

# A Class of Globally Convergent Iteration Functions for the Solution of Polynomial Equations

By J. F. Traub\*

## TABLE OF CONTENTS

1. Introduction.....	113
2. Brief Survey of Results.....	113
3. The $G$ Polynomials.....	114
4. The Generalized $G$ Polynomials.....	116
5. The Formation of Iteration Functions.....	119
6. The Order and Asymptotic Error Constant.....	120
7. Proof of Global Convergence.....	121
8. The Character of the Convergence.....	124
9. Normalized Calculations.....	126
10. Two Algorithms for the Accelerated Calculation of the $G$ Polynomials..	126
11. Iteration Functions for Small Values of $\lambda$ .....	129
12. Relation to Other Methods.....	130
Appendix A. Numerical Examples.....	133
Appendix B. Summary of Formulas for the Construction of Iteration Functions....	135
Appendix C. Notation.....	137
References.....	137

**1. Introduction.** The approximation of a zero of a polynomial  $P$  by an iterative method generally consists of two parts: Obtain a “good” approximation to the zero of interest and then iterate to a solution. Furthermore the classical iteration functions possess a property which may be illustrated by the Newton-Raphson iteration function,  $\varphi = t - P/P'$ . This function is not defined at infinity or at the zeros of  $P'$ .

We introduce a class of new iteration functions which are ratios of polynomials of the same degree and hence defined at infinity. The poles of these rational functions occur at points which cause no difficulty.

The classical iteration functions are given as explicit functions of  $P$  and its derivatives. The new iteration functions are constructed according to a certain algorithm. This construction requires only simple polynomial manipulation which may be performed on a computer.

We shall treat here only the important case that the zeros of  $P$  are distinct and that the dominant zero is real. The extension to multiple zeros, dominant complex zeros, and subdominant zeros will be given in another paper. We shall restrict ourselves to questions relevant to the calculation of zeros. Certain aspects of our investigations which are of broader interest will be reported elsewhere.

**2. Brief Survey of Results.** The  $G$  polynomials are defined as the remainder of a division process. Their importance in our work is due to the property exhibited

---

Received June 14, 1965.

\* Portions of this paper were presented to the American Mathematical Society, New York City, April, 1965 and to IFIP Congress 65, New York City, May, 1965.

in (3.4). A simple recursion for their calculation is given by (3.7). The coefficients of the  $G$  polynomials are also of interest. Two of their important properties are given by (3.17) and (3.18).

The generalized  $G$  polynomials are introduced in Section 4. They are most conveniently generated using (4.7) and (4.8). A determinantal formula is given by (4.19). The  $G$  polynomials and the generalized  $G$  polynomials are used to form iteration functions in Section 5. Some insight into these new iteration functions may be obtained from the discussion at the end of the section. The order and asymptotic error constant are given in Section 6.

The main result of this paper is the theorem of Section 7 which essentially states that the sequence of approximants is guaranteed to converge. A bound on the error after  $i$  iterations is developed at the end of this section. Theorems concerning the monotonicity of the convergence are given in Section 8.

A method of normalized calculation which is important for numerical application is described in Section 9. Algorithms for the accelerated calculation of the  $G$  polynomials are described in Section 10. The iteration functions yield globally convergent approximation sequences only if a certain parameter is chosen sufficiently large. Then the iteration function cannot be prescribed as an explicit function of  $P$  and its derivatives. For small values of this parameter the iteration function can be explicitly given and this leads to an interesting set of iteration functions which are given in Section 11.

In Section 12 we show that many classical methods (nonglobal) may be obtained as special cases of our formulation. A brief survey of the literature ends this section.

Appendix A contains a very brief discussion of computer implementation and gives a number of numerical examples. Important formulas are summarized in Appendix B while notation is summarized in Appendix C.

### 3. The $G$ Polynomials. Let

$$(3.1) \quad P(t) = \sum_{j=0}^n a_{n-j} t^j, \quad a_0 = 1.$$

We assume that  $P(t)$  has real coefficients but this is not essential. Let  $P$  have distinct zeros  $\rho_1, \rho_2, \dots, \rho_n$ , with  $|\rho_1| > |\rho_j|$ ,  $j = 2, 3, \dots, n$ . Thus the dominant zero is real.

Let  $B(t)$  be any polynomial of degree  $m < n$  such that  $B(\rho_1) \neq 0$ . Let  $\lambda$  be a non-negative integer and define polynomials  $G(\lambda, t, B)$  and  $Q(\lambda, t, B)$  as the remainder and quotient of the division of  $B(t)t^\lambda$  by  $P(t)$ . Thus

$$(3.2) \quad B(t)t^\lambda = G(\lambda, t, B) + P(t)Q(\lambda, t, B).$$

For many purposes we need not indicate the dependence of  $G$  and  $Q$  on  $B$  and we write  $G(\lambda, t)$  and  $Q(\lambda, t)$ . Clearly  $G(\lambda, t)$  and  $Q(\lambda, t)$  are of degrees  $n - 1$  and  $m + \lambda - n$  respectively, with  $Q(\lambda, t) \equiv 0$  if  $m + \lambda - n < 0$ . The  $G$  polynomials play a basic role in our work; the  $Q$  polynomials play only an auxiliary role.

From (3.2),

$$G(\lambda, \rho_i) = B(\rho_i)\rho_i^\lambda, \quad i = 1, 2, \dots, n.$$

By Lagrange's interpolation formula,

$$(3.3) \quad G(\lambda, t) = \sum_{i=1}^n B(\rho_i) \rho_i^\lambda L_i(t),$$

where

$$L_i(t) = \frac{P(t)}{t - \rho_i} \frac{1}{P'(\rho_i)}.$$

Hence

$$(3.4) \quad \frac{G(\lambda, t)}{P(t)} = \sum_{i=1}^n \frac{B(\rho_i) \rho_i^\lambda}{P'(\rho_i) (t - \rho_i)}.$$

Let  $\beta(\lambda)$  be the weighted power sum

$$(3.5) \quad \beta(\lambda) = \sum_{i=1}^n \frac{B(\rho_i)}{P'(\rho_i)} \rho_i^\lambda.$$

Let  $E$  be the displacement operator

$$Eh(\lambda) = h(\lambda + 1).$$

Then from (3.4)

$$(3.6) \quad (t - E)G(\lambda, t) = \beta(\lambda)P(t).$$

Clearly  $\beta(\lambda)$  is just the leading coefficient of  $G(\lambda, t)$  which we label  $\alpha_0(\lambda)$ . Since  $G(0, t, B) = B(t)$ , we have the recursion

$$(3.7) \quad G(0, t, B) = B(t), \quad G(\lambda + 1, t, B) = tG(\lambda, t, B) - \alpha_0(\lambda)P(t).$$

This provides a simple recursion for forming the  $G(\lambda, t)$  on a machine or by hand. It is clear that

$$(3.8) \quad P(E)G(\lambda, t) = 0.$$

Forming  $G(\lambda, t)$  by the recursion (3.7) is far simpler than using the recursion (3.8). In hand calculation, the  $G(\lambda, t)$  may be computed by a method of detached coefficients using a movable strip of paper on which the coefficients of  $-P$  have been written. A method for calculating  $G(2\lambda, t)$  directly from  $G(\lambda, t)$ ,  $G(\lambda + 1, t)$ ,  $\dots$ ,  $G(\lambda + n - 1, t)$  is described in Section 10.

We obtain a number of additional properties of the  $G(\lambda, t)$ . From (3.3),

$$(3.9) \quad G(\lambda, t, B) = \sum_{j=0}^{n-1} \alpha_j(\lambda) t^{n-1-j},$$

where

$$(3.10) \quad \alpha_j(\lambda) = \sum_{i=1}^n \frac{B(\rho_i)}{P'(\rho_i)} \rho_i^\lambda A_j(\rho_i).$$

(We should write  $\alpha_j(\lambda, B)$  but will not do so in order to simplify notation.) The  $A_j(t)$  are the "associated polynomials" of  $P$  (Traub [28]) given by

$$(3.11) \quad A_j(t) = \sum_{r=0}^j a_{j-r} t^r.$$

We may also write

$$(3.12) \quad \alpha_j(\lambda) = A_j(E)\beta(\lambda).$$

The expansion of the  $G$  polynomials in terms of the associated polynomials of  $P$  is given by

$$(3.13) \quad G(\lambda, t) = \sum_{j=0}^{n-1} \beta(\lambda + j)A_{n-1-j}(t).$$

In terms of the coefficients  $\alpha_j(\lambda)$ , the recursion on  $G(\lambda, t)$  becomes

$$(3.14) \quad \alpha_j(\lambda + 1) = \alpha_{j+1}(\lambda) - \alpha_0(\lambda)a_{j+1}, \quad j = 0, 1, \dots, n-1, \quad \alpha_n(\lambda) = 0.$$

We note two important properties of the  $\alpha_j(\lambda)$  for the case  $B = 1$ .

Consider the homogeneous linear difference equation

$$(3.15) \quad P(E)y(\lambda) = 0$$

with initial conditions

$$(3.16) \quad E^\mu y(0) = y_\mu, \quad \mu = 0, 1, \dots, n-1.$$

Then (Traub [26], [27]),

$$(3.17) \quad y(\lambda) = \sum_{j=0}^{n-1} y_{n-1-j} \alpha_j(\lambda).$$

That is, the  $\alpha_j(\lambda)$  are the weights of the starting conditions in the general solution. Also,

$$(3.18) \quad \rho_i^\lambda = \sum_{j=0}^{n-1} \rho_i^{n-1-j} \alpha_j(\lambda), \quad i = 1, 2, \dots, n.$$

Hence the  $\alpha_j(\lambda)$ , which may be expressed as polynomials in the coefficients of  $P$ , are the coefficients in the formula for an arbitrary power of a zero of  $P$  in terms of the first  $n-1$  powers.

