x_n = T_h(x_{n-1})$ is well–defined for all $n > 0$ and x_s is an approximate zero for f where $s = S(d, 1/\mu)$.*

More specifically, Smale proved that if $s \geq [100(d + 2)]^9/\mu^7$, then with probability $1-\mu$, x_s is well–defined for suitable h, and x_s is an approximate zero of f.

The scheme suggested by Smale introduces a tradeoff between probability of error and computational complexity. This idea is also useful in studying the complexity of other problems like linear equation solving or linear programming (cf. [Sma85]).

Theorem 1 bounds the probability that certain algorithm fails; now one can try to analyze the "probability of failure" for any Numerical Analysis algorithm. A second, and sometimes more difficult, question is the average behavior of such algorithms. This question underlies many of the Problems in the list stated at the end of [Sma81].

In the forthcoming pages, we will recall many results which are consequences of a fruitful collaboration between Steve Smale and Mike Shub from the mid–eighties to the early nineties. In the univariate case, this collaboration is explicit in [ShSm85, ShSm86] which explored the average complexity of algorithms based on the so–called Global Newton Method ([Sma76, HiSm79]). The main outcome of these two manuscripts can be stated as follows.

Let $P_d(1)$ be the class of all univariate, monic, complex polynomials whose coefficients have absolute value bounded by 1. Namely,

$$P_d(1) = \{f = \sum_{i=0}^{d} a_i X^i : a_d = 1, |a_i| < 1 \; for \; 0 \leq i \leq d-1\}.$$

Let $N_f : z \mapsto z - \frac{f(z)}{f'(z)}$ be Newton's operator. An approximate zero of f is some point $z_0 \in \mathbb{C}^n$ such that the successive iterations of N_f, $z_k = N_f^k(z_0)$

satisfy

$$|\zeta - z_k| \leq \frac{1}{2^{2^k-1}}|\zeta - z_0|,$$

for some actual zero $\zeta \in \mathbb{C}$ of f. Namely, z_0 is in the strong basin of attraction of Newton's operator and converges quickly to an exact zero of f.

Let E be the Euler–Newton iteration scheme (cf. [ShSm85]) and define the following algorithm. For each $f \in P_d(1)$, we consider

$$\varepsilon_f := \frac{1}{(2d)^{4d}} \mid D_f \mid \geq 0,$$

where D_f is the discriminant of f. Let

$$n := \lfloor K(d + \mid \log \varepsilon_f \mid) \rfloor + 1,$$

where K is some universal constant and $\lfloor \cdot \rfloor$ means integer part.

Algorithm (N-E) *Let $f \in P_d(1)$, satisfy $\varepsilon_f > 0$.*

(1) Set $m = 1$.
(2) Choose $z_0 \in \mathbb{C}$, $|z_0| = 3$ at random and set $z_n := E^n(z_0)$. If $|f(z_n)| < \varepsilon_f$ terminate and print: " z_n is an approximate zero".
(3) Otherwise let $m = m + 1$ and go to (2).

Here m is just counting the number of iterations of the algorithm.

Theorem 2 ([ShSm86])**.** *Algorithm (N-E) terminates (and hence produces an approximate zero) with probability* 1 *and the average number of iterations is less than $K_1 d \log d$ where K_1 is a universal constant.*

In the first half of the eighties, there were other studies on the tractability of the problem of numerically–solving univariate polynomial equations. It was a period of great activity on the subject. We may cite, among others, [Sch81, Sch86], [Ren85, Ren87],[Kim85, Kim88], [Pan87] which also stated tractability results of numerical methods that deal with univariate polynomial equation–solving from different perspectives and different models of complexity. We also mention a more recent paper [HSS01], where a different approach is proposed, and it is proved that there is a universal set of initial complex numbers which converge to every solution of normalized polynomials.

At this stage, Problem 17 may be re–stated as follows.

Problem (Problem 17)**.** *Generalize Theorem 2 to the multi–variate case.*

3. Approximate Zero Theory and Multi–variate Case

The multi–variate case is the central problem in Computational Algebraic Geometry. For this problem, the collaboration between Mike Shub and Steve Smale produced important results, which have inspired many other authors afterwards. Some features of their work are:

- The problem is treated from a Numerical Analysis point of view.
- Probability and Approximation are an essential part of the context, so involving Integral Geometry is a must.

- New point estimates are needed. This need leads to the study of the α, β, γ quantities and theorems, and the normalized condition number μ_{norm}. The first three quantities continue Kantorovich's modern treatment of Newton's method.
- A geometric–iterative algorithmic scheme (path–following methods, homotopy) is used to compute approximate zeros.
- A strong incidence with the continuous version of Cook's conjecture appears (Hilbert's Nullstellensatz as $\mathbf{NP}_{\mathbb{C}}$–complete problem).

An extensive account of these ideas may be found in [BCSS98]. Here we just briefly discuss some of them.

In the affine multi–variate case, we deal with multi–variate polynomial mappings:

$$f = (f_1, \ldots, f_n) : \mathbb{C}^n \longrightarrow \mathbb{C}^n,$$

where f_i is a polynomial of degree d_i. Let (d) be the list of degrees and $\mathcal{P}_{(d)}$ the vector space of such f's. We denote $d = \max\{d_i : 1 \leq i \leq n\}$. The solution set is the algebraic variety $V(f)$ defined as the fiber at $0 \in \mathbb{C}^n$ of f. Namely, $V(f) := f^{-1}(\{0\})$. Newton's multi–variate operator N_f is

$$N_f(z) = z - Df(z)^{-1} f(z),$$

where $Df(z)$ is the jacobian matrix.

An (affine) approximate zero of $f \in \mathcal{P}_{(d)}$ with associated zero $\zeta \in V(f)$ is a point $z \in \mathbb{C}^n$ satisfying:

$$||\zeta - N_f^k(z)|| \leq \frac{1}{2^{2^{k-1}}} ||\zeta - z||.$$

We recall here the main definitions and results of Smale's α–Theory in the affine case. We encourage the reader to see the original papers [Sma86, Sma87].

Let $f : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ be analytic, and $z \in \mathbb{C}^n$. Let $D^{(k)} f(z)$ be the k–th derivative of f at z, seen as a multi–linear map. Then, define the following quantities:

$$\gamma(f, z) = \sup_{k>1} \left\| Df(z)^{-1} \frac{D^{(k)} f(z)}{k!} \right\|_2^{\frac{1}{k-1}},$$

$$\beta(f, z) := \|Df(z)^{-1} f(z)\|_2,$$

and

$$\alpha(f, z) = \beta(f, z)\gamma(f, z).$$

Here, $\|\cdot\|_2$ holds for operator norm (or Euclidean norm, if applied to vectors). The following result shows a bound for the radius of a ball where safe and fast convergence of Newton's operator is guaranteed.

Theorem 3 ($\gamma-$Theorem, [Sma86])**.** *Let $\zeta \in \mathbb{C}^n$ be such that $f(\zeta) = 0$ and let $z \in \mathbb{C}^n$ satisfy*

$$\|z - \zeta\| < \frac{3 - \sqrt{7}}{2\,\gamma(f, \zeta)}.$$

Then, z is an approximate zero of f with associated zero ζ.

Note that this statement yields a sufficient condition but not a checkable test: deciding whether z is close enough to some zero ζ or not requires an "a priori" knowledge of the zero ζ. But ζ is the quantity we want to approximate, thus yielding a paradox. Smale avoided this paradox with the following statement.

