

Generalization and Acceleration of an Algorithm of Sebastião e Silva and its Duals

Soon Park Chung

Received November 18, 1974

Summary. In this paper, we extend Householder's [4] generalization of an algorithm of Sebastião e Silva [11] by adding a new elimination rule for defining the sequences which converge to the factors of the given polynomial. We then present the dual algorithm and show that the dual algorithm becomes equivalent to the direct algorithm in the generalized form. Next, we give accelerated forms of these algorithms which are quadratically convergent. We also study the relation of these methods to other methods.

The algorithm of Sebastião e Silva [11] for obtaining zeros of polynomials has been further elaborated by Bauer [1, 2] and generalized by Householder [4]. The purpose of this paper is to investigate generalizations of the algorithm in which different elimination rules are used to get approximations to the factors of the given polynomial and in which the operating polynomial is varied to accelerate the convergence, and to show the relation of these methods to other algorithms. At the outset a theorem will be stated that extends Householder's generalization, in which a new elimination rule appears. In fact, this elimination rule has been used in the dual form of the original algorithm (see [7]), which has been extended to adapt to transcendental functions [13]. Next, we will present the dual algorithm in the generalized form and show that it is equivalent to the direct algorithm. Finally, we will study the accelerated forms of these algorithms which are quadratically convergent, and show how these algorithms can be related to other methods.

The polynomial to be considered will be taken as monic, for convenience:

$$\begin{aligned} f(z) &= z^n + a_{n-1}z^{n-1} + \dots + a_0, \quad a_0 \neq 0 \\ &= (z-r_1)(z-r_2) \dots (z-r_n). \end{aligned} \tag{1}$$

Let

$$f_{i_j \dots k}(z) = f(z) / [(z-r_i)(z-r_j) \dots (z-r_k)].$$

Also, for any polynomial

$$p(z) = p_m z^m + p_{m-1} z^{m-1} + \dots + p_0, \quad p_m \neq 0,$$

it will be understood that

$$p^*(z) = p_m^{-1} p(z).$$

Theorem 1. Let $f(z)$ be given by (1) and let $g(z)$ and $g_0(z)$ be any polynomials of degree $n-1$ at most such that neither $g(z)$, $g'(z)$ nor $g_0(z)$ vanishes for any r_j ,

and $g(r_j) \neq g(r_k)$ for $r_j \neq r_k$. Define recursively

$$g_{v+1}(z) = g(z)g_v(z) - \alpha_v(z)f(z), \quad v=0, 1, 2, \dots, \tag{2}$$

where each $g_v(z)$ is of degree $n-1$ at most. If

$$|g(r_1)| > |g(r_2)| \geq \dots \geq |g(r_n)|,$$

then

$$\lim_{v \rightarrow \infty} g_v^*(z) = f_1(z).$$

Moreover, let

$$g_{v,1}(z) = g_v(z), \quad \text{for each } v,$$

and define the sequences $g_{v,p}(z)$, $p=1, 2, \dots, n$, by any one of the following rules:

Rule 1. Form $g_{v,p+1}(z)$ by eliminating the highest term between $g_{v,p}(z)$ and $g_{v+1,p}(z)$.

Rule 2. Form $g_{v,p+1}(z)$ by eliminating the constant term between $g_{v,p}(z)$ and $g_{v+1,p}(z)$ and dividing by z .

Rule 3. Define $g_{v,p}(z)$ as the following determinant of order p :

$$g_{v,p}(z) = \delta \begin{pmatrix} g_v(z) & g_{v+1}(z) & \dots & g_{v+p-1}(z) \\ g_{v+1}(z) & g_{v+2}(z) & \dots & g_{v+p}(z) \\ \dots & \dots & \dots & \dots \\ g_{v+p-1}(z) & g_{v+p}(z) & \dots & g_{v+2p-2}(z) \end{pmatrix} \cdot [-f(z)]^{-(p-1)}. \tag{3}$$

Then

$$\lim_{v \rightarrow \infty} g_{v,p}^*(z) = f_{12\dots p}(z),$$

if

$$|g(r_1)| \geq \dots \geq |g(r_p)| > |g(r_{p+1})| \geq \dots \geq |g(r_n)|.$$

Proof. Rule 1 und Rule 3 have been proved in [4]. We will verify Rule 2. For simplicity, it will be assumed in the proof that $g_0(z) = 1$, since a different choice would leave the argument unaffected. Let

$$\begin{aligned} \gamma_j &= g(r_j), & G_v(z) &= g^v(z); \\ g_v(z) &= b_{v,0} + b_{v,1}z + \dots + b_{v,n-1}z^{n-1}. \end{aligned}$$

Since $g_v(z)$ is the polynomial of degree $n-1$ at most for which $g_v(r_j) = g^v(r_j) = \gamma_j^v$, $j=1, 2, \dots, n$, we can write, by Newton's formula for interpolation,

$$\begin{aligned} g_v(z) &= G_v(r_n) + (z-r_n) G_v(r_n, r_{n-1}) + \dots + f_{12}(z) G_v(r_n, r_{n-1}, \dots, r_2) \\ &\quad + f_{12}(z) G_v(r_n, r_{n-1}, \dots, r_2, r_1), \end{aligned} \tag{4}$$

where, if the r 's are all distinct,

$$G_v(r_i, r_j, \dots, r_k) = \delta \begin{pmatrix} 1 & r_i & r_i^2 & \dots & G_v(r_i) \\ 1 & r_j & r_j^2 & \dots & G_v(r_j) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & r_k & r_k^2 & \dots & G_v(r_k) \end{pmatrix} \zeta^{1/2}(r_i, r_j, \dots, r_k), \tag{5}$$