Let  $\bar{h}(t)$  denote the polynomial  $h(t)$  divided by its leading coefficient. Thus  $\bar{G}(\lambda, t) = G(\lambda, t)/\beta(\lambda)$  and from (3.3) and (3.5),

$$(3.19) \quad \lim_{\lambda \rightarrow \infty} \bar{G}(\lambda, t) = \frac{P(t)}{t - \rho_1}.$$

From (3.10),

$$(3.20) \quad \lim_{\lambda \rightarrow \infty} \frac{\alpha_j(\lambda + 1)}{\alpha_j(\lambda)} = \rho_1, \quad j = 0, 1, \dots, n-1.$$

Assume that  $A_j(\rho_1) \neq 0$ . Then

$$(3.21) \quad \lim_{\lambda \rightarrow \infty} \frac{\alpha_{j+1}(\lambda)}{\alpha_j(\lambda)} = \frac{A_{j+1}(\rho_1)}{A_j(\rho_1)} = \rho_1 + \frac{a_{j+1}}{A_j(\rho_1)}, \quad j = 0, 1, \dots, n-2.$$

#### 4. The Generalized $G$ Polynomials. Let

$$(4.1) \quad G_p(\lambda, t) = \frac{(-1)^{p-1}}{(p-1)!} P^p(t) \left[ \frac{G(\lambda, t)}{P(t)} \right]^{(p-1)}.$$

The symbols  $G_1$  and  $G$  are used interchangeably. Then

$$(4.2) \quad \frac{G_p(\lambda, t)}{P^p(t)} = \sum_{i=1}^n \frac{c_i \rho_i^\lambda}{(t - \rho_i)^p},$$

with

$$(4.3) \quad c_i = \frac{B(\rho_i)}{P'(\rho_i)}.$$

Hence

$$(4.4) \quad G_p(\lambda, t) = \sum_{i=1}^n B(\rho_i) [P'(\rho_i)]^{p-1} \rho_i^\lambda L_i^p(t)$$

and

$$(4.5) \quad G_p(\lambda, \rho_i) = B(\rho_i) [P'(\rho_i)]^{p-1} \rho_i^\lambda, \quad i = 1, 2, \dots, n.$$

Differentiation of (4.1) yields

$$(4.6) \quad G_{p+1}(\lambda, t) = P'(t)G_p(\lambda, t) - \frac{P(t)}{p} G_p'(\lambda, t), \quad p = 1, 2, \dots.$$

By Leibniz's rule,

$$(4.7) \quad \begin{aligned} G_p(\lambda, t) &= \sum_{k=0}^{p-1} [-P(t)]^{p-1-k} \frac{G^{(p-1-k)}(\lambda, t)}{(p-1-k)!} V_k(t), \\ V_k(t) &= \frac{(-1)^k}{k!} P^{k+1}(t) \left[ \frac{1}{P(t)} \right]^{(k)}. \end{aligned}$$

The  $V_k(t)$  may be calculated recursively by

$$(4.8) \quad V_k(t) = P'(t)V_{k-1}(t) - \frac{P(t)}{k} V_{k-1}'(t), \quad V_0(t) = 1.$$

The  $V_k(t)$  may also be written in terms of Bell polynomials (Traub [25]) as

$$(4.9) \quad V_k(t) = \frac{1}{k!} \sum_{l=0}^k l! (-P)^{k-l} U_{k,l} [P', \dots, P^{(k-l+1)}],$$

where the  $U_{k,l}$ , the coefficients of the  $k$ th Bell polynomial, are defined by

$$B_k(\omega; g_1, \dots, g_k) = \sum_{l=0}^k U_{k,l}(g_1, \dots, g_{k-l+1}) \omega^l.$$

Since

$$(4.10) \quad V_k(t) = G_{k+1}(0, t, 1),$$

we may write

$$(4.11) \quad G_p(\lambda, t, B) = \sum_{k=0}^{p-1} [-P(t)]^{p-1-k} \frac{G^{(p-1-k)}(\lambda, t, B)}{(p-1-k)!} G_{k+1}(0, t, 1),$$

and, in particular,

$$(4.12) \quad G_p(0, t, B) = \sum_{k=0}^{p-1} [-P(t)]^{p-1-k} \frac{B^{(p-1-k)}(t)}{(p-1-k)!} G_{k+1}(0, t, 1).$$

Equations (4.7) and (4.8) give a very practical method for calculating  $G_p(\lambda, t, B)$  on a computing machine. Generalization of the  $G_p(\lambda, t)$  to the case of  $\lambda$  and  $p$  negative will be discussed in a later paper.

We develop a number of relations involving the  $G_p(\lambda, t)$ . From (4.2),

$$(4.13) \quad (t - E)G_p(\lambda, t) = G_{p-1}(\lambda, t)P(t)$$

or

$$(4.14) \quad G_p(\lambda + 1, t) = tG_p(\lambda, t) - G_{p-1}(\lambda, t)P(t).$$

Since  $G_0(\lambda, t) = \beta(\lambda) = \alpha_0(\lambda)$ , this generalizes (3.7) and is valid for  $p = 1, 2, \dots$ . From (4.13),

$$(4.15) \quad (t - E)^p G_p(\lambda, t) = \beta(\lambda)P^p(t)$$

or

$$(4.16) \quad \sum_{j=0}^p (-1)^j C(p, j) G_p(\lambda + j, t) t^{p-j} = \beta(\lambda)P^p(t),$$

where  $C(p, j) = p! / [(p - j)! j!]$ . Also,

$$(4.17) \quad (t^j - E^j)G_p(\lambda, t) = P(t) \sum_{k=0}^{j-1} t^{j-1-k} E^k G_{p-1}(\lambda, t)$$

or

$$(4.18) \quad G_p(\lambda + j, t) = t^j G_p(\lambda, t) - P(t) \sum_{k=0}^{j-1} t^{j-1-k} G_{p-1}(\lambda + k, t).$$

One may easily show that  $G_p(\lambda, t)$  satisfies the determinantal formula,

$$(4.19) \quad G_p(\lambda, t) = \begin{vmatrix} G(\lambda, t) & P & 0 & \cdots & 0 \\ G'(\lambda, t) & P' & P & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \frac{G^{(p-2)}(\lambda, t)}{(p-2)!} & \frac{P^{(p-2)}}{(p-2)!} & \frac{P^{(p-3)}}{(p-3)!} & \cdots & P \\ \frac{G^{(p-1)}(\lambda, t)}{(p-1)!} & \frac{P^{(p-1)}}{(p-1)!} & \frac{P^{(p-2)}}{(p-2)!} & \cdots & P' \end{vmatrix}.$$

Expanding this determinant in elements of the last row shows that

$$(4.20) \quad \begin{aligned} G_p(\lambda, t) &= P' G_{p-1}(\lambda, t) - \frac{P''}{2} P G_{p-2}(\lambda, t) \\ &+ \cdots + (-1)^{p-2} \frac{P^{(p-1)}}{(p-1)!} P^{p-2} G(\lambda, t) + \frac{(-1)^{p-1}}{(p-1)!} P^{p-1} G^{(p-1)}(\lambda, t). \end{aligned}$$

Expansion of the determinant in terms of the first column must yield (4.7).

We derive a number of limit relations for  $G_p(\lambda, t)$ . From (4.2), it follows that

$$\bar{G}_p(\lambda, t) = \frac{G_p(\lambda, t)}{\beta(\lambda)}$$

and

$$\lim_{\lambda \rightarrow \infty} \tilde{G}_p(\lambda, t) = \left( \frac{P(t)}{t - \rho_1} \right)^p.$$

Hence

$$(4.21) \quad \lim_{\lambda \rightarrow \infty} \frac{G_p(\lambda, t)}{G_{p-1}(\lambda, t)} = \frac{P(t)}{t - \rho_1}.$$

Furthermore

$$(4.22) \quad \lim_{p \rightarrow \infty} \frac{G_p(\lambda, t)}{G_{p-1}(\lambda, t)} = \frac{P(t)}{t - \rho_1}$$

provided  $|t - \rho_1| < |t - \rho_j|, j = 2, 3, \dots, n.$

**5. The Formation of Iteration Functions.** For  $\lambda$  and  $p$  fixed, define

$$(5.1) \quad \varphi_p(\lambda, t, B) = \frac{G_p(\lambda + 1, t, B)}{G_p(\lambda, t, B)}.$$

We will often not indicate the dependence of  $\varphi$  on  $B$  explicitly.  $\varphi_p(\lambda, t)$  will be used as an iteration function. That is, we define a sequence by

$$t_{i+1} = \varphi_p(\lambda, t_i), \quad i = 0, 1, \dots.$$

From (4.14) it follows that

$$(5.2) \quad \varphi_p(\lambda, t) = t - P(t) \frac{G_{p-1}(\lambda, t)}{G_p(\lambda, t)}.$$

This exhibits  $\varphi_p(\lambda, t)$  in “canonical form” for an iteration function. It permits the calculation of each member of the sequence as a small correction to the preceding member. In particular,

$$(5.3) \quad \varphi_1(\lambda, t) = t - P(t) \frac{\alpha_0(\lambda)}{G(\lambda, t)},$$

$$(5.4) \quad \varphi_2(\lambda, t) = t - P(t) \frac{G(\lambda, t)}{P'(t)G(\lambda, t) - P(t)G'(\lambda, t)}.$$

For purposes of comparison (but not for calculation) this may be written as

$$(5.5) \quad \varphi_2(\lambda, t) = t - \frac{u(t)}{1 - u(t)W(\lambda, t)},$$

where

$$u(t) = \frac{P(t)}{P'(t)}, \quad W(\lambda, t) = \frac{G'(\lambda, t)}{G(\lambda, t)}.$$

This iteration function may be compared to the Newton-Raphson iteration function  $\varphi(t) = t - u(t)$ . One may also write

$$(5.6) \quad \varphi_2(\lambda, t) = t - u(t) - \frac{u^2(t)W(\lambda, t)}{1 - u(t)W(\lambda, t)}.$$

This exhibits  $\varphi_2(\lambda, t)$  as the Newton-Raphson function plus terms of order  $u^2(t)$ . Similarly the third order iteration function, which is given in Appendix B, may be viewed as a generalization of Halley's function (Traub [24, p. 91]),

$$\varphi(t) = t - \frac{u(t)}{1 - A_2(t)u(t)}, \quad A_2(t) = \frac{P''(t)}{2P'(t)}.$$

We obtain insight into the new iteration functions by observing that if  $Q_1(t) = P(t)/(t - \rho_1)$ , then

$$(5.7) \quad \rho_1 = t - \frac{P(t)}{Q_1(t)}.$$

$Q_1(t)$  is not known but we approximate it as well as we can. Let

$$\varphi(t) = t - \frac{P(t)}{H(t)}$$

be a general iteration function. In the Newton-Raphson iteration function we take  $H(t) = P'(t)$  and hence

$$H(\rho_1) = Q_1(\rho_1).$$

That is,  $H(t)$  and  $Q_1(t)$  are equal only at the zero. On the other hand if we choose

$$H(t) = H(\lambda, t) = \frac{G_p(\lambda, t)}{G_{p-1}(\lambda, t)},$$

then

$$\lim_{\lambda \rightarrow \infty} H(\lambda, t) = Q_1(t)$$

for all  $t$ . Hence for  $\lambda$  sufficiently large,  $H(t)$  approximates  $Q_1(t)$  for all  $t$ .

