Journal of Korean
Data & Information Science Society
2005, Vol. 16, No. 1, pp. 127~136

On Roots of Perturbed Polynomials

Young Kou Park[1)]

Abstract

We will derive some results on the perturbation of roots using Newton's interpolation formula. And we also compare our results with those obtained by Ostrowski by giving some numerical experiments with Wilkinson's polynomials.

Keywords : Perturbed polynomials, Perturbed roots

1. Introduction and Notations

In solving polynomial equations, the values of the coefficients are usually rounded off. In this paper we will approach the problem on the perturbation of roots with the aid of Newton's interpolation formula. And then we will apply the obtained results to root bound problems and give some numerical experiments. Before proceeding, we will give short comments on notations and some known results from the theory of divided differences. The set of all complex numbers is denoted by C. By $B(z_0, \rho)$, we will always mean the closed disk of radius ρ centered at z_0. For any bounded set S in C, $\mathrm{dia}(S) = sup_{z, z_1 \in S}(|z - z_1|)$. where the points $z_0, \ldots, z_n$ lie inside the countour Γ.

Definition 1.1 Let $p(z)$ be a polynomial in the complex variable z. The first divided difference of $p(z)$ is denoted by $p[z_0, z_1]$ and defined by the relation

$$p[z_0, z_1] = \frac{p(z_1) - p(z_0)}{z_1 - z_0}.$$

The n-th divided difference is defined by induction in terms of the $(n-1)$-th

1) Associate Professor, Department of Mathematics, Yeungnam University, Gyongsan, 712-749,
E-mail : ykpark@yu.ac.kr

one by the formula;

$$p[z_0, z_1, \ldots, z_n] = \frac{p[z_0, \ldots, z_{n-2}, z_n] - p[z_0, \ldots, z_{n-2}, z_{n-1}]}{z_n - z_{n-1}}.$$

Lemma 1.2 [Thomson (1933) and Tulovsky (1990)]

$$p[z_0, \ldots, z_n] = \frac{1}{2\pi i} \int_\Gamma \frac{p(z)}{(z-z_0)\cdots(z-z_n)}\, dz,$$

where the points $z_0, \ldots, z_n$ lie inside the countour Γ.

Fact. The interchange of any two of the arguments does not alter the value of the divided difference, which is therefore a symmetric function. Moreover it is clear from this formula that $p[z_0, \ldots, z_n]$ is an analytic function of $z_0, \ldots, z_n$, even if some of them coincide. For $n+1$ coincident arguments z_0, we obtain the equality

$$p[z_0, \ldots, z_n] = \frac{1}{n!} p^{(n)}(z_0).$$

If $p(z)$ is a polynomial of degree n, then by Newton's interpolation formula, $p(z)$ can be reconstructed uniquely from the values of the divided differences at $z_0, \ldots, z_n$ as follows:

$$p(z) = p[z_0] + p[z_0, z_1](z-z_0) + \cdots + p[z_0, \cdots, z_n](z-z_0)\cdots(z-z_{n-1}).$$

Definition 1.3

1) For $p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$ with roots $q_1, \ldots, q_n$,

$$U(p) = \max\{\, |q_i| : i = 1, \ldots, n\}.$$

2) Ψ will be denoted by the class of monic complex polynomials of degree n.

3) $M: \Psi \to R$ is called a root bound functional (rbf) if

$$M(p) \geq U(p) \text{ for } all\ p \in \Psi.$$

4) $M: \Psi \to R$ is called an absolute rbf if $M(p) = M(\tilde{p})$ for

$$p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0,$$

$$\tilde{p}(z) = z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$$

with $|b_i| = |c_i|$ for $all\ i$.

Now some results on this field will be presented without proof. For the proofs and references, see Van der Sluis(1970).

If $p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$, then by Cauchy's Theorem, the unique positive root z_0 of

$$z^n - |b_{n-1}| z^{n-1} - \cdots - |b_1| z - |b_0| = 0$$

is an absolute rbf. For any $p(z)(\neq z^n) \in \Psi$, we denote the corresponding z_0 as $B(p) = z_0$ and also define $B(z^n) = 0$. Then B is the best absolute rbf of all absolute rbfs. While B is optimal, the positive solution $B(p) = z_0$ can't be easily calculated. Therefore, other more computable absolute rbfs which are well-known from the literature are widely used, see Park (1995) and Van der Sluis (1970).