with $\zeta^{1/2}(r_i, r_j, \dots, r_k)$ the Vandermonde determinant, and in the confluent case when certain r 's become equal, certain rows in the numerator and denominator

in (5) are replaced by the derivatives. Now let

$$|g(r_1)| \geq |g(r_2)| \geq \dots \geq |g(r_p)| > |g(r_{p+1})| \geq \dots \geq |g(r_n)|,$$

for $p > 1$. It can be seen [5] that Rule 2 is equivalent to defining $g_{v,p}(z)$ by

$$g_{v,p}(z) = \delta \begin{pmatrix} b_{v,0} & b_{v,1} & \dots & b_{v,p-2} & g_v(z) \\ b_{v+1,0} & b_{v+1,1} & \dots & b_{v+1,p-2} & g_{v+1}(z) \\ \dots & \dots & \dots & \dots & \dots \\ b_{v+p-1,0} & b_{v+p-1,1} & \dots & b_{v+p-1,p-2} & g_{v+p-1}(z) \end{pmatrix} \cdot z^{-(p-1)}. \tag{6}$$

Let

$$\begin{aligned} \phi_v(z) &= f_{12\dots p}(z)G_v(r_n, r_{n-1}, \dots, r_p) + \dots + f_{12}(z)G_v(r_n, r_{n-1}, \dots, r_2) \\ &\quad + f_1(z)G_v(r_n, r_{n-1}, \dots, r_1) \\ &= \phi_{v,0} + \phi_{v,1}z + \dots + \phi_{v,n-1}z^{n-1}. \end{aligned}$$

Since $\phi_v(z)$ dominates all other terms of $g_v(z)$, we know that

$$\delta \begin{pmatrix} \phi_{v,0} & \phi_{v,1} & \dots & \phi_{v,p-2} & \phi_v(z) \\ \phi_{v+1,0} & \phi_{v+1,1} & \dots & \phi_{v+1,p-2} & \phi_{v+1}(z) \\ \dots & \dots & \dots & \dots & \dots \\ \phi_{v+p-1,0} & \phi_{v+p-1,1} & \dots & \phi_{v+p-1,p-2} & \phi_{v+p-1}(z) \end{pmatrix} \cdot z^{-(p-1)}$$

dominates all other terms of $g_{v,p}(z)$ provided only it contains a term in $\gamma_1^v \gamma_2^v \dots \gamma_p^v$. Now $\phi_v(z)$ can be written in the form

$$\phi_v(z) = f_{12\dots p}(z) \cdot \psi_v(z),$$

where

$$\begin{aligned} \psi_v(z) &= G_v(r_n, \dots, r_p) + (z-r_p)G_v(r_n, \dots, r_{p-1}) + \dots \\ &\quad + (z-r_p)(z-r_{p-1}) \dots (z-r_2)G_v(r_n, \dots, r_1). \end{aligned} \tag{7}$$

Let

$$\psi_v(z) = \psi_{v,0} + \psi_{v,1}z + \dots + \psi_{v,p-1}z^{p-1} \tag{8}$$

and

$$f_{12\dots p}(z) = c_0 + c_1z + \dots + c_{n-p}z^{n-p}.$$

We know that $c_0 \neq 0$ since $a_0 \neq 0$. We then have

$$\begin{aligned} &\delta \begin{pmatrix} \phi_{v,0} & \phi_{v,1} & \dots & \phi_{v,p-2} & \phi_v(z) \\ \dots & \dots & \dots & \dots & \dots \\ \phi_{v+p-1,0} & \phi_{v+p-1,1} & \dots & \phi_{v+p-1,p-2} & \phi_{v+p-1}(z) \end{pmatrix} \cdot z^{-(p-1)} \\ &= \delta \begin{pmatrix} \psi_{v,0}c_0 & \psi_{v,1}c_0 + \psi_{v,0}c_1 & \dots & \psi_{v,p-2}c_0 + \dots + \psi_{v,0}c_{p-2} \\ \dots & \dots & \dots & \dots \\ \psi_{v+p-1,0}c_0 & \psi_{v+p-1,1}c_0 + \psi_{v+p-1,0}c_1 & \dots & \psi_{v+p-1,p-2}c_0 + \dots + \psi_{v+p-1,0}c_{p-2} \\ & & & f_{12\dots p}(z)\psi_v(z) \\ & & & \dots \\ & & & f_{12\dots p}(z)\psi_{v+p-1}(z) \end{pmatrix} \cdot z^{-(p-1)} \end{aligned}$$

$$\begin{aligned}
 &= f_{12\dots p}(z) \delta \begin{pmatrix} \psi_{v,0}c_0 & \psi_{v,1}c_0 & \dots & \psi_{v,p-2}c_0 & \psi_v(z) \\ \dots & \dots & \dots & \dots & \dots \\ \psi_{v+p-1,0}c_0 & \psi_{v+p-1,1}c_0 & \dots & \psi_{v+p-1,p-2}c_0 & \psi_{v+p-1}(z) \end{pmatrix} \cdot z^{-(p-1)} \\
 &= f_{12\dots p}(z) c_0^{p-1} \delta \begin{pmatrix} \psi_{v,0} & \psi_{v,1} & \dots & \psi_{v,p-2} & \psi_{v,p-1}z^{p-1} \\ \dots & \dots & \dots & \dots & \dots \\ \psi_{v+p-1,0} & \psi_{v+p-1,1} & \dots & \psi_{v+p-1,p-2} & \psi_{v+p-1,p-1}z^{p-1} \end{pmatrix} \cdot z^{-(p-1)} \\
 &= f_{12\dots p}(z) c_0^{p-1} \delta \begin{pmatrix} \psi_{v,0} & \psi_{v,1} & \dots & \psi_{v,p-1} \\ \dots & \dots & \dots & \dots \\ \psi_{v+p-1,0} & \psi_{v+p-1,1} & \dots & \psi_{v+p-1,p-1} \end{pmatrix}. \tag{9}
 \end{aligned}$$

