Chapter 3
Point Estimation of Simultaneous Methods

In this chapter, we are primarily interested in the construction of computationally verifiable initial conditions and the corresponding convergence analysis of the simultaneous methods presented in Sect. 1.1. These quantitative conditions predict the immediate appearance of the guaranteed and fast convergence of the considered methods. Two original procedures, based on (1) suitable localization theorems for polynomial zeros and (2) the convergence of error sequences, are applied to the most frequently used iterative methods for finding polynomial zeros.

3.1 Point Estimation and Polynomial Equations

As mentioned in Chap. 2, one of the most important problems in solving nonlinear equations is the construction of such initial conditions which provide both the guaranteed and fast convergence of the considered numerical algorithm. Smale's approach from 1981, known as "point estimation theory," examines convergence conditions in solving an equation $f(z) = 0$ using only the information of f at the initial point z_0. In the case of monic algebraic polynomials of the form

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

which are the main subject of our investigation in this chapter and Chaps. 4 and 5, initial conditions should be some functions of polynomial coefficients $\boldsymbol{a} = (a_0, \ldots, a_{n-1})$, its degree n, and initial approximations $\boldsymbol{z}^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)})$. A rather wide class of initial conditions can be represented by the inequality of general form

$$\phi(\boldsymbol{z}^{(0)}, \boldsymbol{a}, n) < 0. \tag{3.1}$$

M. Petković, *Point Estimation of Root Finding Methods*. Lecture Notes in Mathematics 1933,

It is well known that the convergence of any iterative method for finding zeros of a given function is strongly connected with the distribution of its zeros. If these zeros are well separated, almost all algorithms show mainly good convergence properties. Conversely, in the case of very close zeros ("clusters of zeros"), almost all algorithms either fail or work with a big effort. From this short discussion, it is obvious that a measure of separation of zeros should be taken as an argument of the function ϕ given in (3.1). Since the exact zeros are unknown, we restrict ourselves to deal with the minimal distance among initial approximations $d^{(0)} = \min_{j \neq i} |z_i^{(0)} - z_j^{(0)}|$. Furthermore, the closeness of initial approximations to the wanted zeros is also an important parameter, which influences the convergence of the applied method. A measure of this closeness can be suitably expressed by a quantity of the form $h(z) = |P(z)/Q(z)|$, where $Q(z)$ does not vanish when z lies in the neighborhood $\Lambda(\zeta)$ of any zero ζ of P. For example, in the case of simple zeros of a polynomial, the choice

$$Q(z) = P'(z), \quad Q(z) = \prod_{\substack{j=1 \\ j \neq i}}^{n} (z - z_j) \quad \text{or}$$

$$|Q(z)| = |P'(z)|^{-1} \sup_{k>1} \left| \frac{P^{(k)}(z)}{k! P'(z)} \right|^{1/(k-1)} \quad \text{(see Sect. 2.1)}$$

gives satisfactory results. Let us note that, considering algebraic equations, the degree of a polynomial n appears as a natural parameter in (3.1). Therefore, instead of (3.1), we can take the inequality of the form

$$\varphi(h^{(0)}, d^{(0)}, n) < 0, \tag{3.2}$$

where $h^{(0)}$ depends on P and Q at the initial point $\boldsymbol{z}^{(0)}$.

Let $\boldsymbol{I}_n := \{1, \ldots, n\}$ be the index set. For $i \in \boldsymbol{I}_n$ and $m = 0, 1, \ldots$, let us introduce the quantity

$$W_i^{(m)} = \frac{P\big(z_i^{(m)}\big)}{\prod_{\substack{j=1 \\ j \neq i}}^{n} \big(z_i^{(m)} - z_j^{(m)}\big)} \quad (i \in \boldsymbol{I}_n, \ m = 0, 1, \ldots), \tag{3.3}$$

which is often called Weierstrass' correction since it appeared in Weierstrass' paper [187]. In [178], D. Wang and Zhao improved Smale's result for Newton's method and applied it to the Durand–Kerner's method for the simultaneous determination of polynomial zeros (see Sect. 2.5, (2.38), and (2.39)). Their approach led in a natural way to an initial condition of the form

$$w^{(0)} \leq c_n \, d^{(0)}, \tag{3.4}$$

where

$$w^{(0)} = \max_{\substack{1\le i,j\le n \\ i\ne j}} |W_i^{(0)}|, \quad d^{(0)} = \min_{\substack{1\le i,j\le n \\ i\ne j}} |z_i^{(0)} - z_j^{(0)}|.$$

A completely different approach presented in [112] for the same method also led to the condition of the form (3.4). In both cases, the quantity c_n was of the form $c_n = 1/(an+b)$, where a and b are suitably chosen positive constants. It turned out that initial conditions of this form are also suitable for other simultaneous methods for solving polynomial equations, as shown in the subsequent papers [5], [110], [112], [114]–[117], [119]–[121], [123], [132], [133], [136], [137], [140], [150], [151], [178], [195] and the books [20] and [118]. For these reasons, in the convergence analysis of simultaneous methods considered in this book, we will also use initial conditions of the form (3.4). We note that (3.4) is a special case of the condition (3.2). The quantity c_n, which depends only on the polynomial degree n, will be called the *inequality factor*, or the i-factor for brevity. We emphasize that during the last years, special attention has been paid to the increase of the i-factor c_n for the following obvious reason. From (3.4), we notice that a greater value of c_n allows a greater value of $|W_i^{(0)}|$. This means that cruder initial approximations can be chosen, which is of evident interest in practical realizations of numerical algorithms.

The proofs of convergence theorems of the simultaneous methods investigated in this chapter and Chaps. 4 and 5 are based on the inductive arguments. It turns out that the inequality of the form (3.4), with a specific value of c_n depending on the considered method, appears as a connecting link in the chain of inductive steps. Namely, $w^{(0)} \le c_n d^{(0)} \;\Rightarrow\; w^{(1)} \le c_n d^{(1)}$, and one may prove by induction that $w^{(0)} \le c_n d^{(0)}$ implies $w^{(m)} \le c_n d^{(m)}$ for all $m = 0, 1, 2, \ldots$.

In this chapter, we discuss the best possible values of the i-factor c_n appearing in the initial condition (3.4) for some efficient and frequently used iterative methods for the simultaneous determination of polynomial zeros. The reader is referred to Sect. 1.1 for the characteristics (derivation, historical notes, convergence speed) of these methods. We study the choice of "almost optimal" factor c_n. The notion "almost optimal" i-factor arises from (1) the presence of a system of (say) k inequalities and (2) the use of computer arithmetic of finite precision:

(1) In the convergence analysis, it is necessary to provide the validity of k substantial successive inequalities $g_1(c_n) \ge 0, \ldots, g_k(c_n) \ge 0$ (in this order), where all $g_i(c_n)$ are monotonically decreasing functions of c_n (see Fig. 3.1). The optimal value c_n should be determined as the unique solution of the corresponding equations $g_i(c_n) = 0$. Unfortunately, all equations cannot be satisfied simultaneously and we are constrained to find such c_n which makes the inequalities $g_i(c_n) \ge 0$ as sharp as possible. Since $g_i(c_n) \ge 0$ succeeds $g_j(c_n) \ge 0$ for $j < i$, we first find c_n so that the inequality $g_1(c_n) \ge 0$ is as sharp as possible and check the validity

of all remaining inequalities $g_2(c_n) \geq 0, \ldots, g_k(c_n) \geq 0$. If some of them are not valid, we decrease c_n and repeat the process until all inequalities are satisfied. For demonstration, we give a particular example on Fig. 3.1. The third inequality $g_3(c_n) \geq 0$ is not satisfied for $c_n^{(1)}$, so that c_n takes a smaller value $c_n^{(2)}$ satisfying all three inequalities. In practice, the choice of c_n is performed iteratively, using a programming package, in our book *Mathematica* 6.0.

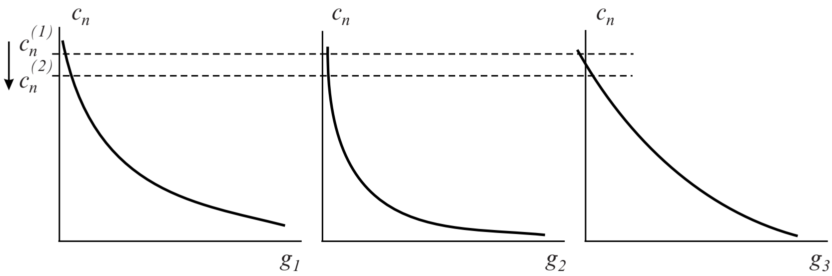

Fig. 3.1 The choice of i-factor c_n iteratively

(2) Since computer arithmetic of finite precision is employed, the optimal value (the exact solution of $g_i(c_n) = 0$, if it exists for some i) cannot be represented exactly, so that c_n should be decreased for a few bits to satisfy the inequalities $g_i(c_n) > 0$. The required conditions (in the form of inequalities $g_i(c_n) \geq 0$) are still satisfied with great accuracy. We stress that this slight decrease of the i-factor c_n with respect to the optimal value is negligible from a practical point of view. For this reason, the constants a and b appearing in $c_n = 1/(an + b)$ are rounded for all methods considered in this book.

The entries of c_n, obtained in this way and presented in this chapter, are increased (and, thus, improved) compared with those given in the literature, which means that newly established initial conditions for the guaranteed convergence of the considered methods are weakened (see Fig. 3.3).

We note that all considerations in this book are given for $n \geq 3$, taking into account that algebraic equations of the order ≤ 2 are trivial and their numerical treatment is unnecessary. In our analysis, we will sometimes omit the iteration index m, and new entries in the later $(m+1)$th iteration will be additionally stressed by the symbol $\widehat{}$ ("hat"). For example, instead of

$$z_i^{(m)}, z_i^{(m+1)}, W_i^{(m)}, W_i^{(m+1)}, d^{(m)}, d^{(m+1)}, N_i^{(m)}, N_i^{(m+1)}, \quad \text{etc.},$$

we will write

$$z_i, \hat{z}_i, W_i, \widehat{W}_i, d, \hat{d}, N_i, \widehat{N}_i.$$

According to this, we denote

$$w = \max_{1 \le i \le n} |W_i|, \quad \widehat{w} = \max_{1 \le i \le n} |\widehat{W}_i|.$$

This denotation will also be used in the subsequent study in Chaps. 4 and 5.

3.2 Guaranteed Convergence: Correction Approach

In this chapter, we present two procedures in the study of the guaranteed convergence of simultaneous methods (1) the approach based on iterative corrections and (2) the approach based on convergent sequences. Both schemes will be applied to the most frequently used simultaneous zero-finding methods in considerable details.

Applying the first method (1), we will deal with a real function $t \mapsto g(t)$ defined on $(0,1)$ by

$$g(t) = \begin{cases} 1 + 2t, & 0 < t \le \dfrac{1}{2} \\ \dfrac{1}{1-t}, & \dfrac{1}{2} < t < 1 \end{cases}$$

The minorizing function of $g(t)$ on $(0,1)$ is given in the following lemma whose proof is elementary.

Lemma 3.1. *Let*

$$s_m\left(t\right) = \sum_{i=0}^{m} t^i + t^m \quad \left(t \in \left(0,1\right),\ m = 1,2,\ldots\right).$$

Then, $s_m\left(t\right) < g\left(t\right)$.

Most of the iterative methods for the simultaneous determination of simple zeros of a polynomial can be expressed in the form

$$z_i^{(m+1)} = z_i^{(m)} - C_i\big(z_1^{(m)}, \ldots, z_n^{(m)}\big) \quad (i \in \boldsymbol{I}_n,\ m = 0,1,\ldots), \tag{3.5}$$

where $z_1^{(m)}, \ldots, z_n^{(m)}$ are some distinct approximations to simple zeros $\zeta_1, \ldots, \zeta_n$, respectively, obtained in the mth iterative step by the method (3.5). In what follows, the term

$$C_i^{(m)} = C_i\big(z_1^{(m)}, \ldots, z_n^{(m)}\big) \quad (i \in \boldsymbol{I}_n)$$

will be called the *iterative correction* or simply the *correction*.

Let $\Lambda(\zeta_i)$ be a sufficiently close neighborhood of the zero ζ_i $(i \in \boldsymbol{I}_n)$. In this book, we consider a class of iterative methods of the form (3.5) with corrections C_i which can be expressed as

$$C_i(z_1, \dots, z_n) = \frac{P(z_i)}{F_i(z_1, \dots, z_n)} \quad (i \in \boldsymbol{I}_n), \tag{3.6}$$

where the function $(z_1, \dots, z_n) \mapsto F_i(z_1, \dots, z_n)$ satisfies the following conditions for each $i \in \boldsymbol{I}_n$ and distinct approximations $z_1, \dots, z_n$:

1° $F_i(\zeta_1, \dots, \zeta_n) \neq 0$,

2° $F_i(z_1, \dots, z_n) \neq 0$ for any $(z_1, \dots, z_n) \in \Lambda(\zeta_1) \times \cdots \times \Lambda(\zeta_n) =: Y$,

3° $F_i(z_1, \dots, z_n)$ is continuous in $\mathbb{C}^n$.

Starting from mutually disjoint approximations $z_1^{(0)}, \dots, z_n^{(0)}$, the iterative method (3.5) produces n sequences of approximations $\{z_i^{(m)}\}$ $(i \in \boldsymbol{I}_n)$ which, under certain convenient conditions, converge to the polynomial zeros. Indeed, if we find the limit values

$$\lim_{m \to \infty} z_i^{(m)} = \zeta_i \quad (i \in \boldsymbol{I}_n),$$

then having in mind (3.6) and the conditions 1°–3°, we obtain from (3.5)

$$\begin{aligned} 0 &= \lim_{m \to \infty} \big(z_i^{(m)} - z_i^{(m+1)}\big) = \lim_{m \to \infty} C_i\big(z_1^{(m)}, \dots, z_n^{(m)}\big) \\ &= \lim_{m \to \infty} \frac{P\big(z_i^{(m)}\big)}{F_i\big(z_1^{(m)}, \dots, z_n^{(m)}\big)} = \frac{P(\zeta_i)}{F_i(\zeta_1, \dots, \zeta_n)} \quad (i \in \boldsymbol{I}_n). \end{aligned}$$

Hence $P(\zeta_i) = 0$, i.e., ζ_i is a zero of the polynomial P.

Theorem 3.1 has the key role in our convergence analysis of simultaneous methods presented in this section and Chap. 4 (see M. Petković, Carstensen, and Trajković [112]).

Theorem 3.1. *Let the iterative method* (3.5) *have the iterative correction of the form* (3.6) *for which the conditions* 1°–3° *hold, and let* $z_1^{(0)}, \dots, z_n^{(0)}$ *be distinct initial approximations to the zeros of* P. *If there exists a real number* $\beta \in (0, 1)$ *such that the following two inequalities*

(i) $\big|C_i^{(m+1)}\big| \le \beta \big|C_i^{(m)}\big| \quad (m = 0, 1, \dots)$,

(ii) $\big|z_i^{(0)} - z_j^{(0)}\big| > g(\beta)\Big(\big|C_i^{(0)}\big| + \big|C_j^{(0)}\big|\Big) \quad (i \neq j, \quad i, j \in \boldsymbol{I}_n)$

are valid, then the iterative method (3.5) *is convergent.*

Proof. Let us define disks $D_i^{(m)} := \{z_i^{(m+1)}; |C_i^{(m)}|\}$ for $i \in \boldsymbol{I}_n$ and $m = 0, 1, \ldots$, where $z_i^{(m+1)}$ and $C_i^{(m)}$ are approximations and corrections appearing in (3.5). Then for a fixed $i \in \boldsymbol{I}_n$, we have

$$\begin{aligned} D_i^{(m)} &= \{z_i^{(m)} - C_i^{(m)}; |C_i^{(m)}|\} = \{z_i^{(m-1)} - C_i^{(m-1)} - C_i^{(m)}; |C_i^{(m)}|\} = \cdots \\ &= \{z_i^{(0)} - C_i^{(0)} - C_i^{(1)} - \cdots - C_i^{(m)}; |C_i^{(m)}|\} \subset \{z_i^{(0)}; r_i^{(m)}\}, \end{aligned}$$

where

$$r_i^{(m)} = |C_i^{(0)}| + \cdots + |C_i^{(m-1)}| + 2|C_i^{(m)}|.$$

Using (i), we find $|C_i^{(k)}| \le \beta^k |C_i^{(0)}|$ ($k = 1, 2, \ldots,\ \beta < 1$) so that, according to Lemma 3.1 and the definition of the function $g(t)$,

$$r_i^{(m)} \le |C_i^{(0)}|(1 + \beta + \cdots + \beta^m + \beta^m) < g(\beta)|C_i^{(0)}|.$$

Therefore, for each $i \in \boldsymbol{I}_n$, we have the inclusion

$$D_i^{(m)} \subset S_i := \{z_i^{(0)}; g(\beta)|C_i^{(0)}|\},$$

which means that the disk S_i contains all the disks $D_i^{(m)}$ ($m = 0, 1, \ldots$). In regard to this and the definition of disks $D_i^{(m)}$, we can illustrate the described situation by Fig. 3.2.

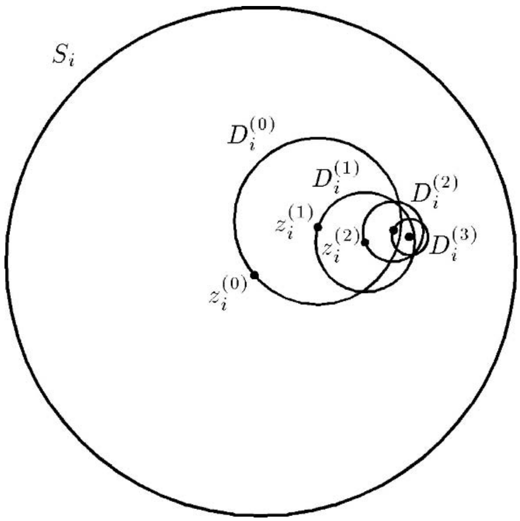

Fig. 3.2 Inclusion disk S_i contains all disks $D_i^{(m)}$

The sequence $\{z_i^{(m)}\}$ of the centers of the disks $D_i^{(m)}$ forms a Cauchy's sequence in the disk $S_i \supset D_i^{(m)}$ ($m = 0, 1, \ldots$). Since the metric subspace S_i

is complete (as a closed set in $\mathbb{C}$), there exists a unique point $z_i^* \in S_i$ such that

$$z_i^{(m)} \to z_i^* \quad \text{as} \quad m \to \infty \quad \text{and} \quad z_i^* \in S_i.$$

Since

$$z_i^{(m+1)} = z_i^{(m)} - \frac{P(z_i^{(m)})}{F(z_1^{(m)}, \ldots, z_n^{(m)})}$$

and $F(z_1^{(m)}, \ldots, z_n^{(m)})$ does not vanish whenever $(z_1, \ldots, z_n) \in Y$, there follows

$$\begin{aligned} |P(z_i^{(m)})| &= |F(z_1^{(m)}, \ldots, z_n^{(m)})(z_i^{(m+1)} - z_i^{(m)})| \\ &\leq |F(z_1^{(m)}, \ldots, z_n^{(m)})|\,|z_i^{(m+1)} - z_i^{(m)}|. \end{aligned}$$

Taking the limit when $m \to \infty$, we obtain

$$|P(z_i^*)| \leq \lim_{m\to\infty} |F(z_1^{(m)}, \ldots, z_n^{(m)})| \lim_{m\to\infty} |z_i^{(m+1)} - z_i^{(m)}| = 0,$$

which means that the limit points $z_1^*, \ldots, z_n^*$ of the sequences $\{z_1^{(m)}\}, \ldots, \{z_n^{(m)}\}$ are, actually, the zeros of the polynomial P. To complete the proof of the theorem, it is necessary to show that each of the sequences $\{z_i^{(m)}\}$ $(i \in \boldsymbol{I}_n)$ converges to one and only one zero of P. Since $z_i^{(m)} \in S_i$ for each $i \in \boldsymbol{I}_n$ and $m = 0, 1, \ldots$, it suffices to prove that the disks $S_1, \ldots, S_n$ are mutually disjoint, i.e. (according to (1.67)),

$$|z_i^{(0)} - z_j^{(0)}| = |\text{mid } S_i - \text{mid } S_j| > \text{rad } S_i + \text{rad } S_j = g(\beta)\big(|C_i^{(0)}| + |C_j^{(0)}|\big) \ \ (i \neq j),$$

which reduces to (ii). □

In this section and Chap. 4, we will apply Theorem 3.1 to some iterative methods for the simultaneous approximation of simple zeros of a polynomial. We will assume that an iterative method is well defined if $F(z_1, \ldots, z_n) \neq 0$ under the stated initial conditions and for each array of approximations $(z_1, \ldots, z_n)$ obtained in the course of the iterative procedure.

