Nova Série
BOLETIM
DA SOCIEDADE BRASILEIRA DE MATEMÁTICA

Bol. Soc. Bras. Mat., Vol. 22, No. 2, 157-177

How to be sure of finding a root of a complex polynomial using Newton's method

Anthony Manning

Abstract. The trouble with Newton's method for finding the roots of a complex polynomial is knowing where to start the iteration. In this paper we apply the theory of rational maps and some estimates based on distortion theorems for univalent functions to find lower bounds, depending only on the degree d, for the size of regions from which the iteration will certainly converge to a root. We can also bound the number of iterations required and we give a method that works for every polynomial and takes at most some constant times $d^2(\log d)^2 \log(d^3/\epsilon)$ iterations to find one root to within an accuracy of ϵ.

1. Preliminaries

Assume $d > 1$ and consider the complex polynomial

$$p(z) = z^d + a_{d-1}z^{d-1} + \cdots + a_1 z + a_0 = (z - \alpha_1)(z - \alpha_2)\ldots(z - \alpha_d).$$

Our problem is to determine the d (not necessarily distinct) roots $\alpha_1, \ldots, \alpha_d$ when the coefficients $a_{d-1}, \ldots, a_0$ are given. By the theory of Galois there is no general formula when $d \geq 5$. Let $\hat{\mathbb{C}}$ denote the Riemann sphere, obtained by compactifying the complex numbers $\mathbb{C}$ with the point ∞. Then Newton's method for finding the roots of p is to iterate the rational map $N: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ given by $N(z) = z - p(z)/p'(z)$. The following properties of N are easy to check.

Proposition 1.1. (i) *The degree of N is the number of distinct roots of p.*

(ii) $N' = pp''/p'^2$.

(iii) *N conjugated by an affine map $z \to az + b$ is the Newton map for the polynomial with roots $\{a\alpha_i + b : 1 \leq i \leq d\}$.*

(iv) *The fixed points of N are $\alpha_1, \ldots, \alpha_d$ and ∞.*

(v) *∞ is repelling with eigenvalue $d/(d-1)$.*

(vi) *A simple root α of p is 'superattractive' with $N'(\alpha) = 0$ and $N''(\alpha) = \frac{p''(\alpha)}{p'(\alpha)}$.*

Received 6 May 1991.

(vii) *A root* α *of* p *of multiplicity* m *is a sink with* $N'(\alpha) = (m-1)/m$.

Assumptions. By (iii) we can make an initial affine change of coordinates so that $a_{d-1} = 0$ and each coefficient has $|a_j| \leq 1$. Having $a_{d-1} = 0$ means that the centre of mass of the roots is at the origin and this will help in Lemma 3.2. When each $|a_j| \leq 1$ each root α_i has $|\alpha_i| < 2$ since, for $|z| \geq 2, |z^d| > |p(z) - z^d|$. It will also be convenient for some of our estimates to assume that $d \geq 10$ although this is not necessary to the theory.

We state our main result as an algorithm.

Theorem 1.2. *There is a constant* $c < 800$ *with the following property. Let* p *be any complex polynomial* $p(z) = z^d + a_{d-2}z^{d-2} + \cdots + a_1 z + a_0$ *with* $d \geq 10$ *for which each* $|a_i| \leq 1$ *and let* $\epsilon > 0$. *Put*

$$R = (19.2 d \log d + 2)\pi, \quad \omega = \exp(2\pi i/R) \text{ and } \rho = 1 - 2\pi/R.$$

Define an array A of points

$$\{\rho^j d\omega^k : 0 \leq j < -(R/2\pi)\log(1 - d^{-1} - 2d^{-2}) + 1, \quad 0 \leq k < R\}$$

and consider the following algorithm.

For each (j,k) *in turn evaluate successively* $w = N^\ell(\rho^j d\omega^k)$ *for* $1 \leq \ell \leq [d\log(d^3/\epsilon)]$. *If at any stage* $|w| > d$ *go to the next value of* (j,k). *If* $|w - N(w)| < \epsilon/d$ *stop.*

Then this algorithm does stop and, when it stops, w *is within* ϵ *of some root of* p. *The number of times* N *has been evaluated is at most* $cd^2(\log d)^2 \log(d^3/\epsilon)$.

Definition. The *immediate basin* $B(\alpha_i)$ of a root α_i is the component of $\{z : N^n(z) \to \alpha_i \text{ as } n \to \infty\}$ containing α_i.

Cayley asked in [3] about these basins and Peitgen and co-workers made some interesting pictures of them, see [10] and [11]. There is some symbolic dynamics for the Newton map of a real polynomial, see [1], [7] and [15].

We wish to start iterating N at a point of the component $B(\alpha_i)$ because in the pictures the other components are much smaller. Choosing an initial point in the Julia set $J(N)$ would give an orbit that might be a periodic repellor or might be 'chaotic' since $N|J(N)$ is topologically transitive. In practice, however, such an orbit might through accumulated errors be thrown off $J(N)$ into the basin of a sink. More serious is the fact that N can have other non-fixed periodic sinks. In Proposition 4 of [19] Smale gave the example that $\{0,1\}$ is a superattractive orbit of period 2 for the Newton map of the polynomial $z^3 - 2z + 2$ and Hurley shows in [6] that there can be as many as $d - 2$ such sinks.

In [18] Smale estimated the number of iterations of $z \to z - hp(z)/p'(z)$ for appropriate h needed to solve a polynomial taken from a large class but avoiding some awkward cases. In [19] with proofs in [16] and [17] Shub and Smale obtained better results from a related algorithm in which h varies. Notice that we are looking for a point within ϵ of a root while they wanted $|p(z)| < \epsilon$.

Our purpose is not, of course, to recommend Newton's method as a practical way of finding roots but to use results from complex analysis and rational maps to understand better the basins of attraction of the various roots and to show that in a worst case analysis this method will find one root in a number of iterates which is a reasonable order in d, lower order in fact than $O(d^3 \log(d/\epsilon))$ found in [8] for an algorithm of PL homotopy type due to Kuhn. Renegar [14] shows that, in the worst case, one cannot hope for an algorithm to do much better.

We refer the reader to [2] for the theory of rational maps and to Chapter 2 of [4] for univalent functions.

In order to take enough starting points for our iteration that we can be certain some of their orbits will go to a root we shall estimate from below the size of the immediate basins of the roots. This is best done near ∞. The features of the rational dynamics specific to Newton's method that we shall use are that there is only one repelling fixed point, namely ∞, and that the immediate basin of a fixed sink accumulates there. For pictures of the way the basins approach ∞ look where $p' = 0$ in Figure 4a in [10] or at Maps 61-66 or 75 in [11]. In the next section we shall construct a model of part of $B(\alpha_i)$ reaching to ∞. In §3 we make our estimates of the width of $B(\alpha_i)$ and the number of iterations required and then use them to prove Theorem 1.2.

2. Constructing a Model

Let V denote the closed convex hull of the roots $\{\alpha_1, \dots, \alpha_d\}$ of p. By our assumption that each $|a_j| \le 1$ we have $V \subset D(0,2)$. $D(w,r)$ will denote the disc with centre w and radius r.

