MATHEMATICS OF COMPUTATION
VOLUME 51, NUMBER 184
OCTOBER 1988, PAGES 707-719

On Approximate Zeros and Rootfinding Algorithms for a Complex Polynomial

By Myong-Hi Kim

Abstract. In this paper we give criteria for a complex number to be an approximate zero of a polynomial f for Newton's method or for the kth-order Euler method. An approximate zero for the kth-order Euler method is an initial point from which the method converges with an order $(k+1)$ Also, we construct families of Newton (and Euler) type algorithms which are surely convergent.

1. Introduction. Newton's method has long been used for solving a nonlinear equation $f(z)=0$. The Newton method attempts to solve $f(z)=0$ by an iteratively defined sequence $z_{n+1} = z_n - f(z_n)/f'(z_n)$, for an initial point z_0. It indeed converges to a root at a fast rate, if it starts with a good initial point. However, not much is known about the region of convergence or of fast convergence, and it is difficult to obtain a priori knowledge of convergence.

In this paper we study the efficiency and the convergence properties of the Newton method and other generalized methods for solving a *polynomial equation* $f(z)=0$. We have two main goals. First, we establish an estimate for a point z_0, which predicts fast convergence of the algorithms starting at z_0. Secondly, we develop a method which is guaranteed to converge, given an arbitrary initial point z_0.

Following Shub and Smale, we consider the following generalized version of the Newton method, called the modified kth-order Euler method.

We recall from elementary complex analysis that for a polynomial f and $z \in \mathbf{C}$ such that $f'(z) \neq 0$, there is a well-defined local inverse branch f_z^{-1} of f such that $f_z^{-1}(f(z)) = z$.

Definition 1.1. For an integer k and a complex number h, the Euler method iteratively defines a sequence $z_{n+1} = E_{k,h,f}(z_n) = T_k f_{z_n}^{-1}((1-h)f(z_n))$ for an initial point z_0, where T_k is the kth-order truncation of $f_{z_n}^{-1}$ considered as a power series about $f(z_n)$.

For brevity we denote $E_{k,h,f}$ by E_k if there is no confusion. Note that $E_{1,1,f}$ gives the Newton method.

We define an approximate zero of f for E_k as follows.

Received October 6, 1986; revised June 3, 1987 and December 28, 1987.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 30E05, 65E05, 68Q15.

Definition 1.2. z_0 is an approximate zero of f for E_k if

$$\frac{|f(z_n)|}{|f(z_0)|} \le \left(\frac{1}{2}\right)^{(k+1)^n}, \tag{1}$$

$$|z_n - \xi| \le c\left(\frac{1}{2}\right)^{(k+1)^n} |z_0 - \xi|, \tag{2}$$

where c is a constant, ξ is a root and $z_{n+1} = E_{k,1,f}(z_n) \to \xi$. □

Note the fast convergence of z_n to ξ. This notion is due to Smale [12].

For a polynomial f and $z \in \mathbf{C}$, let

$$a_{f,z} \equiv \max_{j\ge 2} \left|\frac{f(z)}{f'(z)}\right| \left|\frac{f^{(j)}(z)}{j!f'(z)}\right|^{1/(j-1)}.$$

We show that z is an approximate zero of f for all E_k if $a_{f,z} \le \frac{1}{48}$ (see Theorem 4.4). Recently, Smale [14] has obtained a similar result for $k = 1$ with a better estimate (a constant α_0, in his notation, between $\frac{1}{8}$ and $\frac{1}{7}$) for a more general class of polynomial maps $f\colon \mathbf{C}^n \to \mathbf{C}^n$.

The estimate for $a_{f,z}$ plays an important role even when z is not an approximate zero of f. It suggests the next iterate in the construction of algorithms which produces sure convergence.

We construct two families of modified Euler methods A_k and B_k, which always converge to a root or a critical point of f (see Theorems 5.A and 5.B). θ is called a critical point of f if $f'(\theta) = 0$. As in the work of Shub and Smale ([9], [10]), the idea is to approximate the solution curve $\phi_t(z_0)$ to the Newton vector field $F(z) = -f(z)/f'(z)$ where $\phi_0(z_0) = z_0$. Note that $f(\phi_t(z_0)) = e^{-t}f(z_0)$, a straight line through $f(z_0)$. Hence one can approximate a root by approximating $f^{-1}(e^{-t}f(z_0)\colon t \to \infty)$. To do so, Shub and Smale use the modified Euler method with a fixed step size h in $E_{k,h,f}$ together with a probabilistic estimate on the set of initial points. In our algorithms we use a varying step size h at each point z, where h is given in terms of $a_{f,z}$ and hence related to the radius of convergence of f_z^{-1}. In particular, we show that for any polynomial f and initial point z_0, B_k always produces a sequence z_n converging to a root unless there is a critical value of f on the ray $(0, f(z_0)]$ (see Theorem 5.B). Recently, Shub and Smale [11] have shown that an algorithm similar to A_1 converges to a root for almost all polynomials and for almost all initial points z_0.

We have run some experiments on the algorithm A_1 and other similar algorithms with a starting point 0 and with a supplementary algorithm of Shub and Smale [10] for degenerate cases such as $a_{f,z} \ge 50d^2$. This corresponds to the case where z is near a critical point. Among $(100 \cdot d^2)$ randomly selected polynomials of each degree $d \le 100$ with complex coefficients $|a_i| \le 1$, the average number of iterations to locate an approximate zero or to locate ς such that $|f(\varsigma)| \le 10^{-4}$ is found to be less than 200. Our experimental result is independent of the degree d.

2. Preliminaries. In this section we discuss some preliminary material needed in the later sections on the local behavior of analytic functions.

The main tools used in Section 3 are from the theory of schlicht functions. f is called a schlicht function if $f(0) = 0$, $f'(0) = 1$ and it is univalent on $D_1(0)$, the unit disk at 0. A univalent function is a one-to-one complex analytic function.

To each $z \in \mathbf{C}$ and complex polynomial f such that $f(z) \neq 0$ and $f'(z) \neq 0$, one associates a normalized polynomial σ by means of

$$\sigma(w) = w + \sigma_2 w^2 + \cdots + \sigma_d w^d, \quad \text{where } \sigma_j = \left(\frac{-f(z)}{f'(z)}\right)^{j-1} \frac{f^{(j)}(z)}{j! f'(z)}.$$

Let $R_{f,z}$ be the radius of convergence of f_z^{-1}, considered as a power series at $f(z)$.