**6. The Order and Asymptotic Error Constant.** We study the convergence of the iteration in the neighborhood of  $\rho_1$ . Convergence in the large is studied in Sections 7 and 8.

If there exist real numbers  $r$  and  $C_r(\lambda, B)$  such that

$$(6.1) \quad \lim_{t \rightarrow \rho_1} \frac{\varphi_p(\lambda, t, B) - \rho_1}{(t - \rho_1)^r} = C_r(\lambda, B),$$

then  $r$  is called the order and  $C_r(\lambda, B)$  is called the asymptotic error constant of  $\varphi_p(\lambda, t)$  (Traub [24, p. 9]). From (4.2) and (5.1) it follows that

$$\varphi_p(\lambda, t) = \frac{\sum_{i=1}^n c_i \rho_i^{\lambda+1} / (t - \rho_i)^p}{\sum_{i=1}^n c_i \rho_i^\lambda / (t - \rho_i)^p}.$$

Hence

$$(6.2) \quad \varphi_p(\lambda, t) - \rho_1 = \frac{\sum_{i=2}^n d_i \mu_i^\lambda (\rho_i - \rho_1) \left( \frac{t - \rho_1}{t - \rho_i} \right)^p}{1 + \sum_{i=2}^n d_i \mu_i^\lambda \left( \frac{t - \rho_1}{t - \rho_i} \right)^p},$$



where

$$(6.3) \quad d_i = \frac{c_i}{c_1}, \quad \mu_i = \frac{\rho_i}{\rho_1}, \quad i = 2, 3, \dots, n.$$

(Since  $B(\rho_1) \neq 0, c_1 \neq 0$ .) Hence

$$(6.4) \quad \lim_{t \rightarrow \rho_1} \frac{\varphi_p(\lambda, t) - \rho_1}{(t - \rho_1)^p} = - \sum_{i=2}^n \frac{d_i \mu_i^\lambda}{(\rho_1 - \rho_i)^{p-1}}.$$

Thus the order of  $\varphi_p(\lambda, t)$  is  $p$  and the asymptotic error constant  $C_p(\lambda)$  is given by the right side of (6.4). In particular,

$$(6.5) \quad C_p(\lambda, P') = - \sum_{i=2}^n \frac{\mu_i^\lambda}{(\rho_1 - \rho_i)^{p-1}},$$

$$(6.6) \quad C_p(\lambda, 1) = - \sum_{i=2}^n \frac{P'(\rho_1)}{P'(\rho_i)} \frac{\mu_i^\lambda}{(\rho_1 - \rho_i)^{p-1}}.$$

Observe that

$$\lim_{\lambda \rightarrow \infty} C_p(\lambda) = 0.$$

If we assume that  $|\rho_1| > |\rho_2| > |\rho_j|, j = 3, 4, \dots, n$ , then

$$(6.7) \quad \lim_{\lambda \rightarrow \infty} \frac{C_p(\lambda)}{\mu_2^\lambda} = - \frac{d_2}{(\rho_1 - \rho_2)^{p-1}}.$$

We consider some special cases. It is easy to show that  $\varphi_2(0, t, 1)$  is the Newton-Raphson function. We have

$$(6.8) \quad C_2(0, 1) = - \sum_{i=2}^n \frac{P'(\rho_1)}{P'(\rho_i)} \frac{1}{\rho_1 - \rho_i}.$$

One may show that the right side of (6.8) is identical with  $P''(\rho_1)/[2P'(\rho_1)]$ , which is the standard result. If  $n = 2$ , then

$$C_p(\lambda, P') = - \frac{(\rho_2/\rho_1)^\lambda}{[P'(\rho_1)]^{p-1}},$$

$$C_p(\lambda, 1) = \frac{(\rho_2/\rho_1)^\lambda}{[P'(\rho_1)]^{p-1}}.$$

In particular,

$$C_2(\lambda, 1) = \frac{(\rho_2/\rho_1)^\lambda}{P'(\rho_1)},$$

while the corresponding result for Newton-Raphson is just  $1/P'(\rho_1)$ . Hence for this special case,  $C$  has been decreased precisely by  $(\rho_2/\rho_1)^\lambda$ .

**7. Proof of Global Convergence.** We shall prove the following

**THEOREM.** *Let the zeros  $\rho_i$  of the polynomial  $P$  be distinct with  $|\rho_1| > |\rho_i|, i = 2, 3, \dots, n$ . Let  $t_0$  be an arbitrary point in the extended complex plane such that  $t_0 \neq \rho_2, \rho_3, \dots, \rho_n$  and let  $t_{i+1} = \varphi_p(\lambda, t_i)$ . Then for all sufficiently large but fixed  $\lambda$  the sequence  $t_i$  is defined for all  $i$  and  $t_i \rightarrow \rho_1$ .*

*Remark.* Since  $\rho_1$  is real, one would in practice take  $t_0$  real.

*Remark.* Since  $\varphi_p(\lambda, \rho_j) = \rho_j, j = 1, 2, \dots, n$ , we have an

**ALTERNATIVE FORMULATION.** *Let the zeros  $\rho_1$  of the polynomial  $P$  be distinct with  $|\rho_1| > |\rho_i|, i = 2, 3, \dots, n$ . Let  $t_{i+1} = \varphi_p(\lambda, t_i)$ . Then for all sufficiently large but fixed  $\lambda$  the sequence  $t_i$  is defined for all  $i$  and  $t_i \rightarrow \rho_j$  for some  $j$ .*

Before proving the theorem we prove a number of simple lemmas. It is convenient to introduce the following notation. Let  $K$  be a circle of radius  $R$  centered at  $\rho_1$  so that no other zeros of  $P$  lie within or on  $K$  and let

$$S = \{t \mid |t - \rho_1| < R\}.$$

We have

**LEMMA 1.** *For all sufficiently large  $\lambda, \varphi_p(\lambda, \infty) \in S$ .*

*Proof.* From (6.2),

$$\varphi_p(\lambda, \infty) - \rho_1 = \frac{\sum_{i=2}^n d_i(\rho_i - \rho_1)\mu_i^\lambda}{1 + \sum_{i=2}^n d_i\mu_i^\lambda}.$$

One can clearly choose  $\lambda$  so large that  $|\varphi_p(\lambda, \infty) - \rho_1| < R$ .

**LEMMA 2.** *Let  $t_0 \neq \infty, \rho_2, \rho_3, \dots, \rho_n$ . For all sufficiently large  $\lambda, G_p(\lambda, t_0) \neq 0$ .*

*Proof.* From (4.2),

$$G_p(\lambda, t_0) = c_1\rho_1^\lambda(Q_1(t_0))^p[1 + e(\lambda, t_0)],$$

where  $Q_1(t) = P(t)/(t - \rho_1)$  and where  $e(\lambda, t_0)$  is defined for any  $t_0 \neq \rho_2, \dots, \rho_n$  and

$$\lim_{\lambda \rightarrow \infty} e(\lambda, t_0) = 0.$$

Hence the expression in brackets does not vanish for  $\lambda$  sufficiently large. Since  $Q_1(t_0)$  does not vanish,  $G_p(\lambda, t_0)$  does not vanish for sufficiently large finite  $\lambda$ .

**LEMMA 3.** *Let  $t_0 \neq \infty$ . For all sufficiently large  $\lambda, \varphi_p(\lambda, t_0) \in S$ .*

*Proof.*

$$\varphi(\lambda, t_0) - \rho_1 = \frac{\sum_{i=2}^n d_i\mu_i^\lambda(\rho_i - \rho_1) \left(\frac{t_0 - \rho_1}{t_0 - \rho_i}\right)^p}{1 + \sum_{i=2}^n d_i\mu_i^\lambda \left(\frac{t_0 - \rho_1}{t_0 - \rho_i}\right)^p}.$$

One can clearly choose  $\lambda$  so large that

$$|\varphi_p(\lambda, t_0) - \rho_1| < R.$$

**LEMMA 4.** *For all sufficiently large  $\lambda, G_p(\lambda, t) \neq 0$  for any  $t \in S$ .*

*Proof.* Consider any  $t \in S$ . Then

$$G_p(\lambda, t) = c_1\rho_1^\lambda Q_1^p(t)[1 + e(\lambda, t)],$$

where  $Q_1(t) = P(t)/(t - \rho_1)$  and where  $e(\lambda, t)$  is defined for all  $t \in S$  and  $\lim_{\lambda \rightarrow \infty} e(\lambda, t) = 0$ . Hence the expression in brackets does not vanish for  $\lambda$  sufficiently large. Since  $Q_1(t)$  does not vanish for  $t \in S, G_p(\lambda, t) \neq 0$  for sufficiently large finite  $\lambda$ .