Definition. A root α of p is said to be *exposed* if α is a vertex of the polygon which is the frontier of V and the angle subtended at α by V is at most $\pi(1-2/d)$.

Lemma 2.1. *The polynomial p has at least one exposed root.*

Proof. The sum of the angles subtended at the q, say, distinct roots in the frontier of V is $(q-2)\pi$. If no root is exposed then this sum is more than $q\pi(1-2/d) \geq q\pi - 2\pi$. □

Remark. We shall find it easier to study approach to an exposed root. If all the roots are real only two of them are exposed. When the roots are not collinear it is quite possible that V is a triangle with three roots at the vertices and all the other roots are 'hidden inside'. Friedman [5] has found a lower bound for the area in the basin of an exposed root by working outwards from the root.

Proposition 2.2. (Gauss-Lucas, see e.g. [9].) *The roots of p' are contained in the convex hull V of the roots of p.*

Proof. If $p'(\beta)=0$ then $p'(\beta)/p(\beta) = \sum_1^d 1/(\beta-\alpha_i) = 0$ which is impossible if β is outside V for then $\{\beta-\alpha_i : 1 \leq i \leq d\}$ is in an open half-plane. □

We now investigate how points outside V approach V when we apply the map N.

Proposition 2.3. *If ℓ is a straight line with $V\backslash\ell$ in only one component of $V\backslash\ell$ (for example ℓ is an edge of V), z and V are on opposite sides of ℓ and t is the foot of the perpendicular from z to ℓ then*

$$\operatorname{Re}\{(N(z)-t)/(z-t)\} \leq 1 - d^{-1}.$$

This proposition shows that if z and $N(z)$ are both on the other side of ℓ from V then $N(z)$ is closer to ℓ than z is by a factor of at worst $1-d^{-1}$.

Proof. We have

$$N(z)-t = z-t-p(z)/p'(z) = z-t-\{\sum_1^d (z-\alpha_i)^{-1}\}^{-1}$$

so that

$$(N(z)-t)/(z-t) = 1-\{\sum_1^d (1-(\alpha_i-t)/(z-t))^{-1}\}^{-1}.$$

Now, for each i,

$$\operatorname{Re}\{1-(\alpha_i-t)/(z-t)\} \geq 1$$

so

$$\{1-(\alpha_i-t)/(z-t)\}^{-1} \in \operatorname{Cl} D\left(\frac{1}{2}, \frac{1}{2}\right).$$

Thus

$$\sum_1^d \{1-(\alpha_i - t)/(z-t)\}^{-1} \in \mathrm{Cl}\, D(d/2, d/2)$$

and

$$\mathrm{Re}\,\{\sum_1^d (1-(\alpha_i - t)/(z-t))^{-1}\}^{-1} \geq d^{-1}$$

which gives the desired inequality. □

Proposition 2.4. *There is a branch of N^{-1}, approximately $wd/(d-1)$ for large w, that maps $\hat{\mathbb{C}}\backslash V$ inside itself. Under iteration of this branch $(N^{-1})^n(w) \to \infty$ for each $w \in \hat{\mathbb{C}}\backslash V$.*

Proof. The inverse images of ∞ under $N(z) = z - p(z)/p'(z)$ are the roots of p', contained in V by Proposition 2.2, and ∞ itself. Since $N' = pp''/p'^2$ the critical points of N are the simple roots of p and also the roots of p'', which are all in V by a further application of Proposition 2.2.

Thus the branch of N^{-1} which is approximately $wd/(d-1)$ for large w can be extended along any path in the simply-connected domain $\hat{\mathbb{C}}\backslash V$ provided we do not meet a critical value of N^{-1}, that is provided this branch of N^{-1} does not have $N' = 0$ at $N^{-1}(w)$ for any w in $\hat{\mathbb{C}}\backslash V$. The following topological argument shows that $N^{-1}(\hat{\mathbb{C}}\backslash V) \subset \hat{\mathbb{C}}\backslash V$ which suffices since N' does not vanish outside V.

The quotient topological space $\hat{\mathbb{C}}\backslash V$ obtained from $\hat{\mathbb{C}}$ by identifying the convex set V to a point is homeomorphic to the sphere. Define $M: \hat{\mathbb{C}}\backslash V \to \hat{\mathbb{C}}\backslash V$ by $M(V) = V, M(z) = V$ if the interval from z to $N(z)$ meets V and $M(z) = N(z)$ otherwise.

We claim that M is continuous. If $M(z_1) = N(z_1)$ then $N(z_1) \notin V$ and continuity of N shows that $N(z) \notin V$ for z near z_1 so that $M(z) = N(z)$ and M is continuous at z. Now consider continuity at points of $M^{-1}(V)$. If $z_2 \notin V, M(z_2) = V$ and V and z_2 are contained in a line then $1/(V - z_2) = \{1/(v - z_2): v \in V\}$ is convex so that $\sum_1^d (\alpha_j - z_2)^{-1} \in d/(V - z_2)$ and

$$N(z_2) = z_2 + \{\sum_1^d (\alpha_j - z_2)^{-1}\}^{-1} \in z_2 + (V - z_2)/d = (1 - d^{-1})z_2 + d^{-1}V$$

so that $N(z_2) \in V$ and continuity of N gives $N(z)$ within any prescribed distance of V for z sufficiently close to z_2. Thus $M(z)$ is either V or $N(z)$ close to V, which proves that M is continuous at z_2. If $z_3 \notin V, M(z_3) = V$ and now V and z_3 are not contained in a line then there are (not necessarily unique) roots α_ℓ, α_r for which

$$\arg(\alpha_\ell - z_3) \geq \arg(\alpha_j - z_3) \geq \arg(\alpha_r - z_3) \text{ for } 1 \leq j \leq d$$

(with a choice of the function arg continuous except on a cut from 0 to ∞ that does not meet $V - z_3$. Then, from

$$N(z_3) - z_3 = \{\sum_1^d (\alpha_j - z_3)^{-1}\}^{-1},$$

it follows that

$$\arg(\alpha_\ell - z_3) > \arg(N(z_3) - z_3) > \arg(\alpha_r - z_3)$$

so this same property holds with z_3 replaced throughout by any sufficient close z. If the interval from z_3 to $N(z_3)$ meets V only at $N(z_3)$ then continuity of N gives $M(z) = V$ or $M(z) = N(z)$ near V. If, however, the interior of the interval from z_3 to $N(z_3)$ meets V then this continues to hold for z near z_3 and $M(z) = V$. To prove continuity of M at the point V in $\hat{\mathbb{C}}/V$ we observe that z near V in $\hat{\mathbb{C}}/V$ either has $M(z) = V$ or, like z_3 above, has $M(z)$ between z and V or, like z_2 above, has $M(z)$ again between z and V. Thus M is continuous as claimed.

