A DENSITY THEOREM FOR PURELY ITERATIVE ZERO FINDING METHODS

Joel Friedman

CS-TR-115-87

November 1987

A Density Theorem for Purely Iterative Zero Finding Methods

Joel Friedman
Princeton University

November 12, 1987

Contents

1 Introduction

The goal of this paper is to prove a theorem about the density of points for which a purely iterative root finding method converges to a root.

For $z \in \mathbf{C}$ and $f(z) = \sum_{i=0}^{d} a_i z^i$ consider a map

$$T_f(z) = \frac{P(z, f, f', \ldots, f^{(l)})}{Q(z, f, f', \ldots, f^{(l)})}$$

where P and Q are polynomials over $\mathbf{C}$. For each f, T_f is a map from $\mathbf{C} \cup \{\infty\}$ to itself which we think of as an iteration in a root finding method. We require that

1.

$$T_f(z) = \frac{z^s P_0(f, zf', z^2 f'', \ldots)}{z^{s-1} Q_0(f, zf', z^2 f'', \ldots)} \tag{1.1}$$

where P_0 and Q_0 are homogeneous polynomials of the same degree.

2. $T_f(z)$ depends only on z and the roots $r_1, \ldots, r_d$ of f, and
$$A(T_f(z)) = T_{Af}(Az)$$
for any linear map $A\colon z \mapsto az + b$, where
$$Af(z) = a_d(z - Ar_1)\ldots(z - Ar_d)$$
for
$$f(z) = a_d(z - r_1)\ldots(z - r_d).$$

3. $T_f(r) = r$, $|T'_f(r)| < 1$ for any root r of f.

4. $T_f(\infty) = \infty$, $|T'_f(\infty)| > 1$ for any f of degree > 1.

To measure the density of convergent points for T_f, let P_d denote the polynomials of degree d with roots in the unit ball. For a polynomial f, let
$$\Gamma_{T,f} = \{z : T_f^n(z) \to \text{ a root of } f \text{ as } n \to \infty\}$$
where T_f^n is the n-th iterate of T_f (i.e. $\Gamma_{T,f}$ is the set of points converging to a root of f under the iteration T_f). Let
$$A_{T,f} = |\Gamma_{T,f} \cap B_2(0)|\,.$$
Then $A_{T,f}/4\pi$ is the probability that a random point in $B_2(0)$ converges to a root.

Theorem 1.1 *Let T satisfy (1)-(4). Then for any d there is a $c > 0$ such that*
$$A_{T,f} > c \qquad \forall f \in P_d.$$

The above density theorem was conjectured to hold for Newton's method by Smale in [Sma85]. This conjecture was proven in [Fri86]; the proof used some special properties of Newton's method and explicit bounds on the constants as a function of d were given. The above theorem applies to a much larger class of root finding methods, though no explicit bounds on c are given.

Examples of T satisfying (1)-(4) are

1. Newton's method, $T_f(z) = z - \frac{f}{f'}$.
2. Modified Newton's method, $T_f(z) = z - h\frac{f}{f'}$ with a constant h, $0 < h < 1$.
3. Taylor's Method

$$T_f(z) = z + \sum_{i=1}^{k} \frac{d^i}{dt^i} \left(\frac{\phi_t(z)}{i!} \right) \Bigg|_{t=0} h^i$$

   where $\phi_t(z)$ solves

$$\frac{d\phi_t(z)}{dt} = -\frac{f(z)}{f'(z)}, \qquad \phi_0(z) = z$$

   with k a positive integer and h a positive number sufficiently small (depending on k).

4. Incremental Euler's Method

$$T_f(z) = z + \sum_{i=1}^{k} \frac{(-hf(z))^k}{k!} g^{(k)}(f(z))$$

   with $g = f^{-1}$, k a positive integer, and h positive and sufficiently small.

