A SMALL PROBABILISTIC UNIVERSAL SET OF STARTING POINTS FOR FINDING ROOTS OF COMPLEX POLYNOMIALS BY NEWTON'S METHOD

BÉLA BOLLOBÁS, MALTE LACKMANN, DIERK SCHLEICHER

ABSTRACT. We specify a small set, consisting of $O(d(\log \log d)^2)$ points, that intersects the basins under Newton's method of *all* roots of *all* (suitably normalized) complex polynomials of fixed degrees d, with arbitrarily high probability. This set is an efficient and universal *probabilistic* set of starting points to find all roots of polynomials of degree d using Newton's method; the best known *deterministic* set of starting points consists of $\lceil 1.1d(\log d)^2 \rceil$ points.

1. INTRODUCTION

Newton's root-finding method is as old as analysis, but still not well understood, even in the fundamental case of finding all roots of a polynomial in a single variable. Its local convergence properties are well known; near simple roots convergence is quadratic and thus extremely rapid. However, the global dynamical properties are insufficiently understood so that numerical analysis algorithms often use different global methods, and resort to Newton's method for a final local "polishing" of the roots.

This article is a contribution towards a better understanding of the global properties of Newton's method, applied to polynomials in a single complex variable. Even for polynomials over the reals, and even if all the roots are real, it is often preferable to use complex methods; see Figure 1.

Among the difficulties with Newton's method are the following:

- if an orbit under iteration comes close to a critical point of the polynomial, the Newton map sends the orbit far away near ∞, so that control of the dynamics is lost, and in any case a large number of iterations are required until the orbit comes back to where the roots are;
- there are polynomials with open sets of starting points that do not converge to any root (Smale [S] asked, in 1984, for a classification of such polynomials; an answer is being given by Mikulich in current work);
- the boundary of the basins of convergence for the roots may have positive planar Lebesgue measure (this follows from recent work by Buff and Cheritat on the existence of Julia sets with positive measure
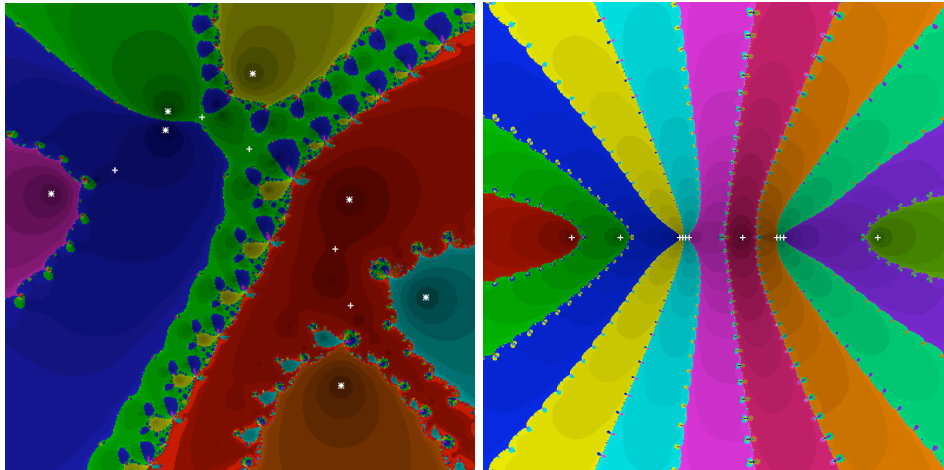
1

FIGURE 1. Dynamical planes of Newton maps of two complex polynomials. Different colors illustrate basins of attraction of different roots; shades of color illustrate different speeds of convergence. It is clearly visible that all immediate basins are unbounded and have one or several channels to ∞ of different widths. Left: a polynomial of degree 7. Right: a polynomial of degree 11 with all roots real. Some of the roots are very close to each other, but away from the disk containing all the roots, the basins and their channels all have almost uniform width, so that finding the real roots using complex methods is much easier.

[BC], combined with Douady and Hubbard's renormalization theory [DH]);

• even if almost every point in $\mathbb{C}$ converges to some root under the Newton iteration, our goal is to find *all* roots of the polynomial, and with bounded complexity. Finding some roots and deflating is usually not an option, because inflation is in general numerically unstable (unless the roots are found in a specific order), and because deflation might not be compatible with the way the polynomial may be specified, or evaluated efficiently (for instance, if the polynomial itself is given by an efficient iteration procedure).

See [Rü] for a recent survey of known results on Newton's method.

Our goal is to turn Newton's method into an efficient algorithm. This goal raises the following questions:

• select a finite number of *good starting points* that are guaranteed to intersect the basins of all roots;
• specify a condition when to stop iterating any of these starting points, because the orbit is either sufficiently close to a root, or the orbit is discarded in favor of some other starting points;

- give a good bound on the complexity of Newton's method to find all roots of the polynomial with prescribed precision.

This article is concerned with the first of these questions. The other tasks are addressed in [Sch1, Sch2]: for instance, roughly speaking, for "most" polynomials p of degree d, properly normalized, our universal set $\mathcal{S}_d$ contains d points that converge to the d different roots of p so that the total number of Newton iterations, for all d roots combined, to achieve an accuracy of ε is at most $O(d^3 \log^3 d) + d \log |\log \varepsilon|$ (or possibly $O(d^2 \log^4 d + d \log |\log \varepsilon|)$). This makes it possible to turn Newton's method from a heuristics into an efficient *algorithm*.

To state our main result, let $\mathcal{P}_d$ be the space of polynomials of degree d, normalized so that all roots are contained in the complex unit disk $\mathbb{D}$.