Now M has degree one as a map of the sphere to itself because $M^{-1}(\infty) = \{\infty\}$. The local degree of M is $+1$ everywhere except on $M^{-1}(V)$ because N' cannot vanish in $\hat{\mathbb{C}}\backslash V$. Thus every point of $\hat{\mathbb{C}}/V\backslash\{V\}$ has exactly one inverse image under M. This gives $N^{-1}: \hat{\mathbb{C}}\backslash V \to \hat{\mathbb{C}}\backslash V$, which for large d is near $wd/(d-1)$.

If $w \notin V$ and ℓ is any line separating w from V then by Proposition 2.3 the perpendicular distance of $(N^{-1})^n(w)$ from ℓ is at least $(1 - d^{-1})^{-n}$ times that of w and this tends to ∞ as required. □

The model for the immediate basin of a root $\alpha = \alpha_1$ will be developed from the standard local model. There are two cases to consider: simple and multiple (or repeated) roots.

Theorem 2.5. *If α is an exposed simple root of p there are domains Q and P containing α and 0 respectively and accumulating on ∞ and 1 respectively and a homeomorphism $h: Q \cup \{\infty\} \to P \cup \{1\}$ with $h(\alpha) = 0, h(\infty) = 1$ and $h|Q$ analytic satisfying hN and gh are defined and equal on Q where*

$g(z) = z^2$. *Moreover, P subtends at the point 1 an angle which, when* $d \geq 10$, *is not less than* $0.948/\log d$.

Proof. The local theory, see for example Theorem 3.4 of [2], shows the existence of a unique injective analytic map h defined on a neighbourhood of α satisfying $h(\alpha) = 0$ and $hN = gh$. In fact

$$h'(\alpha) = \frac{1}{2}N''(\alpha) = \frac{1}{2}p''(\alpha)/p'(\alpha).$$

Notice that $p'(\alpha)$ and $p''(\alpha)$ are non-zero for the exposed root α by Proposition 2.2. Our aim is to extend this map h.

Let α_2 be the root of p nearest to α and put $\sigma = |\alpha - \alpha_2|$. If $|z - \alpha| = \delta\sigma$ and $\delta < 1/(2d)$ then

$$\begin{aligned}\{z - N(z)\}^{-1} &= p'(z)/p(z)\\ &= (z-\alpha)^{-1} + \sum_2^d (z-\alpha_i)^{-1} \in D\Big((z-\alpha)^{-1}, \frac{d-1}{\sigma(1-\delta)}\Big).\end{aligned}$$

Thus $z - N(z) \in D(z - \alpha, \tau)$ and $|N(z) - \alpha| \leq \tau$ where

$$\begin{aligned}\tau &= \{|z-\alpha|^{-1} - (d-1)/(\sigma(1-\delta))\}^{-1} - |z-\alpha|\\ &= \{\delta^{-1}\sigma^{-1} - (d-1)\sigma^{-1}(1-\delta)^{-1}\}^{-1} - \delta\sigma\\ &= \delta\sigma(1-\delta)\{(1-\delta) - (d-1)\delta\}^{-1} - \delta\sigma\\ &= \delta\sigma(1-\delta)(1-d\delta)^{-1} - \delta\sigma\\ &= \delta\sigma(d-1)\delta(1-d\delta)^{-1}.\end{aligned}$$

That means

$$|N(z) - \alpha| \leq (d-1)\delta(1-d\delta)|z-\alpha|.$$

For $\delta = 1/(2d)$ this ratio is $(d-1)/d < 1$ so

$$|N^n(z) - \alpha| \leq \{(d-1)\delta/(1-d\delta)\}^n|z-\alpha| \to 0 \text{ as } n \to \infty.$$

Thus $D(\alpha, \sigma/(2d)) \subset B(\alpha)$.

Claim. For $z \in D(\alpha, \sigma/(2d))$ we claim that $|N(\alpha + t(z-\alpha)) - \alpha|$ is a strictly increasing function of t for $0 \leq t \leq 1$. To prove the claim it is sufficient to show that

$$\text{Re}\,\{(z-\alpha)N'(z)/(N(z)-\alpha)\} > 0 \text{ for } 0 < |z-\alpha| < \sigma/(2d).$$

Now, differentiating the product expression for p twice to obtain p'', we have

$$\begin{aligned}\frac{(z-\alpha)N'(z)}{N(z)-\alpha} &= \frac{(z-\alpha)p(z)p''(z)/(p'(z))^2}{z-\alpha-p(z)/p'(z)}\\ &= \frac{p''(z)/p(z)}{\{p'(z)/p(z)\}\{p'(z)/p(z)-(z-\alpha)^{-1}\}}\\ &= \frac{2\sum_{1\le i<j\le d}(z-\alpha_i)^{-1}(z-\alpha_j)^{-1}}{\{\sum_1^d(z-\alpha_i)^{-1}\}\{\sum_2^d(z-\alpha_j\}^{-1}\}} = \frac{2\sum_{1\le i<j\le d} b_i b_j}{(\sum_1^d b_i)(\sum_2^d b_j)}\\ &= 2-\frac{\sum_2^d b_j}{\sum_1^d b_i}-\frac{\sum_2^d b_k^2}{\sum_1^d b_i \sum_2^d b_j}\end{aligned}$$

where we write b_j for $(z-\alpha_j)^{-1}\sigma\{1-1/(2d)\}$ and use

$$2\sum_{1\le i<j\le d} b_i b_j = 2(\sum_1^d b_i)(\sum_2^d b_j) - (\sum_2^d b_j)^2 - \sum_2^d b_j^2.$$

Note that $|b_1| > 2d-1$ while $|b_j| \le 1$ for $2 \le j \le d$ and that it will suffice to prove

$$\left|\sum_2^d b_j\right| < \left|\sum_1^d b_i\right| \tag{1}$$

and

$$\left|\sum_2^d b_k^2\right| < \left|\sum_1^d b_i\right|\left|\sum_2^d b_j\right|. \tag{2}$$

Now

$$\left|\sum_1^d b_i\right| \ge |b_1| - \left|\sum_2^d b_j\right| > 2d-1-\sum_2^d |b_j| \ge 2d-1-(d-1) = d$$

while

$$\left|\sum_2^d b_j\right| \le \sum_2^d |b_j| \le d-1$$

which gives (1). Also

$$\left|\sum_2^d b_k^2\right| \le \sum_2^d \left|b_j^2\right| = \sum_2^d |b_j|^2 \le \sum_2^d |b_j|.$$

The complex numbers $b_j, 2 \le j \le d$, lie in a wedge with vertex 0 which, like the wedge with vertex z (see Figure 1) containing $\{\alpha_j: 2 \le j \le d\}$, has half its vertex angle to be at most $\theta_d = \pi/2 - \pi/d + \arcsin(1/(2d))$.

The components of the b_j along the bisector of their wedge are all positive and thus

$$\left|\sum_2^d b_j\right| \geq \sum_2^d |b_j| \cos\theta_d.$$

The desired inequality (2) follows now since $d\cos\theta_d > 1$ for $d \geq 2$, so the claim is proved.

