ABOUT THE COMPLEXITY OF THE FUNDAMENTAL THEOREM OF ALGEBRA: AVERAGE COMPLEXITY AND INDEPENDENCE ON THE DEGREE.

MYONG-HI KIM
DEPARTMENT OF MATHEMATICS
SUNY-COLLEGE AT OLD WESTBURY, NY.

ABSTRACT. For any polynomial $f \in P_d(1) = \{\prod_{j=1}^{d}(z-\xi_j) \ : \ |\xi_j| \leq 1\}$, we show that the average number of Newton iterations required for the Smale-Newton homotopy method to find an ϵ-root of f is independent of the degree d. More precisely, let $S_{1+1/d} = \{z : |z| = 1+1/d\}$ be endowed with the uniform measure μ with total mass 1. Then on the average with respect to μ, FIND-ROOT defined below finds an ϵ- root of f with

two random choices of initial points,

$100\lceil 1+|\log\epsilon|\rceil$ Newton iterations,

$100\lceil 1+|\log\epsilon|\rceil$ evaluations of f and f',

where the average is taken over initial points chosen from $S_{1+1/d}$.

In particular, when applied to sparse polynomials $f \in P_d(1)$, on the average with respect to μ, FIND-ROOT finds an ϵ-root of f with

$400K\lceil\log d\rceil\lceil 1+|\log\epsilon|\rceil$ arithmetic operations,

where K is the number of nonzero monomials of f.

1. Introduction and Background Material.

We first describe the algorithm FIND-ROOT. Then we will explain the background material. Hereafter we call z an ϵ-root of f if $|f(z)| \leq \epsilon$. We use $\lceil x \rceil$ to denote the smallest integer exceeding x.

Let $D_r = \{z : |z| < r\}$, $S_r = \{z : |z| = r\}$ and μ be the uniform measure on S_r with total mass 1.

FIND-ROOT
INPUT: $f \in P_d(1)$, $1 > \epsilon > 0$. OUTPUT: an ϵ-root of f.

CYCLE:
1. Choose z_0 at random from $S_{1+1/d}$.
2. Let $N = 50\lceil\log|f(z_0)| - \log\epsilon\rceil$, $w_0 = f(z_0)$, $w_n = \left(\frac{49}{50}\right)^n w_0$. For n from 1 to N do

$$z_{n+1} = z_n - \frac{f(z_n) - w_{n+1}}{f'(z_n)}.$$

1991 *Mathematics Subject Classification.* 68Q25; Secondary 58C10, 65H05, 30C15, 58F08.
Key words and phrases. Fundamental Theorem of Algebra, Newton's method, Approximate Zeros, Average Complexity, Path-lifting Method, Analytic Continuation, Möbius Transformation.

TEST:

1. If $|f(z_N)| \leq \epsilon$ then output z_N.
2. Else Goto CYCLE.

END FIND-ROOT

We may view a sequence of initial points used in FIND-ROOT as an infinite sequence in the product space $\mathcal{S} = \prod_{n=1}^{\infty} S_{1+1/d}$. We endow this product space with the product measure of probability induced by μ. To any sequence $\vec{z} = (z_1, z_2, \cdots) \in \mathcal{S}$ we associate two integers: $\mathcal{N}(\vec{z})$ and $\mathcal{T}(\vec{z})$. We have $\mathcal{N}(\vec{z}) = k$ when CYCLE does not output an ϵ-root when starting with $z_1, \cdots, z_{k-1}$ and outputs an ϵ-root when starting with z_k. We may have $\mathcal{N}(\vec{z}) = \infty$ for certain choices of $\vec{z} \in \mathcal{S}$ but we show that this is a zero probability event. Our second integer $\mathcal{T}(\vec{z})$ counts the total number of evaluations of Newton iterations performed during these k CYCLE's.

Our main theorem is an estimate for these two random variables on the average.

