Annales Univ. Sci. Budapest., Sect. Comp. 10 (1989) 97-108

A PROXIMITY TEST FOR SECOND ORDER DIVISORS

ANNA KREBSZ

Eötvös Loránd University
Department of Numerical Analysis
Budapest, Bogdánfy u. 10/b, H-1117

Abstract "Proximity test method" for solving polynomial equations were introduced by Henrici in [1]. We extended the notion of test for two points and give a test which we calculate by dividing the given polynomial. With the algorithm we get a second degree divisor of the polynomial.

Introduction

Since the widespread application of high-speed computers, there has been a great affort for the construction of globaly convergent algorithms for determination of zeroes of polynomials. It follows from the nature of the problem, that most of these algorithms are based on complex arithmetic [1], [2], [3], even in the case, when the original polynomial is real.

In this paper we give an algorithm which uses only real arithmetic, and gives a second degree divisor for an arbitrary polynomial with real coefficients. The basis of our algorithm is the so called "proximity test method", introduced by P. Henrici, described in [1], but we extended the notation of test for two z_1, z_2 points. Here we give a test which we calculate by dividing the given polynomial, and we prove that the assumptions necessary for being a proximity test are satisfied. By the test we construct a couple of disks around z_1 and z_2 respectively, containing no zeros (see 1. Theorem) and another couple of disks containing at least one zero (see 2. Theorem). These disks form the basis of the search algorithm. The radii of the disks can be bounded by an arbitrary small prescribed number ϵ.

Notations and definitions

Let $P(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0$ be the given polynomial with real coefficients, let $\xi_1, \xi_2, \ldots, \xi_n$ be its zeros and let σ be a number such that $|\xi_i| < \sigma, i = 1, 2, \ldots, n$. Without loss of generality we may assume that $\sigma = 1$. Let D_0^+ denote the set $\{z \in C : |z| \leq 1, Im(z) \geq 0\}$ and D_0^- the set $\{z \in C : |z| \leq 1, Im(z) \leq 0\}$.

$$D_0 := \{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in C^2 : z_1 \in D_0^+, z_2 \in D_0^-$$

and

$$(z_1 = \bar{z}_2 \text{ or } z_1, z_2 \in R)\} \tag{1}$$

Let us divide $P(z)$ by $z^2 + \alpha z + \beta$, where α and β are real numbers, and express the polynomial as

$$\mathbf{P}(\mathbf{z}) = \mathbf{R}(\mathbf{z}, \alpha, \beta)(\mathbf{z}^2 + \alpha\mathbf{z} + \beta) + \mathbf{F}(\alpha, \beta)\mathbf{z} + \mathbf{G}(\alpha, \beta) \tag{2}$$

We approximate two roots of P by the roots of $z^2 + \alpha z + \beta$. In the algorithm we construct a sequence $\{(z_1^{(k)}, z_2^{(k)})^T\}$ and a sequence of disks $\{(D_k^+, D_k^-)\}$ such that

i. $z_1^{(k)}$ and $z_2^{(k)}$ are the roots of $z^2 + \alpha^{(k)}z + \beta^{(k)}$,

ii. $z_1^{(k)}$ is the center of the disk D_k^+ and

$z_2^{(k)}$ is the center of the disk D_k^-,

iii. $D_k^i \subset D_{k-1}^i$ for $i = +, -$ and every $k = 1, 2, \ldots$,

iv. the radii of D_k^+ and D_k^- converges to zero

v. $D_k^i\,(i = +,-)$ contains at least one root of P.

The basic notion of the algorithm is a test, depending on the polynomial P, on a parameter $\rho > 0$ and any points $(z_1, z_2)^T$ such that $(z_1, z_2)^T \in D_0$. We give the test by an inequality. The test gives YES - the inequality holds - at all points $(z_1, z_2)^T$ close enough to a pair of zeroes of P and gives NO - the inequality doesn't hold - at all points far enough from any zeroes. (There may be an in-between region in which the test is either YES or NO.) The parameter ρ regulates the difficulty of the test.

