Japan J. Indust. Appl. Math., 24 (2007), 219–226 Area ⟨2⟩

Tighter Bounds of Errors of Numerical Roots

Tateaki Sasaki

Institute of Mathematics,
University of Tsukuba,
Tsukuba-shi, Ibaraki 305, Japan
E-mail: sasaki@math.tsukuba.ac.jp

Received September 1, 2006
Revised February 14, 2007

Let $P(z)$ be a monic univariate polynomial over $\mathbf{C}$, of degree n and having roots $\zeta_1, \ldots, \zeta_n$. Given approximate roots $z_1, \ldots, z_n$, with $\zeta_i \simeq z_i$ $(i = 1, \ldots, n)$, we derive a very tight upper bound of $|\zeta_i - z_i|$, by assuming that ζ_i has no close root. The bound formula has a similarity with Smale's and Smith's formulas. We also derive a lower bound of $|\zeta_i - z_i|$ and a lower bound of $\min\{|\zeta_j - z_i| \mid j \neq i\}$.

1. Introduction

Let $P(z)$ be a monic univariate polynomial over $\mathbf{C}$, of degree n, having roots $\zeta_1, \ldots, \zeta_n$, and let $z_1, \ldots, z_n$ be its approximate roots determined numerically. There are many researches on error bounds for $z_1, \ldots, z_n$. Kantrovich studied Newton's method for differentiable function $F(z)$ in 1948 and determined a complex disc D_{K} which contains only one root of $F(z)$ (see [4] or [5]). The D_{K} is defined with several constants which are determined by data over a region. For old results, see [5]. In 1970, Smith presented a very useful formula [7]. Let D_i $(1 \leq i \leq n)$ be a disc in the complex plane, with the center at $z = z_i$ and the radius $r_i = n\big|P(z_i)/\prod_{j\neq i}(z_i - z_j)\big|$. Smith's theorem asserts that the union $D_1 \cup \cdots \cup D_n$ contains all the n roots of $P(z)$, and in particular if $D = D_1 \cup \cdots \cup D_m$ is disconnected with $D_{m+1} \cup \cdots \cup D_n$ then D contains exactly m roots. Smith's discs are closely related with Durand-Kerner's method for computing all the roots simultaneously; Durand-Kerner's iteration formula is $z_i := z_i - P(z_i)/\prod_{j\neq i}(z_i - z_j)$. Hence, Smith's formula over-estimates the errors by about n. Since then, many authors improved Smith's formulas, see [2] for example, but this over-estimation is not resolved. We call such formulas Smith-type formulas.

In 1980's, Smale studied Newton's method for analytic function $A(z)$ and determined a disc D_{S} centered at $z = z_i$, which contains only one root of $A(z)$ [6, 1]. Smale developed his theory by using an auxiliary rational function with the following constants (below, $A^{(k)}$ denotes the kth derivative of A).

$$a_1 \stackrel{\text{def}}{=} |A^{(1)}(z_i)|, \quad a \stackrel{\text{def}}{=} \max\left\{ \left| \frac{A^{(k)}(z_i)}{k!a_1} \right|^{1/(k-1)} \,\middle|\, k = 2, 3, \ldots \right\}. \tag{1.1}$$

Smale's error bound is very tight. Since then, many authors derived similar formulas

Work supported in part by Japanese Ministry of Education, Science and Culture under Grants 15300002.

which we call Smale-type formulas. The tightest bound now is obtained by Wang and Han [9]; the same bound was obtained by Inaba and Sasaki from a simple different approach [3].

Then, the following question naturally arises: is there a formula which bounds the errors as accurately as Smale-type formulas by using the numerical roots $z_1, \dots, z_n$ simultaneously? In this short article, we derive such a formula; the idea is simple and the derivation is elementary. The error bound obtained is quite tight and has a similarity with both Smale's and Smith's ones. Using the same idea, we also derive a lower bound of the error $|\zeta_i - z_i|$ and a lower bound of $\min\{|\zeta_j - z_i| \mid j \neq i\}$. These bounds are also quite tight.

