数理解析研究所講究録
1179 巻 2000 年 83-93

On the Number of Iterations of Newton's Method for Complex Polynomials

Dierk Schleicher

October 16, 2000

Abstract

We use methods from conformal geometry to give an explicit estimate for how many iterations of Newton's method it takes to find all the roots of an arbitrary complex polynomial of fixed degree to prescribed precision.

1 Introduction

It is a fundamental problem to find all roots of a complex polynomial p with prescribed precision. The best known method is known as Newton's method: starting with an arbitrary starting point $z \in \mathbb{C}$ iterate the Newton map $z \mapsto N_p(z) = z - p(z)/p'(z)$ until the iterates are sufficiently close to a root. Newton's method is widely used because of its simplicity and because it is efficient in practice: near a simple root, the convergence is quadratic (the number of valid digits doubles in every iteration step), and for "most" starting points of "most" polynomials, it does not take too long to find good approximations of a root (see [Sm] for a survey).

There are problems, though: there are starting points which never converge to a root under the Newton map (for example all the points on the boundaries of the basins of all the roots). There may even be non-empty open sets of starting points which converge to attracting periodic orbits away from the roots. Manning [Ma] has constructed an explicit set of $O(d\log^2 d)$ points of which at least one converges to a root, and it takes $d\log(d^3/\varepsilon)$ Newton iterations to get to distance at most ε from a root. Even if most starting points converge to some root, one often wants to find all the roots (finding some roots and factoring out is not an option for various reasons: it is numerically very instable, so that after several long divisions the roots of the quotient polynomials are far from the correct roots; moreover, the polynomial may be given for example by an iteration which is easy to evaluate but difficult to convert into coefficient form). In [HSS], a set of starting points has been constructed such that every root is found by at least one of these points. This set can contain as little as $1.11\, d\log^2 d$ points.

It has been an open question to find an explicit upper bound for the number of iterations it takes to find all the roots with prescribed precision for arbitrary polynomials using Newton's method. (There are other methods for which efficient bounds on the

[0]**Keywords:** Newton's method, root finding, complex polynomial
[0]**MSC Classification (2000):** 30C10, 30C15, 37F10, 37N30
[0]Dierk Schleicher, Mathematisches Institut, Ludwig-Maximilians-Universität, Theresienstraße 39, D-80333 München, dierk@rz.mathematik.uni-muenchen.de

"cost" of finding all the roots are known [NR], and there are probabilistic results for Newton's method [Mn].)

In our paper, we provide a bound for the number of Newton iteration steps to find all the roots with prescribed precision. This bound is far from efficient, but there is hope that it can be substantially improved; we outline where the efficiency is lost. Here is our main result.

Theorem 1 (The Number of Iterations)
For every degree d and every accuracy $\varepsilon > 0$, there is a number $n = n(d,\varepsilon)$ with the following property: for every polynomial p of degree d, normalized so that all the roots are in $\mathbb{D}$, and every root α of p, there is a point z with $|z| = 2$ such that after at most n iterations of the Newton map, the point z has distance at most ε from α; we have the estimate

$$n = n(d,\varepsilon) \le \frac{9\pi d^4 f_d^2}{\varepsilon^2 \log 2} + 2 + \frac{|\log\varepsilon| + \log 13}{\log 2}$$

with $f_d := \dfrac{d^2(d-1)}{2(2d-1)}\begin{pmatrix} 2d \\ d \end{pmatrix}$.

This estimate is clearly far from effective. For $p(z) = z^d$, we have $N_p(z) = (d-1)z/d$, so for $|z| = 2$ it takes approximately $n = d\log(2/\varepsilon)$ Newton iterations so that $|N_p^{\circ n}(z)| < \varepsilon$. We believe that this is the worst case.

All our polynomials will always be normalized so that their roots are in $\mathbb{D}$; this is no real restriction because it is easy to estimate the largest absolute value of the roots, allowing to rescale the discussion.

We will use p for a polynomial and N_p for its associated Newton map, and $d \ge 2$ will always denote the degree of p.

The paper is organized as follows: in Section 2, we briefly review needed results from [HSS]. We then introduce in Section 3 "clusters of roots", which help to estimate the distance from a given $z \in \mathbb{C}$ to the roots of p. In Section 4 we show that for every root, there is a nearby point in its immediate basin which can be connected to a point z with $|z| = 4$ by a curve well inside the immediate basin (at definite distance to the boundary), and we conclude the proof of the theorem. Finally, in Section 5, we discuss where the efficiency is lost, and how close Newton's method is from being an algorithm.