2 Some Preliminary Results

One of the main tools used will be the Fatou-Julia theory of iterations of rational maps; see [Bla84] for an exposition. We shall use the following consequence of their theory— let $g: \mathbf{C} \cup \{\infty\} \to \mathbf{C} \cup \{\infty\}$ be a rational map. Let z be a repelling fixed point, i.e. $g(z) = z$ and $|g'(z)| > 1$.

Lemma 2.1 *For any $\epsilon > 0$ we have*

$$\bigcup_{n=0}^{\infty} g^n \{B_\epsilon(z)\} = \mathbf{C} - A$$

where A consists of at most two points.

Proof See [Bla84].

For our maps T, we have that ∞ is a repelling fixed point so the lemma can be applied.

From condition (1)-(4) on T it is easy to see that

$$T_f'(\infty) = q(d) = \frac{Q(1, d, d(d-1), \ldots)}{P(1, d, d(d-1), \ldots)}$$

is a rational funciton of d independent of f, and that if r is a k-tuple root, then

$$T_f'(r) = \frac{P(1, k, k(k-1), \ldots)}{Q(1, k, k(k-1), \ldots)} = \frac{1}{q(k)}.$$

For any f we have that in a neighborhood of ∞,

$$T_f(z) = \frac{z}{q(d)} + O(1)$$

and

$$T_f'(z) = \frac{1}{q(d)} + O\left(\frac{1}{|z|}\right)$$

and T_f^{-1} is defined locally. We have

$$\frac{T_f(z)}{z} = \frac{1}{q(d)} + O\left(\frac{1}{|z|}\right)$$

and so for $|z|$ sufficiently large, we have $z_0 = z, z_{-1}, z_{-2}, \ldots$ given by $T_f(z_{-i}) = z_{-i+1}$ has $|z_{-n}|$ growing like $(q(d) - \epsilon)^n$ for any $\epsilon > 0$ depending on how large $|z|$ is, and thus

$$\begin{aligned} \frac{z_{-n}}{z} &= \prod_{i=0}^{n-1} \left(1 - O\left(\frac{1}{|z_{-i}|}\right)\right) q(d) \\ &= q^n(d) \left(1 - \sum O\left(\frac{1}{|z_{-i}|}\right)\right) \\ &= q^n(d) \left(1 - O\left(\frac{1}{|z|}\right)\right), \end{aligned}$$

since the sum of a geometric progression is bounded by a constant times its largest term. The mean value theorem yields for, say, $r < |z|/2$,

$$T_f^n \{B_{r'}(z_{-n})\} \subset B_r(z)$$

with

$$
\begin{aligned}
r' &= r\left(1 - O\left(\frac{1}{|z_{-n+1}|} + \cdots + \frac{1}{|z|}\right)\right) \\
&= r\left(1 - O\left(\frac{1}{|z|}\right)\right).
\end{aligned}
$$

Thus, if we let

$$\tilde{z} = \lim_{n\to\infty} \frac{z_{-n}}{q^n(d)}$$

we have that for any $r < |z|/2$ we have

$$T_f^n \left\{ B_{rq^n(d)/2}\left(\tilde{z}q^n(d)\right)\right\} \subset B_r(z) \tag{2.1}$$

for n sufficiently large (depending on r).

Next we would like to obtain a version of equation 2.1 for polynomials close to f in a certain sense. Fix D and consider the set $\mathcal{F}_{f,\delta,D}$ of polynomials

$$g(z) = (z - s_1)\ldots(z - s_{d+D})$$

with $s_i \in B_\delta(r_i)$ for $1 \leq i \leq d$ and $|s_i| > 1/\delta$ for $i > d$.

Lemma 2.2 *For any sufficiently large z and $r < |z|/2$ there is a c, δ_0 and n_0 such that if $\delta < \delta_0$ and $n > n_0$ we have*

$$T_g^n \left\{ B_{rq^n(d)/2}(\tilde{z}q^n(d))\right\} \subset B_r(z)$$

if

$$|\tilde{z}|q^n(d) < \frac{c}{\delta}$$

for all $g \in \mathcal{F}_{f,\delta,D}$.