2. An inequality for the error

We assume that the numerical roots $z_1, \dots, z_n$ satisfy $z_j \simeq \zeta_j$ $(j = 1, \dots, n)$. We express $P(z)$ as

$$\begin{cases} P(z) = \tilde{P}(z) + \Delta(z), \ \tilde{P}(z) = (z - z_1)\cdots(z - z_n), \\ \deg(\Delta) \le n - 1, \qquad \dfrac{\|\Delta\|}{\|\tilde{P}\|} \ll 1. \end{cases} \tag{2.1}$$

We assume further that ζ_i has no close root. By $A^{(k)}(z)$ we denote $\mathrm{d}^k A(z)/\mathrm{d}z^k$.

Put $\zeta_i = z_i + \varepsilon$, then we have

$$\tilde{P}(z_i + \varepsilon) = \frac{\tilde{P}^{(1)}(z_i)\varepsilon}{1!} + \frac{\tilde{P}^{(2)}(z_i)\varepsilon^2}{2!} + \cdots + \frac{\tilde{P}^{(n)}(z_i)\varepsilon^n}{n!}.$$

Define positive numbers $\tilde{p}_1$ and $\tilde{p}$ as follows.

$$\tilde{p}_1 \stackrel{\text{def}}{=} |\tilde{P}^{(1)}(z_i)|, \quad \tilde{p} \stackrel{\text{def}}{=} \max\left\{ \left| \frac{\tilde{P}^{(2)}(z_i)}{2!\,\tilde{p}_1} \right|^{1/1}, \dots, \left| \frac{\tilde{P}^{(n)}(z_i)}{n!\,\tilde{p}_1} \right|^{1/(n-1)} \right\}. \tag{2.2}$$

Note that $\tilde{p}_1 \neq 0$ because we have assumed that ζ_i has no close root. Assuming $2\tilde{p}|\varepsilon| < 1$, we bound $|\tilde{P}(z_i + \varepsilon)|$ as follows.

$$\begin{aligned} |\tilde{P}(z_i + \varepsilon)| &\ge \tilde{p}_1|\varepsilon| \cdot \left\{ 1 - \frac{|\varepsilon \tilde{P}^{(2)}(z_i)|}{2!\,\tilde{p}_1} - \cdots - \frac{|\varepsilon^{n-1}\tilde{P}^{(n)}(z_i)|}{n!\,\tilde{p}_1} \right\} \\ &\ge \tilde{p}_1|\varepsilon| \cdot \{1 - \tilde{p}|\varepsilon| - \cdots - \tilde{p}^{n-1}|\varepsilon|^{n-1}\} > \tilde{p}_1|\varepsilon| \frac{1 - 2\tilde{p}|\varepsilon|}{1 - \tilde{p}|\varepsilon|}. \end{aligned}$$

Since $P(z_i + \varepsilon) = 0 = \tilde{P}(z_i + \varepsilon) + \Delta(z_i + \varepsilon)$, we obtain the following inequality.

$$\tilde{p}_1|\varepsilon| \cdot \frac{1 - 2\tilde{p}|\varepsilon|}{1 - \tilde{p}|\varepsilon|} \ < \ |\Delta(z_i + \varepsilon)|. \tag{2.3}$$

This is our basic inequality. In the next section, we will obtain a polynomial inequality by bounding $|\Delta(z_i + \varepsilon)|$ suitably.

3. A tighter upper bound of $|\varepsilon|$

Define positive numbers δ_0 and d_0 as follows.

$$\delta_0 \stackrel{\text{def}}{=} |\Delta(z_i)|, \quad d_0 \stackrel{\text{def}}{=} \max\left\{\left|\frac{\Delta^{(1)}(z_i)}{1!\,\delta_0}\right|^{1/1}, \ldots, \left|\frac{\Delta^{(n-1)}(z_i)}{(n-1)!\,\delta_0}\right|^{1/(n-1)}\right\}. \tag{3.1}$$

We assume that $\delta_0 \neq 0$ because if $\delta_0 = 0$ then $z_i = \zeta_i$ hence $\varepsilon = 0$. However, $|\delta_0|$ may be very small hence d_0 may become very large. Therefore, we consider the following three cases.