From (7) and (8), we have

$$\psi_{r,j} = \sum_{i=1}^{p-j} d_{i,j} G_v(r_n, \dots, r_i), \quad j=0, 1, \dots, p-1,$$

for some constants $d_{i,j}$. In particular,

$$d_{p-j,j} = 1, \quad j=0, 1, \dots, p-1.$$

So,

$$\begin{aligned}
 &\delta \begin{pmatrix} \psi_{v,0} & \psi_{v,1} & \dots & \psi_{v,p-1} \\ \dots & \dots & \dots & \dots \\ \psi_{v+p-1,0} & \psi_{v+p-1,1} & \dots & \psi_{v+p-1,p-1} \end{pmatrix} \\
 &= \delta \begin{pmatrix} \sum_{i=1}^p d_{i,0} G_v(r_n, \dots, r_i) & \sum_{i=1}^{p-1} d_{i,1} G_v(r_n, \dots, r_i) & \dots & d_{1,p-1} G_v(r_n, \dots, r_1) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^p d_{i,0} G_{v+p-1}(r_n, \dots, r_i) & \sum_{i=1}^{p-1} d_{i,1} G_{v+p-1}(r_n, \dots, r_i) & \dots & d_{1,p-1} G_{v+p-1}(r_n, \dots, r_1) \end{pmatrix} \\
 &= \delta \begin{pmatrix} d_{p,0} G_v(r_n, \dots, r_p) & d_{p-1,1} G_v(r_n, \dots, r_{p-1}) & \dots & d_{1,p-1} G_v(r_n, \dots, r_1) \\ \dots & \dots & \dots & \dots \\ d_{p,0} G_{v+p-1}(r_n, \dots, r_p) & d_{p-1,1} G_{v+p-1}(r_n, \dots, r_{p-1}) & \dots & d_{1,p-1} G_{v+p-1}(r_n, \dots, r_1) \end{pmatrix} \\
 &= \delta \begin{pmatrix} G_v(r_n, \dots, r_p) & G_v(r_n, \dots, r_{p-1}) & \dots & G_v(r_n, \dots, r_1) \\ \dots & \dots & \dots & \dots \\ G_{v+p-1}(r_n, \dots, r_p) & G_{v+p-1}(r_n, \dots, r_{p-1}) & \dots & G_{v+p-1}(r_n, \dots, r_1) \end{pmatrix}
 \end{aligned}$$

which has been shown [4] to contain a term in $\gamma_1^v \gamma_2^v \dots \gamma_p^v$. This, together with (9), shows that the term in $f_{12\dots p}(z)$ dominates all other terms of $g_{v,p}(z)$ and hence that

$$\lim_{v \rightarrow \infty} g_{v,p}^*(z) = f_{12\dots p}(z).$$

If $g(z) = z$, Theorem 1 is the original algorithm of Sebastião e Silva. As noted in Householder [4], the original algorithm fails when all the zeros of the given polynomial are equal in modulus while this generalization produces at least quadratic factors even for this case. The determinants (3), when $g(z) = z$, can be shown to be equivalent to those given by Bauer, who developed the following

algebraic relations among the polynomials $g_{v,p}(z)$ defined by his determinants:

$$q_{v,p} g_{v+1,p}^*(z) = z g_{v,p}^*(z) - g_{v+1,p-1}^*(z), \tag{10}$$

$$e_{v,p} g_{v,p+1}^*(z) = g_{v+1,p}^*(z) - g_{v,p}^*(z), \tag{11}$$

where $e_{v,p}$ and $q_{v,p}$ are constants such that

$$q_{v+1,p} + e_{v+1,p-1} = q_{v,p} + e_{v,p} \tag{12}$$

$$e_{v+1,p} q_{v+1,p} = e_{v,p} q_{v,p+1}. \tag{13}$$

In fact, (10) is the basis of Bauer's treppeniteration and (11) justifies Rule 1 of Theorem 1 for the case $g(z) = z$. Also the algebraic relations (10) through (13) suggest that Rutishauser's qd algorithm is closely related (in this connection, see Stewart [13]).

The dual of Theorem 1 will be stated in the following corollary.

Corollary 1. *Let $f(z)$ be given by (1) and let $g(z)$ and $g_0(z)$ satisfy all the conditions described in Theorem 1 and let k be the degree of $g(z)$. Define recursively*

$$g_{v+1}(z) = [g(z)]^{-1} [g_v(z) - \phi_v(z)f(z)], \quad v=0, 1, 2, \dots, \tag{14}$$

where each $\phi_v(z)$ is of degree $k-1$ at most such that $g_v(z) - \phi_v(z)f(z)$ is divisible by $g(z)$.

If

$$|g(r_1)| < |g(r_2)| \leq \dots \leq |g(r_n)|$$

then

$$\lim_{v \rightarrow \infty} g_v^*(z) = f_1(z).$$

Moreover, let

$$g_{v,1}(z) = g_v(z),$$

and define the sequences $g_{v,p}(z)$ by any one of the three rules described in Theorem 1. Then

$$\lim_{v \rightarrow \infty} g_{v,p}^*(z) = f_{12\dots p},$$

if

$$|g(r_1)| \leq \dots \leq |g(r_p)| < |g(r_{p+1})| \leq \dots \leq |g(r_n)|.$$

Proof. Since $g(z)$ and $f(z)$ do not have common zeros, there exist polynomials $h(z)$ and $k(z)$ of degrees $n-1$ and $k-1$, respectively, at most, such that

$$h(z)g(z) + k(z)f(z) = 1.$$

Hence

$$h(r_j) = 1/g(r_j), \quad j=1, 2, \dots, n$$

so that the direct algorithm with $h(z)$ is equivalent to the dual algorithm with $g(z)$.