For $f(z) \neq 0$, let $H_{f,z} = R_{f,z}/|f(z)|$; see Figure 2.1. The following lemma is extracted from the work of Shub and Smale (see [9, p. 113]).

LEMMA 2.1. *Let σ^{-1} be the inverse branch of σ taking* 0 *to* 0. *Then*

(1) $\sigma^{-1}(0) = 0$, $\sigma^{-1\prime}(0) = 1$.

(2) *Let $x = f_z^{-1}((1-h)f(z))$ with $|h| < H_{f,z}$. Then*

$$\frac{f(x)}{f(z)} = 1 - \sigma \circ \varepsilon, \quad \textit{where } \varepsilon = \frac{x-z}{F(z)} \textit{ and } F(z) = \frac{-f(z)}{f'(z)}.$$

(3) $f_z^{-1}((1-h)f(z)) = z + F(z)\sigma^{-1}(h)$.

(4) $T_k f_z^{-1}((1-h)f(z)) = z + F(z)T_k\sigma^{-1}(h)$.

(5) *The radius of convergence of σ^{-1} at* 0 *is $R_{\sigma,0} \equiv H_{f,z}$.*

(6) $\frac{1}{H}\sigma^{-1}(Hh)$ *is schlicht, $H \equiv H_{f,z}$.*

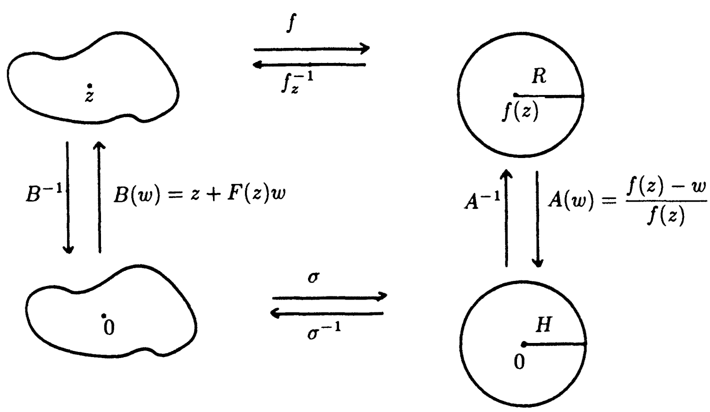

FIGURE 2.1. $R = R_{f,z}$, $H = H_{f,z}$

Proof. (1) is immediate. (2) is from Proposition 2 in [9]. For (3), (4) and (5), see [9, p. 114] and [10, p. 153]. (6) is a trivial consequence of (1) and (5). □

Using Lemma 2.1, we may reformulate Definition 1.1 of $E_{k,h,f}$ as follows.

Definition 2.2. $E_{k,h,f}(z) = z + F(z)T_k\sigma^{-1}(h)$.

We will need the following properties.

LEMMA 2.3 (De Branges' Theorem: Bieberbach conjecture). *Let $g(z) = z + g_2 z^2 + g_3 z^3 + \cdots$ be schlicht. Then $|g_k| \leq k$.*

Proof. See [2]. □

LEMMA 2.4 (Shub and Smale). (1) *Let g be schlicht. Then* $|g(h) - T_k g(h)| \leq (k+1)r^{k+1}/(1-r)^2$, *where* $r = |h| < 1$.

(2) *Let g be univalent on $D_H(0)$, $g(0) = 0$ and $g'(0) = 1$. Then for h with* $r = |h|/H < 1$,

$$|g(h) - T_k g(h)| \leq \frac{H(k+1)r^{k+1}}{(1-r)^2}.$$

Proof. From Lemma 2.3 we have

$$|g(h) - T_k g(h)| \leq \sum_{j=k+1}^{\infty} j r^j \leq r\left(\frac{r^{k+1}}{1-r}\right)' \leq \frac{(k+1)r^{k+1}}{(1-r)^2}.$$

For the second statement, note that $\frac{1}{H}g(Hh)$ is schlicht, and then use (1). □

LEMMA 2.5 (Koebe Distortion Theorem). *Let g be schlicht. Then for* $|h| = r < 1$,

$$\frac{r}{(1+r)^2} \leq |g(h)| \leq \frac{r}{(1-r)^2}, \tag{1}$$

$$\frac{1-r}{(1+r)^3} \leq |g'(h)| \leq \frac{1+r}{(1-r)^3}. \tag{2}$$

Proof. See [4, Vol. 2, pp. 351 and 353]. □

By rescaling, we obtain immediately the following

COROLLARY 2.6. *Let g be univalent on $D_H(0)$ and $g(0) = 0$, $g'(0) = 1$. Let* $r = |h|/H < 1$. *Then*

$$\frac{|h|}{(1+r)^2} \leq |g(h)| \leq \frac{|h|}{(1-r)^2}, \tag{1}$$

$$\frac{1-r}{(1+r)^3} \leq |g'(h)| \leq \frac{1+r}{(1-r)^3}. \tag{2}$$

Proof. Note that $\frac{1}{H}g(Hh)$ is schlicht. Now use Lemma 2.5. □

COROLLARY 2.7. *Let* $x = f_z^{-1}((1-h)f(z)) \equiv z + F(z)\sigma^{-1}(h)$ *for* $|h| < H_{f,z}$. *Then we have*

$$f'(x) = f'(z)\sigma'(\varepsilon) \equiv \frac{f'(z)}{\sigma^{-1\prime}(h)}, \quad \textit{where } \varepsilon = \frac{x-z}{F(z)}, \tag{1}$$

$$|f'(z)|\frac{(1-r)^3}{1+r} \leq |f'(x)| \leq |f'(z)|\frac{(1+r)^3}{1-r}, \quad \textit{where } r = \frac{|h|}{H_{f,z}}. \tag{2}$$

Proof. Recall from Lemma 2.1(2) that $f(x) = f(z)(1-\sigma(\varepsilon))$ and $\sigma(\varepsilon) = h \equiv (f(z) - f(x))/f(z)$. Hence (1) is immediate by taking derivatives of f. (2) follows from Corollary 2.6(2) since $\sigma^{-1}(0) = 0$, ${\sigma^{-1}}'(0) = 1$ and σ^{-1} is univalent in $D_H(0)$. □

We close this section with the following lemma.