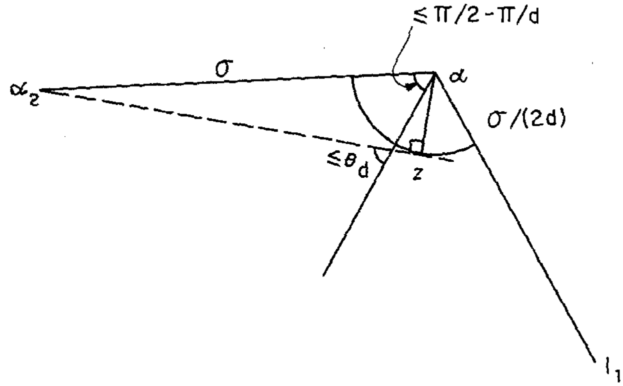

Figure 1

Let Y_k denote $D(\alpha, \{(d-1)/d\}^k \sigma/(2d))$ for $k = 0, 1, 2, \ldots$. We have shown that N maps Y_k inside Y_{k+1}, that N maps rays from α in Y_0 to curves whose distance from α is strictly increasing and consequently that N maps each circle centre α in Y_0 to a curve in Y_1 along which arg increases by 4π. Thus the Riemann surface of $N|D(\alpha, \sigma/(2d))$ has two sheets branched only at α. The conjugacy h is already defined as a univalent map on Y_k for some large k and satisfies $hN = gh$ there. Thus h can be lifted to a univalent map $\tilde{h}$ from a neighbourhood of α in the Riemann surface of $N|D(\alpha, \sigma/(2d))$ onto a neighbourhood of 0 in the Riemann surface of g (branched over 0 and ∞) for which $h = g^{-1}\tilde{h}N$ on Y_k. (Here g^{-1} is univalent and we reject the choice of $\tilde{h}$ that gives $h = -g^{-1}\tilde{h}N$.) Now define h on Y_{k-1} to be $g^{-1}\tilde{h}N$, notice that it is univalent and lift it to a univalent extension $\tilde{h}$ of the map between Riemann surfaces. Continuing this construction inductively we obtain a univalent map h from Y_0 into $D(0,1)$ satisfying $hN = gh$.

In order to estimate hY_0 from inside independently of the distance σ from α to the nearest root α_2 we shall establish the estimate (3) below for $|h'(\alpha)|$. We obtain

$$h'(\alpha) = \frac{1}{2}N''(\alpha) = \frac{1}{2}p''(\alpha)/p'(\alpha) = \sum_2^d (\alpha - \alpha_j)^{-1}$$

from twice differentiating the product expression for p. Let ℓ_1 be the edge of V through α further from α_2. Considering only the component of $(\alpha-\alpha_j)^{-1}$ perpendicular to $\{(\alpha-z)^{-1}: z\in\ell_1\}$ because components parallel to it might cancel out in the sum and then ignoring contributions for $j\neq 2$ we obtain

$$\begin{aligned}|h'(\alpha)| &= \left|\sum_2^d(\alpha-\alpha_j)^{-1}\right| \\ &\geq |\alpha-\alpha_2|^{-1}\sin(2\pi/d) \\ &= \sigma^{-1}\sin(2\pi/d).\end{aligned} \tag{3}$$

Before extending h further consider $hD(\alpha, r\sigma/(2d)) = hQ'$, say, where r is to be chosen with $0<r<1$. By the Growth Theorem, Theorem 2.6 of [4], hQ' certainly contains $D(0,\tau)$, which we denote by P_1, where

$$\tau = r(1+r)^{-2}|h'(\alpha)|\sigma/(2d) \geq r(1+r)^{-2}\sin(2\pi/d)/(2d).$$

With ℓ denoting the bisector of the exterior angle V makes at α we use Q'' to denote the component of $Q'\backslash\ell$ that is disjoint from V. We claim that hQ'' contains all the positive reals in P_1. For this claim notice that

$$(h|D(\alpha,\sigma/(2d)))^{-1}\{t\in\mathbb{R}: t\geq 0\}$$

is a real analytic curve, ρ say, starting at α and invariant under N. Its tangent vector at α is

$$(h'(\alpha))^{-1} = \{\sum_2^d(\alpha-\alpha_j)^{-1}\}^{-1}$$

and this points into Q'' since it is obtained from the vectors $\alpha-\alpha_j$ which each, if translated to α, point into Q''. By Proposition 2.3 applied to the edges of V through α, N moves points of $\ell\backslash\{\alpha\}$ towards V so $N^k(\ell)\cap Q''=\phi$ for $k\geq 0$ and ℓ cannot meet ρ except at α, which proves the claim.

The curve $h(\ell\cap D(\alpha,\sigma/(2d)))$ is the univalent image of a diameter and its direction satisfies the Rotation Theorem (Theorem 3.7 in [4] which does not require $h'(\alpha)=1$) that

$$\left|\arg h'(\alpha+re^{i\theta}\sigma/(2d)) - \arg h'(\alpha)\right| \leq 4\arcsin r$$

when $r<1/\sqrt{2}$. If θ is an argument of the direction of ℓ chosen so that the argument $\theta+\arg h'(\alpha)$ of the tangent vector to $h(\ell)$ at 0 lies in the interval $(0,\pi)$ then the argument of tangent vectors to $h(\ell)$ in P_1 differs from this by at most

$4 \arcsin r$. Thus $h(Q'') \supset P_1'$ defined by

$$P_1' = \{se^{i\varphi} : r^2 \leq s < r, \min(0, \theta + \arg h'(\alpha) - \pi + 4 \arcsin r) < \varphi < \max(0, \theta + \arg h'(\alpha) - 4 \arcsin r)\}.$$

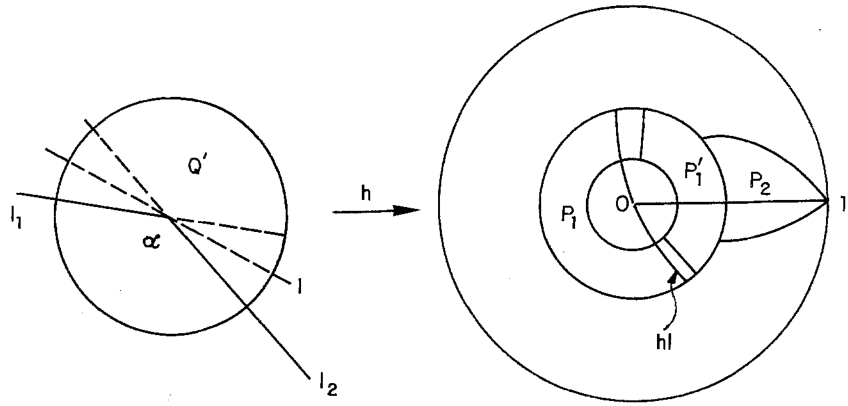

Figure 2

Now define P_2 by

$$P_2 = \{\exp(t \log r + it\varphi) : 0 < t \leq 1, \frac{1}{2} \min(0, \theta + \arg h'(\alpha) - \pi + 4 \arcsin r) < \varphi < \frac{1}{2} \max(0, \theta + \arg h'(\alpha) - 4 \arcsin r)\}.$$

Notice that $P = P_1 \cup P_2$ is an open subset of $D(0,1)$ which is forward invariant for g and that every point of P_2 has just one point of its forward orbit in P_1'. Put $Q_1 = h^{-1}P_1$ and $P_1'' = P_1 \cap gP_2$. Define h from $Q_2 = \bigcup_{n=1}^{\infty} N^{-n}h^{-1}P_1''$ (where N^{-1} is the branch specified in Proposition 2.4 which can be applied because $h^{-1}P_1'' \cap V = \phi$) to $P_2 = \bigcup_{n=1}^{\infty} g^{-n}P_1''$ (where g^{-1} is the branch that maps the real interval $(0,1)$ to itself) as follows. On each $N^{-n}h^{-1}P_1''$ set $h = g^{-n}hN^n$. Then h from $Q = Q_1 \cup Q_2$ to P is analytic because on each $N^{-n}h^{-1}P_1''$ it is a branch of $g^{-(n+1)}hN^{n+1}$. Also $h: Q \to P$ is univalent because it is univalent on Q_1 and on each $N^{-n}h^{-1}P_1''$ with disjoint images.