Main Theorem . *On the average over $\mathcal{S}$, FIND-ROOT terminates with an ϵ-root of f with two random choices of initial points and $100\lceil 1 + |\log \epsilon| \rceil$ Newton iterations,*

We notice that these estimates are independent on the degree d and on the considered polynomial $f \in P_d(1)$.

An upper bound for the arithmetic complexity of this algorithm, for a model of computation over the real numbers, is obtained in multiplying the number of evaluations of f and f' by the cost of such evaluations. This is an important remark because this cost depends mainly on the sparse structure of f with a modest contribution on the degree. Thus, our estimate differs from classical complexity estimates where the degree appears with an exponent greater than or equal to 1. See Renegar [5], Pan [6], Shub-Smale [10] or Pan's survey [7]. For a degree d sparse polynomial f with K monomials we can evaluate f and its derivative in $4K \log d$ arithmetic operations. According to this remark the following is immediate from our MAIN THEOREM.

Corollary 1. *For a sparse polynomial $f \in P_d(1)$ with K monomials, on the average, FIND-ROOT finds an ϵ-root of f within the arithmetic number of operations*

$$400K\lceil \log d \rceil \lceil 1 + |\log \epsilon| \rceil.$$

Now we review background material extracted from Smale (see Chapter 9 in [1]) and Kim-Sutherland [4]. See also Smale [12] and Shub-Smale [8], [9].

A polynomial f of degree d is a branched covering map of degree d. For a regular point z, i.e. if $f'(z) \neq 0$, there is a well defined inverse branch f_z^{-1} of f^{-1} which maps $f(z)$ back to z.

In this paper, we often relate an analytic continuation method starting at $f_{z_0}^{-1}$ to the approximation method of the solution curve (flow line) of Newton vector field starting at z_0 and vice versa.

Let $\psi_{z_0}(t)$ be the solution curve of the Newton vector field,

$$\dot{z} = -\frac{f(z)}{f'(z)},$$

with $\psi_{z_0}(0) = z_0$.

Notice that the Newton Vector field $-\frac{f(z)}{f'(z)}$ has the same solution curves as the gradient vector field (see Lemma 10, [1])

$$-\nabla\frac{1}{2}|f(z)|^2 = -\frac{f(z)}{f'(z)}|f'(z)|^2.$$

We recall that the gradient vector field is orthogonal to level curves. It is easy to see that the solution curve $\psi_{z_0}(t)$ satisfies

$$f(\psi_{z_0}(t)) = e^{-t}f(z_0),$$

wherever it is defined, i.e. except at critical points of f.

In other words, the image of $\phi_{z_0}(t)$ under f is the ray $[0,\ f(z_0)] = \{\lambda f(z_0),\ 0 \le \lambda \le 1\}$ and the image of this ray under $f_{z_0}^{-1}$ is a solution curve to the Newton vector field wherever it is defined.

When $\psi_z(t)$ is defined for all $t \ge 0$,

$$\psi_z(t) \to \xi \text{ for some root } \xi \text{ as } t \to \infty.$$

In all other cases, $\psi_z(t) \to \theta$ for some critical point θ of f as $t \to T^*$ for some finite number T^*. For convenience, we will call a flow line a regular flow line if it is defined for all $-\infty < t < \infty$, and call its image under f a regular ray. We will call all other flow lines critical flow lines and call the image of critical flow lines under f critical rays. We note that critical flow lines also include those $\psi_z(t) \to \xi$ for some root ξ as $t \to \infty$ but $\psi_z(t) \to \theta$ as $t \to T^*$, for some $T^* < 0$.