This algorithm, with $g(z) = z$ and with elimination Rule 2, has been shown to be effective for finding zeros of transcendental functions in a circle of analyticity as well as of polynomials [13]. A more general theorem for transcendental functions can be obtained by extending Corollary 1.

In computing the polynomials $g_v(z)$ in Corollary 1, we have two methods.

First Method: We can find the polynomials $\phi(z)$ and $\psi(z)$ of degrees $n-1$ and $k-1$, respectively, at most, such that

$$\phi(z)g(z) - \psi(z)f(z) = 1.$$

Hence

$$[g_\nu(z)\phi(z)]g(z) - [g_\nu(z)\psi(z)]f(z) = g_\nu(z).$$

Let

$$g_\nu(z)\phi(z) = w(z)f(z) + r(z), \tag{15}$$

and

$$g_\nu(z)\psi(z) = \pi(z)g(z) + s(z),$$

where $r(z)$ and $s(z)$ are polynomials of degrees $n-1$ and $k-1$, respectively, at most. Then

$$[w(z)f(z) + r(z)]g(z) - [\pi(z)g(z) + s(z)]f(z) = g_\nu(z)$$

or

$$[w(z) - \pi(z)]f(z)g(z) + r(z)g(z) - s(z)f(z) = g_\nu(z).$$

Since $g_\nu(z)$ is of degree $n-1$ at most, the term in $f(z)g(z)$ must be zero, that is $w(z) = \pi(z)$. Hence

$$r(z)g(z) - s(z)f(z) = g_\nu(z).$$

Let

$$g_{\nu+1}(z) = r(z), \quad \phi_\nu(z) = s(z).$$

In practice, the polynomials $\phi(z)$ and $\psi(z)$ can be found once and for all, thereafter the sequence of $g_\nu(z)$ can be obtained by

$$g_{\nu+1}(z) = \phi(z)g_\nu(z) - w(z)f(z). \tag{15'}$$

Second method: Let

$$g(z) = (z - \alpha_1)(z - \alpha_2) \dots (z - \alpha_k).$$

For each ν , let

$$g_\nu^{[0]}(z) = g_\nu(z),$$

and define

$$g_\nu^{[i]}(z) = (z - \alpha_i)^{-1} [g_\nu^{[i-1]}(z) - g_\nu^{[i-1]}(\alpha_i) f(z) / f(\alpha_i)]$$

for $i = 1, 2, \dots, k$ and let

$$g_{\nu+1}(z) = g_\nu^{[k]}(z).$$

The verification of the second method is trivial.

In Theorem 1, the rate of convergence of the sequence $g_{\nu,p}(z)$ depends upon the ratio $|g(r_p)/g(r_{p+1})|$, when

$$|g(r_1)| \geq \dots \geq |g(r_p)| > |g(r_{p+1})| \geq \dots \geq |g(r_n)|.$$

Since

$$\lim_{\nu \rightarrow \infty} g_{\nu,p}^*(z) = f_{12\dots p}(z),$$

we can have

$$|g_{\nu,p}(r_p)/g_{\nu,p}(r_{p+1})| > |g(r_p)/g(r_{p+1})|,$$

for a ν sufficiently large. Hence if we redefine the recursion, replacing $g(z)$ by $g_{\nu,p}(z)$, then the convergence should be much faster. In particular, we will examine the case when $g(z)$ is replaced by $g_{\nu,p}(z)$ at each step.

Theorem 2. *In Theorem 1, suppose*

$$|g(r_1)| \geq \dots \geq |g(r_p)| > |g(r_{p-1})| \geq \dots \geq |g(r_n)|.$$

Then

$$\lim_{\nu \rightarrow \infty} g_{\nu, p}^*(z) = f_{12 \dots p}(z).$$

Also, suppose that

$$f_{12 \dots p}(r_i) \neq f_{12 \dots p}(r_j) \quad \text{for } r_i \neq r_j, \quad i, j = 1, 2, \dots, p.$$

Let

$$\sigma_0(z) = g_{M, p}^*(z),$$

for a sufficiently large M such that

$$|\sigma_0(r_j) - f_{12 \dots p}(r_j)| < \varepsilon, \quad j = 1, 2, \dots, n,$$

for a sufficiently small ε . Let

$$g_0(z) = \sigma_0(z)$$

and define $\sigma_\nu(z)$ for $\nu = 0, 1, 2, \dots$, by the following scheme:

Step 1. Let $g_\nu^{[0]}(z) = g_\nu(z)$ and define

$$g_\nu^{[i]}(z) = \sigma_\nu(z) g_\nu^{[i-1]}(z) - \phi_{\nu, i}(z) f(z), \quad i = 1, 2, \dots, i p$$

$$i p = \begin{cases} p & \text{in Rule 1, 2,} \\ 2p-1 & \text{in Rule 3,} \end{cases}$$

where each $g_\nu^{[i]}(z)$ is of degree $n-1$ at most and let

$$g_{\nu-1}(z) = g_\nu^{[1]}(z).$$

Step 2. Let $g_{\nu, 1}^{[i]}(z) = g_\nu^{[i]}(z)$ for each i , and for $p > 1$, define $g_{\nu, p}^{[1]}(z)$ by any one of the following rules:

Rule 1. For $k = 1, 2, \dots, p-1$; $i = 1, 2, \dots, p-k$, form $g_{\nu, k+1}^{[i]}(z)$ by eliminating the highest term between $g_{\nu, k}^{[i]}(z)$ and $g_{\nu, k}^{[i+1]}(z)$.