The convergence analysis of simultaneous methods considered in this section is essentially based on Theorem 3.1 and the four relations connecting the quantities $|W_i|$ (Weierstrass' corrections), d (minimal distance between approximations), and $|C_i|$ (iterative corrections). These relations are referred to as W–D, W–W, C–C, and C–W inequalities according to the quantities involved, and read thus:

$$\text{(W–D):} \qquad w^{(0)} \leq c_n d^{(0)}, \tag{3.7}$$

$$\text{(W–W):} \qquad \big|W_i^{(m+1)}\big| \leq \delta_n \big|W_i^{(m)}\big| \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots), \tag{3.8}$$

$$\text{(C–C):} \qquad \left|C_i^{(m+1)}\right| \le \beta_n \left|C_i^{(m)}\right| \qquad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots), \tag{3.9}$$

$$\text{(C–W):} \qquad \left|C_i^{(m)}\right| \le \lambda_n \frac{\left|W_i^{(m)}\right|}{c_n} \qquad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.10}$$

Here, c_n, δ_n, β_n, and λ_n are real positive constants depending only on the polynomial degree n. The W–D inequality (3.7) defines the initial condition for the guaranteed convergence of an iterative method and plays the main role in the convergence analysis based on the relations (3.7)–(3.10).

The convergence analysis consists of two steps:

1° Starting from the W–D inequality (3.7), derive the W–W inequality (3.8) for each $m = 0, 1, \ldots$. The i-factor c_n has to be chosen so that $\delta_n < 1$ holds. In this way, the convergence of the sequences of Weierstrass' corrections $\{W_i^{(m)}\}$ $(i \in \boldsymbol{I}_n)$ to 0 is ensured.

2° Derive the C–C inequality (3.9) for each $m = 0, 1, \ldots$ under the condition (3.7). The choice of the i-factor c_n must provide the validity of the C–W inequality (3.10) and the inequalities

$$\beta_n < 1 \tag{3.11}$$

and

$$\lambda_n < \frac{1}{2g(\beta_n)}. \tag{3.12}$$

The last requirement arises from the following consideration. Assume that (3.7) implies the inequality (3.10) for all $i \in \boldsymbol{I}_n$. Then using (3.7), we obtain

$$\left|z_i^{(0)} - z_j^{(0)}\right| \ge d^{(0)} \ge \frac{w^{(0)}}{c_n} \ge \frac{\left|C_i^{(0)}\right| + \left|C_j^{(0)}\right|}{2\lambda_n}.$$

Hence, to provide the inequality (ii) in Theorem 3.1, it is necessary to be $1/(2\lambda_n) > g(\beta_n)$ (the inequality (3.12)) where, according to the conditions of Theorem 3.1, the (positive) argument β_n must be less than 1 (the inequality (3.11)). Note that the requirement $\beta_n < 1$ is also necessary to ensure the contraction of the correction terms (see (3.9)) and, thus, the convergence of the considered simultaneous method.

In the subsequent analysis, we will apply the described procedure to some favorable simultaneous methods. This procedure requires certain bounds of the same type and, to avoid the repetition, we give them in the following lemma.

Lemma 3.2. *For distinct complex numbers $z_1, \ldots, z_n$ and $\hat{z}_1, \ldots, \hat{z}_n$, let*

$$d = \min_{\substack{1 \le i,j \le n \\ i \ne j}} |z_i - z_j|, \quad \hat{d} = \min_{\substack{1 \le i,j \le n \\ i \ne j}} |\hat{z}_i - \hat{z}_j| \quad (i \in \boldsymbol{I}_n).$$

If

$$|\hat{z}_i - z_i| \le \lambda_n d \quad (i \in \boldsymbol{I}_n,\, \lambda_n < 1/2), \tag{3.13}$$

then

$$|\hat{z}_i - z_j| \ge (1 - \lambda_n)d \quad (i \in \boldsymbol{I}_n), \tag{3.14}$$

$$|\hat{z}_i - \hat{z}_j| \ge (1 - 2\lambda_n)d \quad (i \in \boldsymbol{I}_n), \tag{3.15}$$

and

$$\left|\prod_{j\ne i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\right| \le \left(1 + \frac{\lambda_n}{1 - 2\lambda_n}\right)^{n-1}. \tag{3.16}$$

Proof. Applying the triangle inequality, we find

$$|\hat{z}_i - z_j| \ge |z_i - z_j| - |\hat{z}_i - z_i| \ge d - \lambda_n d = (1 - \lambda_n)d$$

and

$$|\hat{z}_i - \hat{z}_j| \ge |z_i - z_j| - |\hat{z}_i - z_i| - |\hat{z}_j - z_j| \ge d - \lambda_n d - \lambda_n d = (1 - 2\lambda_n)d. \tag{3.17}$$

From

$$\prod_{j\ne i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j} = \prod_{j\ne i}\left(1 + \frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\right)$$

and

$$\left|\frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\right| \le \frac{\lambda_n d}{(1 - 2\lambda_n)d} = \frac{\lambda_n}{1 - 2\lambda_n},$$

we obtain

$$\left|\prod_{j\ne i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\right| = \prod_{j\ne i}\left|1 + \frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\right| \le \prod_{j\ne i}\left(1 + \left|\frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\right|\right)$$
$$\le \prod_{j\ne i}\left(1 + \frac{\lambda_n}{1 - 2\lambda_n}\right) = \left(1 + \frac{\lambda_n}{1 - 2\lambda_n}\right)^{n-1}. \quad \square$$

Remark 3.1. Since $\hat{d} \le |\hat{z}_i - \hat{z}_j|$, from (3.17) we obtain

$$\hat{d} \le (1 - 2\lambda_n)d. \tag{3.18}$$

In what follows, we apply Theorem 3.1 to the convergence analysis of four frequently used simultaneous zero-finding methods.

The Durand–Kerner's Method

One of the most frequently used iterative methods for the simultaneous determination of simple zeros of a polynomial is the Durand–Kerner's (or Weierstrass') method defined by

$$z_i^{(m+1)} = z_i^{(m)} - W_i^{(m)} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots), \tag{3.19}$$

where $W_i^{(m)}$ is given by (3.3). In this case, the iterative correction term is equal to Weierstrass' correction, i.e., $C_i = W_i = P(z_i)/F_i(z_1, \ldots, z_n)$, where

$$F_i(z_1, \ldots, z_n) = \prod_{\substack{j=1 \\ j \neq i}}^{n} (z_i - z_j) \quad (i \in \boldsymbol{I}_n).$$

To simplify the denotation, we will omit sometimes the iteration index m in the sequel and denote quantities in the subsequent $(m+1)$th iteration by $\widehat{\ }$ ("hat"). It will be always assumed that the polynomial degree n is not smaller than 3.

Lemma 3.3. *Let $z_1, \ldots, z_n$ be distinct approximations and let*

$$w \le c_n d, \tag{3.20}$$

$$c_n \in (0, 0.5), \tag{3.21}$$

$$\delta_n := \frac{(n-1)\, c_n}{1 - c_n} \left(1 + \frac{c_n}{1 - 2c_n}\right)^{n-1} \le 1 - 2c_n \tag{3.22}$$

hold. Then:

(i) $\left|\widehat{W}_i\right| \le \delta_n \left|W_i\right|$.

(ii) $\widehat{w} \le c_n \hat{d}$.

Proof. Let $\lambda_n = c_n$. From (3.19) and (3.20), there follows

$$|\hat{z}_i - z_i| = |W_i| \le w \le c_n d. \tag{3.23}$$

According to this and Lemma 3.2, we obtain

$$|\hat{z}_i - z_j| \ge (1 - c_n) d \tag{3.24}$$

and

$$|\hat{z}_i - \hat{z}_j| \ge (1 - 2c_n) d. \tag{3.25}$$

From the iterative formula (3.19), it follows

$$\frac{W_i}{\hat{z}_i - z_i} = -1,$$

so that

$$\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1 = \frac{W_i}{\hat{z}_i - z_i} + \sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j} + 1 = \sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j}. \tag{3.26}$$

Putting $z = \hat{z}_i$ in the polynomial representation by Lagrange's interpolation formula

$$P(z) = \left(\sum_{j=1}^{n} \frac{W_j}{z - z_j} + 1\right) \prod_{j=1}^{n} (z - z_j), \tag{3.27}$$

we find by (3.26)

$$P(\hat{z}_i) = (\hat{z}_i - z_i)\left(\sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j}\right) \prod_{j \neq i} (\hat{z}_i - z_j).$$

After dividing with $\prod_{j \neq i} (\hat{z}_i - \hat{z}_j)$, one obtains

$$\widehat{W}_i = \frac{P(\hat{z}_i)}{\prod_{j \neq i} (\hat{z}_i - \hat{z}_j)} = (\hat{z}_i - z_i)\left(\sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j}\right) \prod_{j \neq i} \left(1 + \frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\right). \tag{3.28}$$

Using the inequalities (3.20), (3.23)–(3.25) and Lemma 3.2, from (3.28), we estimate

$$\begin{aligned}
\left|\widehat{W}_i\right| &\leq |\hat{z}_i - z_i| \sum_{j \neq i} \frac{|W_j|}{|\hat{z}_i - z_j|} \prod_{j \neq i} \left(1 + \frac{|\hat{z}_j - z_j|}{|\hat{z}_i - \hat{z}_j|}\right) \\
&\leq |W_i| \frac{(n-1)w}{(1-c_n)d} \left(1 + \frac{c_n d}{(1-2c_n)d}\right)^{n-1} \\
&\leq |W_i| \frac{(n-1)c_n}{1-c_n} \left(1 + \frac{c_n}{1-2c_n}\right)^{n-1} \\
&= \delta_n |W_i| .
\end{aligned}$$

This proves the assertion (i) of the lemma.

Since

$$\hat{d} = \min_{\substack{1 \leq i,j \leq n \\ i \neq j}} |\hat{z}_i - \hat{z}_j| ,$$

from (3.25), one obtains

$$\hat{d} \geq (1 - 2\lambda_n)d = (1 - 2c_n)d, \quad \text{i.e.,} \quad d \leq \frac{\hat{d}}{1 - 2c_n}.$$

According to the last inequality and (3.22), we estimate

$$\left|\widehat{W}_i\right| \leq \delta_n |W_i| \leq \delta_n c_n d \leq \frac{\delta_n}{1 - 2c_n} c_n \hat{d} \leq c_n \hat{d}.$$

Therefore, the assertion (ii) holds. □

Theorem 3.2. *Let the assumptions from Lemma* 3.3 *hold. If* $z_1^{(0)}, \ldots, z_n^{(0)}$ *are distinct approximations for which the initial condition*

$$w^{(0)} \le c_n d^{(0)} \tag{3.29}$$

is valid, then the Durand–Kerner's method (3.19) *is convergent.*

Proof. It is sufficient to prove the assertions (i) and (ii) of Theorem 3.1 taking $C_i^{(m)} = W_i^{(m)}$ in this particular case.

According to (ii) of Lemma 3.3, we conclude that (3.29) provides the implication $w^{(0)} \le c_n d^{(0)} \;\Rightarrow\; w^{(1)} \le c_n d^{(1)}$. In a similar way, we show the implication

$$w^{(m)} \le c_n d^{(m)} \implies w^{(m+1)} \le c_n d^{(m+1)},$$

proving by induction that the initial condition (3.29) implies the inequality

$$w^{(m)} \le c_n d^{(m)} \tag{3.30}$$

for each $m = 1, 2, \ldots$. Hence, by (i) of Lemma 3.3, we get

$$\big|W_i^{(m+1)}\big| \le \delta_n \big|W_i^{(m)}\big| = \beta_n |W_i^{(m)}| \tag{3.31}$$

for each $m = 0, 1, \ldots$. Let us note that (3.31) is the W–W inequality of the form (3.8), but also the C–C inequality of the form (3.9) since $C_i = W_i$ in this particular case with $\beta_n = \delta_n$, where δ_n is given by (3.22). Therefore, the assertion (i) holds true.

In a similar way as for (3.25), under the condition (3.29), we prove the inequality

$$\big|z_i^{(m+1)} - z_j^{(m+1)}\big| \ge (1 - 2c_n) d^{(m)} > 0 \quad (i \ne j,\ i, j \in \boldsymbol{I}_n,\ m = 0, 1, \ldots),$$

so that

$$F_i\big(z_1^{(m)}, \ldots, z_n^{(m)}\big) = \prod_{i \ne j} (z_i^{(m)} - z_j^{(m)}) \ne 0$$

in each iteration. Therefore, the Durand–Kerner's method (3.19) is well defined.

Since $\beta_n = \delta_n$, from (3.22), we see that $\beta_n < 1$ (necessary condition (3.11)), and the function g is well defined. To prove (ii) of Theorem 3.1, we have to show that the inequality (3.12) is valid. If $\beta_n \ge 1/2$, then (3.12) becomes

$$\frac{1}{1 - \beta_n} < \frac{1}{2\lambda_n},$$

which is equivalent to (3.22). If $\beta_n < 1/2$, then (3.12) reduces to

$$1 + \beta_n < \frac{1}{2\lambda_n}, \quad \text{i.e.,} \quad \lambda_n = c_n < \frac{1}{2(1 + 2\beta_n)} \in (0.25, 0.5),$$

which holds according to the assumption (3.21) of Lemma 3.3. Since we have proved both assertions (i) and (ii) of Theorem 3.1, we conclude that the Durand–Kerner's method (3.19) is convergent. □

The choice of the "almost optimal" value of c_n is considered in the following lemma.

Lemma 3.4. *The i-factor c_n given by*

$$c_n = \frac{1}{An+B}, \quad A = 1.76325, \quad B = 0.8689425, \tag{3.32}$$

satisfies the conditions (3.21) *and* (3.22)*.*

Proof. Since $c_n \le c_3 \approx 0.16238$, it follows that $c_n \in (0, 0.5)$ and (3.21) holds true.

To prove (3.22), it is sufficient to prove the inequality

$$\eta_n := \frac{\delta_n}{1-2c_n} = \frac{n-1}{1-c_n}\frac{c_n}{1-2c_n}\left(1+\frac{c_n}{1-2c_n}\right)^{n-1} \le 1. \tag{3.33}$$

Since

$$\lim_{n\to\infty}\frac{1}{1-c_n} = 1, \quad \lim_{n\to\infty}\left(1+\frac{c_n}{1-2c_n}\right)^{\dfrac{1-2c_n}{c_n}} = e, \quad \lim_{n\to\infty}\frac{(n-1)\,c_n}{1-2c_n} = \frac{1}{A},$$

where $A = 1.76325$ appears in (3.32), we obtain

$$\lim_{n\to\infty}\eta_n = \frac{1}{A}e^{1/A} < 0.99998 < 1.$$

Since the sequence $\{\eta_n\}$, defined by (3.33), is monotonically increasing for $n \ge 3$, we have $\eta_n < \eta_\infty < 0.99998 < 1$. □

Remark 3.2. The constant $A = 1.76325$ is determined as the reciprocal value of the approximate solution of the equation $xe^x = 1$, and chosen so that it satisfies the inequality $e^{1/A}/A < 1$ (to fulfill the condition $\lim_{n\to\infty}\eta_n < 1$). The use of an approximate solution of the equation $xe^x = 1$ instead of the exact solution (that cannot be represented in floating-point arithmetic of finite precision) just leads to the notion of the "almost optimal" i-factor. Taking a greater number of decimal digits for A (and, consequently, for B, see Remark 3.3), we can make the inequality (3.34) arbitrarily sharp. In this way, we can improve the i-factor c_n to the desired (optimal) extent but, from a practical point of view, such improvement is negligible.

Remark 3.3. Note that the coefficient B in (3.32), not only for the Durand–Kerner's method but also for other methods, is chosen so that the entries δ_n, β_n, and λ_n appearing in the W–W, C–C, and C–W inequalities (3.8)–(3.10) ensure the validity of these inequalities for a particular n, most frequently for $n = 3$. For example, this coefficient for the Durand–Kerner's method is $B = 0.8689425$.

According to Theorem 3.1 and Lemma 3.4, we can state the convergence theorem, which considers initial conditions for the guaranteed convergence of the Durand–Kerner's method.

Theorem 3.3. *The Durand–Kerner's method is convergent under the condition*

$$w^{(0)} < \frac{d^{(0)}}{1.76325n + 0.8689425}. \tag{3.34}$$

Remark 3.4. The sign $<$ ("strongly less") in the inequality (3.34) differs from "$\le$" used in the previous consideration since the concrete choice of A and B in (3.14) yields $\delta_n < 1 - 2c_n$ in (3.22) (also "strongly less"). This is also the case in all remaining methods presented in this book, so that the subsequent situations of this type will not be explained again.

Some authors have considered initial conditions in the form of the inequality

$$\|\boldsymbol{W}^{(0)}\|_1 = \sum_{i=1}^{n} \left|W_i^{(0)}\right| \le \Omega_n d^{(0)}, \quad \boldsymbol{W}^{(0)} = (W_1^{(0)}, \ldots, W_n^{(0)}),$$

instead of the condition (3.7). Obviously, one can take $\Omega_n = n\, c_n$ since (3.29) implies

$$\left|W_i^{(0)}\right| \le c_n d^{(0)} \quad (i = 1, \ldots, n).$$

As already mentioned, the choice of c_n and Ω_n as large as possible permits cruder initial approximations.

We recall some previous ranges concerned with the bounds of Ω_n for $n \ge 3$. X. Wang and Han obtained in [181]

$$\Omega_n = \frac{n}{n-1}\left(3 - 2\sqrt{2}\right) \in (0.1716, 0.2574) \quad (n \ge 3).$$

D. Wang and Zhao improved in [178] the above result yielding the interval

$$\Omega_n \in (0.2044, 0.3241) \quad (n \ge 3).$$

Batra [5] and M. Petković et al. [120] have dealt with $c_n = 1/(2n)$, which gives $\Omega_n = 0.5$. The choice of c_n in this section (see (3.32)) yields

$$\Omega_3 = 3c_3 = 0.48712$$

and

$$\Omega_n \in \left(4c_4, \lim_{n\to\infty} nc_n\right) = \left(0.50493, \frac{1}{A}\right) = (0.50493, 0.56713) \;\; (n \ge 4),$$

which improves all previous results.

The Börsch-Supan's Method

Börsch-Supan's third-order method for the simultaneous approximations of all simple zeros of a polynomial, presented for the first time in [10] and later in [95], is defined by the iterative formula

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots), \tag{3.35}$$

where $W_i^{(m)}$ are given by (3.3) (see Sect. 1.1). This formula has the form (3.5) with the correction

$$C_i(z_1, \ldots, z_n) = \frac{P(z_i)}{F_i(z_1, \ldots, z_n)} \quad (i \in \boldsymbol{I}_n),$$

where

$$F_i(z_1, \ldots, z_n) = \left(1 + \sum_{j \neq i} \frac{W_j}{z_i - z_j}\right) \prod_{j \neq i} (z_i - z_j) \quad (i \in \boldsymbol{I}_n).$$

Before establishing the main convergence theorems, we prove two auxiliary results.