We remind the reader that there are at most $d-1$ critical points of a polynomial f with degree d. At each critical point θ_k of multiplicity m_k there are (m_k+1) Newton flow lines converging to θ_k, as illustrated in the example of $f = 1 + z^3$ of which 0 is a critical point with multiplicity 2. Hence the total number of critical flow lines is at most $\sum_k (m_k+1)$. Notice that $\sum_k (m_k+1) \le 2(d-1)$ where the maximum of $2(d-1)$ is obtained when all critical points θ_k have multiplicity 1. Hence there are at most $2(d-1)$ critical flow lines of the Newton vector field applied to a polynomial of degree d. For any point z_0 not on a critical flow line, $\psi_{z_0}(t)$ is defined for all $-\infty < t < \infty$ and $f_{z_0}^{-1}$ can be analytically continued along the ray starting from $f(z_0)$ down to 0. To obtain a root $\xi = f_{z_0}^{-1}(0)$, one simply uses analytic continuation of $f_{z_0}^{-1}$ by following down the ray from $f(z_0)$ to the origin. We recall that all critical points of f are contained in the convex hull of all roots of f. Further, it is well known that the Newton flow lines intersect transversally with the circles lying outside of the root disk (see Lemma 9.3.11, [1]). This gives the following lemma (see Lemma 9.3.13, [1].

Lemma 1. *Let $r > |\xi_j|$ for all roots ξ_j of f. Then at most at $2(d-1)$ points of S_r lie on a critical flow line.*

Let

$$\Theta' = \{\theta'_1, \dots, \theta'_K\},\ K \le 2(d-1) \tag{1}$$

be the points on S_r which lie on critical flow lines. For convenience, we put θ'_j in counter clockwise order. Note that if $z \in S_r$ and $z \notin \Theta'$, then $\psi_z(t)$ is a regular flow line and hence converges to a root as $t \to \infty$. Our study focuses on how to choose initial points away from those bad points on S_r.

Let ρ_j be the open arc on S_r connecting the two adjacent θ'_j and θ'_{j+1}. Then for all $z \in \rho_j$, $\psi_z(t)$ is a regular flow line and f_z^{-1} can be analytically continued along the infinite ray $L_z = \{\lambda f(z) : -\infty < \lambda < \infty\}$. Further it can be analytically continued to an open infinite wedge,

$$\mho_j = \{f(\psi_z(t)) : z \in \rho_j, -\infty < t < \infty\}. \tag{2}$$

Notice that $\mho_j$ is bounded by two rays, one passing through $f(\theta'_j)$ the other passing though $f(\theta'_{j+1})$, i.e.

$$\mho_j = \{w : \arg(f(\theta'_j)) < \arg(w) < \arg f((\theta'_{j+1}))\}.$$

If any subarc $\rho_{x,y}$ of ρ_j bounded by x and y and $z \in \rho_{x,y}$, then f_z^{-1} can be analytically continued to the open infinite wedge bounded by two rays passing through $f(x)$ and $f(y)$. In summary, we have the following lemma.

Lemma 2. *Suppose that $\rho_{x,y} \subset \rho_j$ is an open arc connecting x and y. Then for any point $z \in \rho_{x,y}$, f_z^{-1} can be analytically continued along the wedge $\mho_{x,y}$ bounded by the rays passing through $f(z_1)$ and $f(z_2)$,*

$$\mho_{x,y} = \{w : \arg(f(x)) < \arg(w) < \arg(f(y))\}.$$

f_z^{-1} can be further analytically continued to $\mho_j$.

When using analytic continuation, we would like to choose z_0 away from critical flow lines, so that $\psi_{z_0}(t)$ stays away from the critical flow lines and $f_{z_0}^{-1}$ has a large radius of convergence. This allows us to track ψ_{z_0} more efficiently.

Let $W_{A,y} = \{w \; : \; |w| < 2|y|, \; |\arg(\frac{w}{y})| < A\}$ be the wedge centered at y with angle A.

