Theoretical Computer Science 133 (1994) 141–164
Elsevier

# Complexity of Bezout's theorem V: Polynomial time

M. Shub*

*IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, USA*

S. Smale*

*Department of Mathematics, University of California, Berkeley, Berkeley, CA 94720, USA*

*Abstract*

Shub, M. and S. Smale, Complexity of Bezout's theorem V: Polynomial time, Theoretical Computer Science 133 (1994) 141–164.

We show that there are algorithms which find an approximate zero of a system of polynomial equations and which function in polynomial time on the average. The number of arithmetic operations is $cN^{4s}$, where $N$ is the input size and $c$ a universal constant.

## 1. Introduction

The main goal of this paper is to show that the problem of finding approximately a zero of a polynomial system of equations can be solved in polynomial time, on the average. The number of arithmetic operations is bounded by $cN^4$, where $N$ is the number of input variables and $c$ is a universal constant.

Let us be more precise. For $d=(d_1,\ldots,d_n)$ each $d_i$ a positive integer, let $\mathscr{H}_{(d)}$ be the linear space of all maps $f\colon \mathbb{C}^{n+1}\to\mathbb{C}^n, f=(f_1,\ldots,f_n)$, where each $f_i$ is a homogeneous polynomial of degree $d_i$.

The notion of an approximate zero $z$ in projective space $P(\mathbb{C}^{n+1})$ of $f$ has been defined in [11, 12, 14, 6] and below. It means that Newton's method converges quadratically, immediately, to an actual zero $\zeta$ of $f$, starting from $z$. Given an approximate zero, an $\varepsilon$ approximation of an actual zero can be obtained with a further $\log|\log\varepsilon|$ number of steps.

*Correspondence to*: M. Shub, IBM T.J. Watson Research Cent., P.O., Box 218, Yorktown Heights, NY 10598, USA. Email: shub@watson.ibm.com, smale@math.berkeley.edu.

*Supported in part by NSF, IBM, and CRM (Barcelona).

*SSDI* 0304-3975(94)E00060-V

A probability measure on the projective space (of lines) $P(\mathscr{H}_{(d)})$ was developed in [5, 11] and "average" below refers to that measure. Let $N=\text{dimension } \mathscr{H}_{(d)}$ as a complex vector space.

**Main Theorem.** *Fixing $d$, the average number of arithmetic operations to find an approximate zero of $f \in P(\mathscr{H}_{(d)})$ is less than $cN^4$, $c$ a universal constant, unless $n \leqslant 4$ or some $d_i = 1$.*

**Remark.** If $n \leqslant 4$ or some $d_i = 1$, we get $cN^5$

The result is also valid in the non-homogeneous case $f\colon \mathbb{C}^n \rightarrow \mathbb{C}^n$.

The import of the Main Theorem can be understood especially clearly in the case of quadratic polynomials. Thus consider the case $d=(2, \ldots, 2)$ of this theorem. Here we have that the average arithmetic complexity is bounded by a polynomial function of the dimension $n$ since an easy count shows that $N \leqslant n^3$. This seems quite surprizing in view of the history of complexity results for polynomial systems (see [1] for references).

The special case of quadratic systems has extra significance in view of the NP-completeness theorems of [1].

Here it is shown that the decision problem "quadratic systems" is NP-complete over $\mathbb{R}$ or over $\mathbb{C}$. This problem is given $k$ quadratic inhomogeneous equations, in $n$ variables, to decide if there is a common zero. For various reasons, it seems unlikely, that there is a polynomial-time algorithm, even with exact arithmetic (in the sense of [1]) for this problem.

Moreover in the recent "weak model" of [4], quadratic systems definitely do not admit a polynomial-time algorithm, so that $\mathrm{P} \neq \mathrm{NP}$, as was shown in [2].

One might reasonably ask about the analog of the Main Theorem for the worst case, rather than the average. In a trivial sense the corresponding conclusion cannot be true since some polynomial systems have no approximate zeros.

But there seems to be a deeper sense in which the result (or a modification thereof) fails for the worst case. The algorithms we use here are robust in that they work well in the presence of round-off error (see [3, 6]).

This could be made more formal, more conceptual, by the introduction of a "$\delta$-machine", see especially [6], but also [8, 9, 17] for background.

A $\delta$-machine is defined to introduce a relative error $\delta$ at each computation node of a [1] machine.

Then it could no doubt be proved that a $\delta$-machine would be sufficient in our Main Theorem, where $\delta > 0$ could be well-estimated.

On the other hand for $n > 1$, polynomial systems may have one-dimensional sets of solutions ("excess components") and that fact seems to imply, that the worst-case complexity problem is undecidable with $\delta$-machines, for any $\delta > 0$. Even the linear case produces an argument. These ideas need formalization and development, but one would expect that to happen in view of [6].

The algorithm of the Main Theorem, developed in [11, 12, 14], is a homotopy method, with steps based on a version (projective) of Newton's method. There is a weak spot in its present use in that the existence of a start system-zero pair $(g, \zeta)$ is proved, but not constructively. Thus the algorithms depends on $d=(d_1, \ldots, d_n)$, and even on a probability of failure $\sigma$. It is not uniform in the sense of [1] in $d$ and $\sigma$.

In Section 2 an obvious candidate for $(g, \zeta)$ is given. If our (highly likely) conjecture stated there is true, then the uniformity of the algorithms is achieved.

The Main Theorem has the following generalization, which includes the case studied in [14]. We say that $z_1, \ldots, z_l$ are $l$ (distinct) approximate zeros of $f \in P(\mathscr{H}_{(d)})$ if they converge under iteration of (projective) Newton's method to $l$ distinct roots $\zeta_1, \ldots, \zeta_l$ of $f$.

**Generalized Main Theorem.** *Fixing $d$, the average number of arithmetic operations to find $\mathscr{D} \geqslant l \geqslant 1$, where $\mathscr{D} = \prod_{i=1}^{n} d_i$ is the Bezout number, approximate zeros of $f \in P(\mathscr{H}_{(d)})$ is less than $cl^2 N^4$, $c$ a universal constant, unless $n \leqslant 4$ or some $d_i = 1$ in which case $cl^2 N^5$ suffice.*

## 2. Main theorem, weak version

Let $\mathscr{H}_{(d)}$ be as in Section 1 and we suppose it is endowed with the Hermitian inner product in [5, 11] invariant under the unitary group $U(n+1)$. Then $S(\mathscr{H}_{(d)})$ denotes the unit sphere in $\mathscr{H}_{(d)}$ and

$$\hat{V} = \{(f, \zeta) \in S(\mathscr{H}_{(d)}) \times P(\mathbb{C}^{n+1}) \mid f(\zeta) = 0\}.$$

Then let $\hat{\Sigma}'$ be the set of singular points of the restriction $\hat{\pi}_1 : \hat{V} \to S(\mathscr{H}_{(d)})$ (i.e. $(f, \zeta)$ such that the derivative $D\hat{\pi}_1 : T_{f,\zeta}(\hat{V}) \to T_f(S(\mathscr{H}_{(d)}))$ is singular). Compare all this with the similar notions for $V$ of [11, 12, 14]. In fact, one has the fibration $\hat{V} \to V$ with fibres $SO(2)$ induced by the fibration $S(\mathscr{H}_{(d)}) \to P(\mathscr{H}_{(d)})$; thus the vertical distance to $\Sigma'$ of [11] defines a similar vertical distance $\rho$ (same notation) and neighborhood $N_\rho(\hat{\Sigma}')$ of $\hat{\Sigma}'$ in $\hat{V}$.

Let $\mathscr{L}_g$ be the space of great circles of $S(\mathscr{H}_{(d)})$ which contain $g \in S(\mathscr{H}_{(d)})$. Let $\psi : S(\mathscr{H}_{(d)}) - \{\pm g\} \to \mathscr{L}_g$, $\psi(f) = L_f$ be the map which sends $f$ into the unique great circle containing $f$ and $g$.

For such an $f$ we may define $\hat{L}_f = \hat{\pi}_1^{-1}(L_f) \subset \hat{V}$. If $L_f \cap \hat{\Sigma} = \emptyset$, where $\hat{\Sigma} = \hat{\pi}_1(\hat{\Sigma}')$ is the discriminant locus [11], then $\hat{L}_f$ is a one-dimensional submanifold in $\hat{V}$ oriented by going from $g$ to $f$ omitting $-g$. If in addition, $\zeta$ is a zero of $g$, then there is a unique arc in $\hat{L}_f$ starting at $(g, \zeta)$ and ending at the first point of $\hat{\pi}_1^{-1}(f)$ met on $\hat{L}_f$. Let us call that arc $\hat{L}(f, g, \zeta)$.

**Remark.** $\hat{L}(f, g, \zeta)$ may be interpreted as a path of zeros of "the homotopy" $tf + (1-t)g$ as $t$ goes from 0 to 1.

Our Hermitian structure on $\mathscr{H}_{(d)}$ induces natural Riemannian metrics and probability measures on $S(\mathscr{H}_{(d)})$, $P(\mathscr{H}_{(d)})$ and $\mathscr{L}_g$; see [12, 14]. Moreover with these measures, the natural maps $S(\mathscr{H}_{(d)}) \to P(\mathscr{H}_{(d)})$, $\psi: S(\mathscr{H}_{(d)}) - \{\pm g\} \to \mathscr{L}_g$ are measure preserving, in the usual sense that meas $\psi^{-1} A = \text{meas}\, A$.

Fixing $(g, \zeta)$ as above let $\sigma = \sigma(\rho, g, \zeta)$, $0 < \sigma < 1$ be the probability that $\hat{L}(f, g, \zeta)$ meets $N_\rho(\hat{\Sigma}')$ for $f \in S(\mathscr{H}_{(d)})$. Later we will see how $\sigma$ may be interpreted as the "probability of failure".

