ANNALES SCIENTIFIQUES DE L'É.N.S.

MIKE SHUB
STEVE SMALE

Computational complexity. On the geometry of polynomials and a theory of cost. I

Annales scientifiques de l'É.N.S. 4e série, tome 18, nº 1 (1985), p. 107-142
<http://www.numdam.org/item?id=ASENS_1985_4_18_1_107_0>

© Gauthier-Villars (Éditions scientifiques et médicales Elsevier), 1985, tous droits réservés.

L'accès aux archives de la revue « Annales scientifiques de l'É.N.S. » (http://www.elsevier.com/locate/ansens) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques
http://www.numdam.org/

Ann. scient. Éc. Norm. Sup.,
4 série, t. 18, 1985, p. 107 à 142.

COMPUTATIONAL COMPLEXITY
On the geometry of polynomials and a theory of cost: Part I

BY MIKE SHUB AND STEVE SMALE (*)

The main goal of this paper is to show that the number of iterations required for a certain fast algorithm to find a zero of a complex polynomial of degree d, is linear in d, provided that an arbitrarily small set of problems is excluded.

The algorithm which does this is an "incremental" version of that of Newton and Euler.

This result is part of a broad study of algorithms for polynomial root finding. Our work could also be considered as an investigation into the geometry of a complex polynomial f. This geometry is especially concerned with the foliation of the complex numbers $\mathbb{C}$ defined by the lifting of rays by f^{-1}. This curve family branches at the critical points θ of f and constitutes solutions to Newton's differential equation

$$\frac{dz}{dt} = -\frac{f(z)}{f'(z)}.$$

Let $P_d(1)$ be the set of polynomials $f(z)=z^d+a_{d-1}z^{d-1}+\ldots+a_1 z+a_0$ satisfying $|a_i| \leqq 1$. Let S_R^1 be the set of complex numbers z_0 with $|z_0|=R$ and D_R^1 be the set of complex numbers z_0 with $|z_0| \leqq R$. Impose Lebesque measure on $S_R^1 \times P_d(1)$ or $D_R^1 \times P_d(1)$ normalized to total measure 1. An approximate zero of a polynomial f is a complex z for which Newton's (and Euler's) method converges in a very strong sense (*see* the Corollary to Theorem 1). As a consequence of Proposition 3 of section 2 one has that the measure of the set of (z, f) in $D_R^1 \times P_d(1)$ such that z is an approximate zero of f is greater than c/d^5 where c is about 1/1,000. This gives an answer to Problem 7 *Smale*.

(*) Both authors received partial support from the N.S.F.

Main Theorem. – *Given* $0<\mu<1$, $d>1$, *there is an* $R=R(\mu, d)>0$ *and an iterative algorithm* $E=E(\mu, d)$ *such that for* (z_0, f) *in* $S^1_R \times P_d(1)$, *with probability* $1-\mu$, *and*

$$s=L_1 d\left(\frac{|\log \mu|}{\mu}\right)^{1+(1/\log d)}+L_2,$$

then $z=(E)^s(z_0)$ *is an approximate zero.*

Here L_1, L_2 are constants, L_1 less than 20, L_2 smaller yet, and $E(\mu, d)$ is a variation of an algorithm that was used by Euler. It is robust and executed with relatively small cost.

The s is the number of steps in the iteration. The main feature of this result is the good estimate for s. It is already difficult to obtain any polynomial bound (in *Smale*, a result is obtained using Newton's method with d^9; this reference also gives some background to our paper here).

The algorithm $E(\mu, d)$ is incremental Euler E_k where we give the increment h as an explicit function of μ, d. It is noteworthy that the k depends on d and is simply the smallest integer greater than log d. From the definition of E_k, k is the number of derivatives evaluated at each step and so depends on d simply. The idea perhaps could be useful in the pratice of equation solving.

One problem the result poses is why choose $|z_0|=R$ which grows like d and contributes the factor d in the estimate of s? It would seem that choosing $|z_0|\leqq 1$ is more sensible, but the corresponding analysis becomes especially difficult.

As it stands now the above theorem, besides requiring the long proof below, depends on mathematics related to the Bieberbach conjecture. The Bieberbach conjecture itself would only slightly improve the constants in the result.

The main theorem of this paper does not distinguish between the intrinsic difficulty of finding an approximate zero for a particular f, and the difficulty of finding a good starting point z_0 for f. In part II of this paper we will separate these problems and show that the k-th incremental Euler algorithms may be adapted to produce probabilistic and deterministic algorithms for finding approximate zeros for $f\in P_d(1)$ with the average number of steps and arithmetic operations required $O(d \log d)$ and $O(d^2(\log d)^2 \log \log d)$ respectively.

Section 1

An *incremental algorithm* is given by a map

$$I_{h,f}: \quad S^2 \to S^2, \qquad z'=I_{h,f}(z)=I(z),$$

parameterized by $0\leqq h\leqq 1$ and complex polynomials f. Here S^2 denotes the Riemann sphere of complex numbers. We suppose always that

$$I_{0,f}(z)=z$$

and so $\mathrm{I}_{h,f}$ is a parameterized family of algorithms starting at the identity. Eventually continuity conditions will be put on I.

To solve $g(z)=0$ by using an incremental algorithm $\mathrm{I}_{h,f}$ one lets z_0 be a complex number, and chooses h appropriately. In a number of situations, the sequence

$$z_n=\mathrm{I}_{h,g}(z_{n-1})=I^n_{h,g}(z_0), \qquad n=1, 2, 3, \ldots,$$

will converge to a solution of $g(z)=0$.

Most of our examples of incremental algorithms are derived from some standard iterative process or scheme for solving either non-linear systems or ordinary differential equations. We frequently call the maps $\mathrm{I}_{h,f}$ iterative or iteration processes or schemes. We will sometimes assume $f(z)\neq 0$, and $f'(z)\neq 0$ when the context requires it, for example to be sure we are not dividing by zero below.

Example 1. — Incremental Newton's method (*see* Smale for background)

$$\mathrm{I}_{h,f}(z)=z-\frac{hf(z)}{f'(z)}.$$

The special case $\mathrm{N}(z)=z-f(z)/f'(z)$ (for $h=1$) is Newton's method.

Example 2_k. — Derived from k-th order Taylor's method, $k=1, 2, \ldots$ (*see Atkinson*)

$$I(z)=z+\sum_{i=1}^{k}\frac{d^i}{dt^i}\left(\frac{\varphi_t(z)}{i!}\right)\Big/_{t=0} h^i,$$

where $\varphi_t(z)$ is the solution of the differential equation

$$\frac{dz}{dt}=\frac{-f(z)}{f'(z)}=\mathrm{F}(z),$$

with initial condition $\varphi_0(z)=z$.

The quantities $(d^i/dt^i)\varphi_t(z)/_{t=0}$ are of course computable from F and its derivatives. Examples 2_1 and Example 1 coincide. Example 2_2 is explicitly given by:

$$\mathrm{I}(z)=z+\mathrm{F}(z)\left(h+\frac{h^2}{2}\mathrm{F}'(z)\right); \qquad \mathrm{F}'=\frac{f''f}{(f')^2}-1.$$

Example 2_3 is :

$$\mathrm{I}(z)=z+\mathrm{F}(z)\left[h+\frac{h^2}{2}\mathrm{F}'(z)+\frac{h^3}{3}((\mathrm{F}'(z))^2+\mathrm{F}''(z)\,\mathrm{F}(z))\right].$$

We continue to write

$$\mathrm{F} \quad \text{for} \quad \frac{-f}{f'}.$$

Example 3. — (derived from simple Runge-Kutta; *see Atkinson*)

$$I(z)=z+\frac{h}{2F(z)}[F(z)+F(z+hF(z))].$$

Example 4. — We have taken this algorithm from *Durand*, p. 25 ff., where it appears with $h=1$

$$I(z)=z-\frac{hf(z)f'(z)}{(f'(z))^2-(1/2)hf(z)f''(z)}=z+F(z)\left(\frac{h}{1-(1/2)h[f(z)f''(z)/(f'(z))^2]}\right).$$

In *Durand*, p. 69 it is pointed out that this iterative method (with $h=1$) has order 3 for simple zeros.

The following construction is important for the next example and is also used throughout our analysis.

Given a polynomial f and a point z such that $f'(z)\neq 0$, denote the map given by a power series for example, which takes $f(z)$ to z and is a compositional inverse to f by f_z^{-1}. Note as in *Smale* that the radius of convergence $r=r(f, z)$ of f_z^{-1} is $|f(z)-f(\theta_*)|$ for some θ_* such that $f'(\theta_*)=0$. Thus $f_z^{-1}: D_r(f(z))\to\mathbb{C}$ is an injective analytic function sending $f(z)$ to z, where $D_r(f(z))$ is the disk of radius r about $f(z)$. It is a "branch" of f^{-1}.

If $f'(z)=0$, let $r(f, z)=0$.

If $f(z)\neq 0$ define $h_1=h_1(f, z)=r(f, z)/|f(z)|$. Thus

$$h_1(f, z)\geqq \min_{\substack{\theta \\ f'(\theta)=0}} \frac{|f(z)-f(\theta)|}{|f(z)|}.$$

If $\theta_*=\theta_*(f, z)$ is one of the critical points θ for which $r=|f(z)-f(\theta_*)|$, then

$$h_1(f, z)=\frac{|f(z)-f(\theta_*)|}{|f(z)|}.$$

Example 5_k. — Incremental Euler, $k=1, 2, \ldots, \infty$.

This is our most important example and we take some time to develop it carefully. The evidence of this paper suggests it is the most appropriate for practically computing zeros of complex polynomials.

If $h<h_1(f, z)$ we may solve the equation

$$\frac{f(z')}{f(z)}=1-h,$$

by setting $z'=f_z^{-1}((1-h)f(z))$.

Expanding around $f(z)$, (Taylor's Series) we obtain

$$(5_\infty) \qquad \mathrm{E}_\infty(z)=z'=z+\sum_{l=1}^{\infty}\frac{(f_z^{-1})f^{(e)}(z)}{l!}(-hf(z))^l.$$

Evaluation of (5_∞) is usually computationally infeasible, so we truncate. Let τ_d be the operation of truncating a power series,

$$\tau_d\left(\sum_{l=0}^{\infty}a_l h^l\right)=\sum_{l=0}^{d}a_l h^l.$$

The k-th incremental Euler E_k or $\mathrm{E}_{k(h,\,f)}$ is given by

$$(5_k) \qquad \mathrm{E}_k(z)=\tau_k f_z^{-1}((1-h)f(z))=z+\sum_{l=1}^{k}\frac{(f_z^{-1})f^{(e)}(z)}{l!}(-hf(z))^l.$$

We end our discussion of this example by a series of remarks:

(1) Suppose $h_1>1$. Then one can put $h=1$ in (5_∞) to obtain a power series representation of a solution to $f(z)=0$.

(2) One can easily write down

$$\mathrm{E}_4(z)=z-\frac{f(z)}{f'(z)}(h-\sigma_2h^2+(2\sigma_2^2-\sigma_3)h^3-(5\sigma_2^3-5\sigma_2\sigma_3+\sigma_4)h^4),$$

adapting the computations of *Durand*, p. 5 and using

$$\sigma_i=(-1)^{i-1}\frac{f^{(i)}(z)\,(f(z))^{i-1}}{i!\,(f'(z))^i}.$$

Note that by keeping only the first k powers of h, $k=1$, 2 or 3, we obtain E_k for those values. In particular E_1 is just incremental Newton of Example 1.

In the literature the algorithms E_k with $h=1$ are sometimes accredited to Euler and sometimes to Shroeder. According to *Durand*, the case E_∞ with $h=1$ was reasoned by Euler. On the other hand both Henrici and Householder refer to Shroeder for algorithms which seem to amount to E_k with $h=1$. Ostrowski says that the power series for the solution goes back to Newton and Euler. Euler in his Caput IX, De Usu Calculi Differentialis in Aequationibus Resolvendis writes, p. 428:

"Si igitur f ponatur designare istum ipsius x
valorem que erit radix aequationis $y=0$, quoniam
x abit in f, si statuator $y=0$, erit per ea,
quae supra [§ 67] sunt demonstrata,

$$f=x-y\frac{dx}{dy}+\frac{y^2\,d\,dx}{2\,dy^2}-\frac{y^3\,d^3x}{6\,dy^3}+\frac{y^4\,d^4x}{24\,dy^4}-, \text{ etc.},$$

In qua aequatione statuitor differentiale dy constans."

which gives the power series solution for the root and is exactly E_∞ with $h=1$. Euler goes on to solve equations iterating E_k with $h=1$ up to $k=5$. We have thus called the algorithms k-th (incremental) Euler and denoted them E_k. *Shroeder* gives a much more systematic and modern treatment of these algorithms, dealing with such questions as convergence and order.