Lemma 3.5. *Let $z_1, \ldots, z_n$ be distinct complex numbers and let*

$$c_n \in \left(0, \frac{1}{n+1}\right) \tag{3.36}$$

and

$$w \le c_n d. \tag{3.37}$$

Then:

(i) $\dfrac{c_n}{\lambda_n} \le \left|1 + \sum_{j \neq i} \dfrac{W_j}{z_i - z_j}\right| \le 2 - \dfrac{c_n}{\lambda_n}$.

(ii) $|\hat{z}_i - z_i| \le \dfrac{\lambda_n}{c_n} |W_i| \le \lambda_n d$.

(iii) $|\hat{z}_i - z_j| \ge (1 - \lambda_n) d$.

(iv) $|\hat{z}_i - \hat{z}_j| \ge (1 - 2\lambda_n) d$.

(v) $\left|\sum_{j=1}^{n} \dfrac{W_j}{\hat{z}_i - z_j} + 1\right| \le \dfrac{(n-1)\lambda_n c_n}{1 - \lambda_n}$.

(vi) $\left|\prod_{j \neq i} \dfrac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\right| \le \left(1 + \dfrac{\lambda_n}{1 - 2\lambda_n}\right)^{n-1}$,

where $\lambda_n = \dfrac{c_n}{1 - (n-1)c_n}$.

Proof. Since $1-2\lambda_n = \dfrac{1-(n+1)c_n}{1-(n-1)c_n}$, from (3.36), it follows $0 < 1-2\lambda_n < 1$, hence $\lambda_n \in (0, 0.5)$. By (3.37) and the definition of d, we obtain

$$\left|1+\sum_{j\neq i}\frac{W_j}{z_i-z_j}\right| \geq 1-\sum_{j\neq i}\frac{|W_j|}{|z_i-z_j|} \geq 1-\frac{(n-1)w}{d} \geq 1-(n-1)\,c_n = \frac{c_n}{\lambda_n}$$

and

$$\left|1+\sum_{j\neq i}\frac{W_j}{z_i-z_j}\right| \leq 1+\frac{(n-1)w}{d} \leq 1+(n-1)\,c_n = 2-\frac{c_n}{\lambda_n},$$

which proves (i). By (i) and (3.37), we prove (ii):

$$|\hat{z}_i - z_i| = \frac{|W_i|}{\left|1+\sum_{j\neq i}\frac{W_j}{z_i-z_j}\right|} \leq \frac{|W_i|}{1-(n-1)\,c_n} = \frac{\lambda_n}{c_n}|W_i| \leq \lambda_n d.$$

The assertions (iii), (iv), and (vi) follow directly according to Lemma 3.2. Omitting the iteration index, from (3.35), we find

$$\frac{W_i}{\hat{z}_i - z_i} = -1-\sum_{j\neq i}\frac{W_j}{z_i-z_j},$$

so that

$$\left|\sum_{j=1}^{n}\frac{W_j}{\hat{z}_i-z_j}+1\right| = \left|\frac{W_i}{\hat{z}_i-z_i}+\sum_{j\neq i}\frac{W_j}{\hat{z}_i-z_j}+1\right| = \left|\sum_{j\neq i}\frac{W_j(z_i-\hat{z}_i)}{(\hat{z}_i-z_j)(z_i-z_j)}\right|.$$

Hence, using (3.37), (ii), and (iii), it follows

$$\left|\sum_{j=1}^{n}\frac{W_j}{\hat{z}_i-z_j}+1\right| \leq |\hat{z}_i-z_i|\sum_{j\neq i}\frac{|W_j|}{|\hat{z}_i-z_j||z_i-z_j|} \leq \frac{(n-1)\lambda_n c_n}{1-\lambda_n},$$

which means that (v) is also true. This completes the proof of the lemma. □

According to Lemma 3.5, we can prove the following assertions.

Lemma 3.6. *Let* $z_1,\ldots,z_n$ *be distinct approximations and let the assumptions* (3.36) *and* (3.37) *of Lemma* 3.5 *hold. In addition, let*

$$\delta_n := \frac{(n-1)\lambda_n^2}{1-\lambda_n}\left(1+\frac{\lambda_n}{1-2\lambda_n}\right)^{n-1} \leq 1-2\lambda_n \tag{3.38}$$

be valid. Then:

(i) $|\widehat{W}_i| \le \delta_n |W_i|$.
(ii) $\widehat{w} \le c_n \hat{d}$.

Proof. Setting $z = \hat{z}_i$ in (3.27), where $\hat{z}_i$ is a new approximation produced by the Börsch-Supan's method (3.35), we obtain

$$P(\hat{z}_i) = (\hat{z}_i - z_i)\left(\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1\right) \prod_{j\neq i} (\hat{z}_i - z_j).$$

After dividing with $\prod_{j\neq i} (\hat{z}_i - \hat{z}_j)$, we get

$$\widehat{W}_i = (\hat{z}_i - z_i)\left(\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1\right) \prod_{j\neq i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}.$$

Using the bounds (ii), (v), and (vi) of Lemma 3.5, we estimate

$$\begin{aligned}\left|\widehat{W}_i\right| &= |\hat{z}_i - z_i| \left|\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1\right| \left|\prod_{j\neq i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\right| \\ &\le |W_i| \frac{(n-1)\lambda_n^2}{1-\lambda_n} \left(1 + \frac{\lambda_n}{1-2\lambda_n}\right)^{n-1},\end{aligned}$$

i.e., $\left|\widehat{W}_i\right| \le \delta_n |W_i|$. Therefore, the assertion (i) holds true.

According to (iv) of Lemma 3.5, there follows

$$\hat{d} \ge (1 - 2\lambda_n) d.$$

This inequality, together with (i) of Lemma 3.6 and (3.38), gives (ii), i.e.,

$$\left|\widehat{W}_i\right| \le \delta_n |W_i| \le \frac{\delta_n}{1-2\lambda_n} c_n \hat{d} \le c_n \hat{d}. \quad \square$$

Theorem 3.4. *Let the assumptions from Lemmas* 3.5 *and* 3.6 *hold and, in addition, let*

$$\beta_n := \left(\frac{2\lambda_n}{c_n} - 1\right)\delta_n < 1 \tag{3.39}$$

and

$$g(\beta_n) < \frac{1}{2\lambda_n}. \tag{3.40}$$

If $z_1^{(0)}, \ldots, z_n^{(0)}$ *are distinct initial approximations satisfying*

$$w^{(0)} \le c_n d^{(0)}, \tag{3.41}$$

then the Börsch-Supan's method (3.35) *is convergent.*

Proof. It is sufficient to prove (i) and (ii) of Theorem 3.1 for the iterative correction given by

$$C_i^{(m)} = \frac{W_i^{(m)}}{1 + \sum_{j \neq i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots)$$

(see (3.35)). By virtue of Lemma 3.6, which holds under the conditions (3.36), (3.38), and (3.41), we can prove by induction that

$$w^{(m+1)} \leq \delta_n w^{(m)} \leq \frac{\delta_n}{1 - 2\lambda_n} c_n d^{(m+1)} \leq c_n d^{(m+1)}$$

holds for each $m = 0, 1, \ldots$.

Starting from the assertion (i) of Lemma 3.5, under the condition (3.41), we prove by induction

$$F_i\big(z_1^{(m)}, \ldots, z_n^{(m)}\big) = \left(1 + \sum_{j \neq i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}\right) \prod_{j \neq i} \Big(z_i^{(m)} - z_j^{(m)}\Big) \neq 0$$

for each $i \in \boldsymbol{I}_n$ and $m = 0, 1, \ldots$. Therefore, the Börsch-Supan's method (3.35) is well defined in each iteration.

Using (i) of Lemma 3.5, we find

$$|C_i| = \frac{|W_i|}{\left|1 + \sum_{j \neq i} \frac{W_j}{z_i - z_j}\right|} \leq \frac{\lambda_n}{c_n} |W_i|, \tag{3.42}$$

so that for the next iterative step we obtain by Lemma 3.5 and (i) of Lemma 3.6

$$\begin{aligned}
\big|\widehat{C}_i\big| &\leq \frac{\lambda_n}{c_n} \big|\widehat{W}_i\big| \leq \frac{\lambda_n \delta_n}{c_n} \frac{|W_i|}{\left|1 + \sum_{j \neq i} \frac{W_j}{z_i - z_j}\right|} \left|1 + \sum_{j \neq i} \frac{W_j}{z_i - z_j}\right| \\
&= \frac{\lambda_n \delta_n}{c_n} |C_i| \left|1 + \sum_{j \neq i} \frac{W_j}{z_i - z_j}\right| \leq \frac{\lambda_n \delta_n}{c_n} \Big(2 - \frac{c_n}{\lambda_n}\Big) |C_i| \\
&= \delta_n \Big(\frac{2\lambda_n}{c_n} - 1\Big) |C_i| = \beta_n |C_i|,
\end{aligned}$$

where $\beta_n < 1$ (the assumption (3.39)). Using the same argumentation, we prove by induction

$$\left|C_i^{(m+1)}\right| \le \beta_n \left|C_i^{(m)}\right|$$

for each $i \in \boldsymbol{I}_n$ and $m = 0, 1, \ldots$.

By (3.41) and (3.42), we estimate

$$\frac{1}{\lambda_n}|C_i^{(0)}| \le \frac{|W_i^{(0)}|}{c_n} \le d^{(0)}.$$

According to this and (3.40), we see that

$$\begin{aligned}\left|z_i^{(0)} - z_j^{(0)}\right| \ge d^{(0)} \ge \frac{w^{(0)}}{c_n} \ge \frac{1}{2\lambda_n}\Big(\left|C_i^{(0)}\right| + \left|C_j^{(0)}\right|\Big) \\ > g(\beta_n)\Big(\left|C_i^{(0)}\right| + \left|C_j^{(0)}\right|\Big)\end{aligned}$$

holds for each $i \ne j$, $i, j \in \boldsymbol{I}_n$. This proves (ii) of Theorem 3.1. The validity of (i) and (ii) of Theorem 3.1 shows that the Börsch-Supan's method (3.35) is convergent under the given conditions. □

The choice of the i-factor c_n is considered in the following lemma.

Lemma 3.7. *The i-factor c_n defined by*

$$c_n = \begin{cases} \dfrac{1}{n + \frac{9}{2}}, & n = 3, 4 \\[2ex] \dfrac{1}{\frac{309}{200}n + 5}, & n \ge 5 \end{cases} \tag{3.43}$$

satisfies the condition of Theorem 3.4.

The proof of this lemma is elementary and it is derived by a simple analysis of the sequences $\{\beta_n\}$ and $\{g(\beta_n)\}$.

According to Lemma 3.7 and Theorem 3.4, we may state the following theorem.

Theorem 3.5. *The Börsch-Supan's method* (3.35) *is convergent under the condition* (3.41)*, where c_n is given by* (3.43)*.*

Tanabe's Method

In Sect. 1.1, we have presented the third-order method, often referred to as Tanabe's method

$$z_i^{(m+1)} = z_i^{(m)} - W_i^{(m)}\left(1 - \sum_{j \ne i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}\right) \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.44}$$

As in the previous cases, before stating initial conditions that ensure the guaranteed convergence of the method (3.44), we give first some necessary bounds using the previously introduced notation and omitting the iteration index for simplicity.

Lemma 3.8. *Let $z_1, \dots, z_n$ be distinct approximations and let*

$$c_n \in \Big(0, \frac{1}{1+\sqrt{2n-1}}\Big). \tag{3.45}$$

If the inequality

$$w \le c_n d \tag{3.46}$$

holds, then for $i, j \in \boldsymbol{I}_n$ we have:

(i) $\dfrac{\lambda_n}{c_n} = 1 + (n-1)c_n \ge \Big|1 - \sum_{j \ne i} \dfrac{W_j}{z_i - z_j}\Big| \ge 1 - (n-1)c_n = 2 - \dfrac{\lambda_n}{c_n}$.

(ii) $|\hat{z}_i - z_i| \le \dfrac{\lambda_n}{c_n}|W_i| \le \lambda_n d$.

(iii) $|\hat{z}_i - z_j| \ge (1-\lambda_n)d$.

(iv) $|\hat{z}_i - \hat{z}_j| \ge (1-2\lambda_n)d$.

(v) $\Big|\sum_{j=1}^{n} \dfrac{W_j}{\hat{z}_i - z_j} + 1\Big| \le \dfrac{(n-1)c_n^2}{(2c_n - \lambda_n)(1-\lambda_n)}\Big(\lambda_n + (n-1)c_n\Big)$.

(vi) $\prod_{j \ne i} \Big|\dfrac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\Big| \le \Big(1 + \dfrac{\lambda_n}{1-2\lambda_n}\Big)^{n-1}$,

where $\lambda_n = (1 + (n-1)c_n)c_n$.

Proof. We omit the proofs of the assertions (i)–(iv) and (vi) since they are quite similar to those given in Lemma 3.5. To prove (v), we first introduce

$$\sigma_i = \sum_{j \ne i} \frac{W_j}{z_i - z_j}.$$

Then

$$|\sigma_i| \le \frac{(n-1)w}{d} \le (n-1)c_n \quad \text{and} \quad \frac{|\sigma_i|}{1-|\sigma_i|} \le \frac{(n-1)c_n}{1-(n-1)c_n}. \tag{3.47}$$

From the iterative formula (3.44), we obtain

$$\frac{W_i}{\hat{z}_i - z_i} = -\frac{1}{1 - \sum_{j \ne i} \dfrac{W_j}{z_i - z_j}},$$

so that by (3.47) it follows

$$\begin{aligned}\Big|\sum_{j=1}^{n}\frac{W_j}{\hat z_i - z_j}+1\Big| &= \Big|\frac{W_i}{\hat z_i - z_i}+\sum_{j\neq i}\frac{W_j}{\hat z_i - z_j}+1\Big| = \Big|1-\frac{1}{1-\sigma_i}+\sum_{j\neq i}\frac{W_j}{\hat z_i - z_j}\Big|\\
&= \frac{1}{|1-\sigma_i|}\Big|\sum_{j\neq i}\frac{W_j}{\hat z_i - z_j}-\sum_{j\neq i}\frac{W_j}{z_i - z_j}-\sigma_i\sum_{j\neq i}\frac{W_j}{z_i - z_j}\Big|\\
&\le \frac{1}{|1-\sigma_i|}\Big|\sum_{j\neq i}\frac{W_j(z_i-\hat z_i)}{(\hat z_i - z_j)(z_i - z_j)}\Big| + \frac{|\sigma_i|}{1-|\sigma_i|}\Big|\sum_{j\neq i}\frac{W_j}{\hat z_i - z_j}\Big|\\
&\le \frac{1}{1-(n-1)c_n}|z_i-\hat z_i|\sum_{j\neq i}\frac{|W_j|}{|\hat z_i - z_j||z_i - z_j|}\\
&\quad + \frac{(n-1)c_n}{1-(n-1)c_n}\sum_{j\neq i}\frac{|W_j|}{|\hat z_i - z_j|}.\end{aligned}$$

Hence, by (ii), (iii), (3.46), and (3.47), we estimate

$$\begin{aligned}\Big|\sum_{j=1}^{n}\frac{W_j}{\hat z_i - z_j}+1\Big| &\le \frac{\lambda d}{1-(n-1)c_n}\,\frac{(n-1)w}{(1-\lambda_n)d\cdot d}+\frac{(n-1)c_n}{1-(n-1)c_n}\,\frac{(n-1)w}{(1-\lambda_n)d}\\
&= \frac{(n-1)c_n^2}{(2c_n-\lambda_n)(1-\lambda_n)}\Big(\lambda_n+(n-1)c_n\Big). \quad \square\end{aligned}$$

Remark 3.5. The inequalities (iv) and (vi) require $2\lambda_n<1$ or $2c_n(1+(n-1)c_n)<1$. This inequality will be satisfied if $c_n < 1/(1+\sqrt{2n-1})$, which is true according to (3.45).

Lemma 3.9. *Let $z_1,\dots,z_n$ be distinct approximations and let the assumptions* (3.45) *and* (3.46) *of Lemma* 3.8 *hold. If the inequality*

$$\delta_n := \frac{(n-1)c_n\lambda_n}{(2c_n-\lambda_n)(1-\lambda_n)}\Big(\lambda_n+(n-1)c_n\Big)\Big(1+\frac{\lambda_n}{1-2\lambda_n}\Big)^{n-1} \le 1-2\lambda_n \quad (3.48)$$

is valid, then:

(i) $|\widehat W_i| \le \delta_n|W_i|$.
(ii) $\widehat w \le c_n\hat d$.

Proof. Putting $z=\hat z_i$ in (3.27), where $\hat z_i$ is a new approximation obtained by Tanabe's method (3.44), and dividing with $\prod_{j\neq i}(\hat z_i-\hat z_j)$, we obtain

$$\widehat W_i = \frac{P(\hat z_i)}{\prod_{j\neq i}(\hat z_i-\hat z_j)} = (\hat z_i - z_i)\Big[\sum_{j=1}^{n}\frac{W_j}{\hat z_i - z_j}+1\Big]\prod_{j\neq i}\frac{\hat z_i - z_j}{\hat z_i-\hat z_j}.$$

From the last relation, we obtain by (ii), (v), and (vi) of Lemma 3.8

$$
\begin{aligned}
|\widehat{W}_i| &= |\hat{z}_i - z_i| \Big| \sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1 \Big| \prod_{j \neq i} \Big| \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j} \Big| \\
&\leq \frac{\lambda_n}{c_n} |W_i| \frac{(n-1)c_n^2}{(2c_n - \lambda_n)(1-\lambda_n)} \Big(\lambda_n + (n-1)c_n\Big) \Big(1 + \frac{\lambda_n}{1 - 2\lambda_n}\Big)^{n-1} \\
&= \delta_n |W_i|,
\end{aligned}
$$

which proves (i).

Using (iv) of Lemma 3.8, we find

$$\hat{d} \geq (1 - 2\lambda_n) d.$$

Combining this inequality with (i) of Lemma 3.9, (3.45), and (3.48), we prove (ii):

$$|\widehat{W}_i| \leq \delta_n |W_i| \leq \delta_n c_n d \leq \frac{\delta_n}{1 - 2\lambda_n} c_n \hat{d} \leq c_n \hat{d}. \quad \square$$

Theorem 3.6. *Let the assumptions of Lemmas* 3.7 *and* 3.8 *be valid and let*

$$\beta_n := \frac{\lambda_n \delta_n}{2c_n - \lambda_n} < 1 \tag{3.49}$$

and

$$g(\beta_n) < \frac{1}{2\lambda_n}. \tag{3.50}$$

If $z_1^{(0)}, \ldots, z_n^{(0)}$ *are distinct initial approximations satisfying*

$$w^{(0)} \leq c_n d^{(0)}, \tag{3.51}$$

then the Tanabe's method (3.44) *is convergent.*

Proof. In Lemma 3.9 (assertion (ii)), we derived the implication $w \leq c_n d \Rightarrow \widehat{w} \leq c_n \hat{d}$. Using a similar procedure, we prove by induction that the initial condition (3.51) implies the inequality $w^{(m)} \leq c_n d^{(m)}$ for each $m = 1, 2, \ldots$. Therefore, all assertions of Lemmas 3.8 and 3.9 are valid for each $m = 1, 2, \ldots$. For example, we have

$$|W_i^{(m+1)}| \leq \delta_n |W_i^{(m)}| \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.52}$$

From the iterative formula (3.44), we see that corrections $C_i^{(m)}$ are given by

$$C_i^{(m)} = W_i^{(m)} \Bigg(1 - \sum_{j \neq i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}\Bigg) \quad (i \in \boldsymbol{I}_n). \tag{3.53}$$

This correction has the required form

$$C_i^{(m)} = P(z_i^{(m)})/F(z_1^{(m)}, \ldots, z_n^{(m)}),$$

where

$$F_i(z_1^{(m)}, \ldots, z_n^{(m)}) = \frac{\displaystyle\prod_{j \neq i} \big(z_i^{(m)} - z_j^{(m)}\big)}{1 - \displaystyle\sum_{j \neq i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n).$$

According to (i) of Lemma 3.8, it follows $F_i(z_1^{(m)}, \ldots, z_n^{(m)}) \neq 0$, which means that the Tanabe's method is well defined in each iteration.