**Theorem 2.1.** *For each $\rho > 0$, there is a $(g, \zeta) \in \hat{V}$ such that*

$$\sigma \leqslant c\rho^2 N^2 n^3 D^{3/2}.$$

Here, as throughout this paper, $c$ is some universal constant. Moreover,

$$D = \max_{i=1,\ldots,n} (d_i).$$

**Proposition 2.2.** *We have $n^3 D^3 \leqslant cN$ unless some $d_i = 1$ or $n \leqslant 4$ in which case there is a slightly weaker estimate.*

The proof is left to the reader (use $\binom{D+n}{n} \leqslant N$).

Using this proposition and [11] for the case $n = 1$, one can use Theorem 2.1 to get the estimate:

$$\sigma \leqslant c\rho^2 N^3.$$

Theorem 2.1 has a ready interpretation in terms of the condition number of $f$,

$$\mu_{g,\zeta}(f) = \max_{(h,z) \in \hat{L}(f, g, \zeta)} \mu_{\text{norm}}(h, z).$$

Here see [11, 12, 14] for $\mu_{\text{norm}}(h, z)$ as well as the condition number theorem

$$\rho(h, z) = \frac{1}{\mu_{\text{norm}}(h, z)}, \quad (h, z) \in V, \quad (\text{or } \hat{V}).$$

Let

$$S_{g,\zeta,\rho} = \{f \in S(\mathscr{H}_{(d)}) - \{\pm g\} \mid \hat{L}(f, g, \zeta) \cap N_\rho(\hat{\Sigma}') \neq \emptyset\}.$$

**Theorem 2.3.** *Given $\rho > 0$, there is a $g \in S(\mathscr{H}_{(d)})$ such that*

$$\frac{1}{\mathscr{D}} \frac{\sum_{\zeta, g(\zeta) = 0} \operatorname{Vol} S_{g,\zeta,\rho}}{\operatorname{Vol} S(\mathscr{H}_{(d)})} \leqslant c\rho^2 N^2 n^3 D^{3/2}.$$

Note that Theorem 2.1 is a consequence of Theorem 2.3, since the left-hand side of Theorem 2.3 is an average over the zeros $\zeta$ of $g$. For at least one $\zeta$, one gets less than the average. Hence there exist a pair $(g,\zeta)\in V$ such that

$$\sigma=\frac{\mathrm{Vol}\, S_{g,\zeta,\rho}}{\mathrm{Vol}\, S(\mathscr{H}_{(d)})}\leqslant c\rho^2 N^2 n^3 D^{3/2}$$

proving Theorem 2.1.

**Conjecture 2.4.** *The pair $(g,\zeta)$ of Theorem* 2.1 *given by $g_i(z)=z_0^{d_i-1}z_i$, $i=1,\dots,n$, and $\zeta=e_0=(1,0,\dots,0)$ makes the conclusion of Theorem* 2.1 *true.*

The truth of this conjecture would make our algorithms more constructive and in fact algorithms in the sense of [1] with input $(d,\sigma,f)$.

**Remark.** Let $S_{g,\rho}=\bigcup_{\substack{\zeta\\ g(\zeta)=0}} S_{g,\zeta,\rho}$. So

$$\mathrm{Vol}\,(S_{g,\rho)}\leqslant\sum_{\zeta}\mathrm{Vol}\, S_{g,\zeta,\rho}.$$

In this way it is seen that Theorem 2.3 is a sharp form of Theorem 2 of [14]. In fact that suggests a proof. Theorem 2.3 is proved in the next section following [14], but with a multiplicity function taken into account.

Next we use the Main Theorem of [11] and Theorem 2.1 to obtain a weak version of our main result.

**Main Theorem** (Weak version). *Let be given a probability of failure $\sigma$, $0<\sigma<1$. Then there exists $(g,\zeta)\in V$ such that a number of projective Newton steps $k$ sufficient to find an approximate zero of input $f\in S(\mathscr{H}_{(d)})$ is*

$$k\leqslant cN^3/\sigma.$$

*(If $n\leqslant 4$, or some $d_i=1$, one obtains only $cN^4/\sigma$.)*

**Proof.** The Main Theorem of [11] asserts that $k\leqslant cD^{3/2}/\rho^2$. But by Theorem 2.1, we can take $\sigma=c\rho^2N^2D^{3/2}n^3$. Thus by Proposition 2.2, we obtain the estimate of our theorem.

The number of arithmetic operations of a projective Newton step can be bounded by $cN$, so we get $cN^4/\sigma$ arithmetic operations. □

This theorem is easier to prove than the Main Theorem, the weakness coming from the factor $1/\sigma$ which cannot be averaged. Thus we must be able to replace $1/\sigma$ by $1/\sigma^{1-\varepsilon}$, Sections 4–7 are devoted to this.

## 3. The proof of Theorem 2.3

As we have noted Theorem 2.3 uses a sharpened form of Theorem 2 of [14]. This suggests a proof. To this end we sharpen accordingly, Theorem C of [12] and then Proposition 4(a) and (b) of [14].

Let $M_{2\rho}: S(\mathscr{H}_{(d)}) \to Z^+$ be defined as follows. $M_{2\rho}(f)$ is the number of roots of $f$ in $N_{2\rho}(\hat{\Sigma}')$ (perhaps $\infty$) or more properly, the cardinality of $\hat{\pi}_1^{-1}(f) \cap N_{2\rho}(\hat{\Sigma}')$. Here we are following the notation of Section 2.

**Theorem 3.1.** *For any* $\rho > 0$

$$\frac{1}{\operatorname{Vol} S(\mathscr{H}_{(d)})} \int_{S(\mathscr{H}_{(d)})} M_{2\rho}(f) \leqslant c\rho^4 n^3 N^2 \mathscr{D}.$$

This is a sharper version of Theorem C of [12], but the same proof works.

Define for $(g, \zeta) \in \hat{V}$, $\mathscr{L}_{g,\zeta,\rho} \subset \mathscr{L}_g$ as follows: Suppose that

$$L \cap \hat{\Sigma} = \emptyset \quad (L, \hat{\Sigma} \subset S(\mathscr{H}_{(d)})).$$

Then $\hat{\pi}_1^{-1}(L) \to L$ is a $\mathscr{D}$-fold covering map and $\hat{\pi}_1^{-1}(L) - \hat{\pi}_1^{-1}(-g)$ consists of $\mathscr{D}$ open arcs in $\hat{V}$. Let $A_{g,\zeta}$ denote the arc among them that contains $\zeta$. Then define:

$$\mathscr{L}_{g,\zeta,\rho} = \{L \in \mathscr{L}_g \mid A_{g,\zeta} \cap N_\rho(\hat{\Sigma}') \neq \emptyset\}.$$

**Lemma 3.2.** *With notation as above,*

$$\frac{\sum_{\zeta, g(\zeta)=0} \operatorname{Vol} S_{g,\zeta,\rho}}{\operatorname{Vol} S(\mathscr{H}_{(d)})} \leqslant \frac{\sum_{\zeta, g(\zeta)=0} \operatorname{Vol} \mathscr{L}_{g,\zeta,\rho}}{\operatorname{Vol} \mathscr{L}_g}.$$

The proof follows from the fact that $\psi: S(\mathscr{H}_{(d)}) - \{\pm g\} \to \mathscr{L}_g$ preserves the probability measures and that

$$S_{g,\zeta,\rho} \subset \psi^{-1}(\mathscr{L}_{g,\zeta,\rho}).$$

**Lemma 3.3.** *For any* $\rho > 0$, $d = (d_1, \ldots, d_n)$ *there is a* $g \in S(\mathscr{H}_{(d)})$ *such that*

$$\frac{\sum_{\zeta, g(\zeta)=0} \operatorname{Vol} \mathscr{L}_{g,\zeta,\rho}}{\operatorname{Vol} \mathscr{L}_g} \leqslant \left(\frac{c\rho^2}{D^{3/2}}\right)^{-1} \frac{\operatorname{Vol} S^1}{\operatorname{Vol} S(\mathscr{H}_{(d)})} \int_{S(\mathscr{H}_{(d)})} M_{2\rho}(f).$$

(*Here* $\operatorname{Vol} S^1 = 2\pi$.)

Note that Lemmas 3.2 and 3.3, and Theorem 3.1 give the proof of Theorem 2.3. One has to just check that the constants come out correctly.

Thus it remains to prove Lemma 3.3. For this we sharpen Propositions 4(a) and 4(b) of [14] as follows.

Let $\mathscr{L}$ denote the space of great circles in $S(\mathscr{H}_{(d)})$.

**Proposition 3.4.** (a) *There is a* $g \in S(\mathscr{H}_{(d)})$ *such that*

$$\frac{1}{\operatorname{Vol} \mathscr{L}_g} \int_{\mathscr{L}_g} \int_{f \in L} M_{2\rho}(f) \leqslant \frac{1}{\operatorname{Vol} \mathscr{L}} \int_{\mathscr{L}} \int_{f \in L} M_{2\rho}(f).$$

(b) *Moreover,*

$$\frac{1}{\operatorname{Vol}\mathscr{L}}\int_{\mathscr{L}}\int_{f\in L}M_{2\rho}(f)=\frac{\operatorname{Vol}S^1}{\operatorname{Vol}S(\mathscr{H}_{(d)})}\int_{S(\mathscr{H}_{(d)})}M_{2\rho}(f).$$

The proof follows so closely that of Propositions 4a, b of [14] that we leave it to the reader.