Now N^{-n} converges uniformly on $h^{-1}P_1''$ to ∞ and g^{-n} converges uniformly on P_1'' to 1 so putting $h(\infty) = 1$ makes h a homeomorphism from $Q \cup \{\infty\}$ to $P \cup \{1\}$.

The angle P subtends at 1 is the angle that $\log P_2$ subtends at 0 which is at least

$$\arctan \frac{\pi - 8\arcsin r}{\log r} \geq \arctan \frac{\pi - 8\arcsin r}{-\log\{r(1+r)^{-2}\sin(2\pi/d)/(2d)\}}$$

and, if we choose $r = 0.050$ this is at least $0.948/\log d$, for $d \geq 10$, as required.□

Theorem 2.6. *Suppose α is an exposed root of p of multiplicity $m > 1$, set $a = 1 - 1/m$ and define $f: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ by $f(z) = z(z+a)/(1+az)$. Then there is an open subset P of $D(0,1)$ homeomorphic to $D(0,1)$ containing 0 and 1 in its closure and subtending an angle of at least 0.124 radians at 1 for which $fP = P$. Also there is Q open in $\mathbb{C}$ and an univalent map h from Q to P extending to a homeomorphism from $Q \cup \{\alpha, \infty\}$ to $P \cup \{0,1\}$ with $h(\alpha) = 0, h(\infty) = 1$ and satisfying $hN = fh$.*

Proof. Notice that $fD(0,1) = D(0,1)$ and 0 and 1 are fixed points of f with $f'(0) = a$ and $f'(1) = 2/(1+a)$. Set $E = D(0,1) \cap \{z: \text{Re } z > 0\}$. We claim that $f|E: E \to f(E)$ is univalent and $f(E) \supset E$ and then define P as the nested intersection of topological discs $\bigcap_{n=1}^{\infty} f^{-n}(E)$. See Figure 3.

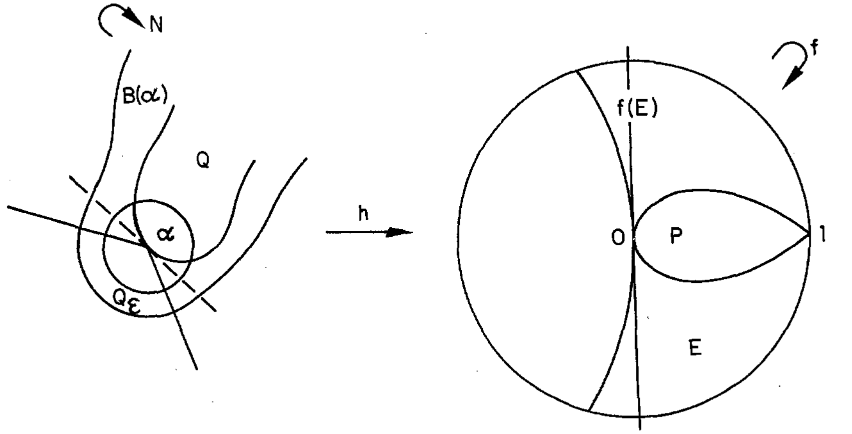

Figure 3

For this claim observe that f is a double cover as a map from the unit circle to itself. Now $f[0,1] = [0,1], f' = 0$ at $z = -m \pm \sqrt{2m-1}$ so $f'|E$ is nowhere zero and $f(E)$ is a subset of $D(0,1)$ symmetrical about the real axis and enclosed by part of the unit circle and $f\{iy: -1 \leq y \leq 1\}$. Also

$$f(iy) = -(1-a^2)y^2/(1+a^2y^2) + iay(1+a^2)/(1+a^2y^2)$$

so that

$$(\partial/\partial y)f(iy) = (1+a^2y^2)^{-2}\{-2(1-a^2)y + ia(1+a^2)(1-a^2y^2)\}$$

which has negative real part and positive imaginary part for $0 < y \leq 1$ since $0 < a < 1$. Thus $f(E) \supset E$ and $f|E$ is univalent as claimed.

Now, since $N'(\alpha) = (m-1)/m = a = f'(0)$, there is an analytic conjugacy h_0 from N to f with $h_0(\alpha) = 0$ defined on a neighbourhood of α according to the local theory, see e.g. Theorem 3.3 of [2]. There is one complex parameter $h_0'(\alpha)$ available in the choice of h_0 and we choose the argument of this so that the line ℓ (which as in Theorem 2.5 is the external bisector of the angle V makes at α) has its image tangent at 0 to the imaginary axis. Then, at least for small enough $\epsilon, Q_\epsilon = \{z: |z-\alpha| < \epsilon, \mathrm{Re}\, h(z) > 0\}$ lies in the complement of V where the branch N^{-1} was defined in Proposition 2.4. For $0 < |z| < 1$ we have $|f(z)| < |z|$ so $f^n(z) \to 0$ for each $z \in P$. For $z \in N^{-n}Q_\epsilon$ put $h_n(z) = (f|E)^{-n}h_0N^n(z)$. This is univalent being a composition of univalent maps. If $z \in N^{-n}Q_\epsilon \cap N^{-r}Q_\epsilon$ with $r > n$ then, by the properties of $h_0, h_0N^n(z) = (f|E)^{-r+n}h_0N^r(z)$ so $h_r(z) = h_n(z)$. Thus defining $h(z)$ as $h_n(z)$ whenever $z \in N^{-n}Q_\epsilon$ gives a univalent map h from $Q = \bigcup_{r=0}^{\infty} \bigcap_{n=r}^{\infty} N^{-n}Q_\epsilon$ with image P. Since $N^{-n}|(Q_\epsilon \backslash D(\alpha, \epsilon/3))$ converges uniformly to ∞ and $f^{-n}|(P\backslash D(0,\epsilon))$ converges uniformly to 1, putting $h(\infty) = 1$ and $h(\alpha) = 0$ gives a homeomorphism $h: Q \cup \{\alpha, \infty\} \to P \cup \{0, 1\}$.