ACKNOWLEDGEMENTS. I have enjoyed many helpful discussions on Newton's method with Walter Bergweiler, Hartje Kriete, Steffen Rohde, Scott Sutherland, and particularly with John Hubbard. The results in this paper were inspired by discussions with Anthony Manning during a PRODYN Euro-conference in Warsay (1999). I would like to thank them all. Special thanks go to Niklas Beisert for a suggestion about the proof of Lemma 9.

2 Background

For every root α of p, the *basin of attraction* is the open set of points $z \in \mathbb{C}$ such that $N_p^{\circ n}(z) \to \alpha$ as $n \to \infty$; the *immediate basin* is the connected component containing α. It is known [Pr, Sh] that every immediate basin U is simply connected, hence conformally isomorphic to the unit disk $\mathbb{D}$. It carries a hyperbolic metric, normalized so that the curvature is -1 everywhere; we will denote it by $d_U(z_1, z_2)$ or simply $d(z_1, z_2)$.

By [HSS, Proposition 6], every immediate basin has some finite number $m \geq 1$ of *channels* to infinity: a channel of the immediate basin U of a root α is an unbounded connected component of $U \setminus \overline{\mathbb{D}}$, and it determines a homotopy class of curves $\gamma\colon [0,1] \to \overline{U}$ with $\gamma(0) = \alpha$, $\gamma(1) = \infty$ and $\gamma([0,1[) \subset U$; see Figure 1. The "width" of such a channel is measured by its *modulus:* near ∞, the quotient of the channel by the dynamics of N_p is conformally isomorphic to a complex annulus, that is the following quotient

$$\{z \in \mathbb{C} : |\mathrm{Im}(z)| < \pi/2\}/(z \equiv z + \tau)\}$$

for some $\tau > 0$. This quotient is a cylinder of height π and circumference τ, and the modulus of the channel is π/τ, the conformal modulus of the cylinder. It is well known that the unique simple closed geodesic in the cylinder (with respect to its hyperbolic metric) has hyperbolic length τ. In [HSS, Proposition 7], it is shown that every root has a channel with modulus at least $\pi/\log d$.

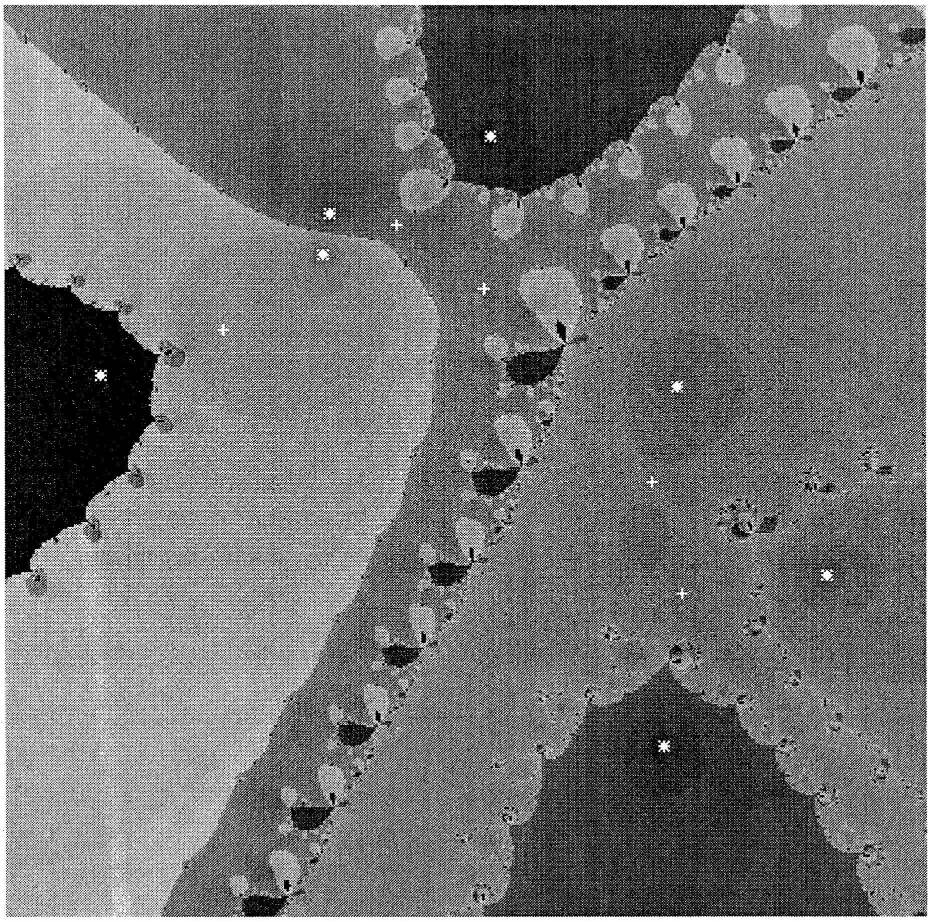

Figure 1: *The dynamics of the Newton map for a polynomial of degree 7: shown are the basins of attraction of the seven roots; it can be seen that the immediate basin of every root has at least one "channel to ∞" along which the immediate basin stretches out to ∞.*

3 Distance to the Root

We define the *round annulus* $A_a(r,R)$ to be the region

$$A_a(r,R) := \{z \in \mathbb{C} : r \leq |z-a| \leq R\} .$$

We say that $z \in \mathbb{C}$ with $|z-a| < r$ is *inside* $A_a(r,R)$ or *surrounded by* $A_a(r,R)$.

