Reports of the Academy of Sciences of the USSR

1957

SovietRxiv

Abstract

Full Text

Reports of the Academy of Sciences of the USSR

1957, Volume 115, No. 3

MATHEMATICS

M. S. GORNSTEIN

NUMERICAL SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS

(Presented by Academician M. V. Keldysh on 6 II 1957)

I. Let it be required to find a simple root Z of the equation $f(z) = 0$, lying in the domain $G: \ |z-\zeta| < \rho$, where ζ is a fixed number; $f(z)$ is a function of a complex variable, regular in the domain G. We represent this equation in the form $z = \varphi_n(z;\zeta)$, where $\varphi_n(z;\zeta) = z - H_n(z;\zeta)f(z)$; $H_n(z;\zeta)$ is a polynomial of degree n in z, whose coefficients depend on ζ. Expand the functions $H_n(z;\zeta)$ and $f(z)$ in powers of $z-\zeta$:

$$H_n(z,\zeta) = \sum_{k=0}^{n} h_{nk}(z-\zeta)^k, \quad f(z) = \sum_{s=0}^{\infty} f_s(z-\zeta)^s \quad \left(f_s = \frac{f^{(s)}(\zeta)}{s!}\right).$$

Then

$$\varphi_n(z;\zeta) = \sum_{k=0}^{\infty}(z-\zeta)^k,$$

where $\varphi_{n0} = \zeta - f_0 h_{n0}$, $\varphi_{n1} = 1 - f_1 h_{n0} - f_0 h_{n1}$,

$$\varphi_{nk} = -\sum_{s=0}^{k} f_{k-s} h_{ns} \quad (k = 2, 3, \dots; \ h_{ns} = 0 \text{ for } s > n),$$

and

$$\varphi_{nk} = \frac{1}{k!} \left.\frac{d^k \varphi_n(z;\zeta)}{dz^k}\right|_{z=\zeta}.$$

Subject the function $\varphi_n(z;\zeta)$ to the following $n+1$ conditions:

$$\left.\frac{d^k \varphi_n(z;\zeta)}{dz^k}\right|_{z=\zeta} = 0 \quad \text{for } k = 1, 2, \dots, n+1,$$

which gives a system of $n+1$ linear equations $\varphi_{nk}=0$ $(k=1,2,\dots,n+1)$ with unknowns $h_{n0},h_{n1},\dots,h_{nn}$. The determinant of this system is

$$g_n=\begin{vmatrix} f_1 & f_0 & 0 & \dots & 0 & 0 \\ f_2 & f_1 & f_0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_n & f_{n-1} & f_{n-2} & \dots & f_1 & f_0 \\ f_{n+1} & f_n & f_{n-1} & \dots & f_2 & f_1 \end{vmatrix}. \tag{1}$$

The solution of the system (under the condition $g_n\neq 0$) will be $h_{nk}=g_{nk}/g_n$ $(k=0,\dots,n)$, where g_{nk} are the algebraic cofactors of the elements of the first row of determinant (1).

The given equation can be written in the form

$$z=z_n+r_n(z;\zeta), \tag{2}$$

where

$$z_n=\zeta-f_0g_{n-1}/g_n, \tag{3}$$

$$r_n(z;\zeta)=\sum_{k=n+2}^{\infty}\varphi_{nk}(z-\zeta)^k. \tag{4}$$

We take z_n as the approximate value of the root and shall call it the **approximation of the n-th order**. The error of the n-th approximation is equal to $r_n(z;\zeta)$.

II. Expanding the determinant g_n by the elements of the first column, we obtain for it the recurrence formula

$$g_n=\sum_{s=0}^{n}(-f_0)^s f_{s+1}g_{n-1-s} \qquad (n=0,1,\dots;\ g_{-1}=1). \tag{5}$$

Using the known properties of determinants, we find

$$g_{nk}=(g_{k-1}g_{n-1}-g_{k-2}g_n)/(-f_0)^k \qquad (k=0,\dots,n;\ g_{-1}=1,\ g_{-2}=0).$$

