ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XI, 2 (1970)

B. BOJANOV (Wrocław)

ON AN ESTIMATION OF THE ROOTS OF ALGEBRAIC EQUATIONS

1. Introduction. In this paper we give a method for determining upper and lower bounds for the unique positive root of the equation

$$x^n = \sum_{r=1}^{n} p_r x^{n-r} \qquad (p_r \geqslant 0,\ r = 1, 2, \ldots, n). \tag{1}$$

The knowledge of this root, or its upper bound, enables us to localize zeros of other polynomials with complex coefficients. This follows from the following theorem by Cauchy.

THEOREM. *All zeros of the polynomial*

$$z^n + a_1 z^{n-1} + \ldots + a_n, \tag{2}$$

with complex coefficients, lie in a circle

$$|z| \leqslant R$$

where R is the positive root of the polynomial (1) *with*

$$p_k = |a_k| \qquad (k = 1, 2, \ldots, n).$$

This theorem plays an essential role in the majority of the known methods for finding a circular region enclosing all zeros of a polynomial. The method proposed here yields, in particular, the Westerfield bound [1] and makes it possible to find an upper bound, which is usually better than, and always at least as good as, the Westerfield bound. The proposed method gives also a lower bound for the positive root of (1).

2. The basic theorem. First, we shall prove some lemmas to be used in the sequel.

LEMMA 1. *If there is*

$$a_{jk} \geqslant 0 \quad (j = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n),$$

$$\sum_{k=1}^{n} a_{jk}^{k} \leqslant M \quad (M > 0;\ j = 1, 2, \ldots, m),$$

$$w_j \geqslant 0 \quad (j = 1, 2, \ldots, m), \quad \sum_{j=1}^{m} w_j = 1,$$

then

$$\sum_{k=1}^{n} \Big[\sum_{j=1}^{m} w_j a_{jk}\Big]^k \leqslant M. \tag{3}$$

The proof is by induction on m. Let $m = 2$ and

$$F(w) = \sum_{k=1}^{n} \big((1-w)a_{1k} + wa_{2k}\big)^k.$$

The second derivative of $F(w)$

$$F''(w) = \sum_{k=2}^{n} k(k-1)(a_{2k}-a_{1k})^2\big((1-w)a_{1k}+wa_{2k}\big)^{k-2},$$

is nonnegative for every $w \in [0, 1]$ and so the function $F(w)$ is convex on $[0,1]$. On the other hand

$$F(0) = \sum_{k=1}^{n} a_{jk}^{k} \leqslant M \quad \text{and} \quad F(1) = \sum_{k=1}^{n} a_{2k}^{k} \leqslant M.$$

Hence $F(w) \leqslant M$ for every $w \in [0, 1]$.

Now we shall show that if the lemma holds for $m = p$, then it holds also for $m = p+1$ (p is a positive integer).

Let

$$\sum_{j=1}^{p+1} w_j = 1$$

and s be such that $1 \leqslant s \leqslant p+1$ and $w_s < 1$. Next we define B_k $(k = 1, 2, \ldots, n)$ by

$$(1-w_s)B_k = \sum_{\substack{j=1\\ j\neq s}}^{p+1} w_j a_{jk}.$$

B_k satisfy the inequality

$$\sum_{k=1}^{n} B_k^k = \sum_{k=1}^{n} \Big[\sum_{\substack{j=1\\ j\neq s}}^{p+1} \frac{w_j}{1-w_s} a_{jk}\Big]^k \leqslant M.$$

This follows from

$$\sum_{\substack{j=1\\ j\neq s}}^{p+1} \frac{w_j}{1-w_s} = 1,$$

and from the inductive supposition for $m = p$. Now

$$\sum_{k=1}^{n} \Big[\sum_{j=1}^{p+1} w_j a_{jk}\Big]^k = \sum_{k=1}^{n} \big((1-w_s) B_k + w_s a_{sk}\big)^k.$$

Since $(1-w_s)+w_s = 1$, we conclude that

$$\sum_{k=1}^{n} \Big[\sum_{j=1}^{p+1} w_j a_{jk}\Big]^k \leqslant M$$

which proves (3) for $m = p+1$.

LEMMA 2. *If $H > 0$ and $H^n \geqslant \sum_{r=1}^{n} p_r H^{n-r}$, then $H \geqslant x$, where x is the unique positive root of the equation* (1).