Theorem 4 ($\alpha-$Theorem, [Sma86])**.** *There is a universal constant α_0 (approximately equal to* 0.130707*) such that the following holds for every $z \in \mathbb{C}^n$: If $\alpha(f,z) < \alpha_0$, then z is an approximate zero of f.*

Indeed, Smale proved this result with greater generality, valid for analytic maps between Banach spaces. We do not treat this general framework here.

Theorems 3 and 4 and their consequences are among the most important results known about Newton's Method. They have been generalized to the underdetermined and overdetermined cases (i.e. $f : \mathbb{C}^n \to \mathbb{C}^m$ with $m \neq n$) [ShSm96, Ded01], analysis on manifolds [Shb93, DPM03, DeSh00], the singular case [DeSh01, GLSY05, GLSY07], diophantine aspects [Mal00, CHMP01, CMSP02, CSP03], error analysis [Mal94] and even to other operators and methods. An exhaustive bibliography would be too extensive for the narrow margins of this manuscript.

Kantorovich's Theory, which also gives convergence criteria for Newton's operator, is a natural precedent to Theorem 3. The criteria for Kantorovich's result involve bounds for the second derivative in a ball containing the zero. The power of Smale's α–theory is that it allows us to decide convergence of the sequence $z_0, z_1, z_2, \ldots$ knowing only the function and its derivatives at z_0.

M. Shub extended the definition of Newton's method to be valid in a homogeneous setting in [Shb93]. Let

$$f = (f_1, \ldots, f_n) : \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^n$$

where f_i is a homogeneous polynomial of degree d_i. Let (d) be the list of degrees and $\mathcal{H}_{(d)}$ the vector space of such f's. Note that $\mathcal{P}_{(d)}$ and $\mathcal{H}_{(d)}$ can be identified as every element of $\mathcal{H}_{(d)}$ can be seen as the homogenization of some $f \in \mathcal{P}_{(d)}$. Note also that if $\zeta \in \mathbb{C}^{n+1}$ is a zero of $f \in \mathcal{H}_{(d)}$ then every point in the complex line $\{\lambda\zeta : \lambda \in \mathbb{C}\}$ is also a zero of f. Thus, it is natural to consider zeros of f as projective points $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$. We will denote by $V_{\mathbb{P}}(f) \subseteq \mathbb{P}(\mathbb{C}^{n+1})$ the (non–empty) projective algebraic variety defined by f. Shub's projective version of Newton's operator is then defined as

$$N_f^{\mathbb{P}}(z) = z - (T_z f)^{-1} f(z), \quad z \in \mathbb{P}(\mathbb{C}^{n+1}),$$

where $T_z f = Df(z) \mid_{z^\perp}$ is the restriction of the differential matrix to the orthogonal complement of z. A projective approximate zero of f is then a point $z \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the successive iterations of $N_f^{\mathbb{P}}$ with initial point z are well–defined and converge quadratically to an actual projective zero $\zeta \in V_{\mathbb{P}}(f)$.

In [ShSm93a], Shub and Smale generalized theorems 3 and 4 to this homogeneous/projective context. Moreover, they transformed point estimates and local convergence results like Theorem 3 into estimates for the complexity of path–following methods. This was achieved by introducing a new

quantity, the (normalized) condition number for polynomial system–solving: Given $f \in \mathcal{H}_{(d)}$ and $z \in \mathbb{P}(\mathbb{C}^{n+1})$, let

$$\mu_{\text{norm}}(f,z) = \|f\|_\Delta \|(T_z f)^{-1} Diag(\|z\|^{d_i-1} d_i^{1/2})\|_2, \tag{3.1}$$

and $\mu_{\text{norm}}(f,z) = +\infty$ if $T_z f$ is not onto. Here, $\|f\|_\Delta$ is the Bombieri–Weyl–Kostlan unitarily invariant norm in $\mathcal{H}_{(d)}$. See [ShSm93a] for a detailed description.

As in Eckardt and Young's Theorem [EcYo36] for the linear case, this condition number is related to the inverse of the distance of some "singular locus". More exactly,

Theorem 5 (Condition Number Theorem,[ShSm93a])**.**

$$\mu_{\text{norm}}(f,\zeta) = \frac{1}{\sin\left(d_R(f,\Sigma_\zeta)\right)},$$

where d_R is the Riemannian distance in $\mathbb{P}(\mathcal{H}_{(d)})$ and

$$\Sigma_\zeta = \{f \in \mathbb{P}(\mathcal{H}_{(d)}) \ : \ f(\zeta) = 0, \text{ and } T_\zeta f \text{ is not an onto mapping}\}.$$

Shub and Smale used the condition number $\mu_{\text{norm}}(f,z)$ to analyze the behavior of a homotopy method to solve systems $f \in \mathcal{H}_{(d)}$. They first proved a statement similar to Theorem 3:

Theorem 6 (μ–Theorem, [ShSm93a])**.** *Let $\zeta \in V_{\mathbb{P}}(f)$ and let $z \in \mathbb{P}(\mathbb{C}^{n+1})$ satisfy*

$$\tan(d_R(z-\zeta)) < \frac{3-\sqrt{7}}{2\, d^{3/2}\, \mu_{\text{norm}}(f,\zeta)}.$$

where d_R is the Riemannian distance in $\mathbb{P}(\mathbb{C}^{n+1})$. Then, z is an approximate zero of f with associated zero ζ.

Consider the solution variety $V = \{(f,\zeta) \in \mathcal{H}_{(d)} \times \mathbb{P}(\mathbb{C}^{n+1}) : f(\zeta) = 0\}$, and the two canonical projections $\pi_1 : V \to \mathcal{H}_{(d)}$ and $\pi_2 : V \longrightarrow \mathbb{P}(\mathbb{C}^{n+1})$.

Let $\Sigma' \subseteq V$ be the set of all critical points of π_1 and $\Sigma := \pi_1(\Sigma')$ the set of critical values. Note that Σ is the classical discriminant variety of Elimination Theory, and

$$\Sigma' = \{(f,\zeta) \in V : \mu(f,\zeta) = \infty\}.$$

In particular, both Σ and Σ' are complex codimension 1 algebraic subvarieties of their respective ambient spaces.

Let $f \in \mathcal{H}_{(d)}$ be a target system to be solved, and let $g \in \mathcal{H}_{(d)}$ be another system that has a known solution $\zeta_0 \in \mathbb{P}_n(\mathbb{C})$. Consider some piecewise $\mathcal{C}^1$ curve $C := \{f_t : t \in [0,1]\}$ joining g and f, so that $f_0 = g$, $f_1 = f$.

Under some regularity hypothesis ($C \cap \Sigma = \emptyset$ suffices), π_1 defines a $\mathcal{D}-$fold covering map $\pi_1 : \pi_1^{-1}(C) \longrightarrow C$ and the curve C can be lifted to a differentiable curve $\{(f_t,\zeta_t), t \in [0,1]\} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$, with initial vertex (g,ζ_0) and such that $f_t(\zeta_t) = 0$, $\forall t \in [0,1]$. This curve will be denoted by $\Gamma(f,g,\zeta_0)$.