We now prove the first part of the theorem which is concerned with the monotonicity of the sequences $\{C_i^{(m)}\}$ $(i \in \boldsymbol{I}_n)$ of corrections. Starting from (3.53) and omitting iteration indices, we find by (ii) of Lemma 3.8 (which is valid under the condition (3.51))

$$|C_i| = |W_i| \Big| 1 - \sum_{j \neq i} \frac{W_j}{z_i - z_j} \Big| \leq \frac{\lambda_n}{c_n} |W_i|. \tag{3.54}$$

According to (3.52)–(3.54) and by the inequalities (i) of Lemma 3.8, we obtain

$$\begin{aligned} |\widehat{C}_i| &\leq \frac{\lambda_n}{c_n} |\widehat{W}_i| \leq \frac{\lambda_n \delta_n}{c_n} |W_i| = \frac{\lambda_n \delta_n}{c_n} \cdot \frac{\left| W_i \Big(1 - \displaystyle\sum_{j \neq i} \frac{W_j}{z_i - z_j}\Big) \right|}{\left| 1 - \displaystyle\sum_{j \neq i} \frac{W_j}{z_i - z_j} \right|} \\ &\leq \frac{\lambda_n \delta_n}{c_n (2 - \lambda_n / c_n)} |C_i| = \beta_n |C_i|, \end{aligned}$$

where $\beta_n < 1$ (assumption (3.49)). By induction, it is proved that the inequality $|C_i^{(m+1)}| \leq \beta_n |C_i^{(m)}|$ holds for each $i = 1, \ldots, n$ and $m = 0, 1, \ldots$.

By (3.51) and (3.54), we estimate

$$\frac{1}{\lambda_n} |C_i^{(0)}| \leq \frac{w^{(0)}}{c_n} \leq d^{(0)}.$$

According to this, (3.50), and (3.51), we conclude that

$$|z_i^{(0)} - z_j^{(0)}| \geq d^{(0)} \geq \frac{w^{(0)}}{c_n} \geq \frac{1}{2\lambda_n} \big(|C_i^{(0)}| + |C_j^{(0)}|\big) > g(\beta_n) \big(|C_i^{(0)}| + |C_j^{(0)}|\big)$$

holds for each $i \neq j, \ i, j \in \boldsymbol{I}_n$. This proves (ii) of Theorem 3.1. □

Lemma 3.10. *The i-factor c_n given by $c_n = 1/(3n)$ satisfies the condition* (3.45), (3.48), (3.49), *and* (3.50).

Proof. Obviously, $c_n = 1/(3n) < 1/\big(1+\sqrt{2n-1}\,\big)$. Furthermore, the sequence $\{\delta_n\}$ is monotonically increasing so that

$$\delta_n < \lim_{n\to+\infty} \delta_n = \frac{2}{9}e^{4/9} \approx 0.3465 < 0.35 \quad \text{for every } n \geq 3.$$

We adopt $\delta_n = 0.35$ and prove that (3.48) holds; indeed,

$$\delta_n = 0.35 < 1 - 2\lambda_n = 1 - \frac{2(4n-1)}{9n^2} \;\left(\geq \frac{59}{81} \approx 0.728\right)$$

for every $n \geq 3$.

For $\delta_n = 0.35$ and $c_n = 1/(3n)$, the sequence $\{\beta_n\}$ defined by

$$\beta_n = \frac{\delta_n \lambda_n}{2c_n - \lambda_n} = \frac{0.35(4n-1)}{2n+1}$$

is monotonically increasing so that

$$\beta_n < \lim_{n\to+\infty} \delta_n = 0.7 < 1 \quad (n \geq 3),$$

which means that (3.49) is valid.

Finally, we check the validity of the inequality (3.50) taking $\beta_n = 0.7$. We obtain

$$g(\beta_n) = g(0.7) = \frac{1}{1-0.7} = 3.333... < \frac{1}{2\lambda_n} = \frac{9n^2}{2(4n-1)} \;\left(\geq \frac{81}{22} \approx 3.68\right),$$

wherefrom we conclude that the inequality (3.50) holds for every $n \geq 3$. □

According to Lemma 3.10 and Theorem 3.1, the following theorem is stated.

Theorem 3.7. *The Tanabe's method* (3.44) *is convergent under condition* (3.51), *where $c_n = 1/(3n)$.*

The Chebyshev-Like Method

In Sect. 1.1, we have presented the iterative fourth-order method of Chebyshev's type

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1+G_{1,i}^{(m)}}\left(1 - \frac{W_i^{(m)}G_{2,i}^{(m)}}{(1+G_{1,i}^{(m)})^2}\right) \quad (i \in \boldsymbol{I}_n,\ m = 0,1,2,\ldots), \tag{3.55}$$

proposed by M. Petković, Tričković, and Đ. Herceg [146]. Before stating initial conditions that guarantee the convergence of the method (3.55), three lemmas which concern some necessary bounds and estimations are given first.

Lemma 3.11. *Let* $z_1, \ldots, z_n$ *be distinct approximations to the zeros* $\zeta_1, \ldots, \zeta_n$ *of a polynomial* P *and let* $\hat{z}_1, \ldots, \hat{z}_n$ *be new approximations obtained by the iterative formula* (3.55)*. If the inequality*

$$w < c_n d, \quad c_n = \frac{2}{5n+3} \quad (n \geq 3) \tag{3.56}$$

holds, then for all $i \in \boldsymbol{I}_n$ *we have:*

(i) $\dfrac{3n+5}{5n+3} < |1 + G_{1,i}| < \dfrac{7n+1}{5n+3}$.

(ii) $|\hat{z}_i - z_i| \leq \dfrac{\lambda_n}{c_n}|W_i| \leq \lambda_n d$, *where* $\lambda_n = \dfrac{2(9n^2+34n+21)}{(3n+5)^3}$.

Proof. According to the definition of the minimal distance d and the inequality (3.56), it follows

$$|G_{1,i}| \leq \sum_{j \neq i} \frac{|W_j|}{|z_i - z_j|} < (n-1)c_n, \quad |G_{2,i}| \leq \sum_{j \neq i} \frac{|W_j|}{|z_i - z_j|^2} < \frac{(n-1)c_n}{d}, \tag{3.57}$$

so that we find

$$|1 + G_{1,i}| \geq 1 - \sum_{j \neq i} \frac{|W_j|}{|z_i - z_j|} > 1 - (n-1)c_n = \frac{3n+5}{5n+3}$$

and

$$|1 + G_{1,i}| \leq 1 + \sum_{j \neq i} \frac{|W_j|}{|z_i - z_j|} < 1 + (n-1)c_n = \frac{7n+1}{5n+3}.$$

Therefore, the assertion (i) of Lemma 3.11 is proved.

Using (i) and (3.57), we estimate

$$\left| \frac{W_i}{1 + G_{1,i}} \right| < \frac{w}{1 - (n-1)c_n} < \frac{2}{3n+5} d \tag{3.58}$$

and

$$\left| \frac{W_i G_{2,i}}{(1 + G_{1,i})^2} \right| < \frac{w(n-1)c_n/d}{(1 - (n-1)c_n)^2} < \frac{c_n^2 (n-1)}{(1 - (n-1)c_n)^2} \leq \frac{4(n-1)}{(3n+5)^2}. \tag{3.59}$$

Using (3.58) and (3.59), we obtain the bound (ii):

$$\begin{aligned}|\hat{z}_i - z_i| = |C_i| &= \left|\frac{W_i}{1+G_{1,i}}\left(1 - \frac{W_i G_{2,i}}{(1+G_{1,i})^2}\right)\right| \\ &\leq \frac{|W_i|}{|1+G_{1,i}|}\left(1 + \frac{|W_i G_{2,i}|}{|1+G_{1,i}|^2}\right) \\ &< |W_i| \cdot \frac{5n+3}{3n+5}\left(1 + \frac{4(n-1)}{(3n+5)^2}\right) \\ &< \frac{2(9n^2+34n+21)}{(3n+5)^3} d = \lambda_n d. \quad \square\end{aligned}$$

According to Lemma 3.2 and the assertion (ii) of Lemma 3.11, under the condition (3.56), we have

$$|\hat{z}_i - z_j| > (1-\lambda_n)d \quad (i,j \in \boldsymbol{I}_n), \tag{3.60}$$

$$|\hat{z}_i - \hat{z}_j| > (1-2\lambda_n)d \quad (i,j \in \boldsymbol{I}_n), \tag{3.61}$$

and

$$\left|\prod_{j\neq i} \frac{\hat{z}_i - z_j}{\hat{z}_i - \hat{z}_j}\right| < \left(1 + \frac{\lambda_n}{1-2\lambda_n}\right)^{n-1}. \tag{3.62}$$

Let us note that the necessary condition $\lambda_n < 1/2$ is satisfied under the condition (3.56).

Remark 3.6. Since (3.61) is valid for arbitrary pair $i, j \in \boldsymbol{I}_n$ and $\lambda_n < 1/2$ if (3.56) holds, there follows

$$\hat{d} = \min_{j\neq i} |\hat{z}_i - \hat{z}_j| > (1-2\lambda_n)d. \tag{3.63}$$

Lemma 3.12. *If the inequality* (3.56) *holds, then*

(i) $|\widehat{W}_i| < 0.22|W_i|$.
(ii) $\widehat{w} < \dfrac{2}{5n+3}\hat{d}$.

Proof. For distinct points $z_1, \ldots, z_n$, we use the polynomial representation (3.27) and putting $z = \hat{z}_i$ in (3.27), we find

$$P(\hat{z}_i) = (\hat{z}_i - z_i)\left(\frac{W_i}{\hat{z}_i - z_i} + \sum_{j\neq i} \frac{W_j}{\hat{z}_i - z_j} + 1\right)\prod_{j\neq i}(\hat{z}_i - z_j).$$

After division with $\prod_{j\neq i}(\hat z_i-\hat z_j)$, we get

$$\widehat{W}_i=\frac{P(\hat z_i)}{\prod_{j\neq i}(\hat z_i-\hat z_j)}=(\hat z_i-z_i)\bigg(\frac{W_i}{\hat z_i-z_i}+\sum_{j\neq i}\frac{W_j}{\hat z_i-z_j}+1\bigg)\prod_{j\neq i}\Big(\frac{\hat z_i-z_j}{\hat z_i-\hat z_j}\Big). \tag{3.64}$$

From the iterative formula (3.55), it follows

$$\frac{W_i}{\hat z_i-z_i}=\frac{-(1+G_{1,i})}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}=-1-\frac{G_{1,i}+\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}. \tag{3.65}$$

Then

$$\frac{W_i}{\hat z_i-z_i}+\sum_{j\neq i}\frac{W_j}{\hat z_i-z_j}+1=-1-\frac{G_{1,i}+\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}+\sum_{j\neq i}\frac{W_j}{\hat z_i-z_j}+1$$

$$=\frac{-\sum_{j\neq i}\dfrac{W_j}{z_i-z_j}+\sum_{j\neq i}\dfrac{W_j}{\hat z_i-z_j}-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}\sum_{j\neq i}\dfrac{W_j}{\hat z_i-z_j}}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}$$

$$=\frac{-(\hat z_i-z_i)\sum_{j\neq i}\dfrac{W_j}{(z_i-z_j)(\hat z_i-z_j)}-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}\Big(1+\sum_{j\neq i}\dfrac{W_j}{\hat z_i-z_j}\Big)}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}.$$

From the last formula, we obtain by (3.59), (3.60), the definition of the minimal distance, and (ii) of Lemma 3.11

$$\Big|\frac{W_i}{\hat z_i-z_i}+\sum_{j\neq i}\frac{W_j}{\hat z_i-z_j}+1\Big|$$

$$\leq\frac{|\hat z_i-z_i|\sum_{j\neq i}\dfrac{|W_j|}{|z_i-z_j||\hat z_i-z_j|}+\Big|\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}\Big|\Big(1+\sum_{j\neq i}\dfrac{|W_j|}{|\hat z_i-z_j|}\Big)}{1-\Big|\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}\Big|}$$

$$<\frac{8(135n^5+594n^4+646n^3-292n^2-821n-262)}{(5n+3)(9n^2+26n+29)(27n^3+117n^2+157n+83)}=y_n. \tag{3.66}$$

Now, starting from (3.64) and taking into account (3.60)–(3.62), (3.66), and the assertions of Lemma 3.11, we obtain

$$\begin{aligned}|\widehat{W}_i| &\le |\hat{z}_i - z_i|\Big|\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1\Big| \prod_{j\ne i}\Big(1 + \frac{|\hat{z}_j - z_j|}{|\hat{z}_i - \hat{z}_j|}\Big)\\ &< \frac{\lambda_n}{c_n}|W_i| y_n \Big(1 + \frac{\lambda_n}{1 - 2\lambda_n}\Big)^{n-1}\\ &= \phi_n |W_i|.\end{aligned}$$

Using the symbolic computation in the programming package *Mathematica* 6.0, we find that the sequence $\{\phi_n\}_{n=3,4,\dots}$ attains its maximum for $n = 5$:

$$\phi_n \le \phi_5 < 0.22, \quad \text{for all } n \ge 3.$$

Therefore, $|\widehat{W}_i| < 0.22|W_i|$ and the assertion (i) is valid.

According to this, (3.63), and the inequality

$$\frac{0.22(3n+5)^3}{27n^3 + 99n^2 + 89n + 41} \le 0.32 < 1,$$

we find

$$|\widehat{W}_i| < 0.22|W_i| < 0.22\,\frac{2d}{5n+3} < 0.22\,\frac{2}{5n+3}\,\frac{(3n+5)^3}{27n^3 + 99n^2 + 89n + 41}\hat{d},$$

wherefrom

$$\widehat{w} < \frac{2}{5n+3}\hat{d},$$

which proves the assertion (ii) of Lemma 3.12. □

Now, we are able to establish the main convergence theorem for the iterative method (3.55).

Theorem 3.8. *If the initial approximations* $z_1^{(0)}, \dots, z_n^{(0)}$ *satisfy the initial condition*

$$w^{(0)} < c_n d^{(0)}, \quad c_n = \frac{2}{5n+3} \quad (n \ge 3), \tag{3.67}$$

then the iterative method (3.55) *is convergent.*

Proof. It is sufficient to prove that the inequalities (i) and (ii) of Theorem 3.1 are valid for the correction

$$C_i^{(m)} = \frac{W_i^{(m)}}{1 + G_{1,i}^{(m)}}\left(1 - \frac{W_i^{(m)} G_{2,i}^{(m)}}{(1 + G_{1,i}^{(m)})^2}\right) \quad (i \in \boldsymbol{I}_n),$$

which appears in the considered method (3.55).

Using Lemma 3.11(i) and (3.59), we find

$$\begin{aligned}|C_i| &= |\hat{z}_i - z_i| = \Big|\frac{W_i}{1+G_{1,i}}\Big(1 - \frac{W_iG_{2,i}}{(1+G_{1,i})^2}\Big)\Big| \\ &< \frac{5n+3}{3n+5}W_i\Big(1+\frac{4(n-1)}{(3n+5)^2}\Big) = \frac{(5n+3)(9n^2+34n+21)}{(3n+5)^3}|W_i| \\ &= x_n|W_i|.\end{aligned}$$

It is easy to show that the sequence $\{x_n\}_{n=3,4,\ldots}$ is monotonically increasing and $x_n < \lim_{m\to\infty} x_n = 5/3$, wherefrom

$$|C_i| < \frac{5}{3}|W_i|. \tag{3.68}$$

In Lemma 3.12 (assertion (ii)), the implication $w < c_n d \Rightarrow \widehat{w} < c_n \hat{d}$ has been proved. Using a similar procedure, we prove by induction that the initial condition (3.67) implies the inequality $w^{(m)} < c_n d^{(m)}$ for each $m = 1, 2, \ldots$. Therefore, by (i) of Lemma 3.12, we obtain

$$|W_i^{(m+1)}| < 0.22|W_i^{(m)}| \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots).$$

According to this and by the inequalities (i) of Lemma 3.11 and (3.68), we obtain (omitting iteration indices)

$$\begin{aligned}|\widehat{C}_i| &= \frac{5}{3}|\widehat{W}_i| < \frac{5}{3}\cdot 0.22|W_i| \\ &= \frac{1.1}{3}\Big|\frac{W_i}{1+G_{1,i}}\Big(1-\frac{W_iG_{2,i}}{(1+G_{1,i})^2}\Big)\Big|\left|\frac{1+G_{1,i}}{1-\dfrac{W_iG_{2,i}}{(1+G_{1,i})^2}}\right| \\ &< \frac{1.1}{3}|C_i|\frac{\dfrac{7n+1}{5n+3}}{1-\dfrac{4(n-1)}{(3n+5)^2}} < 0.52|C_i|.\end{aligned}$$

In this manner, we have proved by induction that the inequality $|C_i^{(m+1)}| < 0.52|C_i^{(m)}|$ holds for each $i = 1, \ldots, n$ and $m = 0, 1, \ldots$. Furthermore, comparing this result with (i) of Theorem 3.1, we see that $\beta = 0.52 < 1$. This yields the first part of the theorem. In addition, according to (3.57), we note that the following is valid:

$$|G_{1,i}| < (n-1)c_n = \frac{2(n-1)}{5n+3} \le \frac{2}{9} < 1,$$

which means that $0 \notin 1 + G_{1,i}$. Using induction and the assertion (ii) of Lemma 3.12, we prove that $0 \notin 1 + G_{1,i}^{(m)}$ for arbitrary iteration index m.

Therefore, under the condition (3.67), the iterative method (3.55) is well defined in each iteration.

To prove (ii) of Theorem 3.8, we first note that $\beta = 0.52$ yields $g(\beta) = \frac{1}{1-0.52} \approx 2.08$. It remains to prove the disjunctivity of the inclusion disks

$$S_1 = \{z_1^{(0)}; 2.08|C_1^{(0)}\}, \ldots, S_n = \{z_n^{(0)}; 2.08|C_n^{(0)}\}.$$

By virtue of (3.68), we have $|C_i^{(0)}| < \frac{5}{3}w^{(0)}$, wherefrom

$$\begin{aligned} d^{(0)} &> \frac{5n+3}{2}w^{(0)} > \frac{5n+3}{2} \cdot \frac{3}{5}|C_i^{(0)}| \geq \frac{3(5n+3)}{20}\big(|C_i^{(0)}| + |C_j^{(0)}|\big) \\ &> g(0.52)\big(|C_i^{(0)}| + |C_j^{(0)}|\big). \end{aligned}$$

This means that

$$\big|z_i^{(0)} - z_j^{(0)}\big| \geq d^{(0)} > g(0.52)\big(\big|C_i^{(0)}\big| + \big|C_j^{(0)}\big|\big) = \operatorname{rad} S_i + \operatorname{rad} S_j.$$

Therefore, the inclusion disks $S_1, \ldots, S_n$ are disjoint, which completes the proof of Theorem 3.8. □

3.3 Guaranteed Convergence: Sequence Approach

In what follows, we will present another concept of the convergence analysis involving initial conditions of the form (3.7) which guarantee the convergence of the considered methods.