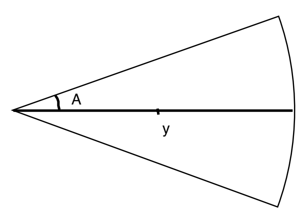

FIGURE 1. $W_{A,y}$

Suppose that $f_{z_0}^{-1}$ is analytic in the wedge $W_{A,f(z_0)}$ centered at $f(z_0)$ with an angle A. Notice that this can happen only if $W_{A,f(z_0)}$ is contained in one of $\mho_j$ defined above. At any point $z = \psi_{z_0}(t)$, $f(z)$ is on the center line of the wedge $W_{A,f(z)}$, and f_z^{-1} is analytic in this wedge with a radius of convergence of at least $|f(z)| \sin A$. Now we try to approximate the point on $\psi_{z_0}(t)$ by solving

$$f(z) - w = 0$$

for w inside of the disk of convergence of f_z^{-1} and also near enough to $f(z)$ so that a single Newton iteration gives a good approximation of $f - w$. It turns out that it is enough to take w with $|w - f(z)|/|f(z)| \geq (\sin A)/19$. More precisely, the following lemma is proven in Smale [11] but we state a somewhat sharper version from Kim-Sutherland [4] Theorem 1.5.

Lemma 3. *Suppose that $f_{z_0}^{-1}$ is analytic in the wedge $W_{A,f(z_0)}$ with an angle A or larger. Let $w_0 = f(z_0)$, $h \leq (\sin A)/19$, $w_n = (1-h)^n$, and*

$$z_{n+1} = z_n - \frac{f(z_n) - w_{n+1}}{f'(z_n)}.$$

Then $|f(z_n) - w_n| \leq hw_n/2$ for all n. In particular,

$$|f(z_n)| \leq (1-h)^n |f(z_0)| \text{ and } |f(z_N)| \leq \epsilon$$

for any $N \geq (\log|f(z_0)| - \log\epsilon)/h$.

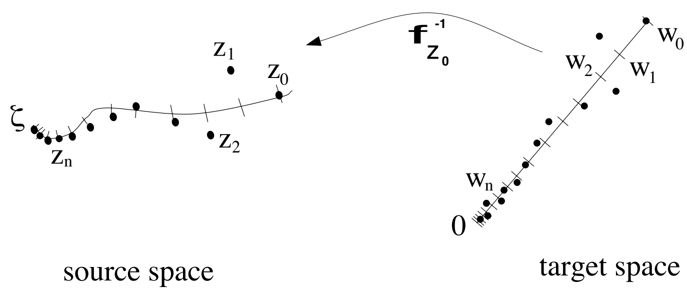

FIGURE 2. Each z_i and $f(z_i)$ are indicated by a black dot in the source and target space respectively. The w_i and $f_{z_0}^{-1}(w_i)$ are each marked by ticks.

Corollary 2. *The step size and the number iterations for the selected angles are given below.*

Angle	*step size h*	*number of iterations N*
$\frac{\pi}{4}$	$h = 1/27$	$27(\log|f(z_0)| - \log\epsilon)$.
$\frac{\pi}{8}$	$h = 1/50$	$50(\log|f(z_0)| - \log\epsilon)$.
$\frac{\pi}{12}$	$h = 1/74$	$74(\log|f(z_0)| - \log\epsilon)$.

¿From now on, we will call z a good initial point if f_z^{-1} is analytic in $W_{A,f(z)}$ with an angle $A \geq \pi/8$. If z is a good initial point, then the homotopy defined in the above lemma converges to a root and hence the CYCLE in FIND-ROOT outputs an ϵ-root. Notice that the iterations defined in CYCLE in FIND-ROOT are those defined in Lemma 3 with $A = \pi/8$, i.e.

$$z_0 = z,\ z_{n+1} = z_n - \frac{f(z_n) - w_{n+1}}{f(z_n)},\ w_n = \left(\frac{49}{50}\right)^n.$$

Our algorithm FIND-ROOT is based on the probability estimate in Corollary 3 in §2 below which says that one out of two randomly chosen points from $S_{1+1/d}$ is a good initial point. In [11], Smale analyzes a similar algorithm based on Lemma 3 which chooses the initial points from the circle of radius 3 for $f \in \{\sum_0^d a_j z^j, |a_j| \leq 1\}$, noting that all the roots of such an f lie in $D_2(0)$. There, radius 3 is chosen to control the wedge angle A between

$f(z_0)$ and the critical values. However, for $|z_0| = 3$, $|f(z_0)| \sim 3^d$ and hence a factor of $d \log 3$ appears in the number of iterations.