The homotopy method constructs a polygonal path that closely follows $\Gamma(f,g,\zeta_0)$. This path has initial vertex (g,ζ_0) and final vertex (f,z), for some $z \in \mathbb{P}_n(\mathbb{C})$, which is the output of the algorithm. The polygonal path is constructed by homotopy steps, each of which is an application of the

projective Newton operator, with an appropriate step size selection. Hence, the method constructs a finite partition $0 = t_0, t_1, \ldots, t_k = 1$ and defines

$$z_0 = \zeta_0, \quad z_{i+1} = N^{\mathbb{P}}_{f_{t_{i+1}}}(z_i). \quad \text{Output} : z_k.$$

The choice of the steps t_i has to be done in a sensible fashion, guaranteeing that every z_i is in the strong basin of attraction of the next system $f_{t_{i+1}}$. This is possible thanks to Theorem 6 above. One can see that high values of μ_{norm} will lead to a small radius for the strong basin of attraction ball, forcing us to choose smaller steps t_i and thus slowing down the process. Reciprocally, small values of μ_{norm} will increase the size of those balls, allowing us to choose greater t_i and thus speeding the algorithm up.

The main result of [ShSm93a] is a qualitative version of this last idea: it provides a bound for the number of steps k needed, in order to guarantee that z_k is an approximate zero of f.

Theorem 7 ([ShSm93a])**.** *The number $NS(f, g, \zeta_0) = k$ of (projective) Newton steps necessary to guarantee that z_k is an approximate zero of f satisfies*

$$NS(f, g, \zeta_0) \leq Cd^{3/2} \max_{t\in[0,1]} \mu_{\text{norm}}(f_t, \zeta_t)^2,$$

$C > 0$ a universal constant.

That is, the number of Newton steps necessary to follow homotopy paths depends mainly on the condition number along the path! This result has been recently improved in [Shu09].

Theorem 8 ([Shu09])**.** *The number $NS(f, g, \zeta_0) = k$ of (projective) Newton steps necessary to guarantee that z_k is an approximate zero of f satisfies*

$$NS(f, g, \zeta_0) \leq Cd^{3/2} \int_0^1 \mu_{\text{norm}}(f_t, \zeta_t) \|(\dot{f}_t, \dot{\zeta}_t)\|_2 \, dt,$$

$C > 0$ a universal constant. Thus, $NS(f, g, \zeta_0)$ is bounded by $Cd^{3/2}$ times the length of the path $(f_t, \zeta_t) \subseteq V$ in the condition metric, obtained by multiplying the canonical Riemannian (product) metric by the condition number μ at each point. The following bound also holds,

$$NS(f, g, \zeta_0) \leq C_2 d^{3/2} \int_0^1 \mu(f_t, \zeta_t)^2 \, dt.$$

Shub and Smale were interested in the average behavior of the method, so they described the probability distribution of μ_{norm} in [ShSm93b]. Let $\mathbb{S}$ be the sphere of radius one in $\mathcal{H}_{(d)}$ with respect to the Bombieri-Weyl Hermitian norm. Let $\mathbb{S}$ be endowed with the inherited Riemmanian structure. As the volume of $\mathbb{S}$ is finite we may introduce a probability distribution that we simply denote by P.

Theorem 9 (cf [ShSm93b])**.** *With these notations*

$$\mathrm{P}\left[f \in \mathbb{S} \, : \, \exists \zeta \in V_{\mathbb{P}}(f), \ \mu_{\text{norm}}(f, \zeta) > \varepsilon^{-1}\right] \leq \frac{1}{4} n^3 (n+1) N (N-1) \mathcal{D} \varepsilon^4,$$

where $N + 1$ is the complex dimension of $\mathcal{H}_{(d)}$.

There are several conceivable ways to choose the path to be lifted f_t. One may prove the existence of curves (f_t, ζ_t) in $V \setminus \Sigma'$ that have small length in the condition metric, and then look for some strategy to explicitly describe the projection on the first coordinate f_t. This is the idea behind Theorem 17 below. However, until now there is not a practical way to describe these "a priori" short curves. For example, explicitly describing the short curves of Theorem 17 requires the knowledge of the target solution ζ of f. We refer the reader to Subsection 5.1 for details.

An easier method is to construct a path f_t in the space of systems $\mathbb{S}$ that avoids the discriminant variety Σ. The existence of lifting curves then suffices to apply the homotopy method described above, and thus approximate the lifting Γ. Sometimes, one can be sure "a priori" that the path f_t will not intersect Σ. Otherwise, one can just use a dimensional argument: Σ has complex codimension 1, thus real codimension 2, and hence most "reasonable" curves f_t will not intersect it.

This philosophy has been used in all path–following methods (also called homotopy continuation methods or simply homotopy methods). See [GaZa79], [Li83, Li87], [Mor87], [SoWa05], [Ver96] for a complementary list of historical references.

Shub and Smale centered their attention on the most simple choice of paths f_t: linear paths, i.e. great circles in $\mathbb{S}$. More specifically, let $f, g \in \mathbb{S}$, $f \neq \pm g$. Then, one can choose the f_t to be the (short) portion of the great circle between g and f. The use of Integral Geometry allowed them to transform probability results like Theorem 9 into much more sophisticated results, like the following one.

Theorem 10 ([ShSm94])**.** *There exists a initial pair $(g, \zeta_0) \in V$ such that, for random $f \in \mathcal{H}_{(d)}$ such that $\|f\| = 1$, the average number of Newton steps necessary to follow the linear homotopy and thus produce an approximate zero of f, is at most $CN^3 \log \mathcal{D}$, C a universal constant. Namely, for some $(g, \zeta_0) \in V$,*

$$\mathrm{E}_{f \in \mathbb{S}}(NS(f, g, \zeta_0)) \leq CN^3 \log \mathcal{D}.$$

(if some $d_i = 1$ or $n \leq 4$ this quantity must be replaced by $CN^4 \log \mathcal{D}$.)

Theorem 10 suggests for the first time that an algorithm may exist to produce approximate zeroes of random systems in average polynomial time. However, [ShSm94] does not show how the pair (g, ζ_0) of Theorem 10 can be constructed, so a practical algorithm cannot be deduced immediately. At this stage, Problem 17 could have the following form:

Problem (Problem 17)**.** *Find an explicit (i.e. algorithmically constructive) description of the initial pair (g_0, ζ_0) that satisfies the thesis of Theorem 10 above.*

In [ShSm94], the following pair was conjectured to satisfy the conditions of Theorem 10 above, and this is still an open question.

$$g(z) = \begin{cases} d_1^{1/2} z_0^{d_1 - 1} z_1, \\ \quad \vdots \\ d_n^{1/2} z_0^{d_n - 1} z_n \end{cases}, \qquad \zeta_0 = (1, 0, \ldots, 0).$$

The pair conjectured in [ShSm94] does not contain these $d_i^{1/2}$ factors. There is, however, some consensus that these extra factors should be added. With these extra factors, this pair satisfies that $\mu_{\mathrm{norm}}(g, e_0) = n^{1/2}$ is minimal.

4. Attacks based on Average Las Vegas algorithms

After a frenetic activity in the first half of the nineties, there were years with no progress on the search for an initial pair. This is the context where Smale proposed Problem 17. Ten years later, the authors of this manuscript tried an alternative approach to Problem 17: since finding the initial point (g_0, ζ_0) of Theorem 10 seems to be hard, let's think probabilistically!

Probabilistic algorithms have been extensively used in Computational Mathematics since the first famous examples in Primality testing (cf. [SoSt77], [Mi76], [Rab80], for instance). Even after the appearance of deterministic primality testings ([AKS04]), probabilistic procedures are still used in practice. The reasons are multiple. On one hand, running probabilistic procedures on a computer does not differ perceptively from running deterministic ones. On the other hand, probabilistic procedures are often faster. The main outcomes of [Sma81] and [ShSm85] (i.e. Theorems 1 and 2 above) describe probabilistic algorithms for the univariate case.