Formally, a

$$\mathbf{T}(\mathbf{P}, \rho, (\mathbf{z}_1, \mathbf{z}_2)^{\mathbf{T}}) : \mathbf{P}_n \mathbf{x}(0, \rho_0] \mathbf{x} \mathbf{D}_0 \rightarrow \{\text{YES}, \text{NO}\}$$

function is called test if there are two functions $\phi, \psi : (0, \rho_0] \rightarrow \mathbf{R}^+$, so that for an arbitrary pair z_1, z_2 the following statement hold

a) if there exist at least two roots of P, ξ_1, ξ_2, so that

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty \leq \Phi(\rho),$$

(the disk $\mid z_1 - w \mid \leq \Phi(\rho)$ and the disk $\mid z_2 - w \mid \leq \Phi(\rho)$ contains 1-1 roots of P,)

then

$$T(P, \rho, (z_1, z_2)^T) = \text{YES}$$

b) if for any pair of roots of P, ξ_1, ξ_2

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty > \psi(\rho),$$

(the disk $\mid z_1 - w \mid \leq \psi(\rho)$ and the disk $\mid z_2 - w \mid \leq \psi(\rho)$ contains no root of P,) then

$$T(P, \rho, (z_1, z_2)^T) = \text{NO}.$$

Here

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty = \max \ \{| z_1 - \xi_1 |, | z_2 - \xi_2 |\}.$$

c) $\lim_{\rho \to 0} \psi(\rho) = 0$ and

d) the function ψ is continuous, strictly monotonically increasing and its range contains the interval $(0,1]$.

Remark: a) and b) evidently imply, that $\psi(\rho) \geq \Phi(\rho)$. We do not require the inequality $\Phi = \psi$.

In the algorithm we shall use the test

$$\begin{aligned} \mathbf{T}(\mathbf{P}, \rho, (\mathbf{z}_1, \mathbf{z}_2)^{\mathbf{T}}) : &= (\max\{| \mathbf{F}(\alpha, \beta) |, | \mathbf{G}(\alpha, \beta) |\} \leq \rho) \\ \text{where } \ \alpha &= -(\mathbf{z}_1 + \mathbf{z}_2) \\ \text{and } \ \beta &= \mathbf{z}_1 \mathbf{z}_2 \qquad \text{real numbers.} \end{aligned} \tag{3}$$

To show that this test has the required properties consider the following theorems.

1. Theorem: Let $\epsilon > 0$ and $\mathbf{z}_1, \mathbf{z}_2$ be the roots of

$$\mathbf{z}^2 + \alpha \mathbf{z} + \beta.$$

If there exist two roots of $\mathbf{P}, \xi_1, \xi_2$, such that

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty \leq \epsilon$$

$$\text{then } | F(\alpha, \beta) | \leq (n-1) 2^{n-1} \epsilon$$

$$\text{and } | G(\alpha, \beta) | \leq n 2^{n-1} \epsilon$$

it follows that $\max\{|F(\alpha,\beta)|, |G(\alpha,\beta)|\} \leq n2^{n-1}\epsilon)$.

Proof: Substituting z_1, z_2 into (2) we get

$$P(z_1) = F \cdot z_1 + G$$
$$P(z_2) = F \cdot z_2 + G$$

If $z_1 \neq z_2$ $F = \frac{P(z_1)-P(z_2)}{z_1-z_2}$ and $G = P(z_1) - F \cdot z_2$. (4)

For the case $z_1 = z_2$ we get from (4) that

$$\lim_{(z_1-z_2)\to 0} F = \lim_{(z_1-z_2)\to 0} \frac{P(z_1 - P(z_2)}{z_1 - z_2} = P'(z_1).$$

The (4) expression can be used in the form

$$F = \frac{\prod_{i=1}^{n}(z_1 - \xi_i) - \prod_{i=1}^{n}(z_2 - \xi_i)}{z_1 - z_2} =$$

$$= (z_1 - \xi_1)(z_1 - \xi_2)\frac{\prod_{i=3}^{n}(z_1 - \xi_i) - \prod_{i=3}^{n}(z_2 - \xi_i)}{(z_1 - z_2)} +$$

$$+ \frac{((z_1 - \xi_1)(z_1 - \xi_2) - (z_2 - \xi_1)(z_2 - \xi_2))}{z_1 - z_2}\prod_{i=3}^{n}(z_2 - \xi_i)$$