Rule 2. For $k = 1, 2, \dots, p-1$; $i = 1, 2, \dots, p-k$, form $g_{\nu, k+1}^{[i]}(z)$ by eliminating the constant term between $g_{\nu, k}^{[i]}(z)$ and $g_{\nu, k}^{[i+1]}(z)$ and dividing by z .

Rule 3. Define $g_{\nu, p}^{[1]}$ by the following determinant:

$$g_{\nu, p}^{[1]}(z) = \delta \begin{pmatrix} g_\nu^{[1]} & g_\nu^{[2]} & \dots & g_\nu^{[p]} \\ g_\nu^{[2]} & g_\nu^{[3]} & \dots & g_\nu^{[p]} \\ \dots & \dots & \dots & \dots \\ g_\nu^{[p]} & g_\nu^{[p]} & \dots & g_\nu^{[2p-1]} \end{pmatrix} [-f(z)]^{-(p-1)}.$$

Step 3. Let $\sigma_{\nu+1}(z) = g_{\nu, p}^{[1]}(z)$.

Then

$$\lim_{\nu \rightarrow \infty} \sigma_\nu(z) = f_{12 \dots p}(z)$$

with the order of convergence 2.

Proof. Rule 3 will be proved here. The reader can easily verify Rule 1 and Rule 2 in a similar way using Newton's formula for interpolation. We will consider the

case when r_i are distinct. Then we have, by Lagrange interpolation,

$$g_\nu(z) = \sum_{i=1}^n \sigma^{\langle \nu-1 \rangle}(r_i) f_i(z) / f_i(r_i),$$

where

$$\sigma^{\langle \nu \rangle}(z) = \sigma_\nu(z) \sigma_{\nu-1}(z) \dots \sigma_1(z) \sigma_0^2(z).$$

So,

$$g_{\nu, p}^{[1]}(z)$$

$$= \delta \left(\begin{array}{ccc} \sum_{i=1}^n \frac{\sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^n \frac{\sigma_\nu(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^n \frac{\sigma_\nu^{\rho-1}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \\ \sum_{i=1}^n \frac{\sigma_\nu(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^n \frac{\sigma_\nu^2(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^n \frac{\sigma_\nu^\rho(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n \frac{\sigma_\nu^{\rho-1}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^n \frac{\sigma_\nu^\rho(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^n \frac{\sigma_\nu^{2\rho-2}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \end{array} \right) \cdot [-f(z)]^{-(\rho-1)}.$$

First we note that

$$\begin{aligned} & \delta \left(\begin{array}{ccc} \sum_{i=1}^p \frac{\sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^p \frac{\sigma_\nu(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^p \frac{\sigma_\nu^{\rho-1}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \\ \sum_{i=1}^p \frac{\sigma_\nu(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^p \frac{\sigma_\nu^2(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^p \frac{\sigma_\nu^\rho(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^p \frac{\sigma_\nu^{\rho-1}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \sum_{i=1}^p \frac{\sigma_\nu^\rho(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} & \dots & \sum_{i=1}^p \frac{\sigma_\nu^{2\rho-2}(r_i) \sigma^{\langle \nu \rangle}(r_i) f_i(z)}{f_i(r_i)} \end{array} \right) \\ &= \delta \left[\begin{array}{c} \left(\begin{array}{cccc} 1 & 1 & \dots & 1 \\ \sigma_\nu(r_1) & \sigma_\nu(r_2) & \dots & \sigma_\nu(r_p) \\ \dots & \dots & \dots & \dots \\ \sigma_\nu^{\rho-1}(r_1) & \sigma_\nu^{\rho-1}(r_2) & \dots & \sigma_\nu^{\rho-1}(r_p) \end{array} \right) \text{diag} \left[\frac{\sigma^{\langle \nu \rangle}(r_1) f_1(z)}{f_1(r_1)}, \dots, \frac{\sigma^{\langle \nu \rangle}(r_p) f_p(z)}{f_p(r_p)} \right] \\ \cdot \left(\begin{array}{cccc} 1 & \sigma_\nu(r_1) & \dots & \sigma_\nu^{\rho-1}(r_1) \\ 1 & \sigma_\nu(r_2) & \dots & \sigma_\nu^{\rho-1}(r_2) \\ \dots & \dots & \dots & \dots \\ 1 & \sigma_\nu(r_p) & \dots & \sigma_\nu^{\rho-1}(r_p) \end{array} \right) \end{array} \right] \\ &= \zeta [\sigma_\nu(r_1), \sigma_\nu(r_2), \dots, \sigma_\nu(r_p)] \frac{\sigma^{\langle \nu \rangle}(r_1) \sigma^{\langle \nu \rangle}(r_2) \dots \sigma^{\langle \nu \rangle}(r_p) f_1(z) f_2(z) \dots f_p(z)}{f_1(r_1) f_2(r_2) \dots f_p(r_p)} \\ &= c f_1(z) f_2(z) \dots f_p(z), \end{aligned}$$

where c is a non-zero constant. Now suppose

$$|\sigma^{\langle \nu \rangle}(r_j)| < \varepsilon_\nu, \quad j = p+1, p+2, \dots, n.$$