2. Perturbation of roots

Now we will introduce the following result which is related to our research on the perturbation of roots. The proof is found in Ostrowski (1960).

Definition 2.1 (Descarte's rule of signs (Pola and Szeo, 1976))
Let $p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$ be a real polynomial (not the zero polynomial) and let v denote the number of sign changes in the sequence $\{b_k\}$ of its non-zero coefficients, and let r denote the number of its real positive roots (each root counted with its proper multiplicity), then $v - r$ is even and non-negative.

Theorem 2.2 [Park (1993)] Let $p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$ with roots $q_1, \ldots, q_n$, $\widetilde{p}(z) = z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$ with roots $\widetilde{q}_1, \ldots, \widetilde{q}_n$,

$$\rho(n) = \left(\sum_{i=1}^{n-1} |b_i - c_i| m^i \right)^{\frac{1}{n}}, \quad m = m(|q_i|, |\widetilde{q}_i|). \tag{2.1}$$

Then the roots q_i and $\widetilde{q}_i$ can be ordered in such a way that

$$|q_i - \widetilde{q}_i| \leq dia(C_i) - \rho(n), \quad i = 1, \ldots, n,$$

where the roots C_i is the connected component containing q_i of $\bigcup_{i=1}^{n} B(q_i, \rho(n))$, and $\widetilde{q}_i$ is the perturbed root of q_i for each i.

Theorem 2.3 Let $p(z)=z^n+b_{n-1}z^{n-1}+\cdots+b_1z+b_0$ with roots $q_1,\ldots,q_n$, $r(z)$ some polynomial of degree $\leq n-1$, $\tilde{p}(z)=z^n+c_{n-1}z^{n-1}+\cdots+c_1z+c_0$ with roots $\tilde{q}_1,\ldots,\tilde{q}_n$. If $k(n)$ is the positive solution of the equation;

$$k^n-|r[q_1]|-|r[q_1,q_2]|k-\cdots-|r[q_1,\cdots,q_n]|k^{n-1}=0 \tag{2.2}$$

Then

(1) $p(z)$ and $\tilde{p}(z)$ have the same number of roots, counting multiplities, in each connected component of the region $G=\bigcup_{i=1}^n B(q_i,k(n))$.

(2) If G_i is the connected component of G containing q_i, then

$$|q_i-\tilde{q}_i|\leq dia(C_i)-k(n),\quad i=1,\ldots,n. \tag{2.3}$$

where $\tilde{q}_i$ is the perturbed root of q_i for each i.

Proof. By Descarte's rule of signs [Henrici (1974) and Pola and Szeo (1976)] we know that the equation (2.2) has only one positive solution. Suppose $k(n)$ is the positive solution of the equation (2.2). For $\delta>0$, let us set $G_\delta=\bigcup_{i=1}^n B(q_i;k(n)+\delta)$. For all z on the boundary of G_δ, we have

$$|z-q_i|\geq k(n)+\delta>k(n)\quad \text{for } all\ i.$$

Newton's interpolation formula gives

$$\begin{aligned} r(z)\ &= r[q_1]+r[q_1,q_2](z-q_1)+\cdots \\ &\quad + r[q_1,\ldots,q_n](z-q_1)\cdots(z-q_{n-1}). \end{aligned}$$

By using the fact that

$$\frac{|r[q_1]|+|r[q_1,q_2]||z-q_1|+\cdots+|r[q_1,\cdots,q_n]||z-q_1|\cdots|z-q_{n-1}|}{|z-q_1|\cdots|z-q_n|}$$

is a decreasing function of $|z-q_i|>0$, we have, for all z on the boundary of G_δ,

$$\begin{aligned} \frac{|r(z)|}{|p(z)|} &\leq \frac{|r[q_1]|+|r[q_1,q_2]|(k(n)+\delta)+\cdots+|r[q_1,\cdots,q_n]|(k(n)+\delta)^{n-1}}{(k(n)+\delta)^n} \\ &< \frac{|r[q_1]|+|r[q_1,q_2]|k(n)+\cdots+|r[q_1,\cdots,q_n]|k(n)^{n-1}}{k(n)^n}=1. \end{aligned}$$

