ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XI, 3 (1970)

H. BOGDANÓW (Wrocław)

APROKSYMACJA PIERWIASTKÓW RÓWNANIA SZEŚCIENNEGO

1. Wstęp. Celem niniejszej pracy jest podanie wzorów przybliżonych na największy pierwiastek równania sześciennego

$$x^3+px+q=0 \tag{1}$$

o współczynnikach p, q rzeczywistych i wyróżniku

$$D=\frac{p^3}{27}+\frac{q^2}{4}$$

ujemnym. Jeżeli ξ oznacza dowolny pierwiastek równania (1), to pozostałe pierwiastki spełniają równanie kwadratowe

$$x^2+\xi x+p+\xi^2=0 \tag{2}$$

(czyli równanie $x^2+\xi x-q/\xi=0$). Tak więc wyznaczenie pozostałych pierwiastków sprowadza się do rozwiązania równania kwadratowego.

Dokładne metody wyznaczania pierwiastków równania (1) w przypadku $D<0$ wymagają np. — oprócz kilku pierwiastkowań — obliczenia wartości funkcji arctg x i funkcji trygonometrycznej. Waszczyszyn i Życzkowski w pracy [3] podają pewne wzory aproksymacyjne na pierwiastki równania (1) dla $D<0$, których zastosowanie wymaga jedynie pierwiastkowania. Maksymalny błąd bezwzględny obliczonych pierwiastków nie przekracza $0.0_3360\sqrt{-p}$.

Niniejsza praca zawiera kilka podobnych, ale otrzymanych inną metodą i dokładniejszych wzorów na największy pierwiastek równania (1) w przypadku $D<0$.

2. Zastosowana metoda. Zakładam dalej, że $D<0$. Jeśli w (1) podstawić

$$x=\sqrt{\frac{-p}{3}}(t+1), \tag{3}$$

a następnie

$$\lambda = \frac{1}{4}\left(2 - \frac{q}{\frac{-p}{3}\sqrt{\frac{-p}{3}}}\right), \tag{4}$$

to otrzyma się równanie

$$\tfrac{1}{4}t^2(t+3) = \lambda \tag{5}$$

z niewiadomą t. Ponieważ

$$\lambda = \frac{1}{2}\left(1 - \frac{q/2}{\sqrt{q^2/4-D}}\right), \tag{4'}$$

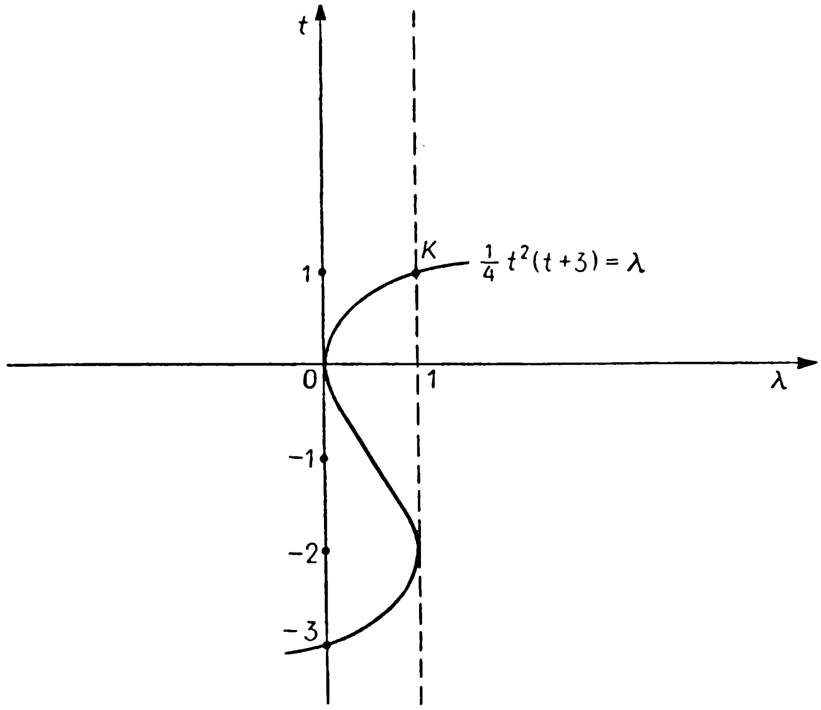
więc $0 < \lambda < 1$.

Największy pierwiastek równania (1) odpowiada największemu pierwiastkowi $t(\lambda)$ równania (5).