LEMMA 5. For all sufficiently large  $\lambda$  and for all  $t \in S$ ,

$$(7.1) \quad |\varphi_p(\lambda, t) - \rho_1| \leq L |t - \rho_1|, \quad L < 1.$$

*Proof.* Since  $\varphi_p(\lambda, \rho_1) = \rho_1$ , the result is true for  $t = \rho_1$ . Assume  $t \neq \rho_1$ . Then

$$\begin{aligned} \frac{\varphi_p(\lambda, t) - \rho_1}{t - \rho_1} &= \frac{G_p(\lambda + 1, t) - \rho_1}{(t - \rho_1)G_p(\lambda, t)} \\ &= (t - \rho_1)^{p-1} \frac{\sum_{i=2}^n d_i \mu_i^\lambda \frac{(\rho_i - \rho_1)}{(t - \rho_i)^p}}{1 + \sum_{i=2}^n d_i \mu_i^\lambda \left(\frac{t - \rho_1}{t - \rho_i}\right)^p}. \end{aligned}$$

Order the  $\rho_i$  so that  $|\rho_2| \geq |\rho_i|, i = 3, 4, \dots, n$ . Hence  $|\mu_2| \geq |\mu_i|$ . Let  $D = \sum_{i=2}^n |d_i|$ . Since  $t$  lies inside the circle  $K$  while  $\rho_2, \rho_3, \dots, \rho_n$  lie outside  $K, |t - \rho_i|^p \geq M^{-1}, i = 2, 3, \dots, n$  for some number  $M$ . Recalling that for all  $t \in S, |t - \rho_1| < R$ , we conclude

$$\left| \frac{\varphi_p(\lambda, t) - \rho_1}{t - \rho_1} \right| \leq \frac{2 |\rho_1| DR^{p-1}M |\mu_2|^\lambda}{1 - R^p DM |\mu_2|^\lambda}.$$

For all sufficiently large  $\lambda$  the right side of this inequality may be made as small as desired. The uniform bound specified in (7.1) therefore holds for all  $t$  in  $S$ .

*Proof of Theorem.* If  $t_0 = \infty$  apply Lemma 1. Hence it is sufficient to consider the case where  $t_0$  is finite. From Lemmas 2 and 3 we may choose  $\lambda$  so large that  $\varphi_p(\lambda, t_0)$  is defined and  $\varphi_p(\lambda, t_0) \in S$ . By Lemmas 4 and 5 we choose  $\lambda$  so large that  $\varphi_p(\lambda, t)$  is defined for all  $t \in S$  and if  $t \in S, \varphi_p(\lambda, t) \in S$ . Hence the sequence  $t_i$  is defined for all  $i$  and since  $|t_{i+1} - \rho_1| \leq L |t_i - \rho_1|, L < 1$ , we conclude that  $t_i \rightarrow \rho_1$ .

In Section 6 we examined the asymptotic behavior of the error. We now obtain bounds on the error  $|t_i - \rho_1|$  for all  $i$ . We use the same notation as in the proof of the preceding theorem. By Lemmas 1 and 3 we may assume that  $\lambda$  is sufficiently large so that  $t_i \in S$  for all  $i > 0$ . Hence

$$\frac{|t_{i+1} - \rho_1|}{|t_i - \rho_1|^p} \leq K_1, \quad i = 1, 2, \dots,$$

where

$$K_1 = \frac{2 |\rho_1| DM |\mu_2|^\lambda}{1 - DMR^p |\mu_2|^\lambda}.$$

Hence,

$$|t_{i+1} - \rho_1| \leq K_1^{(p^i - 1)/(p - 1)} R^{p^i},$$

for  $p \geq 2$ .

Let  $\rho_k$  be the closest zero to  $\rho_1$ . We do not assume that  $\rho_k$  is unique. Choose  $R = \frac{1}{2} |\rho_k - \rho_1|$ . Then for all  $t \in S$ , and for  $j = 2, 3, \dots, n, |t - \rho_j| \geq R$ . Hence we may choose

$$(7.2) \quad M^{-1} = R^p.$$

We may assume that  $\lambda$  is sufficiently large so that

$$DMR^p |\mu_2|^\lambda \leq \frac{1}{2},$$

or using (7.2),  $D |\mu_2|^\lambda \leq \frac{1}{2}$ . If  $B = P'$ ,  $D = n - 1$  and the requirement is simply that  $(n - 1)\mu_2^\lambda \leq \frac{1}{2}$ . Then

$$|t_{i+1} - \rho_1| \leq K^{(p^i-1)/(p-1)} R^{p^i}, \quad K = |\rho_1| DM |\mu_2|^\lambda.$$

Hence

$$(7.3) \quad |t_{i+1} - \rho_1| \leq R \left[ \frac{|\rho_1| D |\mu_2|^\lambda}{R} \right]^{(p^i-1)/(p-1)}, \quad \text{for } p \geq 2,$$

$$(7.4) \quad |t_{i+1} - \rho_1| \leq R \left[ \frac{|\rho_1| D |\mu_2|^\lambda}{R} \right]^i, \quad \text{for } p = 1.$$

If  $B = P'$ , we get the very simple bounds,

$$(7.5) \quad |t_{i+1} - \rho_1| \leq \frac{1}{2} |\rho_k - \rho_1| \left[ \frac{2(n-1) |\rho_1| |\rho_2/\rho_1|^\lambda}{|\rho_k - \rho_1|} \right]^{(p^i-1)/(p-1)},$$

for  $p \geq 2$ ,

$$(7.6) \quad |t_{i+1} - \rho_1| \leq \frac{1}{2} |\rho_k - \rho_1| \left[ \frac{2(n-1) |\rho_1| |\rho_2/\rho_1|^\lambda}{|\rho_k - \rho_1|} \right]^i, \quad \text{for } p = 1.$$

**8. The Character of the Convergence.** We investigate the conditions under which the sequence of approximants converges monotonically or alternatingly. Throughout this section we strengthen our standard hypothesis.

$$|\rho_1| > |\rho_2| \geq |\rho_j|, \quad j = 3, 4, \dots, n$$

to

$$|\rho_1| > |\rho_2| > |\rho_j|, \quad j = 3, 4, \dots, n.$$

We assume that  $t_0$ , the initial approximant, is real and not equal to one of the subdominant zeros.

By Lemmas 1 and 3 of Section 7 we may choose  $\lambda$  so large that  $t_1$  lies in an open disk  $S$  centered at  $\rho_1$  which contains no other zeros of  $P$ . By Lemma 5 of Section 7 we may choose  $\lambda$  so large that for all  $t \in S$ ,

$$\frac{|\varphi_p(\lambda, t) - \rho_1|}{|t - \rho_1|} \leq L < 1.$$

Since

$$\varphi_p(\lambda, t) - t = \varphi_p(\lambda, t) - \rho_1 + \rho_1 - t = -(t - \rho_1) \left[ 1 - \frac{\varphi_p(\lambda, t) - \rho_1}{t - \rho_1} \right],$$

we conclude that for  $\lambda$  sufficiently large and all  $t \in S$ ,

$$(8.1) \quad \text{Sgn} [\varphi_p(\lambda, t) - t] = -\text{Sgn} (t - \rho_1).$$

Hence

$$(8.2) \quad \begin{array}{ll} \text{if } t_i > \rho_1, & \text{then } t_{i+1} < t_i; \\ \text{if } t_i < \rho_1, & \text{then } t_{i+1} > t_i. \end{array}$$

From (6.2) it follows that for  $\lambda$  sufficiently large

$$(8.3) \quad \begin{aligned} \text{Sgn } [\varphi_p(\lambda, t) - \rho_1] &= \text{Sgn} \left[ d_2 \left( \frac{\rho_2}{\rho_1} \right)^\lambda (\rho_2 - \rho_1) \left( \frac{t - \rho_1}{t - \rho_2} \right)^p \right], \\ d_2 &= \frac{B(\rho_2) P'(\rho_1)}{B(\rho_1) P'(\rho_2)}. \end{aligned}$$

From (8.2) and (8.3) we can determine the character of the convergence by imposing conditions on the signs of  $\rho_1$  and  $\rho_2$  or on the evenness or oddness of  $p$  and  $\lambda$ . Rather than pursuing all possible cases we content ourselves with the statement of two typical results. *The importance of these results is that the type of convergence may be chosen at pleasure.*

First we assume  $\rho_1 > 0, \rho_2 > 0$ . This is not a serious restriction for we need only calculate a bound on the magnitude of the real zeros by, say, Lagrange's method (Dickson [7], p. 24) and introduce a change of variables that ensures that all real zeros are positive. Using (8.2) and (8.3) it is easy to prove

**THEOREM.** *Let  $\rho_1 > 0, \rho_2 > 0, B(t) = P'(t)$ . Let  $\lambda$  be sufficiently large. Then if  $p$  is even,  $t_i \uparrow \rho_1$ ; if  $p$  is odd,  $t_i$  converges alternately to  $\rho_1$ .*

We turn to an example of a theorem where conditions are imposed on  $p$  and  $\lambda$ . We have

**THEOREM.** *Choose  $p$  and  $\lambda$  even and let  $\lambda$  be sufficiently large. Let  $B(t) = P'(t)$ . Then if  $\rho_1 - \rho_2 > 0, t_i \uparrow \rho_1$ ; if  $\rho_1 - \rho_2 < 0, t_i \downarrow \rho_1$ . Let  $B(t) = 1$ . Then if  $\rho_1 - \rho_2 > 0, t_i \downarrow \rho_1$ ; if  $\rho_1 - \rho_2 < 0, t_i \uparrow \rho_1$ .*

It is also a simple matter to determine the sign of the asymptotic error constant. It follows from (6.4) that for  $\lambda$  sufficiently large,

$$(8.4) \quad \text{Sgn } C_p(\lambda) = -\text{Sgn} \left[ \frac{d_2(\rho_2/\rho_1)^\lambda}{(\rho_1 - \rho_2)^{p-1}} \right].$$

Again we have two types of results. For example, if  $\rho_1 > 0, \rho_2 > 0$ , we have

$$C_p(\lambda, P') < 0, \quad C_p(\lambda, 1) > 0.$$

If  $\lambda$  is even and  $p$  is odd, we have

$$C_p(\lambda, P') < 0, \quad C_p(\lambda, 1) > 0.$$

Finally, we investigate the character of the convergence of  $\tilde{G}_p(\lambda, t)$  to  $[P/(t - \rho_1)]^p$ . From (4.2),

$$\tilde{G}_p(\lambda, t) - \left( \frac{P(t)}{t - \rho_1} \right)^p = d_2 \mu_2^\lambda \left[ \left( \frac{P(t)}{t - \rho_2} \right)^p - \left( \frac{P(t)}{t - \rho_1} \right)^p \right] [1 + e(\lambda, t)]$$

where  $e(\lambda, t) \rightarrow 0$ . Hence

$$\lim_{\lambda \rightarrow \infty} \frac{\tilde{G}_p(\lambda, t) - \left( \frac{P(t)}{t - \rho_1} \right)^p}{\mu_2^\lambda} = d_2 \left[ \left( \frac{P(t)}{t - \rho_2} \right)^p - \left( \frac{P(t)}{t - \rho_1} \right)^p \right].$$

Furthermore we have

**THEOREM.** *If  $\rho_1/\rho_2 > 0$ , the coefficients of  $\tilde{G}_p(\lambda, t)$  converge to the coefficients of  $[P(t)/(t - \rho_1)]^p$  monotonically as  $\lambda \rightarrow \infty$ , while if  $\rho_1/\rho_2 < 0$  they converge alternately.*

**9. Normalized Calculations.** We noted in Section 3 that if any of the zeros of  $P$  have magnitude greater than unity, then the coefficients of  $G(\lambda, t)$  increase without limit. On the other hand, if all the zeros of  $P$  lie within the unit circle,  $G(\lambda, t)$  converges to the zero polynomial. Let  $\bar{h}(t)$  denote the polynomial  $h(t)$  divided by its leading coefficient. We observed that  $\bar{G}(\lambda, t) \rightarrow P(t)/(t - \rho_1)$ . Hence  $\bar{G}(\lambda, t)$  has well-behaved coefficients. As we shall see,  $\bar{G}_p(\lambda, t)$  is just as easy to calculate as  $G_p(\lambda, t)$ .