One can express g_n directly in terms of f_s,

$$g_n=\sum_{s=0}^{n}(-1)^s A_s^{(n)}f_0^s.$$

However, as n and s increase, the expressions for $A_s^{(n)}$ rapidly become complicated, and it is hardly advantageous to use them for large values of n and s. We write out several of the first values of $A_s^{(n)}$ (the notation $m^{(k)} = m(m-1)\dots(m-k+1)$ has been introduced):

$$\begin{aligned} A_0^{(n)} &= f_1^{n+1}; \\ A_1^{(n)} &= n f_1^{n-1} f_2; \\ A_2^{(n)} &= (n-1)^{(1)} f_1^{n-2} f_3 + (n-1)^{(2)} f_1^{n-3} f_2^2/2!; \end{aligned}$$

$$A_3^{(n)} = (n-2)^{(1)} f_1^{n-3} f_4 + (n-2)^{(2)} f_1^{n-4} f_2 f_3 + (n-2)^{(3)} f_1^{n-5} f_2^3/3!;$$

$$\begin{aligned} A_4^{(n)} = (n-3)^{(1)} f_1^{n-4} f_5 &+ (n-3)^{(2)} f_1^{n-5} (f_2 f_4 + f_3^2/2!) \\ &+ (n-3)^{(3)} f_1^{n-6} f_2^2 f_3/2! + (n-3)^{(4)} f_1^{n-7} f_2^4/4!, \dots \end{aligned}$$

The first expressions for g_n (in terms of f_s) have the form:

$$g_0 = f_1; \qquad g_1 = f_1^2 - f_0 f_2; \qquad g_2 = f_1^3 - 2 f_0 f_1 f_2 + f_0^2 f_3;$$

$$g_3 = f_1^4 - 3 f_0 f_1^2 f_2 + f_0^2 (2 f_1 f_3 + f_2^2) - f_0^3 f_4;$$

$$g_4 = f_1^5 - 4 f_0 f_1^3 f_2 + 3 f_0^2 (f_1 f_3 + f_2^2) - 2 f_0^3 (f_1 f_4 + f_2 f_3) + f_0^4 f_5, \dots$$

III. We introduce the notation

$$g_n / g_{n-1} = a_n \quad (n = 0, 1, \dots); \qquad f_0 / a_n = l_n \quad (n = 0, 1, \dots); \tag{6}$$

$$f_n / f_1 = m_n \quad (n = 2, 3, \dots)$$

(it is assumed that $f_1 \neq 0$). From (5) we obtain

$$a_n = f_1 + \sum_{s=1}^{n} (-f_0)^s f_{s+1} : \frac{g_{n-1}}{g_{n-1-s}} =$$

$$= f_1 + \sum_{s=1}^{n} \frac{(-f_0)^s f_{s+1}}{a_{n-1} \dots a_{n-s}},$$

or

$$a_n = f_1(1 + b_n), \tag{7}$$

where

$$b_n = \sum_{s=1}^{n} (-1)^s l_{n-1} \dots l_{n-s} m_{s+1}. \tag{8}$$

Dividing f_0 by both sides of (7), we obtain

$$l_n = \frac{l_0}{1 + b_n}. \tag{9}$$

Lemma. If the conditions are satisfied (capital letters denote moduli of numbers)

$$L_0 = |f_0/f_1| < 3 - 2\sqrt{2} = 0.171572\dots, \qquad M_{s+1} < \left((3 - 2\sqrt{2})/L_0\right)^s \tag{10}$$

$$(s = 1, 2, \dots)$$

the sequence a_n is bounded, and $a_n \neq 0$ (hence also $g_n \neq 0$) for $n = 0, 1, \dots$.

Proof. First suppose that $L_0 < 1$. Take a number q from the interval $(L_0, 1)$ and another number

$$q_1 = (q - L_0)/(2q - L_0).$$

Obviously, $q_1 < 1$. Require that $q_1 > q$. It is easy to show that, for $L_0 < 3 - 2\sqrt{2}$, one has $q_1 > q$ for all values of q from the specified interval, and for $q = L_0(1 + 1/\sqrt{2})$ the ratio q_1/q assumes the largest possible value:

$$(q_1/q)_{\max} = (2 - 3\sqrt{2})/L_0.$$

By complete induction one can prove that, when conditions (10) are satisfied, $L_n < q$ for $n = 0, 1, 2, \dots$. From (8) we obtain

$$B_n \leq \sum_{s=1}^{n} L_{n-1} \dots L_{n-s} M_{s+1} < \sum_{s=1}^{n} q^s \left(\frac{q_1}{q}\right)^s < \frac{q_1}{1 - q_1} = \sqrt{2} - 1 < 0.5.$$

Hence, taking (7) into account, the assertions of the lemma follow.

