Available online at www.sciencedirect.com
ScienceDirect
Electronic Notes in Theoretical Computer Science
ELSEVIER
Electronic Notes in Theoretical Computer Science 202 (2008) 201–218
www.elsevier.com/locate/entcs

Newton's method and the Computational Complexity of the Fundamental Theorem of Algebra

Prashant Batra[1]

Institute for Computer Technology
Hamburg University of Technology
Hamburg, Germany

Abstract

Several different uses of Newton's method in connection with the Fundamental Theorem of Algebra are pointed out. Theoretical subdivision schemes have been combined with the numerical Newton iteration to yield fast root-approximation methods together with a constructive proof of the fundamental theorem of algebra. The existence of the inverse near a simple zero may be used globally to convert topological methods like path-following via Newton's method to numerical schemes with probabilistic convergence. Finally, fast factoring methods which yield root-approximations are constructed using some algebraic Newton iteration for initial factor approximations.

Keywords: Subdivision schemes, global methods, factorization approach, constructive proofs.

1 Introduction

Weyl used Cauchy's integral theorem to show that a subdivision scheme combined with a root-proximity test is sufficient to obtain approximations of prescribed accuracy to all roots of a polynomial after a bounded number of steps. The method may be considerably accelerated locally by Newton's iteration if a zero or cluster is sufficiently isolated. This yields method of low arithmetic and boolean complexity, and is an interesting example of a theoretical proof generating a practical method considered in Section 2.

Cauchy gave a version of Argand's conitinuity-based proof in his 'Cours d'analyses' notes of 1821. This proof proceeds via a descent in modulus. Given a non-zero, a point with function value of smaller modulus might be determined in polynomial

[1] Email: batra@tuhh.de

doi:10.1016/j.entcs.2008.03.016

time as became clear in a constructivist proof by H.Kneser in 1940. Hirsch/Smale exhibited a never-failing method to approximate roots via Newton's method with a self-adapting Newton-correction but this method suffers from high complexity. Smale posed in 1981 the problem to approximate roots of polynomial equations using only (non-modified) Newton steps for approximation, and in Shub/Smale suggested to approximate solutions via continuity of solutions along a coefficient homotopy. It is well-known that Newton's method fails if a zero of the derivative is produced, and it slows down near multiple roots (i.e. near the set of ill-posed problems). Thus, the convergence of Newton's method (which implies succesful root-approximation) might be measured in probabilistic terms considering the measure of succesful starting points or the distance of homotopies from the manifold of ill-posed problems. There exist no methods which come close to the theoretical lower complexity estimates, and it is not known whether purely iterative higher-dimensional methods exist. We review several results in Section 3.

Very fast methods in terms of arithmetic and bit-complexity are based on the factorization approach for a polynomial P. Building on tight root-moduli bounds a zero-free annulus separating k and $\deg P - k$ roots is determined. The factor F corresponding to k roots of P is approximated, and subsequently refined. The refinement via an algebraic Newton iteration is the workhorse of Schönhage's fundamental study as well as the more recent work of Neff/Reif which exhibits a low complexity exceeding $O(\deg P \log^5 \deg P)$ just by a logarithmic term depending on the coefficient size and factor approximation quality. The connection to Newton's method is outlined in Section 4.

2 Cauchy

Cauchy gave two proofs of the fundamental theorem of algebra. One after Argand's proof sketch of 1806, the other using the integral theorem (somewhat similar to Gauss' 3rd proof of 1816 [16]).

2.1 Cauchy's integral theorem

Theorem 2.1 *(FTA)*
Every algebraic equation over $\mathbb{C}$ with exact degree n, i.e.

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0 \text{ with } a_n \neq 0 \text{ and } a_i \in \mathbb{C}\ \forall i$$

has n roots in $\mathbb{C}$.

Starting from a root modulus bound like

$$1 + \max_{0 \leq i < n} |a_i|/|a_n|$$

(which was obtained by Cauchy in [5]), we obtain a circle of radius R which contains all roots in its interior.

Cauchy's integral theorem and its consequences imply that the increase in argument of $f(z)$ on the circle $|z| = R$, i.e.

$$\int_{|z|=R} f'(z)/f(z)dz$$

is proportional to n, which is the number of roots. This proof seemingly does not exhibit any method to approximate roots. But Cauchy's argument allows to count roots inside a curve by evaluation of the integral of f'/f.

This was used in a (intuitionistic) proof by Weyl to exhibit a subdivision scheme which yields all roots with a specified accuracy after a finite number of steps. The scheme runs as follows: The circle of radius R containing all roots is inscribed in a square of side length $2R$. This first square is *suspect*, i.e. it contains roots. In each level (of increasing precision) the suspect squares are subdivided into four (whence later researchers named this procedure *quadtree construction* [20]). Using a numerical evaluation of the integral f'/f, the connected sets of squares containing roots are marked as suspect.

Approaches to make this Cauchy-Weyl scheme computationally attractive substituted the numerical evaluation of the integral by other tests indicating proximity of roots. Typically, estimates for the root modulus are applied after shifting the origin to the center of a suspect square. A test which estimates the modulus of polynomial roots with but a constant overestimation factor is Turan's proximity test (we state here a sharp version after [3]):

$$\max_{v=1,\ldots,n} \left(\frac{|s_v|}{n}\right)^{1/v} \leq \max_j |\zeta_j| \leq \frac{2}{\sqrt{2}-1} \max_{v=1,\ldots,n} \left(\frac{|s_v|}{n}\right)^{1/v},$$

where the $s_v := \sum \zeta_v$ may be determined via $\sum_{i=0}^{n} a_i \zeta^i = a_n \prod (z - \zeta_i)$ from

$$\begin{aligned}
0 &= a_n \cdot s_1 + a_{n-1} \cdot 1, \\
0 &= a_n \cdot s_2 + a_{n-1} \cdot s_1 + a_{n-2} \cdot 2, \\
0 &= a_n \cdot s_j + \ldots + a_{n-j+1} \cdot s_{j-1} + a_{n-j} \cdot j, \quad j = 1, 2, \ldots, n, \\
0 &= a_n \cdot s_{n+k} + \ldots + a_1 \cdot s_{k+1} + a_0 \cdot s_k, \quad k = 1, 2, \ldots, n\,.
\end{aligned}$$

Using Graeffe's root-squaring process, we might obtain small annular rings for the root moduli. Graeffe's root-squaring proceeds from a monic polynomial $t_i(y)$ with roots $z_1, \ldots, z_n$ to a polynomial $t_{i+1}(y) := t_i(\sqrt{y})t_i(-\sqrt{y})$ with roots $z_1^2, z_2^2, \ldots$ at the cost of a polynomial multiplication (which may be effected via FFT techniques).

