ON THE CRITICAL POINTS OF SOME ITERATION METHODS FOR SOLVING ALGEBRAIC EQUATIONS. GLOBAL CONVERGENCE PROPERTIES[1]

Nikolay Kyurkchiev, Anton Iliev

Abstract. *In this work we give sufficient conditions for k-th approximations of the polynomial roots of $f(x)$ when the Maehly–Aberth–Ehrlich, Werner–Borsch–Supan, Tanabe, Improved Borsch–Supan iteration methods fail on the next step. For these methods all non-attractive sets are found. This is a subsequent improvement of previously developed techniques and known facts. The users of these methods can use the results presented here for software implementation in Distributed Applications and Simulation Environments. Numerical examples with graphics are shown.*

Keywords: polynomial roots, critical initial approximations, Maehly–Aberth–Ehrlich method, Werner–Borsch–Supan method, Tanabe method, Improved Borsch–Supan method, divergent sets

2010 Mathematics Subject Classification: 65H05

1. Introduction

Let f be a monic polynomial of degree n,

$$f(x) := x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with simple roots x_i, $i = 1, 2, \ldots, n$.

Let x_i^k, $i = 1, 2, \ldots, n$, be distinct reasonably close approximations of these zeros.

Using logarithmic derivative of the polynomial Maehly [11] (and later, Borsch–Supan [3], Ehrlich [4], Aberth [1], and others) derived the total–step method:

$$(1) \qquad x_i^{k+1} = x_i^k - \frac{f(x_i^k)}{f'(x_i^k) - f(x_i^k)\displaystyle\sum_{j\neq i}^{n} \frac{1}{x_i^k - x_j^k}}, \qquad \begin{array}{l} i = 1, \ldots, n; \\ k = 0, 1, 2, \ldots . \end{array}$$

There are many practical observations described in scientific studies that method (1) is globally convergent for almost every starting point $\mathbf{x}^0 = (x_1^0, \ldots, x_n^0)$ assuming that the components of $\mathbf{x}^0$ are distinct (see Kyurkchiev [9], Atanassova, Kyurkchiev and Yamamoto [2], Kanno, Kyurkchiev and Yamamoto [8].

[1]This paper is partially supported by project of National Science Fund from 2010.

Let us denote

$$\mu_i^k = \frac{f'(x_i^k) - f(x_i^k)\displaystyle\sum_{j\neq i}^{n}\frac{1}{x_i^k - x_j^k}}{\displaystyle\prod_{j\neq i}^{n}(x_i^k - x_j^k)}, \qquad \begin{array}{l} i = 1, \ldots, n, \\ k = 0, 1, 2, \ldots \ . \end{array}$$

The following was shown by Kyurkchiev and Moskona [10]:

Let x_i^{k+1} be determined by (1) for $i = 1, 2, \ldots, n$ and $k = 0, 1, 2, \ldots$, then the following relations are valid

$$\begin{aligned}
&\sum_{i=1}^{n}\mu_i^k x_i^{k+1} = \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k - a_{n-1},\\
&\sum_{i=1}^{n}\mu_i^k x_i^{k+1}\sum_{j\neq i}^{n} x_j^k = \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k\sum_{j\neq i}^{n} x_j^k + \sum_{l<s}^{n} x_l^k x_s^k + a_{n-2},\\
&\ldots\\
&\sum_{i=1}^{n}\mu_i^k x_i^{k+1}\prod_{j\neq i}^{n} x_j^k = \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k\prod_{j\neq i}^{n} x_j^k + (n-1)\prod_{j=1}^{n} x_j^k + (-1)^n a_0.
\end{aligned} \tag{2}$$

If the sequence of approximations x_i^k satisfies

$$\begin{aligned}
&\sum_{i=1}^{n}(\mu_i^k - 1)x_i^k - a_{n-1} = 0,\\
&\sum_{i=1}^{n}(\mu_i^k - 1)x_i^k\sum_{j\neq i}^{n} x_j^k + \sum_{l<s}^{n} x_l^k x_s^k + a_{n-2} = 0,\\
&\ldots\\
&\sum_{i=1}^{n}(\mu_i^k - 1)x_i^k\prod_{j\neq i}^{n} x_j^k + (n-1)\prod_{j=1}^{n} x_j^k + (-1)^n a_0 = 0,
\end{aligned} \tag{3}$$

then $x^{k+1} = (0, \ldots, 0)^t$ and the method (1) is not defined at the $(k+2)$-th approximation step, i.e. for any monic polynomial $f(x)$ of degree n there exists a set $G_f \subset C^n$ such that the method (1), starting from $x^k \in G_f$ does not converge to the roots of $f(x)$.