We can estimate $|F|$ in the following way:

a)

$$\left|\frac{\prod_{i=3}^{n}(z_1 - \xi_i) - \prod_{i=3}^{n}(z_2 - \xi_i)}{z_1 - z_2}\right| \leq \sum_{i=1}^{n-2}\left|\frac{z_1^i - z_2^i}{z_1 - z_2}\right|\binom{n-2}{i} \leq$$

$$\leq \sum_{i=1}^{n-2}(|z_1|^{i-1} + |z_1|^{i-2}|z_2| + \ldots + |z_2|^{i-1})\binom{n-2}{i} \leq$$

$$\leq \sum_{i=1}^{n-2} i \binom{n-2}{i} < \sum_{i=1}^{n-2} (n-2) \binom{n-2}{i} < (n-2)2^{n-2}$$

b)

$$| z_1 - \xi_1 | \cdot | z_1 - \xi_2 | \leq \epsilon \cdot 2$$

c)

$$\left| \frac{(z_1 - \xi_1)(z_1 - \xi_2) - (z_2 - \xi_1)(z_2 - \xi_2)}{z_1 - z_2} \right| \leq$$

$$\leq | z_1 - \xi_1 + z_2 - \xi_2 | \leq | z_1 - \xi_1 | + | z_1 - \xi_2 | \leq 2\epsilon$$

d)

$$\left| \prod_{i=3}^{n} (z_2 - \xi_i) \right| \leq 2^{n-2}.$$

From a,b,c, and d,

$| F | \leq (n-2)2^{n-1}\epsilon + 2^{n-1}\epsilon = (n-1)2^{n-1}\epsilon$ and
$| G | \leq | P(z_1) | + | z_1 | \cdot | F | \leq 2^{n-1}\epsilon + (n-1)2^{n-1}\epsilon = n2^{n-1}\epsilon.$

1.1 Corollary: Let $\rho > 0$ and z_1, z_2 be the roots of $z^2 + \alpha z + \beta$. If

$$\max\{| F(\alpha, \beta) |, | G(\alpha, \beta) |\} > \rho$$

then for every roots of P, ξ_1, ξ_2

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty > \frac{\rho}{n2^{n-1}}$$

(i.e. no root of P is contained in the disks

$$| z_1 - w | \leq \frac{\rho}{n2^{n-1}} \quad \text{and} \quad | z_2 - w | \leq \frac{\rho}{n2^{n-1}}).$$

Hence the test defined by (3) satisfies the a) property of the test with

$$\mathbf{P}(\rho) = \frac{\rho}{\mathbf{n2^{n-1}}}.$$

For the determination of the ψ function consider the following lemma and theorem.

Lemma: Let H be a connected set in $\mathbf{C}$.

If

$$\sup_{z \in H} | P(z) | \leq c$$

then the diameter of H is less then $4\sqrt[n]{c}$.

Proof: Let d be the diameter of H. For any $\epsilon > 0$ there exist z_0, z^* from H such that $a := | z_0 - z^* | > d - \epsilon$. Using connectivity there must exist for every $0 \leq x \leq a$ a point $z \in H$ for which $z = z_0 + xe^{i\Phi}$. Now

$$| P(z) | = \prod_{k=1}^{n} | z_0 + xe^{i\Phi} - \xi_k | \geq \prod_{k=1}^{n} || xi^{i\varphi} | - | z_0 - \xi_k || \geq$$

$$\geq \prod_{k=1}^{n} | x - \alpha_k | \qquad \text{with} \quad \alpha_k = | z_0 - \xi_k | .$$

Thus $\sup_{z \in H} | P(z) | \geq \sup_{0 \leq x \leq a} | x - \alpha_1 | \cdot \ldots \cdot | x - \alpha_n | .$