Let $z_1^{(m)}, \ldots, z_n^{(m)}$ be approximations to the simple zeros $\zeta_1, \ldots, \zeta_n$ of a polynomial P, generated by some iterative method for the simultaneous determination of zeros at the mth iterative step and let $u_i^{(m)} = z_i^{(m)} - \zeta_i$ $(i \in \boldsymbol{I}_n)$. Our main aim is to study the convergence of the sequences $\{u_1^{(m)}\}, \ldots, \{u_n^{(m)}\}$ under the initial condition (3.7). In our analysis, we will use Corollary 1.1 proved in [118] (see Sect. 1.2).

The point estimation approach presented in this section consists of the following main steps:

1° If $c_n \leq 1/(2n)$ and (3.7) holds, from Corollary 1.1, it follows that the inequalities

$$|u_i^{(0)}| = |z_i^{(0)} - \zeta_i| < \frac{|W_i^{(0)}|}{1 - nc_n} \tag{3.69}$$

are valid for each $i \in \boldsymbol{I}_n$. These inequalities have an important role in the estimation procedure involved in the convergence analysis of the sequences $\big\{z_i^{(m)}\big\}$, produced by the considered simultaneous method.

2° In the next step, we derive the inequalities

$$d < \tau_n \hat{d} \quad \text{and} \quad |\widehat{W}_i| < \beta_n |W_i|,$$

which involve the minimal distances and the Weierstrass' corrections at two successive iterative steps. The i-factor c_n appearing in (3.7) has to be chosen to provide such values of τ_n and β_n which give the following implication

$$w < c_n d \implies \widehat{w} < c_n \hat{d},$$

important in the proof of convergence theorems by induction. Let us note that the above implication will hold if $\tau_n \beta_n < 1$.

3° In the final step, we derive the inequalities of the form

$$|u_i^{(m+1)}| \le \gamma(n, d^{(m)}) |u_i^{(m)}|^p \Bigg(\sum_{\substack{j=1 \\ j \ne i}}^{n} |u_j^{(m)}|^q \Bigg)^r \tag{3.70}$$

for $i = 1, \ldots, n$ and $m = 0, 1, \ldots$, and prove that all sequences $\{|u_1^{(m)}|\}, \ldots, \{|u_n^{(m)}|\}$ tend to 0 under the condition (3.7) (with suitably chosen c_n), which means that $z_i^{(m)} \to \zeta_i$ $(i \in \boldsymbol{I}_n)$. The order of convergence of these sequences is obtained from (3.70) and it is equal to $p + qr$.

To study iterative methods which do not involve Weierstrass' corrections W_i, appearing in the initial conditions of the form (3.7), it is necessary to establish a suitable relation between Newton's correction $P(z_i)/P'(z_i)$ and Weierstrass' correction W_i. Applying the logarithmic derivative to $P(t)$, represented by the Lagrangean interpolation formula (3.27) (for $z = t$), one obtains

$$\frac{P'(t)}{P(t)} = \sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{1}{t - z_j} + \frac{\displaystyle\sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j}{t - z_j} + 1 - (t - z_i) \sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j}{(t - z_j)^2}}{W_i + (t - z_i) \Bigg[\displaystyle\sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j}{t - z_j} + 1 \Bigg]}. \tag{3.71}$$

Putting $t = z_i$ in (3.71), we get Carstensen's identity [15]

$$\frac{P'(z_i)}{P(z_i)} = \sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{1}{z_i - z_j} + \frac{\displaystyle\sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j}{z_i - z_j} + 1}{W_i}. \tag{3.72}$$

In what follows, we will apply the three-stage aforementioned procedure for the convergence analysis of some frequently used simultaneous methods.

The Ehrlich–Aberth's Method

In this part, we use Newton's and Weierstrass' correction, given, respectively, by

$$N_i^{(m)} = \frac{P(z_i^{(m)})}{P'(z_i^{(m)})} \quad \text{and} \quad W_i^{(m)} = \frac{P(z_i^{(m)})}{\prod_{\substack{j=1 \\ j \neq i}}^{n} (z_i^{(m)} - z_j^{(m)})} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots).$$

We are concerned here with one of the most efficient numerical methods for the simultaneous approximation of all zeros of a polynomial, given by the iterative formula

$$z_i^{(m+1)} = z_i^{(m)} - \frac{1}{\dfrac{1}{N_i^{(m)}} - \sum_{\substack{j=1 \\ j \neq i}}^{n} \dfrac{1}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.73}$$

Our aim is to state practically applicable initial conditions of the form (3.7), which enable a guaranteed convergence of the Ehrlich–Aberth's method (3.73). First, we present a lemma concerned with the localization of polynomial zeros.

Lemma 3.13. *Assume that the following condition*

$$w < c_n d, \quad c_n = \begin{cases} \dfrac{1}{2n + 1.4}, & 3 \le n \le 7 \\[2ex] \dfrac{1}{2n}, & n \ge 8 \end{cases} \tag{3.74}$$

is satisfied. Then, each disk $\left\{ z_i; \dfrac{1}{1 - nc_n} |W_i| \right\}$ $(i \in \boldsymbol{I}_n)$ *contains one and only one zero of* P.

The above assertion follows from Corollary 1.1 under the condition (3.74).

Lemma 3.14. *Let* $z_1, \ldots, z_n$ *be distinct approximations to the zeros* $\zeta_1, \ldots, \zeta_n$ *of a polynomial* P *of degree* n *and let* $\hat{z}_1, \ldots, \hat{z}_n$ *be new respective approximations obtained by the Ehrlich–Aberth's method* (3.73). *Then, the following formula is valid:*

$$\widehat{W}_i = -(\hat{z}_i - z_i)^2 \sum_{j \neq i} \frac{W_j}{(\hat{z}_i - z_j)(z_i - z_j)} \prod_{j \neq i} \Big(1 + \frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j} \Big). \tag{3.75}$$

Proof. From the iterative formula (3.73), one obtains

$$\frac{1}{\hat{z}_i - z_i} = \sum_{j \neq i} \frac{1}{z_i - z_j} - \frac{P'(z_i)}{P(z_i)},$$

so that, using (3.72),

$$\begin{aligned}\frac{W_i}{\hat{z}_i - z_i} &= W_i\left(\sum_{j \neq i} \frac{1}{z_i - z_j} - \frac{P'(z_i)}{P(z_i)}\right) = -W_i\left[\frac{1}{W_i}\Big(\sum_{j \neq i} \frac{W_j}{z_i - z_j} + 1\Big)\right]\\ &= -\sum_{j \neq i} \frac{W_j}{z_i - z_j} - 1.\end{aligned}$$

According to this, we have

$$\begin{aligned}\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1 &= \frac{W_i}{\hat{z}_i - z_i} + \sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j} + 1 = -\sum_{j \neq i} \frac{W_j}{z_i - z_j} + \sum_{j \neq i} \frac{W_j}{\hat{z}_i - z_j}\\ &= -(\hat{z}_i - z_i)\sum_{j \neq i} \frac{W_j}{(\hat{z}_i - z_j)(z_i - z_j)}.\end{aligned}$$

Taking into account the last expression, returning to (3.27), we find for $z = \hat{z}_i$

$$\begin{aligned}P(\hat{z}_i) &= \Big(\sum_{j=1}^{n} \frac{W_j}{\hat{z}_i - z_j} + 1\Big)\prod_{j=1}^{n}(\hat{z}_i - z_j)\\ &= -(\hat{z}_i - z_i)^2 \sum_{j \neq i} \frac{W_j}{(\hat{z}_i - z_j)(z_i - z_j)} \prod_{j \neq i}(\hat{z}_i - z_j).\end{aligned}$$

After dividing by $\prod_{j \neq i}(\hat{z}_i - \hat{z}_j)$ and some rearrangement, we obtain (3.75). □

Let us introduce the abbreviations:

$$\rho_n = \frac{1}{1 - nc_n}, \quad \gamma_n = \frac{1}{1 - \rho_n c_n - (n-1)(\rho_n c_n)^2},$$

$$\lambda_n = \rho_n c_n (1 - \rho_n c_n)\gamma_n, \quad \beta_n = \frac{(n-1)\lambda_n^2}{1 - \lambda_n}\Big(1 + \frac{\lambda_n}{1 - 2\lambda_n}\Big)^{n-1}.$$

Lemma 3.15. *Let $\hat{z}_1, \ldots, \hat{z}_n$ be approximations produced by the Ehrlich–Aberth's method* (3.73) *and let $u_i = z_i - \zeta_i$ and $\hat{u}_i = \hat{z}_i - \zeta_i$. If $n \geq 3$ and the inequality* (3.74) *holds, then:*

(i) $d < \dfrac{1}{1 - 2\lambda_n}\hat{d}$.

(ii) $\widehat{w} < \beta_n w$.
(iii) $\widehat{w} < c_n \hat{d}$.
(iv) $|\hat{u}_i| \le \dfrac{\gamma_n}{d^2}|u_i|^2 \sum_{j\ne i} |u_j|$.

Proof. According to the initial condition (3.74) and Lemma 3.13, we have

$$|u_i| = |z_i - \zeta_i| \le \rho_n |W_i| \le \rho_n w < \rho_n c_n d. \tag{3.76}$$

In view of (3.76) and the definition of the minimal distance d, we find

$$|z_j - \zeta_i| \ge |z_j - z_i| - |z_i - \zeta_i| > d - \rho_n c_n d = (1 - \rho_n c_n)d. \tag{3.77}$$

Using the identity

$$\frac{P'(z_i)}{P(z_i)} = \sum_{j=1}^{n} \frac{1}{z_i - \zeta_j} = \frac{1}{u_i} + \sum_{j\ne i} \frac{1}{z_i - \zeta_j}, \tag{3.78}$$

from (3.73), we get

$$\begin{aligned} \hat{u}_i = \hat{z}_i - \zeta_i &= z_i - \zeta_i - \cfrac{1}{\cfrac{1}{u_i} + \sum_{j\ne i} \cfrac{1}{z_i - \zeta_j} - \sum_{j\ne i} \cfrac{1}{z_i - z_j}} \\ &= u_i - \frac{u_i}{1 - u_i S_i} = -\frac{u_i^2 S_i}{1 - u_i S_i}, \end{aligned} \tag{3.79}$$

where $S_i = \sum_{j\ne i} \dfrac{u_j}{(z_i - \zeta_j)(z_i - z_j)}$.

Using the definition of d and the bounds (3.76) and (3.77), we estimate

$$\begin{aligned} |u_i S_i| &\le |u_i| \sum_{j\ne i} \frac{|u_j|}{|z_i - \zeta_j||z_i - z_j|} < \rho_n c_n d \cdot \frac{(n-1)\rho_n c_n d}{(1 - \rho_n c_n)d^2} \\ &= \frac{(\rho_n c_n)^2 (n-1)}{1 - \rho_n c_n}. \end{aligned} \tag{3.80}$$

Now, by (3.76) and (3.80), from (3.73) we find

$$\begin{aligned} |\hat{z}_i - z_i| &= \left|\frac{u_i}{1 - u_i S_i}\right| \le \frac{|u_i|}{1 - |u_i S_i|} < \cfrac{|u_i|}{1 - \cfrac{(\rho_n c_n)^2 (n-1)}{1 - \rho_n c_n}} \\ &< \frac{\rho_n c_n (1 - \rho_n c_n)}{1 - \rho_n c_n - (\rho_n c_n)^2 (n-1)} d = \rho_n c_n (1 - \rho_n c_n)\gamma_n d \\ &= \lambda_n d \end{aligned} \tag{3.81}$$

and also

$$|\hat{z}_i - z_i| < (1 - \rho_n c_n)\gamma_n |u_i| < (1 - \rho_n c_n)\gamma_n \rho_n |W_i|. \tag{3.82}$$

Having in mind (3.81), according to Lemma 3.2, we conclude that the estimates $|\hat{z}_i - z_j| > (1 - \lambda_n)d$ and $|\hat{z}_i - \hat{z}_j| > (1 - 2\lambda_n)d$ ($i \in \boldsymbol{I}_n$) hold. From the last inequality, we find

$$\frac{d}{\hat{d}} < \frac{1}{1 - 2\lambda_n} \quad \text{for every } n \geq 3, \tag{3.83}$$

which proves the assertion (i) of Lemma 3.15.

Using the starting inequality $w/d < c_n$ and the bounds (3.81), (3.82), (3.14), (3.15), and (3.16), we estimate the quantities involved in (3.75):

$$\begin{aligned}
|\widehat{W}_i| &\leq |\hat{z}_i - z_i|^2 \sum_{j \neq i} \frac{|W_j|}{|\hat{z}_i - z_j||z_i - z_j|} \prod_{j \neq i} \Big(1 + \frac{|\hat{z}_j - z_j|}{|\hat{z}_i - \hat{z}_j|}\Big) \\
&< \frac{(n-1)\lambda_n^2}{(1 - \lambda_n)} \Big(1 + \frac{\lambda_n}{1 - 2\lambda_n}\Big)^{n-1} |W_i| \\
&= \beta_n |W_i|.
\end{aligned}$$

Therefore, we have

$$\widehat{w} < \beta_n w \tag{3.84}$$

so that, by (3.74), (3.83), and (3.84), we estimate

$$\widehat{w} < \beta_n w < \beta_n c_n d < \frac{\beta_n}{1 - 2\lambda_n} c_n \hat{d}.$$

Since

$$\frac{\beta_n}{1 - 2\lambda_n} < 0.95 < 1 \quad \text{for all} \quad 3 \leq n \leq 7$$

and

$$\frac{\beta_n}{1 - 2\lambda_n} < 0.78 < 1 \quad \text{for all} \quad n \geq 8,$$

we have

$$\widehat{w} < c_n \hat{d}, \quad n \geq 3.$$

In this way, we have proved the assertions (ii) and (iii) of Lemma 3.15.

Using the previously derived bounds, we find

$$\begin{aligned}
|\hat{u}_i| &\leq \frac{|u_i|^2 |S_i|}{1 - |u_i S_i|} < \frac{|u_i|^2}{1 - \dfrac{(\rho c_n)^2 (n-1)}{1 - \rho_n c_n}} \sum_{j \neq i} \frac{|u_j|}{|z_i - \zeta_j||z_i - z_j|} \\
&< \frac{1 - \rho_n c_n}{1 - \rho_n c_n - (\rho_n c_n)^2 (n-1)} |u_i|^2 \sum_{j \neq i} \frac{|u_j|}{(1 - \rho_n c_n) d^2} \\
&= \frac{1}{(1 - \rho_n c_n - (\rho_n c_n)^2 (n-1)) d^2} |u_i|^2 \sum_{j \neq i} |u_j|,
\end{aligned}$$

wherefrom

$$|\hat{u}_i| < \frac{\gamma_n}{d^2}|u_i|^2 \sum_{j\neq i} |u_j|. \tag{3.85}$$

This strict inequality is derived assuming that $u_i \neq 0$ (see Remark 3.8). If we include the case $u_i = 0$, then it follows

$$|\hat{u}_i| \leq \frac{\gamma_n}{d^2}|u_i|^2 \sum_{j\neq i} |u_j|$$

and the assertion (iv) of Lemma 3.15 is proved. □

Remark 3.7. In what follows, the assertions of the form (i)–(iv) of Lemma 3.15 will be presented for the three other methods, but for different *i*-factor c_n and specific entries of λ_n, β_n, and γ_n.

We now give the convergence theorem for the Ehrlich–Aberth's method (3.73), which involves only initial approximations to the zeros, the polynomial coefficients, and the polynomial degree n.

Theorem 3.9. *Under the initial condition*

$$w^{(0)} < c_n d^{(0)}, \tag{3.86}$$

where c_n *is given by* (3.74)*, the Ehrlich–Aberth's method* (3.73) *is convergent with the cubic convergence.*

Proof. The convergence analysis is based on the estimation procedure of the errors $u_i^{(m)} = z_i^{(m)} - \zeta_i$ $(i \in \boldsymbol{I}_n)$. The proof is by induction with the argumentation used for the inequalities (i)–(iv) of Lemma 3.15. Since the initial condition (3.86) coincides with (3.74), all estimates given in Lemma 3.15 are valid for the index $m = 1$. Actually, this is the part of the proof with respect to $m = 1$. Furthermore, the inequality (iii) again reduces to the condition of the form (3.74) and, therefore, the assertions (i)–(iv) of Lemma 3.15 hold for the next index, and so on. All estimates and bounds for the index m are derived essentially in the same way as for $m = 0$. In fact, the implication

$$w^{(m)} < c_n d^{(m)} \implies w^{(m+1)} < c_n d^{(m+1)}$$

plays the key role in the convergence analysis of the Ehrlich–Aberth's method (3.73) because it involves the initial condition (3.86), which enables the validity of all inequalities given in Lemma 3.15 for all $m = 0, 1, \ldots$. In particular, regarding (3.83) and (3.85), we have

$$\frac{d^{(m)}}{d^{(m+1)}} < \frac{1}{1 - 2\lambda_n} \tag{3.87}$$

and

$$|u_i^{(m+1)}| \le \frac{\gamma_n}{(d^{(m)})^2} |u_i^{(m)}|^2 \sum_{\substack{j=1 \\ j \ne i}}^{n} |u_j^{(m)}| \quad (i \in \boldsymbol{I}_n) \tag{3.88}$$

for each iteration index $m = 0, 1, \ldots$ if (3.86) holds.

Substituting

$$t_i^{(m)} = \Big[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(m)})^2}\Big]^{1/2} |u_i^{(m)}|,$$

the inequalities (3.88) become

$$t_i^{(m+1)} \le \frac{(1-2\lambda_n)d^{(m)}}{(n-1)d^{(m+1)}} \big(t_i^{(m)}\big)^2 \sum_{\substack{j=1 \\ j \ne i}}^{n} t_j^{(m)},$$

wherefrom, by (3.87),

$$t_i^{(m+1)} < \frac{\big(t_i^{(m)}\big)^2}{n-1} \sum_{\substack{j=1 \\ j \ne i}}^{n} t_j^{(m)} \quad (i \in \boldsymbol{I}_n). \tag{3.89}$$

By virtue of (3.76), we find

$$\begin{aligned} t_i^{(0)} &= \sqrt{\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(0)})^2}} |u_i^{(0)}| < \rho_n c_n d^{(0)} \sqrt{\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(0)})^2}} \\ &= \rho_n c_n \sqrt{\frac{(n-1)\gamma_n}{1-2\lambda_n}} \end{aligned}$$

for each $i = 1, \ldots, n$. Taking

$$t = \max_{1 \le i \le n} t_i^{(0)} < \rho_n c_n \sqrt{\frac{(n-1)\gamma_n}{1-2\lambda_n}},$$

we come to the inequalities

$$t_i^{(0)} \le t < 0.571 < 1 \quad (3 \le n \le 7)$$

and

$$t_i^{(0)} \le t < 0.432 < 1 \quad (n \ge 8)$$

for all $i = 1, \ldots, n$. According to this, from (3.89), we conclude that the sequences $\{t_i^{(m)}\}$ (and, consequently, $\{|u_i^{(m)}|\}$) tend to 0 for all $i = 1, \ldots, n$. Therefore, the Ehrlich–Aberth's method (3.73) is convergent.