IV. Denote the lower and upper bounds of the sequence A_n respectively by a and A; $a \leq A_n \leq A$. In addition, suppose that the sequence

$$F_n = |f_n| = |f_n(\zeta)/n!|$$

is bounded, and let

$$F_n \leq F \qquad (n = 2, 3, \dots). \tag{11}$$

Theorem 1. *If conditions (10), (11) and the condition* $\mu < 1$ *are satisfied, where*

$$\mu = A\rho(1 + F/(1 - L_0)A)/a,$$

then the sequence (3), which can be written in the form $z_n = \zeta - l_n$*, converges to a root of the given equation.*

Proof. It is enough to show that $\lim_{n\to\infty} R_n = 0$. From (4) we obtain

$$R_n < \sum_{k=n+2}^{\infty} \Phi_{nk}\rho^k, \qquad \Phi_{nk} \leq \sum_{s=0}^{k} F_{k-s}H_{ns} \leq F\sum_{s=0}^{n} H_{ns}$$

($h_{ns} = 0$ for $s > n$). It can be shown that

$$H_{ns} < \frac{FA^{n-s}(A + F/(1 - L_0))^{s-1}}{(1 - L_0)a^{n+1}} \qquad (s = 1, 2, \dots, n)$$

and

$$H_{n0} < A^n/a^{n+1},$$

so that

$$\Phi_{nk} < F\sum_{s=0}^{n} \frac{FA^{n-s}\big(A + F/(1 - L_0)\big)^{s-1}}{(1 - L_0)a^{n+1}} < \frac{FA^n}{a^{n+1}}\left(1 + \frac{1}{1 - L_0}\frac{F}{A}\right)^n,$$

$$R_n < K\rho\mu^{n+1} \qquad (n > 0), \tag{12}$$

where

$$K = F/A(1 - \rho)(1 + F/(1 - L_0)A), \qquad \mu = A\rho(1 + F/(1 - L_0)A)/a.$$

In view of the condition $\mu < 1$, we shall have $\lim_{n\to\infty} R_n = 0$. The theorem is proved.

The error estimate of the n-th approximation ($n = 1, 2, \dots$) is given by expression (12). As for the zeroth-order approximation, for it we have

$$r_0 = -\frac{1}{f_1}\sum_{k=2}^{\infty} f_k(z - \zeta)^k.$$

Suppose that in this series the first of the numbers f_k that is not equal to zero is f_r ($r \geq 2$). Then we obtain

$$R_0 < (F/(1 - \rho)F_1)\rho^r.$$

Remarks 1. In practice, when computing the error, the following simplifications may be allowed: 1) replace $1 - L_0$ and $1 - \rho$ by unity; 2) take $a = A = F_1 = |f_1|$. Then

$$K = F/(F + F_1), \qquad \mu = (1 + F/F_1)\rho.$$

2. In the case of a real root of a real function one obtains

$$K = F/(F + F_1)(1 + \rho(1 + F_1/F)).$$

3. It can be shown that: *if a sequence* f_k $(k = 0, 1, \dots)$ *has been found corresponding to some number* ζ *satisfying the conditions of Theorem 1, then in some* ρ*-neighborhood of the number* ζ *there lies a root of the equation, and the sequence* z_n *converges to the root.* One may take $\rho \approx L_0$.

Example. $z^3 - 6z^2 + 109z - 306 = 0$. Let $\zeta = 3$. Then $f_0 = -6$, $f_1 = 100$, $f_2 = 3$, $f_3 = 1$, $f_k = 0$ for $k > 3$. One may use formu-

by (3), computing g_n by (5), or by the formula $z_n = \zeta - l_n$, computing b_n by (8) and l_n by (9). Finally, one may compute g_n (up to $n = 4$) from the expressions in item II. We obtain $z_0 = 3.060$; $z_1 = 3.05989$; $z_2 = 3.0598945$; $z_3 = 3.0598945279$; $z_4 = 3.059894527995; \dots$ The errors, computed from the simplified values of K and μ, are: $R_0 < 0.11 \cdot 10^{-3}$; $R_1 < 0.23 \cdot 10^{-5}$; $R_2 < 0.14 \cdot 10^{-6}$; $R_3 < 0.87 \cdot 10^{-8}$; $R_4 < 0.53 \cdot 10^{-10}$. The actual errors are: $0.11 \cdot 10^{-3}$; $0.23 \cdot 10^{-5}$; $0.2 \cdot 10^{-7}$; $0.2 \cdot 10^{-9}$; 10^{-12}.