Case 1: When $|\Delta(z_i)| = O(\|\Delta\|)$. Assuming $d_0|\varepsilon| < 1$, we bound $|\Delta(z_i+\varepsilon)|$ as follows.

$$\begin{aligned}|\Delta(z_i+\varepsilon)| &\le \delta_0 \cdot \left\{1 + \frac{|\varepsilon\Delta^{(1)}(z_i)|}{1!\,\delta_0} + \cdots + \frac{|\varepsilon^{n-1}\Delta^{(n-1)}(z_i)|}{(n-1)!\,\delta_0}\right\} \\ &\le \delta_0 \cdot \{1 + d_0|\varepsilon| + \cdots + (d_0|\varepsilon|)^{n-1}\} < \frac{\delta_0}{1 - d_0|\varepsilon|}.\end{aligned}$$

Substituting this into (2.3) and clearing the denominators, we obtain

$$2\tilde{p}_1\tilde{p}d_0|\varepsilon|^3 < \delta_0 - (\tilde{p}_1 + \tilde{p}\delta_0)|\varepsilon| + \tilde{p}_1(2\tilde{p} + d_0)|\varepsilon|^2. \tag{3.2}$$

Since $|\varepsilon|$ is assumed to be small, the above $|\varepsilon|^3$-term will be very small, so we discard it. Then, since $|\varepsilon|$ must go to zero as $\|\Delta\| \to 0$, we obtain the following upper bound of $|\varepsilon|$; note that, since the $|\varepsilon|^3$-term is positive, the discarding of $|\varepsilon|^3$-term loosens the bound by a little.

$$|\varepsilon| < r_1' \stackrel{\text{def}}{=} \frac{(\tilde{p}_1 + \tilde{p}\delta_0) - \sqrt{(\tilde{p}_1 + \tilde{p}\delta_0)^2 - 4\tilde{p}_1(2\tilde{p} + d_0)\delta_0}}{2\tilde{p}_1(2\tilde{p} + d_0)}. \tag{3.3}$$

Using inequality $\sqrt{1-x} > 1 - x/2 - x^2/2$, which is valid for $0 < x < 1$, and assuming $\tilde{p}_1 \gg \tilde{p}\delta_0$, we bound r_1' as follows.

$$r_1' < \frac{\delta_0}{\tilde{p}_1 + \tilde{p}\delta_0}\left[1 + \frac{\tilde{p}_1(8\tilde{p} + 4d_0)\delta_0}{(\tilde{p}_1 + \tilde{p}\delta_0)^2}\right] < r_1 \stackrel{\text{def}}{=} \frac{\delta_0}{\tilde{p}_1 + \tilde{p}\delta_0}\left[1 + \frac{(8\tilde{p} + 4d_0)\delta_0}{\tilde{p}_1}\right]. \tag{3.4}$$

We note that Durand-Kerner's correction term is $\delta_0/\tilde{p}_1$ and Durand-Kerner's iteration converges quadratically. This means that $|\varepsilon| = \delta_0/\tilde{p}_1 + O(|\varepsilon|^2)$ if $\delta_0 \ll 1$, which is completely consistent with (3.4). Therefore, r_1 is a very tight bound of $|\varepsilon|$ practically. In the next section, we will derive a lower bound of $|\varepsilon|$, too, which will also show that the above bound is very tight.

In deriving (3.3) and (3.4), we set the following conditions.

$$2\tilde{p}|\varepsilon| < 1, \quad d_0|\varepsilon| < 1, \quad (\tilde{p}_1 + \tilde{p}\delta_0)^2 - 4\tilde{p}_1(2\tilde{p} + d_0)\delta_0 > 0, \quad \tilde{p}_1 \gg \tilde{p}\delta_0. \tag{3.5}$$

By assumption, δ_0 is small. Hence, in the above third condition, we discard small term $\tilde{p}^2\delta_0^2$ and replace the condition by a stronger one $\tilde{p}_1^2 - \tilde{p}_1(6\tilde{p} + 4d_0)\delta_0 > 0$, or

Condition-1: $\tilde{p}_1 > (6\tilde{p} + 4d_0)\delta_0$. The fourth condition is satisfied by Condition-1. Using inequality $\sqrt{1-x} > 1 - x$ $(0 < x < 1)$, we find $r'_1 < 2\delta_0/(\tilde{p}_1 + \tilde{p}\delta_0)$. This inequality and Condition-1 give $r'_1 < 2/(7\tilde{p} + 4d_0)$, hence the first and the second conditions are satisfied.