**Corollary 3.5.** *For any $d, \rho$ there is a $g\in S(\mathscr{H}_{(d)})$ such that*

$$\frac{1}{\operatorname{Vol}\mathscr{L}_g}\int_{\mathscr{L}_g}\int_{L}M_{2\rho}(f)\leqslant\frac{\operatorname{Vol}S^1}{\operatorname{Vol}S(\mathscr{H}_{(d)})}\int_{S(\mathscr{H}_{(d)})}M_{2\rho}(f).$$

Let $\tau_\rho:\mathscr{L}_g\to Z^+$ be defined as follows. $\tau_\rho(L)$ is the number of $A_{g,\zeta}$ meeting $N_\rho(\hat{\Sigma}')$. Then from the definitions

$$\sum_{\substack{\zeta\\ g(\zeta)=0}}\operatorname{Vol}\mathscr{L}_{g,\zeta,\rho}=\int_{L\in\mathscr{L}_g}\tau_\rho(L).$$

From the corollary of Theorem 1 of [14] we have

$$\tau_\rho(L)\leqslant\left(\frac{c\rho^2}{D^{3/2}}\right)^{-1}\int_{f\in L}M_{2\rho}(f).$$

Therefore, we obtain for any $g\in S(\mathscr{H}_{(d)})$,

$$\frac{\sum_\zeta\operatorname{Vol}\mathscr{L}_{g,\zeta,\rho}}{\operatorname{Vol}\mathscr{L}_g}\leqslant\frac{1}{\operatorname{Vol}\mathscr{L}_g}\left(\frac{c\rho^2}{D^{3/2}}\right)^{-1}\int_{\mathscr{L}_g}\int_{L}M_{2\rho}(f).$$

The corollary of Proposition 3.4 now finishes the proof. □

## 4. Integral geometry

The goal here is to estimate the volume of certain real algebraic sets. The arguments go back to Crofton and [19], but we use a modern form closer to [12, 14].[1]

The following theorem illustrates what we are doing.

**Theorem 4.1.** *Let $M\subset P(\mathbb{R}^l)$ be a real algebraic variety, given by the vanishing of real homogeneous equations with its complexification having dimension m and degree $\delta$, over $\mathbb{C}$. Then the m-dimensional volume of M is less than or equal to $\delta\operatorname{Vol}P(\mathbb{R}^{m+1})$.*

The affine version can be dealt with by the same methods and in fact is in [15] for the one-dimensional case in $\mathbb{R}^2$.

Let $\hat{V}\subset S(\mathscr{H}_{(d)}\times P(\mathbb{C}^{n+1})$ be as in Section 2 with restrictions of the projections denoted by $\hat{\pi}_1:\hat{V}\to S(\mathscr{H}_{(d)})$, $\hat{\pi}_2:\hat{V}\to P(\mathbb{C}^{n+1})$. Let $L$ be a great circle in $S(\mathscr{H}_{(d)})$, with

[1] *Added in proof*: See also [18].

$L \cap \Sigma = \emptyset$. Then $\hat{\pi}_1^{-1}(L)$ is a one-dimensional (real) submanifold in $\hat{V}$. Also $\hat{\pi}_2(\hat{\pi}_1^{-1}(L))$ is a curve $B$ in $P(\mathbb{C}^{n+1})$.

**Theorem 4.2.** *The length of $B$ is less than or equal to $2\mathscr{D}^2$.*

We sketch some basic results on integration, especially Fubini's theorem, in a Riemannian manifold setting (the Coarea Formula, see [7]).

Suppose $F: X \to Y$ is a surjective map from a Riemannian manifold $X$ to a Riemannian manifold $Y$, and suppose the derivative $DF(x): T_x(X) \to T_{f(x)}(Y)$ is surjective for almost all $x \in X$. The horizontal subspace $H_x$ of $T_x(X)$ is defined as the orthogonal complement to $\ker Df(x)$.

The horizontal derivative of $F$ at $X$ is the restriction of $DF(x)$ to $H_x$. The Normal Jacobian $N_J F(x)$ is the absolute value of the determinant of the horizontal derivative, defined almost everywhere on $X$.

**Example 4.3.** Suppose a compact Lie group $G$ acts transitively and isometrically on a manifold $S$. Fixing $s_0 \in S$, the normal Jacobian of the map $G \to S$, $g \to gs_0$ is a constant.

More generally the following result is seen easily.

**Proposition 4.4.** *Let $F: X \to Y$ be a map of Riemannian manifolds, equivariant under the action of a compact group $G$ of isometries of $X$ and $Y$. If $G$ acts transitively on $X$ then the normal Jacobian is a constant.*

Fubini's theorem takes the following form.

**Coarea Formula.** *Let $F: X \to Y$ be a map of Riemannian manifolds satisfying the above surjectivity conditions. Then for $\varphi: X \to \mathbb{R}$*

$$\int_X \varphi(x) = \int_Y \int_{F^{-1}(y)} \varphi(x) \frac{1}{N_J F(x)}.$$

Here the usual integrability conditions of Fubini's theorem are supposed.

Next suppose that $G$ is a compact Lie group acting transitively and isometrically on the manifold $S$. Let $N$ be a submanifold of $S$ such that the subgroup $I_N$ of $G$ leaving $N$ invariant acts transitively on $N$. Thus the quotient space

$$G_N = G/I_N = \{gN \mid g \in G\}$$

represents the various images of $N$ under applications of elements of $G$.

Our application will be to the case $S$ is real projective space $P(\mathbb{R}^l)$, $G$ is the orthogonal group $O(l)$ and $N$ is $P(\mathbb{R}^{k+1})$ considered as inbedded in $P(\mathbb{R}^l)$ as a coordinate $k$ subspace. In this case $G_N$ can be identified with the Grassmannian $G_k$ of $k$-dimensional linear subspaces of $P(\mathbb{R}^l)$.

Returning to the general setting let $W \subset G_N \times S$ be the submanifold

$$W = \{(gN, s) \mid s \in gN\}.$$

Let $p_1 : W \to G_N$, $p_2 : W \to S$ be the restrictions of the natural projections. The following proposition can be easily proved.

**Proposition 4.5.** *The above $W$ is indeed a submanifold, the product action of $G$ on $G_N \times S$ leaves $W$ invariant, and acts isometrically and transitively on $W$. Moreover, $p_1$ and $p_2$ are equivariant under $G$.*

**Corollary to Propositions 4.4 and 4.5.** *The normal Jacobians of $p_1$ and $p_2$ are constant.*

Since $G$ acts on $S$ it acts also on the tangent bundle $T(S)$ by the derivative. It also acts on the associated bundle $G_m(T(S))$ with fiber, all $m$ planes through the origin in $T_s(S)$. We say that the action of $G$ on $S$ is $m$-transitive if this last action is transitive. Note that in our application, $m$-transitivity is satisfied for all the relevant integers $m$.

**Proposition 4.6.** *Let $M$ be an $m$-dimensional submanifold of $S$. Suppose that $G$ as above is also $m$-transitive on $S$. Let $\tilde{M} = p_1^{-1}(M) \subset W$. Then the restriction $p_2|_{\tilde{M}} : \tilde{M} \to G_N$ has normal Jacobian a constant $c$ (possibly not defined everywhere) depending only on $G$, $S$, $N$ and $m$.*

**Proof.** Define an associated bundle $E(T(W)) = E$ over $W$ of $T(W)$ as follows. Let $w \in W$: we will define the fiber $E_w$ by

$$E_w = \{Dp_2(w)^{-1}(L) \subset T_w \mid L \in G_m(T_{p_2(w)}(S))\}.$$

Then the induced action of $G$ on $E$ is transitive by our hypothesis of $m$-transitivity. Let $\tilde{H}_w$ be the orthogonal space to $\ker Dp_1(w)$ in $E_w$. Then $N_{J p_1}|_{\tilde{M}}(w)$ is the determinant of the restriction of the derivative of $p_1$ to $\tilde{H}_w$. By our transitivity we are finished, noting also that the surjectivity of the derivative holds everywhere if at one point. □

**Theorem 4.7.** *Let $M$ be as above. Then*

$$\operatorname{Vol} M = \frac{c}{\operatorname{Vol} W_0} \int_{gN \in G_N} \operatorname{Vol}(gN \cap M),$$

*where $c$ is the constant of Proposition* 4.6 *and $W_0 = p_2^{-1}(s_0)$ for some fixed $s_0 \in S$.*

The volumes are of course in the appropriate dimensions.

**Proof.** Apply the Coarea Formula and Proposition 4.5 to $p_2$ restricted to $\tilde{M}$ to obtain

$$\operatorname{Vol} \tilde{M} = \operatorname{Vol} M \operatorname{Vol} W_0.$$

Next apply the same argument to $p_1$ restricted to $\tilde{M}$ to obtain

$$\operatorname{Vol}\tilde{M}=\int_{gN\in G_N}\operatorname{Vol}(gN\cap M)\left(\frac{1}{N_{Jp_1}|\tilde{M}}\right).$$

By Proposition 4.6, and by elimination of $\operatorname{Vol}\tilde{M}$ we obtain the result. □

Returning to our special case of projective spaces recall that $G_k$ denotes the space of $k$-linear subspaces of $P(\mathbb{R}^l)$.