We observe that, in general, these divergent sets are not the only divergent ones.

Such critical initial conditions for some methods have been considered in Hristov and Kyurkchiev [5], Hristov, Kyurkchiev and Iliev [6], and in other publications cited there.

2. Main Results

2.1. Critical Points of Maehly–Aberth–Ehrlich

It should be noted that the divergent set (3) leads to $x_1^{k+1}= x_2^{k+1} = \cdots = x_n^{k+1} = 0$ for which method (1) will fail.

We observe that the Maehly–Aberth–Ehrlich method also can not be performed at the $(k+2)$-th step if $x_i^{k+1} = x_j^{k+1}$ for some $1 \le i < j \le n$.

The resulting systems of equations (2) can be written in vector form as:

$$A\mathbf{x}^{k+1} = \mathbf{b},$$

where

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} :=$$

$$= \begin{pmatrix} \mu_1^k & \mu_2^k & \dots & \mu_n^k \\ \mu_1^k \sum\limits_{q\neq 1} x_q^k & \mu_2^k \sum\limits_{q\neq 2} x_q^k & \dots & \mu_n^k \sum\limits_{q\neq n} x_q^k \\ \vdots & \vdots & & \vdots \\ \mu_1^k \prod\limits_{q\neq 1} x_q^k & \mu_2^k \prod\limits_{q\neq 2} x_q^k & \dots & \mu_n^k \prod\limits_{q\neq n} x_q^k \end{pmatrix},$$

$$\det A = \prod_{t=1}^{n} \mu_t^k \prod_{i<j}^{n} (x_i^k - x_j^k) \neq 0,$$

$$\mathbf{x}^{k+1} := \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \vdots \\ x_n^{k+1} \end{pmatrix},$$

$$\mathbf{b} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k - a_{n-1} \\ \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k \sum_{j\neq i}^{n} x_j^k + \sum_{l<s}^{n} x_l^k x_s^k + a_{n-2} \\ \vdots \\ \sum_{i=1}^{n}(\mu_i^k - 1)x_i^k \prod_{j\neq i}^{n} x_j^k + (n-1)\prod_{j=1}^{n} x_j^k + (-1)^n a_0 \end{pmatrix}.$$

We denote

$$\Delta_{sj} := \begin{vmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{s-11} & \dots & a_{s-1j-1} & a_{s-1j+1} & \dots & a_{s-1n} \\ a_{s+11} & \dots & a_{s+1j-1} & a_{s+1j+1} & \dots & a_{s+1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj+1} & \dots & a_{nn} \end{vmatrix}.$$

We have the following theorem.

Theorem. *Suppose that for some $1 \le i < j \le n$, the sequence of approximations $x_1^k, \dots, x_n^k$ satisfies the condition*

$$\sum_{s=1}^{n}(-1)^s b_s \left((-1)^i \Delta_{si} - (-1)^j \Delta_{sj}\right) = 0. \tag{4}$$

Then $x_i^{k+1} = x_j^{k+1}$, and thus, the $(k+2)$-th step of the Maehly–Aberth–Ehrlich method cannot be performed.

The proof follows the ideas given in [5] and [6].

We will outline briefly the proof of this theorem which is the “model” substantially used in the treatment of problems related to the study of global properties of the iterative procedures.