In the thesis of Gregg *Saunders*, these and related algorithms are studied from a little different point of view.

Next we generalize examples 5_k.

Example 6_k. — Generalized Incremental Euler.

First suppose $c_1>0$, c_2, ..., c_k are given real parameters. Let $P(h)=c_1 h+c_2 h^2+\ldots+c_k h_k$. Then if h is small enough so that $|P(h)|<h_1$, we may solve

$$\frac{f(z')}{f(z)}=1-P(h) \quad \text{by } z'=f_z^{-1}((1-P(h))f(z)).$$

This motivates

$$(6_k) \qquad GE_{P,\,k,\,h,\,f}(z)=GE_k(z)=\tau_k f_z^{-1}((1-P(h))\,f(z)),$$

or

$$GE_k(z)=z+\tau_k \sum_{l=1}^{\infty} \frac{(f_z^{-1})^{(l)}_{f(z)}}{l!}((-P(h)\,f(z))^l.$$

Clearly (5_k) is the special case of (6_k) obtained by setting $c_1=1$, $c_i=0$, $i>1$. Moreover example 2_k (derived from k-th order Taylor's method) is obtained by putting $c_i=(-1)^{i-1}/i!$, $i=1, \ldots, k$. This may be seen as follows.

Let $\varphi_t(z)$ be the solution of the differential equation $dz/dt=-f(z)/f'(z)$ with initial condition $\varphi_0(z)=z$. Then it is easy to see that

$$f(\varphi_t(z))=e^{-t} f(z)$$

and thus the k-th order Taylor expansion of $\varphi_t(z)$ is given by

$$\tau_k f_z^{-1}(e^{-t} f(z))=z+\tau_k \sum_{l=1}^{\infty} \frac{(f_z^{-1})^{(l)}}{l!}\left(\sum_{i=1}^{\infty} \frac{(-t)^i}{i!} f(z)\right)^l,$$

where the derivatives of f_z^{-1} are evaluated at $f(z)$.

Returning to the general case, we may express

$$GE_k(z)=z+F\sum_{j=0}^{k-1} P_j h^{j+1},$$

where P_j are polynomials in the c_i and σ_l, where σ_l is defined above.

PROPOSITION 1:

$$P_0=c_1,$$
$$P_1=c_2-\sigma_1 c_1^2,$$
$$P_2=c_3-2\sigma_2 c_1 c_2-(\sigma_3-2\sigma_2^2)c_1^3$$

and $P_k=P_k(\sigma_1, \ldots, \sigma_{k+1}, c_1, \ldots, c_{k+1})$ *has the property: in each term the subscripts of* c_i, σ_j *times the powers to which they are raised sum to k less the number of factors. One can write down* P_k *explicitly inductively, in terms of* P_i, $i<k$, σ_{k+1} *and* c_{k+1}.

The easiest way to prove this proposition is to make a little detour. Introduce the polynomial $\sigma(w)=\sum_{i=1}^{d}\sigma_i w_i$ where as above

$$\sigma_i=(-1)^{(i-1)}\frac{f^{(i)}(z)(f(z))^{i-1}}{i!(f'(z))^i}$$

and f has degree d. This polynomial will be useful later. Note that $\sigma(0)=0$ and $\sigma'(0)=1$. If we write an incremental algorithm

$$z'=I_{h,f}(z) \quad \text{as} \quad z'=I_{h,f}(z)=z+FR(h, f, z).$$

Then by Taylor Series

$$f(z')=f(z)+\sum_{i=1}^{d}\frac{f^{(i)}(z)}{i!}(z'-z)^i$$
$$=f(z)+\sum_{i=1}^{d}\frac{f^{(i)}(z)}{i!}F^i R^i$$
$$=f(z)-f(z)\sum_{i=1}^{d}\sigma_i R^i.$$

Thus:

PROPOSITION 2:

$$\frac{f(z')}{f(z)}=1-\sigma\circ R \qquad \textit{where} \quad R=\frac{I_{h,f}(z)-z}{F}.$$

We apply this to an example.

For

$$I_{h,f}=E_{\infty(h,f)}, \qquad \frac{f(z')}{f(z)}=1-h=1-\sigma\circ R.$$

Thus $h=\sigma\circ \mathrm{R}$. Since $\sigma(0)=0$ and $\sigma'(0)=1$ σ is invertible on some disc to σ^{-1} with $\sigma^{-1}(0)=0$ and $(\sigma^{-1})'(0)=1$. Thus we have

$$\sigma^{-1}(h)=\mathrm{R}=\frac{1}{\mathrm{F}}\sum_{l=1}^{\infty}\frac{(f_z^{-1})_{f(z)}^{(l)}}{l!}(-f(z))^l h^l.$$

If we expand σ^{-1} around 0

$$\sigma^{-1}(h)=\sum_{i=1}^{\infty}\frac{(\sigma^{-1})^{(i)}(0)}{i!}h^i,$$

it follows that $(\sigma^{-1})_0^{(l)}=(1/\mathrm{F})(f_z^{-1})_{(f(z))}^{(l)}(-f(z))^l$ and thus σ^{-1} is defined and injective on the open disk of radius h_1 around 0.

Now we return to Proposition 1.

$$\mathrm{GE}_k(z)=z+\tau_k\sum_{l=1}^{\infty}\frac{(f_z^{-1})_{(f(z))}^{(l)}}{l!}(-f(z))^l(\mathrm{P}(h))^l=z+\mathrm{F}\,\tau_k\sum_{l=1}^{\infty}\frac{(\sigma^{-1})^{(l)}}{l!}(\mathrm{P}(h))^l.$$

Now the proof of the Proposition is easily seen by applying inductive formulas for the coefficients of the inverse of a power to σ *see Durand*, p. 4 and formulas for composition, *see Henrici.*

Our study of incremental algorithms focuses on $f(z')/f(z)$ and its Taylor expansion in h. This "target space" approach becomes clearer as the paper develops.

The idea is to consider the curve $h\to f(\mathrm{I}_{h,\,f}(z))$ in the target space for small positive h. For an ideal algorithm, for all f and z, as h increases from 0, $f(\mathrm{I}_{h,\,f}(z))$ moves along the ray from $f(z)$ to 0. No practical incremental algorithm accomplishes this, but one of the main results of this paper is that some algorithms do this infinitesimally up to any order of contact at $f(z)$ and with a certain uniformity. This motivates the following definition which measures the efficiency of an incremental algorithm.

An *incremental algorithm* $\mathrm{I}_{h,\,f}$ will be said to be of *efficiency* k provided: there exist real constants $\delta>0$, $\mathrm{K}>0$, $c_1, \ldots, c_k$, $c_1>0$ independent of h, f and z such that

$$\frac{f(\mathrm{I}_{h,\,f}(z))}{f(z)}=1-(c_1 h+c_2 h^2+\ldots+c_k h^k)+\mathrm{S}_{k+1}(h),$$

where $|\mathrm{S}_{k+1}(h)|\leqq \mathrm{K}\,h^{k+1}\max(1,\,1/h_1^k)$ for $0<h\leqq\delta(\min(1,\,h_1))$ and $h_1=h_1(f,\,z)$ is defined as above.

In this case we also simply say that the iteration $\mathrm{I}_{h,\,f}$ is of efficency k.

In section 2, we will show that E_k is of efficiency k. In section 3 we use this fact to track the iterates of E_k in the target space. Later we characterize the incremental algorithms of efficiency k. It will follow that all of the above examples are of efficiency k, for appropriate k. Moreover, the set of all (small) incremental algorithms which are polynomials in h of degree k and of efficiency k are precisely those of Example 6_k.

Section 2

The goal of this section is to prove:

THEOREM 1. — *For any k, $0<k\leqq\infty$, the incremental k-th Euler algorithm (example 5_k of section 1) is of efficiency k. More precisely, there is a universal constant $1\leqq B\leqq 1.07$, and for any polynomial f, complex number z with $f'(z)\neq 0$, $f(z)\neq 0$, $z'=E_k(z)=E^{(z)}_{k(h,f)}$*

$$\frac{f(z')}{f(z)}=1-h+Q(h,f,z)\frac{h^{k+1}}{h_1^k},\qquad |Q|\leqq\beta_k(\gamma),$$

where

$$\beta_k(\gamma)=\frac{B(k+1)(1-\gamma)^2}{[(1-\gamma)^2-4\gamma][(1-\gamma)^2-4\gamma(1+B(k+1)\gamma^k)]}.$$

Here $\gamma=h/h_1$ is assumed to satisfy $0<\gamma<\gamma_k$ where γ_k is the first positive number for which the denominator of $\beta_k(\gamma)$ vanishes. Otherwise said, $0<h<\gamma_k h_1$. Here $h_1=h_1(f,z)$ is the function defined in section 1.

We first give a proof of Theorem 1. At the end of the section is some discussion and the first implication of the theorem.

We recall some results form the theory of Schlicht functions, related to the Bieberbach conjecture, *Duren*. One relation to our work comes from the fact that if f is a polynomial, then f_z^{-1} is defined and injective on the disc of radius $r(f,z)$ about $f(z)$.

An analytic function $f: D_1\to\mathbb{C}$ defined on the unit disc, given by the power series $f(z)=\sum_{i=1}^{\infty}a_i z^i$ is called Schlicht if it converges, $a_1=1$, and is one to one for $|z|<1$. Under these conditions there is the classic:

Bieberbach Conjecture: $|a_l|\leqq l$, $l=2, 3, \ldots$

In fact $|a_2|\leqq 2$ (Bieberbach), $|a_3|\leqq$ (Loewner), $|a_i|\leqq i$ for $i\leqq 6$ (*see Duren* for the references).

Bieberbach-Koebe Theorem. — See Hille. The image of a schlicht function contains a disc of radius 1/4 around 0.

Theorem of Littlewood. — See Hille.

$$|a_n|\leqq ne\qquad (e=2.71\ldots).$$

Improved Theorem of Littlewood. — See Duren.

$$|a_n|\leqq 1.07\,n.$$

Loewner Theorem. — See Hayman. If $g(w)=w+b_2w^2+b_3w^3+\ldots$ is the inverse of a schlicht function expressed as a power series near zero, then

$$|b_i|\leqq 2^i\frac{1.3\ldots(2i-1)}{1.2\ldots(i+1)}\leqq 4^{i-1}.$$

Finally we state:

Distortion Theorem of Koebe and Gronwall. — See Hayman, Hille. If f is schlicht then

$$\frac{r}{(1+r)^2} \leqq |f(z)| \leqq \frac{r}{(1-r)^2}, \qquad \text{for} \quad |z| \leqq r,$$

$$\frac{1-r}{(1+r)^3} \leqq |f'(z)| \leqq \frac{1+r}{(1-r)^3}, \qquad |z| \leqq r,$$

these are sharp bounds.

We use B to indicate the best number between 1 and 1.07 (Improved Littlewood Theorem) such that $|a_i| \leqq i\mathrm{B}$ for all i. Thus $\mathrm{B}=1$ if the Bieberbach conjecture is true. Recall that τ_k is truncation. Our main tool in proving Theorem 1 is the following Proposition.

PROPOSITION 1. — *Suppose that f is schlicht and g is its inverse. Then*

$$|g(f(z))-g(\tau_k f(z))| \leqq \frac{\gamma^{k+1}(1-\gamma)^2\,\mathrm{B}(k+1)}{[(1-\gamma)^2-4\gamma][(1-\gamma)^2-4\gamma(1+\mathrm{B}(k+1)\gamma^k)]}$$

where $\gamma=|z|$, for γ less than the first positive root of the denominator.

We use a sequence of lemmas.

LEMMA 1. — *For $|x|<1$, $l_0 \geqq 1$*

$$\sum_{l=l_0}^{\infty} l x^{l-l_0} \leqq \frac{l_0 x^{l_0-1}}{(1-x)^2}.$$

Proof:

$$\sum_{l=l_0}^{\infty} l x^{l-1} = \sum_{l=1}^{\infty} l x^{l-1} - \sum_{l=0}^{l_0-1} l x^{l-1} = \left(\frac{1}{1-x}\right)' - \left(\frac{1-x^{l_0}}{1-x}\right)' = \frac{x^{l_0-1}(l_0-(l_0-1)x)}{(1-x)^2} \leqq \frac{l_0 x^{l_0-1}}{(1-x)^2}.$$

Q.E.D.