It remains to prove that P subtends an angle of at least 0.124 radians at 1. For this consider an f-orbit in (0,1) from some u near 1 to $f^n(u)$ near 0 for some large n. Now $(f|E)^{-n}|D(f^n(u), f^n(u))$ is univalent so, by the Koebe 1/4 Theorem (Theorem 2.3 in [4]) we have

$$D(u, f^n(u)/(4(f^n)'(u))) \subset f^{-n}D(f^n(u), f^n(u)) \subset f^{-n}(E).$$

Since $f^{-n}D(f^n(u), f^n(u))$ is symmetric about the real axis an improvement on the Koebe 1/4 Theorem (obtained by applying it to the composition of the function f^{-n} with a Möbius transformation that fixes the line $\mathrm{Re}\, z = u$ and the point u, see Exercise 32 of Chapter 2 of [4] gives

$$\{u + iv: |v| < \frac{1}{2} f^n(u)/(f^n)'(u)\} \subset f^{-n}D(f^n(u), f^n(u)) \subset f^{-n}(E).$$

Thus if we can find κ with

$$(1-u)\kappa < \frac{1}{2} f^n(u)/(f^n)'(u)$$

for each positive n then

$$\{u + iv : |v| < (1 - u)\kappa\} \subset P.$$

Such κ must satisfy

$$\kappa < f^n(u)/(2(1-u)(f^n)'(u)) = \frac{1}{2}u/(1 - f^n(u)) \prod_{i=0}^{n-1} g(f^i(u))$$

where

$$g(x) = \{f(x)/(xf'(x))\}\{(1-f(x))/(1-x)\} = (x+a)(1+x)/(ax^2+2x+a)$$

on the interval $0 \leq x \leq 1$ and $g(0) = g(1) = 1$. To estimate the product of g along an orbit observe that, for $0 < x < 1$,

$$\begin{aligned} 1/g(x) - 1 &= (1-a)x(1-x)/\{(x+a)(1+x)\} \\ &< 2(1-a)x(1-x)/(1+ax) = 2(x - f(x)) \end{aligned}$$

where this inequality uses $a \geq \frac{1}{2}$.

Thus

$$-\log g(f^i(u)) < 1/g(f^i(u)) - 1 < 2(f^i(u) - f^{i+1}(u))$$

and

$$-\sum_{i=0}^{n-1} \log g(f^i(u)) < 2(u - f^n(u)) < 2$$

so that

$$\prod_{i=0}^{n-1} g(f^i(u)) > e^{-2} > 0.135$$

and we can take $\kappa = 0.135$ for any u in (0,1). Finally the intervals $\{u+iv : |v| < \kappa(1-u)\}$ contained in P subtend, for u near 1, an angle at 1 of at least 0.124 radians or over 7 degrees. □

Remarks. Computer calculations of the shape of P suggest that it subtends an angle at 1 of about 74° for $m = 2$ and a slightly larger angle for larger m. When two simple roots of p coalesce the degree of N drops by one and it would seem that the basins merge while the piece of Julia set and the small components of the domain of equicontinuity that separated them disappear.

3. Estimating the Width of the Basin near ∞

For both simple and multiple exposed roots α we have constructed in Theorems 2.5 and 2.6 a homeomorphism h from a simply connected domain $Q \subset B(\alpha)$

to $P \subset D(0,1)$. In each case Q accumulates on the fixed repellor ∞ and P accumulates on the fixed repellor 1. However, the conjugacy h was not defined on a neighbourhood of ∞. Since $N'(\infty) = d/(d-1)$ while $g'(1) = 2$ or, to model a multiple root, $f'(1) = 2m/(2m-1)$ there cannot be an analytic conjugacy from a neighbourhood of ∞ to a neighbourhood of 1. We shall first linearise N and g around their repelling fixed points, then observe that taking the $\log 2/\log\{d/(d-1)\}$ power conjugates the linear repulsions and then estimate how near our conjugacy h must be to *this* map.

The map log, defined on the right half-plane, conjugates $g(z) = z^2$ with fixed repellor 1 to multiplication by 2 with fixed repellor 0. There is also (see e.g. Theorem 3.3 of [2] for N^{-1}) a unique analytic map L tangent to the identity at ∞ that conjugates N on a (large) neighbourhood of ∞ to its linear part, multiplication by $(d-1)/d$. The map M defined on the left half-plane by

$$M(z) = \rho z^{\log(1-1/d)/\log 2},$$

for some constant ρ, sends 0 to ∞ and conjugates multiplication by 2 to multiplication by $(d-1)/d$. Thus we have a commutative diagram.

$$\begin{array}{ccccccccc}
\hat{\mathbb{C}},\infty & \xleftarrow{L} & Q,\infty & \xrightarrow{h} & P,1 & \xrightarrow{\log} & \mathbb{C},0 & \xrightarrow{M} & \hat{\mathbb{C}},\infty \\
\downarrow{\scriptstyle \times(d-1)/d} & & \downarrow{\scriptstyle N} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle \times 2} & & \downarrow{\scriptstyle \times(d-1)/d} \\
\hat{\mathbb{C}},\infty & \xleftarrow[L]{} & Q,\infty & \xrightarrow[h]{} & P,1 & \xrightarrow[\log]{} & \mathbb{C},0 & \xrightarrow[M]{} & \hat{\mathbb{C}},\infty
\end{array}$$

We shall study the width of Q using LQ. Recall that LQ is a simply-connected domain accumulating on ∞ invariant under multiplication by $(d-1)/d$. Now the open cone $M \log P$ subtends an angle at its vertex ∞ which is at least

$$(0.948/\log d) \times \log\{d/(d-1)\}/\log 2 > 1.36/(d \log d)$$

using Theorem 2.5.

In the case of Theorem 2.6 where we modelled the basin of a root of multiplicity $m > 1$ by $f: D(0,1) \to D(0,1)$ with

$$f(1) = 1 \quad \text{and} \quad f'(1) = 2m/(2m-1)$$

we had a cone in P subtending 0.124 radians. We replace log by a map L_1 that linearises f around 1 and replace M by M_1 where

$$M_1(z) = \rho z^{\log(1-1/d)/\log\{2m/(2m-1)\}}.$$

Thus the open cone $M\log P$ is replaced by M_1L_1P containing an open cone which subtends an angle at ∞ of at least

$$0.124\log\{d/(d-1)\}/\log\{2m/(2m-1)\} \geq 0.124(2m-1)/d \geq 0.372/d.$$

Now, for $d \geq 10$, we have $0.372/d > 0.856/(d\log d)$ so this proves

Lemma 3.1. *If P is the subset of the unit disc used in Theorem 2.5 (2.6) to model part of the basin of an exposed simple (multiple) root then $M\log P(M_1L_1P)$ subtends at ∞ an angle of at least $0.856/(d\log d)$ radians.*

The idea of studying the angle subtended by these basins at ∞ arose after the author investigated experimentally the quartic polynomial with roots $\pm\exp(\pm i\theta)$ for $0 < \theta < \pi/2$ and observed that the angle on a circle such as $|z| = 4$ occupied by each basin was $73\pm 2°$ independent of θ although the 'junk' surrounding the real and imaginary axes varied in width with θ!