Definition 2 (Cluster of Roots)
Suppose that $A_a(r,R)$ contains no roots of p. Let n and N be the number of roots of p inside resp. outside of A such that $n \geq 2$ (with $n+N=d$). If

$$\frac{R-r}{r} \geq \frac{2N}{n-1} \qquad \text{or equivalently} \qquad R \geq \left(\frac{2N}{n-1}+1\right) r ,$$

then we say that the roots inside $A_a(r,R)$ form a cluster of roots.

Theorem 3 (Cluster and Immediate Basins)
Let A be an annulus surrounding a cluster of roots. Let α be some root with immediate basin U. If there is a $z \in U$ which is surrounded by A, then α is also surrounded by A.

PROOF. We write the Newton map as $N_p(z) = z - 1/\tau(z)$, where $\tau\colon \mathbb{C} \to \overline{\mathbb{C}}$ is defined as

$$\tau(z) := \frac{p'(z)}{p(z)} = \sum_{i=1}^{d} \frac{1}{z-\alpha_i} \tag{1}$$

and $\alpha_1, \ldots, \alpha_d$ are the roots of p. The Möbius transform $w \mapsto 1/(r-w)$ maps $\{z \in \mathbb{C}\colon |z| = r\}$ to the vertical line at real part $1/2r$, and it maps $\{z \in \mathbb{C}\colon |z| < r\}$ to the right of this line.

Suppose that A is centered around the origin (i.e. $a = 0$) with inner radius r and outer radius R, and suppose that $z = r$ is real. Then the sum in (1), restricted to roots with $|\alpha_i| < r$, has real part greater than $n/2r$. The remaining roots with $|\alpha_i| > R$ contribute less than $N/(R-r)$, so that $\mathrm{Re}(\tau(z)) > n/2r - N/(R-r)$. If $(R-r)/r \geq 2N/(n-1)$, then $N/(R-r) \leq (n-1)/2r$ and $\mathrm{Re}(\tau(z)) > 1/2r$, so $1/\tau(z)$ is in the disk around r with radius r and $|N_p(z)| < r$.

This means that when $a = 0$, the point $z = r$ on the inner boundary of A is mapped into the bounded complementary component of A. In general, the Newton map respects affine coordinate changes, so we may always suppose that A is centered at 0 and z is on the positive real axis. Under the conditions of the theorem, it follows that any z on the inner circular boundary of A maps into the inner complementary component of A.

If z is in the immediate basin of a root α outside A, then connect z to $N_p(z)$ by a curve within the immediate basin of attraction; iterating forward, we obtain a curve γ from z to a small neighborhood of α. Let w be the last intersection point of γ with the inner boundary of A (so that γ restricts to a curve from w to a neighborhood of α which meets the inner boundary of A only in w). Then $N_p(w)$ is on γ in the bounded complementary component of A, and this is a contradiction. □

Lemma 4 (Near Root Implies Convergence)
For any root α which has distance at least ε to all other roots, the closed disk of radius $\varepsilon/2(d-1)$ around α is contained in the immediate basin of α.

PROOF. This is easy to check using the ideas in the proof of Theorem 3: choosing coordinates so that $\alpha = 0$, $\varepsilon = 2(d-1)$ and $z = 1$, the term $(z-\alpha)^{-1}$ equals 1, while the remaining $d-1$ terms in (1) have absolute value at most $1/2(d-1)$ each or $1/2$ combined, so $\tau(z)$ is in the closed disk around 1 of radius $1/2$ and $1/\tau(z)$ is in a disk with real center which intersects $\mathbb{R}$ in $[2/3, 2]$, and $N_p(z) = z - 1/\tau(z) \in \overline{\mathbb{D}}$. The case $|N_p(z)| = 1$ is possible only if $N_p(z) = -1$ and $\alpha_i = 2(d-1)$ for all $\alpha_i \neq \alpha$, and then $|N_p(-1)| < 1$. □

Lemma 5 (Short Displacement is Close to Root)
If a point z is displaced by ε under the Newton map, then the Euclidean distance between z and some root is at most $d\varepsilon$. Moreover, if z is in an immediate basin, then the Euclidean distance between z and the root it converges to is at most $f_d\,\varepsilon$ (with f_d as in Theorem 1).