In this paper, we analyze the same algorithm when choosing initial points from $S_{1+\frac{1}{d}}$, for f in $P_d(1)$. This choice avoid this extra d factor in our complexity analysis.

The focus of the next section is to estimate the measure of the set of z on S_r such that f_z^{-1} is analytic in $W_{A,f(z)}$, for a given angle A.

2. Lemmas and Proofs.

Recall that z is a good initial point if f_z^{-1} is analytic in $W_{A,f(z)}$ with the angle $A \geq \pi/8$. One of our goals is to estimate the measure of the set of good initial points on $S_{1+1/d}$. The following lemma estimates the rate of change of the argument of $f(z)$ as z travels around $S_{1+1/d}$.

Lemma 4. *Let $r = 1 + \frac{1}{d}$. For $z, z' \in S_r$, suppose that*

$$\arg(z) - \arg(z') = B \geq 0.$$

Then we have

$$\frac{d(d+1)}{2d+1} B \leq \arg(f(z)) - \arg(f(z') \leq d(d+1)B.$$

Proof. First note that

$$\log(f(z)) = log|f(z)| + i \arg f(z).$$

Let us denote $\rho(z) = \Re(\log(f(z)))$ and $\phi(z) = \arg(f(z)) = \Im(\log(f(z)))$. On $|z| = r$, with an abuse of notation, we will also write $\phi(t)$, $\rho(t)$ to denote $\phi(re^{it})$, $\rho(re^{it})$. Then its derivative around the circle $z = re^{it}$ is

$$\frac{d}{dt} \log(f(re^{it}) = \rho'(t) + i\phi'(t) = \frac{izf'(z)}{f(z)} = -\Im \frac{zf'(z)}{f(z)} + i\Re \frac{zf'(z)}{f(z)}.$$

Since $zf'(z)/f(z) = \sum_{j=1}^{d} z/(z - \xi_j)$ counting multiplicity, we have

$$\phi'(t) = \Re \left(\sum_{j=1}^{d} \frac{z}{z - \xi_j} \right), \quad \rho'(t) = -\Im \left(\sum_{j=1}^{d} \frac{z}{z - \xi_j} \right).$$

We apply a Möbius transformation to each term of $\phi'(t)$ to obtain the bounds on $\phi'(t)$. Write $\xi_j = R_j e^{i\alpha_j}$, where $R_j \leq 1$. Let

$$\omega_j(t) = \frac{1}{1 - \frac{R_j}{r} e^{it}}, \quad \delta_j(t) = \frac{z}{z - \xi_j} = \frac{1}{1 - \frac{\xi_j}{z}}.$$

Note that

$$\delta_j(t) = \omega_j(-(t - \alpha_j)).$$

Recall that a Möbius transformation maps a circle to a circle possibly with ∞ radius. Then ω_j maps the real axis to the real axis and the unit circle to the circle passing through the two points

$$\omega_j(\pi) = \frac{d+1}{d+1+dR_j} \geq \frac{d+1}{2d+1}, \qquad \omega_j(0) = \frac{d+1}{2d+1} \leq d+1,$$

and

$$\frac{d+1}{d+1+dR_j} \le \Re(\omega_j(t)) \le d+1,$$

for all t, where "equality" holds only when $R_j = 1$.

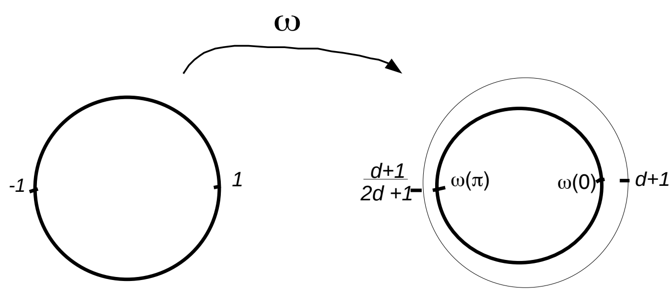

FIGURE 3.