From

$$G(\lambda + 1, t) = tG(\lambda, t) - \alpha_0(\lambda)P(t)$$

we conclude that

$$(9.1) \quad \begin{aligned} \bar{G}(\lambda + 1, t) &= \overline{t\bar{G}(\lambda, t) - P(t)}, & \text{if } \alpha_0(\lambda) \neq 0, \\ \bar{G}(\lambda + 1, t) &= t\bar{G}(\lambda, t), & \text{if } \alpha_0(\lambda) = 0. \end{aligned}$$

If  $B = 1$ , start with  $\bar{G}(n - 1, t, 1) = \overline{t^n - P(t)}$ . Let

$$(9.2) \quad G(\lambda, t) = \sum_{j=0}^{n-1} \gamma_j(\lambda)t^{n-1-j}.$$

Then

$$(9.3) \quad \gamma_j(\lambda + 1) = \frac{\gamma_{j+1}(\lambda) - a_{j+1}}{\gamma_1(\lambda) - a_1}, \quad j = 1, 2, \dots, n - 1, \quad \gamma_n(\lambda) = 0.$$

On a machine with slow division,  $[\gamma_1(\lambda) - a_1]^{-1}$  should be formed first. Since

$$\gamma_1(\lambda) - a_1 = \frac{\beta(\lambda + 1)}{\beta(\lambda)},$$

the denominator of (9.3) cannot vanish for  $\lambda$  sufficiently large and finite.

Since the leading coefficient of  $G_p(\lambda, t)$  is  $\beta(\lambda)$  for  $p = 1, 2, \dots$ , it follows from (4.7) that

$$(9.4) \quad \bar{G}_p(\lambda, t) = \sum_{k=0}^{p-1} [-P(t)]^{p-1-k} \frac{\bar{G}^{(p-1-k)}(\lambda, t)}{(p - 1 - k)!} V_k(t).$$

In terms of the  $\bar{G}_p(\lambda, t)$ , the iteration function is simply

$$(9.5) \quad \varphi_p(\lambda, t) = t - P(t) \frac{\bar{G}_{p-1}(\lambda, t)}{\bar{G}_p(\lambda, t)}.$$

In particular,

$$(9.6) \quad \varphi_1(\lambda, t) = t - \frac{P(t)}{\bar{G}(\lambda, t)},$$

$$(9.7) \quad \varphi_2(\lambda, t) = t - P(t) \frac{\bar{G}(\lambda, t)}{P'(t)\bar{G}(\lambda, t) - P(t)\bar{G}'(\lambda, t)}.$$

**10. Two Algorithms for the Accelerated Calculation of the  $G$  Polynomials.** The calculation of  $G(2\lambda, t)$  from  $G(\lambda, t)$  using the recursion

$$(10.1) \quad G(\lambda + 1, t) = tG(\lambda, t) - \alpha_0(\lambda)P(t)$$

requires about  $\lambda n$  multiplications. In this section we give two algorithms for calculating  $G(2\lambda, t)$  from  $G(\lambda, t)$  in a number of multiplications which is proportional to  $n^2$ . In Algorithm I,  $G(\lambda + 1, t), \dots, G(\lambda + n - 1, t)$  are formed as intermediate quantities while in Algorithm II we proceed from  $G(\lambda, t)$  to  $G(2\lambda, t)$  without forming any intermediate  $G(\lambda + j, t)$ . In both algorithms the vector of coefficients of  $G(2\lambda, t)$  is given by a matrix-vector multiplication.

We must emphasize that the results of this section hold only for  $B(t) = 1$ . Hence

$$\beta(\lambda) = \Omega(\lambda) = \sum_{i=1}^n \frac{\rho_i^\lambda}{P'(\rho_i)}.$$

Furthermore the normalization of Section 9 cannot be applied here.

We first prove a number of identities which are more general than required for the algorithms. We have for  $l = 0, 1, \dots$ ,

$$(10.2) \quad G(2\lambda + l, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda + l)G(\lambda + n - 1 - j, t),$$

$$(10.3) \quad G(2\lambda + l, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda)G(\lambda + l + n - 1 - j, t),$$

where  $\alpha_j(\lambda)$  is defined by

$$(10.4) \quad G(\lambda, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda)t^{n-1-j}.$$

To prove (10.2) we define

$$h(\lambda, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda + l)G(\lambda + n - 1 - j, t).$$

Then

$$\begin{aligned} h(\lambda, \rho_i) &= \sum_{j=0}^{n-1} \alpha_j(\lambda + l)\rho_i^{\lambda+n-1-j} \\ &= \rho_i^\lambda G(\lambda + l, \rho_i) = \rho_i^{2\lambda+l}. \end{aligned}$$

Since  $h(\lambda, t)$  is a polynomial of degree at most  $n - 1$  and since

$$h(\lambda, \rho_i) = \rho_i^{2\lambda+l} = G(2\lambda + l, \rho_i), \quad i = 1, 2, \dots, n,$$

we conclude  $h(\lambda, t) \equiv G(2\lambda + l, t)$ . This proves (10.2) and the proof of (10.3) may be handled similarly.

Define  $C_j(\lambda)$  by

$$(10.5) \quad G^2(\lambda, t) = \sum_{j=0}^{2n-2} C_j(\lambda)t^{2n-2-j}.$$

Then one may easily show that for  $l = 0, 1, \dots$ ,

$$(10.6) \quad G(2\lambda + 2l, t) = \sum_{j=0}^{2n-2} C_j(\lambda + l)G(2n - 2 - j, t).$$

Algorithm I is based on (10.2) or (10.3) with  $l = 0$ , that is,

$$(10.7) \quad G(2\lambda, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda)G(\lambda + n - 1 - j, t)$$

while Algorithm II is based on (10.6) with  $l = 0$ , that is,

$$(10.8) \quad G(2\lambda, t) = \sum_{j=0}^{2n-2} C_j(\lambda)G(2n - 2 - j, t).$$

We first describe Algorithm I. It follows from (10.4) and (10.7) upon comparison of powers of  $t$  that

$$(10.9) \quad \alpha_j(2\lambda) = \sum_{k=0}^{n-1} \alpha_k(\lambda)\alpha_j(\lambda + n - 1 - k), \quad j = 0, 1, \dots, n - 1.$$

Let  $\mathbf{A}(\lambda)$  be the matrix whose columns are the coefficients of  $G(\lambda + k, t)$ ,  $k = 0, 1, \dots, n - 1$ . That is,

$$\mathbf{A}(\lambda) = \begin{pmatrix} \alpha_0(\lambda) & \alpha_0(\lambda + 1) & \cdots & \alpha_0(\lambda + n - 1) \\ \alpha_1(\lambda) & \alpha_1(\lambda + 1) & \cdots & \alpha_1(\lambda + n - 1) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n-1}(\lambda) & \alpha_{n-1}(\lambda + 1) & \cdots & \alpha_{n-1}(\lambda + n - 1) \end{pmatrix}.$$

Let  $\mathbf{N}(\lambda)$  be the first column of  $\mathbf{A}(\lambda)$  and let  $\mathbf{M}(\lambda)$  be the first column in inverted order. Then

$$(10.10) \quad \mathbf{N}(2\lambda) = \mathbf{A}(\lambda)\mathbf{M}(\lambda).$$

Observe that  $\mathbf{N}(2\lambda)$  is just the vector of coefficients of  $G(2\lambda, t)$  and is also the first column of  $\mathbf{A}(2\lambda)$ . From

$$G(\lambda + k, t) = tG(\lambda + k - 1, t) - \alpha_0(\lambda + k - 1)P(t),$$

it follows that

$$(10.11) \quad \begin{aligned} \alpha_{j-1}(\lambda + k) &= \alpha_j(\lambda + k - 1) - a_j\alpha_0(\lambda + k - 1), \\ k = 1, 2, \dots, n - 1, \quad j &= 1, 2, \dots, n, \quad \alpha_n(\lambda + k - 1) = 0. \end{aligned}$$

This permits the computation of the remaining columns of the new  $\mathbf{A}$  matrix from the first column.

To start the computation, we note that  $G(n, t, 1) = t^n - P$ , or

$$(10.12) \quad \alpha_j(n) = -a_{j+1}, \quad j = 0, 1, \dots, n - 1.$$

Algorithm I may be summarized as follows. Initially compute the first column of  $\mathbf{A}(n)$  by (10.12). After this initialization the remaining columns of  $\mathbf{A}$  are computed by (10.11) while the next first column is computed by (10.12). After  $r$  steps  $G(2^r n, t)$  has been computed.

It is easy to see that the computation of  $G(2\lambda, t)$  from  $G(\lambda, t)$  by Algorithm I requires about  $2n^2$  multiplications as contrasted with the  $\lambda n$  multiplications required if recursion (10.1) is used. Hence if  $\lambda > 2n$ , it is cheaper to use Algorithm I. On the other hand, extra routines and memory are required.