LEMMA 2.8. (1) $R_{f,z} = |f(z) - f(\theta^*)| \geq \mathrm{Min}_{f'(\theta)=0}|f(z) - f(\theta)|$ *for some critical point θ^* of f.*

(2) *Let* $x = f_z^{-1}((1-h)f(z))$ *with* $|h|/H_{f,z} < 1$. *Then* $R_{f,x} \geq R_{f,z} - |f(z) - f(x)|$.

(3) *Let* $g = f - y$ *be a translation of* f *by* $y \in \mathbf{C}$. *Then* $E_{k,h',g}(x) = E_{k,h,f}(x)$, *where* $h' = hf(z)/g(z)$.

Proof. For (1), see Lemma 3 in [12].

For (2), we note that by the uniqueness of analytic maps, we have $f_x^{-1} \equiv f_z^{-1}$ on their common domain of definitions. In particular, f_x^{-1} is analytically continued for all w such that $|w - f(z)| < R_{f,z}$. Since $|w-f(z)| < |w-f(x)|+|f(x)-f(z)| < R_{f,z}$, f_x^{-1} is analytic for all w such that $|w - f(x)| < R_{f,z}\text{-}|f(z) - f(x)|$. Hence $R_{f,x} \geq R_{f,z}\text{-}|f(z) - f(x)|$.

For (3), note that $g_z^{-1}(w - y)$ is well defined where $f_z^{-1}(w)$ is well defined and $g_z^{-1}(w - y) = f_z^{-1}(w)$. As power series at $f(z)$ and $g(z)$ respectively, we have

$$f_z^{-1}(f(z) - w) \equiv g_z^{-1}(f(z) - w - y) \equiv g_z^{-1}(g(z) - w),$$

where $w = hf(z) = h'g(z)$ and $h' = h(f(z)/g(z))$. Hence we also have

$$T_k f_z^{-1}((1-h)f(z)) = T_k g_z^{-1}((1-h')g(z)) \quad \text{and} \quad E_{k,h',g}(z) = E_{k,h,f}(z). \quad \square$$

3. Koebe Distortion Theorem and Euler Iteration. We recall that $E_{k,h,f}(z) = T_k f_z^{-1}((1-h)f(z))$. In this section we show that $E_{k,h,f}$ approximates f_z^{-1} with a suitable h, i.e., $E_{k,h,f}(z) = f_z^{-1}(w)$ for w such that $|w - f(z)| < R_{f,z}$ and $R_{f,z}$ is the radius of convergence of f_z^{-1}. In particular, we show that $E_{k,h,f}$ approximates f_z^{-1} for all values on the disk of convergence as $k \uparrow \infty$. The main goal of this section is to prove Theorem 3.2 below.

We recall that $H_{f,z} = R_{f,z}/|f(z)|$, where $R_{f,z}$ denotes the radius of convergence of f_z^{-1} at $f(z)$.

THEOREM 3.1. *Let* $x = f_z^{-1}((1-h)f(z))$. *Assume that*

$$r = \frac{|h|}{H_{f,z}} < 1 \quad \text{and} \quad t \leq |h|\frac{(1-r)^3}{(1+r)^3}.$$

Then $D_{t|F|}(x) \subset f_z^{-1}(D_{|f(z)|s}(f(x)))$, *where*

$$s = \operatorname{Min}\left\{t\frac{(1+r)^3}{(1-r)^3},\ H_{f,z}(1-r)\right\} \quad \text{and} \quad F = \frac{-f(z)}{f'(z)}.$$

The proof will be given later. $\square$

Let

$$B_k(r) = (k+1)\frac{(1+r)^3}{(1-r)^5}r^k$$

and r_k be the smallest positive solution to $B_k(r) = 1$. Note that $B_k(r)$ is increasing on $[0, r_k]$. The condition that $|h| < r_k H_{f,z}$ is crucial for $E_{k,h,f}$ to approximate f_z^{-1}.

THEOREM 3.2. *Let* $z' = E_{k,h,f}(z)$ *with* $r = |h|/H_{f,z} < r_k$. *Then we have* $z' = f_z^{-1}((1-h')f(z))$ *and* $f(z')/f(z) = 1 - h + \varepsilon$, *where* $|\varepsilon| = |h - h'| \leq \operatorname{Min}\{|h|B_k(r), H_{f,z}(1-r)\}$.

A table of approximate values of r_k is given below.

TABLE 3.1

k	1	2	3	4	5	10	177	3303	47400
r_k	.148	.225	.282	.329	.367	.495	.9	.99	.999

Remark. We note that $r_k \uparrow 1$ as $k \uparrow \infty$. In [9], Shub and Smale showed that Theorem 3.2 holds for all $r \le \gamma_k$ where $\gamma_k \uparrow 0.175$ as $k \uparrow \infty$.

Proof of Theorem 3.2. Recall that $z' = E_{k,h,f}(z) = z + F(z)T_k\sigma^{-1}(h)$, where $F(z) = -f(z)/f'(z)$. Let $f(z') = (1-h')f(z)$. Let $x = f_z^{-1}((1-h)f(z)) = z + F(z)\sigma^{-1}(h)$. Then $|z' - x| = |F|\,|\sigma^{-1}(h) - T_k\sigma^{-1}(h)|$. Since σ^{-1} is univalent on $D_H(0)$, we have by Lemma 2.4,

$$t \equiv |\sigma^{-1}(h) - T_k\sigma^{-1}(h)| \le \frac{H_{f,z}(k+1)r^{k+1}}{(1-r)^2}$$
$$= |h|B_k(r)\frac{(1-r)^3}{(1+r)^3} \le |h|\frac{(1-r)^3}{(1+r)^3} \quad \text{for } r < r_k.$$

Now, by Theorem 3.1, we have $z' \in f_z^{-1}(D_{|f(z)|s}(f(x)))$ and

$$|f(z') - f(x)| = |(1-h')f(z) - (1-h)f(z)| = |f(z)|\,|h' - h| \le |f(z)|s.$$

Hence,

$$|\varepsilon| = |h' - h| \le s = \operatorname{Min}\{|h|B_k(r), H_{f,z}(1-r)\}$$

and we have $z' = f_z^{-1}((1-h')f(z))$ and $f(z')/f(z) = 1 - h'$ for some h' with $|h' - h| \le s$. □

We need the following lemmas to prove Theorem 3.1.