PROOF. If all $|z - \alpha_i| > d\varepsilon$, then $\sum(z-\alpha_i)^{-1} < 1/\varepsilon$, and the displacement is more than ε. This proves the first statement.

Suppose that z is in the immediate basin of some root α. Then some root α_1 is at most $d\varepsilon$ away from z. If $\alpha_1 = \alpha$, then we are done; otherwise, there must be another root at distance at most $2d(d-1)\varepsilon$ from α_1 by Lemma 4. Let α_2 be a closest such root. If $\alpha_2 = \alpha$, we are done again; otherwise, by Theorem 3, there must be a third root at distance at most $2d(d-1)\varepsilon[2d-3]$, and then a fourth root at most at distance $2d(d-1)\varepsilon[2d-3][d-2]$, ..., and finally, the last d-th root is at most at distance

$$
\begin{aligned}
&2d(d-1)\varepsilon\prod_{n=2}^{d-1}\left[2\frac{d-n}{n-1}+1\right] = 2d(d-1)\varepsilon\prod_{n=2}^{d-1}\frac{2d-n-1}{n-1}\\
&= 2d(d-1)\varepsilon\frac{(2d-3)!}{(d-1)!(d-2)!} = 2d(d-1)\varepsilon\frac{d^2(d-1)}{2d(2d-1)(2d-2)}\cdot\frac{(2d)!}{d!\,d!}\\
&= \frac{d^2(d-1)}{2(2d-1)}\binom{2d}{d}\varepsilon = f_d\,\varepsilon\ .
\end{aligned}
$$

Therefore, if a point z moves by at most ε in some iteration step and z is in an immediate basin, then the Euclidean distance between z and its root is at most $f_d\varepsilon$ (without any required normalization of the polynomial). □

4 Thick Curves in Channels

Lemma 6 (Hyperbolic Displacement in Iteration)
For every $r \geq 1$, every immediate basin U contains at least one point z with $|z| = r$ such that $d_U(z, N_p(z)) \leq \log d$.

PROOF. Fix a channel with modulus at least $\pi/\log d$ (it exists by [HSS, Proposition 6]; compare Section 2). Restricting the channel to a neighborhood of infinity and taking a quotient by the dynamics of N_p, we obtain an annulus in which the unique simple closed geodesic has length at most $\log d$. Let

$$Q := \{z \in \mathbb{C} : |\mathrm{Im}(z)| < \pi/2, 0 \leq \mathrm{Re}(z) \leq \log d\}\ ,$$

with both vertical ends identified, be a conformal model for the quotient of the channel, and let

$$Q' := \{z \in : |\mathrm{Im}(z)| < \pi/2, 0 \le \mathrm{Re}(z)\}$$

be a semi-infinite cover of Q with natural projection map q. For every $w_0 \in Q$ and every $\varepsilon > 0$, there is a $w_0' \in q^{-1}(w)$ such that in the hyperbolic metric of Q', we have $d_{Q'}(w_0', w_0' + \log d) < \log d + \varepsilon$: if Q' was the universal cover, then the projection map would be a local isometry; in our case, Q' is a truncation of the universal cover. If w_0' is chosen sufficiently far to the right, then the truncation affects only very far parts of the universal cover, and consequently the hyperbolic distance $d_{Q'}(w_0', w_0' + \log d)$ is increased by less than ε.

There is a conformal embedding $j: Q' \to U$ such that $j(w) = N_p(j(w + \log d))$ for all $w \in Q'$ (the right end of Q' corresponds to $\infty \in \partial U$). Since conformal embeddings never increase hyperbolic distances, the point $z_0' := j(w_0' + \log d)$ has $d_U(N_p(z_0'), z_0') < \log d + \varepsilon$, and we may suppose that $|z_0'| > r$. Now we iterate N_p until $z' := N_p^{\circ n}(z_0')$ has $|z'| \le r$; since the conformal map N_p can only decrease hyperbolic distances, we have $d_U(z', N_p(z')) < \log d + \varepsilon$; in fact, by adjusting z_0' properly on the hyperbolic geodesic between z_0' and $N_p(z_0')$, we can arrange things so that $|z'| = r$.