Theorem 1 (Small Probabilistic Universal Set of Starting Points)**.**
For every degree $d \geq 3$, there is an explicit universal probabilistic set $\mathcal{S}_d$ consisting of $O(d(\log \log d)^2)$ starting points so that for every polynomial $p \in \mathcal{P}_d$, the probability is greater than $1/2$ that the immediate basin of each root of p contains at least one point in $\mathcal{S}_d$.

Remark 1. The meaning of an "explicit and universal" probabilistic set is as follows: we give an explicit probability distribution of starting points that depends only on d so that for any $p \in \mathcal{P}_d$, with probability at least $1/2$ all immediate basins contain at least one point in this set. Of course, enlarging this set by a factor of n will increase the probability to at least $1 - 2^{-n}$.

This result is in a similar spirit as [HSS], where a similar explicit universal set of starting points is constructed. It consists of $\lceil 1.1 d (\log d)^2 \rceil$ points and is deterministic. Our new set is significantly smaller than the deterministic set, much closer to the "ideal lower bound" of d points, but we can do so only using a probabilistic set. We believe that there is no deterministic explicit and universal set of starting points with $o(d \log d)$ points.

Our set $\mathcal{S}_d$ is constructed as follows: firstly, we define a define a "fundamental annulus" $V := \{z \in \mathbb{C}: R\sqrt{1 - 1/d} < z < R\}$ for some $R > 1 + \sqrt{2}$, and choose a "deterministic set" of approximately $(16/\pi)d(\log \log d)^2$ points that are distributed on $m = \lceil (2/\pi) \log \log d \rceil$ circles. These circles have radii $R_k = R(1 - 1/d)^{(k-1/2)/2m}$ for $k = 0, 1, , \ldots, m - 1$, and each circle contains $\lceil 4\pi d \lceil (2/\pi) \log \log d \rceil \rceil$ points at equal distances. This construction is in principle the same as in [HSS]. Secondly, we choose a "probabilistic set" of $\lceil (300/\pi)d \log \log d \rceil$ points randomly inside the annulus $A_R = \{z \in \mathbb{C}: R(d-1)/d - 1/d < |z| < R\}$ for some $R \geq 11$. These deterministic and probabilistic sets of points will respectively find "thick" and "thin" roots, as defined below.

Historical Remark. This research has its origins at the 50-th anniversary celebration of the International Mathematical Olympiad (IMO) held in 2009 in Bremen, Germany. One chief goal of this celebration was to bring together olympiad mathematics and research mathematics, and people involved in

both. This paper was authored by one research mathematician who in his youth was one of the first contestants ever at IMOs and in 2009 was a guest of honor at the 50-th IMO, together with one contestant at the 50-th IMO, and one research mathematician who was among the senior organizers of that IMO and its anniversary. This work is thus very much in the spirit of the IMO anniversary, and we are grateful to this anniversary celebration that has brought us together.

2. Channels and Their Moduli

Consider a complex polynomial $p(z) = c \prod_{j=1}^{d} (z - \alpha_j)$ and let $N_p(z) = z - p(z)/p'(z)$ be the associated Newton map. This is a rational map of degree d if all roots of p are distinct, and of lower degree otherwise. Without changing the Newton map, we may suppose that $c = 1$, and after rescaling, we may suppose that all $\alpha_j \in \mathbb{D}$.

For any root α of p, let U_α be the *immediate basin* of α: the basin is the set of all $z \in \mathbb{C}$ that converge to α under iteration of N_p, and the immediate basin is the connected component containing α. It is known that each U_α is simply connected [Pr] and that the restriction of N_p to U_α sends U_α to itself as a proper map of some degree $k + 1 \in \{2, 3, \ldots, d\}$. We will use the construction and some results from [HSS]. If $\varphi \colon U_\alpha \to \mathbb{D}$ is a Riemann map with $\varphi(\alpha) = 0$, then $f := \varphi \circ N_p \circ \varphi^{-1}$ is a proper holomorphic self-map of $\mathbb{D}$ of degree $k + 1$ and thus extends, by Schwarz reflection, to a rational map of degree $k + 1$, and the restriction of f to $\partial \mathbb{D}$ is a covering of $\partial \mathbb{D}$, also of degree $k + 1$. In particular, the restriction of f to $\partial \mathbb{D}$ has $k \geq 1$ fixed points $q_1, \ldots, q_k$. Set $\lambda_i := f'(q_i)$, for $i = 1, 2, \ldots, k$.

The holomorphic fixed point formula (which essentially is the residue theorem for $1/(z - f(z))$; see [M]) implies that

$$(1) \qquad \sum_{i=1}^{k} \frac{1}{\lambda_i - 1} \geq 1$$

(with equality if the root α is simple). Each of these k fixed points gives rise to a *channel* to ∞ in the immediate basin U_α: for our purposes, a channel is an unbounded component B_i of $U_\alpha \setminus \overline{\mathbb{D}}$. Near ∞, each channel is mapped by N_p conformally to itself, and it defines an access to ∞ within U_α that is fixed by N_p. The quotient of B_i by the dynamics of N_p is a conformal annulus with modulus $\mu_i = \pi / \log \lambda_i$.

Choose some positive real number $M < \pi / \log 4 \approx 2.266$ that will be specified later (we will eventually use $M = \pi / \log \log d$ for large d).

We call a root α *thick* if it has a channel with modulus $\mu_i \geq M$, and *thin* if there is no such channel. We will treat these two cases separately.