**Theorem 4.8.** *Let $M\subset P(\mathbb{R}^l)$ be an m-submanifold. Then*

$$\operatorname{Vol}M=\frac{\operatorname{Vol}P(\mathbb{R}^{m+1})}{\operatorname{Vol}P(\mathbb{R}^{m+k-l+2})}\frac{1}{\operatorname{Vol}G_k}\int_{L\in G_k}\operatorname{Vol}(L\cap M).$$

**Proof.** It suffices to prove that

$$\frac{c}{\operatorname{Vol}W_0}=\frac{\operatorname{Vol}P(\mathbb{R}^{m+1})}{\operatorname{Vol}P(\mathbb{R}^{m+k-l+2})}\frac{1}{\operatorname{Vol}G_k}.$$

Apply Theorem 4.7 to $M=P(\mathbb{R}^{m+1})$. So

$$\operatorname{Vol}P(\mathbb{R}^{m+1})=\frac{c}{\operatorname{Vol}W_0}\int_{L\in G_k}\operatorname{Vol}(L\cap P(\mathbb{R}^{m+1})).$$

The theorem follows, noting that $\operatorname{Vol}(L\cap P(\mathbb{R}^{m+1}))=\operatorname{Vol}P(\mathbb{R}^{m+k-l+2})$. □

**Proof of Theorem 4.1.** Let $M_{\mathbb{C}}\subset P(\mathbb{C}^l)$ be the complexification of $M\subset P(\mathbb{R}^l)$ of the theorem. So $M=M_{\mathbb{C}}\cap P(\mathbb{R}^l)$, $P(\mathbb{R}^l)\subset P(\mathbb{C}^l)$. The real dimension of $M$ is less than or equal to $m$ and we suppose that it is $m$. The generic $(l-m-1)$ linear subspace in $P(\mathbb{R}^l)$ meets $M$ transversally and in at most $\delta$ points since its complexification can meet $M_{\mathbb{C}}$ in at most $\delta$ points of transversal intersection. Thus

$$\int_{L\in G_{l-m-1}}\operatorname{Vol}(L\cap M)\leqslant\delta\operatorname{Vol}(G_{l-m-1}).$$

Since $\operatorname{Vol}P(\mathbb{R}^1)=1$, and we are finished by Theorem 4.8. □

**Proof of Theorem 4.2.** Real projective $2n+1$ space $P(\mathbb{R}^{2(n+1)})$ fibers over $P(\mathbb{C}^{n+1})$ with $S^1$ fibres, by the isometric action of the unit complex numbers mod ± 1. Let

$$q:P(\mathbb{R}^{2(n+1)})\to P(\mathbb{C}^{n+1})$$

be this fibration. Denote by $A$, $q^{-1}(B)$, where $B$ is as in Theorem 4.2. Note that $A$ is a surface.

**Lemma 4.9.** (a) *The length of $B$ equals* $(1/\pi)$ Area $A$.

(b) *A generic* $(2n-1)$ *linear subspace of* $P(\mathbb{R}^{2(n+1)})$ *meets $A$ in at most $\mathscr{D}^2$ points.*

**Proof of Theorem 4.2** (*continued*). We first show that Theorem 4.2 follows from Lemma 4.9.

We use Theorem 4.8 just as in the proof of Theorem 4.1. Thus,

$$\text{Area}\, A \leqslant \mathscr{D}^2 \frac{\operatorname{Vol} P(\mathbb{R}^3)}{\operatorname{Vol} P(\mathbb{R}^1)} \leqslant \operatorname{Vol} P(\mathbb{R}^3)\mathscr{D}^2$$

and so length $B \leqslant (\operatorname{Vol} P(\mathbb{R}^3)/\pi)D^2$ proving Theorem 4.2. □

So it remains to prove Lemma 4.9. Part (a) of Lemma 4.9 is again a (rather simple) case of the coarea formula. Now consider (b). The idea is to lift the setting to $P(\mathbb{R}^{2n+2})$ and then complexity.

Observe that $L \subset P(\mathscr{H}_{(d)})$, so $\hat{\pi}_1^{-1} L \subset L \times P(\mathbb{C}^{n+1}) \subset P(\mathscr{H}_{(d)}) \times P(\mathbb{C}^{n+1})$ and of course $\hat{\pi}_1^{-1} L \subset \hat{V}$.

The following diagram helps:

$$\begin{tikzcd}
q_*^{-1}\hat{\pi}_1^{-1}(L) \arrow[d] & = (L \times P(\mathbb{R}^{2n+2})) \cap \hat{\hat{V}} \arrow[r, "\hat{\hat{\pi}}_2"] \arrow[d, "q_*"] & P(\mathbb{R}^{2n+2}) \arrow[d, "q"] \\
\hat{\pi}_1^{-1}(L) & = (L \times P(\mathbb{C}^{n+1})) \cap \hat{V} \arrow[r, "\hat{\pi}_2"] & P(\mathbb{C}^{n+1})
\end{tikzcd}$$

Here $q_*: L \times P(\mathbb{R}^{2n+2}) \to L \times P(\mathbb{C}^{n+1})$ is induced by $q$ and $\hat{\hat{V}} = q_*^{-1}(\hat{V})$. Thus $A = \hat{\hat{\pi}}_2 q_*^{-1} \hat{\pi}_1^{-1} L$.

Now $L$ may be described as $tg_1 + (s-t)g_2$ for particular $g_1, g_2 \in \mathscr{H}_{(d)}$, $t, s \in \mathbb{R}$ and thus $L$ may be identified with $P(\mathbb{R}^2) \subset P(\mathbb{C}^2)$. Then $q_*^{-1} \hat{\pi}_1^{-1} L = X$ may be defined by the $2n$ real homogeneous equations

$$t \operatorname{Re} f + (s-t) \operatorname{Re} g = 0,$$

$$t \operatorname{Im} f + (s-t) \operatorname{Im} g = 0.$$

These equations are homogeneous of degree 1 in $(s, t)$ and $(d_1, \ldots, d_n)$ in $(x_j, y_j)$ where $z_j = x_j + \sqrt{-1}\, y_j$. Complexifiying these equations in $s, t, x_j, y_j$, we obtain a variety $X_{\mathbb{C}}$ in $P(\mathbb{C}^2) \times P(\mathbb{C}^{2(n+1)})$. The generic $2n-1$ linear subspace $K \subset P(\mathbb{R}^{2n+2})$ has the property that the complexification $P(\mathbb{C}^2) \times K_{\mathbb{C}}$ meets $X_{\mathbb{C}}$ in at most $\mathscr{D}^2$ points of transversal intersections (via elementary intersection theory).

This yields the upper bound $\mathscr{D}^2$ for the number of real intersection points, finishing the proof of the Lemma 4.9 (and hence Theorem 4.2). □

## 5. Some approximate zero Theory

In this section we do some of the $\alpha$-theory and approximate zero theory of [15] and [11] in the projective Newton setting. See also [6].

We take a slightly different perspective than [11], but still relying on it in part. In this account we do not attempt to get the best constants.

Let $f:\mathbb{C}^{n+1}\rightarrow\mathbb{C}^n, f=(f_1,\ldots,f_n)$, where $f_i$ is analytic and homogeneous of degree $d_i$, e.g. $f\in\mathscr{H}_{(d)}$.

Recall that the projective Newton method is a map $N_f:\mathbb{C}^{n+1}\rightarrow\mathbb{C}^{n+1}$, defined by

$$X\rightarrow X-(Df(x)|_{\mathrm{Null}_x})^{-1}f(x)$$

and an induced map $N_f:P(\mathbb{C}^{n+1})\rightarrow P(\mathbb{C}^{n+1})$ which we have denoted by the same letter. We also sometimes identify $x$ and its equivalence class in $P(\mathbb{C}^{n+1})$. Also $N_f$ is obviously not defined everywhere, so the above represents a certain abuse of notation.

A point $x\in\mathbb{C}^{n+1}$ or $P(\mathbb{C}^{n+1})$ will be called an *approximate zero*[2] for $f$ if the sequence $x_i$ defined by $x=x_0$ and $N_f(x_i)=x_{i+1}$ is defined for all natural numbers $i$, and there is a zero $\zeta$ of $f$ such that $d_{\mathbf{R}}(x_i,\zeta)\leqq(\frac{1}{2})^{2^i-1}\,d_{\mathbf{R}}(x_0,\zeta)$. Here $d_{\mathbf{R}}$ is the Riemannian distance in $P(\mathbb{C}^{n+1})$ and $\zeta$ is the associated zero of $x$, i.e. $\zeta=\lim x_i$.

The next theorems are devoted to giving criteria for a point to be an approximate zero in terms of invariants $\alpha$, $\beta_0$, $\gamma_0$ and the distance function $d_{\mathbf{R}}$.

$$\beta_0(f,x)=\frac{1}{\|x\|}\,\|(Df(x)|_{\mathrm{Null}\,x})^{-1}f(x)\|,$$

$$\gamma_0(f,x)=\sup_{k\geqq 2}\left(\frac{1}{k!}\|(Df(x)|_{\mathrm{Null}\,x})^{-1}D^kf(x)\|\right)^{1/(k-1)}\|x\|,$$

$$\alpha(f,x)=\beta_0(f,x)\gamma_0(f,x),$$

where $\beta_0,\gamma_0,\alpha$ are taken as infinite if $(Df(x)|_{\mathrm{Null}\,x})^{-1}$ does not exist

**Note.** We have not restricted $D^kf(x)$ to Null $x$ as in [11]. This change does not effect the theorems of [11]; they still hold with this $\gamma_0$.

If $\zeta$ is a simple zero of $f$, i.e. $(Df(\zeta)|_{\mathrm{Null}\,\zeta})^{-1}$ exists, then $\zeta$ is a fixed point of $N_f$. We begin by showing that $N_f$ contracts discs of a certain size in $P(\mathbb{C}^{n+1})$, centered at $\zeta$, towards $\zeta$.

**Theorem 5.1.** *There are constants $c>0$ and $\hat{u}_*>0$ such that given $f$ as above, and $x,\zeta\in P(\mathbb{C}^{n+1})$ with*

$$f(\zeta)=0,$$

$$d_{\mathbf{R}}(x,\zeta)\leqq 1,$$

$$d_{\mathbf{R}}(x,\zeta)\gamma_0(f,\zeta)\leqq\hat{u}_*,$$

*then $\|DN_f(x)\|<cd_{\mathbf{R}}(x,\zeta)\gamma_0(f,\zeta)$. Here $DN_f(x):T_xP(\mathbb{C}^{n+1})\rightarrow T_{N_f(x)}P(\mathbb{C}^{n+1})$ is the derivative.*

**Remark.** Theorem 5.1 is sharper if we maximize $\hat{u}_*$ and minimize $c$. In any case in the sequel we assume $\hat{u}_*\leqq 1/2c$.