Case 2: When $|\Delta(z_i)| \ll |\Delta^{(1)}(z_i)| = O(\|\Delta\|)$; this case occurs rarely. Define positive numbers δ_1 and d_1 as follows.

$$\delta_1 \stackrel{\text{def}}{=} |\Delta^{(1)}(z_i)|, \quad d_1 \stackrel{\text{def}}{=} \max\left\{ \left| \frac{\Delta^{(2)}(z_i)}{2!\,\delta_1} \right|^{1/1}, \ldots, \left| \frac{\Delta^{(n-1)}(z_i)}{(n-1)!\,\delta_1} \right|^{1/(n-2)} \right\}. \tag{3.6}$$

Assuming $d_1|\varepsilon| < 1$, we bound $|\Delta(z_i + \varepsilon)|$ as follows.

$$|\Delta(z_i + \varepsilon)| \le \delta_0 + \delta_1|\varepsilon| \cdot \{1 + d_1|\varepsilon| + \cdots + (d_1|\varepsilon|)^{n-2}\} < \delta_0 + \delta_1 \frac{|\varepsilon|}{1 - d_1|\varepsilon|}.$$

Substituting this into (2.3) and clearing denominators, we obtain

$$\begin{aligned} 2\tilde{p}_1\tilde{p}d_1|\varepsilon|^3 &< \delta_0 - (\tilde{p}_1 + \tilde{p}\delta_0 + d_1\delta_0 - \delta_1)|\varepsilon| \\ &\quad + \{\tilde{p}_1(2\tilde{p} + d_1) + \tilde{p}(d_1\delta_0 - \delta_1)\}|\varepsilon|^2. \end{aligned} \tag{3.7}$$

Discarding the small $|\varepsilon|^3$-term as in Case 1, which loosens the upper bound by a little, we obtain the following upper bound of $|\varepsilon|$.

$$|\varepsilon| \;<\; r'_2 \stackrel{\text{def}}{=} \frac{B - \sqrt{B^2 - 4A\delta_0}}{2A}, \qquad \begin{cases} A = \tilde{p}_1(2\tilde{p} + d_1) + \tilde{p}(d_1\delta_0 - \delta_1), \\ B = \tilde{p}_1 + \tilde{p}\delta_0 + d_1\delta_0 - \delta_1. \end{cases} \tag{3.8}$$

Here, we assumed that $A > 0$ and $B > 0$. Using $\sqrt{1-x} > 1 - x/2 - x^2/2$ $(0 < x < 1)$, and assuming $\tilde{p}_1 - \delta_1 \gg (\tilde{p} + d_1)\delta_0$, we bound r'_2 as follows.

$$r'_2 < \frac{\delta_0}{B}\left(1 + \frac{4A\delta_0}{B^2}\right) < r_2 \stackrel{\text{def}}{=} \frac{\delta_0}{\tilde{p}_1 - \delta_1}\left[1 + \frac{4A\delta_0}{(\tilde{p}_1 - \delta_1)^2}\right]. \tag{3.9}$$

In deriving (3.8) and (3.9), we set the following conditions as well as $A > 0$ and $B > 0$.

$$2\tilde{p}|\varepsilon| < 1, \quad d_1|\varepsilon| < 1, \quad B^2 - 4A\delta_0 > 0, \quad \tilde{p}_1 - \delta_1 \gg (\tilde{p} + d_1)\delta_0. \tag{3.10}$$

We have $B^2 - 4A\delta_0 = \tilde{p}_1\{\tilde{p}_1 - (6\tilde{p} + 2d_1)\delta_0 - 2\delta_1\} + (\tilde{p}\delta_0 - d_1\delta_0 + \delta_1)^2$. Discarding the small term $(\tilde{p}\delta_0 - d_1\delta_0 + \delta_1)^2$, we replace the above third condition by a stronger one, Condition-2: $\tilde{p}_1 > (6\tilde{p} + 2d_1)\delta_0 + 2\delta_1$. Inequalities $A > 0$, $B > 0$ and the fourth condition are satisfied by Condition-2. Using $\sqrt{1-x} > 1 - x$ $(0 < x < 1)$, we find $r'_2 < 2\delta_0/\{\tilde{p}_1 + (\tilde{p} + d_1)\delta_0 - \delta_1\}$. This inequality and Condition-2 give $r'_2 < 2\delta_0/\{(7\tilde{p} + 3d_1)\delta_0 + \delta_1\} < 2/(7\tilde{p} + 3d_1)$. Hence, the first and the second conditions in (3.10) are satisfied.