Since $\phi'(t) = \sum_{j=1}^{d} \Re(\omega_j(-(t-\alpha_j)))$ we have

$$\frac{d(d+1)}{2d+1} \le \phi'(t) \le d(d+1).$$

This gives us the claim since $\phi(t) = \arg(f(re^{it})$ and

$$\min \phi'(t)|\arg(z) - \arg(z')| \le |\arg f(z) - \arg f(z')| \le \max \phi'(t)|\arg(z) - \arg(z')|.$$

□

Of particular interest to us is when $z' = \theta'_j \in \Theta'$. We recall that $\Theta' = \{\theta'_1, \dots, \theta'_K\}$, $K \le 2(d-1)$, is the set of points on S_r which lie on critical flow lines. We put θ'_j in counter clockwise order.

Lemma 5. *Let*

$$Q_A = \{z \in S_r : f_z^{-1} \textit{ can be analytically continued to } W_{A,f(z)}\}.$$

Then for any $0 < A \le \pi/4$,

$$\mu(Q_A) \ge 1 - \frac{8A}{2\pi}.$$

Proof. Let

$$B_A = \{z \in S_r : f_z^{-1} \text{ can not be analytically continued to } W_{A,f(z)}\}.$$

We notice from Lemma 2 that f_z^{-1} can be analytically continued to $W_{A,f(z)}$ if $W_{A,f(z)} \subset \mho_j$ for some j. In other words, f_z^{-1} is not analytic in $W_{A,f(z)}$ only if $|\arg(f(z)) - \arg(f(\theta'_j))| < A$, for some $\theta'_j \in \Theta'$. By Lemma 4, this can happen only if $|\arg(z) - \arg(\theta'_j)| < \frac{2d+1}{d(d+1)}A$ for some j. Hence $B_A \subset \cup_j\{|\arg(z) - \arg(\theta'_j)| < \frac{2d+1}{d(d+1)}A\}$. Since there are at most 2(d-1) such θ'_j we have

$$\mu(B_A) \le \sum_{\theta'_j} \mu\{z \in S_{1+1/d} : |\arg(z) - \arg(\theta'_j)| <$$

$$\frac{2d+1}{d(d+1)}A\} < 2(d-1)2\frac{2d+1}{d(d+1)}\frac{A}{2\pi} < \frac{8A}{2\pi},$$

recalling that μ has a total mass 1 on $S_{1+1/d}$. Consequently we have

$$\mu(Q_A) \geq 1 - \frac{8A}{2\pi}.$$

This completes the proof. □

Corollary 3. *Below we list the lower bounds for the measures of the sest Q_A. Of special interest to us is the set of good initial points,*

$$\mu(Q_{\pi/8}) = \mu\{z \in S_{1+1/d} : z \textit{ is a good initial point }\} \geq 1/2.$$

Angle in radian	*lower bound of $\mu(Q_A)$*
$\pi/6$	*1/3*
$\pi/8$	*1/2*
$\pi/12$	*2/3*

Remark 1. *Note that $z \in Q_A$ is a sufficient condition for the iterational process defined as in Lemma 3 with an angle A to terminate with an ϵ-root, but it is not a necessary condition.*

The following lemma estimates the average value of $|f(z)|$ over the circle of $|z| = r > 1$, when all the roots are inside the unit disk $\overline{D_1}$. Let $r > |\xi_j|$, for all roots ξ_j of f.