Proof Dividing both numerator and denominator by $z^{s-1}g^{\deg(P_0)}$ in condition (1) on T yields

$$T_g(z) = \frac{zP_0(1, z\frac{g'}{g}, z^2\frac{g''}{g}, \ldots)}{Q_0(1, z\frac{g'}{g}, z^2\frac{g''}{g}, \ldots)}.$$

For $|z|$ sufficiently large and, say, $\leq \frac{1}{2\delta}$ we have

$$\begin{aligned}
\left|\frac{f'}{f} - \frac{g'}{g}\right| &\leq \sum_{i=1}^{d}\left|\frac{1}{z-r_i} - \frac{1}{z-s_i}\right| + \sum_{j=d+1}^{d+D}\left|\frac{1}{z-s_j}\right| \\
&= \sum\left|\frac{s_i - r_i}{(z-r_i)(z-s_i)}\right| + \sum\frac{1}{|z-s_i|} \\
&= O\left(\frac{\delta}{|z|^2} + \delta\right).
\end{aligned}$$

Similarly we have

$$\begin{aligned}
\left|\frac{f^{(k)}}{f} - \frac{g^{(k)}}{g}\right| &= \sum_{1\leq i_1\leq\ldots\leq i_k\leq d}\left|\frac{1}{(z-r_{i_1})\ldots(z-r_{i_k})} - \frac{1}{(z-s_{i_1})\ldots(z-s_{i_k})}\right| \\
&+ \sum_{1\leq i_1\leq\ldots\leq i_k\leq d+D,\ i_k>d}\left|\frac{1}{(z-s_{i_1})\ldots(z-s_{i_k})}\right| \\
&= O\left(\frac{\delta}{|z|^{k+1}} + \frac{\delta}{|z|^{k-1}} + \frac{\delta^2}{|z|^{k-2}} + \cdots + \delta^k\right) \\
&= O\left(\frac{\delta}{|z|^{k+1}} + \frac{\delta}{|z|^{k-1}}\right).
\end{aligned}$$

Thus

$$\left|z^k\frac{f^{(k)}}{f} - z^k\frac{g^{(k)}}{g}\right| = O\left(\frac{\delta}{|z|} + \delta|z|\right)$$

and so

$$\begin{aligned}
T_g(z) &= T_f(z)\left(1 + O\left(\frac{\delta}{|z|} + \delta|z|\right)\right), \qquad (2.2)\\
T_g'(z) &= T_f'(z)\left(1 + O\left(\frac{\delta}{|z|} + \delta|z|\right)\right).
\end{aligned}$$

Now fix a z sufficiently large and a small ϵ so that $z_0 = z, z_{-1}, z_{-2}, \ldots$ defined as before grow like a geometric series. Then, using equation 2.2, we see that for δ sufficiently small we have that $y_0 = z, y_{-1}, y_{-2}, \ldots, y_{-n}$ given by $T_g(y_{-i}) = y_{-i+1}$ grows like a geometric series, as long as $|y^{-n}| < c/\delta$ for

c sufficiently small. Then we get

$$\begin{aligned} y_{-n} &= z_{-n}\left(1+\sum_{i=0}^{n-1} O\left(\frac{\delta}{|y_{-i}|}+\delta|y_{-i}|\right)\right) \\ &= z_{-n}\left(1+O\left(\frac{\delta}{|z|}+\delta|y_{-n}|\right)\right). \end{aligned}$$