Thus, we may obtain after N squarings the close estimate

$$1 \leq \max_j |\zeta_j| / \max_{v=1,\ldots,n} \left(\frac{|s_{vN}|}{n} \right)^{1/(vN)} \leq \left(\frac{2}{\sqrt{2}-1} \right)^{1/N}.$$

The number of suspect squares remains bounded by $4n \cdot h$ in iteration step h even if proximity tests are computed with relative eror of 50% [19]. Thus, Pan [19] estimates the cost of the *Cauchy-Weyl-Turan* approach as order of

$$O(n^2 h \log n \log\log n)$$

arithmetic operations to approximate all n zeros within $diam/2^h$, where $diam$ denotes the diameter of the set of all the zeros (viz. [19], p.193).

To speed up the linear process of subdivision, Renegar, and independently, Pan suggested in 1987 to use Newton's iteration.

2.2 Newton iteration

The Newton iteration proceeds from x_i to $x_{i+1} := x_i - f(x_i)/f'(x_i)$ (if $f'(x_i) \neq 0$) to approximate solutions of $f(x) = 0$. Cauchy gave one of the first known *general* convergence criteria.

Theorem 2.2 *(Cauchy; 1829) Given a real polynomial f and a real value x_0. Put $h := -\frac{f(x_0)}{f'(x_0)}$. Let $\mathcal{M}$ denote the maximum of $f''(x)$ over $\mathcal{K} =< x_0 + h, h >$, and assume that*

$$C := \frac{|f(x_0)|\mathcal{M}}{|f'(x_0)|^2} \leq \frac{1}{2}. \tag{1}$$

Then there exists precisely one zero of f in $\mathcal{K}$.

$$\textit{If moreover } \frac{|f(x_0)|\mathcal{M}}{|f'(x_0)| \cdot \min_{x\in\mathcal{K}} |f'(x)|} \leq \frac{1}{2}, \tag{2}$$

the Newton iteration converges to the zero.

The lower bound (2) for the derivative modulus was made obsolete by the analysis of Ostrowski who showed (cf., e.g., [17]) that (1) is a sufficient condition for convergence to an isolated zero,

Ostrowski's condition is fulfilled near a root and quantitative convergence conditions are possible, see, e.g. [32].

Theorem 2.3 *Let $P(x) \in \mathbb{Z}[x]$ be square-free with $\deg P = m$ and $M = 1 + \|P\|_\infty$. If a root X^* of $P(X) = 0$ and the starting point of the Newton iteration, X_0, are at a distance of at most*

$$\delta_0 = [m^{3m+9}(1+M)^{6m}]^{-1}$$

the Newton iteration converges.

We will see below that the condition of the theorem implies that the root approximation lies close to a root, and isolated from all other roots. Generally, it may be shown that the Newton iteration converges to a root X^* from starting points inside a disc D containing X^*, if D has an isolation ratio of at least $2\sqrt{2}+\sqrt{13n}$ [22,18,20]. Estimates of the derivative f' on a domain (as used in Ostrowski's analysis) were replaced by (scaled) estimates of the higher derivatives $f^{(k)}$ at a single, common point in Smale's theory of point estimates. Rheinboldt [23] showed that Ostrowski's convergence condition implies a point estimate condition.

The Newton iteration converges if the updates decay exponentially.

Definition 2.4 Given a polynomial f of degree n and a starting point x_0. Suppose the sequence of Newton iterates $x_{i+1} = x_i - f(x_i)/f'(x_i)$ is well-defined for all $k \geq 0$, and that it holds true that

$$\|x_{k+1} - x_k\| \leq \left(\frac{1}{2}\right)^{2^k-1} \|x_1 - x_0\|.$$

Then we say that x_0 is an *approximate zero.*

From an approximate zero point x_0 the Newton iteration converges exponentially to an actual zero. Approximate zeros may be characterized from point estimates.

Definition 2.5 Given a complex polynomial f of degree n define

$$\alpha(f, x_0) := |\frac{f(x_0)}{f'(x_0)}| \sup_{k=2,\ldots,n} |\frac{f^{(k)}(x_0)}{k! f'(x_0)}|^{1/(k-1)}.$$

Theorem 2.6 *Given a polynomial f of degree n, and a starting point x_0. If*

$$\alpha(f, x_0) < \frac{1}{8},$$

then x_0 is an approximate zero. Thus, the Newton iteration $x_{i+1} = x_i - f(x_i)/f'(x_i)$ converges starting from x_0.

To speed up the Cauchy-Weyl-Turan root approximation scheme, the linear refinement of a root approximation via subdivision of a suspect square is replaced by the Newton iteration whenever it converges at least quadratically. Using either isolation radii or Smale's point estimates it may be shown [20] that only $O(nlog(hn))$

squares in Weyl's quadtree have to be treated. This allows to approximate all n zeros within $diam/2^h$, where $diam$ denotes the diameter of the set of all the zeros, with

$$O((n^2 \log n) \log(hn)) \quad \text{arithmetic operations.}$$

2.3 Root separation

For a square-free polynomial we might also want to give isolating squares for complex roots or isolating intervals for real roots. For integer polynomials, Cauchy established a lower bound for the minimum root separation [5], and employed it in a first real root isolation method via brute force partitioning.