Case 3: When $|\Delta(z_i)|, |\Delta^{(1)}(z_i)| \ll \|\Delta\|$; this case occurs very very rarely.

There exists an integer m, $2 \le m \le n-1$, such that $|\Delta^{(j)}(z_i)| = O(\delta_0)$ for every $j \le m-1$ and $|\Delta^{(m)}(z_i)/m!| \gg \delta_0$. Define positive numbers δ_m and d_m as follows.

$$\delta_m \stackrel{\text{def}}{=} |\Delta^{(m)}(z_i)|, \quad d_m \stackrel{\text{def}}{=} \max\left\{ \left|\frac{\Delta^{(2)}(z_i)}{2!\,\delta_m}\right|^{1/1}, \ldots, \left|\frac{\Delta^{(n-1)}(z_i)}{(n-1)!\,\delta_m}\right|^{1/(n-2)} \right\}. \quad (3.11)$$

Assuming $d_m|\varepsilon| < 1$, we bound $|\Delta(z_i+\varepsilon)|$ as follows.

$$|\Delta(z_i+\varepsilon)| < \delta_0 + \delta_m|\varepsilon| \cdot \{1 + d_m|\varepsilon| + \cdots + (d_m|\varepsilon|)^{n-2}\} < \delta_0 + \frac{\delta_m|\varepsilon|}{1-d_m|\varepsilon|}.$$

Substituting this into (2.3) and clearing denominators, we obtain

$$2\tilde{p}_1\tilde{p}d_m|\varepsilon|^3 < \delta_0 - (\tilde{p}_1+\tilde{p}\delta_0+d_m\delta_0-\delta_m)|\varepsilon| + \{\tilde{p}_1(2\tilde{p}+d_m)+\tilde{p}(d_m\delta_0-\delta_m)\}|\varepsilon|^2. \quad (3.12)$$

Then, as in Case 2, we obtain the following bounds.

$$|\varepsilon| \;<\; r_3' \stackrel{\text{def}}{=} \frac{B-\sqrt{B^2-4A\delta_0}}{2A}, \qquad \begin{cases} A = \tilde{p}_1(2\tilde{p}+d_m)+\tilde{p}(d_m\delta_0-\delta_m), \\ B = \tilde{p}_1+\tilde{p}\delta_0+d_m\delta_0-\delta_m, \end{cases} \quad (3.13)$$

$$r_3' \;<\; \frac{\delta_0}{B}\left(1+\frac{4A\delta_0}{B^2}\right) \;<\; r_3 \stackrel{\text{def}}{=} \frac{\delta_0}{\tilde{p}_1-\delta_m}\left[1+\frac{4A\delta_0}{(\tilde{p}_1-\delta_m)^2}\right]. \quad (3.14)$$

In deriving (3.13) and (3.14), we set the following conditions as well as $A > 0$ and $B > 0$.

$$2\tilde{p}|\varepsilon| < 1, \quad d_m|\varepsilon| < 1, \quad B^2-4A\delta_0 > 0, \quad \tilde{p}_1-\delta_m \gg (\tilde{p}+d_m)\delta_0. \quad (3.15)$$

As in Case 2, these conditions are satisfied by Condition-3: $\tilde{p}_1 > (6\tilde{p}+2d_m)\delta_0+2\delta_m$.

Let us compare the bounds in (3.4), (3.9) and (3.14). We have $\|\Delta\| \ll \|P\|$ practically, hence $\delta_0 \ll 1$. Then,

$$r_1' \simeq \frac{\delta_0}{\tilde{p}_1+\tilde{p}\delta_0}, \quad r_2' \simeq \frac{\delta_0}{\tilde{p}_1+\tilde{p}\delta_0+d_1\delta_0-\delta_1}, \quad r_3' \simeq \frac{\delta_0}{\tilde{p}_1+\tilde{p}\delta_0+d_m\delta_0-\delta_m}.$$

We see $r_2' \lesssim r_1'$ only when $d_1\delta_0 - \delta_1 > 0$. On the other hand, r_2' was derived by assuming $\delta_0 \ll \delta_1$. Therefore, bounds in (3.9) and (3.14) are scarcely useful in practice, so we employ only the bound in (3.4).