This is possible for every $\varepsilon > 0$. As $\varepsilon \searrow 0$, we obtain a sequence of points $z'(\varepsilon) \in U$ with $|z'(\varepsilon)| = r$. Let z'' be any limit point of $z'(\varepsilon)$; then $z'' \in \overline{U}$ and $|z''| = r$. Since points z near ∂U have large $d_U(z, N_p(z))$, any limit point z'' is in U and satisfies the claim. □

Lemma 7 (Hyperbolic and Euclidean Length)
Let $U \subset \mathbb{C}$ be conformally isomorphic to $\mathbb{D}$ and let $\gamma: [a,b] \to U$ be a continuously differentiable curve such that

$$\inf_{t \in [a,b]} \Big\{ |\gamma(t), \partial U| \Big\} = \delta\ .$$

If γ has finite Euclidean length ℓ, then the hyperbolic length of γ is at least $\frac{\log(1+\ell/\delta)}{2}$.

PROOF. Let ρ be the density of the hyperbolic metric with respect to the Euclidean metric, that is the ratio of infinitesimal hyperbolic and Euclidean lengths. Then for any $z \in U$ at distance r from ∂U, we have $1/2r \le \rho(z) \le 2/r$. Suppose that $\gamma: [0,u] \to U$ is parametrized by arc length, so that $|\gamma'| = 1$. Suppose first that the endpoint $\gamma(0)$ is closest to ∂U, hence $|\gamma(0) - \partial U| = \delta$. Then the hyperbolic length of γ is

$$\int_0^\ell \rho(\gamma(t))\,|\gamma'(t)|\,dt \ge \int_0^\ell \frac{1}{2}\,|\gamma(t) - \partial U|^{-1}\,dt \ge \frac{1}{2}\int_0^\ell (\delta + t)^{-1} dt$$
$$= \frac{1}{2}(\log(\ell + \delta) - \log(\delta)) = \frac{1}{2}\log((\ell+\delta)/\delta)\ .$$

If $\gamma(0)$ is not closest to ∂U, cut γ at one of its closest points to ∂U and estimate the hyperbolic length of one part as above. The hyperbolic length of the other part can only be more than what our estimate for the rest would say, because it cannot be further away from the boundary than estimated. □

REMARK. The estimate in the Lemma 7 is realized for the Koebe extremal domain $\mathbb{C} \setminus \mathbb{R}_0^-$, where γ is part of $\mathbb{R}^+$. This is really the worst case: if the Euclidean distance of γ to ∂U is comparable to δ during most of its length, then the hyperbolic length of γ is comparable to ℓ/δ.

Lemma 8 (Thick Curve in Channel)
For $\varepsilon > 0$, let $\delta := \varepsilon/(f_d d^2)$ with f_d from Theorem 1. Then for every root α with immediate basin U and every $r \geq 1$, there are two points $z, w \in U$ with the following properties: $|z| = r$, $|w - \alpha| < \varepsilon$, w is on the N_p-orbit of z, and there is a curve within U connecting z and w which has everywhere distance at least δ from ∂U.

PROOF. By Lemma 6, there is a point z with $|z| = r$ and $d_U(z, N_p(z)) \leq \log d$. Let $z_1 := z$, let (z_n) be its orbit under the Newton iteration and let z_N be the first point on this orbit with $|z_N - \alpha| < \varepsilon$. By Lemma 5, $|z_n - z_{n+1}| > \varepsilon/f_d$ for $1 \leq n < N$. Let γ_n be a differentiable curve of hyperbolic length $\log d$ joining z_n and z_{n+1}; its Euclidean length is then at least $|z_n - z_{n+1}|$. If δ_n is the minimal distance of γ_n to ∂U, then by Lemma 7, we have

$$2 \log d \geq \log(1 + |z_n - z_{n+1}|/\delta_n) \geq \log|z_n - z_{n+1}| - \log \delta_n$$

or

$$\log \delta_n \geq \log|z_n - z_{n+1}| - 2\log d ,$$

hence $\delta_n \geq |z_n - z_{n+1}|/d^2 > \varepsilon/(f_d d^2)$. □

Lemma 9 (Short Thick Curve to Neighborhood of Root)
For any root α and any distance $\varepsilon > 0$, there are two points z, w in the immediate basin of α with $|z| = 2$, $|w - \alpha| < \varepsilon$ such that w is on the N_p-orbit of z and the hyperbolic distance in U between w and z is at most $9\pi d^4 f_d^2/\varepsilon^2$.

PROOF. We take the points z and w from Lemma 8; they can be connected by a curve $\gamma \subset U$ such that the Euclidean distance between γ and ∂U is at least $\delta = \varepsilon/(f_d d^2)$.