Definition 2.7 The *minimum root separation* of an integer polynomial $P(x)$ given as

$$P(x) := \sum_{i=0}^{n} a_i x^i \in \mathbb{Z}[x] = a_n \cdot \prod_{i=1}^{n} (x - \zeta_i), \text{ where } a_n \neq 0,$$

is defined as $sep(P) := \min_{\substack{\zeta_i \neq \zeta_j \\ P(\zeta_i)=P(\zeta_j)=0}} |\zeta_i - \zeta_j|$.Let $s(P) := \sum_{i=0}^{n} |a_i|$.

Cauchy used the polynomial's discriminant,

Definition 2.8

$$discr(P) := a_n^{2n-2} \prod_{1 \leq i \leq n} \prod_{i<j} (\zeta_i - \zeta_j)^2$$

(esp. the fact that $discr(P) \in \mathbb{Z}$) in his memoir of 1829 (to be found in [5]) to derive the following estimate for the root separation.

Proposition 2.9 *Given $P(x) \in \mathbb{Z}[x]$ of degree n it holds true that*

$$sep(P) > \frac{\sqrt{discr(P)}}{|a_n|^{n-1}(2 \cdot s(P))^{\frac{n(n-1)}{2}-1}} \geq \frac{1}{|a_n|^{n-1}(2 \cdot s(P))^{\frac{n(n-1)}{2}-1}}.$$

Estimates of the computational cost of any subdivision or exclusion/inclusion scheme, like Cauchy's, Weyl's, Lehmer's [12], Gargantini/Henrici [6], will depend on good estimates of this root separation. Define the measure $M(P)$ of a polynomial as

$$M(P) := |a_n| \prod_{i=1}^{n} \max\{1, |\zeta_i|\}.$$

From Jensen's inequality, we find $M(P) \leq s(P) =: s$ (a first proof appears in [11]). Using the continous measure $M(P)$ Mahler derived the best published bound [13] from a study of the discriminant.

Theorem 2.10 *(Mahler; 1964) Let $P(x) = \sum_0^n a_i x^i$ be a square-free integer polynomial of size s and degree n. Then*

$$sep(P) > \frac{\sqrt{3} \cdot \sqrt{discr(P)}}{n^{n/2+1} \cdot M(P)^{n-1}} \geq \frac{\sqrt{3} \cdot \sqrt{discr(P)}}{n^{n/2+1} \cdot s^{n-1}}.$$

Considering $\hat{P}(x) := P(x)/gcd(P(x), P'(x))$ we find that a similar result holds true for arbitrary polynomials P.

Corollary 2.11 *Let $P(x) = \sum_0^n a_i x^i$ be an integer polynomial of size s and degree n. Then*

$$sep(P) > \frac{\sqrt{3}}{n^{n/2+1} \cdot M(P)^{n-1}} \geq \frac{\sqrt{3}}{n^{n/2+1} \cdot s^{n-1}}.$$

Open problem: To prove or disprove that

$$L(n) := \limsup_{deg(P)=n, s(P)=s \to \infty} \frac{-\log\ sep(P)}{s} \stackrel{!}{=} n - 1,$$

see [4].

3 Argand

Argand used the existence of $k-th$ roots in $\mathbb{C}$ together with the facts that every polynomial is a continous function on $\mathbb{C}$, and that continous functions on $\mathbb{C}$ attain a minimum over compact sets to give a proof of the fundamental theorem of algebra in 1806. The proof was reworked by Argand in 1813, and became widely known when it was published by Cauchy in 1820.

Argand's main point is that to every $c \in \mathbb{C}$ with $|f(c)| \neq 0$ there exists $c' \in \mathbb{C}$ with $|f(c')| < |f(c)|$. The continous function $|f|$ attains its minimum over a compact set containing all roots, and it would lead to a contradiction if this minimum were unequal to zero.

The existence of a descent in modulus may be shown as follows:

Let $h(z) := f(c+z)/f(c) = 1 + b_k z^k + b_{k+1} z^{k+1} + \ldots + b_n z^n$, and with $g(z) := b_{k+1} z + \ldots + b_n z^{n-k}$ write $h(z) = 1 + b_k z^k + z^k g(z)$. Denote the $k-th$ root of $-1/b_k$ as d, then it holds true for all real $t, 0 < t \leq 1$, that

$$|h(d \cdot t)| \leq |1 - t^k| + |d^k t^k g(dt)| = 1 - t^k + t^k |d^k g(dt)|.$$

As g is a continous function with $g(0) = 0$ there exists (we may choose) a point δ in $[0, 1]$ such that $|d^k g(dt)| < 1/2$ for all $0 < t < \delta$. For all such t we obtain $|h(d \cdot t)| \leq$

$1 - t^k + 1/2t^k < 1$. Thus, there exists $u \in \mathbb{C}$ with $|h(u)| = |f(c+u)/f(c)| < 1$ whenever $f(c) \neq 0$.

This nice analytic proof depends heavily on the notion of continuity, but it may also be turned into an effective computational scheme as H. Kneser showed in 1940. H. Kneser [9] gave his proof in the framework of intuitionism via an auxillary result which enabled a descent in modulus as

$$|f(x_k)| < (1 - \frac{1}{2})^k \cdot \max\{1, |f(x_0)|\}$$

Thus, we may trace the image $|f|$ such that in every considered point the modulus decreases. We may consider two different demands at this point:

i) To follow the curve of decreasing modulus numerically.

ii) Gauss' statement in his fourth and last proof of the FTA that any method capable of proving the existence of all roots simultaneously is of higher distinction:

'Indessen gewinnt ohne Zweifel jede Beweisführung eine höhere Vollendung, wenn nachgewiesen wird, dass sie geeignet ist, das Vorhandensein der sämmtlichen Factoren unmittelbar anschaulich zu machen.'