Theorem 1. *Let $\tilde{p}_1$, $\tilde{p}$, etc. be defined as above. If $\tilde{p}_1 > (6\tilde{p}+4d_0)\delta_0$ then the root ζ_i of $P(z)$ is contained in the circle of radius r, located at $z = z_i$, where*

$$r = \frac{\delta_0}{\tilde{p}_1+\tilde{p}\delta_0}\left[1+\frac{(8\tilde{p}+4d_0)\delta_0}{\tilde{p}_1}\right].$$

4. A tighter lower bound of $|\varepsilon|$

Bounding $|\tilde{P}(z_i+\varepsilon)|$ and $|\Delta(z_i+\varepsilon)|$ oppositely, we obtain a lower bound of $|\varepsilon|$. Assuming $\tilde{p}|\varepsilon|<1$, we bound $|\tilde{P}(z_i+\varepsilon)|$ as follows.

$$|\tilde{P}(z_i+\varepsilon)| \le \tilde{p}_1|\varepsilon|\cdot\{1+\tilde{p}|\varepsilon|+\cdots+\tilde{p}^{n-1}|\varepsilon|^{n-1}\} < \tilde{p}_1\frac{|\varepsilon|}{1-\tilde{p}|\varepsilon|}. \tag{4.1}$$

Assuming $2d_0|\varepsilon|<1$, we bound $|\Delta(z_i+\varepsilon)|$ as follows.

$$|\Delta(z_i+\varepsilon)| \ge \delta_0\cdot\{1-d_0|\varepsilon|-\cdots-d_0^{n-1}|\varepsilon|^{n-1}\} > \delta_0\frac{1-2d_0|\varepsilon|}{1-d_0|\varepsilon|}. \tag{4.2}$$

Bounds (4.1) and (4.2) give inequality $\tilde{p}_1|\varepsilon|/(1-\tilde{p}|\varepsilon|) > \delta_0(1-2d_0|\varepsilon|)/(1-d_0|\varepsilon|)$, or

$$(\tilde{p}_1+2\tilde{p}\delta_0)d_0|\varepsilon|^2-(\tilde{p}_1+\tilde{p}\delta_0+2d_0\delta_0)|\varepsilon|+\delta_0<0. \tag{4.3}$$

This inequality and $\sqrt{1-x} < 1-x/2$ $(0<x<1)$ give the following lower bound of $|\varepsilon|$.

$$|\varepsilon| > \check{r} \stackrel{\text{def}}{=} \frac{\delta_0}{\tilde{p}_1+\tilde{p}\delta_0+2d_0\delta_0}. \tag{4.4}$$

Combining (3.4) with (4.4), we have $\check{r}<|\varepsilon|<r_1$ and

$$r_1-\check{r} \simeq \frac{8\tilde{p}+6d_0}{(\tilde{p}_1+\tilde{p}\delta_0)^2}\delta_0^2 \quad \text{if} \quad \delta_0 \ll 1. \tag{4.5}$$

Therefore, formulas (3.4) and (4.4) bound $|\varepsilon|$ in a very narrow domain in practice.

In deriving (4.4), we set the following conditions.

$$\tilde{p}|\varepsilon|<1, \quad 2d_0|\varepsilon|<1, \quad (\tilde{p}_1+\tilde{p}\delta_0+2d_0\delta_0)^2-4(\tilde{p}_1+2\tilde{p}\delta_0)d_0\delta_0>0.$$

The first and the second conditions are satisfied by Condition-1 given in the previous section. The third condition can be rewritten as $\tilde{p}_1(\tilde{p}_1+2\tilde{p}\delta_0)+(\tilde{p}-2d_0)^2\delta_0^2>0$, hence it holds always. Thus, the above lower bound is valid under Condition-1.