In the case of multi–variate polynomial equation–solving, having a uniform, probabilistic, efficient algorithmic procedure that runs in average polynomial time yields a positive answer to an old question and may, perhaps, open the way to efficient deterministic algorithms that solve the problem.

In the series of papers [BePa08, BePa09a, BePa09b, BePa09c], the authors of this manuscript demonstrated a uniform, probabilistic, efficient algorithm that solves systems of multi–variate polynomial equations both for affine or projective solutions,with a running time polynomial on the size of the input, on the average. That algorithm is a solution to Smale's 17^{th} problem, with a probabilistic component. This section contains a short account of the main results of these works.

4.1. Generalizing Theorem 1 of [Sma81].

We first describe a multi–variate version of the main statement in [Sma81] (Theorem 1 above). Global Newton method is replaced by homotopy continuation. As we have said, the main problem was to find a good set of initial pairs.

A Homotopic Deformation scheme (HD for short) with initial pair $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ and resource function $\varphi : \mathcal{H}_{(d)} \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is an algorithmic scheme based on the following strategy:

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

Perform $\varphi(f, \varepsilon)$ homotopy steps following the segment $(1-t)g+tf$, $t \in [0, 1]$, starting at (g, z_0), where z_0 is an approximate zero of g associated with some zero $\zeta_0 \in V(g)$.

Output:

either failure, or
an approximate zero $z_1 \in \mathbb{P}_n(\mathbb{C})$ of f.

Unless otherwise specified, the homotopy steps are chosen to be equally spaced in the segment $(1-t)g + tf$.

We say that an initial pair $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ is ε-efficient for HD if the HD scheme with initial pair (g, z_0) and resource function

$$\varphi(f, \varepsilon) = 18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2}, \ \ \forall f \in \mathcal{H}_{(d)}, \ \varepsilon > 0,$$

satisfies the following property:
" For a randomly chosen system $f \in \mathcal{H}_{(d)}$, the probability that HD outputs an approximate zero of f is at least $1 - \varepsilon$".

A set $\mathcal{G} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ is called a correct test set (also questor set) for efficient initial pairs if for every $\varepsilon > 0$ the probability that a randomly chosen pair $(g, \zeta) \in \mathcal{G}$ is ε-efficient is greater than

$$1 - \varepsilon.$$

Main Theorem (Weak Version) of [ShSm94] could be read in this context as: "There is a questor set $\mathcal{G}$ for efficient initial pairs with a single element (i.e. $\sharp(\mathcal{G}) = 1$)". Hence, the problem becomes to find such a singleton which is a questor set.

We didn't succeed in solving this "singleton" case, but in [BePa08] we exhibited a subset $\mathcal{G}_{(d)} \subseteq V$ which satisfies the following two properties.

- It is a questor set for efficient initial pairs.
- It is constructible and its elements may be used in a probabilistic version of the HD scheme.

Thus, we can consider the following HD scheme:

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Guess at random* $(g, \zeta) \in \mathcal{G}_{(d)}$.
- *Perform* $\varphi(f, \varepsilon) = 18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2}$ *homotopy steps following the segment* $(1 - t)g + tf$, $t \in [0, 1]$, *starting at* (g, ζ).

Output:

either failure, or
an approximate zero $z \in \mathbb{P}_n(\mathbb{C})$ *of* f.

The following statement generalizes Theorem 1 above to the multi–variate case.

Theorem 11 ([BePa08])**.** *The set* $\mathcal{G}_{(d)}$ *is a questor set for efficient initial pairs in* $\mathcal{H}_{(d)}$. *Moreover, for every positive real number* $\varepsilon > 0$, *the probability that a randomly chosen pair* $(g, e_0) \in \mathcal{G}_{(d)}$ *is* ε*-efficient is greater than*

$$1 - \varepsilon.$$

In particular, for these ε*-efficient pairs* $(g, e_0) \in \mathcal{G}_{(d)}$, *the probability that a randomly chosen input* $f \in \mathcal{H}_{(d)}$ *is solved by HD with initial pair* (g, e_0) *performing* $O(n^5 N^2 d^3 \varepsilon^{-2})$ *homotopy steps is at least*

$$1 - \varepsilon.$$

A less technical but more comprehensive version of this statement is the following one (just replacing ε by $\frac{1}{N}$):

Corollary 1. *There is a Bounded Error Probability* (**BPP**) *Numerical Analysis procedure that solves most systems of multivariate homogeneous polynomial equations with running time polynomial in*

$$n, N, d.$$

In fact, the probability that a randomly chosen system $f \in \mathcal{H}_{(d)}$ is solved by this procedure is greater than

$$1 - \frac{1}{N}.$$

4.2. **Generalizing Theorem 2 of** [ShSm86]**: The average complexity case.** Problem 17 asks about average complexity and the main outcome of [BePa08] does not immediately yield an average complexity estimate. This was achieved for the first time in [BePa09a]. The main innovation was to adapt the inclusion/exclusion homotopy method described in Section 6 of [ShSm94] to the probabilistic algorithm described in Subsection 4.1 above.

Note that the algorithm of Subsection 4.1 considers partitions of $18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2}$ subintervals of the great circles (or segments). Namely, the number of subintervals of the partition is fixed by the input. As we have seen, this suffices to prove that most systems are solved for random choice of $(g, \zeta) \in \mathcal{G}_{(d)}$. In [ShSm94] a path–following method is described that has no "a priori" number of steps. Namely, given a path f_t it creates a partition adapted to that path, and when the method finishes, it always returns an approximate zero of the target system f. As a (necessary) drawback, there are some bad choices of paths f_t for which the method never ends. These bad choices are precisely those whose lifted paths (f_t, ζ_t) intersect Σ'. This is the philosophy used in theorems 7 and 8 above, and it suggests the following method. We will call this kind of path–following strategy "adaptive homotopy method", as it adapts the step size (and hence the resulting partition) to the path–to–follow f_t, instead of fixing "a priori" such partition.

Adaptative Homotopy Method with random initial pair (ahmr)

Input: $f \in \mathcal{H}_{(d)}$.

- *Guess at random* $(g, \zeta) \in \mathcal{G}_{(d)}$.
- *Perform the adaptive homotopy method following the segment* $(1-t)g + tf$, $t \in [0,1]$, *starting at* (g, ζ).

Output: an approximate zero $z \in \mathbb{P}(\mathbb{C}^{n+1})$ *of* f.

Note that this algorithm may never give an answer, if $(f_t, \zeta_t) \cap \Sigma' \neq \emptyset$. Nonetheless, as we have seen by the Integral Geometry argument, the probability that this happens is 0.

The combination of the ideas described in Section 4.1 and this adaptative homotopy method yields the following answer to Smale's 17^{th} Problem.