LEMMA 2. — *Let $g(x)=\sum b_i z^i$ be a convergent power series with $b_0=0$ and $|b_1|=1$. Let $a=\max_{l>1}|b_l|^{(1/l-1)}$. Then*

$$\left|\frac{g^{(l)}}{l!}(x)\right| \leqq a^{l-1}\left(\frac{1}{1-a|x|}\right)^{l+1} \qquad \textit{for} \quad |x|<\frac{1}{a}.$$

Proof. — Note

$$g^{(l)}(x)=\sum_{j=l}^{\infty} j(j-1)\ldots(j-l+1)\,b_j x^{j-l}$$

and

$$|g^{(l)}(x)| \leqq \sum_{j=l}^{\infty} j(j-1)\ldots(j-l+1)\,a^{j-1}\,x^{j-l}.$$

So if $y=a|x|<1$, then

$$\begin{aligned}|g^{(l)}(x)| &\leqq a^{l-1}\sum_{j=l}^{\infty} j(j-1)\ldots(j-l+1)\,y^{j-l}\\ &\leqq a^{l-1}\frac{d^l}{dy^l}\sum_{j=0} y^j = a^{l-1}\frac{d^l}{dy^l}\left(\frac{1}{1-y}\right)\\ &\leqq a^{l-1}\frac{l!}{(1-y)^{l+1}}.\end{aligned}$$

Divide by l!

Q.E.D.

LEMMA 3. — *Let $g(w)=\sum b_i w^i$ be a convergent power series with $b_0=0$, $|b_1|=1$ and let $a=\max_{l>1}|b_l|^{1/(l-1)}$. Let x, $w\in\mathbb{C}$, and b, c be positive numbers. Suppose $(1+c)ab<1$, $|x|\leqq b$, $|w-x|\leqq bc$.*
Then

$$|g(w)-g(x)| \leqq \frac{bc}{(1-ab)(1-(1+c)ab)}.$$

Proof. — Consider the Taylor expansion of g at x

$$g(w)-g(x)=\sum_{l=1}^{\infty}\frac{g^{(l)}(x)}{l!}(w-x)^l.$$

Apply Lemma 2 to obtain

$$\begin{aligned}|g(w)-g(x)| &\leqq \sum_{l=1}^{\infty} a^{l-1}\left(\frac{1}{1-a|x|}\right)^{l+1}|w-x|^l\\ &\leqq \sum_{l=1}^{\infty} a^{l-1}\left(\frac{1}{1-a|x|}\right)^{l+1}(bc)^l=\left(\frac{1}{1-a|x|}\right)^2(bc)\sum_{l=1}^{\infty}\left(\frac{abc}{1-a|x|}\right)^{l-1}\\ &=bc\left(\frac{1}{1-a|x|}\right)^2\frac{1}{(1-(abc/(1-a|x|)))}=\left(\frac{bc}{1-a|x|}\right)\left(\frac{1}{1-a|x|-abc}\right).\end{aligned}$$

Q.E.D.

Now we can give the proof of Proposition 1. Let $f(z)=\sum_{i=1}^{\infty} a_i z^i$. By the Koebe-Gronwall Distortion Theorem,

$$|f(z)| \leqq \frac{\gamma}{(1-\gamma)^2} \qquad \text{where} \quad \gamma=|z|.$$

Then

$$|f(z)-\tau_k f(z)| \leqq \sum_{j=k+1}^{\infty} |a_j| \gamma^j \leqq B \sum_{j=k+1}^{\infty} j \gamma^j,$$

by the extended Littlewood Theorem where $1 \leqq B \leqq 1.07$. In turn this is less than $(B(k+1)\gamma^{k+1})/(1-\gamma)^2$ by Lemma 1.

Now we apply Lemma 3, where we may take $a=4$ by the Loewner estimate. Let $c=B\gamma^k(k+1)$ and $b=\gamma/(1-\gamma)^2$ so that our above argument shows that $|f(z)-\tau_k f(z)| \leqq bc$. Moreover $(1+c)ab<1$. Therefore Lemma 3 applies and yields the proposition.

We generalize the proposition slightly to the case of $|z|<r$ rather than $|z|<1$.

COROLLARY. — *Let $f(z)=\sum a_i z^i$ be a one to one analytic function for $|z|<r$ with $a_0=0$, $a_1=1$ and let g be its inverse. Then*

$$|g(f(z))-g(\tau_k f(z))| \leqq \frac{r\gamma^{k+1}(1-\gamma)^2 B(k+1)}{[(1-\gamma)^2-4\gamma][(1-\gamma)^2-4\gamma(1+B(k+1)\gamma^k)]},$$

for $\gamma=|z|/r$ less than the first positive root of the denominator.

Proof. — The map $z \to 1/r\, f(rz)$ is schlicht with inverse $w \to (1/r)g(rw)$. Apply the Proposition to estimate

$$\left|\frac{1}{r}g(f(rz))-\frac{1}{r}g(\tau_k f(rz))\right|.$$

Q.E.D.

Now we use the polynomial σ associated to f and z from section 1 together with the last Corollary to obtain Theorem 1. From Proposition 2 of section 1 and the discussion after it, one has

$$\frac{f(z')}{f(z)}=1-\sigma\circ R, \qquad R=\tau_k\sigma^{-1}(h).$$

Thus

$$\frac{f(z')}{f(z)}=1-\sigma\tau_k\sigma^{-1}(h)=1-h-(\sigma\tau_k(\sigma^{-1}h)-\sigma(\sigma^{-1}h)).$$

Since σ^{-1} is an analytic function, one to one on a disc of radius $h_1=h_1(f, z)$ about 0, we can apply the preceding Corollary to obtain Theorem 1.

Remark 1. — If the Bieberbach conjecture is true, $B=1$, *see* above and *Duren*. In any case, using results of Bieberbach, Loewner, Shiffer and others B can be replaced by something less than 1.07 and the 4's can be replaced by a smaller number.

Remark 2. — The case of Theorem 1 for $k=1$, incremental Newton's method, was also studied in Smale. Unfortunately the theorem here doesn't imply that one.

The main applications of Theorem 1 are in sections 3 and 4. In the meantime we give a Corollary which for the classical case of $h=1$, gives new convergence criteria.

LEMMA 4. — $\beta'_k(\gamma)>0$ *for* $0<\gamma<\gamma_k$.

Proof. — Use the quotient rule to differentiate $\beta_k(\gamma)$. Write the numerator as

$$2\,B(k+1)(1-\gamma)[(2+2\gamma)((1-\gamma)^2-4\gamma(1+B(k+1)\gamma^k)+(1-\gamma)((1-\gamma)^2-4\gamma)[(1-\gamma) \\ +2(1+(k+1)B\gamma^k)+2B\gamma k(k+1)\gamma^{k-1})]$$

and all the terms are positive since $0<\gamma<\gamma_k$.

Note that γ_k increases with k and tends to the first root of $(1-\gamma)^2-4\gamma=0$ which is $3-\sqrt{8}=.171572$.

Here we tabulate some values of γ_k calculated to six decimal places by Gerry Roskes, who also did the other calculations in this paper with $B=1.07$.

	$k=1$	2	3	4	5	6	7	8	9	10
γ_k	.141454	.161948	.169095	.171019	.171457	.171549	.171568	.171571	.171572	.171572

We will see that Theorem 1 is interesting even for the case $h=1$.

In this case $z'=E_{k(1,\,f)}(z)$, and

$$\left|\frac{f(z')}{f(z)}\right| \leqq \beta_k(\gamma)\gamma^k \qquad \text{where} \quad \gamma=\frac{1}{h_1},$$

$\gamma<\bar{\gamma}_k$, where $\bar{\gamma}_k$ is described as follows.

For $\gamma\in(0,\gamma_k)$ let $\alpha_k(\gamma)=\beta_k(\gamma)\gamma^k$. Then from Lemma 4 it follows that $\alpha_k(\gamma)$ is an increasing function of γ. Let $\bar{\gamma}_k$ be the unique solution of $\alpha_k(\gamma)=1$. Clearly $\bar{\gamma}_k$ increases with k and tends to $3-\sqrt{8}$. One calculates to six decimals.

$k=$	1	2	3	4	5
$\bar{\gamma}_k$	.093870	.132654	.152367	.162345	.167258
$k=$	6	7	8	9	10
$\bar{\gamma}_k$	.169602	.170689	.171182	.171401	.171498

Define $\rho_f=\min\limits_{\substack{\theta \\ f'(\theta)=0}} |f(\theta)|$.

COROLLARY 1. — *Let* $k>0$, *and write for brevity* $\bar{\gamma}=\bar{\gamma}_k$. *Suppose a polynomial* f *and a complex number* z *satisfy* $|f(z)|=b(\bar{\gamma}/(1+\bar{\gamma}))\rho_f$ *for some* $b<1$. *Then with* $h=1$, $(E_k)^l(z)=z_l$ *is defined for all* l *and* $z_l\to z^*$ as $l\to\infty$ *with* $f(z^*)=0$. *Moreover* $|f(z_l)|\leqq C|f(z_{l-1})|^{k+1}$ *all* $l>0$ ($k+1$ *order convergence*) *with* $C=(b/|f(z_0)|)^k$. *Finally* $|f(z_l)|\leqq b^{((k+1)^l)}(\bar{\gamma}/(1+\bar{\gamma}))\rho_f$, *all* $l>0$.

Proof. — Let $\gamma=|f(z)|/|f(z)-f(\theta_*)|$ where $\rho_f=|f(\theta_*)|$, $f'(\theta_*)=0$. Then $\gamma<\bar{\gamma}$ since

$$|f(z)-\mathrm{f}(\theta_*)|>\frac{1}{1+\bar{\gamma}}\rho_f$$

and

$$\frac{|f(z)|}{|f(z)-f(\theta_*)|}<\left(\frac{\gamma}{1+\bar{\gamma}}\rho_f\right)\left(\frac{1}{(1/(1+\bar{\gamma}))\,\rho_f}\right)=\bar{\gamma}.$$

We may take $h=1$ in Theorem 1. Thus

$$\left|\frac{f(z')}{f(z)}\right|<\beta_k(\gamma)\,\gamma^k<\beta_k(\bar{\gamma})\,\bar{\gamma}^k=1,$$

using Lemma 4 and the definition $\bar{\gamma}$ above. An obvious induction then yields the first part of Corollary 1.

Also

$$\left|\frac{f(z')}{f(z)}\right|\leqq\beta_k(\bar{\gamma})\frac{|f(z)|^k}{|f(z)-f(\theta_*)|}\leqq\mathrm{C}\,|f(z)|^k,$$

where

$$\mathrm{C}=\beta_k(\bar{\gamma})\left(\frac{1+\bar{\gamma}}{\rho_f}\right)^k=\left(\frac{b}{|f(z)|}\right)^k.$$

Next let $y_0=b\bar{\gamma}$ and define inductively $y_l=(y_{l-1})^{k+1}\,\beta_k(\bar{\gamma})$. Consider

$$(1_l)\qquad |f(z_l)|\leqq y_l\frac{\rho_f}{1+\bar{\gamma}},\qquad l=0,\,1,\,2,\,\ldots$$

$$(2_l)\qquad \gamma_l\leqq y_l,\qquad l=0,\,1,\,2,\,\ldots$$

where

$$\gamma_l=\frac{|f(z_l)|}{\min\limits_{\substack{\theta\\ f'(\theta)=0}}|f(z_l)-f(y)|}.$$

Then (1_0) is true by definition. We proceed inductively showing (1_l) implies (2_l) and then (l_{l-1}) *and* (2_{l-1}) imply (1_l).

Thus

$$\gamma_l\leqq\left(\gamma_l\frac{\rho_f}{(1+\bar{\gamma})}\right)\frac{1}{\min\limits_{\theta}|f(z_l)-f(\theta)|}\leqq y_l\frac{f}{(1+\bar{\gamma})}\left(\frac{1+\bar{\gamma}}{\rho_f}\right)\leqq y_l.$$

Finally

$$|f(z_l)| \leqq \beta_k(\bar{\gamma})\gamma_{l-1}^k |f(z_{l-1})| \leqq \beta_k(\bar{\gamma})\gamma_{l-1}^k \left(\frac{\rho_f}{1+\bar{\gamma}}\right) \leqq y_l \frac{\rho_f}{(1+\bar{\gamma})}.$$

This proves (1_l) and (2_l) all l.