Using the chain rule we have

$$\begin{aligned} (T_g^n)'(w) &= \prod_{i=0}^{n-1} T_g'(T_g^i(w)) \\ &= \left(\frac{1}{q(d)}\right)^n \left(1+O\left(\frac{\delta}{|T_g^n(w)|}+\delta|w|\right)\right) \end{aligned}$$

assuming $|T_g^n(w)|$ is sufficiently large and $|w| \leq c/\delta$. The mean value theorem then implies

$$T_g^n\left\{B_{r'}(z_{-n})\right\} \subset B_r(z)$$

where

$$r' = rq^n(d)\left(1+O\left(\frac{\delta}{|z|}+\delta|z_{-n}|\right)\right).$$

Hence, as before, we get that for sufficiently large n,

$$T_g^n\left\{B_{rq^n(d)/2}(\tilde{z}q^n(d))\right\} \subset B_r(z)$$

as long as $|\tilde{z}|q^n(d) < \frac{c}{\delta}$ for c sufficiently small.

3 Successive Normalizations

The difficulty in proving theorem 1.1 is that $A_{T,f}$ is not necessarily continuous when f has multiple roots. Let $f_1, f_2, \ldots$ be a sequence in P_d for which

$$\lim_{n\to\infty} A_{T,f_n} = \inf_{f\in P_d} A_{T,f}.$$

By passing to a subsequence we may assume that

$$f_n(z) = (z-r_1^n)^{e_1}\ldots(z-r_{k_0}^n)^{e_{k_0}}$$

with

$$e_1 + \cdots + e_{k_0} = d$$

and

$$r_i^n \neq r_j^n \qquad \forall n, \quad i < j \leq k_0.$$

By passing to a subsequence we can assume

$$r_i^n \to r_i \qquad \text{as } n \to \infty.$$

If any r_i is isolated, i.e. for some i we have $r_j \neq r_i$ for $j \neq i$, then we could show by continuity in f of T_f that for some $\delta > 0$ we have

$$B_\delta(r_i^n) \subset \Gamma_{T,f_n}$$

for all n sufficiently large, and thus

$$\inf_{f \in P_d} A_{T,f} > 0$$

(the details of the argument appear as part of the proof later in this section). If not, we can assume

$$r_1 = r_2 = \ldots = r_{k_1}$$

and $r_j \neq r_1$ for $j > k_1$. We will now analyze more carefully the way in which $r_1^n, \ldots, r_{k_1}^n$ converge to r_1.

For $z_1, \ldots, z_m \in \mathbf{C}$ not all the same, we define the *normalization of* $z_1, \ldots, z_m$ *centered at* z_1 to be the unique linear map

$$g(z) = az + b, \qquad a \in \mathbf{R},\ a > 0,\ b \in \mathbf{C}$$

such that

$$\sum_{i=1}^{m} g(z_i) = 0,$$

and $g(z_1) = 0$.

By passing to a subsequence we can assume that

1. the normalizations $g_n(z) = a_n z + b_n$ centered at r_1^n have $g_n(r_i^n) \to s_i$ as $n \to \infty$, and

2.
$$q_1^{\lfloor -\log_{q_1} a_n \rfloor} a_n \to a \tag{3.1}$$

as $n \to \infty$ for some $a \in [1/q_1, 1]$ where

$$q_1 = \left(\sum_{i=1}^{k_1} e_i \right)$$

and where $\lfloor a \rfloor$ denotes the largest integer $\leq a$.

Clearly

$$\sum_{i<j} |s_i - s_j| = 1,$$

and so we have

$$s_1 = \cdots = s_{k_2}$$

and $s_j \neq s_1$ for $j > k_2$ where $k_2 < k_1$. In other words, by normalizing we separate the first k_1 roots into smaller groups. By repeated normalization we will finally separate r_1^n from all other r_i^n's. Now we start with the deepest level of normalization and work up, proving a density lower bound for each level.