Since

$$\begin{aligned} g_{\nu, p}^{[1]}(z) &= c f_1(z) f_2(z) \dots f_p(z) [-f(z)]^{-(\rho-1)} + \sum_{j=p+1}^n 0(\sigma^{\langle \nu \rangle}(r_j)) \\ &= \pm c f_{12} \dots f_p(z) + \sum_{j=p+1}^n 0(\sigma^{\langle \nu \rangle}(r_j)), \end{aligned}$$

we have

$$\sigma_{\nu+1}(z) = f_{12\dots p}(z) + \sum_{j=p+1}^n O(\sigma^{(p)}(r_j)),$$

and hence

$$|\sigma_{\nu+1}(z) - f_{12\dots p}(z)| \leq O(\epsilon_\nu),$$

and

$$|\sigma_{\nu+1}(r_j)| \leq O(\epsilon_\nu), \quad j = p+1, p+2, \dots, n.$$

Similarly,

$$\begin{aligned} \sigma_{\nu+2}(z) &= f_{12\dots p}(z) + \sum_{j=p+1}^n O(\sigma^{(p+1)}(r_j)) \\ &= f_{12\dots p}(z) + \sum_{j=p+1}^n O(\sigma_{\nu+1}(r_j) \sigma^{(p)}(r_j)), \end{aligned}$$

and hence

$$|\sigma_{\nu+2}(z) - f_{12\dots p}(z)| \leq O(\epsilon_\nu^2).$$

Since

$$|\sigma_0(r_j)| < \epsilon, \quad j = p+1, p+2, \dots, n,$$

where ϵ is a sufficiently small number, we can see that

$$\lim_{\nu \rightarrow \infty} \sigma_\nu(z) = f_{12\dots p}(z)$$

with the order of convergence 2. The confluent case can be handled in a similar way.

Theorem 2 is particularly useful when $p=1$ or $p=2$. Rule 1 or Rule 2 of Theorem 2 requires p evaluations of $g_\nu^{[i]}$ and $\frac{(p-1)p}{2}$ eliminations of the highest or lowest term between two polynomials to obtain each iterate, while Rule 1 or Rule 2 of Theorem 1 requires 1 evaluation of g_ν and $p-1$ eliminations. Also, Rule 3 of Theorem 2 requires $2p-1$ evaluations of $g_\nu^{[i]}$ while Rule 3 of Theorem 1 requires 1 evaluation of g_ν . Thus, when $p=1$, the amount of work involved is exactly the same in both algorithms, and, for that case, the recursion of Theorem 2 is

$$g_{\nu+1}(z) = g_\nu^2(z) \pmod{f},$$

while that of Theorem 1, for the choices $g(z) = z$ and $g_0(z) = 1$, is

$$g_{2\nu}(z) = g_\nu^2(z) \pmod{f}.$$

The latter case is discussed in detail by Bauer, and this shows that the original algorithm can be made quadratically convergent for $p=1$.

The dual form of Theorem 2 can also be obtained naturally.

Corollary 2. *In Corollary 1, suppose*

$$|g(r_1)| \leq \dots \leq |g(r_p)| < |g(r_{p+1})| \leq \dots \leq |g(r_n)|.$$

Then

$$\lim_{\nu \rightarrow \infty} g_{\nu,p}^*(z) = f_{12\dots p}(z).$$

Let $\sigma_0(z) = g_{M,p}^*(z)$, for an M sufficiently large and let

$$\varrho_0(z) = \frac{1}{\sigma_0(z)} [f(z) - \gamma_0(z)],$$

where $\gamma_0(z)$ is the remainder and $\varrho_0(z)$ is the quotient when $f(z)$ is divided by $\sigma_0(z)$. Then $\varrho_0(z)$ is sufficiently close to $(z - r_1)(z - r_2) \dots (z - r_p)$. Define $\varrho_v(z)$ and $\sigma_v(z)$, for $v = 0, 1, 2, \dots$, by the following scheme:

Step 1. Let $g_v^{[0]}(z) = g_v(z)$ and define

$$g_v^{[i]}(z) = \frac{1}{\varrho_v(z)} [g_v^{[i-1]}(z) - \phi_{v,i}(z) f(z)], \quad i = 1, 2, \dots, i \neq p,$$

where $\phi_{v,i}(z)$ is a polynomial of degree $p - 1$ at most such that $g_v^{[i-1]}(z) - \phi_{v,i}(z) f(z)$ is divisible by $\varrho_v(z)$ and let

$$g_{v+1}(z) = g_v^{[1]}(z).$$

Step 2. Same as Step 2 in Theorem 2.

Step 3. Let

$$\sigma_{v+1}(z) = g_{v,p}^{[1]*}(z)$$

and let

$$\varrho_{v+1}(z) = \frac{1}{\sigma_{v+1}(z)} [f(z) - \gamma_{v+1}(z)], \tag{16}$$

where $\gamma_{v+1}(z)$ is the remainder and $\varrho_{v+1}(z)$ is the quotient when $f(z)$ is divided by $\sigma_{v+1}(z)$. Then

$$\lim_{v \rightarrow \infty} \sigma_v(z) = f_{12 \dots p}(z)$$

and

$$\lim_{v \rightarrow \infty} \varrho_v(z) = (z - r_1)(z - r_2) \dots (z - r_p),$$

with the order of convergence 2.