One now checks inductively that $y_l = b^{(k+1)^l}\bar{\gamma}$ and thus the statement of the Corollary.

Q.E.D.

Note the very rapid decrease of $|f(z_l)|$ a function of l. This justifies the following definition.

We will call z an *approximate zero* for f *relative to* k if and only if $|f(z)| < [\bar{\gamma}_k/(1+\bar{\gamma}_k)]\,\rho_f$. An approximate zero for f relative to all $k>0$ will be called simply an *approximate zero* so z is an approximate zero of f if and only if $|f(z)| < [\bar{\gamma}_1/(1+\bar{\gamma}_1)]\,\rho_f$. Note by the computations below, if $|f(z)| < (1/12)\,\rho_f$ then z is an an approximate zero.

We tabulate some values of $\bar{\gamma}_k/(1+\bar{\gamma}_k)$. Note that $\bar{\gamma}_k/(1+\bar{\gamma}_k)$ is increasing and tends to $(3-\sqrt{8})/(1+3-\sqrt{8}) = .146446$ to six decimals.

k	1	2	3	4	5
$\dfrac{\bar{\gamma}_k}{1+\bar{\gamma}_k}$	.085815	.117117	.132221	.139670	.143291
k	6	7	8	9	10
$\dfrac{\bar{\gamma}_k}{1+\bar{\gamma}_k}$	.145008	.145802	.146161	.146321	.146392

So in particular $1/6 > \gamma_k/(1+\bar{\gamma}_k) > 1/12$ and for $k \geqq 5$, $1/6 > \bar{\gamma}_k/(1+\bar{\gamma}_k) > 1/7$.

We have applied Proposition 1 and its Corollary to σ and σ^{-1} to conclude Theorem 1. We could as well apply this reasoning to a general analytic function g defined on a domain Ω and its inverse g^{-1}. *The resulting general statement is:*

THEOREM 1 *a.* — *Let* $z \in \Omega \subset \mathbb{C}$ *and let* $g : \Omega \to \mathbb{C}$ *be analytic. Suppose* g_z^{-1} *is defined on a disc* D *of radius* $R(g, z)$. *There are constants* c_k *and* K_k depending on k ($c_k \approx .1$ *and* $K_k \approx k$) *such that if* $|w - g(z)| < c_k R(g, z)$ *and* $g g_z^{-1}(w) = w$ *then*

$$|g((t_k g_z^{-1})(w)) - w| \leqq \frac{K_k |w - g(z)|^{k+1}}{R(g, z)^k}.$$

If we suppose that $\rho_f > 0$, then for any root ξ of f, $f'(\xi) \neq 0$ and f_ξ^{-1} may be uniquely analytically continued along any ray starting from 0 as long as the inverse image of this ray doesn't run into a critical ppint of f. Thus f_ξ^{-1} may be analytically continued to the entire complex plane minus k radial slits from $f(\theta_{i_1}), \ldots, f(\theta_{i_k})$ to ∞ for some minimal collection $\theta_{i_1}, \ldots, \theta_{i_k}$ of critical points of f. We shall denote this domain by $S_{\xi, f}$. For $d > 1$, $1 \leqq k \leqq d-1$. We use f_ξ^{-1}: $S_{\xi, f} \to \mathbb{C}$ to denote the analytic continuation. It is quite clear that the images of the $f_{\xi_i}^{-1}$ are disjoint over the roots ξ_i of f since the inverse image of a ray by $f_{\xi_i}^{-1}$ is a solution curve of the Newton differential equation $dz/dt=-f(z)/f'(z)$ in $\mathrm{R}^2=\mathbb{C}$ which terminates at ξ_i. Consider now a general $z\in\mathbb{C}$. If f_z^{-1} can be analytically continued along the ray from $f(z)$ to 0 then $f_z^{-1}=f_\xi^{-1}$ in a neighborhood of this ray for some root ξ of f. We analytically continue $f_z^{-1}: \mathrm{S}_{z,f}\to\mathbb{C}$ by setting it equal to f_ξ^{-1} and $\mathrm{S}_{z,f}=\mathrm{S}_{\xi,f}$.

We let $\rho_{f,z}$ be the radius of convergence of $f_z^{-1}: \mathrm{S}_{z,f}\to\mathbb{C}$ around 0. Then $\rho_f=\min \rho_{f,z}$. We note that if $z'\in$ Image $f_z^{-1}: \mathrm{S}_{z,f}\to\mathbb{C}$ then $f_{z'}^{-1}=f_z^{-1}$ and $\mathrm{S}_{z',f}=\mathrm{S}_{z,f}$.

The discussion above allows an improvement of the Corollary to Theorem 1, with the help of Proposition 4 below.

COROLLARY 2. — *Corollary* 1 *to Theorem* 1 *remains true with* ρ_f *replaced by* $\rho_{f,z}$ *and* $z_*=f_z^{-1}(0)$ *where* $f_z^{-1}: \mathrm{S}_{z,f}\to\mathbb{C}$.

Thus the notion of approximate zero may also be extended to z satisfying $|f(z)|<\bar{\gamma}(1+\bar{\gamma})\rho_{f,z}$.

PROPOSITION 2. — *Let* ξ_i, $i=1, \ldots, k$ *be distinct simple roots of* $f(z)=z^d+a_{d-1}z^{d-1}+\ldots+a_0$. *Then the discs of radius*

$$\frac{1}{4}\frac{\bar{\gamma}}{1+\bar{\gamma}}\frac{\rho_{f,\xi_i}}{|f'(\xi_i)|},$$

centered at ξ_i *are disjoint and consist of approximate zeros.*

For the proof of the Proposition we require a slight extension of the Bieberbach-Koebe Theorem.

LEMMA 5. — *Let f be a one to one analytic function defined on a disc of radius r, then the image of f contains a disc of radius $|f'(0)|r/4$ around $f(0)$.*

Proof. — $g(w)=(1/f'(0)r)f(rw)-f(0)$, $|w|<1$, is schlicht. So image g contains a disc of radius 1/4 around 0 and it follows that the image of f contains a disc of radius $|f'(0)|r/4$ around $f(0)$.

Proof of Proposition 2. — By Lemma 5 the image by $f_{\xi_i}^{-1}$ of the disk of radius $[\bar{\gamma}/(1+\bar{\gamma})]\rho_{f,\xi_i}$ centered at o contains a disc of radius

$$\frac{1}{4}\frac{\bar{\gamma}}{1+\bar{\gamma}}\rho_{f,\xi_i}\cdot|(f_{\xi_i}^{-1})'(0)|=\frac{1}{4}\frac{\bar{\gamma}}{1+\bar{\gamma}}\frac{\rho_{f,\xi_i}}{|f'(\xi_i)|}.$$

And these discs consist of approximate zeros.

It has already been noted that the images of the $f_{\xi_i}^{-1}: \mathrm{S}_{\xi_i,f}\to\mathbb{C}$ are disjoint for distinct ξ_i.

PROBLEM. — Let $\mathrm{P}_d(1)=\{f\,|\,f(z)=z^d+a_{d-1}z^{d-1}+\ldots+a_0 \text{ with } |a_i|\leqq 1\}$. What is the distribution of

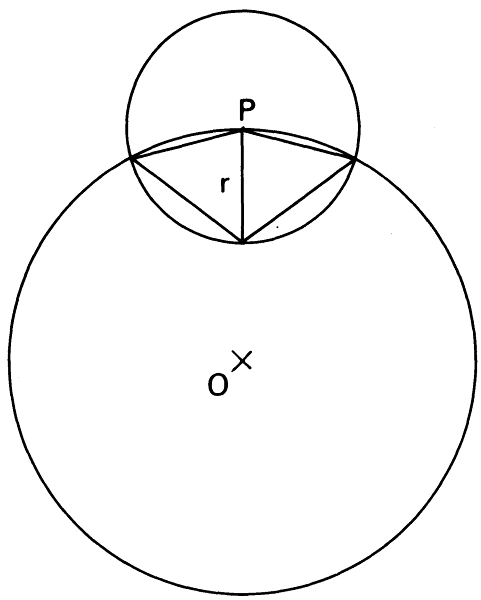
$$\sum_{\substack{\xi\\ f(\xi)=0}}\left(\frac{\rho_{f,\xi}}{|f'(\xi)|}\right)^2 \quad \text{in } \mathrm{P}_d(1)?$$

We now give a simple estimate of the area of the set of approximate zeros of $f \in P_d(1)$ which are contained in the unit disc.

LEMMA 6. — *Suppose that p is a point in the unit disc on* $\mathbf{R}^2$ *and* $0<r<1$. *Then the area of the intersection of the disk of radius r around p and the unit-disk is greater than*

$$\frac{r^2\sqrt{4-r^2}}{2}.$$

Proof. — The worst case occurs for p on the boundary of the unit disk. Thus we may assume that $p=(0, 1)$.



The x-coordinates of the points of intersections are $\pm(r\sqrt{4-r^2})/2$. Thus the area of the two triangles contained in the intersections is

$$2\left(\frac{r}{2}\cdot\frac{r\sqrt{4-r^2}}{2}\right)=\frac{r^2\sqrt{4-r^2}}{2}.$$

Q.E.D.

PROPOSITION 3. — *Suppose that* $f \in P_d(1)$ *and* $\rho_f>0$. *Then the area of the set of points z in the unit disc such that* $f(z)<(\bar{\gamma}/(1+\bar{\gamma}))\,\rho_f$ *is at least*

$$.003\left(\frac{\rho_f}{d(d+1)}\right)^2.$$

Proof. — There is a root ξ of f in the closed unit disk since the product of the root is $\leqq 1$.

$$|f'(\xi)| \leqq d+(d-1)+\ldots+1=\frac{d(d+1)}{2}.$$

Thus

$$\left|\frac{1}{f'(\xi)}\right| \geqq \frac{2}{d(d+1)}.$$

Now apply Lemma 5 to produce a disc centered at ξ, consisting of approximate zeros of f and of radius

$$\frac{1}{2}\frac{\bar{\gamma}}{1+\bar{\gamma}}\rho_f\frac{1}{d(d+1)}.$$

Use Lemma 6 to estimate the area, with the help of a hand held calculator. Since there is a critical point θ of f in the unit disc $\rho_f \leqq d+1$ which simplifies the expression in the radical.

It is convenient here to prove a Proposition which will be used in the proof of Theorem 2 and which can be used in the proof of Corollary 2. First we need a slight alteration of some estimates we have already used.

LEMMA 7. — *Let $f(z)=z+a_2 z^2+\ldots$ be a $1-1$ analytic function defined in the disc of radius h_*. For $|z|<h^*$, let $\gamma=|z|/h_*$. Then*

(*a*) $$|f(z)| \leqq \frac{|z|}{(1-\gamma)^2},$$

(*b*) $$|f(z)-\tau_k f(z)| \leqq \frac{h_* \mathrm{B}(k+1)\gamma^{k+1}}{(1-\gamma)^2}.$$

Let σ be the polynomial defined in section I and let $\mathrm{D}(h_*)$ be the disc of radius h_* around 0.

LEMMA 8. — *Suppose that $0<h_* \leqq h_1(f, z)$ and that $h=\gamma h_*$ for some $0<\gamma<\gamma_k$. Then $\tau_k \sigma^{-1}(h) \in \sigma^{-1}(\mathrm{D}(h_*))$.*

Proof. — By Lemma 7 applied to σ^{-1} on the disc of radius h_*

$$|\sigma^{-1}(h)| \leqq \frac{h}{(1-\gamma)^2} \quad \text{and} \quad |\sigma^{-1}(h)-\tau_k \sigma^{-1}(h)| \leqq \frac{h_* \mathrm{B}(k+1)\gamma^{k+1}}{(1-\gamma)^2}.$$

Thus

$$|\tau_k \sigma^{-1}(h)| \leqq \frac{h}{(1-\gamma)^2}+\frac{h_* \mathrm{B}(k+1)\gamma^{k+1}}{(1-\gamma)^2}.$$

By Lemma 5, $\sigma^{-1}(\mathrm{D}(h_*))$ contains a disc of radius $h_*/4$ around 0. Thus $\tau_k \sigma^{-1}(h) \in \sigma^{-1}(\mathrm{D}(h_*))$ as long as

$$\frac{h}{(1-\gamma)^2} + \frac{h_* \mathrm{B}(k+1)\gamma^{k+1}}{(1-\gamma)^2} < \frac{h_*}{4},$$

or $(1-\gamma)^2 > 4\gamma(1+\mathrm{B}(k+1)\gamma^{k+1})$. The first point of equality is the definition of γ_k.