Taking into account that the quantity $d^{(m)}$, which appears in (3.88), is bounded (see the proof of Theorem 5.1) and tends to $\min_{i \ne j} |\zeta_i - \zeta_j|$ and setting

$$u^{(m)} = \max_{1\le i\le n} |u_i^{(m)}|,$$

from (3.88), we obtain

$$|u_i^{(m+1)}| \le u^{(m+1)} < \frac{(n-1)\gamma_n}{d^{(m)}}|u^{(m)}|^3,$$

which proves the cubic convergence. □

Remark 3.8. As usual in the convergence analysis of iterative methods (see, e.g., [48]), we could assume that the errors $u_i^{(m)}=z_i^{(m)} - \zeta_i$ $(i \in \boldsymbol{I}_n)$ do not reach 0 for a finite m. However, if $u_i^{(m_0)}=0$ for some indices $i_1,\dots,i_k$ and $m_0 \ge 0$, we just take $z_{i_1}^{(m_0)},\dots,z_{i_k}^{(m_0)}$ as approximations to the zeros $\zeta_{i_1},\dots,\zeta_{i_k}$ and do not iterate further for the indices $i_1,\dots,i_k$. If the sequences $\{u_i^{(m)}\}$ $(i \in \boldsymbol{I}_n \setminus \{i_1,\dots,i_k\})$ have the order of convergence q, then obviously the sequences $\{u_{i_1}^{(m)}\},\dots,\{u_{i_k}^{(m)}\}$ converge with the convergence rate at least q. This remark refers not only to the iterative method (3.73) but also to all methods considered in this book. For this reason, we do not discuss this point further.

The Ehrlich–Aberth's Method with Newton's Corrections

The convergence of the Ehrlich–Aberth's method (3.1) can be accelerated using Newton's corrections $N_i^{(m)} = P(z_i^{(m)})/P'(z_i^{(m)})$ $(i \in \boldsymbol{I}_n,\ m = 0,1,\dots)$. In this way, the following method for the simultaneous approximation of all simple zeros of a given polynomial P can be established

$$z_i^{(m+1)} = z_i^{(m)} - \frac{1}{\dfrac{1}{N_i^{(m)}} - \displaystyle\sum_{j\ne i}\frac{1}{z_i^{(m)} - z_j^{(m)} + N_j^{(m)}}} \quad (i \in \boldsymbol{I}_n), \tag{3.90}$$

where $m = 0,1,\dots$, see Sect. 1.1. This method will be briefly called the EAN method.

From Corollary 1.1, the following lemma can be stated.

Lemma 3.16. *Let $z_1,\dots,z_n$ be distinct numbers satisfying the inequality*

$$w < c_n d, \quad c_n = \begin{cases} \dfrac{1}{2.2n+1.9}, & 3 \le n \le 21 \\ \dfrac{1}{2.2n}, & n \ge 22 \end{cases}. \tag{3.91}$$

Then, the disks $\left\{z_1; \dfrac{1}{1-nc_n}|W_1|\right\},\dots,\left\{z_n; \dfrac{1}{1-nc_n}|W_n|\right\}$ are mutually disjoint and each of them contains exactly one zero of a polynomial P.

We now give the expression for the improved Weierstrass' correction $\widehat{W}_i$.

Lemma 3.17. *Let $z_1, \ldots, z_n$ be distinct approximations to the zeros $\zeta_1, \ldots, \zeta_n$ of a polynomial P of degree n and let $\hat{z}_1, \ldots, \hat{z}_n$ be new respective approximations obtained by the EAN method* (3.90). *Then, the following formula is valid:*

$$\widehat{W}_i = -(\hat{z}_i - z_i)\Big(W_i \Sigma_{N,i} + (\hat{z}_i - z_i)\Sigma_{W,i}\Big) \prod_{j \neq i} \Big(1 + \frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\Big), \tag{3.92}$$

where

$$\Sigma_{N,i} = \sum_{j \neq i} \frac{N_j}{(z_i - z_j + N_j)(z_i - z_j)}, \quad \Sigma_{W,i} = \sum_{j \neq i} \frac{W_j}{(\hat{z}_i - z_j)(z_i - z_j)}.$$

The relation (3.92) is obtained by combining the Lagrangean interpolation formula (3.27) for $z = \hat{z}_i$, the iterative formula (3.90), and the identity (3.72). Since the proving technique of Lemma 3.17 is a variation on earlier procedure applied in the proof of Lemma 3.14, we shall pass over it lightly. The complete proof can be found in [119].

We introduce the abbreviations:

$$\rho_n = \frac{1}{1 - nc_n}, \quad \delta_n = 1 - \rho_n c_n - (n-1)\rho_n c_n,$$

$$\alpha_n = (1 - \rho_n c_n)((1 - \rho_n c_n)^2 - (n-1)\rho_n c_n),$$

$$\gamma_n = \frac{n-1}{\alpha_n - (n-1)^2(\rho_n c_n)^3}, \quad \lambda_n = \frac{\alpha_n \gamma_n \rho_n c_n}{n-1},$$

$$\beta_n = \lambda_n (n-1)\Big(\frac{(1 - \rho_n c_n)^2 \rho_n c_n}{\alpha_n} + \frac{\lambda_n}{1 - \lambda_n}\Big)\Big(1 + \frac{\lambda_n}{1 - 2\lambda_n}\Big)^{n-1}.$$

Lemma 3.18. *Let $\hat{z}_1, \ldots, \hat{z}_n$ be approximations generated by the EAN method* (3.90) *and let $u_i = z_i - \zeta_i$, $\hat{u}_i = \hat{z}_i - \zeta_i$. If $n \geq 3$ and the inequality* (3.91) *holds, then:*

(i) $d < \dfrac{1}{1 - 2\lambda_n}\hat{d}$.
(ii) $\widehat{w} < \beta_n w$.
(iii) $\widehat{w} < c_n \hat{d}$.
(iv) $|\hat{u}_i| \leq \dfrac{\gamma_n}{d^3} |u_i|^2 \displaystyle\sum_{j \neq i} |u_j|^2$.

Proof. In regard to (3.91) and Lemma 3.16, we have $\zeta_i \in \Big\{z_i; \dfrac{1}{1 - nc_n}|W_i|\Big\}$ $(i \in \boldsymbol{I}_n)$, so that

$$|u_i| = |z_i - \zeta_i| \leq \rho_n |W_i| \leq \rho_n w < \rho_n c_n d. \tag{3.93}$$

According to this and the definition of the minimal distance d, we find

$$|z_j - \zeta_i| \geq |z_j - z_i| - |z_i - \zeta_i| > d - \rho_n c_n d = (1 - \rho_n c_n)d. \tag{3.94}$$

Using the identity (3.78) and the estimates (3.93) and (3.94), we obtain

$$\begin{aligned}\left|\frac{P'(z_i)}{P(z_i)}\right| &= \left|\sum_{j=1}^{n} \frac{1}{z_i - \zeta_j}\right| \geq \frac{1}{|z_i - \zeta_i|} - \sum_{j \neq i} \frac{1}{|z_i - \zeta_j|} > \frac{1}{\rho_n c_n d} - \frac{n-1}{(1 - \rho_n c_n)d} \\ &= \frac{1 - \rho_n c_n - (n-1)\rho_n c_n}{(1 - \rho_n c_n)\rho_n c_n d} = \frac{\delta_n}{(1 - \rho_n c_n)\rho_n c_n d}.\end{aligned}$$

Hence

$$|N_i| = \left|\frac{P(z_i)}{P'(z_i)}\right| < \frac{(1 - \rho_n c_n)\rho_n c_n d}{\delta_n}, \tag{3.95}$$

so that

$$\begin{aligned}|z_i - z_j + N_j| &\geq |z_i - z_j| - |N_j| > d - \frac{(1 - \rho_n c_n)\rho_n c_n d}{\delta_n} \\ &= \frac{(1 - \rho_n c_n)^2 - (n-1)\rho_n c_n}{\delta_n} d = \frac{\alpha_n}{\delta_n (1 - \rho_n c_n)} d. \end{aligned} \tag{3.96}$$

Let us introduce

$$S_i = \sum_{j \neq i} \frac{N_j - u_j}{(z_i - \zeta_j)(z_i - z_j + N_j)}, \quad h_j = \sum_{k \neq j} \frac{1}{z_j - \zeta_k}.$$

We start from the iterative formula (3.90) and use the identity (3.78) to find

$$\begin{aligned}\hat{u}_i = \hat{z}_i - \zeta_i &= z_i - \zeta_i - \frac{1}{\dfrac{1}{u_i} + \displaystyle\sum_{j \neq i} \frac{1}{z_i - \zeta_j} - \sum_{j \neq i} \frac{1}{z_i - z_j + N_j}} \\ &= u_i - \frac{u_i}{1 + u_i \displaystyle\sum_{j \neq i} \frac{N_j - u_j}{(z_i - \zeta_j)(z_i - z_j + N_j)}} \\ &= u_i - \frac{u_i}{1 + u_i S_i} = \frac{u_i^2 S_i}{1 + u_i S_i}. \end{aligned} \tag{3.97}$$

Furthermore, we find

$$N_j = \frac{u_j}{1 + u_j h_j}, \quad N_j - u_j = -\frac{u_j^2 h_j}{1 + u_j h_j}, \quad S_i = -\sum_{j \neq i} \frac{\dfrac{u_j^2 h_j}{1 + u_j h_j}}{(z_i - \zeta_j)(z_i - z_j + N_j)}.$$

Using (3.93) and the inequality

$$|h_j| = \Big|\sum_{k\neq j} \frac{1}{z_j - \zeta_k}\Big| < \frac{n-1}{(1-\rho_n c_n)d},$$

we find

$$\Big|\frac{h_j}{1+u_j h_j}\Big| \le \frac{|h_j|}{1-|u_j||h_j|} < \frac{\dfrac{n-1}{(1-\rho_n c_n)d}}{1-\rho_n c_n d\dfrac{n-1}{(1-\rho_n c_n)d}} = \frac{n-1}{\delta_n d}. \tag{3.98}$$

Combining (3.93), (3.94), (3.96), and (3.98), we obtain

$$\begin{aligned} |u_i S_i| &\le |u_i| \sum_{j\neq i} \frac{|u_j|^2 \Big|\dfrac{h_j}{1+u_j h_j}\Big|}{|z_i - \zeta_j||z_i - z_j + N_j|} \\ &< \rho_n c_n d \cdot \frac{(n-1)(\rho_n c_n d)^2 \dfrac{n-1}{\delta_n d}}{(1-\rho_n c_n)d\, \dfrac{\alpha_n}{\delta_n(1-\rho_n c_n)} d} \\ &= \frac{(n-1)^2(\rho_n c_n)^3}{\alpha_n}. \end{aligned} \tag{3.99}$$

Using (3.93) and (3.99), from (3.90), we find

$$\begin{aligned} |\hat{z}_i - z_i| &= \Big|\frac{u_i}{1+u_i S_i}\Big| \le \frac{|u_i|}{1-|u_i S_i|} < \frac{|u_i|}{1-\dfrac{(n-1)^2(\rho_n c_n)^3}{\alpha_n}} \\ &= \frac{\alpha_n}{\alpha_n - (n-1)^2(\rho_n c_n)^3}|u_i| < \frac{\alpha_n \rho_n c_n \gamma_n}{n-1} d = \lambda_n d \end{aligned}$$

and

$$|\hat{z}_i - z_i| < \frac{\alpha_n}{\alpha_n - (n-1)^2(\rho_n c_n)^3}|u_i| < \frac{\alpha_n \rho_n \gamma_n}{n-1}|W_i| = \frac{\lambda_n}{c_n}|W_i| < \lambda_n d. \tag{3.100}$$

Since (3.100) holds, we apply Lemma 3.2 and obtain

$$d < \frac{1}{1-2\lambda_n}\hat{d} \tag{3.101}$$

from (3.15). Thus, the assertion (i) of Lemma 3.18 is valid.

Using the starting inequality $w/d < c_n$ and the bounds (3.95), (3.96), (3.100), (3.14), and (3.15), for $n \ge 3$, we estimate the quantities appearing in (3.92):

$$|W_i||\Sigma_{N,i}| < w\,\frac{(n-1)\dfrac{(1-\rho_n c_n)\rho_n c_n d}{\delta_n}}{\dfrac{\alpha_n}{\delta_n(1-\rho_n c_n)}d^2} < \frac{(n-1)(1-\rho_n c_n)^2\rho_n c_n^2}{\alpha_n},$$

$$|\hat{z}_i - z_i||\Sigma_{W,i}| < \lambda_n d\,\frac{(n-1)c_n d}{(1-\lambda_n)d\cdot d} < \frac{(n-1)\lambda_n c_n}{1-\lambda_n}.$$

According to the last two bounds and (3.16), from (3.92), we estimate

$$\begin{aligned}|\widehat{W}_i| &\le |\hat{z}_i - z_i|\Big(|W_i||\Sigma_{N,i}| + |\hat{z}_i - z_i||\Sigma_{W,i}|\Big)\Big|\prod_{j\neq i}\Big(1+\frac{\hat{z}_j - z_j}{\hat{z}_i - \hat{z}_j}\Big)\Big| \\ &< \frac{\lambda_n}{c_n}|W_i|\Big(\frac{(n-1)(1-\rho_n c_n)^2\rho_n c_n^2}{\alpha_n} + \frac{(n-1)\lambda_n c_n}{1-\lambda_n}\Big)\Big(1+\frac{\lambda_n}{1+2\lambda_n}\Big)^{n-1} \\ &= \beta_n|W_i| \le \beta_n w,\end{aligned}$$

i.e.,

$$\widehat{w} < \beta_n w. \tag{3.102}$$

Therefore, we have proved the assertion (ii) of Lemma 3.18.

Since

$$\frac{\beta_n}{1-2\lambda_n} < 0.942 \qquad \text{for} \quad 3 \le n \le 21$$

and

$$\frac{\beta_n}{1-2\lambda_n} < 0.943 \qquad \text{for} \quad n \ge 22,$$

starting from (3.102), by (3.91) and (3.101), we find

$$\widehat{w} < \beta_n w < \beta_n c_n d < \frac{\beta_n}{1-2\lambda_n}\cdot c_n\hat{d} < c_n\hat{d},$$

which means that the implication $w < c_n d \;\Rightarrow\; \widehat{w} < c_n d$ holds. This proves (iii) of Lemma 3.18.

Using the above bounds, from (3.97), we obtain

$$\begin{aligned}|\hat{u}_i| &\le \frac{|u_i|^2|S_i|}{1-|u_i S_i|} < \frac{\alpha_n}{\alpha_n - (n-1)^2(\rho_n c_n)^3}|u_i|^2\sum_{j\neq i}\frac{|u_j|^2\Big|\dfrac{h_j}{1+u_j h_j}\Big|}{|z_i - \zeta_j||z_i - z_j + N_j|} \\ &< \frac{\alpha_n}{\alpha_n - (n-1)^2(\rho_n c_n)^3}\,\frac{\dfrac{n-1}{\delta_n d}}{(1-\rho_n c_n)d\,\dfrac{\alpha_n}{\delta_n(1-\rho_n c_n)}d}|u_i|^2\sum_{j\neq i}|u_j|^2,\end{aligned}$$

wherefrom (taking into account Remark 3.8)

$$|\hat{u}_i| \le \frac{\gamma_n}{d^3}|u_i|^2 \sum_{j\ne i} |u_j|^2,$$

which proves (iv) of Lemma 3.18. □

Now, we give the convergence theorem for the EAN method (3.90).

Theorem 3.10. *Let P be a polynomial of degree $n \ge 3$ with simple zeros. If the initial condition*

$$w^{(0)} < c_n d^{(0)} \tag{3.103}$$

holds, where c_n is given by (3.91)*, then the EAN method* (3.90) *is convergent with the order of convergence 4.*

Proof. Similarly to the proof of Theorem 3.9, we apply induction with the argumentation used for the inequalities (i)–(iv) of Lemma 3.18. According to (3.103) and (3.91), all estimates given in Lemma 3.18 are valid for the index $m = 1$. We notice that the inequality (iii) coincides with the condition of the form (3.103), and hence, the assertions (i)–(iv) of Lemma 3.18 are valid for the next index, etc. The implication

$$w^{(m)} < c_n d^{(m)} \implies w^{(m+1)} < c_n d^{(m+1)}$$

provides the validity of all inequalities given in Lemma 3.18 for all $m = 0, 1, \ldots$. In particular, we have

$$\frac{d^{(m)}}{d^{(m+1)}} < \frac{1}{1 - 2\lambda_n} \tag{3.104}$$

and

$$|u_i^{(m+1)}| \le \frac{\gamma_n}{(d^{(m)})^3}|u_i^{(m)}|^2 \sum_{\substack{j=1 \\ j\ne i}}^{n} |u_j^{(m)}|^2 \quad (i \in \boldsymbol{I}_n) \tag{3.105}$$

for each iteration index $m = 0, 1, \ldots,$ where

$$\gamma_n = \frac{n-1}{\alpha_n - (n-1)^2(\rho_n c_n)^3}.$$

Substituting

$$t_i^{(m)} = \left[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(m)})^3}\right]^{1/3} |u_i^{(m)}|$$

into (3.105) yields

$$t_i^{(m+1)} \le \frac{(1-2\lambda_n)d^{(m)}}{(n-1)d^{(m+1)}} \big(t_i^{(m)}\big)^2 \sum_{\substack{j=1 \\ j\ne i}}^{n} \big(t_j^{(m)}\big)^2 \quad (i \in \boldsymbol{I}_n).$$

Hence, using (3.104), we obtain

$$t_i^{(m+1)} < \frac{1}{n-1}[t_i^{(m)}]^2 \sum_{\substack{j=1 \\ j\neq i}}^{n} [t_j^{(m)}]^2 \quad (i \in \boldsymbol{I}_n). \tag{3.106}$$

Using (3.93), we find

$$\begin{aligned} t_i^{(0)} &= \left[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(0)})^3}\right]^{1/3} |u_i^{(0)}| < \rho_n c_n d^{(0)} \left[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(0)})^3}\right]^{1/3} \\ &= \rho_n c_n \left[\frac{(n-1)\gamma_n}{1-2\lambda_n}\right]^{1/3}. \end{aligned}$$

Taking $t = \max_{1\leq i\leq n} t_i^{(0)}$ yields

$$t_i^{(0)} \leq t < 0.626 < 1 \quad (3 \leq n \leq 21)$$

and

$$t_i^{(0)} \leq t < 0.640 < 1 \quad (n \geq 22)$$

for each $i = 1, \ldots, n$. In regard to this, we conclude from (3.106) that the sequences $\{t_i^{(m)}\}$ and $\{|u_i^{(m)}|\}$ tend to 0 for all $i = 1, \ldots, n$, meaning that $z_i^{(m)} \to \zeta_i$. Therefore, the EAN method (3.90) is convergent. Besides, taking into account that the quantity $d^{(m)}$ appearing in (3.105) is bounded and tends to $\min_{\substack{1\leq i,j\leq n \\ i\neq j}} |\zeta_i - \zeta_j|$ and setting

$$u^{(m)} = \max_{1\leq i\leq n} |u_i^{(m)}|,$$

from (3.105), we obtain

$$|u_i^{(m+1)}| \leq u^{(m+1)} < (n-1)\frac{\gamma_n}{(d^{(m)})^3}\big(u^{(m)}\big)^4,$$

which means that the order of convergence of the EAN method is 4. □

The Börsch-Supan's Method with Weierstrass' Correction

The cubically convergent Börsch-Supan's method

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \sum_{j\neq i} \dfrac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots),$$

presented in [10], can be accelerated by using Weierstrass' corrections $W_i^{(m)} = P(z_i^{(m)})/\prod_{j\neq i}(z_i^{(m)}-z_j^{(m)})$. In this manner, we obtain the following iterative formula (see Nourein [95])

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1+\displaystyle\sum_{j\neq i}\frac{W_j^{(m)}}{z_i^{(m)}-W_i^{(m)}-z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m=0,1,\ldots). \tag{3.107}$$

The order of convergence of the Börsch-Supan's method with Weierstrass' corrections (3.107) is *4* (see, e.g., [16], [188]). For brevity, the method (3.107) will be referred to as the BSW method.