From Rouche's Theorem, we see that $p(z)$ and $\tilde{p}(z)$ have the same number of roots in each connected component of $\widetilde{G}_\delta$, where $\widetilde{G}_\delta$ is the interior of G_δ Since δ is arbitrary, we can see that $p(z)$ and $\tilde{p}(z)$ have the same number of roots, counting multiplicities, in each connected component of G. Moreover, from the well known fact that the roots of a polynomial are continuous functions of the coefficients, the connected component of containing q_i has also its perturbed root $\tilde{q}_i$. This proves the first statement (1).

For the second part (2), if C_i is the connected component containing q_i, then from the above proof, the perturbed root $\tilde{q}_i$ of q_i lies in C_i. So we see that $|q_i - \tilde{q}_i| \le dia(C_i) - k(n)$ because C_i is the union of some closed disks of radius $k(n)$.

As a corollary of Theorem 2.3, we will derive Cauchy's Theorem on bounding the roots of polynomial.

Theorem 2.4 Let $\tilde{p}(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_0$. If $k(n)$ is *the positive solution of the equation*

$$k^n = |b_0| + |b_1|k + \cdots + |b_{n-1}|k^{n-1}, \tag{2.4}$$

then all roots of $\tilde{p}(z)$ *lie in the disk* $B(0, k(n))$.

Proof. Suppose that $k(n)$ is the positive solution of the equation (2.4). From Theorem 2.3, let us set $p(z) = z^n$ and $r(z) = b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$. Since each coefficient $b_i = r[0, \ldots, 0]$, the number of 0 in the bracket is $i+1$, by (1) of Theorem 2.3 we can see that $B(0, k(n))$ contains all roots of $\tilde{p}(z)$.

Remark. (1) Let q_n, be the largest absolute value of the roots of $p(z)$. Then the estimate $\rho(n)$ in Theorem 2.2 increases as $|q_n|$ increases. However, the positive solution $k(n)$ of the equation (2.2) does not change when $|q_n|$ does increase because $r[q_1, q_2, \ldots, q_{n-1}, q_n] = r[q_1, p_2, \ldots, q_{n-1}, q_n']$ for any q_n'.

(2) Theorem 2.2 says that because of the $1/n$ exponent, the bounds between q_i and $\tilde{q}_i$ are weak in general for small perturbing polynomial $r(z)$.

(3) If $k(n)$ is the positive solution of the equation (2.2), then from the following example, we can see that the estimate $|q_i - \tilde{q}_i| \le k(n)$ does not hold in general.

Consider $p(z) = z(z - 0.01)(z + 1.8)$ with roots $q_1 = 0$, $q_2 = 0.01$, $q_3 = -1.8$.

Now take $r(z) = -1$ then $p(z) + r(z) = z(z-0.01)(z+1.8) - 1$ with roots $\widetilde{q_1} \approx 0.6$, $\widetilde{q_2} \approx -1.2 + 0.3i$, $\widetilde{q_3} \approx -1.2 - 0.3i$.

From the equation (2.2), we get the positive solution $k(3) = 1$. Therefore it is impossible to get $|q_i - \widetilde{q}_i| \le 1$ because $|q_2 - \widetilde{q_2}| \le 1.3 > 1$.

(4) Theorem 2.3 does not always give better estimate than Theorem 2.2. For example, let $q > 0$,

$$p(z) = (z-q)^2 = z^2 - 2qz + q^2, \quad \tilde{p}(z) = (z+q)^2 = z^2 + 2qz + q^2.$$

The maximum of the absolute values of the roots of $p(z)$ and $\tilde{p}(z)$ is q, $r(z) = \tilde{p}(z) - p(z) = 4qz$ and $\rho(2) = 2q$ in Theorem 2.2. Note that this estimate is the actual distance between the roots. On the other hand $r[q] = 4q^2$ and $r[q, q] = 4q$. The positive root $k(2)$ of $k^2 = 4q + 4qk$ is clearly $> 2q$. Taking $p(z) = (z-q)^n$ and $\tilde{p}(z) = (z+q)^n$ gives an example of degree n which the result from Theorem 2.2 is better.