- We will explicitly (and thus deterministically) construct a set of $\lceil 4\pi d \lceil 2/M \rceil^2 \rceil$ points that is guaranteed to intersect each channel of a root with modulus greater than M. This set will thus suffice to "find" all thick roots.
- The advantage of thin roots is that even though the individual channels have small moduli, the total area of these channels within any fundamental domain of the Newton dynamics is greater than in the thick case: each channel has little area, but there are more channels in this case. We show that if $\lceil 300\, d \log d / M e^{\pi/M} \rceil$ points are distributed randomly in a certain fundamental annulus of the Newton dynamics, then the probability that any immediate basin U_{α_i} contains one of these points is at least $1 - 1/2d$, and the probability that the immediate basins of *all* thin roots contain such a point is at least $1/2$ because there are no more than d thin roots.

Remark 2. If α is a thin root, then all $\mu_i < M$, hence all $\lambda_i - 1 = e^{\pi/\mu_i} - 1 > e^{\pi/M} - 1$, so by (1), the number k of channels of a thin root is strictly greater than $e^{\pi/M} - 1$. But the mapping degree of U_α equals $k + 1$, so U_α must contain $k > e^{\pi/M} - 1$ of the at most $2d - 2$ critical points of N_p, and thus the number of thin roots is at most $(2d - 2)/(e^{\pi/M} - 1)$. In the end, we will use $M = \pi / \log \log d$, so the number of thin roots will be at most $(2d - 2)/(\log d - 1)$: most roots will be thick.

If there are thin roots, then we can estimate

$$(2) \qquad\qquad e^{\pi/M} < k + 1 \le d \;;$$

in particular, there are no thin roots at all if $M \le \pi / \log d$.

A *conformal quadrilateral* is a Riemann domain $Q \subset \mathbb{C}$ with two distinguished connected and disjoint subsets of the boundary. In our setting, the boundary of Q may not be a topological curve, but the two distinguished boundary subsets will be; we will call them *distinguished boundary arcs*. Then there is a unique $h > 0$ so that the domain $Q_h := \{z \in \mathbb{C} \colon 0 < \operatorname{Im} z < 1, 0 < \operatorname{Re} z < h\}$ has a Riemann map $\varphi \colon Q \to Q_h$ that maps the two distinguished boundary arcs onto the two horizontal sides of Q_h (the Riemann map may not extend continuously to the boundary of Q, but it does so near the two distinguished boundary arcs; the general framework of extremal length using curve families works even if the boundaries are not curves). The value h is defined as the *conformal modulus* of the quadrilateral Q with respect to the two boundary subsets, and denoted $\operatorname{mod}(Q)$; it is invariant for Riemann maps that respect the distinguished boundary subsets [A].

Identifying the two distinguished boundary arcs, we obtain a complex annulus (a doubly connected Riemann surface) with modulus $\operatorname{mod}(Q)$ or less (the exact modulus depends on how the boundaries are identified).
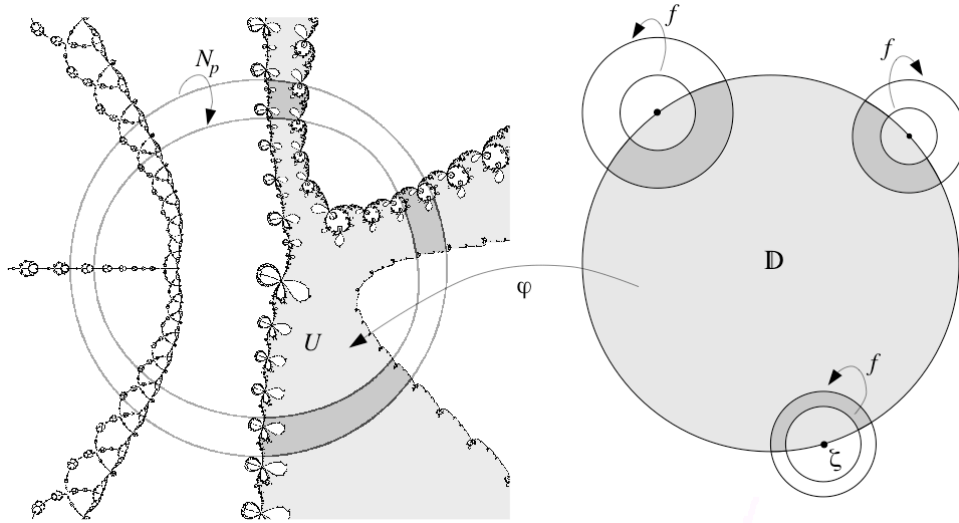
FIGURE 2. Left: the dynamics of Newton's method for some complex polynomial. Highlighted is the immediate basin of attraction of one root, with fundamental domains within the channels shaded. Also shown is the circle at radius R and its image, which is a topological (but not geometric) circle. Right: the complex unit disk $\mathbb{D}$ provides a conformal model for the Newton dynamics of the immediate basin. (Picture taken from [HSS].)

3. HITTING THICK ROOTS

In this section, we will construct an explicit and deterministic set of starting points that is guaranteed to intersect the basins of all thick roots. Our arguments are essentially the same as in [HSS, Section 5], except that we no longer need to find all roots, but only the thick ones.

If $R > (d+1)/(d-1)$ and C_R is the circle of radius R centered at the origin, then N_p maps C_R homeomorphically onto some topological circle around $\mathbb{D}$, and there is some $\kappa > 0$ so that the round annulus $V_{R,\kappa,d} = \{z \in \mathbb{C} \colon R((d-1)/d)^\kappa < |z| < R\}$ is contained in the topological annulus between C_R and $N_p(C_R)$; specifically, if $R \geq 1 + \sqrt{2}$, then $\kappa \geq 1/2$ for all d. If R tends to ∞, then κ tends to 1. All this is [HSS, Lemmas 4 and 12]; see also Figure 2.