3.1 Global Newton methods after Hirsch, Smale and Shub

Another constructive version of Argand's proof (similar to Kneser's [9]) was given by Hirsch and Smale in ([7], see Section 6, p.303 ff.). In the same paper, they showed that for a modified Newton method

$$x_{i+1} = x_i - h_i \cdot f(x_i)/f'(x_i).$$

there exist open subsets W such that the 'incessantly' modified Newton method converges to a point $\hat{x}$ with $|f(\hat{x})| < \epsilon$ from every starting point inside W. The modification proceeds by halving the Newton-step length whenever the derivative vanishes or the modulus exceeds ϵ. Thus, a modulus descent is forced in every step but the running time is not polynomially bounded in terms of degree and modulus bound.

To avoid modifications of Newton's method Shub/Smale [27] considered to choose starting points at random. Choosing a point at random might yield a sequence of Newton iterates converging to some zero, or might lead to a breakdown of the method at a zero of the derivative.

3.1.1 Newton steps at large

Suppose we use a Lesbesgue measure on the continous set $P_n(1)$ of polynomials of degree no larger than n and with coefficients not exceeding 1 in modulus. Renegar [21], and later Smale [28] gave probability estimates for succesful Newton iteration.

Theorem 3.1 *A number of Newton steps sufficient to find an approximate zero of $f(x)$ with probability σ is*

$$100(n+2)^7/(1-\sigma).$$

3.1.2 Restarted generalized Newton method

Re-starting is often used in numerical schemes. It works rather well with higher order generalizations of Newton's method. In the words of S.Smale [29]: 'Choose a starting point at random [...]. Then apply some variation of Newton's method iteratively for a while. If that does not work, pick another starting point at random and repeat. [...] At this time Mike Shub and I [26] have a result which shows that for a polynomial of one complex variable, this method works in fact relatively quickly; six random choices are sufficient on the average.' With f_z^{-1} the branch of the inverse of f which takes $f(z)$ into z, given as an analytic function in a neighbourhood of $f(z)$ (provided $f'(z) \neq 0$), define

$$E_{k,h,f} = T_k(f_z^{-1}((1-h)f(z))),$$

where T_k denotes the power series expansion truncated at degree k. With $h := 1/512$ and $k = k(\epsilon) := \lceil \max(\log|\log\epsilon|, \log n)\rceil$ define $E_\epsilon(\cdot) := E_{k,h,f}(\cdot)$. The following algorithm was proposed in [26]:

Algorithm $(N-E)_\epsilon$ Let $f \in P_n(1)$ and $N = 512(n + |\log\epsilon|)$.

- Choose $z_0 \in \mathbb{C}, |z_0| = 3$ at random and set for $j = 1, 2, \ldots$ $z_i = E_\epsilon(z_{i-1})$. Terminate if at some stage $|f(z_j)| < \epsilon$.
- If $j = N$, return to the first step.

Theorem 3.2 *For each f and ϵ algorithm $(N-E)_\epsilon$ terminates with probability one and produces a z satisfying $|f(z)| < \epsilon$. The average number of cycles is less than or equal to 6. Hence, the average number of iterations is less than $6K(n + |\log\epsilon|)$.*

3.1.3 Negative analysis

A radical approach to root approximation is to fix some arbitrary coefficient homotopy P_t, connecting a chosen $P_0(z) = a_n \prod(z - \xi_i)$ and the target polynomial $P =: P_1$ and to to do one step along the homotopy curve in coefficient space after one root iteration step. Such path following procedure has been called purely iter-

ative by Smale, it may be described as a rational endomorphism on the polynomial coefficients, the iteration data, and the path homotopy.

In the framework of purely iterative schemes, McMullen studied [14] the rational schemes like Newton's regarding their global convergence properties. It transpired that non-convergence of rational iterations from a starting point is not a rare occasion.

Theorem 3.3 *There exists no one-dimensional, rational polynomial root approximation iteration which is* globally convergent*, i.e. which converges for every starting point outside a fixed set of measure zero.*

This result enhances Gauss' remark (quoted above) on the importance of simultaneous root approximation, and strengthens the case for path-following methods. Following a curve taking Newton-steps, or trying to approximate roots via Weierstraß' path-following is sensitive to the distance to ill-posedness.

3.2 Homotopy Path following

Shub and Smale showed that we may trace zeros along a homotopy in principle using Newton steps [27]. To state this result we need some definitions of theirs.

Consider a homotopy from f_0 to f_1, i.e. a continous family of holomorphic functions $f_t : \mathbb{C} \to \mathbb{C}, 0 \leq t \leq 1$. An *associated path* is a continous map from $[0,1] \times \mathbb{C} \to \mathbb{C}, t \to \zeta_t$ where

- $f_t(\zeta_t) = 0$
- $Df_t(\zeta_t)$ is an isomorphism.

Given a subdivision $T = \{t_0 = 0, t_1, \ldots, t_k = 1\}, t_i < t_{i+1}, |T| = k$ employ Newton's method for every t_i as

$$x_i = N_{f_{t_i}}(x_{i-1}).$$

It is said that *Newton's method follows the homotopy path* $\{f_t, \zeta_t\}$ if all x_i are well-defined, all iterates x_i are appromiate zeros of their appropriate f_{t_i}, quantified as $\alpha(f_{t_i}, x_i) < \frac{1}{4}(13 - 3\sqrt{17})$, and even more specifically the x_i are supposed to be approximate zeros to the actual zeros ζ_{t_i}. For a homotopy path $F = \{f_t, \zeta_t\}$ define $L = L(F)$ to be the length in the metric d_P of the curve f_t. Define the condition number of the homotopy path as

$$\mu = \mu(F) = \max_{0 \leq t \leq 1} \mu(f_t, \zeta_t) = \max_{0 \leq t \leq 1} \max\{1; \|f_t\| \|Df(\zeta_t)^{-1}\sqrt{n}(|\zeta|^2 + 1)^{n-1}\}.$$

Theorem 3.4 *(Shub/Smale) Let $F = \{f_t, \zeta_t\}$ be a homotopy path. Let*

$$k \geq \frac{Ln^{3/2}}{0.11}\mu^2.$$

Then k Newton steps are sufficient to follow the path $\zeta_t, 0 \le t \le 1$.