[2] In [15] there are approximate zeros of the first and second kind. Our approximate zeros here correspond to the second kind.

Let $B_r(x)$ denote the closed ball of radius $r$ around $x$.

**Corollary 5.2.** *If* $r < \min(1, \hat{u}_*/\gamma_0(f, \zeta))$, *then* $N_f : B_r(\zeta) \to P(\mathbb{C}^{n+1})$ *is well defined and* $N_f(B_r(\zeta)) \subset B_r(\zeta)$. *It is a contraction with contraction constant* $cr\gamma_0(f, \zeta)$, *so* $N_f(B_r(\zeta)) \subset B_{r'}(\zeta)$, $r' = cr^2\gamma_0(f, \zeta)$.

**Proof.** Observe that $B_r(\zeta)$ is convex. By Theorem 5.1 the length of the image of a curve of length $L$ in $B_r(\zeta)$ is at most $cr\gamma_0(f, \zeta)L$. This establishes the assertions on the contraction constant. Applying the contraction estimate to the straight lines from $\zeta$ to $x$, $x \in B_r(\zeta)$, shows that $N_f(B_r(\zeta)) \subset B_{r'}(\zeta)$, $r' = cr^2\gamma_0(f, \zeta)$. As $cr\gamma_0(f, \zeta) < 1$, $N_f$ is indeed a contraction and $N_f(B_r(\zeta)) \subset B_r(\zeta)$. □

Contraction mappings have a convenient property, which we pause to record.

Let $X$ be a complete metric space with metric $d$, $\phi : X \to X$ a contraction map with contraction constant $k$ and unique fixed point $p$.

**Proposition 5.3.** *Let* $\phi, X, p, k$ *be as above. Then for any* $x \in X$

$$\frac{1}{1+k} d(x, \phi(x)) \leqslant d(x, p) \leqslant \frac{1}{1-k} d(x, \phi(x)).$$

**Proof.** Both inequalities are standard. We prove only the left-hand one:

$$d(x, \phi(x)) \leqslant d(x, p) + d(\phi(x), p) \leqslant (1+k) d(x, p). \qquad \square$$

**Corollary 5.4.** *If* $f(\zeta) = 0$, $d_{\mathrm{R}}(x, \zeta) < 1$ *and* $d_{\mathrm{R}}(x, \zeta)\gamma_0(f, \zeta) < \hat{u}_*$. *Then x is an approximate zero of f with associated zero* $\zeta$.

**Proof.** By Corollary 5.2.

$$d_{\mathrm{R}}(N_f(x), \zeta) \leqslant c d_{\mathrm{R}}(x, \zeta)^2 \gamma_0(f, \zeta)$$

and by induction

$$d_{\mathrm{R}}(N_f^k(x), \zeta) \leqslant (c d_{\mathrm{R}}(x, \zeta)\gamma_0(f, \zeta))^{2^k - 1} d_{\mathrm{R}}(x, \zeta).$$

Since $\hat{u}_* \leqslant 1/2c$, $c d_{\mathrm{R}}(x, \zeta)\gamma_0(f, \zeta) \leqslant \frac{1}{2}$. □

Theorem 5.1 follows immediately from the next two propositions which are of independent interest.

**Proposition 5.5.** *Let f be as above. For* $x \in P(\mathbb{C}^{n+1})$

$$\| DN_f(x) \| \leqslant 2\alpha(f, x).$$

**Proof.** Let

$$E_x = x + \mathrm{Null}\, x,$$

$$E_{N_f(x)} = N_f(x) + \mathrm{Null}\, N_f(x).$$

We use $E_x$ and $E_{N_f(x)}$ as charts for $P(\mathbb{C}^{n+1})$ at $x$ and $N_f(x)$, respectively, in the obvious way. □

Let $v \in \operatorname{Null} x$.

$$DN_f(x)(v) = \pi(Df(x)|_{\operatorname{Null} x})^{-1} D_x Df(x)|_{\operatorname{Null} x}(v)(Df(x)|_{\operatorname{Null} x})^{-1} f(x),$$

where $\pi$ is the orthogonal projection onto the null space of $N_f(x)$.

Thus $\| DN_f(x)v \|_{\mathbb{C}^{n+1}} \leqslant 2\gamma(f,x)\beta(f,x)\| v \|_{\mathbb{C}^{n+1}}$. Recall that $N_f(x) \in x + \operatorname{Null} x$ so $\| N_f(x) \|_{\mathbb{C}^{n+1}} \geqslant \| x \|_{\mathbb{C}^{n+1}}$.

Now

$$\begin{aligned} \| DN_f(x)v \|_{T_{N_f(x)}P(\mathbb{C}^{n+1})} &= \frac{\| DN_f(x)v \|_{\mathbb{C}^{n+1}}}{\| N_f(x) \|_{\mathbb{C}^{n+1}}} \\ &\leqslant \frac{\| DN_f(x)v \|_{\mathbb{C}^{n+1}}}{\| x \|_{\mathbb{C}^{n+1}}} \leqslant 2\alpha(f,x)\frac{\| v \|_{\mathbb{C}^{n+1}}}{\| x \|_{\mathbb{C}^{n+1}}} \\ &= 2\alpha(f,x)\| v \|_{T_x P(\mathbb{C}^{n+1})}. \qquad \square \end{aligned}$$

For the next proposition we use a simple geometry lemma.

**Lemma 5.6.** *Let* $x, y \in \mathbb{C}^{n+1} - \{0\}$. If $\mathrm{d}_{\mathrm{R}}(x,y) < 1$ *in* $P(\mathbb{C}^{n+1})$ *then* $x \in \lambda y + \operatorname{Null} \lambda y$ *for some* $\lambda$ *and* $\| x - \lambda y \| / \| \lambda y \| = \tan d_{\mathrm{R}}(x,y)$.

**Proof.** $\| x - \lambda y \| / \| x \| = \sin d_{\mathrm{R}}(x,y)$ by Proposition 4, Section 1 of [17]. So $\| x - \lambda y \| / \| \lambda y \| = \tan d_{\mathrm{R}}(x,y)$. □

**Proposition 5.7.** *There exist constants* $c > 0$, $\hat{u}_* > 0$ *such that, if*

$$d_{\mathrm{R}}(x,\zeta) < 1$$

*and*

$$d_{\mathrm{R}}(x,\zeta)\gamma_0(f,\zeta) < \hat{u}_*.$$

*Then* $2\alpha(f,x) < c d_{\mathrm{R}}(x,\zeta)\gamma_0(f,\zeta)$.

**Proof.** It follows from the lemma that

$$d_{\mathrm{R}}(x,\zeta) \leqslant \frac{\| x - \lambda\zeta \|}{\| \lambda\zeta \|} \leqslant \tan(1) d_{\mathrm{R}}(x,\zeta),$$

where $x \in \lambda\zeta + \operatorname{Null} \lambda\zeta$. Also, $\gamma_0(f,\lambda\zeta) = \gamma_0(f,\zeta)$ so

$$d_{\mathrm{R}}(x,\zeta)\gamma_0(f,\zeta) \quad \text{and} \quad u = \frac{\| x - \lambda\zeta \|}{\| \lambda\zeta \|}\gamma_0(f,g)$$

also differ at most by a multiple of $\tan(1)$.

Now apply Proposition 2, Section III-2 of [11] to conclude that $2\alpha(f,x)<2(\kappa^2 u/\psi(u)^2)$; $\kappa=\kappa(u)$ and $\psi(u)$ are close to one for $u$ small so we are done. Here we have been using the notation of [11]. □

Next we prove a version of the $\alpha$-theorem.

**Definition 5.8.** *Let* $\hat{\gamma}_0(f,x)=\max(\gamma_0(f,x),1)$ and $\hat{\alpha}(f,x)=\hat{\gamma}_0(f,x)\beta_0(f,x)$.

**Theorem 5.9.** (Projective $\alpha$-theorem). *There is an* $\hat{\alpha}_{\text{proj}}>0$ *such that if* $\hat{\alpha}(f,x)<\hat{\alpha}_{\text{proj}}$ *then* $x$ *is an approximate zero of* $f$.

Compare this to [6].

In fact, we will prove a little more.

**Proposition 5.10.** *There are constants* $\alpha_*>0$, $c>0$ *such that if* $\alpha(f,x)<\alpha_*$ *then there is a zero* $\zeta$ *of* $f$ *and* $d_{\mathbf{R}}(x,\zeta)\gamma_0(f,\zeta)<c\alpha(f,x)$

First we prove Theorem 5.9 from Proposition 5.10.

**Proof of Theorem 5.9.** Just let $\hat{\alpha}_{\text{proj}}=\min(\alpha_*,u_*/c)$ and apply Proposition 5.10 and Corollary 5.4. □

**Proof of Proposition 5.10.** By the Domination Theorem (following the notation of that theorem) (Theorem 2, Section I-2 of [11] for $\alpha(f,x)<\alpha_0$ there is a zero $\zeta$ of $f$ in $E_x=x+\operatorname{Null} x$ such that

$$\|x-\zeta\|<\frac{\tau(\alpha(f,x))}{\gamma(f,x)}.$$

Thus,

$$\frac{\|x-\zeta\|}{\|x\|}\gamma_0(f,x)<\tau(\alpha(f,x)). \qquad (*)$$

Now if $\hat{\alpha}(f,x)$ is small so is $\beta_0(f,x)$ and so is $\|x-\zeta\|/\|x\|$ by the Domination Theorem again. Thus $\|x-\zeta\|/\|x\|$ and $d_{\mathbf{R}}(x,\zeta)$ differ by a multiplicative constant close to one and $d_{\mathbf{R}}(x,\zeta)$ is also small. Since $\zeta$ is a zero, $\ker Df(\zeta)^{-1}=\operatorname{Null}(\zeta)$ and

$$\|(Df(\zeta)|_{\operatorname{Null}\zeta})^{-1}Df(\zeta)|_{\operatorname{Null}x}\|\leqslant 1.$$

Hence in Proposition 2, Section III-2 of [11], $\kappa$ may be taken as 1 and

$$\gamma_0(f,\zeta)\leqslant\frac{\gamma_0(f,x)}{\psi(\tau(\alpha(f,x))(1-\tau(\alpha(f,x))))}.$$

Substituting in $(*)$

$$d_{\mathbf{R}}(x,\zeta)\gamma_0(f,,\zeta)\leqslant c\alpha(f,x)$$

for $\alpha(f,x)$ small enough. Finally, $\alpha(f,x)<\hat{\alpha}(f,x)$ so if $\hat{\alpha}_{\text{proj}}$ is small enough we are done. □

## 6. The homotopy

The goal of this section is to give the proof of Theorem 6.1 below.