We will use the round annulus $V = V_{R,\kappa,d}$ with $R \geq 1 + \sqrt{2}$ and $\kappa = 1/2$ (if we use larger values of R, then we can take larger values of κ, and our bounds will eventually be slightly better; however, in practice these starting points would be further away from the roots, and the iteration would take longer).

Remark 3. The modulus of V is $|\log((d-1)/d)|/4\pi > 1/4\pi d$.

Consider some channel B_i. We want to define Q_i as "the part of the channel B_i within V". If each of the two boundary circles of V intersects B_i in a single connected arc, we set $Q_i := B_i \cap V$. However, if $B_i \setminus V$ has more than two connected components, we need to be more careful. Consider the intersection of B_i with C_R, the outer boundary of V. Let γ be any connected component in this intersection. It separates U_α into two components, one of which contains the root α; then γ will be called an *essential* boundary arc of $B_i \cap C_R$ if the component of $U_\alpha \setminus \gamma$ not containing α is unbounded: this means that γ separates the unbounded part of the channel B_i from the root. At least one component of $B_i \cap C_R$ is essential; choose one such essential component γ, let $\gamma' := N_p(\gamma)$, and let Q_i' be the subset of U_α that is bounded by γ and γ' (if B_i intersects C_R and equivalently $N_p(C_R)$ in only one component, then Q_i' is the part of B_i between C_R and $N_p(C_R)$; in general, the difference may consist of some number of bounded components). Then Q_i' is a fundamental domain of B_i by the dynamics; when viewed as a quadrilateral with distinguished boundary arcs γ and γ', then $\mathrm{mod}(Q_i') \geq \mathrm{mod}(B_i) = \mu_i$ (Q_i' is a quadrilateral, the modulus of B_i is defined using the quotient annulus of B_i by the dynamics).

Now let $C_{R'}$ be the inner boundary circle of V and consider all essential arcs of intersection of $B_i \cap C_{R'}$. If there is only one, then let γ'' be this essential arc. If there are several, then they are totally ordered (because they all separate α in U_α from the unbounded component of $B_i \setminus V$). Let γ'' be the outermost component that separates α from γ (i.e., the one closest to γ), and let Q_i be the component of $B_i \setminus (\gamma \cup \gamma'')$ that is bounded by γ and γ''. This is a conformal quadrilateral with $Q_i \subset Q_i'$, and with γ and γ'' as distinguished boundary arcs, and we have $\mathrm{mod}(Q_i) \geq \mathrm{mod}(Q_i') \geq \mu_i$.

Our task will be to distribute sufficiently many points into V so that we hit quadrilaterals $Q_i \subset V$ with moduli bounded below.

Lemma 2. *Let $S = \{z \in \mathbb{C} \colon -1/2 < \mathrm{Re}\, z < 1/2\}$ and let $Q \subset \mathbb{C}$ be a quadrilateral whose two distinguished boundary arcs are on the two vertical sides of S, one on each. Suppose that Q is disjoint from the set $i\mathbb{Z}$. Then the modulus of Q is at most 2.*

Proof. This is an easy extremal length exercise [A]. There is an integer $n \in \mathbb{Z}$ so that any curve in Q connecting the two distinguished boundary arcs must intersect the segment $[ni, (n+1)i]$. Without loss of generality, suppose that $n = 0$.

Let $B := \{z \in S \colon -1/2 < \mathrm{Im}\, z < 3/2\}$ and let ρ be the characteristic function of B. Then for any curve $\gamma \subset Q$ connecting the two distinguished boundary arcs, its intersection with B has length at least 1. Since $\int_{\mathbb{C}} \rho^2 \, dx \, dy = 2$, it follows that $\mathrm{mod}(Q) \leq 2$. $\qquad\square$

Remark 4. The bound of 2 is not sharp. It is not hard to calculate the exact bound [A], but we are not optimizing constant factors here.
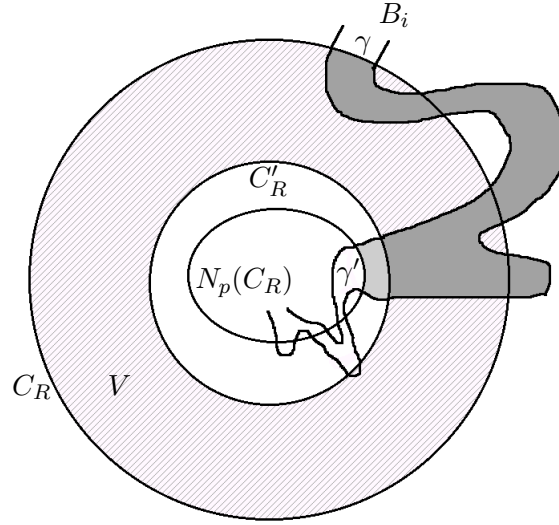
FIGURE 3. The annulus V (hatched). Its outer boundary circle is C_R; the image $N_p(C_r)$ is a topological circle within the bounded complementary component of V. Also shown is a channel B_i; it intersects C_R in four arcs, three of which are essential. Shaded is the quadrilateral Q'_i which is bounded by two essential arcs, one on C_R and one on $N_p(C_R)$; it is a fundamental domain of B_i modulo N_p. The quadrilateral $Q_i \subset Q'_i$ is shaded darker: it is bounded by two essential arcs on ∂V, but may not be contained in V.