Theorem 12 ([BePa09a])**.** *Algorithm* ahmr *terminates (and hence produces an approximate zero) with probability* 1 *and the average number of arithmetic operations is*

$$O(n^6 N^3 d^3 \log_2 d \log_2 \mathcal{D}),$$

where $\mathcal{D} = \prod_{i=1}^{n} d_i$ *is the Bézout number.*

Moreover, one can modify slightly AHMR *to compute not only projective, but also affine approximate zeros of polynomial systems, with a running time of the same order.*

More specifically, the kind of algorithm that we obtain belongs to the class **Average ZPP** (for **Z**ero error probability, **P**robabilistic, Average **P**olynomial Time), or equivalently **Average Las Vegas**. Namely, it satisfies the following properties:

- The algorithm is probabilistic: For a fixed input f, it may output either FAILURE or some information.
- For $f \notin \Sigma$, the probability that the algorithm provides an answer different to FAILURE is 1.
- If an output is given by the algorithm, it is a correct answer.
- The expected value of the running time is bounded by a polynomial in the input length.

See [BePa09a] for more precise details.

However, the result in [Shu09] (as stated in Theorem 8 above) suggests that the expected complexity estimates must be improved to some quantity which is better than the $O(n^6N^3d^3\log_2 d\log_2 \mathcal{D})$ bound of Theorem 12 above. This was achieved in [BePa09b].

Theorem 13 ([BePa09b])**.** *Algorithm* AHMR *– after changing $\mathcal{G}_{(d)}$ by a similar set $\mathcal{U}_{(d)}$ – terminates (and hence produces an approximate zero) with probability* 1 *and the average number of Newton steps is $O(nNd^{3/2})$. The average number of arithmetic operations is $O(d^{3/2}n^2N^2\log(nd))$.*

Moreover, a random choice of $(g,\zeta) \in \mathcal{U}_{(d)}$ satisfies the thesis of Theorem 10 above, with probability at least $1/2$*. Namely, with probability at least* $1/2$ *we have*

$$\mathrm{E}_{f\in\mathbb{S}}(NS(f,g,\zeta)) \leq CnNd^{3/2}.$$

The new idea was to choose a new questor set $\mathcal{U}_{(d)}$ which, in fact, is the solution variety V with a special distribution, and to use the new estimates described in [Shu09].

The process for randomly choosing an initial pair $(g,\zeta_0) \in \mathcal{U}_{(d)}$ is as follows: Choose at random a full rank $n \times (n+1)$ matrix M, and compute its solution ζ_0. Then, construct a polynomial system with solution ζ_0 whose "linear part" at ζ_0 is given by M and add a higher degree non-linear term h, chosen at random from the vector space defined by $h(\zeta_0) = 0$ and $Dh(\zeta_0) = 0$. Linear and non-linear parts must be correctly weighted. The precise description of this process requires the introduction of some extra notation and is beyond the scope of this paper.

Here is a very brief sketch of the proof of Theorem 13: the main technical result of [BePa09b] is

Theorem 14 ([BePa09b])**.** *Let $\mathcal{U}_{(d)}$ be equal to V, but with the probability distribution inherited from the process described above. Then,*

$$\mathrm{E}_{(g,\theta)\in\mathcal{U}_{(d)}}[\Theta(g,\theta)] = \frac{1}{\mathcal{D}}\mathrm{E}_{f\in\mathbb{S}}\left[\sum\nolimits_{\zeta\in V_{\mathbb{P}}(g)}\Theta(f,\zeta)\right],$$

for any projective measurable mapping $\Theta : V \longrightarrow [0,\infty)$.

That is to say, choosing a random pair $(g, \zeta_0) \in \mathcal{U}_{(d)}$ is equivalent to choosing a random system $g \in \mathbb{S}$, and then choosing one of its solutions ζ_0 with the uniform distribution. One can combine Integral Geometry results and average studies of the condition number like Theorem 9 above to prove that a random root ζ_0 of a random system $g \in \mathbb{S}$ provides a pair (g, ζ_0) that most likely satisfies the thesis of Theorem 10. The main consequence of Theorem 14 is that this apparently difficult process of choosing can be substituted by another one which is computationally doable: choosing a pair $(g, \zeta_0) \in \mathcal{U}_{(d)}$. Theorem 13 follows after a careful study of these ideas. We refer the reader to [BePa09b] for further details.

We also mention two extensions of Theorem 13. One of them is oriented toward the search of several solutions.

Theorem 15 ([BePa09c])**.** *Consider the method* AHMR *above, with random initial pair* $(g, \zeta_0) \in \mathcal{U}_{(d)}$*. Then,*

- *Fixed* $f \notin \Sigma$*, every solution* $\zeta \in V(f)$ *is equally probable as an output of this algorithm. Namely, the Shannon entropy of the algorithm is maximal.*
- *For* $f \notin \Sigma$ *and* $s \geq 1$*, execute the algorithm* s *times on the same input* f *. Then, the probability that the algorithm approximates at least* k *different zeros of* f *is greater than*

$$1 - \binom{\mathcal{D}}{k-1} e^{-s\left(1-\frac{k-1}{\mathcal{D}}\right)}.$$

Thus, randomized linear homotopy methods can be used for the search of more than one solution.

Further work due to the first author of this paper and Mike Shub proved that the variance (and some other higher moments) of the randomized linear homotopy algorithm is also finite and, moreover, polynomial on the size of the input N.

Theorem 16 ([BeSh09b])**.** *Let* Var *denote variance. For* $f \notin \Sigma$ *let* $NS(f)$ *be the average number of homotopy steps needed by algorithm* ANHR *starting at a random pair* $(g, \zeta) \in \mathcal{U}_{(d)}$*. Then,*

$$\mathrm{Var}_{f \in \mathbb{S}}(NS(f)) \leq C d^3 n^2 N^2 \ln \mathcal{D},$$

C a universal constant.

5. Open Problems

5.1. **Find a deterministic version of the algorithms described in Subsection 4.2 above.** Smale's 17^{th} Problem demands a uniform algorithm, and theorems 12 and 13 prove the existence of such an algorithm: ANHR. ANHR is of a probabilistic nature, as are many other popular methods in Numerical Analysis, Computer Science and Computational Mathematics.

However, one may still ask for a *deterministic* and uniform algorithm for Smale's 17^{th} Problem. Currently, we have no answer for that question, although solving the Conjecture about the initial pair described in [ShSm94] and recalled at the end of Section 3 would be one way to resolve this issue. Toward this end, the first author of this paper has obtained the following statement with M. Shub.

Theorem 17 ([BeSh09a]). *Let (g, e_0) be the pair of Shub & Smale's Conjecture above. For every pair $(f, \zeta) \in V$ such that $\mu_{\text{norm}}(f, \zeta) < \infty$, there exists a curve $\Gamma_t \subseteq V$ joining (f, ζ) and (g, e_0), and such that its length in the condition metric is at most*

$$9nd^{3/2} + 2\sqrt{n} \ln \left(\frac{\mu_{\text{norm}}(f, \zeta)}{\sqrt{n}} \right).$$

Moreover, the average number of homotopy steps necessary to solve random systems $f \in \mathbb{S}$ and following these short curves, is $O(nd^3 \ln N)$.

Note that this proves that the lifting curves associated with linear homotopy are not the shortest ones. However, the curves used to prove this statement are up until now not constructible without the knowledge of the zero we want to compute. Thus, this result does not define a new algorithm, but it points out that if we can find those short curves, which are geodesics in the condition metric, then the method could be fastened w.r.t. the linear homotopy. Further developments of this idea are ongoing. For example, the condition metric is being studied from the point of view of convex and non-smooth analysis, see [BDMS07, BoDe09].