Throughout this section we suppose that $(f_t,\zeta_t)$ is a curve in $\hat{V}-\hat{\Sigma}'$, $0\leqslant t\leqslant 1$. Except for Proposition 6.2 and Lemma 6.3, we assume moreover that $f_t$ can be represented as $f_t=tf+(1-t)g$ for some $f,g\in S(\mathscr{H}_{(d)}))$. Let $\hat{\gamma}$ be an upper bound for 1 and $\gamma_0(f,\zeta_t)$, $0\leqslant t\leqslant 1$.

**Theorem 6.1.** *With $f_t$ as above, there is a partition*

$$t_0=0, \qquad t_i<t_{i+1}, \qquad t_k=1$$

*with*

$$x_0=\zeta_0, \qquad x_i=N_{f_{t_i}}(x_{i-1}), \quad i=1,\dots,k$$

*well-defined for each i and $x_i$ is an approximate zero of $f_{t_i}$, with associated zero $\zeta_{t_i}$. Also*

$$k\leqslant c\mu_{(g,\zeta)}(f)(1+D^{3/2}L),$$

*where L is the length of the curve $\zeta_t$. Moreover, (as we will see) $t_i$ can be easily calculated at $t_{i-1}$.*

Recall that $N_{f_t}$ *is given by projective Newton's method.*

Towards the proof we have the following proposition.

**Proposition 6.2.** *There exist universal constants $\alpha^*$, $u^*$ with the following property. Suppose*

$$t,\ t+\Delta t\in[0,1], \quad x_t\in P(\mathbb{C}^{n+1})$$

*satisfy:*

$$\hat{\gamma}d_{\mathbf{R}}(x_t,\zeta_t)\leqslant u^*,$$

$$\hat{\gamma}\beta_0(f_{t'},x_t)\leqslant\alpha^* \quad \textit{for all } t'\in[t,t+\Delta t],$$

*then for all $t'\in[t,t+\Delta t]$, $\hat{\gamma}d_{\mathbf{R}}(x',\zeta_{t'})\leqslant u^*$, where $x'=N_{f_{t'}}(x_t)$ and $x_t$ is an approximate zero of $f_{t'}$ with associated zero $\zeta_{t'}$.*

*Moreover, given any positive constant K we may take $u^*\leqslant K\alpha^*$. In fact, $K=1/48$ in what follows.*

**Proof.** As long as $\alpha^* < \hat{\alpha}_{\mathrm{proj}}$, $x_t$ is an approximate zero for $f_{t'}$ from Theorem 5.9. That $\zeta_{t'}$ is the associated zero is a simple continuity argument. As usual there is a constant $K_1$ close to one such that,

$$d_{\mathbf{R}}(N_{f_{t'}}(x_t), x_t) \leqslant K_1 \beta_0(f_{t'}, x_t) \leqslant K_1 \alpha^*/\hat{\gamma}.$$

So from Proposition 5.3,

$$d_{\mathbf{R}}(x_t, \zeta_{t'}) \leqslant 2K_1 \alpha^*/\hat{\gamma}$$

and by Corollary 5.2

$$d_{\mathbf{R}}(N_{f'_t}(x_t), \zeta_{t'}) \leqslant c\left(\frac{2K_1\alpha^*}{\hat{\gamma}}\right)^2 \hat{\gamma} = \frac{(2K_1\alpha^*)^2 c}{\hat{\gamma}}$$

$c$ the constant of Corollary 5.3.

Choose $\alpha^*$, $u^*$ so that $2K_1^2 c(\alpha^*)^2 < u^*$ and $u^* < K\alpha^*$. □

**Lemma 6.3.** *There are universal constants $K > 0$, $u_{**} > 0$ with the following property. If $f \in P(\mathscr{H}_{(d)})$, $f(\zeta) = 0$ and*

$$d_{\mathbf{R}}(x, \zeta)\gamma_0(f, \zeta) \leqslant u_{**},$$

$$d_{\mathbf{R}}(x, \zeta) \leqslant 1/D,$$

*then $\mu_{\mathrm{norm}}(f, x) \leqslant K \mu_{\mathrm{norm}}(f, \zeta)$.*

For the proof we may take $x \in \zeta + \mathrm{Null}\,\zeta$,

Following [11] and the notation there,

$$\begin{aligned}
\mu_{\mathrm{norm}}(f, x) &= \|f\| \, \|(Df(x)|_{\mathrm{Null}\,x})^{-1} \Delta(d_i^{1/2}) \Delta(\|x\|^{d_i})\| \\
&\leqslant \|f\| \, \|(Df(x)|_{\mathrm{Null}\,x})^{-1} Df(x)|_{\mathrm{Null}\,\zeta}\| \, \|(Df(x)|_{\mathrm{Null}\,\zeta})^{-1} Df(\zeta)|_{\mathrm{Null}\,\zeta}\| \\
&\quad \left\| (Df(\zeta)|_{\mathrm{Null}\,\zeta})^{-1} \Delta(d_i^{1/2}) \Delta(\|\zeta\|^{d_i}) \right\| \, \left\| \Delta\left(\left(\frac{\|x\|}{\|\zeta\|}\right)^{d_i}\right) \right\| \\
&= \mu_{\mathrm{norm}}(f, \zeta) \, \|(Df(x)|_{\mathrm{Null}\,\mathrm{x}})^{-1} Df(x)|_{Null\,\zeta}\| \\
&\quad \times \|(Df(x)|_{\mathrm{Null}\,\zeta})^{-1} Df(\zeta)|_{\mathrm{Null}\,\zeta}\| \\
&\quad \left\| \Delta\left(\left(\frac{\|x\|}{\|\zeta\|}\right)^{d_i}\right) \right\| .
\end{aligned}$$

Let $r_0 = \|x - \zeta\|/\|\zeta\|$ and $u = r_0\gamma_0(f, \zeta)$ so $r_0$ and $d_{\mathbf{R}}(x, \zeta)$ differ by a multiplicative constant close to 1, and the same for $u$ and $d_{\mathbf{R}}(x, \zeta)\gamma_0(f, g)$.

By Proposition 1, Section III-2 of [11]

$$\|(Df(x)|_{\mathrm{Null}\,x})^{-1} Df(x)|_{\mathrm{Null}\,\zeta}\| \leqslant \frac{(1 + r_0^2)^{1/2}}{1 - r_0\left(\dfrac{(2-u)u}{\psi(u)}\right)}.$$

By Lemma 3(2), Section II-1 of [11]

$$\|(Df(x)|_{\mathrm{Null}\,\zeta})^{-1}\,Df(\zeta)|_{\mathrm{Null}\,\zeta}\| \leqslant \frac{(1-u)^2}{\psi(u)}$$

and these quantities are both bounded as soon as $d_{\mathbf{R}}(x,\zeta)\gamma_0(f,g)$ is small enough.

Finally

$$\frac{\|x\|}{\|\zeta\|}=(1+r_0^2)^{1/2}\leqslant\left(1+\left(\frac{K_1}{D}\right)^2\right)^{1/2} \quad \text{for } K_1 \text{ near } 1$$

so $(\|x\|/\|\zeta\|)^{d_i}$ is also bounded and we are done.

Now let

$$W_t=\|(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}((f-g)(x_t))\|,$$
$$B_t=\|(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}D(f-g)(x_t))\|.$$

**Proposition 6.4.** (a) $B_t\leqslant 2\mu_{\mathrm{norm}}(f_t,x_t)$.

(b) $\beta^-\leqslant\beta_0(f_{t'},x_t)\leqslant\beta^+$,

*where*

$$\beta^\pm=\frac{\Delta t W_t+\beta_0(f_t,x_t)}{1\mp\Delta t B_t}, \quad \Delta t=|t'-t|$$

*as long as* $\Delta t\leqslant 1/B_t$.

**Proof.** We may assume $\|x_t\|=1$. Then

$$\begin{aligned}B_t&\leqslant\|(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}\|\,\|D(f-g)(x_t)\|\\&\leqslant\mu(f_t,x_t)(\|Df(x_t)\|+\|Dg(x_t)\|)\\&\leqslant 2\mu(f_t,x_t)\end{aligned}$$

using Proposition 2, Section III-1 of [11]. Finally,

$$\mu(f_t,x_t)\leqslant\mu_{\mathrm{norm}}(f_t,x_t)$$

proving (a).

Since

$$\beta_0(f_{t'},x_t)=\|(t'-t)(Df_{t'}(x_t)|_{\mathrm{Null}\,x_t})^{-1}(f-g)(x_t)+(Df_{t'}(x_t)|_{\mathrm{Null}\,x_t})^{-1}f(x_t)\|.$$

□

Proposition 6.4(b) follows from the following lemma.