For $a \in \mathbb{C}$ and $r > 0$, let $D_r(a)$ be the open disk around a with radius r. There are finitely many points $b_1, b_2, \ldots, b_k \in \partial U$ such that

$$B := (U \cap \overline{D}_{2+\delta}(0)) \setminus \bigcup_{j=1}^{k} D_\delta(b_j)$$

is compact, and every $b \in B$ has distance at least $\delta/2$ to ∂U. Then $\gamma \subset B$. Let γ' be the (Euclidean) shortest curve within B connecting w and z. If $z', z'' \in \gamma'$ are such that between z' and z'' on γ' there is no point on ∂U, then γ' connects z' and z'' along a straight line, or γ' would not be shortest. It follows that γ' consists of finitely many straight line segments and arcs on $\partial D_\delta(b_j)$ so that any two consecutive line segments are connected by arcs and vice versa, in such a way that γ' is continuously differentiable.

For each line segment on γ', consider the two parallel line segments of equal length at distance $\delta/2$ on both sides; these span a rectangle of area $\ell'\delta$, where ℓ' is the common length of the line segments. For each arc segment of γ' on $\partial D_\delta(b_j)$, let ℓ'' be its Euclidean length; within the sector of the disk $D_{3\delta/2}(b_j)$ spanned by the given arc segment, consider the region between radii $3\delta/2$ and $\delta/2$ (the "$\delta/2$-arc tube around the arc segment"). This is part of the $\delta/2$-neighborhood of the arc segment, and an easy calculation shows that its area is $\ell''\delta$. If the total length of γ' is ℓ, then the sum of the areas of all associated rectangles and arc tubes is $\ell\delta$ (compare Figure 2).

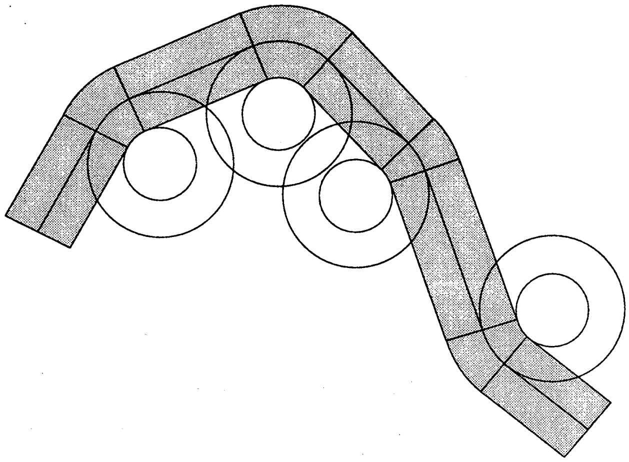

Figure 2: *Illustration of the rectangles and arc tubes around the curve γ': shown are four disks $D_\delta(b_j)$ (large circles) as well as the curve γ', which is surrounded by rectangles and arc tubes.*

Our next claim is that the rectangles and arc tubes have disjoint interiors. To see this, fix some point $z_0 \in \mathbb{C}$ with $|z_0| \leq 2$, let $D := D_{\delta/2}(z_0)$ and $D' := D \setminus \cup_j \overline{D}_\delta(b_j)$. If z_0 is contained in a rectangle or arc tube, then γ' must intersect D', so D' cannot be empty; since ∂U cannot intersect D', every connected component of D' which intersects B is contained in B. Suppose first that D' is connected, so $D' \subset B$. Then the shortest curve in B connecting two points in D' runs through D', so γ' intersects ∂D in no more than two points. It follows that z_0 is contained in the interior of at most one rectangle or arc tube.

If D' is not connected, it consists of at most two connected components: if no two disks $\overline{D}_\delta(b_j)$ and $\overline{D}_\delta(b_{j'})$ meet within D, then D' is connected; if $\overline{D}_\delta(b_j)$ and $\overline{D}_\delta(b_{j'})$ intersect, then $D \setminus (\overline{D}_\delta(b_j) \cup \overline{D}_\delta(b_{j'}))$ has at most two connected components. At least one connected component of $D \setminus (\overline{D}_\delta(b_j) \cup \overline{D}_\delta(b_{j'}))$ meets ∂D at arc length less than $30°$ (in the extremal symmetric situation where b_j, $b_{j'}$ and z_0 are collinear and $|b_j - z_0| = |b_{j'} - z_0| = \delta$, both boundary arcs have length $2\arcsin(1/4) < \arcsin(1/2) = 30°$). It follows that at most one connected component of $D \setminus (\overline{D}_\delta(b_j) \cup \overline{D}_\delta(b_{j'}))$ can intersect γ', and again z_0 is contained in the interior of at most one rectangle or arc tube.