Proof. Let

$$\varphi(y) = y^n - \sum_{r=1}^{n} p_r y^{n-r},$$

$\varphi(y) < 0$ for $0 < y < x$ because x is the unique positive root of the equation (1) and $\varphi(0) \leqslant 0$. Then from $\varphi(H) \geqslant 0$ and $H > 0$ we have

$$H \geqslant x.$$

LEMMA 3. *Let y_m be a positive root of the equation*

$$y^m = \sum_{r=1}^{m} y^{m-r},$$

and let $y(n; k_1, k_2, \ldots, k_m)$ be the unique positive root of the equation

$$y^n = \sum_{r=1}^{m} y^{n-k_r},$$

where $1 \leqslant k_1 < k_2 < \ldots < k_m \leqslant n$. Then

$$y(n; k_1, k_2, \ldots, k_m) \leqslant ym.$$

Proof. From the definition

$$y_m^m = \sum_{r=1}^{m} y_m^{m-r}.$$

This implies

$$y_m^n = \sum_{r=1}^{m} y_m^{n-r}.$$

Since $y_m \geqslant 1$ and $r \leqslant k_r$ $(r = 1, 2, \ldots, m)$, we can write

$$y_m^n \geqslant \sum_{r=1}^{m} y_m^{n-k_r}.$$

From lemma 2 follows

$$y(n; k_1, k_2, \ldots, k_m) \leqslant y_m.$$

The basic result of this paper is the following

THEOREM 1. *If x_j are positive roots of the equations*

$$x^n = a_{j1}x^{n-1} + a_{j2}^2 x^{n-2} + \ldots + a_{jn}^n$$

$$(a_{j1}, \ldots, a_{jn} \geqslant 0;\ j = 1, 2, \ldots, m),$$

then the positive root Z of the equation

$$x^n = \sum_{k=1}^{n} \Big[\sum_{j=1}^{m} a_{jk}\Big]^k x^{n-k}$$

satisfies the inequality

$$Z \leqslant x_1 + x_2 + \ldots + x_m.$$

Proof. From the equality

$$x_j^n = \sum_{k=1}^{n} a_{jk}^k x_j^{n-k}$$

it follows that

$$\sum_{k=1}^{n} \left[\frac{a_{jk}}{x_j}\right]^k = 1.$$

From lemma 1, with $w_j = x_j/(x_1 + x_2 + \ldots + x_m)$ and $a_{jk} = a_{jk}/x_j$, we obtain

$$\sum_{k=1}^{n} \left[\sum_{j=1}^{m} \frac{x_j}{x_1 + x_2 + \ldots + x_m} \frac{a_{jk}}{x_j}\right]^k \leqslant 1,$$

$$\sum_{k=1}^{n} \frac{\Big[\sum_{j=1}^{m} a_{jk}\Big]^k}{\Big[\sum_{j=1}^{m} x_j\Big]^k} \leqslant 1,$$

$$\sum_{k=1}^{n} \Big[\sum_{j=1}^{m} a_{jk}\Big]^k \Big[\sum_{j=1}^{m} x_j\Big]^{n-k} \leqslant \Big[\sum_{j=1}^{m} x_j\Big]^n.$$

Now applying lemma 2, we get

$$Z \leqslant \sum_{j=1}^{m} x_j$$

which ends the proof of theorem 1.

3. Applications. One of the applications of theorem 1 is a new, very simple proof of the following theorem of Westerfield:

THEOREM. *Let x_0 be the unique positive root of the equation*

$$x^n = \sum_{k=1}^{n} p_k x^{n-k} \qquad (p_k \geqslant 0,\ k = 1, 2, \ldots, n),$$

and let positive quantities

$$\sqrt[k]{p_k} \qquad (k = 1, 2, \ldots, n),$$

after being arranged in order of decreasing magnitudes, form a sequence

$$q_1 \geqslant q_2 \geqslant \ldots \geqslant q_n.$$

Then x_0 satisfies the inequality

$$x_0 \leqslant \sum_{k=1}^{n} q_r g_r,$$

where

$$g_1 = y_1, \qquad g_r = y_r - y_{r-1} \qquad (r = 2, 3, \ldots, n),$$

and where y_k is the unique positive root of the equation

$$y^k = \sum_{r=1}^{k} y^{k-r} \qquad (k = 1, 2, \ldots, n).$$

Proof. Apply theorem 1 to the equations

$$x^n = \sum_{l=1}^{j} (q_j - q_{j+1})^{k_l} x^{n-k_l} \qquad (j = 1, 2, \ldots, n),$$