THEOREM 2. *Let $\tilde{p}_1$, $\tilde{p}$, etc. be defined as in **3**. If $\tilde{p}_1>(6\tilde{p}+4d_0)\delta_0$ then the root ζ_i of $P(z)$ is outside of the circle of radius $\delta_0/(\tilde{p}_1+\tilde{p}\delta_0+2d_0\delta_0)$, located at $z=z_i$.*

REMARK. In Cases 2 and 3, we can bound $|\Delta(z_i+\varepsilon)|$ as

$$|\Delta(z_i+\varepsilon)| \ge \delta_0-\delta_1|\varepsilon|\cdot\{1+d_1|\varepsilon|+\cdots+(d_1|\varepsilon|)^{n-2}\} > \delta_0-\delta_1\frac{|\varepsilon|}{1-d_1|\varepsilon|},$$

$$|\Delta(z_i+\varepsilon)| \ge \delta_0-\delta_m|\varepsilon|\cdot\{1+d_m|\varepsilon|+\cdots+(d_m|\varepsilon|)^{n-2}\} > \delta_0-\delta_m\frac{|\varepsilon|}{1-d_m|\varepsilon|}.$$

Then, we obtain the following lower bounds, respectively.

$$|\varepsilon| > \frac{\delta_0}{\tilde{p}_1 + \tilde{p}\delta_0 + d_1\delta_0 + \delta_1}, \quad |\varepsilon| > \frac{\delta_0}{\tilde{p}_1 + \tilde{p}\delta_0 + d_m\delta_0 + \delta_m}.$$

Once again, these lower bounds are not much different from $\check{r}$ in (4.4).

5. A lower bound of the distance to other roots

Inequality in (3.2), with the $|\varepsilon|^3$-term discarded, gives the following lower bound of $|\varepsilon|$; note that the discarding of $|\varepsilon|^3$-term loosens the lower bound by a little.

$$|\varepsilon| > R' \stackrel{\text{def}}{=} \frac{(\tilde{p}_1 + \tilde{p}\delta_0) + \sqrt{(\tilde{p}_1 + \tilde{p}\delta_0)^2 - 4\tilde{p}_1(2\tilde{p} + d_0)\delta_0}}{2\tilde{p}_1(2\tilde{p} + d_0)}. \tag{5.1}$$

In deriving the above bound, we assumed the first three conditions in (3.5) which are satisfied by Condition-1 given in **3**. In addition to the first two conditions in (3.5), we require

$$2\tilde{p}R' < 1 \quad \text{and} \quad d_0R' < 1. \tag{5.2}$$

Since $r'_1 + R' = (\tilde{p}_1 + \tilde{p}\delta_0)/\tilde{p}_1(2\tilde{p} + d_0)$, we bound R' as $R' < (\tilde{p}_1 + \tilde{p}\delta_0)/\tilde{p}_1(2\tilde{p} + d_0)$. Then,

$$2\tilde{p}R' < \frac{1 + \tilde{p}\delta_0/\tilde{p}_1}{1 + d_0/2\tilde{p}}, \quad d_0R' < \frac{1 + \tilde{p}\delta_0/\tilde{p}_1}{1 + 2\tilde{p}/d_0}.$$

Hence, the inequalities in (5.2) are satisfied so long as $\tilde{p}\delta_0/\tilde{p}_1 < d_0/2\tilde{p}$ and $\tilde{p}\delta_0/\tilde{p}_1 < 2\tilde{p}/d_0$, respectively, or $\tilde{p}_1 > (2\tilde{p}^2/d_0)\delta_0$ and $\tilde{p}_1 > d_0\delta_0/2$. Combining these conditions with Condition-1, we obtain $\tilde{p}_1 > \max\{6\tilde{p} + 4d_0, 2\tilde{p}^2/d_0\} \times \delta_0$.

We next show that the bound in (5.1) is a lower bound of the distance $|\zeta_j - z_i|$ ($\forall j \neq i$). Suppose the origin is moved to $z = z_i$, and put

$$\tilde{P}(z + z_i) \stackrel{\text{def}}{=} \tilde{p}_1 z \cdot \tilde{Q}(z) = \tilde{p}_1 z \cdot (1 + q_1 z + \cdots + q_{n-1}z^{n-1}).$$

We see $\tilde{p}_1 = \tilde{P}^{(1)}(z_i)$ and $\tilde{p} = \max\{|q_1|^{1/1}, \ldots, |q_{n-1}|^{1/(n-1)}\}$. Suppose further that we apply the scale transformation $z \mapsto z/\tilde{p}$ to $\tilde{Q}(z)$, then we obtain

$$\tilde{Q}\left(\frac{z}{\tilde{p}}\right) = 1 + \hat{q}_1 z + \cdots + \hat{q}_{n-1}z^{n-1}, \quad \max\{|\hat{q}_1|, \ldots, |\hat{q}_{n-1}|\} = 1.$$