We turn to Algorithm II which is based on (10.8). Since

$$G(j, t, 1) = t^j, \quad j = 0, 1, \dots, n - 1,$$

we have

$$(10.13) \quad \alpha_j(2\lambda) = C_{n-1+j}(\lambda) + \sum_{k=0}^{n-2} \alpha_j(n+k)C_{n-2-k}(\lambda), \quad j = 0, 1, \dots, n - 1.$$



Let  $\mathbf{B}$  be the matrix whose columns are the coefficients of  $G(n + k, t)$ ,  $k = 0, 1, \dots, n - 2$ . That is,  $\mathbf{B}$  is just  $\mathbf{A}(n)$  without its last column. Let  $\mathbf{N}(\lambda)$  be defined as before and let

$$L(\lambda) = \begin{pmatrix} C_{n-2}(\lambda) \\ \vdots \\ C_0(\lambda) \end{pmatrix}, \quad K(\lambda) = \begin{pmatrix} C_{n-1}(\lambda) \\ \vdots \\ C_{2n-2}(\lambda) \end{pmatrix}.$$

Then

$$(10.14) \quad \mathbf{N}(2\lambda) = \mathbf{K}(\lambda) + \mathbf{B}L(\lambda).$$

Observe that  $\mathbf{B}$  is independent of  $\lambda$  and need be formed just once for any problem.

Algorithm II may be summarized as follows. Compute the first column of  $\mathbf{B}$  by (10.12) and the remaining columns by (10.11) with  $\lambda = n$ . Square  $G(n, t) = t^n - P(t)$  and use the coefficients to calculate  $G(2n, t)$  by (10.13). Square this polynomial (which is again of degree  $n - 1$ ) and use the coefficients to calculate  $G(4n, t)$ , etc.

It is easy to see that the computation of  $G(2\lambda, t)$  from  $G(\lambda, t)$  by Algorithm II requires about  $\frac{3}{2} n^2$  multiplications per step, where multiplications by the number 2 have been neglected. The calculation of  $\mathbf{B}$  requires about  $n^2$  multiplications but this need be done only once. Thus Algorithm II requires fewer multiplications than Algorithm I but requires a polynomial squaring routine.

**11. Iteration Functions for Small Values of  $\lambda$ .** By the Theorem of Section 7 we know that for  $\lambda$  sufficiently large and fixed the iteration functions we have studied generate sequences which are guaranteed to converge. These iteration functions are not given by explicit formulas depending on  $P$  and its derivatives. In this section we confine ourselves to values of  $\lambda \leq n$ , and  $B = 1$ . These iteration functions are given by explicit formulas. Naturally these methods will not be globally convergent.

The case  $\lambda = 0$ , which leads to standard methods for solving polynomial and transcendental equations, is considered in Section 12. We turn to the case  $\lambda = \eta$ ,  $0 < \eta \leq n - 1$ . Since  $G(\eta, t, 1) = t^\eta$  we have

$$(11.1) \quad \varphi_2(\eta, t, 1) = t - \frac{tP(t)}{tP'(t) - \eta P(t)},$$

$$(11.2) \quad \varphi_3(\eta, t, 1) = t - \frac{tP(t)[tP'(t) - \eta P(t)]}{t^2T(t) - \eta tP(t)P'(t) + \frac{1}{2}\eta(\eta - 1)P^2(t)},$$

$$T(t) = [P'(t)]^2 - \frac{1}{2}P(t)P''(t), \quad \eta > 1.$$

We would usually take the largest permissible value of  $\eta$ , namely  $\eta = n - 1$ .

We consider the case  $\lambda = n$ . Then  $G(n, t, 1) = t^n - P(t)$  and we have

$$(11.3) \quad \varphi_2(n, t, 1) = t - \frac{P(t)[t^n - P(t)]}{t^{n-1}[tP'(t) - nP(t)]}.$$

Observe the similarity among (11.1), (11.3) and the Newton-Raphson function  $\varphi = t - P(t)/P'(t)$ . Although (11.1) and (11.3) require very little extra computation, they are distinctly superior. This may be seen from the asymptotic

error constant

$$C_p(\lambda, B) = -\sum_{i=2}^n \frac{d_i \mu_i^\lambda}{(\rho_1 - \rho_i)^{p-1}}.$$

Then for the Newton-Raphson function

$$C_2(0, 1) = -\sum_{i=1}^n \frac{P'(\rho_i)}{P'(\rho_i)} \frac{1}{\rho_1 - \rho_i},$$

while for (11.1),

$$C_2(\eta, 1) = -\sum_{i=1}^n \frac{P'(\rho_i)}{P'(\rho_i)} \frac{\mu_i^\eta}{\rho_1 - \rho_i},$$

and for (11.3),

$$C_2(n, 1) = -\sum_{i=1}^n \frac{P'(\rho_i)}{P'(\rho_i)} \frac{\mu_i^n}{\rho_1 - \rho_i}.$$

**12. Relation to Other Methods.** The iteration functions that we have been studying are of the form

$$\begin{aligned} \varphi_p(\lambda, t, B) &= \frac{G_p(\lambda + 1, t, B)}{G_p(\lambda, t, B)} = t - P(t) \frac{G_{p-1}(\lambda, t, B)}{G_p(\lambda, t, B)} \\ (12.1) \quad &= \frac{\sum_{i=1}^n c_i \rho_i^{\lambda+1} / (t - \rho_i)^p}{\sum_{i=1}^n c_i \rho_i^\lambda / (t - \rho_i)^p}, \quad c_i = \frac{B(\rho_i)}{P'(\rho_i)}. \end{aligned}$$

By assigning certain values to  $\lambda, p, t$  we obtain many classical methods for solving equations as special cases of (12.1).

We consider first the case  $B = 1, \lambda = 0$ . From (4.7),

$$(12.2) \quad \varphi_p(0, t, 1) = t + (p - 1) \frac{(1/P)^{(p-2)}}{(1/P)^{(p-1)}}.$$

This is an often-rediscovered sequence of iteration functions. In particular, with  $u = P/P', A_i = P^{(i)}/i!P'$ ,

$$\begin{aligned} \varphi_2(0, t, 1) &= t - u, \\ \varphi_3(0, t, 1) &= t - \frac{u}{1 - A_2 u}, \\ \varphi_4(0, t, 1) &= t - \frac{u(1 - A_2 u)}{1 - 2A_2 u + A_3 u^2}. \end{aligned}$$

From (4.19) it follows that

$$(12.3) \quad G_p(0, t, 1) = \begin{vmatrix} P' & P & \cdots & 0 \\ \frac{1}{2}P'' & P' & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \frac{P^{(p-1)}}{(p-1)!} & \frac{P^{(p-2)}}{(p-2)!} & \cdots & P' \end{vmatrix}.$$

This permits  $\varphi_p(0, t, 1) = t - P(t)G_{p-1}(0, t, 1)/G_p(0, t, 1)$  to be expressed in terms of determinants.

Some insight into these iteration functions can be achieved by observing that

$$\varphi_p(0, t, 1) = \frac{\sum_{i=1}^n \rho_i / P'(\rho_i) (t - \rho_i)^p}{\sum_{i=1}^n 1 / P'(\rho_i) (t - \rho_i)^p}.$$

If  $t$  is sufficiently close to  $\rho_1$ , then all but the first term in the numerator and denominator may be ignored and we have  $\varphi_p(0, t, 1) \sim \rho_1$ . Although these processes are of  $p$ th order, they are guaranteed to converge to  $\rho_1$  only for  $t$  sufficiently close to  $\rho_1$ . The advantage of  $\varphi_p(\lambda, t)$  is that for  $\lambda$  sufficiently large,  $\varphi_p(\lambda, t) \sim \rho_1$  for  $t$  arbitrary.

Next, we observe that

$$(12.4) \quad \varphi_p(\lambda, \infty) = \frac{\sum_{i=1}^n c_i \rho_i^{\lambda+1}}{\sum_{i=1}^n c_i \rho_i^\lambda} = \delta(\lambda),$$

with

$$\delta(\lambda) = \beta(\lambda + 1) / \beta(\lambda).$$

Then  $\delta(\lambda)$  is just the  $\lambda$ th approximant to  $\rho_1$  in the method of Bernoulli. That is, if we generate a sequence  $y(\lambda)$  by

$$P(E)y(\lambda) = 0, \quad y(\mu) = y_\mu, \quad \mu = 0, 1, \dots, n - 1,$$

then  $y(\lambda) \equiv \delta(\lambda)$  for some choice of  $y_0, y_1, \dots, y_{n-1}$ . Hence Bernoulli's method is equivalent to calculating  $\varphi_p(\lambda, \infty)$  for increasing  $\lambda$  and this is a first order process for calculating  $\rho_1$ .

Observe that

$$(12.5) \quad \varphi_p(0, 0) = \delta(-p).$$

If

$$|\rho_n| < |\rho_{n-i}|, \quad i = 1, 2, \dots, n - 1,$$

then

$$\lim_{p \rightarrow \infty} \varphi_p(0, 0) = \rho_n.$$

Observe that  $[\varphi_{p+1}(0, 0)]^{-1}$  is the  $p$ th Bernoulli iterant to  $\rho_n^{-1}$ .

From (12.1) and (12.3)

$$(12.6) \quad \begin{aligned} \varphi_p(0, 0, 1) &= -a_n \frac{G_{p-1}(0, 0, 1)}{G_p(0, 0, 1)}, \\ G_p(0, 0, 1) &= \begin{vmatrix} a_{n-1} & a_n & \cdots & 0 \\ a_{n-2} & a_{n-1} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ a_{n+1-p} & a_{n+2-p} & \cdots & a_{n-1} \end{vmatrix}, \end{aligned}$$

where  $a_{-k} = 0$ ,  $k > 0$ . Hence  $\varphi_p(0, 0, 1)$  gives an approximation to the smallest zero as the ratio of two determinants. This approximation appears due to Furstenau [10], [11]. Whittaker [29] rewrites (12.6) as a series.

We propose to calculate the sequence of polynomials  $G(\lambda, t)$  by

$$G(\lambda + 1, t, B) = tG(\lambda, t, B) - \alpha_0(\lambda)P(t), \quad G(0, t, B) = B(t).$$

However, the  $G(\lambda, t)$  could also be calculated by the recursion  $P(E)G(\lambda, t) = 0$ . Indeed it is easy to prove that for  $G(0, t)$  arbitrary there exist unique polynomials  $y_0(t), y_1(t), \dots, y_{n-1}(t)$  of degree at most  $n - 1$  such that if

$$P(E)y(\lambda, t) = 0, \quad y(\mu, t) = y_\mu(t), \quad \mu = 0, 1, \dots, n - 1,$$

then

$$y(\lambda, t) \equiv G(\lambda, t), \quad \lambda = 0, 1, 2, \dots.$$

Hence the  $G(\lambda, t)$ , which play such a vital part in our work, may be thought of as a Bernoulli sequence whose initial terms depend on  $t$ .