**Lemma 6.5.**

(a) $\|(Df_{t'}(x_t)|_{\mathrm{Null}\,x_t})^{-1}Df_t(x_t)|_{\mathrm{Null}\,x_t}\|\leqslant\dfrac{1}{1-|t'-t|\,\|B_t\|}$.

(b) $\|(Df_{t'}(x_t)|_{\mathrm{Null}\,x_t})^{-1}Df_t(x_t)|_{\mathrm{Null}\,x_t}\|_{\min}\geqslant\dfrac{1}{1+|t'-t|\,\|B_t\|}$.

**Proof.**

$$Df_{t'}(x_t)|_{\mathrm{Null}\,x_t} = Df_t(x_t)|_{\mathrm{Null}\,x_t} + (t'-t)D(f-g)(x_t)|_{\mathrm{Null}\,x_t},$$

so

$$(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}Df_{t'}(x_t) = I + (t'-t)(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}D(f-g)(x_t).$$

Now the minimum norm

$$\|(Df_{t'}(x_t)|_{\mathrm{Null}\,x_t})^{-1}Df_t(x_t)|_{\mathrm{Null}\,x_t}\|_{\min} = \frac{1}{\|(Df_t(x_t)|_{\mathrm{Null}\,x_t})^{-1}Df_{t'}(x_t)|_{\mathrm{Null}\,x_t}\|}$$
$$\geqslant \frac{1}{1+|t'-t|B_t},$$

which proves Lemma 6.5(b). Part (a) follows from the additional fact that if $\|I-A^{-1}B\|<c<1$ then $\|B^{-1}A\|<1/(1-c)$. □

Define

$$\hat{\mu} = \mu_{g,\zeta}(f) = \max_t \mu_{\mathrm{norm}}(f_t, \zeta_t).$$

Choose $\hat{\gamma} = D^{3/2}\hat{\mu}$. This is permissible by Proposition 3, Section I-3 of [11]. Set $\alpha^{**} = \min(\alpha^*, u^{**})$, $\alpha^*, u^{**}$ of Proposition 6.2, Lemma 6.3, respectively. Also $\Delta t = |t'-t|$.

**Proposition 6.6.** *With notation as above, there exists a universal constant c as follows. Given t with*

$$\hat{\gamma} d_{\mathrm{R}}(x_t, \zeta_t) \leqslant u^*,$$

*then there is a* $\Delta t$ *such that*

$$\beta^+(\Delta t) \leqslant \alpha^{**}/\hat{\gamma}$$

*and*

$$\Delta t \geqslant \frac{c}{\hat{\mu}} \text{ or else } d_{\mathrm{R}}(\zeta_t, \zeta_{t+\Delta t}) \geqslant c\alpha^{**}/\gamma.$$

In fact $\Delta t$ is easily computed as will be seen. Also there is an obvious adjustment to make in case $t+\Delta t>1$.

**Proof.** If

$$\beta^+|_{\Delta t = 1/2\beta_t} \leqslant \alpha^{**}/\hat{\gamma}$$

let $\Delta t = 1/2B_t$. Then by Proposition 6.4,

$$\beta(f_{t'}, x_t) \leqslant \alpha^{**}/\hat{\gamma}.$$

Moreover by the same proposition, $B_t \leqslant 2\mu_{\text{norm}}(f_t, x_t)$. Then by Lemma 6.3 we obtain

$$\Delta t = \frac{1}{2B_t} \geqslant \frac{1}{4\mu_{\text{norm}}(f_t, x_t)} \geqslant \frac{1}{4K\mu_{\text{norm}}(f_t, \zeta_t)}.$$

Otherwise let $\Delta t$ be the solution of

$$\beta^+(\Delta t) = \alpha^{**}/\hat{\gamma}.$$

Then $\Delta t \leqslant 1/2B_t$ and it only remains to show that

$$d_{\mathrm{R}}(\zeta_t, \zeta_{t'}) \geqslant c\alpha^{**}/\hat{\gamma}.$$

**Lemma 6.7.** *Under the conditions of Proposition* 6.6,

(a) $d_{\mathrm{R}}(x_t, \zeta_t),\ d_{\mathrm{R}}(x_{t'}, \zeta_{t'}) \leqslant \dfrac{1}{48}\dfrac{\alpha^{**}}{\hat{\gamma}}$,

(b) $\beta_0(f_t, x_t) \leqslant \dfrac{1}{8}\dfrac{\alpha^{**}}{\hat{\gamma}}$,

(c) $\beta^-(\Delta t) \geqslant \dfrac{1}{6}\dfrac{\alpha^{**}}{\hat{\gamma}}$,

(d) $d_{\mathrm{R}}(x_t, x_{t'}) \geqslant \dfrac{1}{12}\dfrac{\alpha^{**}}{\hat{\gamma}}$,

(e) $d_{\mathrm{R}}(\zeta_t, \zeta_{t'}) \geqslant \dfrac{1}{24}\dfrac{\alpha^{**}}{\hat{\gamma}}$.

**Proof.** Since (e) gives our proof of Proposition 6.6 it only remains to prove Lemma 6.7. The first part of (a) is in the hypothesis and since (Proposition 6.4)

$$\beta_0(f_{t'}, x_t) \leqslant \beta^+(\Delta t) = \alpha^{**}/\hat{\gamma},$$

Proposition 6.2 yields the second part of (a).

Use (a) (first part), that $\beta_0(f_t, \zeta_t) = 0$ and Proposition 2, Section II-I of [11] to easily obtain (b). For (c) we argue as follows.

Recall $\Delta t B_t \leqslant \frac{1}{2}$. Since

$$\beta^+(\Delta t) = \frac{\Delta t W_t + \beta_0(f_t, \zeta_t)}{1 - \Delta t B_t} = \frac{\alpha^{**}}{\hat{\gamma}},$$

using (b),

$$\Delta t W_t \geqslant \frac{1}{2}\frac{\alpha^{**}}{\hat{\gamma}} - \beta_0(f_t, \zeta_t) \geqslant \frac{3}{8}\frac{\alpha^{**}}{\hat{\gamma}}.$$

Therefore,

$$\beta^-(\Delta t) = \frac{\Delta t W_t - \beta_0(f_t, \zeta_t)}{1 + \Delta t B_t} \geqslant \frac{2}{3}\left(\frac{3}{8} - \frac{1}{8}\right)\frac{\alpha^{**}}{\hat{\gamma}} \geqslant \frac{1}{6}\frac{\alpha^{**}}{\hat{\gamma}}.$$

We obtain (d) using Proposition 6.4, $\beta_0(f_{t'}, x_t) \geqslant \beta^-(\Delta t)$ and (c). This uses the definition of $\beta_0$ as the Newton vector and the exponential map.

Finally, (e) follows from

$$d_{\mathbf{R}}(\zeta_t, \zeta_{t'}) \geqslant d_{\mathbf{R}}(x_t, x_{t'}) - d_{\mathbf{R}}(x_t, \zeta_t) - d_{\mathbf{R}}(x_{t'}, \zeta_{t'})$$

using (a) and (d). □

**Proof of Theorem 6.1.** We use Proposition 6.2 and 6.6. Let $t_0 = 0$, $t_i = t_{i-1} + \Delta t$ according to Proposition 6.6. So at each step $\Delta t$ satisfies one of the alternatives in (b).

Since

$$\sum d_{\mathbf{R}}(\zeta_{t_i}, \zeta_{t_{i-1}}) \geqslant L, \quad \hat{\mu} = \mu_{(g,\zeta)}(f).$$

We get the result. □

## 7. The main theorem

The goal of this section is to prove the Main Theorem of Section 1. To this end we first prove two theorems on the number of projective Newton steps sufficient to find a zero.

**Theorem 7.1.** *Fix $d = (d_1, \ldots, d_n)$ and a probability of failure $\sigma$, $0 < \sigma < 1$. Then there exists $(g, \zeta) \in \hat{V}$ such that the number $k$ of projective Newton steps, starting from $(g, \zeta)$, sufficient to find an approximate zero of input $f \in S(\mathscr{H}_{(d)})$ is*

$$k \leqslant \frac{cN^3}{\sigma^{1-\varepsilon}}, \qquad \varepsilon = \frac{1}{\log \mathscr{D}}$$

*(or $cN^4/(\sigma^{1-\varepsilon})$ if some $d_i = 1$ or $n \leqslant 4$).*

Thus the set of $f$ where the algorithm fails to produce such an approximate zero in $k$ steps has probability measure less than $\sigma$.

For the proof we have the following result.

**Proposition 7.2.** *Fix $(d) = (d_1, \ldots, d_n)$ and suppose $(g, \zeta) \in \hat{V}$, $f \in S(\mathscr{H}_{(d)}) - \{\pm g\}$ and $0 \leqslant \bar{\varepsilon} \leqslant 1$ are given. Then*

$$k \leqslant c(\mu_{(g,\zeta)}(f))^{2-\bar{\varepsilon}}(\mathscr{D}^2)^{\bar{\varepsilon}} D^{3/2}$$

*steps of projective Newton's method are sufficient to produce an approximate zero of $f$.*

In Section 2 we have defined $\mu_{g,\zeta}(f)$ and an arc $\hat{L}(f, g, \zeta)$ in $\hat{V} \subset S(\mathscr{H}_{(d)}) \times P(\mathbb{C}^{n+1})$. Let $L$ be the length of $\hat{\pi}_2(\hat{L}(f, g, \zeta))$.

**Lemma 7.3.** (a) $L \leqslant 2\mathscr{D}^2$.
(b) $L \leqslant \pi \mu_{g,\zeta}(f)$.

Part (a) of the lemma is a consequence of Theorem 4.2. Part (b) is a projective space version of the Proposition, Section 1D of [14]. The proof is the same noting that $\mu(h, z) \leqslant \mu_{\text{norm}}(h, z)$ for all $(h, z) \in \hat{V}$, and that the length of a great circle in $S(\mathscr{H}_{(d)})$ is $2\pi$.