Let us introduce the abbreviations:

$$\rho_n = \frac{1}{1-nc_n}, \quad \gamma_n = \frac{\rho_n(1+\rho_n c_n)^{2n-2}}{(1-\rho_n c_n)^2},$$

$$\lambda_n = \rho_n c_n(1-c_n), \quad \beta_n = \frac{\lambda_n\rho_n c_n^2(n-1)^2}{(1-\lambda_n)(1-c_n)}\Big(1+\frac{\lambda_n}{1-2\lambda_n}\Big)^{n-1}.$$

Lemma 3.19. *Let* $\hat z_1,\ldots,\hat z_n$ *be approximations obtained by the iterative method* (3.107) *and let* $\hat u_i = \hat z_i - \zeta_i$, $\hat d = \min_{i\neq j}|\hat z_i - \hat z_j|$, *and* $\widehat{w} = \max_{1\le i\le n}|\widehat{W}_i|$. *If the inequality*

$$w < c_n d, \qquad c_n = \begin{cases} \dfrac{1}{2n+1}, & 3 \le n \le 13 \\[2mm] \dfrac{1}{2n}, & n \ge 14 \end{cases} \tag{3.108}$$

holds, then:

(i) $\widehat{w} < \beta_n w$.
(ii) $d < \dfrac{1}{1-2\lambda_n}\hat d$.
(iii) $|u_i| < \rho_n c_n d$.
(iv) $\widehat{w} < c_n \hat d$.
(v) $|\hat u_i| \le \dfrac{\gamma_n}{d^3}|u_i|^2\Big(\sum_{j\neq i}|u_j|\Big)^2$.

The proof of this lemma is strikingly similar to that of Lemmas 3.15 and 3.18 and will be omitted.

Now, we establish initial conditions of practical interest, which guarantee the convergence of the BSW method (3.107).

Theorem 3.11. *If the initial condition given by*

$$w^{(0)} < c_n d^{(0)} \tag{3.109}$$

is satisfied, where c_n *is given by* (3.108), *then the iterative method* (3.107) *is convergent with the order of convergence 4.*

Proof. The proof of this theorem is based on the assertions of Lemma 3.19 with the help of the previously presented technique. As in the already stated convergence theorems, the proof goes by induction. By the same argumentation as in the previous proofs, the initial condition (3.109) provides the validity of the inequality $w^{(m)} < c_n d^{(m)}$ for all $m \geq 0$, and hence, the inequalities (i)–(iv) of Lemma 3.19 also hold for all $m \geq 0$. In particular (according to Lemma 3.19(i)), we have

$$\frac{d^{(m)}}{d^{(m+1)}} < \frac{1}{1-2\lambda_n} \tag{3.110}$$

and, with regard to Lemma 3.19(iv),

$$|u_i^{(m+1)}| \leq \frac{\gamma_n}{(d^{(m)})^3}|u_i^{(m)}|^2\Bigg(\sum_{\substack{j=1\\ j\neq i}}^{n}|u_j^{(m)}|\Bigg)^2 \tag{3.111}$$

for each $i \in \boldsymbol{I}_n$ and all $m = 0, 1, \ldots$.

Substituting

$$t_i^{(m)} = \left[\frac{(n-1)^2\gamma_n}{(1-2\lambda_n)(d^{(m)})^3}\right]^{1/3}|u_i^{(m)}|$$

into (3.111) and using (3.110), we obtain

$$t_i^{(m+1)} < \frac{1}{(n-1)^2}\big(t_i^{(m)}\big)^2\Bigg(\sum_{\substack{j=1\\ j\neq i}}^{n} t_j^{(m)}\Bigg)^2. \tag{3.112}$$

By the assertion (ii) of Lemma 3.19 for the first iteration ($m = 0$), we have

$$t_i^{(0)} = \left[\frac{(n-1)^2\gamma_n}{(1-2\lambda_n)(d^{(0)})^3}\right]^{1/3}|u_i^{(0)}| < \rho_n c_n\left[\frac{(n-1)^2\gamma_n}{1-2\lambda_n}\right]^{1/3}. \tag{3.113}$$

Putting $t = \max_i t_i^{(0)}$, we find from (3.113) that $t_i^{(0)} \leq t < 0.988 < 1$ for $3 \leq n \leq 13$, and $t_i^{(0)} \leq t < 0.999 < 1$ for $n \geq 14$, for each $i = 1, \ldots, n$. According to this, we infer from (3.112) that the sequences $\{t_i^{(m)}\}$ (and, consequently, $\{|u_i^{(m)}|\}$) tend to 0 for all $i = 1, \ldots, n$. Hence, the BSW method (3.107) is convergent.

Putting $u^{(m)} = \max\limits_{1\leq i\leq n} |u_i^{(m)}|$, from (3.111), we get

$$u^{(m+1)} < \frac{\gamma_n}{(d^{(m)})^3}(n-1)^2\big(u^{(m)}\big)^4,$$

which means that the order of convergence of the BSW method is 4. □

The Halley-Like Method

Using a concept based on Bell's polynomials, X. Wang and Zheng [182] established a family of iterative methods of the order of convergence $k+2$, where k is the highest order of the derivative of P appearing in the generalized iterative formula, see Sect. 1.1. For $k=1$, this family gives the Ehrlich–Aberth's method (3.73), and for $k=2$ produces the following iterative method of the fourth order for the simultaneous approximation of all simple zeros of a polynomial P

$$z_i^{(m+1)} = z_i^{(m)} - \frac{1}{f(z_i^{(m)}) - \dfrac{P(z_i^{(m)})}{2P'(z_i^{(m)})}\Big(\Big[S_{1,i}^{(m)}\Big]^2 + S_{2,i}^{(m)}\Big)} \quad (i \in \boldsymbol{I}_n,\ m=0,1,\ldots), \tag{3.114}$$

where

$$f(z) = \frac{P'(z)}{P(z)} - \frac{P''(z)}{2P'(z)}, \quad S_{k,i}^{(m)} = \sum_{j\neq i} \frac{1}{\big(z_i^{(m)} - z_j^{(m)}\big)^k} \quad (k=1,2).$$

Since the function $f(z)$ appears in the well-known Halley's iterative method

$$\hat{z}_i = z_i - \frac{1}{\dfrac{P'(z_i)}{P(z_i)} - \dfrac{P''(z_i)}{2P'(z_i)}} = z_i - \frac{1}{f(z_i)},$$

we could say that the method (3.114) is of Halley's type. In the literature, the method (3.114) is sometimes called the Wang–Zheng's method.

The convergence analysis of the Halley-like method (3.114) is similar to that given previously in this section (see also the paper by M. Petković and Đ. Herceg [117]), so it will be presented in short.

Let us introduce the following abbreviations:

$$\rho_n = \frac{1}{1-nc_n}, \quad \eta_n = \frac{2(1-n\rho_n c_n)}{1-\rho_n c_n} - \frac{n(n-1)(\rho_n c_n)^3(2-\rho_n c_n)}{(1-\rho_n c_n)^2},$$

$$\lambda_n = \frac{2\rho_n c_n(1-\rho_n c_n + (n-1)\rho_n c_n)}{(1-\rho_n c_n)\eta_n}, \quad \gamma_n = \frac{n(2-\rho_n c_n)}{\eta_n(1-\rho_n c_n)^2}.$$

Lemma 3.20. *Let $\hat{z}_1,\ldots,\hat{z}_n$ be approximations generated by the iterative method* (3.114) *and let $\hat{u}_i = \hat{z}_i - \zeta_i$, $\hat{d} = \min_{i\neq j} |\hat{z}_i - \hat{z}_j|$, and $\widehat{w} = \max_{1\le i\le n} |\widehat{W}_i|$. If the inequality*

$$w < c_n d, \qquad c_n = \begin{cases} \dfrac{1}{3n+2.4}, & 3 \le n \le 20 \\[2ex] \dfrac{1}{3n}, & n \ge 21 \end{cases} \tag{3.115}$$

holds, then:

(i) $d < \dfrac{1}{1-2\lambda_n}\hat{d}$.
(ii) $|u_i| < \rho_n c_n d$.
(iii) $\widehat{w} < c_n \hat{d}$.
(iv) $|\hat{u}_i| \le \dfrac{\gamma_n}{d^3}|u_i|^3 \sum_{j\neq i} |u_j|$.

The proof of this lemma is similar to the proofs of Lemmas 3.15 and 3.18.

We now give the convergence theorem for the iterative method (3.114) under computationally verifiable initial conditions.

Theorem 3.12. *Let P be a polynomial of degree $n \ge 3$ with simple zeros. If the initial condition*

$$w^{(0)} < c_n d^{(0)} \tag{3.116}$$

holds, where c_n is given by (3.115)*, then the Halley-like method* (3.114) *is convergent with the fourth order of convergence.*

Proof. The proof of this theorem goes in a similar way to the previous cases using the assertions of Lemma 3.20. By virtue of the implication (iii) of Lemma 3.20 (i.e., $w < c_n d \;\Rightarrow\; \hat{w} < c_n \hat{d}$), we conclude by induction that the initial condition (3.116) implies the inequality $w^{(m)} < c_n d^{(m)}$ for each $m = 1, 2, \ldots$. For this reason, the assertions of Lemma 3.20 are valid for all $m \ge 0$. In particular (according to (i) and (iv) of Lemma 3.20), we have

$$\frac{d^{(m)}}{d^{(m+1)}} < \frac{1}{1-2\lambda_n} \tag{3.117}$$

and

$$|u_i^{(m+1)}| \le \frac{\gamma_n}{(d^{(m)})^3}|u_i^{(m)}|^3 \sum_{\substack{j=1\\ j\neq i}}^{n} |u_j^{(m)}| \quad (i \in \boldsymbol{I}_n) \tag{3.118}$$

for each iteration index $m = 0, 1, \ldots$.

Substituting

$$t_i^{(m)} = \left[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(m)})^3}\right]^{1/3} |u_i^{(m)}|$$

into (3.118) gives

$$t_i^{(m+1)} \le \frac{(1-2\lambda_n)d^{(m)}}{(n-1)d^{(m+1)}}\big(t_i^{(m)}\big)^3 \sum_{\substack{j=1\\ j\neq i}}^{n} t_j^{(m)} \quad (i \in \boldsymbol{I}_n).$$

Hence, using (3.117), we arrive at

$$t_i^{(m+1)} < \frac{1}{n-1}\big(t_i^{(m)}\big)^3 \sum_{\substack{j=1\\ j\neq i}}^{n} t_j^{(m)} \quad (i \in \boldsymbol{I}_n). \tag{3.119}$$

Since $|u_i^{(0)}| < \rho_n c_n d^{(0)}$ (assertion (ii) of Lemma 3.20), we may write

$$t_i^{(0)} = \left[\frac{(n-1)\gamma_n}{(1-2\lambda_n)(d^{(0)})^3}\right]^{1/3} |u_i^{(0)}| < \rho_n c_n \left[\frac{(n-1)\gamma_n}{1-2\lambda_n}\right]^{1/3}$$

for each $i = 1, \ldots, n$. Let $t_i^{(0)} \leq \max_i t_i^{(0)} = t$. Then

$$t < \rho_n c_n \left[\frac{(n-1)\gamma_n}{1-2\lambda_n}\right]^{1/3} < 0.310 \quad \text{for} \quad 3 \leq n \leq 20$$

and

$$t < 0.239 \qquad \text{for} \quad n \geq 21,$$

i.e., $t_i^{(0)} \leq t < 1$ for all $i = 1, \ldots, n$. Hence, we conclude from (3.119) that the sequences $\{t_i^{(m)}\}$ (and, consequently, $\{|u_i^{(m)}|\}$) tend to 0 for all $i = 1, \ldots, n$. Therefore, $z_i^{(m)} \to \zeta_i$ ($i \in \boldsymbol{I}_n$) and the method (3.114) is convergent.

Finally, from (3.118), there follows

$$|u_i^{(m+1)}| \leq u^{(m+1)} < (n-1)\frac{\gamma_n}{(d^{(m)})^3}\left(u^{(m)}\right)^4,$$

where $u^{(m)} = \max_{1 \leq i \leq n} |u_i^{(m)}|$. Therefore, the convergence order of the Halley-like method (3.114) is 4. □

Some Computational Aspects

In this section, we have improved the convergence conditions of four root finding methods. For the purpose of comparison, let us introduce the *normalized* i-factor $\Omega_n = n \cdot c_n$. The former Ω_n for the considered methods, found in the recent papers cited in Sect. 1.1, and the improved (new) Ω_n, proposed in this section, are given in Table 3.1.

Table 3.1 The entries of normalized i-factors

	Former Ω_n	New Ω_n
Ehrlich–Aberth's method (3.73)	$\dfrac{n}{2n+3}$	$\begin{cases} \dfrac{n}{2n+1.4} & (3 \leq n \leq 7), \\ 1/2 & (n \geq 8) \end{cases}$
EAN method (3.90)	$\dfrac{1}{3}$	$\begin{cases} \dfrac{n}{2.2n+1.9} & (3 \leq n \leq 21), \\ 1/2.2 & (n \geq 22) \end{cases}$
BSW method (3.107)	$\dfrac{n}{2n+2}$	$\begin{cases} \dfrac{n}{2n+1} & (3 \leq n \leq 13), \\ 1/2 & (n \geq 14) \end{cases}$
Halley-like method (3.114)	$\dfrac{1}{4}$	$\begin{cases} \dfrac{n}{3n+2.4} & (3 \leq n \leq 20), \\ 1/3 & (n \geq 21) \end{cases}$

To compare the former $\Omega_n = nc_n$ with the improved Ω_n, we introduce a percentage measure of the improvement

$$r\% = \frac{\Omega_n^{(\text{new})} - \Omega_n^{(\text{former})}}{\Omega_n^{(\text{former})}} \cdot 100.$$

Following Table 3.1, we calculated $r\%$ for $n \in [3, 30]$ and displayed $r\%$ in Fig. 3.3 as a function of n for each of the four considered methods. From Fig. 3.3, we observe that we have significantly improved i-factors c_n, especially for the EAN method (3.90) and Halley-like method (3.114).

The values of the i-factor c_n, given in the corresponding convergence theorems for the considered iterative methods, are mainly of theoretical importance. We were constrained to take smaller values of c_n to enable the validity

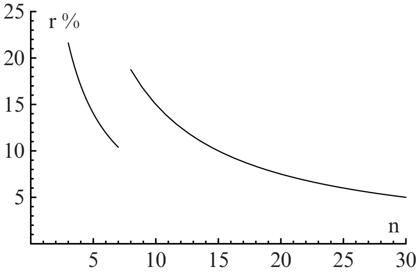

Fig. 3.3 (**a**) $r\%$: the method (3.73)

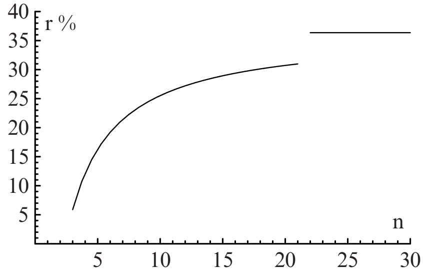

Fig. 3.3 (**b**) $r\%$: the method (3.90)

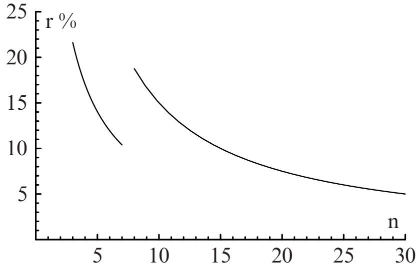

Fig. 3.3 (**c**) $r\%$: the method (3.107)

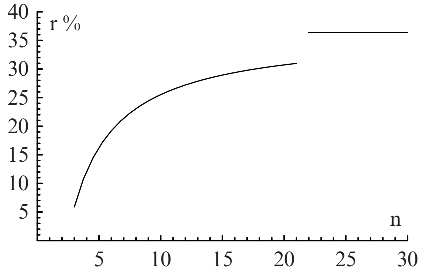

Fig. 3.3 (**d**) $r\%$: the method (3.114)

of inequalities appearing in the convergence analysis. However, these theoretical values of c_n can be suitably applied in ranking the considered methods regarding (1) their initial conditions for the guaranteed convergence and (2) convergence behavior in practice.

As mentioned in [118], in practical implementation of simultaneous root finding methods, we may take greater c_n related to that given in the convergence theorems and still preserve both guaranteed and fast convergence. The determination of the range of values of i-factor c_n providing favorable features (guaranteed and fast convergence) is a very difficult task, and practical experiments are the only means for obtaining some information on its range. We have tested the considered methods in examples of many algebraic polynomials with degree up to 20, taking initial approximations in such a way that the i-factor took the values kc_n for $k = 1$ (theoretical entry applied in the stated initial conditions) and for $k = 1.5$, 2, 3, 5, and 10. The stopping criterion was given by the inequality

$$\max_{1 \le i \le n} |z_i^{(m)} - \zeta_i| < 10^{-15}.$$

In Table 3.2, we give the average number of iterations (rounded to one decimal place), needed to satisfy this criterion.

From Table 3.2, we observe that the new i-factor not greater than $2c_n$ mainly preserves the convergence rate related to the theoretical value c_n given in the presented convergence theorems. The entry $3c_n$ is rather acceptable

Table 3.2 The average number of iterations as the i-factor increases

	c_n	$1.5c_n$	$2c_n$	$3c_n$	$5c_n$	$10c_n$
Ehrlich–Aberth's method (3.73)	3.9	4	4.2	5.4	7.3	13.3
EAN method (3.90)	3.1	3.2	3.4	5.1	6.1	10.2
BSW method (3.107)	3	3.1	3.3	4.3	5.8	9.8
Halley-like method (3.114)	3.2	3.4	4.2	5.5	6.7	10.7

from a practical point of view, while the choice of $5c_n$ doubles the number of iterations. Finally, the value $10c_n$ significantly decreases the convergence rate of all considered methods, although still provides the convergence.

3.4 A Posteriori Error Bound Methods

In this section, we combine good properties of iterative methods with fast convergence and a posteriori error bounds given in Corollary 1.1, based on Carstensen's results [13] on Gerschgorin's disks, to construct efficient inclusion methods for polynomial complex zeros. Simultaneous determination of both centers and radii leads to iterative error bound methods, which enjoy very convenient property of enclosing zeros at each iteration. This class of methods possesses a high computational efficiency since it requires less numerical operations compared with standard interval methods realized in interval arithmetic (see M. Petković and L. Petković [132]). Numerical experiments demonstrate equal or even better convergence behavior of these methods than the corresponding circular interval methods. In this section, the main attention is devoted to the construction of inclusion error bound methods. We will also give a review of some properties of these methods, including the convergence rate, efficient implementation, and initial conditions for the guaranteed convergence.

Corollary 1.1 given in Chap. 1 may be expressed in the following form.

Corollary 3.1. *Let P be an algebraic polynomial with simple (real or complex) zeros. Under the condition $w < c_n d$ ($c_n \le 1/(2n)$), each of disks D_i defined by*

$$D_i = \Big\{z_i; \frac{|W_i(z_i)|}{1 - nc_n}\Big\} = \big\{z_i; \rho_i\big\} \quad (i \in \boldsymbol{I}_n)$$

contains exactly one zero of P.

If the centers z_i of disks D_i are calculated by an iterative method, then we can generate sequences of disks $D_i^{(m)}$ $(m = 0, 1, \ldots)$ whose radii $\rho_i^{(m)} = W_i^{(m)}/(1 - nc_n)$ converge to 0 under some suitable conditions. It should be noted that only those methods which use quantities already calculated in the previous iterative step (in our case, the corrections W_i) enable a high computational efficiency. For this reason, we restrict our choice to the class of derivative-free methods which deal with Weierstrass' corrections, so-called *W-class*. The following most frequently used simultaneous methods from the W-class will be considered.