A formula which bounds the roots from below tells us that any root $\hat{\zeta}$ of $\tilde{Q}(z/\tilde{p})$ is not less than $1/(1 + \max\{|\hat{q}_1|, \ldots, |\hat{q}_{n-1}|\}) = 1/2$. Hence, we have $|\zeta| \geq 1/2\tilde{p}$ for any root ζ of $\tilde{Q}(z)$. Note that the roots of $\tilde{Q}(z)$ are $z_j - z_i$ ($j \neq i$). Consider the limiting case $\Delta \to 0$, which means $P(z) \to \tilde{P}(z)$ and $\zeta_j \to z_j$ ($j = 1, \ldots, n$), we have $r'_1 \to 0$ and $R' \to 1/2\tilde{p}$. Since the roots move continuously as Δ goes to 0 continuously, R' must be a lower bound of $|\zeta_j - z_i|$ ($\forall j \neq i$).

Using inequality $\sqrt{1-x} > 1-x$ $(0 < x < 1)$, we can bound R' as follows.

$$R' > \frac{\tilde{p}_1 + \tilde{p}\delta_0}{\tilde{p}_1(2\tilde{p} + d_0)} \left[1 - \frac{2\tilde{p}_1(2\tilde{p} + d_0)\delta_0}{(\tilde{p}_1 + \tilde{p}\delta_0)^2} \right] > R \stackrel{\text{def}}{=} \frac{1}{2\tilde{p} + d_0} \left[1 - \frac{(4\tilde{p} + 2d_0)\delta_0}{\tilde{p}_1} \right]. \quad (5.3)$$

Theorem 3. *Let $\tilde{p}_1$, $\tilde{p}$, etc. be defined as in* **3**. *If $\tilde{p}_1 > \max\{6\tilde{p} + 4d_0, 2\tilde{p}^2/d_0\} \times \delta_0$ then any root ζ_j $(j \neq i)$ of $P(z)$ is outside of the circle of radius $\{1 - (4\tilde{p} + 2d_0)\delta_0/\tilde{p}_1\}/(2\tilde{p} + d_0)$, located at $z = z_i$.*

6. Concluding remarks

Although the error bounds obtained in this article is very tight, they will be not useful practically, because the bounds are time-consuming to compute; from the practical viewpoint, Smith's formula is most useful. The idea in this article, however, will be useful because it is quite simple. In fact, the author has applied the idea successfully to bounding the error of close roots computed numerically in [8]; suppose $P(z)$ has a well-separated cluster of m close roots and let $C(z)$ be a factor of $P(z)$ corresponding to the m close roots, then the error $|\varepsilon|$ of any close root can be bounded by the size of the close-root cluster and the quantities determined by only $C(z)$.

References

[1] L. Blum, F. Cucker, M. Shub and S. Smale, Complexity and Real Computation. Springer-Verlag, New York, 1998.

[2] L. Elsner, A remark on simultaneous inclusions of the zeros of a polynomial by Gerschgorin's theorem. Numer. Math. **21** (1973), 425–427.

[3] D. Inaba and T. Sasaki, Certification of analytic continuation of algebraic function. Proc. CASC 2004 (Computer Algebra in Scientific Computing), V.G. Ganzha, E.W. Mayr and E.V. Vorozhtsov (eds.). Technishe Universität München Press, 2004, 249–260.

[4] L. Kantrovich and G. Akilov, Functional Analysis in Normed Spaces (translated from Russian). MacMillan, 1964.

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

[6] S. Smale, Newton's method estimates from data at one point. In R. Ewing, K. Gross, and C. Martin (eds.), The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics. Springer-Verlag, 1986.

[7] B.T. Smith, Error bounds for zeros of a polynomial based on Gerschgorin's theorems. J. ACM, **17** (1970), 661–674.

[8] T. Sasaki and A. Terui, Computing clustered close-roots of univariate polynomial. Preprint of Univ. Tsukuba (2007).

[9] X.H. Wang and D.F. Han, On dominating sequence method in the point estimate and Smale theorem. Sci. China (Series A), **33** (1990), 135–144.