LEMMA 3.3. (1) *Let g be univalent on $D_R(z)$. Then $D_{Rt}(g(z)) \subset g(D_{Rs}(z)) \subset D_{Ru}(g(z))$, for any $s < 1$, $t = s|g'(z)|/(1+s)^2$ and $u = s|g'(z)|/(1-s)^2$.*

(2) *Suppose that g is univalent on $D_H(0)$, $g(0) = 0$ and $g'(0) = 1$. Let $z \in D_H(0)$, where $r = |z|/H$. Then for $s \le 1 - r$ we have $D_{Ht}(g(z)) \subset g(D_{Hs}(z)) \subset D_{Hu}(g(z))$, where $t = ((1-r)^3/(1+r)^3)\cdot s/(1-r+s)^2$ and $u = ((1+r)/(1-r))\cdot s/(1-r-s)^2$.*

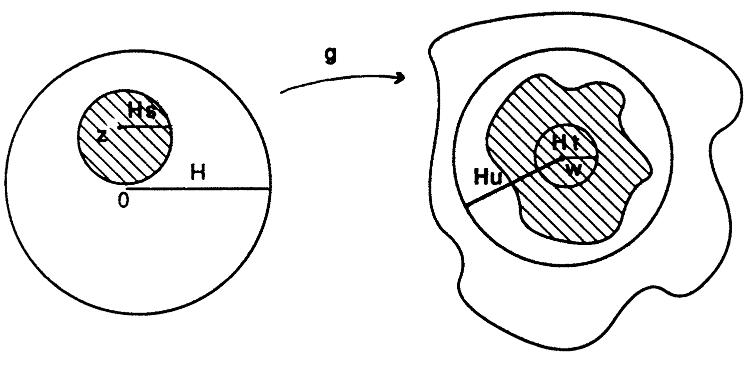

FIGURE 3.2. $w = g(z)$

Proof. (1) Let

$$\psi(h) = \frac{1}{Rg'(z)}(g(z+Rh) - g(z)).$$

Then it is easy to see that ψ is schlicht. Hence by Lemma 2.5, $\delta/(1+\delta)^2 \le |\psi(h)| \le \delta/(1-\delta)^2$ for $|h| = \delta$, so that we have

$$D_{R\mu}(g(z)) \subset g(D_{R\delta}(z)) \subset D_{R\eta}(g(z)),$$

where $R\mu = \delta R|g'(z)|/(1+\delta)^2$, $R\eta = \delta R|g'(z)|/(1-\delta)^2$. By setting $t = \mu$, $s = \delta$ and $u = \eta$, (1) is established.

For (2), note that g is univalent on $D_R(z)$, where $R = H(1-r)$, and hence $(1-r)/(1+r)^3 \leq |g'(z)| \leq (1+r)/(1-r)^3$ by Corollary 2.6. Let $s = (1-r)\delta$. Then we have

$$R\mu \geq \frac{\delta H(1-r)}{(1+\delta)^2}\frac{1-r}{(1+r)^3} = \frac{Hs}{(1-r+s)^2}\frac{(1-r)^3}{(1+r)^3} \equiv Ht,$$
$$R\eta \leq \frac{\delta H(1-r)}{(1-\delta)^2}\frac{1+r}{(1-r)^3} \leq \frac{Hs}{(1-r-s)^2}\frac{1+r}{1-r} \equiv Hu,$$

where

$$t = \frac{(1-r)^3}{(1+r)^3}\frac{s}{(1-r+s)^2} \quad \text{and} \quad u = \frac{1+r}{1-r}\frac{s}{(1-r-s)^2}.$$

Hence we have $D_{Ht}(g(z)) \subset g(D_{Hs}(z)) \subset D_{Hu}(g(z))$. □

Lemma 3.3 gives the following Quarter Theorem at an arbitrary point $z \in D_H(0)$.

COROLLARY 3.4. *Let g be univalent on $D_H(0)$ and $g(0) = 0$ and $g'(0) = 1$. For $z \in D_H(0)$, let $r = |z|/H$. Then $D_{Ht}(g(z)) \subset g(D_{H(1-r)}(z))$, where $t = \frac{1}{4}((1-r)^2/(1+r)^3)$.*

Proof. Use $s = 1-r$ in Lemma 3.3(2). □

In Lemma 3.3(2) we will also need to estimate s as a function of t.

COROLLARY 3.5. *Suppose $t \leq r((1-r)^3/(1+r)^3)$, where $r = |z|/H < 1$. Then $D_{Ht}(g(z)) \subset g(D_{Hs}(z))$, where $s = \mathrm{Min}\{t((1+r)^3/(1-r)^3),\ 1-r\}$.*

Proof. Since $r(1-r) \leq \frac{1}{4}$, we have $t \leq \frac{1}{4}((1-r)^2/(1+r)^3)$. Hence by Corollary 3.4 we have $D_{Ht}(g(z)) \subset D_{H(1-r)}(g(z))$. Let $t' = ((1-r)^3/(1+r)^3)\cdot s/(1-r+s)^2$. Since $s \leq 1-r$, Lemma 3.3(2) shows that $D_{Ht'}(g(z)) \subset g(D_{Hs}(z))$. However, since $s \leq t((1-r)^3/(1+r)^3) \leq r$, we have

$$t' = \frac{(1-r)^3}{(1+r)^3}\frac{s}{(1-r+s)^3} \geq \frac{(1-r)^3}{(1+r)^3}s = t.$$

Consequently, $D_{Ht}(g(z)) \subset D_{Ht'}(g(z)) \subset g(D_{Hs}(z))$, as claimed. □

The proof of Theorem 3.1 now follows easily from Corollary 3.5.