where numbers k_l are such that

$$q_l = \sqrt[k_l]{p_{k_l}}$$

and where $q_{n+1} = 0$. We obtain

$$\begin{aligned} (4) \qquad x_0 \leqslant (q_1 - q_2) y(n; k_1) + (q_2 - q_3) y(n; k_1, k_2) + \ldots \\ \ldots + (q_{n-1} - q_n) y(n; k_1, k_2, \ldots, k_{n-1}) + \\ + q_n y(n; k_1, k_2, \ldots, k_n). \end{aligned}$$

Hence by lemma 3

$$x_0 \leqslant (q_1-q_2)y_1+(q_2-q_3)y_2+\dots+(q_{n-1}-q_n)y_{n-1}+q_ny_n =$$
$$= q_1y_1+q_2(y_2-y_1)+\dots+q_n(y_n-y_{n-1}) = q_1g_1+q_2g_2+\dots+q_ng_n,$$

and this completes the proof.

Proceeding on a similar way, we can prove a generalization of Westerfield's theorem. Let x_0 be the unique root of the equation (1), and let $c_1, c_2, \dots, c_n$ be arbitrary positive numbers. We arrange the quantities

$$\sqrt[k]{c_kp_k} \quad (k = 1, 2, \dots, n)$$

in order of decreasing magnitudes:

$$q_1 \geqslant q_2 \geqslant \dots \geqslant q_n.$$

Denote by $y(n; R_1, R_2, \dots, R_m)$ the positive root of the equation

$$y^n = \sum_{j=1}^{m} \frac{1}{c_{R_j}} y^{n-R_j},$$

where $1 \leqslant R_1 < \dots < R_m \leqslant n$. Let $y_m = \max_{R_j} y(n; R_1, R_2, \dots, R_m)$. We shall prove the following

THEOREM 2. *For any system of n positive numbers $c_1, c_2, \dots, c_n$*

$$x_0 \leqslant \sum_{r=1}^{n} q_rg_r,$$

where

$$g_1 = y_1, \quad g_r = y_r-y_{r-1} \quad (r = 2, 3, \dots, n).$$

Proof. Apply theorem 1 to the equations

$$x^n = \sum_{l=1}^{j} \frac{1}{c_{k_l}} (q_j-q_{j+1})^{k_l} x^{n-k_l} \quad (j = 1, 2, \dots, n),$$

where numbers k_l $(l = 1, 2, \dots, n)$ are such that

$$q_l = \sqrt[k_l]{c_{k_l}p_{k_l}}$$

and $q_{n+1} = 0$. We obtain

$$x_0 \leqslant (q_1-q_2)y(n; k_1)+(q_2-q_3)y(n; k_1, k_2)+\dots$$
$$\dots+(q_{n-1}-q_n)y(n; k_1, k_2, \dots, k_{n-1})+q_ny(n; k_1, k_2, \dots, k_n)$$
$$\leqslant (q_1-q_2)y_1+(q_2-q_3)y_2+\dots+(q_{n-1}-q_n)y_{n-1}+q_ny_n$$
$$= q_1y_1+q_2(y_2-y_1)+\dots+q_n(y_n-y_{n-1}) = q_1g_1+q_2g_2+\dots+q_ng_n$$

which ends the proof of theorem 2.

In the special case $c_r = 1$ $(r = 1, 2, \ldots, n)$, we obtain the Westerfield bound.

Theorem 2 suggests a question. Let W be the set of all equations (1) with the root 1. Denote by $F(c_1, c_2, \ldots, c_n: r)$ the bound resulting from theorem 2 for fixed $c_1, c_2, \ldots, c_n$ and $r \in W$. Let

$$H(c_1, c_2, \ldots, c_n) = \sup_{r \in W} F(c_1, c_2, \ldots, c_n; r).$$

Now the question is, for what system of $c_r (r = 1, 2, \ldots, n)$ $H(c_1, c_2, \ldots, c_n)$ is minimal.