Można obliczyć, że

$$t(\lambda) = 2\sin\left(\frac{2\pi}{3} + \frac{1}{3}\arcsin(1-2\lambda)\right) - 1. \tag{6}$$

Wykresem funkcji $t = t(\lambda)$, $0 \leqslant \lambda \leqslant 1$, jest część OK krzywej na rys 1.



Rys. 1

Funkcję (6) aproksymuję w przedziale $\langle 0,1 \rangle$.

Ze względu na jakościowe podobieństwo wykresu rozważanej funkcji do wykresu funkcji $\sqrt{\lambda}$, wybieram funkcje aproksymujące postaci

$$W_{mn}(\lambda) = A_m(\lambda) + B_n(\lambda)\sqrt{\lambda}, \tag{7}$$

gdzie $A_m(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m$, $B_n(\lambda) = b_0 + b_1\lambda + \ldots + b_n\lambda^n$ są wielomianami stopnia m i n.

Ponieważ funkcje $1, \lambda, \lambda^2, \ldots, \lambda^m, \sqrt{\lambda}, \lambda\sqrt{\lambda}, \ldots, \lambda^n\sqrt{\lambda}$ tworzą układ Czebyszewa w przedziale $0 < \lambda \leqslant 1$ [1], więc współczynniki wielomianów $A_m(\lambda)$ i $B_n(\lambda)$ można wyznaczyć za pomocą drugiego algorytmu Remeza [2].

W pierwszym etapie obliczeń aproksymowano pierwiastek (6) funkcjami postaci (7) na zbiorze Λ punktów $\lambda_k = \frac{1}{4}t_k^2(t_k+3)$, gdzie $t_k = k/128$, $k = 0, 1, 2, \ldots, 128$. W taki sposób otrzymano dla ustalonych par wartości m, n

(a) współczynniki $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$ funkcji optymalnej,

(b) błędy ε_{mn} aproksymacji na zbiorze Λ,

(c) alternans $\lambda_{p_1}, \lambda_{p_2}, \ldots, \lambda_{p_{(m+n+3)}}$.

Obliczenia wykonano z podwójną dokładnością (20 cyfr mantysy) na maszynie cyfrowej Elliott 803.

Tworzone w algorytmie Remeza układy równań liniowych były rozwiązywane metodą elementów głównych.

3. Wyniki obliczeń. Uzyskane wyniki są zawarte w tablicach 1 i 2. Ponieważ liczba działań potrzebnych do obliczenia wartości $W_{mn}(\lambda)$ jest zależna od sumy $\delta = m+n$, więc praktyczne znaczenie mogą mieć te wzory, którym przy stałym $m+n$ odpowiada najmniejszy błąd ε_{mn}. Wartości optymalne wyróżniono kursywą. Z tablicy 1 wynika, że badany błąd ε_{mn} dla ustalonego $\delta = m+n$ ma zapewne zawsze najmniejszą wartość, jeśli $m = \delta - [\delta/2]$, $n = [\delta/2]$.

Aby poprawić wartości punktów alternansu i otrzymać dokładniejsze wartości współczynników $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$, w drugim etapie obliczeń aproksymowano funkcję (6) funkcjami postaci (7) przy m, n

TABLICA 1. (Próby ustalenia m, n optymalnych; 1 etap obliczeń)

$m+n$	m	n	ε_{mn}	$m+n$	m	n	ε_{mn}
2	1	1	*$.13980_{10}-3$*		2	4	$.12075_{10}-6$
3	2	1	*$.16641_{10}-4$*	7	6	1	$.16395_{10}-6$
	1	2	$.30509_{10}-4$		5	2	$.15095_{10}-7$
4	3	1	$.35038_{10}-5$		4	3	*$.56749_{10}-8$*
	2	2	*$.21311_{10}-5$*		3	4	$.78470_{10}-8$
	1	3	$.97592_{10}-5$	8	5	3	$.11032_{10}-8$
5	4	1	$.10328_{10}-5$		4	4	*$.82534_{10}-9$*
	3	2	*$.28618_{10}-6$*		3	5	$.19844_{10}-8$
	2	3	$.43654_{10}-6$	9	6	3	$.27027_{10}-9$
	1	4	$.39516_{10}-5$		5	4	*$.12184_{10}-9$*
6	5	1	$.38066_{10}-6$		4	5	$.15883_{10}-9$
	4	2	$.57337_{10}-7$	10	5	5	*$.18287_{10}-10$*
	3	3	*$.39783_{10}-7$*	11	6	5	*$.27796_{10}-11$*

optymalnych na zbiorze punktów

$$\lambda_{kl} = \tfrac{1}{4} t_{kl}^2 (t_{kl}+3),$$

gdzie

$$t_{kl} = (k+l\cdot 2^{-h})/128, \quad k = p_1, p_2, \ldots, p_{m+n+3},$$
$$l = 0, \pm 1, \pm 2, \ldots, \pm 2^{h-1}, \quad h = 3 \text{ lub } 4.$$