Let the deepest level be ℓ, and let

$$h_n(r_i^n) \to t_i \qquad \text{for } 1 \leq i \leq k_\ell$$

where h_n is the normalization of $r_1^n, \ldots, r_{k_\ell}^n$ centered at r_1^n. We have

$$\sum_{i<j} |t_i - t_j| = 1,$$

$t_1 = 0$, and $t_i \neq t_1$ if $i > 1$. Consider

$$\tilde{f}(z) = (z - t_1)^{e_1} \ldots (z - t_{k_\ell})^{e_{k_\ell}}.$$

Since $T_{\tilde{f}}(t_1) = t_1$, $|T'_{\tilde{f}}(t_1)| < 1$, and ∞ is a repelling fixed point for $T_{\tilde{f}}$ we have open sets E, arbitrarily near ∞, such that $T_{\tilde{f}}^n\{E\} \to t_1$ as $n \to \infty$. Take a point z large enough so that lemma 2.2 holds, with $B_\epsilon(z)$ converging to t_1 under $T_{\tilde{f}}$ for some $\epsilon > 0$. We have

$$B_{\epsilon q_\ell^m/2}(\tilde{z} q_\ell^m) \subset \Gamma_{T,\tilde{f}}$$

for m sufficiently large where $\tilde{z}$ is as in lemma 2.2 and

$$q_\ell = q\left(\sum_{i=1}^{k_\ell} e_i\right).$$

Let h'_n be the normalization of the $\ell-1$-th level, i.e. of $r_1^n, \ldots, r_{k_{\ell-1}}^n$ centered at r_1^n,

$$h'_n(z) = a'_n z + b'_n$$

and let

$$h_n(z) = a_n z + b_n.$$

We have that

$$\frac{a_n}{a'_n} q_\ell^{\lfloor -\log_{q_\ell}(a_n/a'_n)\rfloor} \to a$$

as $n \to \infty$ for some $a \in [\frac{1}{q_\ell}, 1]$ (at each level we normalize and pass to a subsequence satisfying a condition analogous to that of equation 3.1 as well as the preceeding condition). We want to prove that

$$B_{\epsilon_0}(z_0) \subset \Gamma_{T,h'_n f_n} \tag{3.2}$$

for all sufficiently large n, where

$$\begin{aligned} z_0 &= \tilde{z} a q_\ell^{-M} \\ \epsilon_0 &= \epsilon a q_\ell^{-M}/4 \end{aligned}$$

for some positive integer M. To see this, consider first

$$h_n f_n(z) = (z - h_n(r_1^n))^{e_1} \ldots (z - h_n(r_{k_\ell}^n))^{e_{k_\ell}}.$$

We claim that for n sufficiently large we have

$$B_\epsilon(z) \subset \Gamma_{T,h_n f_n}.$$

To see this, we note that for some small $\eta > 0$ we have

$$|z - t_1| \leq \eta \Longrightarrow |T_{\tilde{f}}(z) - t_1| \leq (1-\mu)|z - t_1|$$

for some $\mu > 0$, and that for some large N,

$$T_{\tilde{f}}^N \{B_\epsilon(z)\} \subset B_{\eta/2}(t_1).$$

Estimating as in lemma 2.2 (note that for any δ we have $h_n f_n \in \mathcal{F}_{\tilde{f},\delta,D}$ for n sufficiently large and $D = d - q_\ell$) we get that for n sufficiently large

$$\begin{aligned} |z - t_1| \leq \eta \quad &\Longrightarrow \quad |T_{h_n f_n}(z) - t_1| \leq (1 - \mu/2)|z - t_1| \\ &\Longrightarrow \quad z \in \Gamma_{T,h_n f_n} \end{aligned}$$

and that

$$T^N_{h_n f_n}\left\{B_\epsilon(z)\right\} \subset B_\eta(t_1) \subset \Gamma_{T,h_n f_n}$$

using $h_n(r^n_1) = t_1$ and that for any $y \in B_\epsilon(z)$ we have $y, T_{\tilde{f}}(y), T^2_{\tilde{f}}(y), \ldots$ stays away from the r^n_i's with $i > 1$. Now we apply lemma 2.2 to conclude that for m sufficiently large we have