Lemma 3. *If V is subdivided into at least $2/M$ concentric and conformally equivalent subannuli, and at least $4\pi d \lceil 2/M \rceil$ points are distributed onto the core circles of all subannuli, so that the points on all circles are equidistributed, then each quadrilateral Q_i with modulus at least M contains at least one of these points.*

Proof. Let $m := \lceil 2/M \rceil$ and subdivide V into m concentric and conformally equivalent subannuli $V_1, \ldots, V_m$, ordered by decreasing radii (so that $V_k = \{z \in V \colon R\beta^k < |z| < R\beta^{k-1}\}$ for $\beta = (1 - 1/d)^{1/2m})$. Write Q for Q_i; this is a quadrilateral for which the two distinguished boundary arcs are on ∂V, one on each boundary component of V.

Subdivide Q into quadrilaterals $Q'_1, \ldots, Q'_m$ as follows, similarly as above. The common boundary circle of V_j and V_{j+1} may intersect Q in several arcs; such an arc is essential if it separates the root α from the unbounded component of $B_i \setminus V$. Use an essential arc to separate Q'_j from Q'_{j+1}, for $j = 1, 2, \ldots, m-1$. (In the special case that $B_i \cap \partial V_j$ only has two connected components, then simply $Q'_j = B_i \cap V_j$.)

By the Grötzsch inequality, one of the quadrilaterals Q_j has modulus $\mathrm{mod}(Q'_j) \geq m \cdot \mathrm{mod}(Q) \geq \lceil 2/M \rceil M \geq 2$. Supposing for now that $0 \notin Q_j$
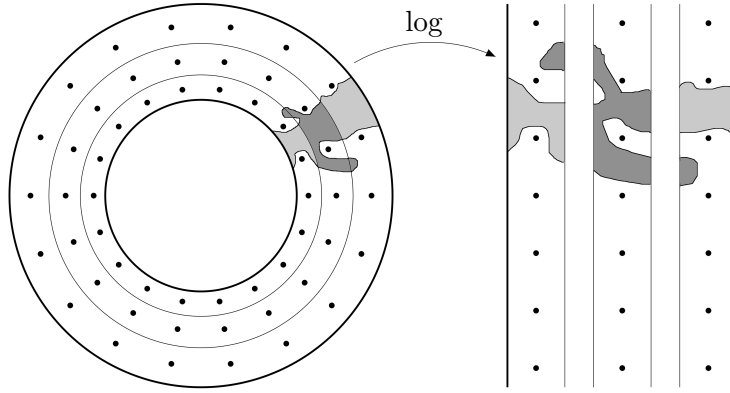
FIGURE 4. The annulus V is subdivided into $M = 3$ concentric subannuli, all of equal moduli. The logarithm unfolds these annuli to vertical strips (moved apart to show them separately). Highlighted is the intersection of one channel with V. Notice that the channel might intersect any circle centered at 0 more than once, but only one arc of intersection is "essential". The quadrilateral in the channel corresponding to the middle subannulus is shown in a darker shade: notice that it intersects the other subannuli as well.

and taking logarithms, the annulus V_j becomes an infinite vertical strip of width $|\log((d-1)/d)|/2m > 1/2md$, and Q_j becomes a quadrilateral that connects the two boundary sides of the strip; see Figure 4.

By Lemma 2, appropriately rescaled, each annulus of modulus 2 intersects the central vertical line within this strip in a straight line segment of length at least $1/2md$. Therefore, placing an infinite sequence of points on any vertical line within the strip so that adjacent points have distance less than $1/2md$, one can be sure that at least one of these points intersects the annulus. The exponential map projects the strip back onto V_j as a universal cover and has period $2\pi i$, so the required number of points on V_j is $4\pi md = 4\pi d \lceil 2/M \rceil$.

If Q_j happens to contain the point $z = 0$, then one cannot take the log of Q_j; but one can take the log of $Q_j \cap V_j$ and transport the function ρ in the proof of Lemma 2 into $Q_j \cap V_j$. This suffices for the conclusion to remain valid. $\qquad\square$

Corollary 4 (Deterministic Starting Points for Thick Roots)**.**
For every d there is an an explicit set consisting of $\lceil 4\pi d \lceil 2/M \rceil \rceil \lceil 2/M \rceil \approx 16\pi d/M^2$ points in V so that for each $p \in \mathcal{P}_d$ and each thick root of p, at least one point in $\mathcal{P}_d$ is contained in the immediate basin of this root.

Proof. Using the construction described in Lemma 3, we have $m = \lceil 2/M \rceil$ circles, and each circle contains $\lceil 4\pi d \lceil 2/M \rceil \rceil$ points. Hence the total number

of required points is as claimed. These points intersect each quadrilateral Q_i and thus the immediate basin of each thick root. □

4. Hitting Thin Roots

Our goal in this case is to find a good lower bound for the area of the union of all channels of any root, guaranteeing us that we will hit one of the channels with high probability if we distribute sufficiently many points randomly on a specified annulus. The area of intersection of a channel with modulus μ_i with an annulus will be bounded below by some multiple of μ_i, so the total area of intersection of an immediate basin with the annulus will be proportional to $\sum \mu_i$, summed over all channels of the root. We thus start with a lower bound for $\sum_{i=1}^{k} \mu_i$.