Proof of Theorem 3.1. Suppose that $z = f_z^{-1}((1-h')f(z))$ where $|z'-x| \leq |F|t$. Since

$$\begin{aligned} z' = f_z^{-1}((1-h')f(z)) &= z + F(z)\sigma^{-1}(h') \\ &= z + F(z)\sigma^{-1}(h) + F(z)(\sigma^{-1}(h') - \sigma^{-1}(h)) \\ &= x + F(z)(\sigma^{-1}(h') - \sigma^{-1}(h)), \end{aligned}$$

we have

$$\begin{aligned} |z'-x| &= |F|\,|\sigma^{-1}(h') - \sigma^{-1}(h)| \\ &\leq |F|t = |F|Ht', \quad \text{where } t' = \frac{t}{H} \leq \frac{|h|}{H}\frac{(1-r)^3}{(1+r)^3} = \frac{r(1-r)^3}{(1+r)^3}, \end{aligned}$$

by the hypothesis. Hence, by Corollary 3.5, we have $|h'-h| \leq Hs'$, where $s' = \mathrm{Min}\{t'((1+r)^3/(1-r)^3),\ 1-r\}$. Now by setting $s = Hs'$ we have the claim. □

4. Domain of Injectivity and a Notion of an Approximate Zero. The main goal of this section is to give a criterion to determine an approximate zero of a polynomial f for the modified Euler method. Hereafter we will denote $E_{k,h,f}$ by E_k if there is no confusion.

Definition. z_0 is an approximate zero of f for E_k if

$$\frac{|f(z_n)|}{|f(z_0)|} \leq \left(\frac{1}{2}\right)^{(k+1)^n}, \tag{1}$$

$$|z_n - \xi| \leq c\left(\frac{1}{2}\right)^{(k+1)^n} |z_0 - \xi|, \tag{2}$$

where $z_n = E^n_{k,1,f}(z_0) \to \xi$ and c is a constant.

We will need the following estimate of the domain of injectivity, which itself is quite interesting.

THEOREM 4.1. *Let $g(z) = z + a_2z^2 + \cdots$ be a power series and ψ be the compositional inverse of g taking 0 to 0. Let $a = \sup_i |a_i|^{1/(i-1)}$. Then ψ is well defined, analytic and one-to-one on $D_R(0)$, where $(3-\sqrt{8})/a \leq R$.*

Proof. Suppose that $|g(z) - z| < r$ on $|z| = r$. Then 0 is the only root of g in $D_r(0)$ by Rouché's Theorem. It follows that (see [1, Theorem 11, p. 131]) the inverse map ψ is well defined on $g(D_r(0))$. In particular, ψ is well defined on $D_R(0)$, where $R = \mathrm{Min}_{|z|=r} |g(z)|$. Now,

$$\begin{aligned} |g(z)| &= |z|\,|1 + a_2z + a_3z^2 + \cdots| \\ &\geq r|1 - ((ar) + (ar)^2 + (ar)^3 + \cdots)| \\ &\geq r\left(1 - \frac{ar}{1-ar}\right) \quad \text{on } |z| = r. \end{aligned}$$

But $r(1 - ar/(1-ar))$ achieves the maximum $(3-\sqrt{8})/a$ when $r = (2-\sqrt{2})/2a$. Also note that

$$\begin{aligned} |g(z) - z| &= |a_2z^2 + a_3z^3 + \cdots| = |z|\,|a_2z + a_3z^2 + \cdots| \\ &\leq r\frac{ar}{1-ar} < r, \quad \text{on } |z| = \frac{2-\sqrt{2}}{2a}. \end{aligned}$$

Hence ψ is well defined and injective on $D_R(0)$, where $R = (3-\sqrt{8})/a \approx 1/5.83a > 1/6a$. □

Remark 4.2. The corresponding upper bound $R \leq 4/a$ is obtained in [12, p. 9, Extended Loewner's Theorem]. For a polynomial f and $z \in \mathbf{C}$ we define

$$a_{f,z} \equiv \max_{j\geq 2} \left|\frac{f(z)}{f'(z)}\right| \left|\frac{f^{(j)}(z)}{j!f'(z)}\right|^{1/(j-1)}.$$

We apply Theorem 4.1 to a polynomial.

COROLLARY 4.3. *Let f be a polynomial of degree d and z be a complex number such that $f'(z) \neq 0$, and $f(z) \neq 0$. Let f_z^{-1} be the inverse branch of f such that $f_z^{-1}(f(z)) = z$. Then f_z^{-1}, as a power series at $f(z)$ has a radius of convergence $R_{f,z}$ satisfying $(3-\sqrt{8})/a \leq R_{f,z}/|f(z)| \leq 4/a$.*

Proof. Let σ be the polynomial associated with f as in Lemma 2.1. Since the radius of convergence of σ^{-1} at 0 is $H_{f,z} = R_{f,z}/|f(z)|$ by Lemma 2.1, we have the claim by the previous theorem. □

We now come to one of the main results.

THEOREM 4.4. *If $a_{f,z_0} \leq 1/48$, then z_0 is an approximate zero of f for E_k for all k. In other words, we have*

$$\frac{|f(z_n)|}{|f(z_0)|} \leq \left(\frac{1}{2}\right)^{(k+1)^n}, \tag{1}$$

$$|z_n - \xi| \leq 4\left(\frac{1}{2}\right)^{(k+1)^n} |z_0 - \xi|, \tag{2}$$

where ξ is a root and $z_{n+1} = E_{k,1,f}(z_n) \to \xi$.

Proof. We will proceed with the proof by an induction on n. For simplicity, we denote $R_n = R_{f,z_n}$, $f_n = f(z_n)$, $f'_n = f'(z_n)$, $H_n = R_n/|f_n|$ and $F_n = -f(z_n)/f'(z_n)$.

Claim 1. $|f_1|/|f_0| \leq (\frac{1}{2})^{k+1}$, for all k.