For certain special choices of  $B(t)$ , we can easily calculate the corresponding  $y_\mu(t)$ . If  $B = 1$ , then  $y_\mu(t) = t^\mu$ . If  $B(t) = P'(t)$ , then

$$y_\mu(t) = \sum_{k=0}^{n-1} A_{n-1-k}(t)s(\mu + k),$$

where  $s(r)$  is the Newton sum  $s(r) = \sum_{i=1}^n \rho_i^r$ .

We indicate some of the important papers concerning the solution of polynomial equations which are related to our material. It is remarkable considering the variety of motivating ideas for methods of solving polynomial equations that so many may be formulated in terms of the power series expansion of ratios of polynomials. (Many of the papers concern themselves with the calculation of the zeros of analytic functions rather than polynomials.) The theorem underlying these methods, which is due to König [14], concerns the coefficients of a function which has just one simple pole on its circle of convergence.

In the method of Bernoulli (D. Bernoulli [5], Euler [8], Lagrange [17]) these coefficients are calculated by a recurrence relation. Simpson [23] describes what amounts to Bernoulli's method and anticipates some of the later work. Furstenau [10], [11] expresses the coefficients as the ratios of determinants and Whittaker [29] expresses them as a series. See also Meyer [18], Runge [20], and Cohn [6].

Related to these methods are those which use the coefficients to derive iteration functions. Much of the basic work in this area is due to Schröder [21]. This line of inquiry has been continued by Hamilton [12], Zajta [30], and Kulik [15], [16], among others.

Overall views may be obtained in Householder [13, Chapter 3] and Munro [19]. A perceptive paper by Fry [9] is also recommended.

Sebastião e Silva [22] uses the remainder of the division of  $t^\lambda$  by  $P$  to obtain a factorization of  $P$  and he states what we have called Algorithm II, without proof. His work has been continued by Aparo [1], [2]. In an interesting series of papers, Bauer [3], [4] obtains related results from a different point of view. Sebastião e Silva, Aparo, and Bauer do not construct iteration functions.

There is an extensive literature (Jacobi, Furstenau, Hadamard, Aitken, Se-

bastião e Silva, Rutishauser, Bauer) on the calculation of equimodular and subdominant zeros which we do not review here.

APPENDIX A. NUMERICAL EXAMPLES

The approximation of zeros by the methods of this paper is a two step process. First an iteration function is formed and then a sequence of approximants is calculated. The iteration function is easily constructed on a computer. All that is required are a few subroutines for performing simple manipulations of polynomials such as differentiation, addition, and multiplication. The programming may be done in any of a number of convenient procedure oriented languages. The examples reported here were obtained from a program written in FORTRAN.

The only choices of  $B(t)$  which have been tried are  $B(t) = 1$  and  $B(t) = P'$ . The appropriate choice of  $\lambda$  depends on the ratio of the magnitude of the largest subdominant zeros to the dominant zero. In the following numerical examples  $\lambda$  was chosen as a fairly small integral multiple of  $n$ .

*Example 1.*

$$P(t) = t^3 - 6t^2 + 11t - 6, \quad \rho_1 = 3, \quad \rho_2 = 2, \quad \rho_3 = 1.$$

Use

$$\varphi_1(24, t, 1) = t - \frac{P(t)}{\bar{G}(24, t, 1)}.$$

Then

$$\bar{G}(24, t, 1) = t^2 - 2.99988117950632t + 1.99988117951340.$$

Take  $t_0 = 100,000$ . Then

$$\begin{aligned} t_1 &= 3.00012 \\ t_2 &= 3.000000014 \\ t_3 &= 3.000000000017. \end{aligned}$$

*Example 2.* Same polynomial as above. Use

$$\varphi_1(24, t, P') = t - \frac{P(t)}{\bar{G}(24, t, P')}.$$

Then

$$\bar{G}(24, t, P') = t^2 - 3.00005939967062t + 2.00005939967770.$$

Let  $t_0 = 100,000$ . Then

$$\begin{aligned} t_1 &= 2.999941 \\ t_2 &= 3.0000000035 \\ t_3 &= 2.9999999999979. \end{aligned}$$

*Example 3.*

$$P(t) = t^4 - 46t^3 + 528t^2 - 1090t + 2175,$$

$$\rho_1 = 29, \quad \rho_2 = 15, \quad \rho_{3,4} = 1 \pm 2i.$$

Use

$$\varphi_2(16, t, 1) = t - P(t) \frac{\tilde{G}(16, t, 1)}{\tilde{G}_2(16, t, 1)}.$$

Then

$$\begin{aligned} \tilde{G}(16, t, 1) &= t^3 - 17.0003674589376t^2 + 35.0007349178752t \\ &\quad - 75.0018372946880, \\ \tilde{G}_2(16, t, 1) &= t^6 - 34.0007349178752t^5 + 359.019107864755t^4 \\ &\quad - 1340.07496162327t^3 + 3775.24105306307t^2 \\ &\quad - 5250.34173681196t + 5625.40420483136. \end{aligned}$$

Take  $t_0 = 100,000$ . Then

$$\begin{aligned} t_1 &= 28.99963 \\ t_2 &= 28.9999999999997. \end{aligned}$$

*Example 4.*

$$\begin{aligned} P(t) = \prod_{i=1}^8 (t - i) &= t^8 - 36t^7 + 546t^6 - 4536t^5 + 22449t^4 - 67284t^3 \\ &\quad + 118124t^2 - 109584t + 40320. \end{aligned}$$

Use

$$\varphi_1(32, t, 1) = t - \frac{P(t)}{\tilde{G}(32, t, 1)}.$$

Then

$$\begin{aligned} \tilde{G}(32, t, 1) &= t^7 - 27.8967511565157t^6 + 319.836370519674t^5 \\ &\quad - 1942.00032787128t^4 + 6693.50209993064t^3 - 12965.3556514612t^2 \\ &\quad + 12887.1250096890t - 4966.21074965027. \end{aligned}$$

Let  $t_0 = 100$ . Then

$$\begin{aligned} t_1 &= 8.10 \\ t_2 &= 8.0089 \\ t_3 &= 8.00084 \\ t_4 &= 8.000081 \\ t_5 &= 8.0000077 \\ t_6 &= 8.00000073 \\ t_7 &= 8.000000070 \\ t_8 &= 8.0000000067 \end{aligned}$$

$$t_9 = 8.00000000064$$

$$t_{10} = 8.00000000061.$$

*Example 5.*

$$P(t) = t^3 - 5.01t^2 + 8.03t - 4.02, \quad \rho_1 = 2.01, \quad \rho_2 = 2, \quad \rho_3 = 1.$$

Use

$$\varphi_2(48, t, P') = t - P(t) \frac{\tilde{G}(48, t, P')}{\tilde{G}_2(48, t, P')}.$$

Then

$$\begin{aligned} \tilde{G}(48, t, P') &= t^2 - 3.00440433725533t + 2.00440433725533 \\ \tilde{G}_2(48, t, P') &= t^4 - 6.00880867451066t^3 + 13.0357787414152t^2 \\ &\quad - 12.0441314592984t + 4.01766139239389. \end{aligned}$$

Let  $t_0 = 100,000$ . Then

$$\begin{aligned} t_1 &= 2.0056 \\ t_2 &= 2.0067 \\ t_3 &= 2.0084 \\ t_4 &= 2.0097 \\ t_5 &= 2.0099940 \\ t_6 &= 2.009999972 \\ t_7 &= 2.0099999999993. \end{aligned}$$

APPENDIX B. SUMMARY OF FORMULAS FOR THE CONSTRUCTION OF ITERATION FUNCTIONS

For the convenience of the reader who is primarily interested in using the new iteration functions, we summarize the important formulas.

**1. Unnormalized Calculation (Sections 3, 4, 5).** Let the degree of  $B(t)$  be less than or equal to  $n - 1$  with  $B(\rho_1) \neq 0$ . Let

$$G(0, t, B) = B(t), \quad G(\lambda + 1, t, B) = tG(\lambda, t, B) - \alpha_0(\lambda)P(t),$$

where  $\alpha_0(\lambda)$  is the leading coefficient of  $G(\lambda, t, B)$ . Let

$$G_p(\lambda, t, B) = \sum_{k=0}^{p-1-k} [-P(t)]^{p-1-k} \frac{G^{(p-1-k)}(\lambda, t, B)}{(p-1-k)!} V_k(t),$$

where

$$V_0(t) = 1, \quad V_k(t) = P'(t)V_{k-1}(t) - \frac{P(t)}{k} V'_{k-1}(t).$$

Let

$$\varphi_p(\lambda, t, B) = t - P(t) \frac{G_{p-1}(\lambda, t, B)}{G_p(\lambda, t, B)}.$$

In particular,

$$\varphi_1(\lambda, t) = t - P(t) \frac{\alpha_0(\lambda)}{G(\lambda, t)},$$

$$\varphi_2(\lambda, t) = t - P(t) \frac{G(\lambda, t)}{P'(t)G(\lambda, t) - P(t)G'(\lambda, t)},$$

$$\varphi_3(\lambda, t) = t - P(t) \frac{[P'(t)G(\lambda, t) - P(t)G'(\lambda, t)]}{T(t)G(\lambda, t) - P(t)P'(t)G'(\lambda, t) + \frac{1}{2}P^2(t)G''(\lambda, t)}$$

$$T(t) = [P'(t)]^2 - \frac{1}{2}P(t)P''(t).$$

**2. Normalized Calculation (Section 9).** Let  $\bar{h}(t)$  denote the polynomial  $h(t)$  divided by its leading coefficient. Then

$$\bar{G}(\lambda + 1, t) = \overline{t\bar{G}(\lambda, t) - P(t)}, \quad \text{if } \alpha_0(\lambda) \neq 0,$$

$$\bar{G}(\lambda + 1, t) = t\bar{G}(\lambda, t) \quad \text{if } \alpha_0(\lambda) = 0.$$

$$\bar{G}_p(\lambda, t) = \sum_{k=0}^{p-1-k} [-P(t)]^{p-1-k} \frac{\bar{G}^{(p-1-k)}(\lambda, t)}{(p-1-k)!} V_k(t),$$

$$\varphi_p(\lambda, t) = t - P(t) \frac{\bar{G}_{p-1}(\lambda, t)}{\bar{G}_p(\lambda, t)}.$$