Now take some w of large modulus on the line bisecting the cone $M\log P$. By an appropriate choice of ρ we can ensure that the map S defined near ∞ as the composition $Lh^{-1}\exp M^{-1}$ fixes w. Notice that S is analytic and univalent from $M\log P$ to LQ. Since S commutes with multiplication by $(d-1)/d$ it fixes the orbit of w under that map. Consider the restriction of $|S|$ to the straight line joining $(d-1)w/d$ to w. There is some z_0 in this interval with $|S(z_0)| = |z_0|$ and $|S'(z_0)| \geq |S|'(z_0) \geq 1$. This implies that LQ contains a certain sized disc centre $S(z_0)$. (In the case where α is an exposed multiple root we make a similar construction of z_0.) Before estimating the size of such a disc in Proposition 3.4 we shall bring it in from near ∞ to a moderate distance from 0 so first we study the distortion caused by iterating N.

Lemma 3.2. $|N'(z)z/N(z)| \geq (1-2.2|z|^{-2})^2$ *provided* $|z| \geq 10$.

Proof. If x denotes $|z|^{-1}$ then

$$\left|\frac{N'(z)}{N(z)/z}\right| = \left|\frac{zp(z)p''(z)}{p'(z)\{zp'(z)-p(z)\}}\right|$$
$$= \left|\frac{\{1+a_{d-2}z^{-2}+\cdots\}\{1+a_{d-2}((d-2)(d-3)/(d(d-1)))z^{-2}+\cdots\}}{\{1+a_{d-2}((d-2)/d)z^{-2}+\cdots\}\{1+a_{d-2}((d-3)/(d-1))z^{-2}+\cdots\}}\right|$$
$$\geq (1-x^2-x^3-\cdots)^2/(1+x^2+x^3+\cdots)^2$$
$$= \{(1-x-x^2)/(1-x+x^2)\}^2 \geq (1-2.2x^2)^2. \quad \square$$

Lemma 3.3. *If $|z|$ is large and n is the least number with $|N^n(z)| < d$ while*

$d \geq 10$ *then* $|(N^n)'(z)z/N^n(z)| > 0.69$.

Proof. If, in Proposition 2.3, we use for ℓ the tangent to the circle with centre 0 and radius 2 at $2N^{n-1}(z)/|N^{n-1}(z)|$ we obtain

$$\left|N^i(z)\right| - 2 \geq (d-2)(1-d^{-1})^{i-(n-1)} \text{ for } 0 \leq i < n.$$

The last Lemma now gives

$$\begin{aligned}
|(N^n)'(z)z/N^n(z)| &= \prod_{i=0}^{n-1} \left|N'(N^i(z))N^i(z)/N^{i+1}(z)\right| \\
&\geq \prod (1 - 2.2)\left|N^i(z)\right|^{-2})^2 \\
&\geq \exp 2 \sum_{i=0}^{n-1} \log(1 - 2.2\{2 + (d-2)(1-d^{-1})^{i-(n-1)}\}^{-2}) \\
&\geq \exp(-2 \times 1.02 \times 2.2 \sum_{i=0}^{n-1} \{2 + (d-2)(1-d^{-1})^{i-(n-1)}\}^{-2}) \\
&\geq \exp(-4.488(d-2)^{-2}/\{1-(1-d^{-1})^2\}) \\
&\geq \exp(-4.488/12.16) > 0.69
\end{aligned}$$

where we have used $d \geq 10$. □

Proposition 3.4. *There is a point t with $|t| \leq d$ and $\left|N^{-1}(t)\right| > d$ for which Q contains a disc with centre t and radius* $0.0738|t|/(d \log d)$.

Proof. We put $t = N^n L^{-1} S(z_0)$ where z_0, large enough for us to have $(L^{-1})'$ very near 1, was found earlier in this section with $|S'(z_0)| \geq 1$ and n gives the first iterate for which $|t| \leq d$. Now the disc centre z_0 and radius $|z_0| \sin(0.856/(2d \log d))$ is contained in the cone $M \log P$ or in $M_1 L_1 P$ where S is defined. The restriction of $N^n L^{-1} S$ to this disc is a univalent map so, by the Koebe 1/4 Theorem (Theorem 2.3 of [4]), Q contains a disc centre t and radius

$$\begin{aligned}
(1/4)|z_0| \sin(\frac{0.856}{2d \log d})\left|(N^n)'(L^{-1}S(z_0))\right| &> 1/4 \times 0.69|t| \sin(\frac{0.856}{2d \log d}) \\
&> 0.0738|t|/(d \log d). \quad \square
\end{aligned}$$

Next we shall bound the number of iterates required to bring a point found in Q in this way to within some prescribed distance ϵ of our root α.

Proposition 3.5. *If $u \in Q = h^{-1}P$ and $|u| \leq d$ then $\left|N^{m-1}(u) - N^m(u)\right| < \epsilon/d$ when $m = [d \log(d^3/\epsilon)] + 1$.*

Proof. When α is a simple exposed root we treat separately the two cases where the point $hN^m(u)$ in P belongs to $P_2 = P\backslash P_1$, the exponential of a cone, or to the small disc P_1. Let ℓ_1 and ℓ_2 denote the edges of V that meet at α and let ℓ denote the bisector of the exterior angle ℓ_1 and ℓ_2 make there. In the first case the orbit $N^j(u)$ for $0 \leq j \leq m$ does not cross ℓ and so either does not cross the extended edge ℓ_1 or ℓ_2, say the former. (See Figure 2.) We use the statement in Proposition 2.3 that N reduces perpendicular distance from ℓ_1 and from ℓ_2 by a factor of at least $1 - d^{-1}$. For u this perpendicular distance is at most d so $N^{m-1}(u)$ is at most

$$\begin{aligned} d(1-d^{-1})^{[d\log(d^3/\epsilon)]} &= d\exp\{\log(1-d^{-1})[d\log(d^3/\epsilon)]\} \\ &< d\exp\{-d^{-1}d\log(2d^2/(\epsilon\sin(\pi/d)))\} = \epsilon/(2d)\sin(\pi/d) \end{aligned}$$

beyond ℓ_1 and the same for ℓ_2. Since π/d is a lower bound for the angle ℓ makes with ℓ_1 and ℓ_2 we have $|N^{m-1}(u) - \alpha| < \epsilon/(2d)$. A similar calculation shows that $|N^m(u) - \alpha| < \epsilon/(2d)$ and so $|N^{m-1}(u) - N^m(u)| < \epsilon/d$ as required.