In order to find good paths for the homotopy method, one may have to take into account the topology of the solution variety V. This has been recently studied in [BeSh09c], proving for example that the first homotopy group of V is relatively small. In fact, it is often trivial and hence any curve can be smoothly deformed into a length-minimizing geodesic. Some higher homotopy groups are also computed in [BeSh09c]: In the case that $n > 1$,

	$\mathbb{K} = \mathbb{R}$ n and $d_1 + \cdots + d_n - 1$ even	$\mathbb{K} = \mathbb{R}$ other cases	$\mathbb{K} = \mathbb{C}$
$\pi_0(V)$	$\{0, 1\}$	$\{0\}$	$\{0\}$
$\pi_1(V)$	8 elements	4 elements	$\mathbb{Z}/a\mathbb{Z}$
$\pi_2(V)$	$\{0\}$	$\{0\}$	$\mathbb{Z}$
$\pi_k(V), k \geq 3$	$\pi_k(\mathcal{SO}_{n+1})$	$\pi_k(\mathcal{SO}_{n+1})$	$\pi_k(\mathcal{SU}_{n+1})$

where $a = \gcd(n, d_1 + \cdots + d_n - 1)$ and $\mathbb{Z}/a\mathbb{Z}$ is the finite cyclic group of a elements, and $\mathcal{SO}_{n+1}$, $\mathcal{SU}_{n+1}$ are the special orthogonal/unitary groups of dimension $n + 1$.

5.2. **Find an Efficient Numerical Analysis Method for Real Polynomial Equation Solving.** As we pointed out, one of the motivations of Steve Smale's initial studies on the complexity of polynomial equation–solving was the computation of equilibria, and in particular, this means computing real solutions of real systems of equations.

Homotopy continuation methods are not easily adapted to solve real systems of equations. Firstly, because the (real) discriminant variety Σ has (real) codimension one, and hence, random choices of great circles will most likely intersect Σ. There is thus no dimensional argument that grants the existence of lifting curves to be followed by Newton's method or any other. Secondly, the behavior of real equations and real solutions is much more erratic than complex ones. Studies on the expected number of real solutions of a randomly chosen, real system were already initiated in [ShSm93b]. Other studies on the probability distribution of certain estimates concerning real

solving may be found in [Cuc99], [AzWs05],[CKMW08], [BoPa08]. Algorithmic methods are described in [CuSm99], [CKMW08], [BePa09c], but none of them has been shown to be efficient. Symbolic methods, like those described in [BGHP05] (cf. also references therein), are known to be more efficient than the ones presently derived from approximate zero theory in the real case. However, numerical methods are usually expected to be faster!

5.3. **Over–Determined Systems and Hilbert's Nullstellensatz.** There are generally three different cases in polynomial system solving: well–determined (i.e. equal number of equations and unknowns), under–determined (more unknowns than equations) and over–determined (more equations than unknowns). In the previous pages, we have focused on the well–determined case, although many of the results can be easily generalized to the under–determined case. In terms of Algebraic Geometry, these two cases are called smooth complete intersection. However, a central question is the over–determined case. This is just an algorithmic version of Hilbert's Nullstellensatz which has been shown to be $\mathbf{NP}_{\mathbb{C}}$–complete in the theory of real Turing machines (cf. [BSS89] or [BCSS98]). The problem can be stated as follows:

Problem (Hilbert's Nullstellensatz)**.** *Find an algorithm, efficient on the average, that solves the following problem:*
Given $f_1, \ldots, f_{n+1}$ homogeneous polynomials of respective degrees $d_1, \ldots, d_{n+1}$,

(1) Decide whether their have a common zero in $\mathbb{P}_n(\mathbb{C})$.
(2) If this were true, find an approximate zero of each common zero.

Solving the first part of this problem would imply having an (almost) positive answer to Cook's Conjecture. For the moment, the fastest numerical analysis method that we know requires exponential average time (in fact, linear in the Bézout number $\mathcal{D} := \prod_{i=1}^{n} d_i$, cf. [BePa09c]), which is not tractable at all. Even if we assume that our input system $f_1, \ldots, f_{n+1}$ belongs to the algebraic variety of consistent systems (i.e. those sharing a common zero), we do not know how to compute efficiently the (most likely) unique common solution. Less is known about the tractability of this question from the Numerical Analysis approach. For references on the symbolic/geometric approach to this problem, see [HeMo93, Par95, GiHe01, HMPS00, KPS01, CGHMP03].

5.4. **Adapting the algorithms to other data structures.** Another drawback of the algorithms and methods described above is their dependence on the data structure chosen to write down the input polynomials: Dense Encoding of Polynomials. The average complexity estimates are polynomial in N which is the number of coefficients, assuming that none are zero. However, less is known about how to adapt these results to subclasses of input systems.

Typical classes of polynomial equations to be solved are not given in dense encoding. They are, for instance, families of multi–homogeneous or sparse polynomials, few–nomials (only a few non–zero coefficients) or polynomials given by straight–line programs. All these cases refer to subvarieties and submanifolds $\mathcal{I} \subseteq \mathcal{H}_{(d)}$ of positive complex co–dimension. As we only know

average complexity estimates, the behavior of the algorithm on these zero–measure subsets is unknown.

Some progress regarding the probability analysis of the condition number in the sparse case, is due to G. Malajovich and M. Rojas [MaRo00, MaRo04].

5.5. **Lower complexity bounds for the homotopy method.** J. P. Dedieu and Smale also studied lower bounds for the number of Newton steps necessary to perform the homotopy [DeSm98]. Consider $f : \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^m$ homogeneous, and define a Newton Continuation Method Sequence as

$$(f_i, \zeta_i), \;\; 0 \leq i \leq k, \;\; f_i(\zeta_i) = 0, \;\; \alpha(f_{i+1}, \zeta_i) \leq \alpha_0 \text{ with assoc. zero } \zeta_{i+1}.$$

A detailed analysis shows that in these conditions, homotopy with initial pair (f_0, ζ_0) yields an approximate zero of f_k with associated zero ζ_k. Then,

Theorem 18 ([DeSm98])**.** *If $(f_i, \zeta_i), 1 \leq i \leq k$ is a Newton Continuation Method Sequence, then*

$$k \geq c \max\left\{1, \frac{d-1}{2}\right\} d_R(\zeta_0, \zeta_k).$$

Moreover, for homogeneous f let Σ_f be the set of $x \in \mathbb{C}^{n+1}$ such that $rank(Df(x))$ is not maximal. assume that $d_R(\zeta_i, \Sigma_{f_i}) \leq \varepsilon$, $0 \leq i \leq k$. Then,

$$k \geq c\varepsilon^{-1} d_R(\zeta_0, \zeta_k).$$

The main tool used to measure complexity of homotopy methods, i.e. the condition number μ_{norm}, appears somehow implicitly in this lower bound, as it may be commensurable with $d_R(\zeta_i, \Sigma_{f_i})^{-1}$. For example both quantities are infinity if ζ_i is a singular solution of f_i. However, we still have that factor $d_R(\zeta_0, \zeta_k)$ that may be very small.

One may wonder if the condition number is more directly related to lower bounds. If we add the concept of stability to the Newton Continuation Method Sequence, it looks like the condition number $\mu_{\text{norm}}(f_i, \zeta_i)$ should play a role, as condition number and stability do have a relation. So, a natural question is

Conjecture. *If $(f_i, \zeta_i), 1 \leq i \leq k$ is a "Stable" Newton Continuation Method Sequence (in some sense to be specified), then*

$$k \geq c_1 \max\{\log \mu_{\text{norm}}(f_i, \zeta_i)\}^{c_2},$$

$c_1, c_2 > 0$ some universal constants.