V. **Generalization of Newton's process.** Denote by $z_{n_1 n_2}$ the approximation of order n_2 under the condition that z_{n_1} is taken as the initial value of the root; in general, denote by $z_{n_1 \dots n_{m-1} n_m}$ the approximation of order n_m under the condition that the initial value of the root is $z_{n_1 \dots n_{m-1}}$. It is assumed that $n_m \geq 0$ $(m = 1, 2, \dots)$.

Theorem 2. *If the initial value of the root* ζ *satisfies the conditions of Theorem 1, then the sequence* $z_{n_1}, z_{n_1 n_2}, z_{n_1 n_2 n_3}, \dots$ *converges to the root of the equation.*

Proof. We shall first assume that all $n_k > 0$. We have $|z - z_{n_1}| = R_{n_1} < K\rho\mu^{n_1+1}$. Take $z_{n_1} = \zeta_1$ and, for the value ζ_1, find $z_{n_1 n_2}$. We have $\rho_1 = K\rho\mu^{n_1+1}$, $|z - z_{n_1 n_2}| < K_1 \rho_1 \mu_1^{n_2+2}$, or $R_{n_1 n_2} < K_1 K^{n_2+2} \rho\mu^{(n_1+2)(n_2+2)-1}$. After m such steps we obtain $z_{n_1 \dots n_m}$ with error

$$R_{n_1 \dots n_m} < K_{m-1} K_{m-2}^{n_m+2} \cdots K^{(n_m+2)(n_{m-1}+2)\cdots(n_2+2)} \rho\mu^{[(n_1+2)\cdots(n_m+2)]-1}. \qquad (13)$$

Since all $K < 1$ and $\mu < 1$, it follows that $\lim_{m\to\infty} R_{n_1 \dots n_m} = 0$, and the theorem is proved for this case. It is not difficult to show that the theorem will also be true in the case when zeros occur among the numbers n_k, if $F_\rho/(1 - \rho)F_1 < 1$.

Remarks. 1. For a rough determination of the error, one may put $K_1 = K_2 = \cdots = K$ in (13). Then we obtain

$$R_{n_1, \dots, n_m} < K^{1+(n_m+2)+\cdots+[(n_m+2)\cdots(n_2+2)]} \rho\mu^{[(n_1+2)\cdots(n_m+2)]-1}.$$

In the case $n_1 = n_2 = \cdots = n$, we obtain

$$R_{n_m} < K^{\frac{(n+2)^m-1}{n+1}} \rho \mu^{(n+2)^m-1}.$$

2. In the case $n_1 = n_2 = \cdots = 0$ (Newton' s process), the error estimate for the m-th step will be $R_{0^m} < K^{2^m-1} \rho^{2^m-1} r$, where $K = F/(1-\rho)F_1$, or, in simplified form, $K = F/F_1$.

3. Theorem 2 can also be applied in the case of a non-analytic function having derivatives up to order $n+2$ $(n \geq 0)$, if for the initial value ζ the conditions of Theorem 1 are satisfied up to and including order $n+2$. In this case one must have $n_m \leq n$ $(m = 1, 2, \ldots)$.

VI. **Remarks.** The first approximation was first published in (^{2}), and then was rediscovered by other authors (^{3,4}). Approximations up to third order are found in (^{5,6}). Approximations of arbitrary order, expressed by the determinant (1), are given in ($^{7-9}$). In (^{9}) there is also formula (5) of the present paper. In all the cited works the question of the existence of $\lim_{n\to\infty} z_n$ is not posed; nor is the question of convergence to the root of the generalized Newton process resolved, except for one case: $n_1 = n_2 = \cdots = 1$ (^{10}). As for the question of convergence of Newton' s process itself, there is an extensive literature, which I shall not touch upon here.

Received
10 XII 1956

CITED LITERATURE

[1] M. S. Gornshtein, DAN, 78, No. 2 (1951).
[2] Halley, Phil. Trans. Roy. Soc., 18 (1694).
[3] J. S. Frame, Am. Math. Monthly, 51 (1944).
[4] H. S. Wall, Am. Math. Monthly, 55 (1948).
[5] E. Shroder, Math. Ann., 2 (1870).
[6] J. Kiss, Acta Techn. Acad. Sci. Hung., 7, No. 3-4 (1954).
[7] H. W. Richmond, Lond. Math. Soc., 19, No. 73 (1944).
[8] A. P. Domoryad, Tr. Sredneaziatsk. univ., nov. ser., 7, No. 36 (1954).
[9] H. J. Hamilton, Am. Math. Monthly, 57, No. 8 (1950).
[10] G. S. Salekhov, DAN, 82, No. 4 (1954).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.