PROPOSITION 4. — *Suppose that* $0 < h_* \leqq h_1(f, z)$ *and that* $h = \gamma h_*$ *for some* $0 < \gamma < \gamma_k$. *Let* $z' = \mathrm{E}_{k(h, f)}(z)$. *Then there is a complex number* h' *with* $|h'| < h_*$ *such that*

$$z' = f_z^{-1}((1-h') f(z)),$$

where $f_z^{-1} : \mathrm{D}(\mathrm{h}_*) \to \mathbb{C}$.

Proof. — We use the definition of E_k and its relation to σ^{-1} as discussed in section 1. $z' = z - \mathrm{F}\,\tau_k(\sigma^{-1}(h))$. By Lemma 8, $\tau_k(\sigma^{-1}(h)) = \sigma^{-1}(h')$ with $|h'| < h_*$. Thus

$$z' = z + \mathrm{F}\,\sigma^{-1}(h') = f_z^{-1}((1-h') f(z)),$$

by the discussion after the statement of Proposition 2 of section 1.

Section 3

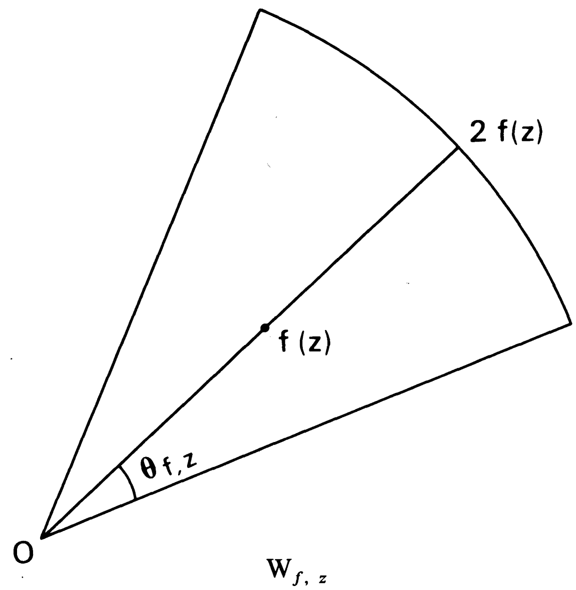
Let f be a polynomial z a complex number. Then recall from section 1 that f_z^{-1} is a function taking $f(z)$ to z, given by a power series on a disk of radius $r(f, z)$ about $f(z)$. Let us call that disk $\mathrm{D}_{f,\,z}$, so $f_z^{-1} : \mathrm{D}_{f,\,z} \to \mathbb{C}$ is analytic.

On the other hand one can consider other domains for f_z^{-1}. In particular define $\mathrm{W}_{f,\,z}$ to be such a domain which is a wedge shaped circular sector as follows. For $0 < \alpha \leqq \pi/2$ let

$$\mathrm{W}_{f,\,z,\,\alpha} = \left\{ w \in \mathbb{C} \;\middle|\; 0 < |w| < 2|f(z)|,\ \left|\arg \frac{w}{f(z)}\right| < \alpha \right\}.$$

Then define $\mathrm{W}_{f,\,z}$ to be the largest of the $\mathrm{W}_{f,\,z,\,\alpha}$ on which f_z^{-1} is analytic and let $\theta_{f,\,z}$ be the corresponding α. Note that if $\theta_{f,\,z} < \pi/2$ then a critical value $f(\theta)$ lies on a side of $\mathrm{W}_{f,\,z}$. It is clear that $\mathrm{W}_{f,\,z} \subset \mathrm{S}_{z,\,f}$ which was defined in section 2, and that $\mathrm{W}_{f,\,z}$ is the largest $\mathrm{W}_{f,\,z,\,\alpha}$ contained in $\mathrm{S}_{z,\,f}$.

Our algorithms attempt to make this analytic continuation a computationally feasible process. What we do in this section is to show that the Euler algorithms yield a sequence of iterates z_n whose values $f(z_n)$ remain in an appropriate wedge-shaped region $\mathrm{W}_{f,\,z}$.



$W_{f,\,z}$

Let

$$K(k)=\frac{(k+1)^{(k+1)/k}}{K\,\bar{\gamma}_k(1-\bar{\gamma}_k)^{1/k}},$$

where $\bar{\gamma}_k$ is the constant given in section 2.

THEOREM 2. — *Suppose given a polynomial f and complex number z_0 and a real number $L>0$ such that $|f(z_0)|>L$ and $\Theta=\theta_{f,\,z_0}>0$. Let $c=1/\Theta\log|f(z_0)|/L$. Then there is*

$$h_0 \geqq \frac{\sin\Theta}{K(k)(c+1)^{1/k}}$$

with this property: For each h, $0<h\leqq h_0$, there is some

$$n\leqq\frac{1}{h}\left[\log\frac{|f(z_0)|}{L}+\frac{\Theta}{k+1}\right]$$

such that $|f(z_n)|<L$ where $z_n=E^n_{k(h,\,f)}(z_0)$.

Remark. — We actually prove a stronger statement. If $a=k/(k+1)\sin\Theta$, one can take

$$n=\left\lceil\frac{1}{h}\,\frac{\log|f(z_0)|}{L}\left(\frac{1}{1-\alpha_k(h/a)}\right)\right\rceil$$

in the theorem where $\lceil\gamma\rceil$ denotes the smallest integer greater than or equal to x. See Lemma 1 below. We also are using the functions $\alpha_k(\gamma)$ introduced in section 2, which are defined and increasing in the range $0<\gamma<\bar{\gamma}_k$.

Note that n, the number of steps depends crucially on the constant (function of k only) $K(k)$.

Here we tabulate some values of $K(k)$ to six decimal places.
$K(k)$ decreases to $1/(3-\sqrt{8})=5.82842$.

k	1	2	3	4	5
$K(k)$	47.0262	21.0297	14.6779	12.0350	10.6492

k	6	7	8	9	10
$K(k)$	9.81331	9.25572	8.85456	8.54935	8.30772

Note that for $k=1$ or Newton's method the value of $K(1)$ shows that Theorem 2 in that case is significantly worse than the results achieved in Smale.

COROLLARY TO THEOREM 2. — *In addition to the hypotheses of Theorem 2, suppose that* $L<(\bar{\gamma}_k/(1+\bar{\gamma}_k))\,\rho_f$. *Then for all* $0<h\leqq h_0$ *in Theorem 2 and all*

$$n\geqq\frac{1}{h}\left(\log\frac{|f(z_0)|}{L}+\frac{\Theta}{k+1}\right),\qquad |f(z_n)|<L.$$

This is simply a consequence of Theorem 2 and the proof of Corollary 1, of Theorem 1.

The proof of Theorem 2 uses heavily Theorem 1 and partially generalizes Theorem 3 of Smale (the case $k=1$, or Newton's method).

LEMMA 1. — *For any* $c>0$, $0<a<1$, *and* $\alpha_k(\gamma)$ *as above, there is a unique solution*

$$h=h_0 \quad\text{of}\quad (k+1)c+1=\frac{1-h}{\alpha_k(h/a)}$$

in the range $0<h_0<a\bar{\gamma}_k$.
This h_0 *satisfies*

$$h_0\geqq\frac{a(k+1)}{k\,K(k)(c+1)^{1/k}},$$

Furthermore for

$$0<h\leqq h_0,\qquad \frac{1}{1-\alpha_k(h/a)}\leqq 1+\frac{1}{c(k+1)}.$$

Proof. — The right hand of the equation for h_0 decreases monotonically from ∞ to below 1 as h goes from 0 to $a\bar{\gamma}_k$ (Lemma 4 of section 2). Since $(k+1)c+1>1$, this yields the unique solution h_0.

Since $\alpha_k(h/a)\leqq\alpha_k(h_0/a)$ for $h<h_0$, using the defining equation for h_0 we have

$$\frac{1}{1-\alpha_k(h/a)}\leqq\frac{1}{1-\alpha_k(h_0/a)}=\frac{(k+1)c+1}{(k+1)c+h_0}\leqq 1+\frac{1}{c(k+1)}.$$

This gives the last statement of Lemma 1. It remains to show

$$h_0 \geqq \frac{a(k+1)}{k\,\mathrm{K}(k)(c+1)^{1/k}}.$$

Note $h_0 \leqq a\bar{\gamma}_k < \bar{\gamma}_k$ since $a<1$. Then use the defining equation for h_0 to obtain

$$\alpha_k\left(\frac{h_0}{a}\right)=\frac{1-h_0}{k(c+1)+1} \geqq \frac{1-\bar{\gamma}_k}{k(c+1)+1}.$$

From the definition $\alpha_k(\gamma)=\beta_k(\gamma)\gamma^k$ of section 2, we have

$$\left(\frac{h_0}{a}\right)^k \beta_k\left(\frac{h_0}{a}\right)=\alpha_k\left(\frac{h_0}{a}\right) \geqq \frac{1-\bar{\gamma}_k}{k(c+1)+1},$$

so

$$\left(\frac{h_0}{a}\right)^k \geqq \frac{1-\bar{\gamma}}{(k(c+1)+1)}\,\frac{1}{\beta_k(h_0/a)} \geqq \frac{1-\bar{\gamma}_k}{(k(c+1)+1)}\,\frac{1}{\beta_k(\bar{\gamma}_k)} = \frac{1-\bar{\gamma}_k}{(k(c+1)+1)}\bar{\gamma}_k^k.$$

By taking k-th roots,

$$\frac{h_0}{a} \geqq \frac{\bar{\gamma}_k(1-\bar{\gamma}_k)^{1/k}}{(k(c+1)+1)^{1/k}} \geqq \frac{\bar{\gamma}_k(1-\bar{\gamma}_k)^{1/k}}{(k+1)^{1/k}(c+1)^{1/k}} \geqq \frac{1}{\mathrm{K}(k)}\,\frac{\mathrm{k}+1}{\mathrm{k}}\,\frac{1}{(c+1)^{1/k}}$$

and thus our desired inequality.

Q.E.D.

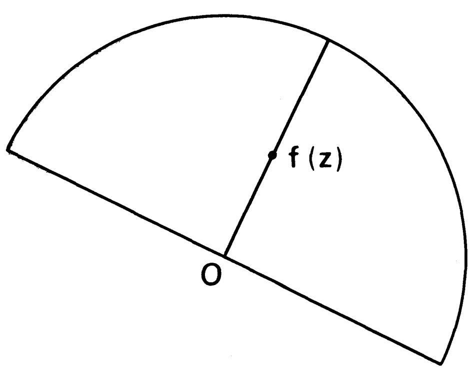
The last part of Lemma 1 explains the remark after Proposition 2 since by it,

$$\frac{1}{h}\log\frac{|f(z_0)|}{\mathrm{L}}\left(\frac{1}{1-\alpha_k(h/a)}\right) \leqq \frac{1}{h}\left(\log\frac{|f(z_0)|}{\mathrm{L}}+\frac{\Theta}{k+1}\right).$$

In fact we proceed to prove Theorem 2 in the form of the n in this remark.

Lemma 2. — $h_{1\,(f,\,z)} \geqq \sin\theta_{f,\,z}$.

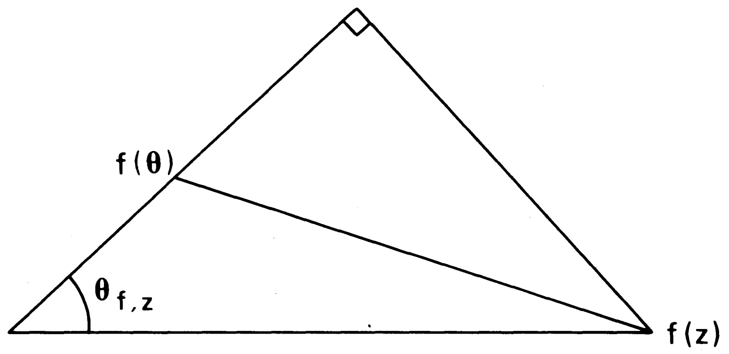
Proof. — If $\theta_{f,\,z}=\pi/2$ then the wedge is a semi-circle of



radius $2|f(z)|$ and the circle of radius $|f(z)|$ about $f(z)$ is contained in the wedge. Thus

$$h_{1(f,\,z)} \geqq \frac{|f(z)|}{|f(z)|} = 1 = \sin\frac{\pi}{2}.$$

If $\theta_{f,\,z} < \pi/2$, fix a critical point θ which maps to the boundary of the wedge.