We denote

$$A_j := \begin{pmatrix} a_{11} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j-1} & b_2 & a_{2j+1} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj-1} & b_n & a_{nj+1} & \dots & a_{nn} \end{pmatrix}.$$

Clearly,

$$\det A_j = \sum_{s=1}^{n} (-1)^{j+s} b_s \Delta_{sj}, \;\; j = 1, 2, \dots, n. \tag{5}$$

We note that, if $x_i^{k+1} = x_j^{k+1}$, then by Cramer's formula

$$\frac{\det A_i}{\det A} = \frac{\det A_j}{\det A}.$$

($\det A \neq 0$), from (5) we have

$$\sum_{s=1}^{n} (-1)^{s+i} b_s \Delta_{si} - \sum_{s=1}^{n} (-1)^{s+j} b_s \Delta_{sj} = 0$$

and we arrive to the formula (4), which completes the proof of the theorem.

The set D_f of the non-attractive starting points is the set of points satisfying equations (4).

The facts for sufficient conditions for k-th approximations of the zeros of $f(x)$ under which the Maehly–Aberth–Ehrlich method fails on the next step are given in [14] by Valchanov, Iliev, Kyurkchiev.

2.2. Critical Points of Werner–Borsch–Supan Method

By Werner's notation [15]

$$\sigma_i^k = \frac{f(x_i^k)}{\prod_{j \neq i}^{n} (x_i^k - x_j^k)}, \;\; i = 1, \dots, n; \; k = 0, 1, \dots$$

we consider the following method (Werner, or Borsch-Supan method)

$$x_i^{k+1} = x_i^k - \frac{\sigma_i^k}{1 + \sum_{j \neq i}^{n} \frac{\sigma_j^k}{x_i^k - x_j^k}}, \qquad \begin{array}{l} i = 1, \dots, n; \\ k = 0, 1, 2, \dots . \end{array} \tag{6}$$

Suppose that for some $1 \le i < j \le n$, the sequence of approximations $x_1^k, \dots, x_n^k$ satisfies the condition

$$x_i^k - \frac{\sigma_i^k}{1 + \sum\limits_{s \neq i}^{n} \frac{\sigma_s^k}{x_i^k - x_s^k}} = x_j^k - \frac{\sigma_j^k}{1 + \sum\limits_{s \neq j}^{n} \frac{\sigma_s^k}{x_j^k - x_s^k}} \tag{7}$$

then $x_i^{k+1} = x_j^{k+1}$, and thus, the $(k+2)$-th step of the Werner–Borsch-Supan method (6) cannot be performed.

The set D_f of the non-attractive starting points is the set of points satisfying equations (7).

Numerical Examples

1. For illustration, we consider non-attractive set D_f in the example of the equation

$$f(x) = x^2 + x + 1 = 0. \tag{8}$$

The non-attractive set D_f, is given by (see (7))

$$D_f := \frac{x^2 + x + 1}{x^2 - 2xy - y - 1} + \frac{y^2 + y + 1}{y^2 - 2xy - x - 1} - 1 = 0$$

(where $x = x_1^k$, $y = x_2^k$) and displayed in Fig. 1 (ContourPlot), Fig. 2 (Plot3D).