The rectangles and arc tubes around γ' have total area $\ell\delta$. They are all disjoint and contained within the disk of radius $2 + \delta/2$ around the origin, which has total area $\pi(2 + \delta/2)^2$; it follows that

$$\ell \leq \frac{\pi(2+\delta/2)^2}{\delta} = \frac{4\pi}{\delta} + 2\pi + \frac{\pi\delta}{4} = \frac{4\pi}{\delta}\left(1 + \frac{\delta}{2} + \frac{\delta^2}{16}\right) .$$

Since the hyperbolic length of γ' is decreasing as a function of δ (for $0 < \delta < 1$ which are the values of interest to us) and $\delta \geq \varepsilon/(f_d d^2)$, the hyperbolic length of γ' is at most

$$\begin{aligned} 2\ell/\delta &\leq 8\pi\left(\frac{f_d d^2}{\varepsilon}\right)^2 \left(1 + \frac{\varepsilon}{2 f_d d^2} + \frac{\varepsilon^2}{16 f_d^2 d^4}\right) \\ &\leq 8\pi\left(\frac{f_d d^2}{\varepsilon}\right)^2 \left(1 + \frac{1}{2\cdot 16} + \frac{1}{16\cdot 16^2}\right) \leq 9\pi\left(\frac{f_d d^2}{\varepsilon}\right)^2 . \end{aligned}$$

(using $\varepsilon/(f_d d^2) \leq 1/16$ because $d \geq 2$, and only $\varepsilon \leq 1$ is of interest). □

We are now ready to prove the main result.

PROOF OF THEOREM 1. Take the points z and w from Lemma 9 and let $\varphi: U \to \mathbb{D}$ be a Riemann map with $\varphi(\alpha) = 0$. Let $z' := \varphi(z)$, $w' := \varphi(w)$ and $R := |\varphi(z)|$, $r := |\varphi(w)|$, and write $R = 1 - A$, $r = 1 - a$.

The hyperbolic distance in $\mathbb{D}$ between r and R is known to be

$$\log\left(\frac{(1-R)/(1+R)}{(1-r)/(1+r)}\right) = \log(A/a) + \log\frac{1+r}{1+R} > \log(A/a) - \log 2 \ .$$

The hyperbolic distance between z' and w' in $\mathbb{D}$, and that between z and w in U, is then at least $\log(A/a) - \log 2$, so we conclude from Lemma 9 that

$$\log(A/a) \le 9\pi d^4 f_d^2/\varepsilon^2 + \log 2 \ .$$

On the other hand, supposing that α is a simple root, we know that the Newton dynamics transported into $\mathbb{D}$ is a Blaschke product with a superattracting fixed point at 0; if we write this as $f := \varphi \circ N_p \circ \varphi^{-1}$, then we know that $|f(x)| = |x^2| \cdot |r(x)|$ with $|r(x)| \le 1$ for all $x \in \mathbb{D}$ and

$$|f^{\circ n}(z')| \le R^{2^n} \ .$$

If $w = N_p^{\circ n}(z)$, then $w' = f^{\circ n}(z')$ and $R^{2^n} = r$, hence

$$\begin{aligned} n\log 2 \ &\le\ \log\left(\frac{\log r}{\log R}\right) = \log\frac{|\log(1-a)|}{|\log(1-A)|} < \log\left(\frac{\frac{a}{1-a}}{A}\right) \\ &=\ \log(a/A) + |\log(1-a)| \le 9\pi d^4 f_d^2/\varepsilon^2 + \log 2 + |\log r| \end{aligned}$$

(using $|\log(1-a)| < a/(1-a)$ for $0 < a < 1$). This number n blows up if r is very close to 0: in the extreme case $r = 0$, the hyperbolic distance of the starting point z to the root α itself is bounded (which is the best possible case for fast iteration of Newton's map), but it takes forever until we hit the root exactly ($r = 0$ means $w = \alpha$). Recall that w is the first point on the orbit of z with $|w - \alpha| < \varepsilon$. Let w_1 be the last point on the orbit of z before w_1 (so that $N_p(w_1) = w$ and $|w_1 - \alpha| \ge \varepsilon$).

We will give a lower bound for $d(w_1, \alpha)$ within U: any $x \in U \cap D_2(0)$ has $|x - \partial U| < 3$, so the density ρ of the hyperbolic metric is at least $1/6$ on $U \cap D_2(0)$. Since $|w_1 - \alpha| \ge \varepsilon$, we have $d_U(w_1, \alpha) \ge \varepsilon/6$. Under the Riemann map $\varphi: U \to \mathbb{D}$, let $r' := |\varphi(w_1)|$. The hyperbolic distance becomes

$$d(0, \varphi(w_1)) = \log\left(\frac{1+r'}{1-r'}\right) \ge \varepsilon/6 \qquad \text{or} \qquad 1 + \frac{2r'}{1-r'} \ge \exp(\varepsilon/6) > 1 + \varepsilon/6 \ ,$$

which implies $r'(2 + \varepsilon/6) > \varepsilon/6$ or $r' > \varepsilon/13$ (for $\varepsilon < 1$).