Returning to the proof of Proposition 7.2, we have from Theorem 6.1 that

$$c\mu_{(g,\zeta)}(f)LD^{3/2}$$

steps of projective Newton suffice. Hence by the lemma (b),

$$c\mu_{(g,\zeta)}(f)^{2-\bar{\varepsilon}}L^{\bar{\varepsilon}}D^{3/2}$$

steps suffice.

Finally, from Lemma 7.3(a),

$$c\mu_{(g,\zeta)}(f)^{2-\bar{\varepsilon}}\mathscr{D}^{2\bar{\varepsilon}}D^{3/2}$$

steps suffice. This proves Proposition 7.2 □

**Proof of Theorem 7.1.** In Proposition 7.2 take $\varepsilon = 2/\log \mathscr{D}$ so that $\mathscr{D}^{2\varepsilon}$ is a universal constant.

By Proposition 7.2 we need to show there exists $(g, \zeta) \in \hat{V}$ such that the function $\mu_{(g,\zeta)}(f)^{2-\bar{\varepsilon}}\mathscr{D}^{2\bar{\varepsilon}}D^{3/2}$ is bounded above by $cN^3/(\sigma^{1-\varepsilon})$ for a subset of $f \in S(\mathscr{H}_{(d)})$ of probability measure at least $1-\sigma$.

Solve the equation $\sigma = c\rho^2 N^2 n^3 D^{3/2}$ for $\rho$. Apply Theorem 2.1, for this $\rho$ and the condition number theorem to conclude the existence of $(g, \zeta)$ such that

$$(\mu_{(g,\zeta)}(f))^{2(1-\varepsilon)} \leqslant \frac{cN^2 n^3 D^{3/2}}{\sigma^{1-\varepsilon}}$$

for all $f$ in a subset of probability measure at least $1-\sigma$. Now Proposition 7.2, Section 2 and a little arithmetic finish the proof. □

**Theorem 7.4.** *The average number of projective Newton steps sufficient to find an approximate zero of $f \in S(\mathscr{H}_{(d)})$ is less than or equal to $c(\log \mathscr{D})N^3$. ($c \log \mathscr{D} N^4$ if some $d_i = 1$ or $n \leqslant 4$).*

In Theorem 7.4 we employ a quasi-algorithm. This construction fails to be an algorithm because it employs an infinite sequence $(g_i, \zeta_i) \in \hat{V}$, $i = 1, 2, 3, \ldots$ without exhibiting them.

The idea is quite simple. Start with $(g, \zeta)$ as in Theorem 7.1 to insure a "small" chance of failure say $\sigma = \frac{1}{2}$ initially. If on input $f$ the algorithm fails, halve the chance of failure and start over.

More formally let parameters of our quasi-algorithm $(g_i,\zeta_i)\in\hat{V}$ be given by Theorem 7.1, with probability of failure $\sigma_i=1/2^i$, $i=1,2,3,4$.

Let $K(\sigma)=cN^3/(\sigma^{1-\varepsilon})$ (or $cN^4/(\sigma^{(1-\varepsilon)})$ if some $d_i=1$ or $n\leqslant 4$).

For input $f$, $i=1$, do $K(\sigma_i)$ projective Newton steps for the homotopy $(1-t)g_i+tf$ starting at $\zeta_0$ as in Theorem 6.1.

If $X_{K(\sigma_i)}$ is an approximate zero of $f$ by the alpha test, Theorem 5.9, halt and output $X_{K(\sigma_i)}$.

If not set $i=i+1$ and repeat (some $f$).

The average number of steps of this algorithms is less than or equal to

$$\sum_{i=1}^{\infty}\sigma_i\left(\sum_{j\leqslant i}K(\sigma_j)\right)\leqslant c'\sum_{i=1}^{\infty}\sigma_i K(\sigma_i)\leqslant c''\int_0^1 K(\sigma)\,\mathrm{d}\sigma.$$

The first inequality follows by summing a geometric series. For the second note that $\sigma_i=|\sigma_i-\sigma_{i-1}|$, $i=1,\dots$ with $\sigma_0=1$, that $K$ is monotone on the interval $[\sigma_i,\sigma_{i-1}]$ and $K(\sigma_i)/K(\sigma_{i-1})=2^{(1-\varepsilon)}$ so the Riemann sum

$$\sum_{i=1}^{\infty}|\sigma_i-\sigma_{i-1}|K(\sigma_i)\leqslant 2^{(1-\varepsilon)}\int_0^1 K(\sigma)\,\mathrm{d}\sigma.$$

Finally, $\int_0^1 K(\sigma)\,\mathrm{d}\sigma=c\log\mathscr{D}N^3$ (or $\leqq\log\mathscr{D}N^4$ if some $d_i=1$ or $n\leqslant 4$).

See [10] for more arguments of this sort.

**Proof of the Main Theorem.** To prove the Main Theorem, we need only make the passage from the number of Newton steps to the number of arithmetic operations. This argument uses well-known facts from numerical analysis about the number of arithmetic operations needed for approximations, for solving linear problems, etc. We omit the details.

**Proof of Generalized Main Theorem.** We sketch some of the changes necessary for the proof.

Theorem 2.1 has the following version which also follows from Theorem 2.3.

Fixing $g\in S(\mathscr{H}_{(d)})$ and $\zeta_1,\dots,\zeta_1\in P(\mathbb{C}^{n+1})$, $l$ distinct zeros of $q$. Let $\sigma_l=\sigma_l(\rho,g,\zeta_1,\dots,\zeta_1)$ $0<\sigma_l<1$ be the probability for $f\in S(\mathscr{H}_{(d)})$ that for some $i=1,\dots,l$, $\hat{L}(f,g,\zeta_i)$ meets $N_\zeta(\hat{\Sigma}')$.

**Theorem 2.3** (New version). *For each $\rho>0$ and $l$, $1\leqslant l\leqslant\mathscr{D}$ there is a $g\in S(\mathscr{H}_{(d)})$ and distinct zeros $\zeta_1,\dots,\zeta_l\in P(\mathbb{C}^{n+1})$* of $g$ *such that*

$$\sigma_l\leqslant\rho^2 n^3 D^{3/2}l.$$

Now we can apply Proposition 7.2 to each of the $l$ homotopies starting at $(g,\zeta_i)$, $i=1,\dots,l$ as in the proof of Theorem 7.1. To prove an $l$-zero version (change an approximate zero to $l$ approximate zeros and $k\leqslant cN^3/(\sigma^{1-\varepsilon})$ to $k\leqslant cN^3l^2/(\sigma^{1-\varepsilon})$) one factor of $l$ is for the probability estimate the other because we follow $l$ homotopies. The $l$-zero version of Theorem 7.4 follows similarly.

## References

[1] L. Blum, M. Shub and S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, Recursive functions and Universal Machines, *Bull. Amer. Math. Soc.* **21** (1989) 1–46.

[2] F. Cucker, M. Shub and S. Smale, Separation of Complexity classes in Koiran's weak model, preprint, 1993.

[3] M.-H. Kim, Error analysis and bit complexity: polynomial root finding problem, Part I, preprint (Bellcore, Morristown, NJ, 1988).

[4] P. Koiran, A weak version of the Blum, Shub and Smale model, 34th Found. of Comp. Sci. 1993, 486–495.

[5] E. Kostlan, Random polynomials and the statistical fundamental theorem of algebra, preprint, Univ. of Hawaii, 1987.

[6] G. Malajovich-Muñoz, On the complexity of path-following Newton algorithms for solving systems of polynomial equations with integer coefficients, Ph.D. Thesis, University of California at Berkeley, 1993.

[7] F. Morgan, Geometric Measure Theory, A Beginners Guide (Academic Press, New York, 1988).

[8] D. Priest, On properties of floating point arithmetics: numerical stability and cost of accurate computations, Ph.D. Thesis, University of California at Berkeley, 1992.

[9] M. Shub, On the work of Steve Smale on the theory of computation, in: M. Hirsch, J. Marsden and M. Shub, eds., From *Topology to Computation: Proceedings of the Smalefest* (Springer, Berlin, 1993) 281–301.

[10] M. Shub and S. Smale, Computational complexity, on the geometry of polynomials and a theory of cost: Part II, *SIAM J. Comput.* **15** (1986) 145–161.

[11] M. Shub and S. Smale, Complexity of Bezout's theorem I: Geometric aspects *J. Amer. Math. Soc.* **6**, (1993) 459–501.

[12] M. Shub and S. Smale, Complexity of Bezout's theorem II: Volumes and probabilities, in: F. Eyssette and A. Galligo, eds., Computational Algebraic Geometry, Progress in Mathematics, Vol. 109 (Birkhäuser, Basel, 1993) 267–285.

[13] M. Shub and S. Smale, Complexity of Bezout's theorem III: Condition number and packing, *J. Complexity* **9** (1993) 4–14.

[14] M. Shub and S. Smale, Complexity and Bezout's theorem IV: Probability of Success, Extensions, *SIAM J. Numer. Anal.*, to appear.

[15] S. Smale, The fundamental theorem of algebra and complexity theory, *Bull. Amer. Math. Soc.* (N.S.) **4** (1981) 1–36.

[16] S. Smale, Algorithms for solving equations, *Proc. Internat. Congr. Mathematicians*, Berkeley, CA (American Mathematical Society, Providence, RI1986) 172–195.

[17] S. Smale, Some remarks on the foundations of numerical analysis, *SIAM Rev.* **32** (1990) 211–220.

## References added in proof

[18] R. Howard, *The Kinematic Formula in Riemannian Homogeneous Spaces*, Memoirs of the AMS, Vol. 509 (AMS, Providence, RI, 1993).

[19] L.A. Sautalo, *Integral Geometry and Geometric Probability* (Addison-Wesley, Reading, MA, 1976).