Proof. It suffices to show that if

$$\sigma_{v+1}(z) = f_{12 \dots p}(z) + O(\varepsilon_{v+1}), \tag{17}$$

then

$$\varrho_{v+1}(z) = (z - r_1)(z - r_2) \dots (z - r_p) + O(\varepsilon_{v+1}). \tag{18}$$

From (16) and (17), we have that

$$\begin{aligned} \gamma_{v+1}(z) &= f(z) - \sigma_{v+1}(z) \varrho_{v+1}(z) \\ &= f(z) - [f_{12 \dots p}(z) + O(\varepsilon_{v+1})] \varrho_{v+1}(z) \\ &= f_{12 \dots p}(z) [(z - r_1)(z - r_2) \dots (z - r_p) - \varrho_{v+1}(z)] + O(\varepsilon_{v+1}). \end{aligned}$$

But, $\gamma_{v+1}(z)$ is of degree $n - p - 1$ at most while $f_{12 \dots p}(z)$ is of degree $n - p$. Thus $[(z - r_1)(z - r_2) \dots (z - r_p) - \varrho_{v+1}(z)]$ must be $O(\varepsilon_{v+1})$, which is (18). Also, note that $\gamma_{v+1}(z) = O(\varepsilon_{v+1})$.

In Corollary 2, when $p = 1$, we can see that

$$\sigma_v(z) = g_v^*(z),$$

and

$$f(z)/g_{\nu+1}^*(z) = \varrho_{\nu+1}(z) + O(\varepsilon_{\nu+1}). \tag{19}$$

Let

$$\varrho_{\nu}(z) = z - \beta_{\nu} \quad \text{for } \nu = 0, 1, 2, \dots \tag{20}$$

Consider the sequence α_{ν} defined by

$$\alpha_{\nu+1} = \alpha_{\nu} - f(\alpha_{\nu})/g_{\nu+1}^*(\alpha_{\nu}), \tag{21}$$

and let

$$q_{\nu}(z) = z - \alpha_{\nu}, \quad \nu = 0, 1, 2, \dots \tag{22}$$

Then

$$\begin{aligned} q_{\nu+1}(z) &= z - \alpha_{\nu+1} \\ &= z - \alpha_{\nu} + f(\alpha_{\nu})/g_{\nu+1}^*(\alpha_{\nu}) \quad \text{[by (21)]} \\ &= z - \alpha_{\nu} + \alpha_{\nu} - \beta_{\nu+1} + O(\varepsilon_{\nu+1}) \quad \text{[by (19 and (20))]} \\ &= \varrho_{\nu+1}(z) + O(\varepsilon_{\nu+1}). \quad \text{[by (20)]} \end{aligned}$$

Thus we can replace $\varrho_{\nu}(z)$ by $q_{\nu}(z)$ in Corollary 2, for $p=1$, without affecting the convergence. This is in fact the algorithm of Jenkins and Traub [8]:

$$g_{\nu+1}(z) = \frac{1}{z - \alpha_{\nu}} [g_{\nu}(z) - \phi_{\nu} f(z)],$$

where α_{ν} is defined by (21). Bauer and Samelson [3] give a similar algorithm replacing (21) by

$$\alpha_{\nu+1} = \alpha_{\nu} - f(\alpha_{\nu})/g_{\nu}^*(\alpha_{\nu}),$$

which requires somewhat less arithmetic per step but somewhat more steps. The algorithm of Jenkins and Traub using quadratic iteration [9] can also be regarded as a variant of Corollary 2 with $p=2$, although the formulation of the sequence $\varrho_{\nu}(z)$ is somewhat different.

Stewart [12] extended the algorithm of Jenkins and Traub [8] for $p \geq 2$: Given monic polynomials $\varrho_{\nu}(z)$ and $\sigma_{\nu}(z)$ of degree p and $n-p$ which are approximate divisors of $f(z)$, define

$$\sigma_{\nu+1}(z) = \frac{1}{\varrho_{\nu}(z)} [f(z) - \phi_{\nu}(z) \sigma_{\nu}(z)],$$

where $\phi_{\nu}(z)$ is of degree $n-p-1$ at most, so chosen that $f(z) - \phi_{\nu}(z) \sigma_{\nu}(z)$ is divisible by $\varrho_{\nu}(z)$. By interchanging $\varrho_{\nu}(z)$ and $\sigma_{\nu}(z)$, an analogous rule can be obtained for forming $\varrho_{\nu+1}(z)$. This algorithm is in fact identical with Samelson's factorization method [10] and the exact connection between these two algorithms is given in [6]. In Corollary 2, we can see that

$$\sigma_{\nu}(z) = \psi_{\nu}(z) g_{\nu}(z) - \theta_{\nu}(z) f(z), \quad \nu = 0, 1, 2, \dots,$$

where $\psi_{\nu}(z)$ and $\theta_{\nu}(z)$ are polynomials of degree p^2-p and p^2-p-1 , respectively. Since

$$\begin{aligned} \sigma_{\nu+1}(z) &= \psi_{\nu+1}(z) g_{\nu+1}(z) - \theta_{\nu+1}(z) f(z) \\ &= \psi_{\nu+1}(z) \cdot \frac{1}{\varrho_{\nu}(z)} [g_{\nu}(z) - \phi_{\nu.1}(z) f(z)] - \theta_{\nu+1}(z) f(z), \end{aligned}$$

we have

$$\begin{aligned} \varrho_\nu(z) \sigma_{\nu+1}(z) &= \psi_{\nu+1}(z) g_\nu(z) - [\psi_{\nu+1}(z) \phi_{\nu,1}(z) + \theta_{\nu+1}(z)] f(z) \\ &= \psi_{\nu+1}(z) \cdot \frac{\sigma_\nu(z) + \theta_\nu(z) f(z)}{\psi_\nu(z)} \pmod{f}. \end{aligned}$$

So,

$$\varrho_\nu(r_i) \sigma_{\nu+1}(r_i) = \psi_{\nu+1}(r_i) \cdot \sigma_\nu(r_i) / \psi_\nu(r_i), \quad i = 1, 2, \dots, n.$$