This is not a constructive result (see [31] for detailed remarks), it merely states that there exists a partitioning of $[0,1]$ of some *associated path* which allows the Newton iteration to trace from ζ_0 to ζ_1 producing approximate zeros all throughout.

3.3 Weierstraß

In the 1859 session of the Preußische Akademie der Wissenschaften Weierstrass presented a proof of the fundamental theorem of algebra. Its publication was postponed, as Weierstraß saw the dependency on continuity as a serious short-coming of the proof. In 1891, the proof was revised and published. It depends on a simultaneous Newton iteration for the first n symmetric functions of the n polynomial roots. The iteration

$$z_k^{(n+1)} = z_k^{(n)} - P(z_k^{(n)})/(a_n \prod_{i \neq k}(z_k^{(n)} - z_i^{(n)})) \tag{3}$$

was given explicitly with local convergence criterion in Weierstrass work; it has been re-discovered in the 20th century by Durand, Dochev, Kerner and others.

Weierstrass' convergence criterion depends on the root-separation together with the quality of the approximation to the roots (i.e. on 'unattainable data'). We state here an improved criterion of the same type [25].

Theorem 3.5 *(Dochev; 1962) Let P be a square-free polynomial of degree n with roots ζ_i. The Weierstraß-Dochev-Durand-Kerner iteration (3) converges with starting values $z_i^{(0)} =: z_i$ if*

$$|z_i - \zeta_i| \le \frac{\sqrt[n-1]{2} - 1}{2\sqrt[n-1]{2} - 1} sep(P) \tag{4}$$

Following the paradigm of Smale's point estimates we derive convergence conditions using attainable data from the iteration process (see, e.g., [1]).

Theorem 3.6 *Let P be a square-free polynomial of degree n with leading coefficient a_n. The Weierstraß-Dochev-Durand-Kerner iteration (3) converges with starting values $z_i^{(0)} =: z_i$ if*

$$\max_i |\frac{P(z_i)}{a_n \prod_{i \neq k}(z_i - z_k)}| \le \frac{\min_{i \neq j}|z_i - z_j|}{2n} \tag{5}$$

If the simultaneous iteration is used along a coefficient homotopy, a criterion like (5) would have to be evaluated in each move along the homotopy path. The allowable step-size along a homotopy from $P_0 = x^n - a_0$ to $P_1 = P$ can be estimated via

matching bounds (see (6) below) for the perturbed polynomials. Replacing the minimum root-separation by its best known lower bound due to Mahler we obtain, all in all, an exponential estimate for the number of Newton steps (see, e.g. the analysis in [30]).

Theorem 3.7 *Given a polynomial P with leading coefficient 1 which has pairwise distinct roots. Suppose that all polynomials of the homotopy*

$$H(t,x) := t \cdot p(x) + (1-t) \cdot (x^n - c), c \neq 0,\ t \in [0,1]$$

have parwise distinct roots. Denote the minimum over the minimum root separation of all polynomials $H_{\hat{t}}(x) := H(\hat{t},x)$ by V. The length of the homotopy curve is denoted by L. To approximate all roots of P with precision ϵ,

$$C_W \frac{L}{V} \cdot n + \log\log\frac{1}{\epsilon}\ \textit{Weierstraß iteration steps}$$

along the homotopy are sufficient, where C_W is a constant independent of n.

Explicit determination of V could be done via minimum root-separation estimates (after appropriate re-scaling each intermediate polynomial to integer coefficients). An arithmetic complexity estimate for path-following in this fashion would contain at least a term $O(s^n)$. It is probably instructive to compare this to the (non-constructive) minimum number of steps estimated by Shub/Smale.

4 Factoring a polynomial

As the FTA predicts n roots for a polynomial $P(x) = x^n + \sum_{i=0}^{n-1} a_i x^i = \prod(x - \zeta_i)$ of degree n we may try to produce n linear factors of P.

If the product of n factors $L_j = (x - \lambda_j)$ approximates P in terms of a coefficient vector norm ,$|P - L_1 \cdot L_2 \cdot \ldots \cdot L_N|_1 < \epsilon$, we might use perturbation estimates to obtain root approximation estimates.

We make here a simplified statement of a perturbation result published in [2] but already contained in the unpublished report [24] of Schönhage's.

Lemma 4.1 *Given two polynomials $f(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$,$g(x) = x^n + \sum_{i=0}^{n-1} b_i x^i$ with roots ζ_v and λ_v, respectively. A bound for both sets of roots is*

$$\gamma := 2 \cdot \max_v \{|a_k|^{1/(n-k)}; |b_k|^{1/(n-k)}\}.$$

Then the roots may be enumerated such that

$$\max|\zeta_v - \lambda_v| \leq 4 \cdot \sqrt[n]{\sum_{k=1}^{n} |b_k - a_k| \gamma^k} \tag{6}$$

To state more estimates relating to factorization we introduce the norm

$$\|u(x)\| := \sum_{i=0}^{n} |u_i| \text{ for } u(x) = \sum_{i=0}^{n} u_i x^i.$$

The effect of recursive splitting in terms of the coefficient norm may be estimated using the following results.

Lemma 4.2 *Let* $\|P(x) - f_1(x)f_2(x)\dots f_k(x)\| \leq k \cdot \epsilon\|P(x)\|/n$, $\|f_1 - f(x)g(x)\| \leq \epsilon_1\|f_1(x)\|$ *and* $\epsilon_1 \leq \epsilon\|p(x)\|/\left(n\prod_{i=1}^{k}\|f_i(x)\|\right)$. *Then* $\|P(x) - f(x)g(x)f_2(x)\dots f_k(x)\| \leq (k+1)\epsilon\|p(x)\|/n$.