3. Applications and Numerical Results

As background on this problem, we may consider the Gershgorin disk theorem [Henrici (1974) and Johnson and Riess (1982)].

Let $A = (a_{ij})$ be a complex matrix of order n and define the absolute off-diagonal row and column sums by

$$r_k = \sum_{j(\neq k)=1}^{n} |a_{k_j}| \quad \text{and} \quad c_k = \sum_{j(\neq k)=1}^{n} |a_{j_k}|,$$

respectively. For $k = 1, \ldots, n$, set

$$\overline{R_k} = \{z : |z - a_{kk}| \le r_k\} \quad \text{and} \quad \overline{C_k} = \{z : |z - a_{kk}| \le c_k\}.$$

Then

(1) If λ is any eigenvalue of A, then $\lambda \in \overline{C_k}$ for some k and $\lambda \in \overline{R_j}$ for some j.

(2) Each component of the set $\cup_{k=1}^{n} \overline{R_k} (\cup_{k=1}^{n} \overline{C_k})$ contains as many eigenvalues of A as points a_{ii}.

To apply the Gershgorin Theorem to the above problem, we need to find a matrix A so that the characteristic polynomial of $A = (a_{ij})$, $a_{ij} \in a_1, \ldots, a_t$ for each i, corresponds to

$$p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0.$$

However, there is little hope that a matrix can be easily calculated unless all $a_i = 0$. But, as a special case we note that

$$p(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$$

is the characteristic polynomial of the matrix,

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ -b_0 & -b_1 & \cdot & \cdot & \cdot & \cdot & \cdot & -b_{n-1} \end{pmatrix}$$

Using the Gerschgorin Theorem, we have the following results.

If $\overline{C_0} = \{z \mid |z| \le |b_0|\}$, $\overline{C_k} = \{z \mid |z| \le 1 + |b_k|\}$, $1 \le k \le n-2$, $\overline{C_{n-1}} = \{z \mid |z + b_{n-1}| \le 1\}$, then

(1) All the roots of $p(z)$ lie in $\bigcup_{i=1}^{n-1} \overline{C_i}$,

(2) All the roots of $p(z)$ lie in the circle $|z| \le \max\left(1, \sum_{i=0}^{n-1} |b_i|\right)$.

Now we will give some numerical experiments with Wilkinson's polynomials for the results obtained in section 2. And using well known absolute rbfs, we also compare our results Theorem 2.3, with Theorem 2.2 obtained by Ostrowski (1960).

We note that the positive solution $k(n)$ of (2.2) can't be easily calculated. However, we can find an approximation of $k(n)$ by using previously discussed root bound functionals (rbfs) of polynomials.

Let $\tilde{k}(n)$ be a root bound of the equation (2.2) obtained by using absolute rbfs.

In order to apply (2.1), we introduce a way to find a convenient estimate of $\rho(n)$ as follows; Let $\tilde{m} = \max\{M(p), M(\tilde{p})\}$, where M is an absolute rbf.

$$\tilde{\rho}(n) = \left(\sum_{i=1}^{n-1} |b_i - c_i| \ \tilde{m}^i \right)^{\frac{1}{n}}. \tag{3.1}$$

So we have

$$| q_i - \widetilde{q}_i | \leq dia(\widetilde{C}_i) - \tilde{\rho}(n), \quad i = 1, \ldots, n$$

where $\widetilde{C}_i$ is the connected component containing q_i of $\bigcup_{i=1}^{n} B(q_i, \tilde{\rho}(n))$.

As application of Theorem 2.3, we will consider the following problem. If $p(z) = z^n + b_{n-1} z^{n-1} + \cdots + b_1 z + b_0$ is given, then for any given t-points $a_1, \ldots, a_t$ in C, we can find disks $B(a_i, \rho)$ such that the union of all disks $B(a_i, \rho)$ contains all roots of $p(z)$.

Let us consider $\tilde{p}(z) = z^4 - 12z^3 + 40z^2 - 20z + 24$ with roots

$0.17+0.81\,i$,	$0.17-0.81\,i$,
$5.83+1.16\,i$,	$5.83-1.16\,i$.