Back to our estimate for n. If $r > \varepsilon/13$, we use the formula above; otherwise, we iterate until we find w_1 and add one iteration. In any case, the number n can be estimated as

$$n < \frac{9\pi d^4 f_d^2}{\varepsilon^2 \log 2} + 2 + \frac{|\log \varepsilon/13|}{\log 2} \ . \tag{2}$$

This is what we claimed, at least when the root is simple. For a multiple root, we can perturb the polynomial slightly so that the root becomes simple. The claim of the theorem is independent of the size of the perturbation, so it holds for arbitrarily small perturbations and hence, by continuity, also for multiple roots. □

REMARK. It may be surprising that the result transfers so easily from simple to multiple roots, although the immediate basin undergoes drastic changes when several roots collapse into a multiple root. What happens is that, from the point of view of the simple root, the points z and w move to infinity in the hyperbolic metric of the immediate basin, but the distance between them does not change (in fact, if several immediate basins merge, there is more space and the estimates get only better).

5 Discussion

WHERE THE INEFFICIENCIES COME FROM.
The main source for the extremely inefficient upper bound is the nested estimate of clusters in Theorem 3: if there "almost" was a cluster, it is thrown away and adds a factor in the estimates, and this can possibly happen repeatedly. However, if there "almost" is a cluster, this means something about the approximate geometry of clusters: the extreme cases can never happen simultaneously.

Another inefficiency of the estimates, relatively minor in comparison, is the following: in the estimate of δ in Lemma 7, we used the extreme case that only one point on the curve was close to the boundary, but the following estimates assumed that all points on the curve were equally far from the boundary of the immediate basin (however, this affects only a relatively minor factor of the estimate).

The remaining estimates do not contribute much to the inefficiencies because they are bounded factors away from the sharp estimates, and they are used only once.

TURNING NEWTON'S METHOD INTO AN ALGORITHM.
Often, Newton's method is called "Newton's algorithm", but for an algorithm one needs to know that it terminates. The task is to find, given the degree d of the polynomial and a required accuracy $\varepsilon > 0$, a set $\{\alpha'_1, \alpha'_2, \ldots, \alpha'_d\}$ of complex numbers such that there is a bijection σ to the set of exact roots $\{\alpha_1, \alpha_2, \ldots, \alpha_d\}$ with $|\alpha_k - \alpha'_{\sigma(k)}| < \varepsilon$ for all k.

The overall strategy will be to use a (small) set of starting points, for example as given in [HSS], so that we can be certain that each root is found by at least one of these starting points. The question is: which of the starting points should be used or discarded? Our estimate for the number of iterations gives a partial answer: we can tell how many iterations any "good" starting point takes at most to get ε-close to "its" root, and how close a point is from some root. If after the specified number of iterations a starting point is still not close to some root then it should be discarded. That leaves at least d approximations with error bounds; some roots will be approximated by several starting points, so we have to tell which roots are (to within our required precision) multiple roots and which are simple.

We will not discuss this here. Our last remark closes the gap between our existence result of a good starting point z with $|z| = 2$ for every root, and the result in [HSS] which specifies a finite number of starting points at $|z| = 2$ which find every root after an unspecified number of iterations. If we use three times as many roots as specified in [HSS], we can be certain to hit the middle third of every channel, so that every root has a starting point z in its immediate basin U for which $d_U(z, N_p(z)) < 3 \log d$ (compare Lemma 6). The factor d^2 in the estimate of δ (Lemma 8) now turns into d^6, so the number $n(d, \varepsilon)$ increases by a factor of d^4.

References

[HSS] John Hubbard, Dierk Schleicher, and Scott Sutherland: *How to really find roots of polynomials by Newton's method.* Manuscript, to appear (2000).

[Ma] Anthony Manning: *How to be sure of finding a root of a complex polynomial using Newton's method.* Boletim da Sociedade Brasileira de Matematica, Vol. 22, No. 2 (1992), 157–177.

[Mn] Alexis Marin, *Géométrie des polynômes, coût global moyen de la méthode de Newton.* Séminaire Bourbaki, Astérisque 152–153 (1987).

[NR] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery: *Numerical Recipes* (2nd ed.). Cambridge University Press (1992).

[Pr] Feliks Przytycki: *Remarks on the simple connectedness of basins of sinks for iterations of rational maps.* In: Dynamical Systems and Ergodic Theory, ed. K. Krzyzewski. Polish Scientific Publishers, Warszawa (1989), 229–235.

[Sm] Steven Smale: *Algorithms for solving equations.* Proceedings of the International Congress of Mathematicians (1986).

[Sh] Mitsuhiro Shishikura: *Connectivity of the Julia set and fixed point.* Preprint, Institut des Hautes Etudes Scientifiques IHES/M/90/37 (1990).