Lemma 4.3 *For a polynomial factorization of a polynomial* $P(x)$ *of positive degree* n, *say* $P(x) = \prod_{i=1}^{k} f_i(x)$, *the following coefficient norm inequalities hold true.*

$$\|P(x)\| \leq \prod_{j=1}^{k} \|f_j(x)\| \leq 2^{n-1}\|P(x)\|.$$

Thus, if $P(x) \sim f_1(x)f_2(x)$ is split subsequently a norm-bound linear in ϵ and $\|p(x)\|$ may be maintained if the refined splitting is controlled by $\epsilon_1 = \epsilon/(n2^n)$.

To obtain an initial splitting a zero-free annulus A is determined such that A separates k inner zeros in D_{in} from $n-k$ lying in the unlimited component of $\mathbb{C} \setminus A$. If the smallest root has modulus less than 0.5 and the largest root modulus exceeds 2, then by appropriate re-scaling and shifting/scaling a zero-free annulus around $|z| = 1$ may be found. Otherwise, either $p^*(x) := x^n p(1/x)$ or $p(x)$ have all roots in $|z| \leq 2$. Shifting to the center of gravity of zeros together with appropriate re-scaling and approximate root-modulus determination leads to a polynomial with largest root inside $0.98 < |w| < 1$, and hence a diameter of the root set $u_1, u_2, \dots, u_n$ larger than 0.98. This allows to choose geometrically a new center v such that $|v - u_1|/|v - u_n| > 4/3$, and shifting and rescaling yield the zero-free annulus A.

The factor F of $P = F \cdot G$ corresponding to the k zeros in D_{in} is approximated by F^*. If the approximation quality is such that $\|F(x) - F^*(x)\| \leq 2^{-C \cdot h \cdot n}$ for a fixed constant C, then G^* is obtained via high-precision division of f by F^*. An approximation F^* may be obtained via the power sums $s_{m,k} = \sum z_i^m$ of the k zeros lying inside D_{in}. Instead of the Newton identities Cauchy's integral theorem may be used for the approximation of the $s_{m,k}$. Denoting by Γ a circle lying concentric inside A we have

$$s_{m,k} = \frac{1}{2\pi i}\int_{\Gamma} \frac{x^m p'(x)}{p(x)} dx.$$

Using quadrature formulas in Q equally spaced points on Γ the numerical integration may be obtained via three FFTs on Q points.

Proposition 4.4 *Suppose the inner boundary of A is a circle of radius r, and the outer boundary a circle of radius R with $(R-r)/r \geq \hat{c}/n^{\hat{d}}$ with positive constants $\hat{c}, \hat{d}$. To obtain a splitting F^*G^* of P with*

$$\|P(x) - F^*(x)G^*(x)\| \leq 2^{-C \cdot nh}\|P(x)\|$$

order of $(h+n)n^2$ Boolean operations are sufficient.

To improve the initial factorization F^*G^* of P a correction is determined. The correction to the initial factorization is a pair of polynomials f, g of degree $k-1$ and $n-k-1$, respectively, chosen such that

$$\begin{aligned} P - FG &= fG + gF \text{ or equivalently} \\ \frac{P-FG}{FG} &= \frac{f}{F} + \frac{g}{G} \, (\deg f < \deg F, \deg g < \deg G) \end{aligned}$$

If $F_{new} = F + f, G_{new} = G + g$ then

$$F_{new}G_{new} - P = (F+f)(G+g) - P = (FG - P) + fG + gF + fg = fg$$

so that

$$\|F_{new}G_{new} - P\| \leq \|f\|\|g\|.$$

The relations determining f and g allow a norm estimate. Determination of terms f, g could in principle be done by Euclid's algorithm (incurring high bit-complexity). We sketch Schönhage's [24] technique in the following.

4.1 The algebraic Newton iteration

A factorization $f_1 \cdot f_2 \cdot \ldots \cdot f_k = f$ with degrees $n_1, \ldots, n_k$ such that $\sum n_i = n$ may be considered (viz. [8]) as a root of the mapping $\alpha : \Pi_{n_1} \times \Pi_{n_2} \times \ldots \times \Pi_{n_k}, \alpha(p_1, p_2, \ldots, p_k) \to P - p_1 \cdot p_2 \cdot \ldots \cdot p_k$. With

$$\frac{1}{p_1 \ldots p_k} \sim \frac{h_1}{p_1} + \ldots + \frac{h_k}{p_k}$$

we might consider the h_k as an approximate encoding of the inverse Jacobian $J_\alpha(p_1, \ldots, p_k)^{-1}$ of α. Defining f_j by $f_j \equiv (h_j \cdot p) \mod p_j$ updates may be defined as $\hat{p_j} := p_j + f_j$.

Updating the factors implies updating of the Jacobian. This is approximated as follows. Suppose

$$d := 1 - \sum_{j=1}^{k} h_j \prod_{i=1; i \neq j}^{k} p_i$$

is the initial defect of the set $(h_1, \ldots, h_k)$. Replacing h_j by $\hat{h_j} := h_j \cdot (1+d) \mod p_j$ implies a quadratic relation between the new and the old defect.