Set $a_i = \frac{1}{\lambda_i - 1}$, so that $\sum_{i=1}^{k} a_i \geq 1$. We have

$$\mu_i = \frac{\pi}{\log \lambda_i} = \frac{\pi}{\log(1 + \frac{1}{a_i})} .$$

Since $\mu_i < M$ for all i, we get that $a_i < 1/(e^{\pi/M} - 1)$ for all i.

We want to find a lower bound for

$$\sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} \frac{\pi}{\log(1 + 1/a_i)}$$

subject to the conditions $\sum_{i=1}^{k} a_i \geq 1$ and $a_i < 1/(e^{\pi/M} - 1)$.

Lemma 5. *The function $f \colon \mathbb{R}^+ \to \mathbb{R}^+$, $f(x) = \pi/\log(1 + 1/x)$ is strictly monotonically increasing and concave (i.e., its graph is above the line segment through any two points on it).*

Proof. It suffices to prove that f' is positive and monotonically decreasing. This is a straightforward exercise. □

Lemma 6. *If $\mu_i < M$ for all $i \in \{1, \ldots k\}$, then $\sum_{i=1}^{k} \mu_i > \frac{1}{2} M e^{\pi/M}$.*

Proof. Without loss of generality, assume that $a_1 \geq a_2 \geq \ldots \geq a_k$, and that $\sum a_i = 1$. We now consider the sequence $(b_1, \ldots b_k)$ defined by

$$b_i = \begin{cases} \frac{1}{e^{\pi/M}-1} & \text{if } i \leq \lfloor e^{\pi/M} - 1 \rfloor \\ 1 - \frac{\lfloor e^{\pi/M}-1\rfloor}{e^{\pi/M}-1} & \text{if } i = \lfloor e^{\pi/M} - 1 \rfloor + 1 \\ 0 & \text{if } i > \lfloor e^{\pi/M} - 1 \rfloor + 1 . \end{cases}$$

Then we also have $\sum b_i = 1$, and since all $a_i < \frac{1}{e^{\pi/M}-1}$, it follows that the sequence $(b_1, b_2, \ldots b_k)$ majorizes the sequence $(a_1, a_2, \ldots a_k)$, in the sense that

$$\sum_{i=1}^{m} b_i \geq \sum_{i=1}^{m} a_i$$

for all $m \in \{1, 2, \ldots, k\}$, with equality for $m = k$. Since the function f is concave by Lemma 5, we get from Karamata's inequality (see [HLP, Thm. 108]) that $\sum f(a_i) \geq \sum f(b_i)$ and thus

$$\sum_{i=1}^{k} f(a_i) \geq \sum_{i=1}^{\lfloor e^{\pi/M} - 1 \rfloor} f(b_i) = \lfloor e^{\pi/M} - 1 \rfloor \cdot f\left(\frac{1}{e^{\pi/M} - 1}\right) > \left(e^{\pi/M} - 2\right) M.$$

Since $M < \frac{\pi}{\log 4}$, we have $e^{\pi/M} > 4$ and thus

$$\sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} f(a_i) > M(e^{\pi/M} - 2) > \frac{1}{2} M e^{\pi/M}$$

as claimed. $\qquad \square$

Let $\psi \colon (\overline{\mathbb{C}} \setminus \overline{\mathbb{D}}) \to \overline{\mathbb{C}}$ be a linearizing map near ∞ of N_p, i.e., $\psi(N_p(z)) = \psi(z)(d-1)/d$ with $\psi(\infty) = \infty$, and normalize so that $\psi(z)/z \to 1$ as $z \to \infty$. Let

$$W_R := \{w \in \mathbb{C} \colon R(d-1)/d < |w| < R\}$$

be a fundamental domain in linearizing coordinates.

Lemma 7. *For any channel B_i, we have*

$$|\psi(B_i) \cap W_R| \geq \mathrm{mod}(B_i) R^2/d^2 \ .$$

Proof. This is another elementary exercise using extremal length: fix a channel B_i and let $B := \psi(B_i) \cap W_R$. By conformal invariance, the modulus of B_i equals the modulus of B where the boundaries are identified by multiplication by $(d-1)/d$, and this is

$$(\mathrm{mod}\, B_i)^{-1} = (\mathrm{mod}\, B)^{-1} = \sup_{\rho} \inf_{\gamma} \frac{\ell^2(\gamma)}{\|\rho^2\|_B} \ ,$$

where $\rho \colon B \to \mathbb{R}^+$ are measurable functions, $\gamma \colon [0, 1] \to \overline{B}$ are smooth curves with $\gamma(1) = \gamma(0)(d-1)/d$, and $\ell(\gamma) = \int_0^1 \rho(\gamma(t)) |\gamma'(t)| \, dt$.

We simply set $\rho \equiv 1|_B$ (the characteristic function of B). If A denotes the Euclidean area of B, then $\|\rho^2\|_B = A$. The two boundary circles of W_R have radii R and $R(d-1)/d$, so $\ell(\gamma) \geq R/d$. Therefore, $1/\mathrm{mod}\, B \geq R^2/d^2 A$ or $A \geq \mathrm{mod}(B) R^2/d^2 = \mathrm{mod}(B_i) R^2/d^2$. $\qquad \square$

Lemma 8. *For $R \geq 5$, the intersection of the annulus*

$$A_R = \{z \in \mathbb{C} \colon R(d-1)/d - 1/d < |z| < R\}$$

with a channel of modulus μ has area at least

$$\frac{\mu}{d^2} \cdot \frac{(R-1)^2 (R-3)^2}{4(R+1)^2} \ .$$

Proof. Consider the circle $C_R := \{z \in \mathbb{C} \colon |z| = R\}$, and the image $C'_R := N_p(C_R)$. Then C'_R is another topological circle with absolute values between $R(d-1)/d - 1/d = R - (R+1)/d \geq (R-1)/2 \geq 2$ and $R(d-1)/d + 1/d = R - (R-1)/d < R$. Let Z_R be the annulus bounded by C_R and C'_R; it is a fundamental domain for the Newton dynamics, and we have $Z_R \subset A_R$. Consider a channel B and set $B_R := B \cap Z_R$; this is a fundamental domain of the channel, but not necessarily connected.