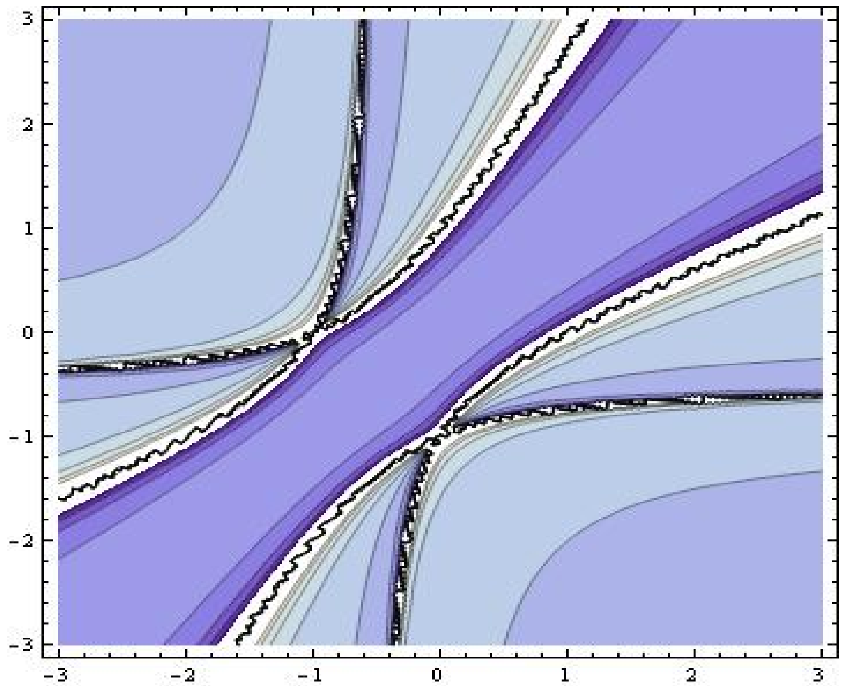

Fig. 1

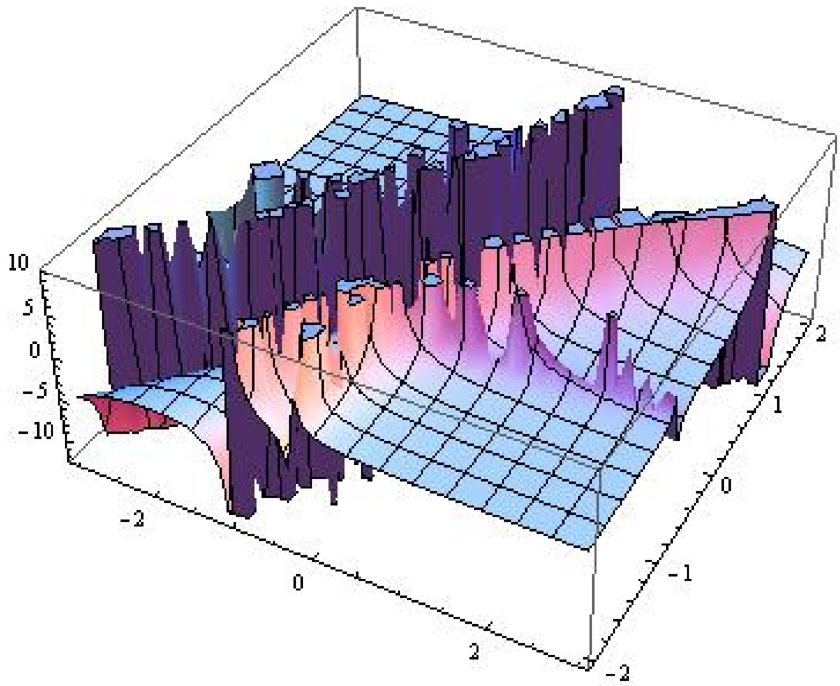

Fig. 2

2. For the equation

$$f(x) = x^3 + 1 = 0 \tag{9}$$

the non-attractive set D_f, is given by (see (7))

$$D_f := -1 + \frac{(x^3+1)(x-z)(y-z)}{(x-y)^2(x-z)^2(y-z) - (y^3+1)(x-z)^2 + (z^3+1)(x-y)^2} + \frac{(y^3+1)(x-z)(y-z)}{(x-y)^2(y-z)^2(x-z) - (x^3+1)(y-z)^2 + (z^3+1)(x-y)^2} = 0,$$

(where $x = x_1^k$, $y = x_2^k$, $z = x_3^k$) and displayed in Fig. 3 (ContourPlot3D).