Another application of theorem 1 gives a lower bound for the unique positive root of equation (1). Using the same notation as in Westerfield's theorem, we shall state the following

THEOREM 3.

$$x_0 \geqslant \sum_{r=1}^{n} q_r g_{n+1-r}.$$

Proof. Let y be the positive root of equation

$$x^n = (q_1 - p_1)x^{n-1} + (q_1 - \sqrt{p_2})^2 x^{n-2} + \ldots + (q_1 - \sqrt[n]{p_n})^n.$$

Since $q_1 y_n$ is the positive root of the equation

$$x^n = \sum_{k=1}^{n} q_1^k x^{n-k},$$

then, applying theorem 1, it is easy to see that

$$q_1 y_n \leqslant x_0 + y.$$

From Westerfield's theorem, we have

$$y \leqslant (q_1 - q_n) g_1 + (q_1 - q_{n-1}) g_2 + \ldots + (q_1 - q_2) g_{n-1}$$

$$= q_1 y_{n-1} - \sum_{r=2}^{n} q_r g_{n+1-r}.$$

Hence

$$q_1 y_n \leqslant x_0 + q_1 y_{n-1} - \sum_{r=2}^{n} q_r g_{n+1-r},$$

$$x_0 \geqslant \sum_{r=2}^{n} q_r g_{n+1-r} + q_1 (y_n - y_{n-1}) = \sum_{r=1}^{n} q_r g_{n+1-r},$$

and so the theorem is proved.

An example. Applying theorem 3 to the equation

$$x^3 = x^2 + 3x + 9$$

with the positive root $x = 3$ we obtain

$$x \geqslant 2{,}52.$$

By Westerfield's theorem, applied to the equation

$$\frac{1}{y^3} = \frac{1}{y^2} + 3\frac{1}{y} + 9,$$

we have

$$y \leqslant 0.9$$

Since $1/x = y$, hence $x \geqslant 1.11$.

Now, we shall generalize theorem 3 in the same way as Westerfield's theorem.

Generalization of theorem 3. *Let x_0, g_r, q_r and c_r $(r = 1, 2, \ldots, n)$ be defined as in theorem 2. Then*

$$x_0 \geqslant \sum_{r=1}^{n} q_r g_{n+1-r}.$$

Proof. Denote by y the positive root of the equation

$$x^n = \sum_{k=1}^{n} \left[\frac{q_1}{\sqrt[k]{c_k}} - \sqrt[k]{p_k} \right]^k x^{n-k}.$$

Applying theorem 1, we get

$$q_1 y_n \leqslant x_0 + y.$$

By theorem 2

$$y \leqslant \sum_{r=2}^{n} (q_1 - q_r) g_{n+1-r}$$

hence

$$q_1 y_n - \sum_{r=2}^{n} (q_1 - q_r) g_{n+1-r} \leqslant x_0, \qquad \sum_{r=1}^{n} q_r g_{n+1-r} \leqslant x_0.$$

In the proof of Westerfield's theorem we have obtained the inequality

$$x_0 \leqslant (q_1 - q_2) y(n; k_1) + (q_2 - q_3) y(n; k_1, k_2) + \ldots$$
$$+ (q_{n-1} - q_n) y(n; k_1, k_2, \ldots, k_{n-1}) + q_n y(n; k_1, k_2, \ldots, k_n).$$

This bound is better, or at least not worse, than Westerfield's bound.

To ilustrate the use of inequality (4), we may apply it to the equation

(5) $$x^9 = x^8 + 512x^6 + 16x^5 + 16807x^4 + x^2 + 1.$$

In this case we have

$$n = 9, \quad q_1 = 8, \quad q_2 = 7, \quad q_3 = 2, \quad q_4 = q_5 = q_6 = 1,$$
$$q_7 = q_8 = q_9 = 0, \quad k_1 = 3, \quad k_2 = 5, \quad k_3 = 4, \quad k_4 = 9,$$
$$k_5 = 7, \quad k_6 = 1,$$

so we get

$$x_0 \leqslant y(9;3) + 5y(9;3,5) + y(9;3,5,4) + y(9;3,5,4,9,7,1).$$

Since

$$y(9;3) = 1,$$
$$y(9;3,5) < 1.21,$$
$$y(9;3,5,4) < 1.33,$$
$$y(9;3,5,4,9,7,1) < 1.98,$$

then

$$x_0 < 10.36.$$

Applying Westerfield's method we obtain

$$x_0 < 13.$$

Actually we have

$$x_0 \approx 9.265.$$

To simplify the application of inequality (4), we provide the reader with a table of the positive roots of the equations:

$$y^n = \sum_{r=1}^{m} y^{m-r}$$

for $m = 2, 3, \ldots, n$ and $n = 2, 3, \ldots, 10$.