We note that $a_{f,z_0} \leq 1/48$ implies by Corollary 4.3 that

$$\frac{1}{H_0} = \frac{|f_0|}{R_0} < \frac{1}{48}\frac{1}{3-\sqrt{8}} < \frac{1}{8.23} < 0.122 < r_k$$

for all k (see Table 3.1). Hence we apply Theorem 3.2 with $h = 1$, and we have

$$z_1 = f_{z_0}^{-1}(f(z_1)) \quad \text{and} \quad \frac{|f_1|}{|f_0|} \leq B_k\left(\frac{1}{H_0}\right) \leq \left(\frac{1}{2}\right)^{k+1},$$

by noting that

$$B_k\left(\frac{1}{H_0}\right) < B_k(0.122) = \frac{(k+1)(1+0.122)^3(0.122)^k}{(1-0.122)^5} < \left(\frac{1}{2}\right)^{k+1} \quad \text{for } k \geq 2.$$

For $k = 1$, we recall from Lemma 2.1 that $f(z_1)/f(z_0) = 1 - \sigma \circ \varepsilon$, where $\varepsilon = (z_1 - z_0)/F_0 = 1$. Since

$$|1 - \sigma(1)| = |\sigma_2 + \sigma_3 + \cdots + \sigma_d| \leq \frac{a}{1-a} \leq \frac{1}{47} \leq \left(\frac{1}{2}\right)^2,$$

we have that $|f_1|/|f_0| \leq (\frac{1}{2})^{k+1}$ for all k as claimed. It is useful for the next claim to note that $|f_1|/|f_0| \leq 1/8$.

Claim 2. Suppose $|f_n|/|f_0| \leq (\frac{1}{2})^{(k+1)^n}$. Then $|f_{n+1}|/|f_0| \leq (\frac{1}{2})^{(k+1)^{n+1}}$. First note that $R_n \geq R_0 - \big||f_n| - |f_0|\big|$ by Lemma 2.8(2). Since $R_0/|f_0| \geq 8.23$ and $|f_n|/|f_0| \leq 1/8$ for all n and k, we have $R_n \geq 8.23|f_0| - \frac{9}{8}|f_0| \geq 7|f_0|$ for all n and k. Hence we have

$$\frac{1}{H_n} = \frac{|f_n|}{R_n} = \frac{|f_0|}{R_n}\frac{|f_n|}{|f_0|} \leq \frac{1}{7}\frac{|f_n|}{|f_0|} \leq \frac{1}{7}\left(\frac{1}{2}\right)^{(k+1)^n} \leq r_k\left(\frac{1}{2}\right)^{(k+1)^n}$$

for all k and n. Now, applying Theorem 3.2 with $h = 1$, we have $|f_{n+1}|/|f_n| \leq B_k(1/H_n)$. Since

$$B_k(r) = (k+1)\frac{(1+r)^3 r^k}{(1-r)^5} < B_k(r_k)\left(\frac{r}{r_k}\right)^k < \left(\frac{r}{r_k}\right)^k$$

for $r < r_k$ we have

$$\frac{|f_{n+1}|}{|f_n|} \le B_k\left(\frac{1}{H_n}\right) \le \left(\left(\frac{1}{2}\right)^{(k+1)^n}\right)^k = \left(\frac{1}{2}\right)^{k(k+1)^n}.$$

Hence we have

$$\frac{|f_{n+1}|}{|f_0|} = \frac{|f_{n+1}|}{|f_n|}\frac{|f_n|}{|f_0|} \le \left(\frac{1}{2}\right)^{k(k+1)^n}\left(\frac{1}{2}\right)^{(k+1)^n} = \left(\frac{1}{2}\right)^{(k+1)^{n+1}}$$

as claimed.

Claim 3. $|z_n - \xi| \le 4(\frac{1}{2})^{(k+1)^n}|z_0 - \xi|$.

First note that z_n defined here has $H_n > 8 > 1/r_k$ (see the proof of Claim 2). Hence σ^{-1} is well defined at all z_n, and we have $\xi = z_n + F(z_n)\sigma^{-1}(1)$ and $z_n - \xi = F(z_n)\sigma^{-1}(1)$, where σ is the polynomial associated with f and z_n.

By Corollary 2.6(1) we note that

$$\frac{|F(z_n)|}{(1+1/H_n)^2} \le |z_n - \xi| \le \frac{|F(z_n)|}{(1-1/H_n)^2}.$$

Hence we have

$$\begin{aligned}|z_n - \xi| &\le \frac{|F(z_n)|}{(1-1/H_n)^2} \le \frac{|F(z_n)|}{(1-1/H_n)^2}\frac{(1+1/H_0)^2}{|F(z_0)|}|z_0 - \xi| \\ &= \frac{|f_n|}{|f_0|}\frac{|f_0'|}{|f_n'|}\frac{(1+1/H_0)^2}{(1-1/H_n)^2}|z_0 - \xi|.\end{aligned}$$

We note that $(1+1/H_0)^2/(1-1/H_n)^2 \le 1.5$, since $1/H_0 \le 0.122$ and $1/H_n \le 1/28$ (see the proof of Claim 2). Further, we claim that $|f_0'|/|f_n'| \le (1+1/7)^3/(1-1/7) \le 1.8$, so that we have $|z_n - \xi| \le 4(\frac{1}{2})^{(k+1)^n}|z_0 - \xi|$. To see this, note that $z_n = f_{z_0}^{-1}((1-h)f(z_0))$ for $|h| = |f(z_n) - f(z_0)|/|f_0| \le 9/8$ and $|h|/H_0 \le (9/8)/8.23 = 1/7$. Now apply Corollary 2.7(2) with $r = 1/7$; we have $|f_0'|/|f_n'| \le (1+r)/(1-r)^3 \le 1.8$. Hence we have completed Claim 3. □

5. Algorithms. The main goal of this section is to construct new algorithms to find a root of a polynomial. Applied to any polynomial f, these new algorithms always converge to a root or a critical point of f. The underlying idea is that, for an initial point z_0, one analytically continues $f_{z_0}^{-1}$ toward 0 in a radial direction as long as it is possible. The idea used to determine the approximate zero in Section 3 is also useful.

As mentioned in Section 1, the radius of convergence (or equivalently, $a_{f,z}$) plays an important role as a successive overrelaxation parameter in our algorithms.

Recall that

$$a_{f,z} = \max_{j\ge 2}\left|\frac{f(z)}{f'(z)}\right|\left|\frac{f^{(j)}(z)}{j!f'(z)}\right|^{1/(j-1)}$$

from Section 4.