Lemma 6.

$$\int_0^1 \log|(f(re^{2\pi it})|dt = d\log r.$$

Proof. Since f has all its roots in $\overline{D_1}$,

$$\frac{f(z)}{z^d} = 1 + O(\frac{1}{z}) \text{ as } |z| \to \infty,$$

so that

$$\log\frac{f(z)}{z^d} = O(\frac{1}{z})$$

is holomorphic outside of $\overline{D_1}$. Now,

$$\frac{1}{z}(\log f(z) - \log z^d) = O(\frac{1}{z^2})$$

$$\int_L (\log f(z) - \log z^d)\frac{dz}{z} = 0,$$

where L is any loop in $\{|z| > 1\}$. Take $z = re^{2\pi it}$, then $\frac{dz}{z} = 2\pi i\, dt$. By taking the real part, we have

$$\int_0^1 \log|f(re^{2\pi it})|dt - \int_0^1 \log|z|^d dt = 0.$$

Since $\int_0^1 \log r^d dt = d\log r$, we have the claim. □

Remark 2. *For $f \in P_d(1)$, $r = 1 + 1/d$, $d \geq 2$, notice that*

$$\int_0^1 \log |(f(re^{2\pi it})|dt = d\log(1 + \frac{1}{d}) \leq 1.$$

In fact, $d\log(1+\frac{1}{d}) \uparrow 1$ as $d \uparrow \infty$.

Proof of Main Theorem. Let G be the set of $z \in S_{1+1/d}$ such that a CYCLE with an initial point z outputs an ϵ-root. Let us denote $B = S_{1+1/d} - G$, $\mu(G) = p$ and $\mu(B) = q = 1-p$. Since any good initial point is in G we have that $p \geq 1/2$ by Corollary 3. Since the selection of an initial point in each cycle is an independent of the prior selections, each cycle is an independent event. Hence every CYCLE terminates with an ϵ-root with the probability p and fails with probability $q = 1 - p$.

We may view a sequence of initial points chosen in FIND-ROOT, as $\vec{z} = (z_1, z_2, \cdots) \in \mathcal{S} = \prod_{n=1}^{\infty} S_{1+1/d}$. We endow this product space with the product measure of probability induced by μ. To any sequence $\vec{z} = (z_1, z_2, \cdots) \in \mathcal{S}$ we associate two integers: $\mathcal{T}(\vec{z})$ and $\mathcal{N}(\vec{z})$. We have $\mathcal{N}(\vec{z}) = k$ when CYCLE does not output an ϵ-root when starting with $z_1, \cdots, z_{k-1}$ and outputs an ϵ-root when starting with z_k. Our second integer $\mathcal{T}(\vec{z})$ counts the total number of Newton iterations performed during these k CYCLE's.

Let Ω_k be the set of sequences $\vec{z} \in \mathcal{S}$ such that the first occurrence of $z_j \in G$ is k, i.e.

$$\Omega_k = \{\vec{z} \;:\; z_1 \notin G, \cdots, z_{k-1} \notin G,\; z_k \in G\} = B^{k-1} \times G \times \prod_{n=k+1}^{\infty} S_{1+1/d}.$$

Then for $\vec{z} \in \Omega_K$, we have

$$\mathcal{N}(\vec{z}) = k \;\text{ and }\; \mathcal{T}(z_1, \cdots, z_n, \cdots) = \sum_{n=1}^{K} T(z_n),$$

where $T(z)$ is the number of Newton iterations FIND-ROOT uses when starting with the initial point z. We note that $\mu(\Omega_k) = q^{k-1}p$. We also note that $\mathcal{S}$ is equal to the disjoint union of the sets Ω_k up to a set of measure equal to zero. Thus

$$\int \mathcal{N}(\vec{z})\, d\vec{z} = \sum_{k=1}^{\infty} \int_{\Omega_k} k\, d\vec{z} = \sum_{k=1}^{\infty} k\, \mu(\Omega_k) = \sum_{k=1}^{\infty} k\, q^{k-1}\, p = \frac{1}{p} \leq 2.$$