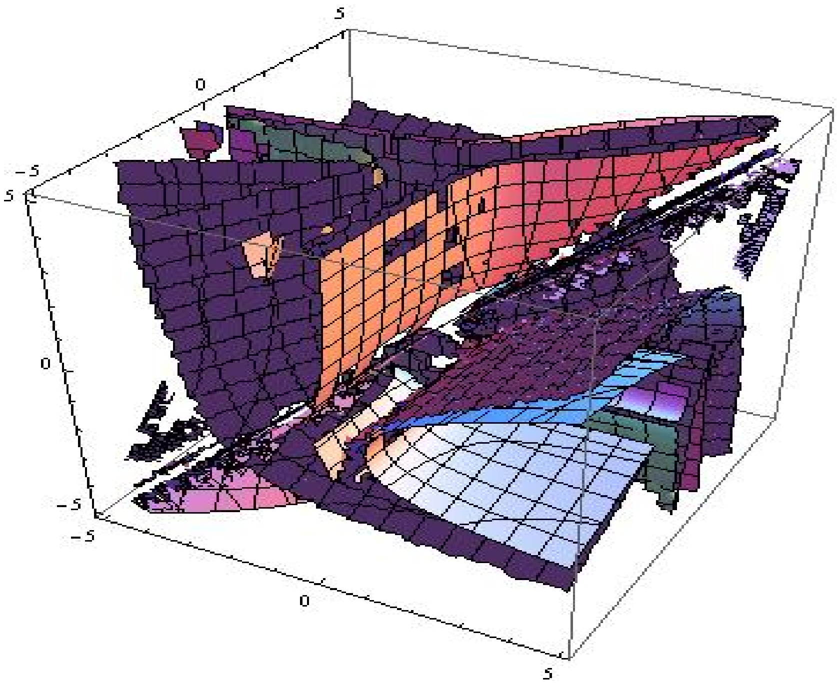

Fig. 3

2.3. Critical Points of Tanabe Method

By Werner's notation [15]

$$\sigma_i^k = \frac{f(x_i^k)}{\prod_{j\neq i}^{n}(x_i^k - x_j^k)}\,, \ \ i = 1, \dots, n;\ k = 0, 1, \dots$$

we consider the following method (Tanabe method [13]):

$$x_i^{k+1} = x_i^k - \sigma_i^k \left(1 - \sum_{j\neq i}^{n} \frac{\sigma_j^k}{x_i^k - x_j^k}\right), \qquad \begin{array}{l} i = 1, \dots, n; \\ k = 0, 1, 2, \dots\ . \end{array} \tag{10}$$

The following relations are valid:

$$
\begin{aligned}
&\sum_{i=1}^{n} x_i^{k+1} = -a_{n-1},\\
&\sum_{i=1}^{n} x_i^{k+1} \sum_{j\neq i}^{n} x_j^k = \sum_{l<s}^{n} x_l^k x_s^k - \sum_{l<s}^{n} \sigma_l^k \sigma_s^k + a_{n-2},\\
(11)\qquad &\sum_{i=1}^{n} x_i^{k+1} \sum_{l,s\neq i, l<s}^{n} x_l^k x_s^k = 2\sum_{l<s<t}^{n} x_l^k x_s^k x_t^k - \sum_{l<s}^{n} \sigma_l^k \sigma_s^k \sum_{p\neq l,s}^{n} x_p^k - a_{n-3},\\
&\ldots\\
&\sum_{i=1}^{n} x_i^{k+1} \prod_{j\neq i}^{n} x_j^k = (n-1)\prod_{j=1}^{n} x_j^k - \sum_{l<s}^{n} \sigma_l^k \sigma_s^k \prod_{p\neq l,s}^{n} x_p^k + (-1)^n a_0.
\end{aligned}
$$

The proof of (11) is based on the divided difference properties and Euler's identity:

$$
\sum_{i=1}^{n} \frac{(x_i^k)^t}{\prod_{j\neq i}^{n}(x_i^k - x_j^k)} = \begin{cases} \sum_{i=1}^{n} x_i^k, & t = n,\\ 1, & t = n-1,\\ 0, & 0 \le t \le n-2. \end{cases}
$$

The resulting systems of equations (11) can be written in vector form:

$$
A\mathbf{x}^{k+1} = \mathbf{b},
$$

where

$$
A := \begin{pmatrix} 1 & 1 & \ldots & 1\\ \sum_{j\neq 1} x_j^k & \sum_{j\neq 2} x_j^k & \ldots & \sum_{j\neq n} x_j^k\\ \vdots & \vdots & \ddots & \vdots\\ \prod_{j\neq 1} x_j^k & \prod_{j\neq 2} x_j^k & \ldots & \prod_{j\neq n} x_j^k \end{pmatrix},
$$

$$
\det A = \prod_{i<j}^{n} (x_i^k - x_j^k) \neq 0,
$$

$$\mathbf{b} := \begin{pmatrix} -a_{n-1} \\ \\ \sum\limits_{l<s} x_l^k x_s^k - \sum\limits_{l<s} \sigma_l^k \sigma_s^k + a_{n-2} \\ \vdots \\ (n-1)\prod\limits_{j=1} x_j^k - \sum\limits_{l<s} \sigma_l^k \sigma_s^k \prod\limits_{p\neq l,s} x_p^k + (-1)^n a_0 \end{pmatrix}.$$

It is easy to conform that if (4) is satisfied when we choose the vector $\mathbf{b}$ by the way described here, then $x_i^{k+1} = x_j^{k+1}$, and thus, the $(k+2)$-th step of the Tanabe's method cannot be performed.