Consider again the linearizing function $\psi \colon \overline{\mathbb{C}} \setminus \overline{\mathbb{D}} \to \overline{\mathbb{C}}$ of N_p, normalized as $\psi(\infty) = \infty$ and $\psi(z)/z \to 1$ as $z \to \infty$. The Koebe distortion theorem in this normalization yields

$$\frac{|z| - 1}{|z|(|z| + 1)} \leq \left| \frac{\psi'(z)}{\psi(z)} \right| \leq \frac{|z| + 1}{|z|(|z| - 1)} \ .$$

Define the sets

$$B_n := \left\{ z \in B_R \colon R\left(\frac{d}{d-1}\right)^{n-1} < |\psi(z)| < R\left(\frac{d}{d-1}\right)^n \right\}$$

for $n \in \mathbb{Z}$. Each area element in B_n is mapped into W_R by the map $z \mapsto \psi(z)((d-1)/d)^n$ with derivative

$$|\psi'(z)| \left(\frac{d-1}{d}\right)^n < R\frac{|\psi'(z)|}{|\psi(z)|} < \frac{R}{|z|} \cdot \frac{|z| + 1}{|z| - 1} < \frac{2R}{R-1} \cdot \frac{R+1}{(R-3)} \ ,$$

where we used the Koebe theorem in the second inequality and then $|z| \geq (R-1)/2$. This yields a diffeomorphism from B_R to $\psi(B) \cap W_R$, except for discontinuities at the finitely many boundary arcs of the B_n.

The set $\psi(B)$ intersects W_R in a set of area $R^2 \mathrm{mod}(B)/d^2$ by Lemma 7, and areas in B_n are distorted by a factor of no more than the square of the derivative. This implies that

$$|B_R| > \frac{(R-1)^2(R-3)^2}{4d^2(R+1)^2} \mathrm{mod}(B)$$

as claimed. $\qquad\qquad\square$

Lemma 9. *Let $R \geq 5$ and consider the annulus A_R defined as in Lemma 8. If*

$$\left\lceil \frac{32\pi d \log d}{Me^{\pi/M}} \cdot \frac{R(R+1)^3}{(R-1)^2(R-3)^2} \right\rceil$$

points are randomly and independently distributed in A_R, then for any polynomial $p \in \mathcal{P}_d$, each M-thin root has at least one of these points in its immediate basin with probability $1/2$.

Proof. The area of all channels within A_R of any fixed thin root is at least $\left((R-1)^2(R-3)^2/4d^2(R+1)^2\right) \sum \mu_i$ by Lemma 8, and $\sum \mu_i > \frac{1}{2}Me^{\pi/M}$ by Lemma 6. A simple calculation shows that the area of A_R is less than

$2\pi R(R+1)/d$. Therefore, the probability that a point chosen randomly in A_R will lie in one of the channels of this root is at least

$$p = \frac{Me^{\pi/M}}{16\pi d} \cdot \frac{(R-1)^2(R-3)^2}{R(R+1)^3} \ .$$

Now, suppose that we distribute some (large) number K of points on the annulus A_R, randomly and independently. Then the probability that we do *not* hit one of the channels of some fixed thin root will be at most $(1-p)^K$. Since there are at most d thin roots, the probability that there is some thin root the channels of which are not hit is at most $d(1-p)^K$. We need to make K large enough so that $d\,(1-p)^K < \frac{1}{2}$, hence we need $K > \log(1/2d)/\log(1-p)$.

Since $\log(1-p) < -p$ and $\log(1/2d) > \log(1/d^2) = -2\log d$, we have

$$\frac{\log\left(\frac{1}{2d}\right)}{\log(1-p)} < \frac{\log\frac{1}{d^2}}{-p} = \frac{2\log d}{p} = \frac{32\pi d\log d}{Me^{\pi/M}} \cdot \frac{R(R+1)^3}{(R-1)^2(R-3)^2} \ ,$$

so it suffices to distribute this number of points within the annulus at random so that, with probability at least $1/2$, at least one channel of each thin root is hit. $\square$

Remark 5. Increasing the radius R will decrease the necessary number of points to asymptotically $32\pi d\log d/Me^{\pi/M}$ for large R. The disadvantage is that the required number of iterations will be very large until the roots are reached.

5. Conclusion

Proof of Theorem 1. We have to distribute $16\pi d/M^2$ points within the annulus V by the algorithm described in Section 3 to be sure that all thick roots are found. To hit the thin roots, we consider the annulus A_R defined as in Lemma 8, where we choose $R = 11$ (see Remark 5) so that $R(R+1)^3/(R-1)^2(R-3)^2 = 2.97$; we thus have to randomly distribute $32 \cdot 2.97\pi d\log d/Me^{\pi/M} < 300d\log d/Me^{\pi/M}$ points inside the annulus A_R to get all the thin roots with probability at least $1/2$ (in both statements, we ignored the condition that we need to round up certain numbers).