### 3. Accelerated Calculation of $G(\lambda, t, 1)$ (Section 10).

ALGORITHM I.

$$\alpha_j(n) = -a_{j+1}, \quad j = 0, 1, \dots, n-1,$$

$$\alpha_{j-1}(\lambda + k) = \alpha_j(\lambda + k - 1) - a_j\alpha_0(\lambda + k - 1), \quad k = 1, 2, \dots, n-1,$$

$$j = 1, 2, \dots, n, \quad \alpha_n(\lambda + k - 1) = 0,$$

$$\alpha_j(2\lambda) = \sum_{k=0}^{n-1} \alpha_k(\lambda)\alpha_j(\lambda + n - 1 - k), \quad j = 0, 1, \dots, n-1.$$

ALGORITHM II.

$$\alpha_j(n) = -a_{j+1}, \quad j = 0, 1, \dots, n-1,$$

$$\alpha_{j-1}(n + k) = \alpha_j(n + k - 1) - a_j\alpha_0(n + k - 1), \quad k = 1, 2, \dots, n-2,$$

$$j = 1, 2, \dots, n, \quad \alpha_n(n + k - 1) = 0,$$

$$\alpha_j(2\lambda) = C_{n-1+j}(\lambda) + \sum_{k=0}^{n-2} \alpha_j(n + k)C_{n-2-k}(\lambda), \quad j = 0, 1, \dots, n-1,$$

where

$$G^2(\lambda, t) = \sum_{j=0}^{2n-2} C_j(\lambda)t^{2n-2-j}.$$



4. Iteration Functions for Small Values of  $\lambda$  (Section 11).

$$\varphi_2(n - 1, t, 1) = t - \frac{P(t)}{tP'(t) - (n - 1)P(t)},$$

$$\varphi_2(n, t, 1) = t - \frac{P(t)[t^n - P(t)]}{t^{n-1}[tP'(t) - nP(t)]}.$$

APPENDIX C. NOTATION

1.  $P(t) = \sum_{j=0}^n a_{n-j}t^j = \prod_{j=0}^n (t - \rho_j), \quad |\rho_1| > |\rho_j|, \quad j = 2, 3, \dots, n.$

2.  $\frac{B(t)t^\lambda}{P(t)} = \frac{G(\lambda, t, B)}{P(t)} + Q(\lambda, t, B).$

3.  $L_i(t) = \frac{P(t)}{t - \rho_i} \frac{1}{P'(\rho_i)}.$

4.  $Eh(\lambda) = h(\lambda + 1).$

5.  $\beta(\lambda) = \sum_{i=1}^n \frac{B(\rho_i)}{P'(\rho_i)} \rho_i^\lambda, \quad \Omega(\lambda) = \sum_{i=1}^n \frac{\rho_i^\lambda}{P'(\rho_i)}.$

6.  $G(\lambda, t) = \sum_{j=0}^{n-1} \alpha_j(\lambda)t^{n-1-j}.$

7.  $A_k(t) = \sum_{j=0}^k a_{k-j}t^j.$

8.  $\bar{h}(t)$ : The polynomial  $h(t)$  divided by its leading coefficient.

9.  $c_i = B(\rho_i)/P'(\rho_i).$

10.  $G_p(\lambda, t, B) = \frac{(-1)^{p-1}}{(p-1)!} P^p(t) \left[ \frac{G(\lambda, t, B)}{P(t)} \right]^{(p-1)}, \quad G_1 \equiv G, \quad G_0 \equiv \beta(\lambda).$

11.  $V_k(t) = \frac{(-1)^k}{k!} P^{k+1}(t) \left[ \frac{1}{P(t)} \right]^{(k)}.$

12.  $\varphi_p(\lambda, t, B) = \frac{G_p(\lambda + 1, t, B)}{G_p(\lambda, t, B)}.$

13.  $u(t) = P(t)/P'(t).$

14. If

$$\lim_{t \rightarrow \rho_1} \frac{\varphi_r(\lambda, t, B) - \rho_1}{(t - \rho_1)^r} = C_r(\lambda, B),$$

then  $r$  is the order and  $C_r(\lambda, B)$  is the asymptotic error constant.

15.  $d_i = c_i/c_1, \quad \mu_i = \rho_i/\rho_1.$

Bell Telephone Laboratories, Incorporated  
Murray Hill, New Jersey

1. E. APARO, "Applicazione di un nuovo metodo per la risoluzione numerica delle equazioni algebriche," *Bol. Soc. Portuguesa Mat. Sér. A*, v. 1, 1948, pp. 49-57. MR 9, 621.

2. E. APARO, "Un procedimento iterativo per la risoluzione numerica delle equazioni algebriche," *Ricerca Sci.*, v. 24, 1954, pp. 1003-1005. MR 15, 900.
3. F. L. BAUER, "Beiträge zur Entwicklung numerischer Verfahren für programmgesteuerte Rechenanlagen.
- a) I. "Quadratisch konvergente Durchführung der Bernoulli-Jacobischen Methode zur Nullstellenbestimmung von Polynomen," *S.-B. Math.-Nat. Kl. Bayer. Akad. Wiss.*, v. 1954, 1955, pp. 275-303. MR 18, 151.
- b) II. Direkte Faktorisierung eines Polynoms, *S.-B. Math.-Nat. Kl. Bayer. Akad. Wiss.* v. 1956, 1957, pp. 163-203. MR 19, 686.
4. F. L. BAUER, "Das Verfahren der abgekürzten Iteration für algebraische Eigenwertprobleme, insbesondere zur Nullstellenbestimmung eines Polynoms," *Z. Angew. Math. Phys.*, v. 7, 1956, pp. 17-32. MR 18, 154.
5. D. BERNOULLI, "Observationes de Seriebus," *Comment. Acad. Sc. Petrop.*, v. 3, 1728, pp. 85-100.
6. F. COHN, "Über die in recurrirender Weise gebildeten Grössen und ihren Zusammenhang mit den Algebraischen Gleichungen," *Math. Ann.*, v. 44, 1894, pp. 473-538.
7. L. E. DICKSON, *New First Course in the Theory of Equations*, Wiley, New York, 1939. MR 1, 1.
8. L. EULER, "De usu serierum recurrentium in radicibus aequationum indagandis," *Opera Omnia*, Ser. 1, v. 8, pp. 339-361.
9. T. C. FRY, "Some numerical methods for locating roots of polynomials," *Quart. Appl. Math.*, v. 3, 1945, pp. 89-105. MR 7, 83.
10. E. FURSTENAU, *Darstellung der reellen Wurzeln algebraischer Gleichungen durch determinanten der Coefficienten*, Marburg, 1860.
11. E. FURSTENAU, *Neue Methode zur Darstellung und Berechnung der imaginären Wurzeln algebraischen Gleichungen*, Marburg, 1867.
12. H. J. HAMILTON, "A type of variation on Newton's method," *Amer. Math. Monthly*, v. 57, 1950, pp. 517-522. MR 12, 537.
13. A. S. HOUSEHOLDER, *Principles of Numerical Analysis*, McGraw-Hill, New York, 1953. MR 15, 470.
14. J. KÖNIG, "Über eine Eigenschaft der Potenzreihen," *Math. Ann.*, v. 23, 1884, pp. 447-449.
15. S. KULIK, "A method for approximating the zeros of analytic functions," *Duke Math. J.*, v. 24, 1957, pp. 137-141. MR 19, 257.
16. S. KULIK, "On the solution of algebraic equations," *Proc. Amer. Math. Soc.*, v. 10, 1959, pp. 185-192.
17. J. L. LAGRANGE, *De la Résolution des Équations Numériques*, Note 6, Paris, 1798.
18. F. MEYER, "Zur Auflösung der Gleichungen," *Math. Ann.*, v. 33, 1880, pp. 511-524.
19. W. D. MUNRO, "Some iterative methods for determining zeros of functions of a complex variable," *Pacific J. Math.*, v. 9, 1959, pp. 555-566. MR 21 #6089.
20. C. RUNGE, "Entwicklung der Wurzeln einer algebraischen Gleichung in Summen von Rationalen Functionen der Coefficienten," *Acta Math.*, v. 6, 1885, pp. 305-318.
21. E. SCHRÖDER, "Über unendlich viele Algorithmen zur Auflösung der Gleichungen," *Math. Ann.*, v. 2, 1870, pp. 317-365.
22. J. SEBASTIÃO E SILVA, "Sur une méthode d'approximation semblable à celle de Gräffe," *Portugal. Math.*, v. 2, 1941, pp. 271-279. MR 3, 235.
23. T. SIMPSON, *Treatise of Algebra*, 5th Edition, London, 1782.
24. J. F. TRAUB, *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964. MR 29 #6607
25. J. F. TRAUB, "On Lagrange-Hermite interpolation," *J. Soc. Indust. Appl. Math.*, v. 12, 1964, pp. 886-891.
26. J. F. TRAUB, "Generalized sequences with applications to the discrete calculus," *Math. Comp.*, v. 19, 1965, pp. 177-200.
27. J. F. TRAUB, "Solution of linear difference and differential equations," *Bull. Amer. Math. Soc.*, v. 71, 1965, pp. 538-541.
28. J. F. TRAUB, Associated polynomials and the solution of linear problems. (To appear.)
29. E. T. WHITTAKER, "A formula for the solution of algebraic or transcendental equations," *Proc. Edinburgh Math. Soc.*, v. 36, 1918, pp. 103-106.
30. A. ZAJTA, "Untersuchungen über die Verallgemeinerungen der Newton-Raphsonschen Wurzelapproximation. a) I. *Acta Tech. Acad. Sci. Hungar.*, v. 15, 1956, pp. 233-260. MR 18, 415. b) II. *Acta Tech. Acad. Sci. Hungar.*, v. 19, 1957, pp. 25-60. MR 19, 983.