Using the transformation $y = \frac{x}{a/2} - 1$ and $\beta_k = \frac{\alpha_k}{a/2} - 1$ we get

$$\sup_{z \in H} | P(z) | \geq \left[\frac{a}{2}\right]^n \sup_{-1 \leq y \leq 1} | y - \beta_1 | \cdot \ldots \cdot | y - \beta_n | .$$

The supremum on the right hand side has its smallest value for the Chebysev polynomial, therefore

$$\sup_{z\in H} | P(z) | \geq \left[\frac{a}{2}\right]^n \frac{1}{2^{n-1}} = \frac{(d-\epsilon)^n}{2^{2n-1}} > \frac{(d-\epsilon)^n}{4^n}$$

for any $\epsilon > 0$. $\epsilon \to 0$ implies

$$c \geq \sup_{z\in H} | P(z) | > \left[\frac{d}{4}\right]^n.$$

which is equivalent to $d < 4\sqrt[n]{c}$.

2. Theorem: Let $0 < \epsilon < 1/3$ and
$\tilde{\mathbf{P}}(\mathbf{z}) = \mathbf{R}(\mathbf{z}, \alpha, \beta)(\mathbf{z}^2 + \alpha\mathbf{z} + +\beta)$. If

$$\max\{| \mathbf{F}(\alpha,\beta) |, | \mathbf{G}(\alpha,\beta) |\} \leq \epsilon$$

then the distance of the roots of P and $\tilde{\mathbf{P}}$ is less then $4\sqrt[n]{3\epsilon}$, i.e.

$$\min_{\pi \text{ perm}} \max_{i=1}^{n} | \xi_i - \mathbf{z}_{\pi(i)} | < 4\sqrt[n]{3\epsilon}).$$

where $\mathbf{z}_k$ denotes the roots of $\tilde{\mathbf{P}}$).

Proof: Let $q(z) := \tilde{P}(z) - P(z) = -Fz + G$ (i.e. $P + q = \tilde{P}$). Let us define the following set

$$H := \{z \in \mathbf{C} : | P(z) | \leq 3\epsilon\}.$$

Let us denote the components of H by $H_1, H_2, \ldots, H_s$. We shall compare the roots of P and $\tilde{P}$ in H_i. H_i is a connected set in $\mathbf{C}$ and $\sup_{z\in H_i} | P(z) | \leq 3\epsilon$, thus it follows from the lemma that the diameter of H_i is less then $4\sqrt[n]{3\epsilon}$. Let Γ_i be the boundary of H_i. For $z \in \Gamma_i$

1. $| P(z) |= 3\epsilon$,
2. $| z |< 2$ (since $| z |\geq 2$ would imply $3\epsilon =| P(z) |= | z - \xi_1 | \ldots | z - \xi_n |\geq 1^n = 1$ which is a contradiction)
3. $| q(z) |< 3\epsilon$.

Since P has no root in Γ_i and for $z \in \Gamma_i$ $| q(z) |<| P(z) |$ it follows from Rouché's theorem (see [4]) that the number of roots of $\tilde{P} = P + q$ in H_i is the same as that of P. The distance of the roots are less then $4\sqrt[n]{3\epsilon}$. Thus the theorem holds.

2.1. Corollary:

Let $0 < \epsilon < 1/3$ and let z_1, z_2 be the roots of $z^2 + \alpha z + \beta$. If

$$\max\{| F(\alpha,\beta) |,| G(\alpha,\beta) |\} \leq \epsilon$$

then there exist two roots of P, ξ_1, ξ_2, such that

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty < 4\sqrt[n]{3\epsilon}.$$

2.2. Corollary:

Let $0 < \epsilon < 1/3$ and let z_1, z_2 be the roots of $z^2 + \alpha z + \beta$. if for every pair of roots of P, ξ_1, ξ_2,

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_\infty > 4\sqrt[n]{3\epsilon}.$$

then $\max\{| F(\alpha,\beta) |,| G(\alpha,\beta) |\} > \epsilon$.

The last corollary means that the test defined by (3) satisfies b) with

$\psi(\rho) = 4\sqrt[n]{3\rho}$.