This gives us a total of

$$P(M) = \frac{16\pi d}{M^2} + \frac{300d\log d}{Me^{\pi/M}}$$

points to be chosen to hit the channels of all roots with probability at least $1/2$. In particular, setting $M = \pi/\log\log d$, it suffices to use at least

$$P\left(\frac{\pi}{\log\log d}\right) = \frac{16d}{\pi}(\log\log d)^2 + \frac{300}{\pi}d\log\log d = O(d\,(\log\log d)^2)$$

points. $\square$

Remark 6. Strictly speaking, this proof only works for $d > e^4 \approx 54$ as we claimed in the beginning that $M < \pi/\log 4$ and finally chose $M = \pi/\log\log d$. However, we only need this to simplify some term in the proof of Lemma 6; for $2 \leq \log d < 4$, by being a little bit more careful in the proof of Lemma 6 one can even get slightly better constants, whereas for $1 \leq \log d \leq 2$ one has to choose another value for M to get the same final upper bound.

Remark 7. Of course, the probability $\frac{1}{2}$ can be replaced by any probability p with $0 < p < 1$ if we carry out the algorithm for thin roots described above $\frac{\log(1-p)}{\log 1/2}$ times independently. This will not even change the leading order of the complexity of the whole algorithm as the thin roots only account for a relatively small part of the points needed (a factor $\log\log d$ less than the thick roots).

Remark 8. At several places, we preferred the easy argument over optimal numerical values, as far as a constant factors were concerned. If one were to optimize these factors, it would involve the following places.

The thick roots have the higher complexity, so asymptotically it is most important to optimize constants here. In Lemma 2, the modulus of a quadrilateral is estimated only roughly using a simple argument. The precise value of this quadrilateral can be determined using elliptic integrals; this has been done in [HSS] in an analogous situation. One could then optimize the number of circles and the number of points on them: taking more (or fewer) circles would allow us to use fewer (more) points on each of them, and there is an optimal value of circles that minimizes the total number of points.

For thin roots, we used the estimate $e^{\pi/M} - 2 > e^{\pi/M}/2$ at the end of the proof of Lemma 6, and for large d this loses a factor of 2. Moreover, at the end of the proof of Lemma 9 we lost another asymptotic factor 2 by estimating $\log(2d) < 2\log d$. Finally, there is a certain loss in the estimation of probabilities of hitting the d different basins; these probabilities are not quite additive as estimated. Our estimates in the thin case are thus roughly a factor 4 away from being optimal.

Remark 9. Since the complexities of the deterministic and the probabilistic parts are different, it is tempting to reduce the total complexity by choosing a value of M different from $\pi/\log\log d$ so that both partial complexities become closer to each other. Slight improvements are indeed possible that way, but the gain seems to be minimal. For example, one has

$$P\left(\frac{\pi}{(\log\log d)^{1-1/(1+\log\log d)}}\right) = O\left(d\,(\log\log d)^{2-2/(1+\log\log d)}\right).$$

In this case, the deterministic term is still much bigger than the probabilistic one. Such calculations seem to become much more complicated with relatively little gain.

Moreover, we have not used the condition $\sum_{\alpha_i} k_i \leq 2d - 2$ coming from the total number of "free" critical points. We believe that the effect of incorporating this condition will be marginal.

References

[A]   Lars Ahlfors, *Lectures on quasiconformal mappings*, Second edition. University Lecture Series **38**. American Mathematical Society, Providence, RI, 2006.

[BC]   Xavier Buff, Arnaud Chéritat, *Ensembles de Julia quadratiques de mesure de Lebesgue strictement positive*. C. R. Acad. Sci. Paris **341** 11 (2005), 669–674.

[DH]   Adrien Douady, John Hubbard, *On the dynamics of polynomial-like maps*. Ann. Sci. Ec. Norm. (4) **18** 2 (1985), 277–343.

[HLP]   Godfrey H. Hardy, John E. Littlewood, George Pólya, *Inequalities*, Second edition. Cambridge University Press, Cambridge 1952.

[HSS]   John Hubbard, Dierk Schleicher, Scott Sutherland, *How to find all roots of complex polynomials by Newton's method*. Inventiones Mathematicae **146** (2001), 1–33.

[M]   John Milnor, *Dynamics in one complex variable*, third edition. Annals of Mathematics Studies **160**, Princeton University Press, Princeton, NJ, 2006.

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

[Rü]   Johannes Rückert, *Rational and transcendental Newton maps*. In: Holomorphic dynamics and renormalization, Fields Inst. Commun. **53**, Amer. Math. Soc., Providence, RI, 2008, pp. 197–211.

[Sch1] Dierk Schleicher, *Newton's method as a dynamical system: efficient root finding of polynomials and the Riemann ζ function*. In: Holomorphic dynamics and renormalization, M. Lyubich and M. Yampolski (eds,) Fields Inst. Commun. **53**, Amer. Math. Soc., Providence, RI, 2008, pp. 213–224.

[Sch2] Dierk Schleicher, *The global dynamics of Newton's method for polynomials is efficient*. Manuscript, in preparation.

[S]   Stephen Smale, *On the efficiency of algorithms of analysis*. Bulletin of the American Mathematical Society (New Series) **13** 2 (1985), 87–121.

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK; and Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA.
    E-mail address: `B.Bollobas@dpmms.cam.ac.uk`

Immenkorv 13, D-24582 Bordesholm, Germany
    E-mail address: `malte.lackmann@web.de`

Research I, Jacobs University Bremen, Postfach 750 561, D-28725 Bremen, Germany
    E-mail address: `dierk@jacobs-university.de`