Let $\Phi_\nu(z)$ be a polynomial of degree $p-1$ such that

$$\Phi_\nu(r_i) = \psi_{\nu+1}(r_i) / \psi_\nu(r_i), \quad i = 1, 2, \dots, p.$$

Then $\Phi_\nu(z) \sigma_\nu(z)$ is a polynomial of degree $n-1$ such that

$$\varrho_\nu(r_i) \sigma_{\nu+1}(r_i) = \Phi_\nu(r_i) \sigma_\nu(r_i), \quad i = 1, 2, \dots, p, \tag{23}$$

and since

$$\lim_{\nu \rightarrow \infty} \sigma_\nu(z) = f_{12\dots p}(z),$$

we have

$$\varrho_\nu(r_i) \sigma_{\nu+1}(r_i) = \Phi_\nu(r_i) \sigma_\nu(r_i) + \varepsilon_\nu, \quad i = p+1, \dots, n, \tag{24}$$

where ε_ν can be arbitrarily small for ν sufficiently large. Since $\varrho_\nu(z) \sigma_{\nu+1}(z)$ is the monic polynomial of degree n such that (23) and (24) hold, we have

$$\varrho_\nu(z) \sigma_{\nu+1}(z) = f(z) - \Phi_\nu(z) \sigma_\nu(z) + \varepsilon_\nu,$$

which shows that Corollary 2 is almost equivalent to Stewart's method and hence to Samelson's. The formulation of the sequence $\varrho_\nu(z)$ is, however, somewhat different: We define in Corollary 2

$$\varrho_{\nu+1}(z) = \frac{1}{\sigma_{\nu+1}(z)} [f(z) - \gamma_{\nu+1}(z)],$$

while Samelson's method uses

$$\varrho_{\nu+1}(z) = \frac{1}{\sigma_\nu(z)} [f(z) - \varphi_\nu(z) \varrho_\nu(z)].$$

We will conclude this paper with a few comments. By the original algorithm of Sebastião, we can factor a given polynomial $f(z)$ as a product of polynomials each of which has equimodular zeros. As we do this, however, if the sequence $g_{\nu,p}$ turns out to converge, at some point of the recursion, by redefining the recursion replacing $g(z) = z$ by $g(z) = g_{\nu,p}(z)$ and using one of the generalized forms of the algorithm, we can accelerate the convergence of $g_{\nu,p}^*(z)$ to $f_{12\dots p}(z)$. We can even use the accelerated form in Theorem 2, which may be, however, impractical for $p > 2$. Now suppose that all zeros of $f(z)$ are equimodular and let α be any nonzero constant which is not a zero of $f(z)$. Then no more than two zeros of $f(z)$ can be equidistant from α , and if we set $g(z) = z - \alpha$ in the generalization, Theorem 1, then either the sequence $g_{\nu,1}(z)$ or the sequence $g_{\nu,2}(z)$ will converge. In either case, the accelerated form of the algorithm in Theorem 2 can be used to give faster convergence.

Acknowledgement. The author wishes to express her thanks to Dr. Alston Householder for his valuable suggestions in preparing this paper.

References

1. Bauer, F. L.: Beiträge zur Entwicklung numerischer Verfahren für programmgesteuerte Rechenanlagen. I. Quadratisch konvergente Durchführung der Bernoulli-Jacobischen Methoden zur Nullstellenbestimmung von Polynomen. Bayer. Akad. Wiss. Math. nat. Kl. S.-B., 275-303 (1954)
2. Bauer, F. L.: Beiträge zur Entwicklung numerischer Verfahren für programmgesteuerte Rechenanlagen. II. Direkte Faktorisierung eines Polynoms. Bayer. Akad. Wiss. Math. Nat. Kl. S.-B., 163-203 (1956)
3. Bauer, F. L., Samelson, K.: Polynomkerne und Iterationsverfahren. Math. Z. **67**, 93-98 (1957)
4. Householder, A. S.: Generalizations of an Algorithm of Sebastião e Silva. Numer. Math. **16**, 375-382 (1971)
5. Householder, A. S.: Multigradients and the Zeros of Transcendental Functions. Linear Algebra and Its Application **4**, 175-182 (1971)
6. Householder, A. S., Stewart, G. W.: Comments on "Some Iterations for Factoring Polynomials", Numer. Math. **13**, 470-471 (1969)
7. Householder, A. S., Stewart, G. W.: The Numerical Factorization of a Polynomial. SIAM Review **13**, 38-46 (1971)
8. Jenkins, M. A., Traub, J. F.: A Three-Stage Variable-Shift Iteration for Polynomial Zeros and Its Relation to Generalized Rayleigh Iteration. Numer. Math. **14**, 252-263 (1970)
9. Jenkins, M. A., Traub, J. F.: A Three-Stage Algorithm for Real Polynomials Using Quadratic Iteration, SIAM J. Numer. Anal. **7**, 545-566 (1970)
10. Samelson, K.: Faktorisierung von Polynomen durch funktionale Iteration. Bayer. Akad. Wiss., Math. Nat. Kl., Abhandlungen 95 (1959)
11. Sebastião e Silva, J.: Sur une méthode d'approximation semblable a celle de Graeffe, Portugal. Math. **2**, 271-279 (1941)
12. Stewart, G. W.: Some Iterations for Factoring a Polynomial. Numer. Math. **13**, 458-470 (1969)
13. Stewart, G. W.: On a Companion Operator For Analytic Functions. Numer. Math. **18**, 26-43 (1971)

Soon Park Chung,
Ass. Prof.
Dept. of Mathematics
Univ. of Michigan
Dearborn, Michigan 48128
U.S.A.