It is clear that ψ also satisfies c) and d).

The Search Algorithm

Let T be the test defined in (3) and let $\{\tau_k\}$ be a monotonous sequence of positive numbers converging to zero such that $\tau_0 = 1$. We shall describe an algorithm for constructing a sequence of points $\{(z_1^{(k)}, z_2^{(k)})^T\}$ such that

1. $(z_1^{(k)}, z_2^{(k)})^T \in D_0$
2. each of the disks
   $D_k^+ := \{z : | z - z_1^{(k)} | \leq \tau_k\} \cap D_0^+$ and
   $D_k^- := \{z : | z - z_2^{(k)} | \leq \tau_k\} \cap D_0^-$
   contains at least one root of P for every $k = 0, 1, \ldots$

Let $(z_1^{(0)}, z_2^{(0)})^T = (0,0)^T$. Then both D_0^+ and D_0^- contains a root. The algorithm now proceeds by induction. Suppose we have found $(z_1^{(k-1)}, z_2^{(k-1)})^T$ such that D_{k-1}^+ contains a root and D_{k-1}^- also. To construct the k-th approximation we cover the set D_{k-1}^+ with closed disks of radius $\leq \epsilon_k$ (which will be given later). The covering disk centers of D_{k-1}^- we get from the centers of D_{k-1}^+ with conjugation. We apply the T test with $\rho = \rho_k$ at each elements of the following set, which contains pairs of covering disk centers:

$$\Omega := \left\{ (z_1, z_2)^T : \begin{array}{l} z_1 \text{ covering disk center in } D_{k-1}^+, \\ z_2 \text{ covering disk center in } D_{k-1}^- \\ \text{and } (z_2 = \bar{z}_1 \text{ or } z_1, z_2 \in \mathbf{R}) \end{array} \right\} \quad i.e.\ \Omega \subset D_0.$$

The parameters ϵ_k and ρ_k are chosen such that the following two conditions are met:

i) If both of the covering disks with center z_1 and z_2, $(z_1, z_2)^T \in \Omega$ contains a root of P, then the test gives YES at $(z_1, z_2)^T$.

ii) If the test gives YES at $(z_1, z_2)^T \in \Omega$, then both of the disks $| z - z_1 | \leq \tau_k$ and $| z - z_2 | \leq \tau_k$ contains a root of P.

Condition i) is satisfied if $\epsilon_k \leq \Phi(\rho_k)$ and condition ii) is satisfied if $\psi(\rho_k) \leq \tau_k$. Thus both conditions are fulfilled if

$$\rho_k = \psi^{-1}(\tau_k) \quad \text{and}$$
$$\epsilon_k = \Phi(\rho_k) = \Phi(\psi^{-1}(\tau_k))$$

where ψ^{-1} denotes the inverse function of ψ.

At least one pair of the covering disk (with centers from Ω) contains a root-pair of P, since both D^+_{k-1} and D^-_{k-1} contains one, and all are contained in $D^+_0 \cup D^-_0$. Thus by i) the T test gives YES with $\rho = \rho_k$ and least one center-pair from Ω. We choose $(z_1^{(k)}, z_2^{(k)})^T$ to be the first pair the test gives YES at. Actually there is no guarantee for each disk of radius ϵ_k surrounding $z_1^{(k)}$ and $z_2^{(k)}$ to contain a root but by ii) D^+_k and D^-_k must contain at least one.

References

[1] **Henrici, P.:** Applied and Computational Complex Analysis. Vol. 1. (Wiley, New York. 1974.) 500-513.

[2] **Henrici, P.:** Topics in computational complex analysis: I. Methods of descent polynomial equations in: Computational Aspects of Complex Analysis (D. Reidel Publishing Company, 1983.) 133-147.

[3] **Murota, K.:** Global convergence of a modified Newton iteration for algebraic equations SIAM J. Numer. Anal. 19. (1982.) 793-799.

[4] **Hurwitz-Courant:** Funktionentheorie, Springer, Berlin, 1929. (Grundlehren der Mathematischen Wissenschaften)

Received: July 27, 1989.