Suppose that for some $1 \leq i < j \leq n$, the sequence of approximations $x_1^k, \dots, x_n^k$ satisfies the condition

$$x_i^k - \sigma_i^k \left(1 - \sum_{s\neq i}^{n} \frac{\sigma_s^k}{x_i^k - x_s^k}\right) = x_j^k - \sigma_j^k \left(1 - \sum_{s\neq j}^{n} \frac{\sigma_s^k}{x_j^k - x_s^k}\right) \tag{12}$$

then $x_i^{k+1} = x_j^{k+1}$, and thus, the $(k+2)$-th step of the Tanabe's method (11) cannot be performed.

The set D_f of the non-attractive starting points is the set of points satisfying equations (12).

Numerical Example

3. For illustration, we consider non-attractive set D_f in the example of the equation

$$f(x) = x^2 + x + 1 = 0. \tag{13}$$

The non-attractive set D_f, is given by (see (12))

$$D_f := (x-y)^4 - (x-y)^2(x^2+y^2+x+y+2) - 2(x^2+x+1)(y^2+y+1) = 0$$

(where $x = x_1^k$, $y = x_2^k$) and displayed in Fig. 3 (Plot3D).

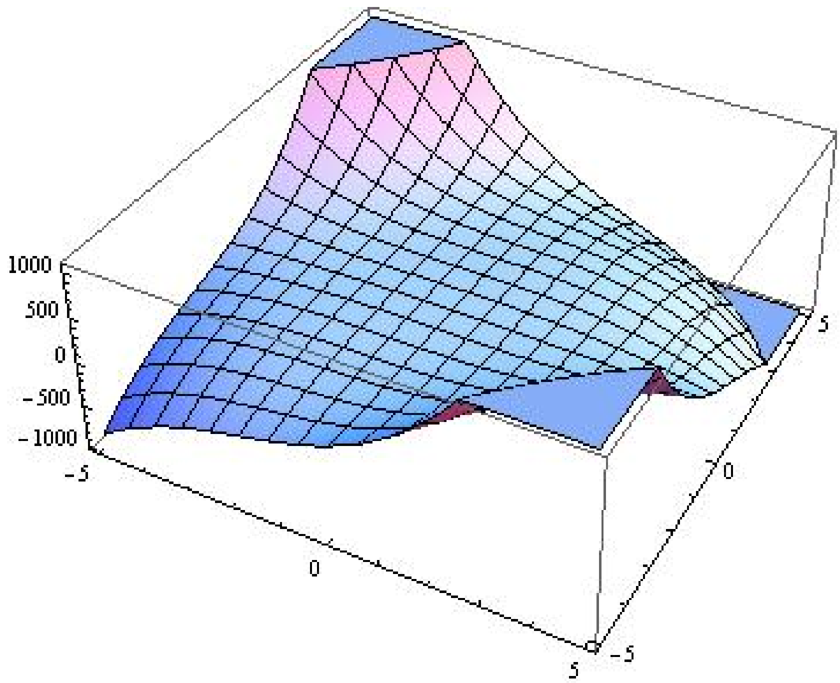

Fig. 3

2.4. Critical Points of Improved Borsch-Supan Method

By Werner's notation [15]

$$\sigma_i^k = \frac{f(x_i^k)}{\prod\limits_{j \neq i}^{n} (x_i^k - x_j^k)}, \quad i = 1, \ldots, n; \; k = 0, 1, \ldots$$

we consider the following method (improved Borsch–Supan method):

$$x_i^{k+1} = x_i^k - \frac{\sigma_i^k}{1 + \sum\limits_{j \neq i}^{n} \frac{\sigma_j^k}{x_i^k - x_j^k - \sigma_i^k}}, \qquad \begin{array}{l} i = 1, \ldots, n; \\ k = 0, 1, 2, \ldots \; . \end{array} \tag{14}$$