Now we describe the algorithms.

ALGORITHM A_k. For a polynomial f and a complex number $z_0 \in \mathbf{C}$, define iteratively,

$$z_{n+1} = E_{k,h_n,f}(z_n), \quad \text{where } h_n = \text{Min}\left(1, \frac{1}{48a_{f,z_n}}\right).$$

For example, if $k = 1$ we have $z_{n+1} = z_n - h_n(f(z_n)/f'(z_n))$. □

ALGORITHM B_k. For a polynomial f and $z_0 \in \mathbf{C}$, let $w_0 = f(z_0)$. Define iteratively

$$z_{n+1} = E_{k,1,g_n}(z_n),$$

where $g_n = f - w_{n+1}$, $w_{n+1} = (1 - h_n)w_n$, and $h_n = \mathrm{Min}(1, 1/1800a_{f,z_n})$.

Remark. Note that

$$z_{n+1} = E_{k,1,g_n}(z_n) = E_{k,h,f}(z_n),$$

where $h = (f(z_n) - w_{n+1})/f(z_n)$ by Lemma 2.8(3). For example, if $k = 1$ we have $z_{n+1} = z_n - (f(z_n) - w_{n+1})/f'(z_n)$.

THEOREM 5.A. *z_n in Algorithm A_k always converges to a root or a critical point of f.*

Proof. Note that once $a_{f,z_n} \leq 1/48$ (i.e., $h_n = 1$) then z_n is an approximate zero of f and converges to a root of f by Theorem 4.4. We may assume that $a_{f,z_n} > 1/48$ and hence $h_n \equiv 1/48a_{f,z_n} < \frac{1}{8}H_{f,z_n}$ by Corollary 4.3. Applying Theorem 3.2 with h_n, we obtain $|f(z_{n+1})|/|f(z_n)| \leq 1 - h'_n$, where $|h'_n - h_n| \leq B_k(\frac{1}{8})h_n \leq \frac{3}{4}h_n$ for all k. Inductively one has $f(z_N)/f(z_0) = \prod^N(1 - h'_n)$, where $|h'_n - h_n| \leq \frac{3}{4}h_n$. Notice that $|f(z_n)|/|f(z_0)|$ converges always since it is decreasing. We will show that z_n converges to a critical point of f, if $|f(z_n)|/|f(z_0)|$ converges to a nonzero number. Recall from the theory of infinite products that this implies that $\sum b_n$ is bounded, where $1 - b_n = |1 - h'_n|$. Note that $\sum h_n$ and $\sum |h'_n|$ are also bounded since $b_n \geq h_n - |h'_n - h_n| \geq \frac{1}{4}h_n \geq \frac{1}{16}|h'_n|$ by (1). Again by the theory of infinite products we have $\prod(1 - h'_n) \to w$, a nonzero complex number. This w is a critical value of f since $h_n \to 0$ and $|f'(z_n)| \to 0$ by the definitions of h_n and a_{f,z_n}. Further, we claim that $|z_{n+1} - z_n| \to 0$ and z_n converges to a critical point θ. To see this, just note that

$$\begin{aligned} a_{f,z} &= \max_{j=2,\ldots,d} \left|\frac{f(z)}{f'(z)}\right| \left|\frac{f^{(j)}(z)}{j!f'(z)}\right|^{1/(j-1)} \geq \left|\frac{f(z)}{f'(z)}\right| \left|\frac{f^{(d)}(z)}{d!f'(z)}\right|^{1/(d-1)} \\ &\geq \left|\frac{f(z)}{f'(z)}\right| \left|\frac{1}{f'(z)}\right|^{1/(d-1)}. \end{aligned}$$

Hence

$$|z_{n+1} - z_n| = h_n \frac{|f(z_n)|}{|f'(z_n)|} \leq \frac{1}{48}|f'(z_n)|^{1/(d-1)} \to 0.$$

Since there are finite preimages of w, we conclude that $z_n \to \theta$ where $w = f(\theta)$. □

THEOREM 5.B. *For any polynomial f and $z_0 \in \mathbf{C}$, z_n in Algorithm B_k converges to a root or a critical point of f. Further, z_n converges to a root unless there is a critical value of f on the ray $(0, f(z_0)]$.*

Proof. It is easy to see that once $h_n = 1$ (i.e., $a_{f,z_n} \leq 1/1800 \leq 1/48$) then z_n is an approximate zero of f and hence z_n converges to a root of f by Theorem 4.4. Note that if $H_n \geq 7200$ then $h_n = 1$, and z_n is an approximate zero by Corollary 4.3. We will show inductively that either z_n is an approximate zero or z_n satisfies the bound

(1) $$\frac{w_n}{f(z_n)} = 1 + \varepsilon_n, \qquad |\varepsilon_n| \leq \frac{H_n}{14400} \leq \frac{1}{2}.$$

For simplicity, we denote $f_n = f(z_n)$, $R_n = R_{f,z_n}$, $H_n = R_n/|f_n|$. We claim that (1) completes the proof: Recall that $R_n = |f_n - f(\theta^*)|$ for some critical point θ^* by Lemma 2.8(1) and that $H_{f,z_n}/7200 \le h_n \le H_{f,z_n}/308$ for $h_n < 1$ by Corollary 4.3. Now in the case $h_n < 1$,

$$\begin{aligned}\frac{|w_n - f(\theta^*)|}{|w_n|} &= \frac{|f_n|}{|w_n|}\frac{|w_n - f_n| + |f_n - f(\theta^*)|}{|f_n|} \le \frac{1}{|1+\varepsilon_n|}(|\varepsilon_n| + H_n) \\ &\le 2\left(\frac{H_n}{14400} + H_n\right) \quad \text{since } |\varepsilon_n| \le \frac{1}{2} \\ &\le 2.5\ H_n \le 20000 h_n.\end{aligned}$$