Then $h_1(f, z) = |f(z) - f(\theta)| / |f(z)| \geqq \sin\theta_{f,\,z_0}$.

Q.E.D.

From trigonometry, one has:

LEMMA 3. — *For* $0 \leqq x < \pi/2$, $0 \leqq \alpha \leqq 1$,

(*a*) $|\arctan(x)| \leqq x$ *and*

(*b*) $\sin\alpha x \geqq \alpha\sin x$.

The next step is to use Theorem 1, section 2. We write

$$\gamma = \frac{h}{h_1}; \qquad 0 < \gamma \leqq \bar{\gamma}_k, \qquad \beta_k(\gamma)\,\gamma^k = \alpha_k(\gamma) \leqq \alpha_k(\bar{\gamma}_k) = 1.$$

Then Theorem 1 asserts

$$\begin{aligned}\frac{f(z')}{f(z)} &= 1 - h + Q(h, f, z)\,\gamma^k h\\ &= 1 - (1 - Q(h, f, z)\,\gamma^k)\,h,\end{aligned}$$

where $|Q(h, f, z)| \leqq \beta_k(\gamma)$.

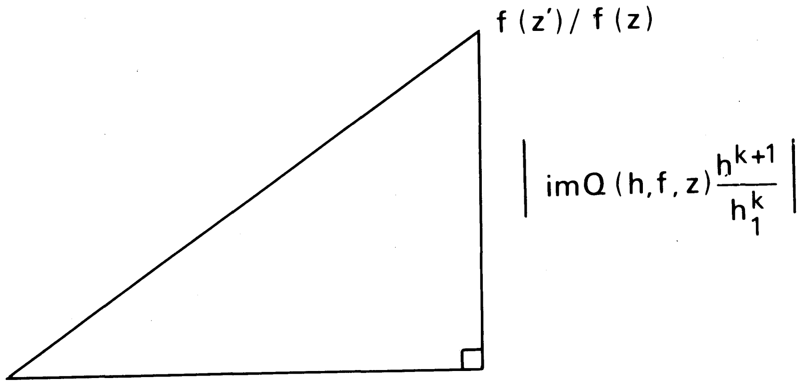
Already this yields the first part of the following lemma.

LEMMA 4. — *Let* $z' = E_{k(h,\,f)}(z)$ *where* E_k *is the* k-th *order Euler incremental algorithm. Then with the above notation for* $0 < \gamma < \bar{\gamma}_k$

(*a*) $$\left|\frac{f(z')}{f(z)}\right| \leqq 1 - (1 - \alpha_k(\gamma))\,h,$$

(*b*) $$\left|\arg\frac{f(z')}{f(z)}\right| < \frac{\alpha_k(\gamma)\,h}{1 - (1 + \alpha_k(\gamma))\,h}.$$

The proof of (*b*) goes by the consequence of Theorem 1 above and is aided by the diagram.



$$\left|1-h+\operatorname{Re} Q(h, f, z)\frac{h^{k+1}}{h_1^k}\right|,$$

$$\left|\arg\frac{f(z')}{f(z)}\right| = \arc\tan\frac{|\operatorname{im} Q(h, f, z)(h^{k+1})/h_1^k|}{1-h+\operatorname{Re} Q(h, f, z)(h^{k+1})/h_1^k}$$
$$< \frac{\beta_k(\gamma)\gamma^k h}{1-h-\beta_k(\gamma)\gamma^k h} = \frac{\alpha_k(\gamma) h}{1-h-\alpha_k(\gamma) h}.$$

We have used Lemma 3; this proves Lemma 4.

Let

$$a = \frac{k}{k+1}\sin\theta_{f,\, z_0} \qquad \text{and} \qquad \delta = \frac{h}{a}.$$

LEMMA 5. — *Let* $0<\delta<\bar{\gamma}_k$ *and* $z_n = \mathrm{E}^n_{k(h,\, f)}(z_0)$. *Then for*

$$0 \leqq n \leqq \frac{(1-(1+\alpha_k(\delta) h))}{(k+1)\alpha_k(\delta) h}\theta_{f,\, z_0},$$

we have

$$\theta_{f,\, z_n} \geqq \frac{k}{k+1}\theta_{f,\, z_0}, \qquad \textit{and} \qquad |f(z_{n+1})| \leqq (1-(1-\alpha_k(\delta)) h)^{n+1} |f(z_0)|.$$

The proof goes by induction on n, the case $n=0$ given by Lemma 4 *a*.

First note that if $\theta_{f,\, z_n} \geqq [k/(k+1)]\theta_{f,\, z_0}$ then by Lemmas 2 and 3 *b*,

$$h_1(f, z_n) \geqq \sin\frac{k}{k+1}\theta_{f,\, z_0} > \frac{k}{k+1}\sin\theta_{f,\, z_0} \quad \text{so } \gamma = \frac{h}{h_1(f, z_n)} < \delta$$

and

$$\left|\arg\frac{f(z_{n+1})}{f(z_n)}\right| < \frac{\alpha_k(\gamma) h}{1-(1+\alpha_k(\gamma)) h} < \frac{\alpha_k(\delta) h}{1-(1+\alpha_k(\delta)) h}$$

and

$$\left|\frac{f(z_{n+1})}{f(z_n)}\right| \leqq 1-(1-\alpha_k(\gamma))h < 1-(1-\alpha_k(\delta))h.$$

Using Lemma 4 applied to z_n instead of z. If $\theta_{f,\, z_n} \geqq k/(k+1)\,\theta_{f,\, z_0}$ then it is also true that

$$\theta_{f,\, z_{n+1}} \geqq \theta_{f,\, z_n} - \left|\arg\frac{f(z_{n+1})}{f(z_n)}\right|,$$

for: By Proposition 4 of section 2 with $h_* = k/(k+1)\sin\theta_{f,\, z_0}$ there is a complex number h' with $|h'| < h_*$ such that

$$z_{n+1} = f_{z_n}^{-1}((1-h')f(z))).$$

Thus $(1-h')f(z) \in W_{f,\, z_n}$ and $z_{n+1} \in \text{Image } f_{z_n}^{-1}$ where $f_{z_n}^{-1} : S_{z_n,\, f} \to \mathbb{C}$. It follows from the discussion of $S_{z,\, f}$ in section 2 and following the definition of $\theta_{f,\, z}$ in section 3, that $S_{z_n,\, f} = S_{z_{n+1},\, f}$. It is then immediate from the definition of $\theta_{f,\, z}$ that

$$\theta_{f,\, z_{n+1}} > \theta_{f,\, z_n} - \left|\arg\frac{f(z_{n+1})}{f(z_n)}\right|.$$

Thus we may proceed by induction as long as we are sure that

$$\theta_{f,\, z_n} \geqq \frac{k}{k+1}\theta_{f,\, z_0}$$

or as long as

$$\theta_{f,\, z_0} - \frac{n\,\alpha_k(\delta)\,h}{1-(1+\alpha_k(\delta))\,h} \geqq \frac{k}{k+1}\theta_{f,\, z_0},$$

that is: as long as

$$0 \leqq n \leqq \frac{1-(1+\alpha_k(\delta))\,h}{(k+1)\,\alpha_k(\delta)\,h}\theta_{f,\, z_0}.$$

With these lemmas done, we are now prepared to give the proof of Theorem 2, n as in the remark following that theorem.

Use Lemma 1 with

$$a = \frac{k}{k+1}\sin\Theta \qquad \text{and} \qquad c = \frac{1}{\Theta}\log\frac{|f(z_0)|}{L},$$

to obtain h_0 of Theorem 2. This h_0 will thus satisfy the inequality of Theorem 2. Now we have to show that for $0<h\leqq h_0$ and

$$n=\left[\frac{1}{h}\log\frac{|f(z_0)|}{L}\left(\frac{1}{1-\alpha_k(h/a)}\right)\right] \quad \text{then } |f(z_n)|<L.$$

We need to apply Lemma 5 [with $(n-1)$ replacing n] to obtain

$|f(z_n)|\leqq(1-(1-\alpha_k(\delta))h)^n|f(z_0)|\leqq L.$

For this it is required that

$$n\log(1-(1-\alpha_k(\delta))h)\leqq\log\left|\frac{L}{f(z_0)}\right|,$$

or

$$n\geqq\frac{\log|f(z_0)|/L}{|\log(1-(1-\alpha_k(\delta))h|}$$

and

(A) $$n\geqq\frac{\log|f(z_0)|/L}{(1-\alpha_k(\delta))h}$$

suffices.

On the other hand Lemma 5 demands an upper bound on $n-1$, namely

(B) $$\frac{(1-(1+\alpha_k(\delta))h)\theta_{f,z_0}}{(k+1)\alpha_k(\delta)h}\geqq n-1.$$

Thus if δ, h $(\delta=h/a)$ satisfy

(C) $$\frac{(1-(1+\alpha_k(\delta))h)\theta_{f,z_0}}{(k+1)\alpha_k(\delta)h}\geqq\frac{\log|f(z_0)|/L}{(1-\alpha_k(\delta))h},$$

then there is an integer n which satisfies both (A) and (B). We now find $\delta's$ which make (C) hold, or

$$\frac{(1-(1+\alpha_k(\delta))h)}{(k+1)\alpha_k(\delta)}\geqq\frac{\log|f(z_0)|/L}{\theta_{f,z_0}(1-\alpha_k(\delta))}=\frac{c}{1-\alpha_k(\delta)}.$$

Now multiplying by $1-\alpha_k(\delta)$ and $k+1$ it suffices to have

$$\frac{(1-\alpha_k(\delta))(1-(1+\alpha_k(\delta))h}{\alpha_k(\delta)}\geqq(k+1)c.$$

$$\frac{(1-h-\alpha_k h-\alpha_k(\delta)+\alpha_k h+\alpha_k(\delta)^2 h)}{\alpha_k(\delta)}\geqq(k+1)c.$$

So it suffices to have

$$\frac{1-h}{\alpha_k(\delta)}-1 \geqq (k+1)c \qquad \text{or} \qquad \frac{1-h}{\alpha_k(\delta)} \geqq (k+1)c+1.$$

But this last exactly corresponds to our defining equation (Lemma 1) for h_0, so that $h \leqq h_0$ guarantees the inequality is satisfied. This finishes the proof of Theorem 2.

Section 4

The Corollary of Theorem 2 shows that the crucial ingredients for producing an approximate zero for f, starting at z_0 and iterating $E_{k(h,\ f)}$, are

$$\theta_{f,\ z_0}, \ \log|f(z_0)| \qquad \text{and} \qquad |\log \rho_f|.$$

In this section we tend to increase $\theta_{f,\ z_0}$ at the expense of increasing $|f(z_0)|$. Doing so we gain considerably. In particular, we prove in this section:

THEOREM 3. — *There exist universal (small) positive constants* K_1, K_2 *with the following true. Given integers* $k>0$, *and* $d>1$, *and a real number* $1>\mu>0$: *There are* R, h *such that:*

$$\textit{If } (z_0, f) \in S_R^1 \times P_d(1) \quad \textit{then} \quad z_s = (E_{(k,\ h,\ f)})^s (z_0)$$

will be an approximate zero of f, *with probability* $1-\mu$ *for any*

$$s \geqq K_1 \left(d \frac{|\log \mu|}{\mu} \right)^{1+1/k} + K_2.$$

Here $S_R^1 = \{ z \in \mathbb{C} \,\|\, z| = R \}$, $P_d(1)$ is the space of all complex polynomials

$$f(z) = \sum_{i=}^{d} a_i z^i \quad \text{with} \quad a_d = 1 \qquad \text{and} \qquad |a_i| \leqq 1 \quad \text{for } i=1, 2, \ldots, d-1$$

and μ is normalized Lebesgue measure on the product $S_R^1 \times P_d(1)$.

COROLLARY. — *Given* $d>1$, $1>\mu>0$ *there are* R, h *such that:*

If $(z_0, f) \in S_R^1 \times P_d(1)$ *then* $z_s = E^s_{[\log d](h,\ f)}(z_0)$ *is an approximate zero of* f *with probability* $1-\mu$ *for any*

$$s \geqq e K_1 d \left(\frac{|\log \mu|}{\mu} \right)^{1+1/[\log d]} + K_2.$$

Proof. — Let k the smallest integer greater than log d in Theorem 4, then $d^{(k+1)/k} \leqq ed$.

If the constants K_1, K_2 in Theorem 3 and its corrollary were allowed to depend on k, K_1 would be approximately $K(k)$ of section 3 and K_2 even smaller. Recall that $K(k)$ decreases rapidly with k to about 7. One obviously can construct an appropriate algorithm for Theorem 3 and its Corollary.