To improve a factorization $F \cdot G \sim P$ corrections f, g are sought with

$$P - FG = fG + gF \quad \text{or} \quad \frac{P - FG}{FG} = \frac{f}{F} + \frac{g}{G}. \tag{7}$$

Schönhage's approach to approximate such partial fraction decompostion is as follows. Suppose with a D of small norm we have

$$\begin{aligned} HG + QF = 1 - D, \quad HG \sim 1 - D \mod F \\ \text{then } \hat{H} := H(1 + D) \sim 1 - D^2 \mod F. \end{aligned}$$

If $\hat{H}$ is restricted to be a polynomial of degree $k - 1$ it is uniquely determined. Polynomials f and g of degree not exceeding $k - 1$ and $n - k - 1$ respectively may be uniquely determined from

$$\begin{aligned} f \sim \hat{H}P \sim \hat{H}(P - FG) \mod F \\ (P - FG) - fG = gF + R \quad \text{with} \quad \deg R \leq k - 1. \end{aligned}$$

With a zero-free annulus (after transformations) around $E := \{z : |z| = 1\}$ a lower bound for $\mu := \min_{z \in E} |P(z)|$ may be established. The polynomial is then normalized such that $\|P\| = 1$. If

$$\begin{aligned} F(x) = z^k + \phi z^{k-1} + \ldots + \phi_k, \quad G(x) = a_n(z^{n-k} + \ldots) \\ \|FG - P\| \leq \mu/8 \end{aligned}$$

then $\|FG\| \leq 1 + \mu/8 \leq 9/8 \cdot \|P\|$ and $\|F\| < 2^k$, $\|F\| < \frac{9}{8} 2^{n-k}$. Schoenhage's analysis shows that if $|D| \leq \mu^2/(k^2 2^{2n})$, where $\mu = \min_{z \in E} |P(z)|$ a this iteration has exponential convergence.

The norms of f and g are bounded using the relation $P - FG = fG + gF$, and the estimates on the zero-free annulus. Schönhage's analysis in [24] yields a complexity estimate in terms of the bit complexity (we state a reformulation after Neff and Reif below as Proposition 4.5). It was noted in [24] that the algorithm might suffer from non-balanced factors, e.g. the splitting into factors of degree $n - 1$ and 1, respectively. Such non-balancing might incur $n - 1$ splittings in total.

The basic splitting technique from [24] was supplemented in [15] by a balanced splitting of f into $f_1 \cdot f_2 \cdot f_3$ such that either $\max_{i=1,2,3} \deg f_i \leq n/2$ or if $\deg f_k > n/2$ then $f_k = f_4 f_5 f_6$ with $\max_{i=4,5,6} \deg f_i \leq n/2$. This allows to apply the algorithm recursively to polynomials of degree $n/2$.

Neff/Reif [15] considered the space $\mathcal{P}_n$ of monic polynomials $z^n + \sum_{i=0}^{n-1} c_i z^i$ with non-trivial coefficients bounded as $|c_i| < 2^m$, where m is minimal. The output precision is measured in terms of μ such that

$$|z_i - w_i| \leq 2^{-\mu}.$$

(This has no relation with the previous use of μ. The cluster isolation is measured in terms of δ-isolation: A disk $D = D(z_0, R)$ is called δ-isolated for a polynomial f if there are no roots in the annulus

$$T_D = D(z_0; (1+\delta)R) \setminus D(z_0; (1+\delta)^{-1} \cdot R).$$

Using appropriate transformations of the matching bounds (6) and factor estimates in Lemmas 4.2, 4.3 Schönhage's method to construct and refine an initial splitting is the method supporting the following result.

Proposition 4.5 *Suppose we are given a δ-isolated ($\delta \geq 0.4$) disk D containing k roots ($1 \leq k < n$) of the polynomial $f \in \mathcal{P}_n$. Then there exists an $O(n \log^2 n \log(m+\mu))$ arithmetic algorithm for computing a factorization of f into approximate factors $\tilde{f}_1, \tilde{f}_2$ corresponding to roots inside and outside D, respectively, such that the distance of $\tilde{f}_i$ to f_i ($f = f_1 \cdot f_2$) is at most*

$$|f_i - \tilde{f}_i| < 2^{(2n+nm)-(\mu+2n+n\max\{n;m\}+\log\log n)}$$

If the δ-isolation is not sufficiently strong the following chain of ideas is used:

Graeffe's root squaring allows to obtain an isolation ratio of $4/10$ after at most $k = \lceil \log \delta \rceil - 1$ squarings. Under the assumption that $m \geq n$, and if after $i-1$ squarings the moduli of f_i lie inside $[2^{-m/2}, 2^{m/2}]$ but outside after the i-th squaring a $4/10$-isolated disk exists. Thus, the coefficients of all squared polynomials are bounded by 2^m. From the factors of f_k the factors of f reconstructed using a partial GCD computation. The partial GCD is computed via structured determinants. Neff and Reif obtain the following.

Theorem 4.6 *Suppose we are given a δ-isolated ($\delta > 0$) disk D containing k roots ($1 \leq k < n$) of the polynomial $f \in \mathcal{P}_n$. Then there exists an $O(n \log^2 n \log(m+\mu) \log 1/\delta)$ arithmetic algorithm for computing a factorization of f into approximate factors $\tilde{f}_1, \tilde{f}_2$ corresponding to roots inside and outside D, respectively, such that the distance of $\tilde{f}_i$ to f_i ($f = f_1 \cdot f_2$) is at most*

$$|f_i - \tilde{f}_i| < 2^{(2n+nm)-(\mu+2n+n\max\{n;m\}+\log\log n)}$$

Turan's proximity test is used to obtain high-precision inclusions $[L_i, U_i]$ of the root moduli $r_i = |\zeta_i|$, where the roots are numbered sucht that $r_1 < r_2 < \ldots$. The inclusion quality determines different root regions according to which the polynomial is split. If the relative separation $(L_{i+1} - U_i)/L_{i+1}$ is sufficiently large, i.e. greater than $1/(168n^2)$ while $U_i/L_i \leq 1 + 1/(168n^2)$, and $\lfloor n/4 \rfloor \leq i \leq \lceil 3n/4 \rceil$, we may use the partial GCD based splitting producing factors of degree no less than $n/4$. If the relative separation of root moduli is not sufficient, a new center is determined after consideration of several root regions. Comparing different annuli and circles containg the roots a new *balanced splitting* point is determined. Neff/Reif's techniques allow recursive balanced splitting, and yield the following result.