Tablica 2. przedstawia wartości współczynników $a_0, a_1, \ldots, a_m$, $b_0, b_1, \ldots, b_n$ dla optymalnych wartości m, n.

TABLICA 2. (Współczynniki $a_0, a_1, \ldots, a_m$, $b_0, b_1, \ldots, b_n$ i błędy ε_{mn} w przypadkach m, n optymalnych; alternans).

$m = 1,\ n = 1$	$\varepsilon = .13983_{10}-3$
$a_0 =$ .000140	0
$a_1 =$ −.192081	.14794922
	.50146485
$b_0 =$ 1.149587	.85302735
$b_1 =$ .042494	1

$m = 2,\ n = 1$	$\varepsilon = .16650_{10}-4$
$a_0 =$ .0000167	0
$a_1 =$ −.2136784	.09960938
$a_2 =$ −.0179175	.35156250
	.66210938
$b_0 =$ 1.1537658	.90625000
$b_1 =$ .0777968	1

$m = 2,\ n = 2$	$\varepsilon = .21316_{10}-5$
$a_0 =$ .00000213	0
$a_1 =$ −.22000257	.06933594
$a_2 =$ −.03897492	.25781250
	.50976563
$b_0 =$ 1.15453058	.75683594
$b_1 =$ .09591487	.93554688
$b_2 =$ .00853204	1

$m = 3,\ n = 2$	$\varepsilon = .28640_{10}-6$
$a_0 =$ .000000286	0
$a_1 =$ −.221679604	.05175781
$a_2 =$ −.053351974	.19628906
$a_3 =$ −.004368483	.40039063
	.62304688
$b_0 =$ 1.154669796	.81933594
$b_1 =$ .103238760	.95312500
$b_2 =$ .021490931	1

$m = 3,\ n = 3$	$\varepsilon = .39831_{10}-7$
$a_0 =$ .0000000398	0
$a_1 =$ −.2220954066	.04003906
$a_2 =$ −.0607790353	.15429688
$a_3 =$ −.0125029123	.32128906
	.51464844
$b_0 =$ 1.1546950034	.70312500
$b_1 =$ .1057917505	.86035156
$b_2 =$ .0325431768	.96386719
$b_3 =$ .0023474235	1

$m = 4,\ n = 3$	$\varepsilon = .56851_{10}-8$
$a_0 =$ .00000000569	0
$a_1 =$ −.22219359590	.03222656
$a_2 =$ −.06399260481	.12402344
$a_3 =$ −.02082742371	.26269531
$a_4 =$ −.00130576402	.42871094
	.60253906
$b_0 =$ 1.15469954586	.76171875
$b_1 =$ .10659513934	.88964844
$b_2 =$ .03950076127	.97167969
$b_3 =$ .00752393060	1

$m = 4,\ n = 4$	$\varepsilon = .82801_{10}-9$
$a_0 =$ .000000000828	0
$a_1 =$ −.222215936078	.02636719
$a_2 =$ −.065219844013	.10205078
$a_3 =$ −.027010532318	.21826172
$a_4 =$ −.004633095139	.36181641
	.51757813
$b_0 =$ 1.154700361009	.66992188
$b_1 =$ .106829557021	.80468750
$b_2 =$ .043099039923	.90966797
$b_3 =$ .013705014360	.97705078
$b_4 =$ .000745435235	1

TABLICA 2 c.d.