$$T^m_{h_n f_n}\left\{B_{\epsilon q^m_\ell/2}(\tilde{z} q^m_\ell)\right\} \subset B_\epsilon(z) \subset \Gamma_{T,h_n f_n}$$

so that

$$B_{\epsilon q^m_\ell/2}(\tilde{z} q^m_\ell) \subset \Gamma_{T,h_n f_n}$$

as long as $|\tilde{z}| q^m_\ell < c/\delta$ for some c sufficiently small, where $1/\delta$ is a lower bound on $h_n(r^n_i)$ for $i > k_\ell$. Rescaling by a factor of a_n/a'_n and translating appropriately we get

$$B_{\epsilon q^m_\ell a_n/(2a'_n)}(\tilde{z} q^m_\ell a_n/a'_n) \subset \Gamma_{T,h'_n f_n}$$

if

$$|\tilde{z}| q^m_\ell a_n/a'_n < c \min_{i>k_\ell} h'_n(r^n_i) < c'. \tag{3.3}$$

Taking

$$m(n) = \lfloor \log_{q_\ell} \frac{a'_n}{a_n} \rfloor - M$$

where M is sufficiently large to ensure equation 3.3 holds, we get that for sufficiently large n,

$$B_{\epsilon a q_\ell^{-M}/4}(\tilde{z} a q_\ell^{-M}) \subset \Gamma_{T,h'_n f},$$

the 4 in $\epsilon a q_\ell^{-M}/4$ appearing to account for the fact that

$$\frac{a_n}{a'_n} q_\ell^{m(n)}$$

approaches, rather that equals, $a q_\ell^{-M}$ as $n \to \infty$. Thus equation 3.2 is established.

Now that we have a statement of the form

$$B_{\epsilon_0}(z_0) \subset \Gamma_{T,h'_n f_n},$$

we proceed to get a statement of the form

$$B_{\epsilon_1}(z_1) \subset \Gamma_{T,h''_n f_n},$$

where h''_n is the normalization at the $\ell-2$-th level, i.e. the normalization of $r_1^n, \ldots, r_{k_{\ell-2}}^n$ centered at z_1^n. To do this we consider

$$\hat{f}(z) = (z-t_1)^{e_1} \ldots (z - t_{k_{\ell-1}})^{e_{k_{\ell-1}}}.$$

Using lemma 2.1 we can find an arbitrarily large z with an ϵ so that for some N

$$T_{\hat{f}}^N \{B_\epsilon(z)\} \subset B_{\epsilon_0}(z_0).$$

Now we repeat the argument of before to conclude

$$T_{h'_n f_n}^N \{B_\epsilon(z)\} \subset B_{\epsilon_0}(z_0)$$

i.e.

$$B_\epsilon(z) \subset \Gamma_{T,h'_n f_n}$$

for n sufficiently large, and that

$$T_{h''_n f_n}^{m'(n)} \{B_{\epsilon_1}(z_1)\} \subset \Gamma_{T,h''_n f_n}$$

for some $m'(n)$ and fixed ϵ_1, z_1.

Repeating the above argument $\ell - 2$ more times yields that for all n sufficiently large we have

$$B_\epsilon(z) \subset \Gamma_{T,f_n}$$

for some fixed ϵ and z with z very near r_1^n. Hence

$$\lim_{n\to\infty} A_{T,f_n} > \pi\epsilon^2 > 0$$

and theorem 1.1 is proven.

References

[Bla84] Paul Blanchard. Complex analytic dynamics on the riemann sphere. *Bull. of the AMS*, 11(1):85–141, July 1984.

[Fri86] Joel Friedman. On the convergence of newton's method. In *27th Annual Symposium on Foundations of Computer Science*, pages 153–161, 1986.

[Sma85] Steve Smale. On the efficiency of algorithms of analysis. *Bull. of the AMS*, 13(2):87–121, Oct. 1985.