Suppose that for some $1 \leq i < j \leq n$, the sequence of approximations $x_1^k, \ldots, x_n^k$ satisfies the condition

$$x_i^k - \frac{\sigma_i^k}{1 + \sum\limits_{s \neq i}^{n} \frac{\sigma_s^k}{x_i^k - x_s^k - \sigma_i^k}} = x_j^k - \frac{\sigma_j^k}{1 + \sum\limits_{s \neq j}^{n} \frac{\sigma_s^k}{x_j^k - x_s^k - \sigma_j^k}} \tag{15}$$

then $x_i^{k+1} = x_j^{k+1}$, and thus, the $(k+2)$-th step of the improved Werner–Borsch-Supan method (14) cannot be performed.

The set D_f of the non-attractive starting points is the set of points satisfying equations (15).

Numerical Example

4. For illustration, we consider non-attractive set D_f in the example of the equation

$$f(x) = x^2 + x + 1 = 0. \tag{16}$$

The non-attractive set D_f, is given by (see (15))

$D_f := (x-y)^4 - 2(x-y)^2(x^2+y^2+x+y+2) + (x^2+x+1)^2 + (y^2+y+1)^2 = 0$

(where $x = x_1^k$, $y = x_2^k$) and displayed in Fig. 4 (Plot3D).

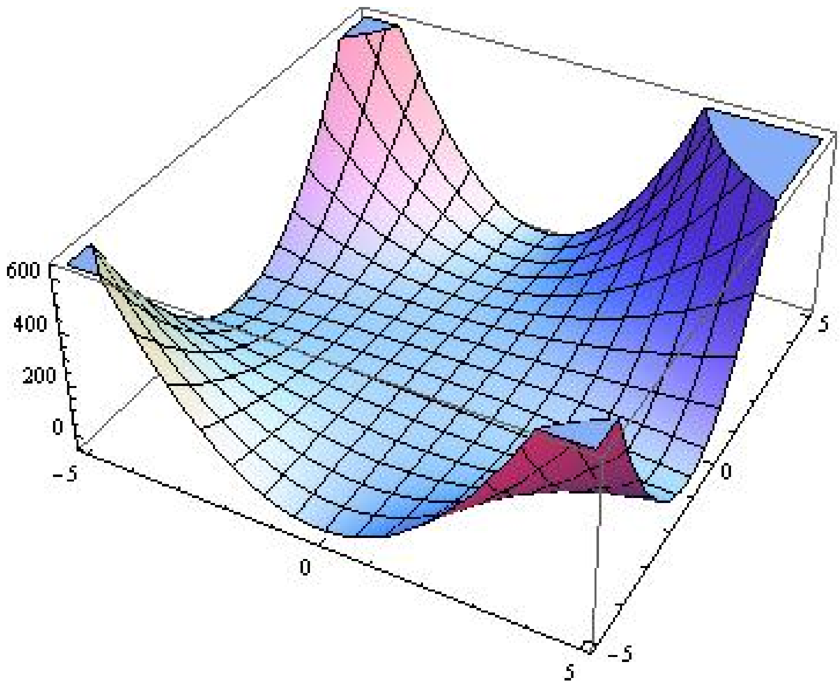

Fig. 4

3. Concluding Remarks

In connection to this topic many scientists have investigated issues connected with practical software implementations of these methods: K. Dochev, J. Dvorchuk, V. Popov, Bl. Sendov, P. Barnev, J. Herzberger, G. Alefeld, Sh. Zheng, T. Yamamoto, Q. Fang, S. Kanno, L. Atanassova, M. Petkovic, V. Hristov, P. Marinov, K. Mahdi, L. Petkovic, N. Kyurkchiev, E. Moskona, A. Iliev, N. Valchanov, A. Andreev, S. Tashev, S. Markov, N. Dimitrova, G. Nedzhibov, M. Petkov, Kh. Semerdzhiev, P. Proinov, I. Makrelov, S. Tamburov, E. Angelova, I. Angelov and many others.