$m = 5, n = 4$	$\varepsilon = .12256_{10}-9$	$m = 5,\ n = 5$	$\varepsilon = .18376_{10}-10$
$a_0 = .000\,000\,000\,123$	0	$a_0 = .000\,000\,000\,018$	0
$a_1 = -.222\,220\,872\,538$	.02197266	$a_1 = -.222\,221\,937\,609$	.01660156
$a_2 = -.065\,646\,687\,794$	.08544922	$a_2 = -.065\,784\,671\,122$	.07128906
$a_3 = -.030\,732\,502\,628$	.18408203	$a_3 = -.032\,657\,221\,457$	.15625000
$a_4 = -.009\,174\,709\,807$	.30810547	$a_4 = -.013\,579\,178\,609$	.26562500
$a_5 = -.000\,434\,249\,851$	.44726563	$a_5 = -.001\,838\,549\,381$	.38964844
	.58935547		.52148438
$b_0 = 1.154\,700\,506\,782$	.72216797	$b_0 = 1.154\,700\,532\,758$	.64746094
$b_1 = .106\,894\,088\,283$	.83740235	$b_1 = .106\,911\,044\,232$	.76464844
$b_2 = .044\,716\,228\,363$	.92529297	$b_2 = .045\,369\,592\,387$	.86132813
$b_3 = .018\,997\,577\,196$	.98095703	$b_3 = .022\,632\,791\,105$	.93750000
$b_4 = .002\,900\,621\,754$	1	$b_4 = .006\,210\,517\,352$	.98437500
		$b_5 = .000\,257\,080\,345$	1

Otrzymane wzory na największy pierwiastek równania (1) mają postać

$$x(\lambda) = \sqrt{\frac{-p}{3}}\,[1+a_0+a_1\lambda+\ldots+a_m\lambda^m+(b_0+b_1\lambda+\ldots+b_n\lambda^n)\sqrt{\lambda}], \tag{8}$$

gdzie

$$\lambda = \frac{1}{4}\left(2-\frac{q}{(-p/3)\sqrt{-p/3}}\right) = \frac{1}{2}\left(1-\frac{q/2}{\sqrt{q^2/4-D}}\right),$$

a współczynniki $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$ mają wartości przytoczone w tablicy 2.

Błąd bezwzględny obliczonego pierwiastka nie przekracza $\sqrt{-p/3}\cdot\varepsilon$, gdzie ε jest odpowiednią wartością z tablicy 2.

Pozostałe pierwiastki równania sześciennego (1) można wyrazić za pomocą wzorów

$$x_{2,3} = \frac{-x}{2}\pm\sqrt{-p-3\left(\frac{-x}{2}\right)^2}$$

na pierwiastki równania (2).

Prace cytowane

[1] G. G. Lorentz, *Approximation of functions*, New York 1966, str. 23-26.

[2] S. Paszkowski, *The theory of uniform approximation, I, Non-asymptotic theoretical problems*, Rozprawy Mat. 26 (1962), rozdz. IV.

[3] Z. Waszczyszyn i M. Życzkowski, *Wzory aproksymacyjne na pierwiastki rzeczywiste równania stopnia trzeciego*, Zastosow. Matem. 8 (1966), str. 243-254.

KATEDRA METOD NUMERYCZNYCH
UNIWERSYTETU WROCŁAWSKIEGO

Praca wpłynęła 14. 12. 1968

Г. БОГДАНОВ (Вроцлав)

ПРИБЛИЖЕНИЕ КОРНЕЙ КУБИЧЕСКОГО УРАВНЕНИЯ

РЕЗЮМЕ

В работе рассмотрены выражения (8) аппроксимирующие наибольший корень кубического уравнения (1) с вещественными коэффициентами p, q и отрицательным дискриминантом. Параметр λ в формулах (8) равен

$$\lambda = \frac{1}{4}\left(2 - \frac{q}{\frac{-p}{3}\sqrt{\frac{-p}{3}}}\right).$$

В работе приведены коэффициенты $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$ выражений (8), наилучшие в смысле равномерной аппроксимации корня при фиксированном значении суммы $m+n$. Абсолютная погрешность корня уравнения (1), вычисленного по формуле (8) не превышает $\varepsilon\sqrt{-p/3}$, где ε определено в таблице 2.

H. BOGDANÓW (Wrocław)

AN APPROXIMATION OF THE ROOTS OF A CUBIC EQUATION

SUMMARY

Considered is the approximation formula (8) giving the maximum root of cubic equation (1) with real coefficients p, q and negative discriminant. The parameter λ in (8) is given by

$$\lambda = \frac{1}{4}\left(2 - \frac{q}{\frac{-p}{3}\sqrt{\frac{-p}{3}}}\right).$$

The paper gives the coefficients $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$ of formula (8) which are optimum in the sense of a uniform approximation for a fixed sum $m+n$. The absolute error of the root of equation (1) calculated from formula (8) does not exceed $\varepsilon\sqrt{-p/3}$, where ε is given in table 2.