References

[AKS04] M. Agrawal, N. Kayal, N. Saxena. *PRIMES is in P.* Annals of Mathematics **160** (2004) 781-793.

[AzWs05] J. Azais, M. Wschebor. *On the roots of a random system of equations. The Shub-Smale theorem and some extensions.* Foundations of Computational Mathematics **5** (2005) 125–144.

[BGHP05] B. Bank, M. Giusti, J. Heintz, L. Pardo. *Generalized polar varieties: Geometry and algorithms.* J. Complexity **21** (2005) 377412.

[BeWe93] T. Becker, V. Weispfenning. Grbner bases. A computational approach to commutative algebra. In cooperation with Heinz Kredel. Graduate Texts in Mathematics, **141**. Springer-Verlag, New York, 1993.

[BDMS07] C Beltrán, J.P. Dedieu, G. Malajovich, and M. Shub, *Convexity properties of the condition number*, to appear (2007).

[BePa06] C. Beltrán and L.M. Pardo, *On the complexity of non–universal polynomial equation solving: old and new results.*, Foundations of Computational Mathematics: Santander 2005. L. Pardo, A. Pinkus, E. Süli, M. Todd editors., Cambridge University Press, 2006, pp. 1–35.

[BePa08] C. Beltrán and L.M. Pardo,, *On Smale's 17th problem: a probabilistic positive solution*, Found. Comput. Math. **8** (2008), no. 1, 1–43.

[BePa09a] C. Beltrán and L.M. Pardo,, *Smale's 17th problem: Average polynomial time to compute affine and projective solutions*, J. Amer. Math. Soc. **22** (2009), 363–385.

[BePa09b] C. Beltrán and L.M. Pardo,, *Fast linear homotopy to find approximate zeros of polynomial systems*, To appear (2009).

[BePa09c] C. Beltrán and L.M. Pardo,, *Computing several zeros of polynomial systems: A complexity analysis and shannon's entropy*, To appear (2009).

[BeSh09a] C. Beltrán and M. Shub, *Complexity of Bezout's Theorem VII: Distance estimates in the condition metric*, Found. Comput. Math. **9** (2009), no. 2, 179–195.

[BeSh09b] C. Beltrán and M. Shub, *A note on the finite variance of the averaging function for polynomial equation solving*, To appear in Found. Comput. Math.

[BeSh09c] C. Beltrán and M. Shub, *On the Geometry and Topology of the Solution Variety for Polynomial System Solving*, To appear (2009).

[Blm04] L. Blum. *Computing over the reals: where Turing meets Newton.* Notices Amer. Math. Soc. **51** (2004) 1024–1034.

[BCSS98] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, Springer-Verlag, New York, 1998.

[BSS89] L. Blum, M. Shub, S. Smale. *On a Theory of Computation and Complexity over the Real Numbers; NP Completeness, Recursive Functions and Universal Machines*. Bull. Amer. Math. Soc. (New Series) **21** (1989) pp. 1–46.

[BoDe09] P. Boito and J.P. Dedieu, *The condition metric in the space of rectangular full rank matrices*, To Appear (2009).

[BoPa08] C. E. Borges, L. M. Pardo. *On the probability distribution of data at points in real complete intersection varieties*. Journal of Complexity **24** (2008) 492–523

[CGHMP03] D. Castro, M. Giusti, J Heintz, G. Matera, L. M. Pardo. *The hardness of polynomial equation solving.* J. Found. Comput. Math. **3-4** (2003) 347–420.

[CHMP01] D. Castro, K. Hägele, J.E. Morais, L. M. Pardo. *Kronecker's and Newton's approaches to solving: a first comparison.* Journal of Complexity **17** (2001) 212–303.

[CMSP02] D. Castro, J.L. Montaña, J. San Martin, L. M. Pardo.*The Distribution of the Condition Number of Rational Data of Bounded bit length*. Foundations of Comput. Math. **2** (2002) 1–52.

[CSP03] D. Castro, J. San Martin, L. M. Pardo.*Systems of Rational Homogeneous Polynomial Equations have Polynomial Size Approximate Zeros*. Journal of Complexity **19** (2003) 161–209.

[Cuc99] F. Cucker.*Approximate zeros and condition numbers*. J. Complexity **15** (1999) 214-226.

[CKMW08] F. Cucker, T. Krick, G. Malajovich, M. Wschebor. *A numerical algorithm for zero counting, I: Complexity and accuracy*. Journal of Complexity **24** (2008) 582–605.

[CuSm99] F. Cucker, S. Smale.*Complexity estimates depending on condition and round-off error*. J. ACM **46** (1999) 113-184.

[Ded01] J.P. Dedieu, *Newton's method and some complexity aspects of the zero-finding problem*, Foundations of computational mathematics (Oxford, 1999), London Math. Soc. Lecture Note Ser., vol. 284, Cambridge Univ. Press, Cambridge, 2001, pp. 45–67.

[DPM03] J.P. Dedieu, P. Priouret, and G. Malajovich, *Newton's method on Riemannian manifolds: convariant alpha theory*, IMA J. Numer. Anal. **23** (2003), no. 3, 395–419.

[DeSh00] J.P. Dedieu and M. Shub, *Multihomogeneous Newton methods*, Math. Comp. **69** (2000), no. 231, 1071–1098 (electronic).

[DeSh01] J.P. Dedieu and M. Shub, *On simple double zeros and badly conditioned zeros of analytic functions of n variables*, Math. Comp. **70** (2001), no. 233, 319–327.

[DeSm98] J.P. Dedieu and S. Smale, *Some lower bounds for the complexity of continuation methods*, J. Complexity **14** (1998), no. 4, 454–465.

[DuLe08] C. Durvye, G. Lecerf, *A concise proof of the Kronecker polynomial system solver from scratch.* Expo. Math. **26** (2008) 101–139.

[EcYo36] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika **1** (1936), 211-218.

[GaZa79] C.B. Garcia, W.I. Zangwill. *Finding all solutions to polynomial systems and other systems of equations*. Math. Programming **16** (1979) 159–176.

[GiHe01] M. Giusti, J. Heintz.*Kronecker's smart, little black boxes.* In Foundations of computational mathematics (Oxford, 1999), London Math. Soc. Lecture Note Ser., **284**, Cambridge Univ. Press, 69–104, 2001.

[GLSY05] M. Giusti, G. Lecerf, B. Salvy and J.-C. Yakoubsohn. *On Location and Approximation of Clusters of Zeros of Analytic Functions*. Foundations of Comput. Mat. **5** (2005) 257–311.

[GLSY07] M. Giusti, G. Lecerf, B. Salvy and J.-C. Yakoubsohn. *On Location and Approximation of Clusters of Zeros: Case of Embedding Dimension One*. Foundations of Comput. Mat. **7** (2007) 1–58.

[HMPS00] K. Hägele, J.E. Morais, L.M. Pardo, M. Sombra, *On the intrinsic complexity of the arithmetic Nullstellensatz,* J. Pure Appl. Algebra **146-2** (2000) 103–183.

[HeMo93] J. Heintz, J. Morgenstern *On the intrinsic complexity of elimination theory,* J. Complexity **9-4** (1993) 471–498.

[Her26] G. Hermann. *Der Frage der endlich vielen Schritte in der Theorie der Polynomideale,* Math. Ann. **95** (1926) 736–788.