To prove Theorem 3 we want to estimate the normalized Lebesgue measure of the set in $S^1_R \times P_d(1)$ where ρ_f is small or θ_{f, z_0} is small.

$$Y_{\alpha, \sigma, R} = \{ (z_0, f) \in S^1_R \times P_d(1) \mid \rho_f < \alpha \text{ or } \theta_{f, z_0} < \sigma \}.$$

PROPOSITION 1. — *For* $R > 2$, $\sigma < \pi/2$,

$$\mu(Y_{\alpha, \sigma, R}) \leqq (d-1)\alpha^2 + \frac{2}{\pi}\left(\frac{d-1}{d}\right)\left(\sigma + 2 \text{ arc sin } \frac{1}{R-1}\right).$$

First we prove Theorem 3 from the Proposition.

Besides k, and d we are given μ, $0 < \mu < 1$. The following equations are easily solved for α, R, and σ.

$$(d-1)\alpha^2 = \frac{\mu}{10},$$

$$\frac{4}{\pi}\left(\frac{d-1}{d}\right) \text{ arc sin } \frac{1}{R-1} = \frac{\mu}{10},$$

$$\frac{2}{\pi}\left|\frac{d-1}{d}\right| \sigma = \frac{4\mu}{5},$$

In fact

$$\alpha = \left(\frac{\mu}{10}\left(\frac{1}{d-1}\right)\right)^{1/2},$$

$$R = \left[\sin \frac{\mu\pi}{40}\left(\frac{d}{d-1}\right)\right]^{-1} + 1,$$

$$\sigma = \frac{2\pi\mu}{5}\left(\frac{d}{d-1}\right).$$

From Proposition 1 it follows that $\mu(Y_{\alpha, \sigma, R}) < \mu$ i.e. the measure of $Y_{\alpha, \sigma, R}$ is less than the given μ using the probability measure on $S^1_R \times P_d(1)$. Thus (z_0, f) in $S^1_R \times P_d(1)$ is not in $Y_{\sigma, \alpha, R}$ with probability $1-\mu$. For such z_0, f, $\rho_f > \alpha$ and $\theta_{f, z_0} > \sigma$. We apply the Corollary of Theorem 2 to obtain an approximate zero of f.

To find the K_1, K_2 and estimates of Theorem 3, one calculates those quantities from that Corollary. We don't carry out the straightforward calculation here, but indicate how it goes. In the Corollary (Theorem 2) the number of steps s (denoted by n in the corollary) is given as a function of $|f(z_0)|$, θ, and ρ_f (instead of L); keep k fixed throughout.

But s in Theorem 3 is given as a function of d and μ. Thus it is required to see how d and μ depend on $|f(z_0)|$, and ρ_f. This goes as follows.

From the construction of $Y_{\alpha,\sigma,R}$ we may replace θ by σ and ρ_f by α. Moreover since

$$|f(z_0)| = \left|\sum_{i=0}^{d} a_i z_0^i\right| \leqq \sum_{i=0}^{d} R^i = \frac{R^{d+1}-1}{R-1},$$

we may replace $|f(z_0)|$ by $(R^{d+1}-1)/(R-1)$. These remarks yield the number of steps given by the Corollary as a function now of σ, α, and R. By substituting the values of σ, α, and R given by the above equations, yields s in terms of μ and d. This function s of μ and d simplifies to give Theorem 3.

It remains to prove Proposition 1. We have

PROPOSITION 2. — (Smale)

$$\mathrm{Vol}\{f \in P_d(1) \mid \rho_f < \alpha\} \leqq (d-1)\alpha^2$$

Where Vol means normalized Lebesgue measure in $P_d(1)$. [*Actually Smale writes* $d\alpha^2$ *but the proof gives* $(d-1)\alpha^2$.]

Thus the remaining work of this section is to prove:

PROPOSITION 3. — *For* $\alpha < \pi/2$, $f \in P_d(1)$, $R > 2$,

$$\mathrm{Vol}\{z_0 \in S_R^1 \mid \theta_{f,z_0} < \alpha\} \leqq \frac{2}{\pi}\left(\frac{d-1}{d}\right)\left(\alpha + 2 \arcsin \frac{1}{R-1}\right).$$

In fact $\{z_0 \in S_R^1 \mid \theta_{f,z_0} < \alpha\}$ *is contained in at most* $2(d-1)$ *arcs of* S_R^1 *of angle* $2/d\,[\alpha + 2 \arcsin 1/(R-1)]$.

Here Vol again is normalized Lebesgue measure so that $V(S_R^1) = 1$. Note that Proposition 1 follows at once from Propositions 2 and 3.

The following lemma is essentially the Gauss-Lucas Theorem, *see Henrici*. Recall that for two complex numbers z, w, $\mathrm{Re}(\bar{z}w)$ is the usual inner product of z and w as vectors in R^2.

LEMMA 1. — *Let* $f = z^d + a_{d-1} z^{d-1} + \ldots + a_0$ *be a complex polynomial and let* S_R^1 *contain all the roots of f in its interior. Then for any* $z \in S_R^1$

$$\mathrm{Re}\left(\bar{z}\frac{f(z)}{f'(z)}\right) > 0.$$

Moreover the Newton differential equation $dz/dt = -f(z)/f'(z)$ *is transversal to* S_R^1 *and points inward.*

Proof. — Let $w_1, \ldots, w_n$ be the roots of f. Then the w_i are in the half plane defined by the tangent space to the circle at z, i. e. $\mathrm{Re}(\bar{z}w_i) < \bar{z}z$.

It follows that

$$\mathrm{Re}(\bar{z}(z-w_i)) > 0,$$

$$\mathrm{Re}\left(\frac{z}{z-w_i}\right) > 0,$$

$$\sum_i \mathrm{Re}\left(\frac{z}{z-w_i}\right) > 0,$$

$$\operatorname{Re}\left(z\frac{f(z)}{f'(z)}\right)>0$$

and

$$\operatorname{Re}\left(\bar{z}\frac{f'(z)}{f(z)}\right)>0.$$

For the last part if $z\in S^1_R$, then the inner product

$$\left(z, \frac{-f(z)}{f'(z)}\right)=\operatorname{Re}\left(\frac{-\bar{z}f(z)}{f'(z)}\right)<0.$$

LEMMA 2. — *Suppose* $f\in P_d(1)$, $R\geqq 2$. *Let* $z_1, z_2\in S^1_R$ *such that*

$$\left|\arg\frac{z_1}{z_2}\right|<\beta<\frac{\pi}{d}.$$

Then

$$d\beta+2\arcsin\frac{1}{R-1}\geqq\left|\arg\frac{f(z_1)}{f(z_2)}\right|\geqq d\beta-2\arcsin\frac{1}{R-1}.$$

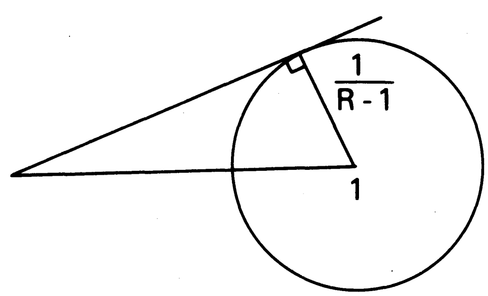
Proof.

$$\frac{f(z)}{z^d}=1+\frac{a_{d-1}z^{d-1}+\ldots+a_0}{z^d}$$

and for $z\in S^1_R$

$$\frac{\left|\sum_{j=0}^{d-1}a_jz^j\right|}{|z^d|}\leqq\frac{\sum_{j=0}^{d-1}R^j}{R^d}=\frac{R^{d-1}-1}{(R-1)/R^d}<\frac{1}{R-1}.$$

Thus $f(z)/z^d$ is inside the circle of radius $1/(R-1)$ centered by 1, and



$$\left|\arg\frac{f(z)}{z^d}\right|\leqq\arcsin\frac{1}{R-1}.$$

Now

$$\frac{f(z_1)}{f(z_2)}=\frac{z_1^d(f(z_1)/z_1^d)}{z_2^d(f(z_2)/z_2^d)}=\left(\frac{z_1}{z_2}\right)^d\frac{f(z_1)/z_1^d}{f(z_2)/z_2^d}$$

and the lemma follows from the fact that the argument of a product is additive and the triangle inequality.

Let θ be a critical point of f. Let L be the ray through $f(\theta)$, $L=\{\lambda f(\theta) \mid \lambda>0\}$ and Σ_θ the component of $f^{-1}(L)$ which contains θ. Let $\sum = \bigcup_{\substack{\theta \\ f'(\theta)=0}} \sum_\theta$.

LEMMA 3. — *If* $R \geqq 2$, $f \in P_d(1)$ *then* $\sum \cap S^1_R$ *is a set of at most* $2(d-1)$ *points.*

Proof. — Recall from section one that the inverse images of rays by f are solution curves of the differential equation $dz/dt=(-f(z))/f'(z)$ which by Lemma 1 are transversal to S^1_R for any $R \geqq 2$. It follows that for any fixed θ_i such that $f'(\theta_i)=0$, $\sum_{\theta_i} \cap S^1_R$ consists of at most k_i+1 points where k_i is the multiplicity of θ_i *as a root of* f'. Thus

$$\sum_{\substack{\theta_i \\ f'(\theta_i)=0}} k_i=d-1 \qquad \text{and} \qquad \sum_{\substack{\theta_i \\ f'(\theta_i)=0}} (k_i+1) \leqq 2(d-1).$$

Equality is only obtained here if each critical point θ_i is a simple zero of f' i. e. if each $k_i=1$.

Consider now the map $f_\xi^{-1}: S_{\xi, f} \to \mathbb{C}$ for ξ a root of f (see section 2). Recall that $S_{\xi, f}=\mathbb{C}-Q$ where Q consists of parts of certain rays. Now given an angle α, let $U_{\xi, \alpha}$ be the set of all numbers w in $S_{\xi, \alpha}$ such that $\arg w/q<\alpha$, some $q \in Q$. It follows that:

LEMMA 4. — *Let* $\alpha<\pi/2$. *If* z *is not in any* $f_\xi^{-1}(U_{\xi, \alpha})$ *for any* ξ *with* $f(\xi)=0$, *then* $\theta_{f, z}>\alpha$.

Moreover:

LEMMA 5. — *Let* $\alpha<\pi/2$, *and*

$$N(\alpha)=S^1_R \cap \bigcup_{\substack{\xi \\ f(\xi)=0}} f_\xi^{-1}(U_{\xi, \alpha}).$$

Then

$$\operatorname{Vol} N(\alpha) \leqq \frac{2(d-1)}{\pi d}\left(\alpha+2 \arcsin \frac{1}{R-1}\right).$$

Here Vol refers to normalized Lebesgue measure in S^1_R.

Proof. — Keep in mind that the f_ξ^{-1} have disjoint images. Using Lemma 3 it follows that $N(\alpha)$ is contained in the set of at most $2(d-1)$ intervals $N(\alpha, z_0)$, $z_0 \in \sum \cap S^1_R$, about the $2(d-1)$ points of $\Sigma \cap S^1_R$. The problem is to estimate the measure of these intervals. This is accomplished as follows. If $z \in N(\alpha, z_0)$, then $|\arg f(z)/f(z_0)|<\alpha$. By Lemma 2,

$$\left|\arg \frac{z}{z_0}\right| d-\alpha \leqq 2 \arcsin \frac{1}{R-1} \quad \text{so} \quad \left|\arg \frac{z}{z_0}\right| \leqq \frac{1}{d}\left(\alpha+2 \arcsin \frac{1}{R-1}\right).$$

Since there are two sides of z_0 in $N(\alpha, z_0)$ the total angle of $N(\alpha)$ is is bounded by $4(d-1)/d[\alpha+2 \arcsin 1/(R-1)]$. Dividing by 2π yields Lemma 5. With Lemma 4 that gives us Proposition 3.

Section 5

The goal of this section is to extend Theorem 2 to any incremental algorithm of efficiency k. Since the proof is similar to the proof of Theorem 2, we are brief.