In the second case let k denote the least number with $hN^k(u) \in P_1$. By the construction of P_2 (see Figure 2) we have $N^k(u)$ on the other side of ℓ from V and hence on the other side of either ℓ_1 or ℓ_2, say the former. Since ℓ makes an angle of at least π/d with ℓ_1 we have $\left|N^k(u) - \alpha\right|$ at most $1/\sin(\pi/d)$ times the perpendicular distance from $N^k(u)$ to ℓ_1 and that perpendicular distance is at most $d(1-d^{-1})^k$. From time k to time $m-1$ the orbit of u is modelled by a g-orbit that lies in P_1. Since P_1 was constructed as 0.05 times the size of a disc in $D(0,1)$, distance to 0 in $D(0,1)$ is reduced there by each application of the map g by a factor of at most 0.05. Moreover, distance to 0 in P_1 is proportional to distance to α in $h^{-1}P_1$ to within a factor that, according to Theorem 2.6 of [4], can only vary between $(1-r)^2/(1+r)^2$ and its inverse where $r = 0.05$. Thus distance to α in $h^{-1}P_1$ is certainly reduced by a factor $1 - d^{-1}$ with each application of N and so

$$\left|N^{m-1}(u) - \alpha\right| \leq d(1-d^{-1})^k(\sin(\pi/d))^{-1}(1-d^{-1})^{m-1-k} < \epsilon/(2d).$$

Now $N^m(u)$ is closer still to α and so $|N^{m-1}(u) - N^m(u)| < \epsilon/d$ as required.

There is a further case, where α is a multiple root. But here the f-orbit in $D(0,1)$ always has positive real part as in Figure 3 and the N-orbit in Q stays on the other side of ℓ from V so the calculations in the first case above apply again. □

In [20] Sutherland was able to model the immediate basins of all roots but not to bound the number of iterates required to approach them.

Proposition 3.6. *If* $|N(z) - z| < \epsilon/d$ *then, for some root* α_j*, we have* $|z - \alpha_j| < \epsilon$.

Proof. $d/\epsilon < |N(z) - z|^{-1} = \left|\frac{p'(z)}{p(z)}\right| = \left|\sum_1^d (z - \alpha_i)^{-1}\right| \leq \sum_1^d |z - \alpha_i|^{-1}$ so certainly $1/\epsilon < |z - \alpha_j|^{-1}$ for some j and then $|z - \alpha_j| < \epsilon$ as required. □

The case $p(z) = z^d$ shows that this estimate is sharp.

Lemma 3.7. *The image under* N *of the circle with centre 0 and radius* d *is outside the circle of radius* $d - 1 - 2/d$.

Proof. $z - N(z) = \{\Sigma(z - \alpha_j)^{-1}\}^{-1} \in \{\Sigma(D(z,2))^{-1}\}^{-1} = d^{-1}D(z,2) = D(z/d, 2/d)$ and so has modulus at most $1 + 2/d$ when $|z| = d$. □

We can now prove our main result.

Proof of Theorem 1.2. The algorithm iterates N up to $[d\log(d^3/\epsilon)] + 1$ times, the number required in Proposition 3.5, on an array A of points in the annulus $d - 1 - 2d^{-1} \leq |z| \leq d$. According to Lemma 3.7 the point t found in Proposition 3.4 lies in this annulus. We claim that at least one point, $\rho^J d\omega^K$ say, in the array lies in the disc centre t and radius $0.0738|t|/(d\log d)$ which is contained in the domain Q in the basin $B(\alpha)$ of our chosen exposed root α. For this claim it is sufficient to consider the case where $\rho d \leq |t| \leq d$ and $0 \leq \arg t \leq 2\pi/R$. Three sides of this region have length $2\pi/R$ and the fourth is shorter. Thus one of the four vertices will be in the disc centre t provided the radius of this disc is at least the circumradius $\pi d\sqrt{2}/R$ of a square with sides of length $2\pi/R$. But

$$|t| \geq d(1 - 2\pi/R) \text{ and } R > (19.2d\log d + 2)\pi$$

so

$$0.0738|t|/(d\log d) > \pi d\sqrt{2}/R$$

as required.

As we consider successively the iterates $w = N^\ell(\rho^J d\omega^K)$ of $\rho^J d\omega^K$ we have $|w - N(w)| < \epsilon/d$ for some $\ell \leq m - 1$ by Proposition 3.5 so the algorithm stops with this value of w. Then, by Proposition 3.6, w is within ϵ of some root of p as required. Along this orbit, as we found in Proposition 3.5, the iterates approach and then remain near V. Thus if at some stage in the algorithm $|w| > d$ we are not on the orbit of $\rho^J d\omega^K$ and can safely proceed to the next point of the array. By this device we avoid the awkward possibility than an iterate could prove too large for the computer to handle.

The number of times N is evaluated in this algorithm is at most $[d\log(d^3/\epsilon)+1]$ at each of the $R\{-R/(2\pi)\log(1-d^{-1}-2d^{-2})+1\}$ points in A and, provided $d \geq 10$, this is less than $800d^2(\log d)^2\log(d^3/\epsilon)$. □

References

[1] B. Barna, "Uber die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischen Gleichungen II, Publ. Math. Debrecen, 4 (1956), 384-397.

[2] P. Blanchard, Complex analytic dynamics on the Riemann sphere, Bull. Amer. Math. Soc., 11 (1984), 85-141.

[3] A. Cayley, On the Newton-Fourier imaginary problem, Proc. Cam. Phil. Soc., 3 (1880), 231-232 or Collected Mathematical Papers XI, p. 143 Cambridge University Press 1896.

[4] P. Duren, Univalent functions, Springer, New York, 1983.

[5] J. Friedman, On the convergence of Newton's method, Journal of Complexity, 5 (1989), 12-33.

[6] M. Hurley, Multiple attractors in Newton's method, Ergod. Th. & Dynam. Sys., 6 (1986), 561-569.

[7] M. Hurley & C. Martin, Newton's algorithm and chaotic dynamical systems, SIAM J. Math. Anal., 15 (1984), 238-252.

[8] H.W. Kuhn, Z. Wang & S. Xu, On the cost of computing roots of polynomials, Math. Programming, 28 (1984), 156-163.

[9] M. Marden, Geometry of polynomials, Amer. Math. Soc., Providence R.I., 1966.

[10] H.-O. Peitgen, D. Saupe & F.v. Haeseler, Cayley's problem and Julia sets, Math. Intelligencer, 6 (1984), 11-20.

[11] H.-O. Peitgen & P.H. Richter, The beauty of fractals, Springer, Berlin, 1986.

[12] G. Peters & J.H. Wilkinson, Practical problems arising in the solution of polynomial equations, Jour. Inst. Math. Appl., 8 (1971), 16-35.

[13] F. 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, Warsaw, 1989, 229-235.

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

[15] D. Saari & J. Urenko, Newton's method, circle maps and chaotic motion, Amer. Math. Monthly, 91 (1984), 3-17.

[16] M. Shub & S. Smale, Computational complexity: on the geometry of polynomials and a theory of cost I, Ann. Sci. Ecole Norm. Sup., 18 (1985), 107-142.

[17] M. Shub & S. Smale, Computational complexity: on the geometry of polynomials and a theory of cost II, SIAM J. Comp., 15 (1986), 145-161.

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

[19] S. Smale, On the efficiency of algorithms of analysis, Bull. Amer. Math. Soc., 13 (1985), 87-121.

[20] S. Sutherland, Finding roots of complex polynomials with Newton's method, Thesis, Boston University, 1989.

Anthony Manning
Mathematics Institute
University of Warwick
Coventry CV4 7AL
England