Theorem 4.7 *Given a monic polynomial* $P(z) = z^n + \sum_{i=0}^{n-1} c_i z^i$ *of degree* n *with coefficients bounded as* $|c_i| < 2^m$ *. Linear factors* $f_1, f_2, \ldots, f_n$ *such that*

$$|P - \prod_{i=1}^{n} f_i| < 2^{-\mu}$$

may be determined by an algorithm of arithmetic complexity

$$O(n \log^5 n \log(m + \mu)).$$

References

[1] Batra, P. Improvement of a convergence condition for the Durand-Kerner iteration. *JCAM*, 96:117–125, 1998.

[2] Bhatia, R.; L. Elsner; G. Krause Bounds for the Variation of the Roots of a Polynomial and the Eigenvalues of a Matrix. *Linear Algebra and its Applications*, 142:195–209, 1990.

[3] Buckholtz, J. D. Sums of powers of complex numbers. *J.Math.Anal.Apl.*, 17:269–279, 1967.

[4] Bugeaud, Y.; M. Mignotte On the distance between roots of integer polynomials. *Proc. Edinburgh Mathematical Society*, 47:553–556, 2004.

[5] Cauchy, A.L. *Oeuvres complètes d'Augustin Cauchy*, volume 3 of 2^e *série.*, chapter Sur la résolution numérique des équations, pages 378–426. Gauthier-Villars, Paris, reprint of 1897 edition, 1921.

[6] Gargantini, I. ; P. Henrici Circular arithmetic and the determination of polynomial zeros. *Numerische Mathematik*, 18:305 – 320, 1972.

[7] Hirsch, Morris W.; Stephen Smale On Algorithms for Solving $f(x) = 0$. *Communications on Pure and Applied Mathematics*, XXXII:281–312, 1979.

[8] Kirrinnis, Peter. Partial Fraction Decomposition in $\mathbb{C}(z)$ and Simultaneous Newton Iteration for Factorization in $\mathbb{C}[z]$. *Journal of Complexity*, 14:378–444, 1998.

[9] Kneser, Hellmuth Der Fundamentalsatz der Algebra und der Intuitionismus. *Math. Zeitschrift*, 46:287–302, 1940.

[10] Kneser, Martin Ergänzung zu einer Arbeit von Hellmuth Kneser über den Fundamentalsatz der Algebra. *Math. Zeitschrift*, 177:285–287, 1981.

[11] Landau, E. Sur quelques théorèmes de *M. Pétrovitch* relatifs aux zéros des fonctions analytiques. *Bull. Soc. Math. Franc.*, 33:251–261, 1905.

[12] Lehmer, D. H. A machine method for solving polynomial equations. *Journal of the ACM*, 8:151 –162, 1961.

[13] Mahler, K. An inequality for the discriminant of a polynomial. *Michigan Math. Journal*, 11:257–262, 1964.

[14] McMullen, C. Families of rational maps and iterative root-finding algorithms. *Annals of Mathematics*, 125:467–493, 1987.

[15] Neff, C. Andrew; John H. Reif. An efficient algorithm for the complex roots problem. *Journal of Complexity*, 12:81–115, 1996.

[16] Netto, E. *Die vier Gauss'schen Beweise für die Zerlegung ganzer algebraischer Funktionen in reelle Factoren ersten oder zweiten Grades (1799-1849)* , volume 14 of *OSTWALD's KLASSIKER DER EXAKTEN WISSENSCHAFTEN.* Akademische Verlagsgesellschaft m.b.H. in Leipzig; Verlag von Wilhelm Engelmann, Leipzig und Berlin, 3rd edition, 1913.

[17] Ostrowski, A. *Solution of Equations in Euclidean and Banach Spaces.* Academic Press, New York, 1973.

[18] Pan, V. Y. New techniques for approximating complex zeros. In *Proeedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, Arlington,Va.*, pages 260–270. ACM, 1994.

[19] Pan, V. Y. Solving a polynomial equation: Some history and recent progress. *SIAM Review*, 39(2):187–220, 1997.

[20] Pan, V. Y. On approximating complex polynomial zeros: Modified quadtree (Weyl's) construction and improved Newton's iteration. Technical Report 2894, INRIA, Sophia-Antipolis, 1996.

[21] Renegar, James On the efficiency of Newton's method in approximating all zeros of sytems of polynomials. *Mathematics of Operations Research*, 12:121–148, 1987.

[22] Renegar, James On the worst-case arithmetic complexity of approximating zeros of polynomials. *Journal of Complexity*, 13:90–113, 1987.

[23] Rheinboldt, W. C. On a theorem of S.Smale about Newton's method for analytic mappings. *Appl. Math. Lett.*, 1(1):69–72, 1988.

[24] Schönhage, A. The fundamental theorem of algebra in term of computational complexity. !Preliminary report, URL: http://www.informatik.uni-bonn.de/~schoe/fdthmrep.ps.gz, August 1982.

[25] Sendov, Bl. ; A. Andreev ; N. Kjurkchiev Numerical solution of polynomial equations. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis*, volume III, pages 625–778. Elsevier Science, Amsterdam, 1994.

[26] Shub, M. ; S. Smale Computational Complexity: On the Geometry of Polynomials and a Theory of Cost: II. *SIAM J. Computing*, 15(1):145 – 161, 1986.

[27] Shub, M. ; S. Smale Complexity of Bezout's Theorem. I: Geometric Aspects. *Journal of the AMS*, 6(2):459 – 501, 1993.

[28] Smale, Steve. The Fundamental Theorem of Algebra and Complexity Theory. *Bull. AMS (N.S.)*, 4(1):1 – 36, 1981.

[29] Smale, Steve. On the Efficiency of Algorithms of Analysis. *Bull. AMS (N.S.)*, 13(2):87 – 121, 1985.

[30] Tilli, P. Polynomial root finding by means of continuation. *Computing*, 59(4):307–324, 1997.

[31] Wang, Xinghua; G. Shen; D. Han. Some Remarks on Smale's " Algorithms for Solving Equations". *Acta Mathematica Sinica, New Series*, 8(4):337–348, 1992.

[32] Yap, C.K. *Fundamental Problems of Algorithmic Algebra*. Oxford University Press, New York, N.Y., 2000.