The Durand–Kerner's or Weierstrass' method [32], [72], shorter the W method, the convergence order 2:

$$z_i^{(m+1)} = z_i^{(m)} - W_i^{(m)} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.120}$$

The Börsch-Supan's method [10], shorter the BS method, the convergence order 3:

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \displaystyle\sum_{j \ne i} \frac{W_j^{(m)}}{z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.121}$$

The Börsch-Supan's method with Weierstrass' correction [95], shorter the BSW method, the convergence order 4:

$$z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \displaystyle\sum_{j \ne i} \frac{W_j^{(m)}}{z_i^{(m)} - W_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.122}$$

Let us note that $W_i^{(m)} = W(z_i^{(m)})$, see (1.17).

Let $z_1^{(0)}, \ldots, z_n^{(0)}$ be given initial approximations and let

$$z_i^{(m)} = \varPhi_W(z_i^{(m-1)}) \quad (i \in \boldsymbol{I}_n,\ m = 1, 2, \ldots) \tag{3.123}$$

be a derivative-free iterative method based on Weierstrass' corrections (belonging to the W-class), which is indicated by the subscript index "W." For example, the methods (3.120)–(3.122) belong to the W-class. Other iterative methods of Weierstrass' class are given in [34], [124], [131], [146], and [196].

Combining the results of Corollary 3.1 and (3.123), we can state the following inclusion method in a general form.

A posteriori error bound method. A posteriori error bound method (shorter PEB method) is defined by the sequences of disks $\{D_i^{(m)}\}$ $(i \in \boldsymbol{I}_n)$,

$$\begin{aligned} &D_i^{(0)} = \left\{ z_i^{(0)}; \frac{|W(z_i^{(0)})|}{1 - nc_n} \right\}, \\ &D_i^{(m)} = \{z_i^{(m)}; \rho_i^{(m)}\}, \qquad (i \in \boldsymbol{I}_n,\ m = 1, 2, \ldots), \\ &z_i^{(m)} = \Phi_W(z_i^{(m-1)}), \quad \rho_i^{(m)} = \frac{|W(z_i^{(m)})|}{1 - nc_n}, \end{aligned} \tag{3.124}$$

assuming that the initial condition $w^{(0)} < c_n d^{(0)}$ (with $c_n \le 1/(2n)$) holds.

The proposed method, defined by the sequences of disks given by (3.124), may be regarded as a quasi-interval method, which differs structurally from standard interval methods that deal with disks as arguments. For comparison, let us present the following circular interval methods which do not use the polynomial derivatives.

The Weierstrass-like interval method [183], the order 2:

$$Z_i^{(m+1)} = z_i^{(m)} - \frac{P(z_i^{(m)})}{\displaystyle\prod_{\substack{j=1 \\ j \ne i}}^{n} \left(z_i^{(m)} - Z_j^{(m)}\right)} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.125}$$

The Börsch-Supan-like interval method [107], the order 3:

$$Z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \displaystyle\sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j^{(m)}}{Z_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.126}$$

The Börsch-Supan-like interval method with Weierstrass' correction [111], the order 4 (with the centered inversion (1.63)):

$$Z_i^{(m+1)} = z_i^{(m)} - \frac{W_i^{(m)}}{1 + \displaystyle\sum_{\substack{j=1 \\ j \ne i}}^{n} \frac{W_j^{(m)}}{Z_i^{(m)} - W_i^{(m)} - z_j^{(m)}}} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots). \tag{3.127}$$

All of the methods (3.124)–(3.127) possess the crucial inclusion property: each of the produced disks contains exactly one zero in each iteration. In this manner, not only very close zero approximations (given by the centers of disks) but also the upper error bounds for the zeros (given by the radii of disks) are obtained. More about interval methods for solving polynomial equations can be found in the books by M. Petković [109] and M. Petković and L. Petković [129].

Studying the convergence of error bounds produced by (3.124), the following important tasks arise:

1. Determine the convergence order of a posteriori error bound method when the centers $z_i^{(m)}$ of disks

$$D_i^{(m)} = \left\{ z_i^{(m)}; \frac{|W(z_i^{(m)})|}{1 - nc_n} \right\} \quad (i \in \boldsymbol{I}_n,\ m = 0, 1, \ldots) \tag{3.128}$$

   are calculated by an iterative method of order k (≥ 2).
2. State computationally verifiable initial condition that guarantees the convergence of the sequences of radii $\{\operatorname{rad} D_i^{(m)}\}$. We note that this problem, very important in the theory and practice of iterative processes in general, is a part of Smale's point estimation theory [165] which has attracted a considerable attention during the last two decades (see [118] and Chap. 2 for details). As mentioned in the previous sections, initial conditions in the case of algebraic polynomials should depend only on attainable data – initial approximations, polynomial degree, and polynomial coefficients.
3. Compare the computational efficiencies of the PEB methods and the existing circular interval methods (given, for instance, by (3.125)–(3.127)). Which of these two classes of methods is more efficient?
4. Using numerical experiments, compare the size of inclusion disks produced by the PEB methods and the corresponding interval methods (3.125)–(3.127). Whether the construction of PEB methods is justified?

The study of these subjects was the main goal of the paper [122]. Here, we give the final results and conclusions of this investigation in short.

Assume that the following inequality

$$w^{(0)} < c_n d^{(0)} \tag{3.129}$$

holds, where c_n is given by

$$c_n = \begin{cases} \dfrac{1}{2n}, & \text{the W method [2] and BS method [42],} \\[2ex] \dfrac{1}{2n+1}, & \text{the BSW method [52].} \end{cases} \tag{3.130}$$

Then, the following three methods from the W-class are convergent: the Durand–Kerner's method (3.120) (for the proof, see Batra [5]), the Börsch-Supan's method (3.121) (M. Petković and Đ. Herceg [117]), and the Börsch-Supan's method with Weierstrass' correction (3.122) (see [60], [140]). The corresponding inequalities of the form

$$|W_i^{(m+1)}| < \delta_n |W_i^{(m)}| \quad (\delta_n < 1)$$

are the composite parts of Lemmas 3.3(i), 3.6(i), and 3.19(i) under the condition (3.129) for specific entries c_n given by (3.130). This means that the sequences $\{|W_i^{(m)}|\}$ $(i \in \boldsymbol{I}_n)$ are convergent and tend to 0. Hence, the sequences of radii $\{\rho^{(m)}|\}$ $(i \in \boldsymbol{I}_n)$ are also convergent and tend to 0 under the condition (3.129). The convergence rate of the PEB methods based on the iterative methods (3.120)–(3.122) was studied in [122], where the following assertions were proved.

Theorem 3.13. *The PEB method* (3.124)*, based on the Durand–Kerner's method* (3.120)*, converges quadratically if the initial condition* (3.129) *holds, where* $c_n = 1/(2n)$.

Theorem 3.14. *The PEB method* (3.124)*, based on the Börsch-Supan's method* (3.121)*, converges cubically if the initial condition* (3.129) *holds, where* $c_n = 1/(2n)$.

Theorem 3.15. *The PEB method* (3.124)*, based on the Börsch-Supan's method with Weierstrass' corrections* (3.122)*, converges with the order 4 if the initial condition* (3.129) *holds, where* $c_n = 1/(2n+1)$.

We emphasize that the initial condition (3.129) (with c_n given by (3.130)) that guarantees the convergence of the PEB methods (3.124)–(3.120), (3.124)–(3.121), and (3.124)–(3.122) depends only on attainable data, which is of great practical importance.

Computational Aspects

In the continuation of this section, we give some practical aspects in the implementation of the proposed methods. As mentioned above, the computational cost significantly decreases if the quantities $W_i^{(0)}, W_i^{(1)}, \ldots$ $(i \in \boldsymbol{I}_n)$, necessary in the calculation of the radii $\rho_i^{(m)} = |W_i^{(m)}|/(1 - nc_n)$, are applied in the calculation of the centers $z_i^{(m+1)}$ defined by the employed iterative formula from the W-class. Regarding the iterative formulae (3.120)–(3.122), we observe that this requirement is satisfied. A general calculating procedure can be described by the following algorithm.

Calculating Procedure (I)

Given $z_1^{(0)}, \dots, z_n^{(0)}$ *and the tolerance parameter* τ.
Set $m = 0$.

1° *Calculate Weierstrass' corrections* $W_1^{(m)}, \dots, W_n^{(m)}$ *at the points* $z_1^{(m)}, \dots, z_n^{(m)}$.

2° *Calculate the radii* $\rho_i^{(m)} = |W_i^{(m)}|/(1 - nc_n) \quad (i = 1, \dots, n)$.

3° *If* $\max_{1 \le i \le n} \rho_i^{(m)} < \tau$, *then STOP*
otherwise, GO TO 4°.

4° *Calculate the new approximations* $z_1^{(m+1)}, \dots, z_n^{(m+1)}$ *by a suitable iterative formula from the W-class (for instance, by* (3.120), (3.121), *or* (3.122)*).*

5° *Set* $m := m + 1$ *and GO TO the step* 1°.

Following the procedure (I), we have realized many numerical examples and, for demonstration, we select the following one.

Example 3.1. We considered the polynomial

$$\begin{aligned} P(z) = {} & z^{13} - (5+5\mathrm{i})z^{12} + (5+25\mathrm{i})z^{11} + (15-55\mathrm{i})z^{10} - (66-75\mathrm{i})z^9 \\ & + 90z^8 - z^5 + (5+5\mathrm{i})z^4 - (5+25\mathrm{i})z^3 - (15-55\mathrm{i})z^2 \\ & + (66-75\mathrm{i})z - 90 \\ = {} & (z-3)(z^8-1)(z^2-2z+5)(z-2\mathrm{i})(z-3\mathrm{i}). \end{aligned}$$

Starting from sufficiently close initial approximations $z_1^{(0)}, \dots, z_{13}^{(0)}$, we first calculated the radii $\rho_i^{(0)} = |W(z_i^{(0)}|/(1 - nc_n)$ of initial disks $D_1^{(0)}, \dots, D_{13}^{(0)}$. These disks were applied in the implementation of a posteriori error bound methods (3.124) as well as interval methods (3.125)–(3.127). We obtained $\max \rho_i^{(0)} = 0.3961$ for the methods (3.125), (I-W), (3.126), (I-BS) and $\max \rho_i^{(0)} = 0.3819$ for (3.127) and (I-BSW). The approximations $z_i^{(m)}$ $(m \ge 1)$ were calculated by the iterative formulae (3.120)–(3.122) and the corresponding inclusion methods are referred to as (I-W), (I-BS), and (I-BSW), respectively. The largest radii of the disks obtained in the first four iterations may be found in Table 3.3, where $A(-q)$ means $A \times 10^{-q}$.

Table 3.3 Resulting disks obtained by Procedure I

Methods	$\max \rho_i^{(1)}$	$\max \rho_i^{(2)}$	$\max \rho_i^{(3)}$	$\max \rho_i^{(4)}$
(I-W) (3.124)–(3.120)	1.26(−1)	1.74(−2)	1.33(−4)	1.59(−8)
Interval W (3.125)	1.05	No inclusions	–	–
(I-BS) (3.124)–(3.121)	1.83(−2)	3.61(−6)	1.32(−17)	7.06(−52)
Interval BS (3.126)	1.99(−1)	2.41(−4)	2.39(−15)	2.38(−49)
(I-BSW) (3.124)–(3.122)	6.94(−3)	4.92(−10)	1.83(−38)	3.20(−152)
Interval BSW (3.127)	2.98(−1)	1.47(−5)	1.81(−24)	2.68(−100)

In our calculation, we employed multiprecision arithmetic in *Mathematica* 6.0 since the tested methods converge extremely fast producing very small disks. From Table 3.3, we observe that the PEB methods are equal or better than the corresponding methods (of the same order) (3.125)–(3.127) realized in complex interval arithmetic. A number of numerical experiments showed similar convergence behavior of the tested methods.

The Weierstrass' interval method (3.125) exhibits rather poor results. The explanation lies in the fact that this method uses the product of disks which is not an exact operation in circular arithmetic and produces oversized disks (see Sect. 1.3).

Calculation Procedure (I) assumes the knowledge of initial approximations $z_1^{(0)}, \ldots, z_n^{(0)}$ in advance. The determination of these approximations is usually realized by a slowly convergent multistage composite algorithm. Sometimes, the following simple approach gives good results in practice.

Calculating Procedure (II)

1° *Find the disk centered at the origin with the radius*

$$R = 2 \max_{1 \le k \le n} \left| a_{n-k} \right|^{1/k} \quad \text{(see (1.58) or (5.72))},$$

which contains all zeros of the polynomial $P(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$.

2° *Calculate Aberth's initial approximations* [1]

$$z_\nu^{(0)} = -\frac{a_{n-1}}{n} + r_0 \exp(\mathrm{i}\theta_\nu), \quad \mathrm{i} = \sqrt{-1}, \quad \theta_\nu = \frac{\pi}{n}\Big(2\nu - \frac{3}{2}\Big) \quad (\nu = 1, \ldots, n),$$

equidistantly distributed along the circle $|z + a_{n-1}/n| = r_0,\ r_0 \le R$ *(see Sect.* 4.4*).*

3° *Apply the simultaneous method* (3.120) *or* (3.121) *starting with Aberth's approximations; stop the iterative process when the condition*

$$\max_{1 \le i \le n} |W(z_i^{(m)})| < c_n \min_{i \ne j} |z_i^{(m)} - z_j^{(m)}| \tag{3.131}$$

is satisfied.

4°–8° *The same as the steps* 1°–5° *of Procedure I.*

We applied Procedure II on the following example.

Example 3.2. To find approximations to the zeros of the polynomial

$$z^{15} + z^{14} + 1 = 0$$

satisfying the condition (3.131) (with $c_n = 1/(2n)$), we applied the Börsch-Supan's method (3.121) with Aberth's initial approximations taking $a_{n-1} = 1$, $n = 15$, and $r_0 = 2$. The condition (3.131) was satisfied after seven iterative steps. The obtained approximations were used to start the PEB methods (I-W), (I-BS), and (I-BSW). After three iterations, we obtained disks whose largest radii are given in Table 3.4.

Table 3.4 Resulting disks obtained by (I-W), (I-BS), and (I-BSW): Procedure II

Methods	$\max \rho_i^{(0)}$	$\max \rho_i^{(1)}$	$\max \rho_i^{(2)}$
(I-W) (3.124)–(3.120)	1.51(−3)	3.79(−6)	2.27(−11)
(I-BS) (3.124)–(3.121)	1.51(−3)	4.10(−9)	8.31(−26)
(I-BSW) (3.124)–(3.122)	1.46(−3)	9.64(−12)	1.60(−44)

From Tables 3.3 and 3.4, we observe that the results obtained by the methods (I-W), (I-BS), and (I-BSW) coincide with the theoretical results given in Corollary 3.1 and Theorems 3.13–3.15; in other words, the order of convergence in practice matches very well the order expressed in Theorems 3.13–3.15.

At the beginning of the section, we mentioned that the PEB methods require less numerical operations compared with their counterparts in complex interval arithmetic. In Table 3.5, we give the total number of numerical operations per one iteration, reduced to real arithmetic operations. We have used the following abbreviations:

– $AS(n)$ (total number of additions and subtractions)
– $M(n)$ (multiplications)
– $D(n)$ (divisions)

Table 3.5 The number of basic operations

	$AS(n)$	$M(n)$	$D(n)$
(I-W) (3.124)–(3.120)	$8n^2 + n$	$8n^2 + 2n$	$2n$
Interval W (3.125)	$22n^2 - 6n$	$25n^2 - 6n$	$8n^2 - n$
(I-BS) (3.124)–(3.121)	$15n^2 - 6n$	$14n^2 + 2n$	$2n^2 + 2n$
Interval BS (3.126)	$23n^2 - 4n$	$23n^2 + 2n$	$7n^2 + 2n$
(I-BSW) (3.124)–(3.122)	$15n^2 - 4n$	$14n^2 + 2n$	$2n^2 + 2n$
Interval BSW (3.127)	$23n^2 - 2n$	$23n^2 + 2n$	$7n^2 + 2n$

From Table 3.5, we observe that the PEB methods require significantly less numerical operations with respect to the corresponding interval methods. One of the reasons for this advantage is the use of the already calculated Weierstrass' corrections W_i in the evaluation of the radii ρ_i.

Parallel Implementation

It is worth noting that the error bound method (3.124) for the simultaneous determination of all zeros of a polynomial is very suitable for the implementation on parallel computers since it runs in several identical versions. In this manner, a great deal of computation can be executed simultaneously. An analysis of total running time of a parallel iteration and the determination of the optimal number of processors points to some undoubted advantages of the implementation of simultaneous methods on parallel processing computers, see, e.g., [22]–[24], [44], [115]. The parallel processing becomes of great interest to speed up the determination of zeros when one should treat polynomials with degree 100 and higher, appearing in mathematical models in scientific engineering, including digital signal processing or automatic control [66], [92].

The model of parallel implementation is as follows: It is assumed that the number of processors k $(\leq n)$ is given in advance. Let

$$\begin{aligned}\boldsymbol{W}^{(m)} &= \big(W_1^{(m)},\dots,W_n^{(m)}\big),\\ \boldsymbol{\rho}^{(m)} &= \big(\rho_1^{(m)},\dots,\rho_n^{(m)}\big),\\ \boldsymbol{z}^{(m)} &= \big(z_1^{(m)},\dots,z_n^{(m)}\big)\end{aligned}$$

denote vectors at the mth iterative step, where $\rho_i^{(m)} = |W(z_i^{(m)})|/(1-nc_n)$, and $z_i^{(m)}$ is obtained by the iterative formula $z_i^{(m)} = \Phi_W(z_i^{(m-1)})$ $(i \in \boldsymbol{I}_n)$. The starting vector $\boldsymbol{z}^{(0)}$ is computed by all processors $C_1,\dots,C_k$ using some suitable globally convergent method based on a subdivided procedure and the inclusion annulus $\{z : r \leq |z| \leq R\}$ which contains all zeros, given later by (4.72).

In the next stage, each step consists of sharing the calculation of $W_i^{(m)}$, $\rho_i^{(m)}$, and $z_i^{(m+1)}$ among the processors and in updating their data through a broadcast procedure (shorter $BCAST(\boldsymbol{W}^{(m)},\ \boldsymbol{\rho}^{(m)})$, $BCAST(\boldsymbol{z}^{(m+1)})$). As in [23], let $I_1,\dots,I_k$ be disjunctive partitions of the set $\{1,\dots,n\}$, where $\cup I_j = \{1,\dots,n\}$. To provide good load balancing between the processors, the index sets $I_1,\dots,I_k$ are chosen so that the number of their components $\mathbb{N}(I_j)$ $(j=1,\dots,k)$ is determined as $\mathbb{N}(I_j) \leq \left[\frac{n}{k}\right]$. At the mth iterative step, the processor C_j $(j = 1,...,k)$ computes $W_i^{(m)}, \rho_i^{(m)}$, and, if necessary, $z_i^{(m+1)}$ for all $i \in I_j$ and then it transmits these values to all other processors using a broadcast procedure. The program terminates when the stopping criterion is satisfied, say, if for a given tolerance τ the inequality

$$\max_{1\leq i\leq n}\big|\rho_i^{(m)}\big| < \tau$$

holds. A program written in pseudocode for a parallel implementation of the error bound method (3.124) is given below.

Program A POSTERIORI ERROR BOUND METHOD
begin
    for all $j = 1,\dots,k$ **do** determination of the approximations $\boldsymbol{z}^{(0)}$;
    $m := 0$
    C :=false

```
    do
        for all j = 1,...,k do in parallel
        begin
            Compute W_i^(m), i ∈ I_j;
            Compute ρ_i^(m), i ∈ I_j;
            Communication: BCAST(W^(m), ρ^(m));
        end
        if max_{1≤i≤n} ρ_i^(m) < τ; C :=true
        else
            m := m + 1
            for all j = 1,...,k do in parallel
            begin
                Compute z_i^(m), i ∈ I_j, by (3.123);
                Communication: BCAST(z^(m));
            end
        endif
    until C
    OUTPUT z^(m), ρ^(m)
end
```