Let

$$\Lambda_{f, z_0} = \min_{\substack{\theta,\ f'(\theta)=0 \\ |f(\theta)|<2|f(z_0)|}} \left| \arg \frac{f(z_0)}{f(\theta)} \right|.$$

THEOREM 4. — *Suppose that an incremental algorithm* I *has efficiency* k. *Then there is a constant* K *depending only on I such that.*

If $\Lambda_{f, z_0}>0$ and $|f(z_0)|>L>0$, then there is an h given explicitly such that $|f(z_n)|<L$ for

$$n=K\left\lceil \frac{\log |f(z_0)|/L}{\Lambda_{f, z_0}}+1 \right\rceil^{1+1/k},$$

where $z_n=(I_{h, f})^n(z_0)$.

Λ_{f, z_0} coincides with θ_{f, z_0} in *Smale* but not with the θ_{f, z_0} we have used above. For our θ_{f, z_0}, $\Lambda_{f, z_0} \leqq \theta_{f, z_0}$ $\lceil x \rceil$ means the smallest integer greater than or equal to x. We take less care about the constants here than in the proof of Theorem 2, although some estimate is given at the end of the proof of the theorem. Recall that $I_{h, f}$ is of efficiency k provided that there exist real constants $\delta>0$, $K>0$, $c_1, \ldots, c_k$, $c_1>0$ independent of h, f and z such that

$$\frac{f(I_{h, f}(z))}{f(z)} = 1-(c_1 h+\ldots+c_k h^k)+S_{k+1}(h)$$

where

$$|S_{k+1}(h)| \leqq K h^{k+1} \max\left(1, \frac{1}{h_1^k}\right) \qquad \text{for} \quad 0<h<\delta \min(1, h_1).$$

LEMMA 1. — *There is a constant* a, $1 \geqq a \geqq 0$, *depending only on* I *such that:*
If $0<h<a \min(1, h_1)$ *then*

$$\left|\frac{f(z')}{f(z)}\right| < 1-\frac{c_1 h}{2} \qquad \textit{and} \qquad \left|\arg \frac{f(z')}{f(z)}\right| < 2K h^{k+1} \max\left(1, \frac{1}{h_1^k}\right)$$

Proof. — Let

$$a=\min\left(1,\ \delta,\ \frac{1}{3c_1},\ \frac{c_1}{4\sum_{i=2}^{k}|c_i|},\ \left(\frac{c_1}{4K}\right)^{1/k}\right).$$

Then $f(z')/f(z)=1-c_1\,h-\sum_{i=2}^{k} c_i h^i+S_{k+1}(h)$ where for $h<a\min(1,\ h_1)$

$$\left|\sum_{i=2}^{k} c_i h^i\right|\leqq h^2\sum_{i=2}^{k}|c_i|<ha\sum_{i=2}^{k}|c_i|\leqq h\frac{c_1}{4},$$

$$|S_{k+1}(h)|\leqq hK\max\left(h_k,\ \left(\frac{h}{h_1}\right)^k\right)\leqq hK\,a^k<h\frac{c_1}{4}.$$

Thus

$$\frac{f(z')}{f(z)}=1-c_1 h+\alpha \quad \text{where } |\alpha|<\frac{hc_1}{2} \quad \text{and} \quad |\operatorname{im}\alpha|\leqq|S_{k+1}(h)|.$$

So

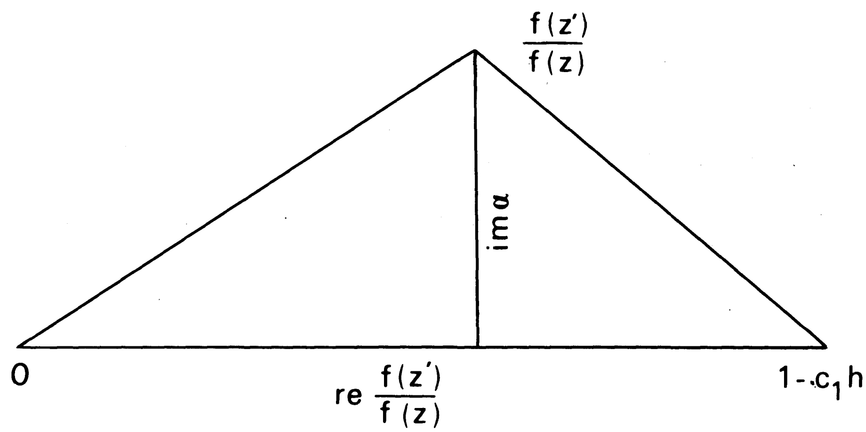
$$\left|\frac{f(z')}{f(z)}\right|\leqq 1-\frac{c_1 h}{2}.$$

Since

$$\frac{3}{2}c_1 h<\frac{3}{2}c_1 a\leqq\frac{1}{2},\qquad 0<1-\frac{3}{2}c_1 h<re\frac{f(z')}{f(z)}.$$

Now

$$\left|\arg\frac{f(z')}{f(z)}\right|=\arc\tan\left(\frac{|\operatorname{im}\alpha|}{re(f(z')/f(z))}\right).$$



and by Lemma 3 of section 3

$$\left|\arg \frac{f(z')}{f(z)}\right| \leqq \frac{|\lim \alpha|}{\mathrm{re}(f(z')/f(z))} \leqq \frac{h^{k+1} \mathrm{K} \max(1, 1/h_1^k)}{1-(3/2) \mathrm{c}_1 \mathrm{h}} \leqq 2 \mathrm{K} h^{k+1} \max\left(1, \frac{1}{h_1^k}\right)$$

LEMMA 2. — *For* $0<\theta \leqq \pi/2$, $1<\theta/\sin\theta \leqq \pi/2$.

Proof.

$$\frac{\pi}{2} = \frac{\pi/2}{\sin \pi/2}, \qquad \lim_{\theta \to 0} \frac{\theta}{\sin\theta} = 1$$

and there are no critical points of $\theta/\sin\theta$ in the open interval since $(\theta/\sin\theta)' = (1/\sin\theta)(1-\theta \operatorname{cotan}\theta)$ and $\theta \operatorname{cotan}\theta < 1$ for $0<\theta<\pi/2$ since arc tan $x<x$ in this range.

Let $h^* = a/2 \sin \Lambda_{f, z_0}$.

LEMMA 3. — *Let* $0<h<h^*$. — *Let* $z_n = \mathrm{I}_{h, f}^n(z_0)$. *Then for all*

$$0 \leqq n-1 \leqq \frac{\Lambda_{f, z_0}}{2} \frac{1}{2\mathrm{K}} \frac{h^{*k}}{h^{k+1}}, \qquad \Lambda_{f, z_{n-1}} > \frac{\Lambda_{f, z_0}}{2}$$

and

$$\frac{|f(z_n)|}{|f(z_0)|} < \left(1 - \frac{c_1 h}{2}\right)^n.$$

Proof. — First note that if $\Lambda_{f, z_k} \geqq 1/2 \Lambda_{f, z_0}$ then

$$h_{1(f, z_k)} \geqq \sin 1/2 \Lambda_{f, z_0} > 1/2 \; sin \; \Lambda_{f, z_0}$$

by Lemmas 2 and 3 of section 3 so

$$h^* = \frac{a}{2} \sin \Lambda_{f, z_0} < a \min(1, h_{1(f, z_k)})$$

and

$$\left|\arg \frac{f(z_{k+1})}{f(z_k)}\right| < 2 \mathrm{K} h^{k+1} \max\left(1, \frac{1}{h^{*k}}\right) = 2 \mathrm{K} \frac{h^{k+1}}{h^{*k}}.$$

We may proceed by induction as long as

$$\Lambda_{f, z_0} - (n-1) \frac{2 \mathrm{K} h^{k+1}}{h^{*k}} > \frac{\Lambda_{f, z_0}}{2} \qquad \text{or} \qquad n-1 \leqq \frac{\Lambda_{f, z_0}}{2} \frac{1}{2\mathrm{K}} \frac{h^{*k}}{h^{k+1}}.$$

Proof of Theorem 2.

$$h^* = \frac{a}{2} \sin \Lambda_{f, z_0} \geqq \frac{a}{2} \frac{2}{\pi} \Lambda_{f, z_0} = \frac{a}{\pi} \Lambda_{f, z_0}.$$

Let $0<h<h^*$. Then for all

$$n-1\leqq\frac{1}{4\mathrm{K}}\left(\frac{a}{\pi}\right)^k\frac{\Lambda_{f,\,z_0}^{k+1}}{h^{k+1}}\quad\text{we have}\quad\left|\frac{f(z_n)}{f(z_0)}\right|<\left(1-\frac{c_1 h}{2}\right)^n.$$

So to have an integer n such that

$$|f(z_n)|<\left(1-\frac{c_1 h}{2}\right)^n|f(z_0)|\leqq\mathrm{L}\quad\text{we want}\quad\left(1-\frac{c_1 h}{2}\right)^n\leqq\frac{\mathrm{L}}{|f(z_0)|}$$

or

$$n\log\left(1-\frac{c_1 h}{2}\right)\leqq\log\frac{\mathrm{L}}{|f(z_0)}|.$$

Thus it suffices to have

$$n\geqq\frac{\log|f(z_0)|/\mathrm{L}}{c_1 h/2}$$

To take n this large it suffices to have

$$\frac{1}{4\mathrm{K}}\left(\frac{a}{\pi}\right)^k\frac{\Lambda_{f,\,z_0}^{k+1}}{h^{k+1}}>n-1,$$

so it suffices to have

$$\frac{1}{4\mathrm{K}}\left(\frac{a}{\pi}\right)^k\frac{\Lambda_{f,\,z_0}^{k+1}}{h^{k+1}}\geqq\frac{\log|f(z_0)|/\mathrm{L}}{c_1 h/2}$$

or

$$h^k\leqq\frac{(c_1/8\,\mathrm{K})\,(a/\pi)^k\,\Lambda_{f,\,z_0}^{k+1}}{\log|f(z_0)|/\mathrm{L}}.$$

So we have proven. If

$$0<h\leqq\left(\frac{c_1}{8\,\mathrm{K}}\right)^{1/k}\frac{a}{\pi}\left(\frac{\Lambda_{f,\,z_0}^{k+1}}{\log|f(z_0)|/\mathrm{L}}\right)^{1/k},$$

then $|f(z_n)|<\mathrm{L}$ for

$$n=\left\lceil\frac{(2/c_1)\,|\log|f(z)|/\mathrm{L}}{h}\right\rceil.$$

Take

$$h=\left(\frac{c_1}{8\,\mathrm{K}}\right)^{1/k}\frac{a}{\pi}\left(\frac{\Lambda_{f,\,z_0}^{k+1}}{\log|f(z_0)|/\mathrm{L}}\right)^{1/k}$$

which gives $|f(z_n)| < L$ for

$$n=\left[\left(\frac{8\,K}{c_1}\right)^{1/k}\frac{2\,\pi}{ac_1}\left(\frac{\log|f(z_0)|/L}{\Lambda_{f,\,z_0}}\right)^{1+1/k}\right].$$

Q.E.D.

REFERENCES

K. Atkinson, *An Introduction to Numerical Analysis*, Wiley, New York, 1978.

E. Durand, *Solutions Numériques de Équations Algébriques*, Masson, Paris, 1960.

P. Duren, *Coefficients of Univalent Functions* (*Bull. Amer. Math. Soc.*, Vol. 83, No. 5, 1977, pp. 891-911).

L. Euler, *Institutiones Calculi Differentialis II, exp. IX. Opera Omnia*, Serie I, Vol. X, pp. 422-455.

W. Hayman, *Multivalent Functions*, Cambridge Univ. Press : Cambridge, Eng., 1958.

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

E. Hille, *Analytic Function Theory*, II Ginn, Boston, 1962.

A. S. Householder, *The Numerical Treatment of a Single Nonlinear equation*, McGraw-Hill, New York, 1970.

S. Lang, *Algebra*, Addison-Wesley, Reading, Mass, 1965.

A. Ostrowski, *Solutions of equations in Euclidean and Banach Spaces*, Academic Press, New York, 1973.

G. Saunders, *Thesis*, U. C. Berkeley, 1982.

E. Schröder, *Ueber unendlich viele Algorithmen zur Auflösing der Gleichungen* (*Math. Ann.* 2, 1870, pp. 317-365).

S. Smale, *The Fundamental Theorem of Algebra and Complexity Theory* (*Bull, Amer. Math. Soc.*, Vol. 4, No. 1, 1981, pp. 1-36.

(Manuscrit reçu le 24 mai 1983,
révisé le 4 janvier 1984).)

M. Shub,
Mathematics Dept, Queens College
and the Graduate School of C.U.N.Y.,
33 West 42 street,
New York,
N.Y. 10036-80099, U.S.A.

S. Smale,
Mathematics Dept,
University of California,
Berkeley, Cal. 94720.