[Hi1893] D. Hilbert. *Über die vollen Invariantensysteme,* Math. Ann. **42** (1893), pp. 313–373.

[HiSm79] M.W. Hirsch, S. Smale, *On algorithms for solving $f(x) = 0$*, Comm. Pure Appl. Math. **32** (1979), no. 3, 281–313.

[HSS01] J. Hubbard, D. Schleicher, S. Sutherland, *How to find all roots of complex polynomials by Newton's method*, Inventiones Mathematicae **146** (2001), no. 1, 1–33.

[Kim85] M.H. Kim. *Computational complexity of the Euler type algorithms for the roots of complez polynomials*. PhD thesis, The City University of New York, 1985.

[Kim88] M.H. Kim. *On approximate zeros and rootfinding algorithms for a complex polynomial.* Math. Comp., **51** (1988) 707–719.

[Kö1903] J. König.*Einleitung in die allgemeine Theorie der algebraischen Grözen*. Druck und Verlag von B.G. Teubner,Leipzig., 1903.

[Kri04] T. Krick. *Straight-line programs in polynomial equation solving.* In Foundations of computational mathematics: Minneapolis 2002, London Math. Soc. Lecture Note Ser., 312 (2004) 96–136

[KPS01] T. Krick, L.M. Pardo, M. Sombra *Sharp estimates for the arithmetic Nullstellensatz.* Duke Math. J. **109-3** (2001) 521–598

[Kr1882] L. Kronecker. *Grundzüge einer arithmetischen theorie de algebraischen grössen.* J. reine angew. Math., **92** (1882) 1–122.

[Li83] T.Y. Li. *On Chow, Mallet–Paret and Yorke homotopy for solving systems of polynomials*. Bulle of the Ints. of Math. Acad. Sin. **11** (1983) 433–437.

[Li87] T.Y. Li. *Solving Polynomial Systems*. The Math. Intelligencer **9** (1987) 33-39.

[Mac16] F.S. Macaulay. "The Algebraic Theory of Modular Systems". Cambridge University Press, 1916.

[Mal94] G. Malajovich. *On generalized Newton algorithms: quadratic convergence, path–following and error analysis*. Theoret. Comput. Sci. **133-1** (1994) 65–84.

[Mal00] G. Malajovich. *Condition number bounds for problems with integer coefficients.* J. of Complexity **16** (2000) 529–551.

[MaRo00] G. Malajovich, M. Rojas. Random sparse polynomial systems. 2000. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

[MaRo04] G. Malajovich, M. Rojas. High probability analysis of the condition number of sparse polynomial systems. *Theoret. Comput. Sci.*, 315(2-3):524–555, 2004.

[Mi76] G.L. Miller. *Riemann's hypothesis and tests for primality.* J. Comput. Syst. Sci. **13** (1976) 300–317.

[MiSt05] E. Miller, B. Sturmfels. Combinatorial commutative algebra. Graduate Texts in Mathematics, **227**. Springer-Verlag, New York, 2005.

[Mor05] T. Mora. Solving polynomial equation systems. II. Macaulay's paradigm and Gröbner technology. Encyclopedia of Mathematics and its Applications, **99**. Cambridge University Press, Cambridge, 2005.

[Mor87] A. Morgan. *Solving Polynomial Systems using continuation for engineering and scientific problems.* Prentice–Hall, Englewood Cliffs, N.J., 1987.

[Pan87] V. Y. Pan. *Algebraic Complexity of Computing Polynomial Zeros.* Computers & Math (with Applications) **14** (1987) 285–304.

[Par95] L. M. Pardo. *How lower and upper complexity bounds meet in elimination theory.* In G. Cohen, M. Giusti, and T. Mora, editors, *Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995)*, volume **948** of *Lecture Notes in Computer Science*, pages 33–69. Springer, Berlin, 1995.

[Rab80] M.O. Rabin. *Probabilistic algorithms for testing primality.* J. Number Theory **12** (1980) 128–138.

[Ren85] J. Renegar. On the cost of approximating all roots of a complex polynomial. *Math. Programming*, 32(3):319–336, 1985.

[Ren87] J. Renegar. *On the worst-case arithmetic complexity of approximating zeros of polynomials.* Journal of Complexity, 3(2)(1987) 90–113.

[Sch81] A. Schönhage. *The fundamental theorem of algebra in terms of computational complexity.* Preliminary report, Mathematisches Institut der Universität Tübingen, 1981.

[Sch86] A. Schönhage. Equation solving in terms of computational complexity. In *Proceedings of the International Congress of Mathematicians*, volume 3, page 40, 1986.

[Shb93] M. Shub, *Some remarks on Bezout's theorem and complexity theory*, From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990) (New York), Springer, 1993, pp. 443–455.

[Shu09] ———, *Complexity of Bézout's theorem. VI: Geodesics in the condition (number) metric*, Found. Comput. Math. **9** (2009), no. 2, 171–178.

[ShSm85] M. Shub and S. Smale, *Computational complexity. On the geometry of polynomials and a theory of cost. I*, Ann. Sci. École Norm. Sup. (4) **18** (1985), no. 1, 107–142.

[ShSm86] M. Shub and S. Smale, *Computational complexity: on the geometry of polynomials and a theory of cost. II*, SIAM J. Comput. **15** (1986), no. 1, 145–161.

[ShSm93a] M. Shub and S. Smale, *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc. **6** (1993), no. 2, 459–501.

[ShSm93b] 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.

[ShSm96] M. Shub and S. Smale, *Complexity of Bezout's theorem. IV. Probability of success; extensions*, SIAM J. Numer. Anal. **33** (1996), no. 1, 128–148.

[ShSm94] M. Shub and S. Smale, *Complexity of Bezout's theorem. V. Polynomial time*, Theoret. Comput. Sci. **133** (1994), no. 1, 141–164, Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

[Sma76] Steve Smale, *A Convergent Process of Price Adjustment and Glocal Newton Method.* J. Math. Econom. **3** (1976) 107–120.

[Sma81] S. Smale, *The fundamental theorem of algebra and complexity theory*, Bull. Amer. Math. Soc. (N.S.) **4** (1981), no. 1, 1–36.

[Sma85] S. Smale, *On the efficiency of algorithms of analysis*, Bull. Amer. Math. Soc. (N.S.) **13** (1985), no. 2, 87–121.

[Sma86] S. Smale, *Newton's method estimates from data at one point*, The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985), Springer, New York, 1986, pp. 185–196.

[Sma87] S. Smale, *Algorithms for solving equations*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986) (Providence, RI), Amer. Math. Soc., 1987, pp. 172–195.

[Sma00] S. Smale, *Mathematical Problems for the Next Century.*Mathematics: frontiers and perspectives, 271–294, Amer. Math. Soc., Providence, RI, 2000.

[SoSt77] R. Solovay, V. Strassen. *A fast Monte Carlo test for primality.* SIAM J. on Comput. **6** (1977) 84–85.

[SoWa05] A.J. Sommese, C.W. Wampler. *The numerical solution of systems of polynomials. Arising in engineering and science.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005.

[Ver96] J. Verschelde, *Homotopy Continuation Methods for Solving Polynomial Systems.* Ph.D. Thesis, Katholieke Universiteit Leuven, 1996.

C. Beltrán and L.M. Pardo : Depto. de Matemáticas, Est. y Computación, Facultad de Ciencias, Universidad de Cantabria, E-39071, SANTANDER, Spain

E-mail address: carlos.beltran@unican.es, luis.pardo@unican.es