TABLE of the roots $y(n; n-1, \ldots, n-m+1)$

m \ n	10	9	8	7	6	5	4	3	2
10	1.999								
9	1.612	1.998							
8	1.451	1.608	1.997						
7	1.355	1.444	1.602	1.992					
6	1.286	1.343	1.432	1.590	1.984				
5	1.230	1.269	1.325	1.413	1.571	1.966			
4	1.180	1.207	1.244	1.297	1.381	1.535	1.928		
3	1.131	1.149	1.172	1.204	1.250	1.325	1.466	1.840	
2	1.076	1.086	1.097	1.113	1.135	1.168	1.220	1.325	1.619

These roots may be used as upper bounds for the roots

$$y(n;k_1,k_2,\dots,k_m) \quad (n \leqslant 10).$$

In fact, it is easy to see that

$$y(n;k_1,k_2,\dots,k_m) \leqslant y(s;s,s-1,\dots,s-m+1),$$

where $s = j+m-1$ and $j = \min(k_1,k_2,\dots,k_m)$.

An example. We apply it to the equation (5). Since

$$y(9;3) = 1,$$
$$y(9;3,5) < y(4;4,3) < 1.220,$$
$$y(9;3,5,4) < y(5;5,4,3) < 1.325,$$
$$y(9;3,5,4,9,7,1) < y(6;6,5,4,3,2,1) < 1.984,$$

then

$$x_0 < 10.409.$$

The bound (4) for the positive root of equation

$$x^n = q_n x^{n-1} + q_{n-1}^2 x^{n-2} + \dots + q_1^n,$$

where $n \leqslant 8$ and $q_1 \geqslant q_2 \geqslant \dots \geqslant q_n$, is

$$x_0 \leqslant \delta_1 + 1.097\delta_2 + 1.172\delta_3 + 1.244\delta_4 + 1.325\delta_5 + 1.432\delta_6 + 1.602\delta_7 + 1.997\delta_8,$$

where $\delta_i = q_i - q_{i+1}$ $(i = 1,2,\dots,7)$ and $\delta_8 = q_8$. For Westerfield's bound we obtain

$$x_0 \leqslant \delta_1 + 1.619\delta_2 + 1.840\delta_3 + 1.928\delta_4 + 1.966\delta_5 + 1.984\delta_6 + 1.992\delta_7 + 1.997\delta_8.$$

Reference

[1] E. C. Westerfield, *A new bound for thd zeros of polynomials,* Amer. Math. Monthly 40 (1933), pp. 18-23.

Received on 15. 6. 1968

B. BOJANOV (Wrocław)

SZACOWANIE PIERWIASTKÓW RÓWNAŃ ALGEBRAICZNYCH

STRESZCZENIE

Opisano metodę szacowania z góry dodatniego pierwiastka równania (1) za pomocą dodatnich pierwiastków równań tego samego typu (twierdzenie 1). Podano nowy dowód twierdzenia Westerfielda [1] i pewne jego uogólnienie (twierdzenie 2).

Otrzymana nierówność (4) pozwala uzyskać oszacowania lepsze, albo przynajmniej nie gorsze, niż oszacowanie Westerfielda. Uzyskano również oszacowanie z dołu dodatniego pierwiastka równania (1).

Б. БОЯНОВ (Вроцлав)

ОБ ОЦЕНКЕ КОРНЕЙ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

РЕЗЮМЕ

В работе описан метод оценки сверху положительного корня уравнения (1) с помощью положительных корней уравнений того же вида (теорема 1). Дается новое доказательство теоремы Вестерфилда [1] и ее обобщение (теорема 2). Из полученного неравенства (4) вытекают оценки корней, которые лучше или по крайней мере не хуже оценок Вестерфилда. Получена также оценка снизу положительного корня уравнения (1).