When 0 and 5 are given, we want to find $k(4)$ so that $B(0, k(4)) \cup B(5, k(4))$ contains all roots of $\tilde{p}(z)$. To apply Theorem 2.3, choose

$$p(z) = z^2(z-5)^2 = z^4 - 10z^3 + 25z^2,$$
then
$$r(z) = \tilde{p}(z) - p(z) = -2z^3 + 15z^2 - 20z - 24.$$

From $r[0] = 24$, $r[0,0] = 20$, $r[0,0,5] = -5$, $r[0,0,5,5] = 2$, we get the following equation by (2.2);

$$k^4 - 2k^3 - 5k^2 - 20k - 24 = 0,$$

which has the positive solution

$$k(4) \approx 4.43, \quad \tilde{k}(4) \approx 5.43 \text{ by rbf.}$$

Therefore, all roots of $\tilde{p}(z)$ lie in $B(0, 4.43) \cup B(5, 4.43)$. From Theorem 2.2, we obtain

$$\rho(4) \approx 5.75, \quad \tilde{\rho}(4) = 15.34$$

Consider Wilkinson's polynomial [see Wilkinson (1963)],
$w(z) = (z-1)(z-2)\cdots(z-n)$ and
$r(z) = 0.002 z^{n-1}$.

$$\beta(n) = \max(|q_i - \tilde{q}_i|). \tag{3.2}$$

Let us recall that $k(n)$ is the positive solution of the equation;

$$k^n - |r[q_1]| - |r[q_1, q_2]|k - \cdots - |r[q_1, \cdots, q_n]|k^{n-1} = 0. \tag{3.3}$$

And recall that

$$\rho(n) = \left(\sum_{i=1}^{n-1} |b_i - c_i| m^i\right)^{\frac{1}{n}}, \quad m = \max(|q_i|, |\tilde{q}_i|). \tag{3.4}$$

Then we obtain the following numerical results.

<Numerical Experiments with Wilkinson's polynomial>

deg $w(z)$	$\rho(n)$	$\tilde{\rho}(n)$	$k(n)$	$\tilde{k}(n)$	$\beta(n)$
6	1.6	6.1	0.8	1.2	0.3
8	2.8	10.6	1.6	2.6	1.1
10	4.5	27.5	2.7	4.2	2.1
12	6.5	45.2	4.0	6.6	3.3
15	10.1	81.4	6.7	10.8	5.5
20	17.8	171.3	12.7	20.1	10.3

Note that we used the absolute rbf $S(p)$ to estimate $\tilde{k}(n)$ for $n > 7$, and $R(p)$ for $n = 6, 7$, and used $L(p)$ to estimate $\tilde{\rho}(n)$ for the best estimate, for the references of the absolute rbfs, see Park (1995) and Van der Sluis (1970).

References

1. Henrici, P. (1974). *Applied and Computational Complex Analysis, Vol .1*, John Wiley and Sons.
2. Johnson, L. and Riess, R. (1982). *Numerical Analysis, 2-nd ed.*, Addison- Wesley Pub .Co.
3. Milne-Thomson, L. (1933). *The Calculus of Finite Differences*, Macmillan and Co.,London.
4. Ostrowski, A. M. (1960). *Solutions of Equations and Systems of Equations*, Academic press, New York and London.
5. Park, Y. K. (1993). *On Perturbation and Location of Roots of a Polynomials by Newton's Interpolation Formula*, Ph.D.Thesis, Oregon State Univ. U.S.A..
6. Park, Y. K. (1995). On perturbation of roots of polynomials by

Newton's interpolation formula, *Journal of the Korean Mathematical Society*, 32, 61-76.
7. Pola, G. and Szeo, G. (1976). *Problems and Theorems in Analysis II*, Springer-verlog, New York, Heidelberg Berlin.
8. Tulovsky, V. (1990). On perturbations of roots of polynomials, *Journal of D'Analyse Mathematique*, 54, 77-89.
9. Van der Sluis (1970). Upper bounds for roots of polynomials, *Numeical. Mathematics,* 15, 260-262.
10. Wilkinson, J. H. (1963). *Rounding Errors in Algebraic Process*, Prentice-Hall, Englewood Cliffs, N, J.

[received date : Nov. 2004, accepted date : Jan. 2005]