Using the same argument as in Theorem 5.A, $w_n = \prod_{m=1}^{n}(1-h_m)$ converges to a nonzero number only if $h_n \to 0$ and hence only if $|w_n - f(\theta^*)| \to 0$. Since there is no critical value on $(0, w_0]$, this is possible only if $w_n \to f(\theta^*) = 0$. Again using the same argument as in Theorem 5.A, we conclude that $z_n \to \theta^*$ where $f(\theta^*) = 0$. Now we start an induction to show (1). Suppose $f_n/w_n = 1 + \varepsilon_n$, $|\varepsilon_n| \le H_n/14400 \le 1/2$. Then we will show that either z_{n+1} is an approximate zero or it satisfies $f_{n+1}/w_{n+1} = 1 + \varepsilon_{n+1}$, $|\varepsilon_{n+1}| \le H_{n+1}/14400 \le 1/2$. Recall that $z_{n+1} = E_{k,h,f}(z_n)$, where $h = (f_n - w_{n+1})/f_n$ and $w_{n+1} = (1-h_n)w_n$. Note that

$$\begin{aligned}|h| &= \frac{|f_n - w_{n+1}|}{|f_n|} = \frac{|f_n - (1-h_n)w_n|}{|f_n|} = \frac{|f_n - (1-h_n)(1+\varepsilon_n)f_n|}{|f_n|} \\ &= |1 - (1-h_n)(1+\varepsilon_n)| = |h_n - \varepsilon_n(1-h_n)| \le h_n + |\varepsilon_n| \\ &\le \frac{H_n}{308} + \frac{H_n}{14400} \le \frac{H_n}{300}.\end{aligned}$$

Applying Theorem 3.2 to z_n with h, we have

$$\begin{aligned}\frac{f_{n+1}}{f_n} &= 1 - h + h\delta, \quad \text{where } |\delta| \le B_k\left(\frac{1}{300}\right) \le \frac{1}{145} \text{ for all } k, \\ \frac{f_{n+1}}{w_{n+1}} &= 1 + \frac{h\delta}{1-h}, \quad \text{since } w_{n+1} = (1-h)f_n,\end{aligned}$$

and

$$\frac{w_{n+1}}{f_{n+1}} = 1 + \varepsilon_{n+1} = \frac{1}{1+\mu}, \qquad \mu = \frac{h\delta}{1-h}.$$

Note that

$$\begin{aligned}H_{n+1} = \frac{R_{n+1}}{|f_{n+1}|} &\ge \left|\frac{f_n}{f_{n+1}}\right| \frac{R_n - |f_n - f_{n+1}|}{|f_n|} \quad \text{by Lemma 2.8(2)} \\ &\ge \frac{1}{|1-h+h\delta|}|H_n - |h - h\delta|| \\ &\ge \frac{1}{|1-h+h\delta|}\left(H_n - \frac{H_n}{300}\left(1 + \frac{1}{145}\right)\right) \\ &\ge \frac{296}{297}\frac{H_n}{|1-h+h\delta|}.\end{aligned}$$

Now

$$\begin{aligned}|\mu| &= \left|\frac{h\delta}{1-h}\right| \le \frac{\frac{H_n}{300}\frac{1}{145}}{|1-h|} \le \left|\frac{1-h+h\delta}{1-h}\right| H_{n+1}\frac{297}{296}\frac{1}{300}\frac{1}{145} \\ &= |1+\mu|\frac{H_{n+1}}{43000}.\end{aligned}$$

Note that if $|\mu| \geq \frac{1}{4}$, then

$$H_{n+1} \geq \frac{|\mu|}{|1+\mu|} 43000 \geq 7200$$

and hence z_{n+1} is an approximate zero. If z_{n+1} is not an approximate zero then $H_{n+1} < 7200$ and $|\mu| < \frac{1}{4}$. Hence we have

$$|\varepsilon_{n+1}| \leq 2|\mu| \leq 2 \cdot \frac{5}{4} \cdot \frac{H_{n+1}}{43000} \leq \frac{H_{n+1}}{30000} \leq \frac{H_{n+1}}{14400} \leq \frac{1}{2}. \quad \square$$

Acknowledgment. This work is partly extracted from the author's thesis at CUNY and I would like to thank my advisor, Mike Shub, for stimulating discussions and encouragement throughout graduate school. I also would like to thank Professor Steve Smale for his interest in this work. I would like to express my very special thanks to the referee of this paper for his patience, for his careful reading, and for his helpful comments.

Department of Mathematics
University of Southern California
Los Angeles, California 90089

1. L. V. AHLFORS, *Complex Analysis*, International Series in Pure and Applied Math., 1979.

2. L. DE BRANGES, "A proof of the Bieberbach conjecture," *Acta Math.*, v. 154, 1985, pp. 137–152.

3. P. HENRICI, *Applied and Computational Complex Analysis*, Wiley, New York, 1977.

4. E. HILLE, *Analytic Function Theory*, Ginn, Boston, Mass., 1962.

5. M. KIM, *Computational Complexity of the Euler Type Algorithms for the Roots of Complex Polynomials*, Thesis, City University of New York, 1985.

6. M. KIM, "Improved criteria for an approximate zero and experimental results on the polynomial root finding problem," in preparation.

7. M. MARDEN, *Geometry of Polynomials*, Math. Surveys, No. 3, Amer. Math. Soc., Providence, R. I., 1966.

→ C. MCMULLEN, "Families of rational maps and iterative root-finding algorithms," *Ann. of Math.*, v. 125, 1987, pp. 467–494.

9. M. SHUB & S. SMALE, "Computational complexity: On the geometry of polynomials and a theory of cost; Part I," *Ann. Sci. École Norm. Sup. Ser.* (4), v. 18, 1985, pp. 107–142.

10. M. SHUB & S. SMALE, "Computational complexity: On the geometry of polynomials and a theory of cost; Part II," *SIAM J. Comput.*, v. 15, 1986, pp. 145–161.

11. M. SHUB & S. SMALE, "On the existence of generally convergent algorithms," *J. of Complexity*, v. 2, 1986, pp. 2–11.

12. S. SMALE, "The fundamental theorem of algebra and complexity theory," *Bull. Amer. Math. Soc.*, v. 4, 1981, pp. 1–36.

13. S. SMALE, "On the efficiency of algorithms of analysis," *Bull. Amer. Math. Soc.*, v. 13, 1985, pp. 87–121.

14. S. SMALE, "Newton's method estimates from data at one point," *The Merging Disciplines: New Directions in Pure, Applied and Computational Mathematics*, pp. 185–196, Springer-Verlag, 1986.