For the second part, notice that

$$T(z) = 50(log|f(z)| - \log \epsilon)$$

by Corollary 2 and

$$\int_S T(z)\, dz = 1 - \log \epsilon$$

by Remark 2. This gives

$$\begin{aligned}
\int \mathcal{T}(\vec{z})\, d\vec{z} &= \sum_{k=1}^{\infty} \int_{\Omega_k} \sum_{j=1}^{k} T(z_j)\, d\vec{z} = \sum_{k=1}^{\infty} \int_{B^{k-1}\times G\times \prod_{k+1}^{\infty} S_{1+1/d}} \sum_{j=1}^{k} T(z_j)\, d\vec{z} \\
&= \sum_{k=1}^{\infty} \left(\sum_{j=1}^{k-1} \int_B T(z_j)\, dz_j\, \mu \left(B^{k-2} \times G \times \prod_{k+1}^{\infty} S_{1+1/d} \right) \right) \\
&+ \int_G T(z_k)\, dz_k\, \mu \left(B^{k-1} \times \prod_{k+1}^{\infty} S_{1+1/d} \right) \\
&= \sum_{k=1}^{\infty} \left(\sum_{j=1}^{k-1} q^{k-2} p \int_B T(z_j)\, dz_j \right) + q^{k-1} p \int_G T(z_k)\, dz_k \\
&= \frac{p}{(1-q)^2} \int_B T(z)\, dz + \frac{1}{1-q} \int_G T(z)\, dz \\
&= \frac{1}{p} \left[\int_G T(z)\, dz + \int_B T(z)\, dz \right] = \frac{1}{p} \int_{S_{1+1/d}} T(z)\, dz \\
&\le 2 \cdot 50(1 - \log \epsilon) = 100(1 - \log \epsilon).
\end{aligned}$$

Since each Newton iteration evaluates one f and one f' $\mathcal{T}(\vec{z})$ counts the total number of evaluations of f and f' and we are done. □

Proof of Corollary. Compute $z, z^2, \cdots, z^{2^{\lceil \log d \rceil - 1}}$ with $\lceil \log d \rceil$ multiplications. Then computing z^n costs at most $\lceil \log d \rceil$ multiplications, except $z^{2^{\lceil \log d \rceil} - 1}$ which costs $\lceil \log d \rceil$. Then there are $K-1$ additions and K multiplications by coefficients in evaluation of a polynomial. Hence evaluation of f costs at most $2K\lceil \log d \rceil$ arithmetic operations. Each iteration evaluates f and f', hence the average number of arithmetic operations used is less than or equal to $4K\lceil \log d \rceil \times (100 + 100|\log \epsilon|) = 400K\lceil \log d \rceil (1 + |\log \epsilon|)$. □

References

[1] L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and Real Computation* , Springer (1997).

[2] J. Friedman, *Random Polynomials and Approximate Zeros of Newton's Method* , SIAM J. Comput., (1990), pp 1068-1099.

[3] J. Hubbard, D. Schleicher, S. Sutherland, *How to really Find Roots of Polynomials by Newton's Method*, preprint, 1998.

[4] M. Kim, S. Sutherland, *Polynomial root-finding algorithms and branched covers* , SIAM J. Comput. Vol. 23, No. 2 (1994), pp 415-436.

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

[6] V. Pan, *Optimal (up to polylog factors) sequential and parallel complexity of approximate evaluations of polynomial zeros*, 27th Annual ACM Symp. on the Theory of Computing, pp 741-750.

[7] V. Pan, *Solving a Polynomial Equation: Some History and Recent Progress*, SIAM Review, 39 (1997) pp 187-220.

[8] M. Shub, S. Smale, *Computational Complexity: On the geometry of polynomials and a theory of cost I*, Ann. Sci. Ecole Norm. Sup., 18 (1985), pp 107-142.

[9] M. Shub, S. Smale, *Computational Complexity: On the geometry of polynomials and a theory of cost II*, SIAM J. Comput., 15 (1986), pp 145-161.
[10] Shub, M. and S. Smale, *Complexity of Bézout's Theorem V : Polynomial Time*, Theoretical Computer Science, 133 (1994) 141-164.
[11] S. Smale, *The Fundamental Theorem of Algebra and Complexity Theory, Part I*, Bull. Amer. Math Soc., 4 (1981), pp 1-36.
[12] S. Smale, *On the efficiency of algorithms of analysis*, Bull. Amer. Math Soc., 13 (1985), pp 87-121.