We note that D_f is an algebraic manifold of high degree. The results for D_f obtained here can be accepted in the direction that they are all non-attractive sets that leads to divergent of Maehly–Aberth–Ehrlich, Werner–Borsch–Supan, Tanabe, Improved Borsch–Supan methods eventually with exception of elements which belongs eventually to strongly connected cyclic graph.

The accumulation of "parasite digits" on some iteration step can lead to obtaining such x_i^k, lying on the already described non-attractive sets D_f.

Convergence of any iterative method for finding zeros [12] of a given function is connected with the distances between its roots. If these zeros are sufficiently separated, all iteration schemes demonstrate good convergence. In the case of very close zeros ("clusters of zeros") all algorithms fail or work with big efforts.

References

[1] O. Aberth: Iteration method for finding all zeros of a polynomial simultaneously, *Math. Comput.*, **27**, (1973), 339–344.

[2] L. Atanassova, N. Kyurkchiev, T. Yamamoto: Methods for computing all roots of a polynomial simultaneously - known results and open problems, *Computing [Suppl.]*, **16**, (2002), 23–35.

[3] W. Borsch–Supan: A posteriori error bound for the zeros of polynomials, *Numer. Math.*, **5**, (1963), 380–398.

[4] L. Ehrlich: A modified Newton method for polynomials, *Comm. ACM*, **10**, (1967), 107–108.

[5] V. Hristov, N. Kyurkchiev: A note on the globally convergent properties of the Weierstrass-Dochev method, *Approximation Theory, A volume dedicated to Bl. Sendov (ed. B. Boyanov)*, DARBA, Sofia, (2002), 231–240.

[6] V. Hristov, N. Kyurkchiev, A. Iliev: *Global convergence properties of the SOR-Weierstrass method*, LNCS: Numerical methods and applications (I. Dimov, S. Dimova, N. Kolkovska eds.), Springer-Verlag GmbH Berlin Heidelberg, 6046, (2010), (accepted).

[7] A. Iliev, N. Kyurkchiev: *Nontrivial Methods in Numerical Analysis. Selected Topics in Numerical Analysis*, LAP LAMBERT Academic Publishing GmbH & Co. KG, (2010), (accepted).

[8] S. Kanno, N. Kyurkchiev, T. Yamamoto: On some methods for the simultaneous determination of polynomial zeros, *Japan J. Indust. Appl. Math.*, **13**, (1996), 267–288.

[9] N. Kyurkchiev: *Initial approximations and root finding methods*, Mathematical Research, Wiley-VCH Verlag Berlin GmbH, Berlin, 104, (1998).

[10] N. Kyurkchiev, E. Moskona: On the critical points of Aberth's method, *In: Mathematics and Education in Mathematics. Proc. of the XXVI Conference of the Union of the Bulgarian Mathematicians*, (1997), 131–138.

[11] H. Maehly: Zur iterativen Auflosung algebraischer Gleichungen,*Z. Angew. Math. Phys.*, **5**, (1954), 260–263.

[12] M. Petkovic, D. Herceg, S. Ilic: *Point Estimation Theory and its Applications*, Institute of Mathematics, Novi Sad, (1997).

[13] K. Tanabe: Behavior of the sequence around multiple zeros generated by some simultaneous methods for solving algebraic equations, *Tech. Rep. Inf. Process. Numer. Anal.* 4–2, (1983), 1–6.

[14] N. Valchanov, A. Iliev, N. Kyurkchiev: On the critical points of Maehly – Aberth – Ehrlich method. Global convergence properties, *Int. J. Pure Appl. Math.*, (2010), (accepted).

[15] W. Werner: On the simultaneous determination of polynomial roots, *Springer Lecture Notes*, **953**, (1982), 188–202.

Nikolay Kyurkchiev, Anton Iliev
Faculty of Mathematics and Informatics
University of Plovdiv
236, Bulgaria Blvd.
4003 Plovdiv, Bulgaria
e-mail: nkyurk@uni-plovdiv.bg, aii@uni-plovdiv.bg