

A Generalization of the Jenkins-Traub Method

By J. A. Ford

Abstract. A class of methods for finding zeros of polynomials is derived which depends upon an arbitrary parameter ρ . The Jenkins-Traub algorithm is a special case, corresponding to the choice $\rho = \infty$. Global convergence is proved for large and small values of ρ and a duality between pairs of members is exhibited. Finally, we show that many members of the class (including the Jenkins-Traub method) converge with R -order at least 2.618 . . . , which improves upon the result obtained by Jenkins and Traub [3].

1. Introduction. A well-known and effective method for computing the zeros of a polynomial has been presented by Jenkins and Traub [3]. Their algorithm consists of three stages, the first of which simply serves to accentuate the smaller zeros. The second stage implements a fixed-shift process which isolates one of the zeros, while the third stage may be viewed as a variable-shift process which converges very quickly to the zero. The third stage may also be viewed as a Newton-Raphson iteration performed on a sequence of rational functions. Another interpretation of the third stage arises from showing that it is equivalent to a generalized Rayleigh iteration applied to the companion matrix of the polynomial. Jenkins and Traub [3] have, in a detailed analysis of their method, shown it to be globally convergent and, furthermore, that the convergence to a zero is faster than second-order.

In this paper, we shall define and investigate a class of methods for the solution of polynomials which contains the Jenkins-Traub method as a special case. We shall analyze both the global and local convergence of members of the class and, in particular, obtain an improved result concerning the local convergence of the Jenkins-Traub method.

We consider a polynomial $P(z)$, where

$$(1.1) \quad P(z) \equiv \sum_{i=0}^n a_i z^i, \quad a_0 a_n \neq 0.$$

(Note the reversal of the order of the coefficients from that used by Jenkins and Traub.) We denote the *distinct* zeros of P by $\{\rho_i\}_{i=1}^j$ and their respective multiplicities by $\{m_i\}_{i=1}^j$, so that

$$(1.2) \quad P(z) \equiv a_n \prod_{i=1}^j (z - \rho_i)^{m_i}, \quad \sum_{i=1}^j m_i = n.$$

Finally, we denote the deflated polynomials obtained from P by

$$(1.3) \quad P_i(z) \equiv P(z)/(z - \rho_i), \quad i = 1(1)j.$$

Received June 18, 1975; revised May 10, 1976.

AMS (MOS) subject classifications (1970). Primary 65H05.

Copyright © 1977, American Mathematical Society

2. The Class of Methods. The third stage of the Jenkins-Traub method consists of the following combined iteration:

$$(2.1) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z - s_\lambda)} \left\{ H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s_\lambda)}{P(s_\lambda)} P(z) \right\},$$

$$(2.2) \quad s_{\lambda+1} = s_\lambda + \frac{H^{(\lambda)}(s_\lambda)}{H^{(\lambda+1)}(s_\lambda)}.$$

The polynomials $\{H^{(\lambda)}(z)\}$ are all of degree $n - 1$. We generalize (2.1) by including a polynomial $q^{(\lambda)}(z)$:

$$(2.3) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z - s_\lambda)} \left\{ q^{(\lambda)}(z)H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s_\lambda)}{P(s_\lambda)} P(z) \right\}.$$

We retain the requirement that each polynomial $H^{(\lambda)}(z)$ should have degree $n - 1$, so that the polynomial $q^{(\lambda)}(z)$ must have degree 1. Also, since $H^{(\lambda+1)}(z)$ is to be a polynomial, we have

$$(2.4) \quad q^{(\lambda)}(s_\lambda) = 1.$$

It follows that $q^{(\lambda)}(z)$ may be written in the form

$$(2.5) \quad q^{(\lambda)}(z) \equiv (z - \rho)/(s_\lambda - \rho),$$

where ρ is an arbitrary parameter. From Eqs. (2.1) and (2.3), we see that the Jenkins-Traub method corresponds to the choice $\rho = \infty$. To derive the generalized form of the iteration (2.2), we assume that the sequence $\{s_\lambda\}$ is converging to the zero ρ_i , say. From Eqs. (2.3) and (2.5), we have

$$(2.6) \quad H^{(\lambda+1)}(\rho_i) = \frac{(\rho_i - \rho)}{(\rho_i - s_\lambda)} \cdot \frac{1}{(s_\lambda - \rho)} \cdot H^{(\lambda)}(\rho_i),$$

and hence

$$(2.7) \quad \rho_i = s_\lambda + \frac{H^{(\lambda)}(\rho_i)}{H^{(\lambda+1)}(\rho_i) - H^{(\lambda)}(\rho_i)(s_\lambda - \rho)^{-1}}.$$

Therefore, since $s_\lambda \rightarrow \rho_i$, we obtain the iteration

$$(2.8) \quad s_{\lambda+1} = s_\lambda + \frac{H^{(\lambda)}(s_\lambda)}{H^{(\lambda+1)}(s_\lambda) - H^{(\lambda)}(s_\lambda)(s_\lambda - \rho)^{-1}}.$$

Again, the Jenkins-Traub iteration (2.2) corresponds to $\rho = \infty$. We remark, at this stage, that a more rigorous derivation of the iterations (2.3) and (2.8) will be described in Section 5, where the relationship of these schemes to generalized Rayleigh iteration is discussed. We also note that the iterative scheme (2.8) may be derived as the result of a Newton-Raphson process (see Theorem 4, Section 6).

We are now able to write down the complete generalized three-stage method.

Stage One (no shift).

$$(2.9) \quad H^{(0)}(z) \equiv P'(z),$$

$$(2.10) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{z} \left\{ \frac{(z - \rho)}{(-\rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(0)}{P(0)} P(z) \right\}.$$

(2.10) is applied for $\lambda = 0(1)M - 1$, where M is an appropriate integer (Jenkins and Traub [3] choose $M = 5$).

Stage Two (fixed shift).

$$(2.11) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z - s)} \left\{ \frac{(z - \rho)}{(s - \rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s)}{P(s)} P(z) \right\}, \quad \lambda = M(1)L - 2,$$

where L will depend on the progress of the method.

Stage Three (variable shift).

$$(2.12) \quad s_{L-1} = s,$$

$$(2.13) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z - s_\lambda)} \left\{ \frac{(z - \rho)}{(s_\lambda - \rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s_\lambda)}{P(s_\lambda)} P(z) \right\},$$

$$(2.14) \quad s_{\lambda+1} = s_\lambda + \frac{H^{(\lambda)}(s_\lambda)}{H^{(\lambda+1)}(s_\lambda) - H^{(\lambda)}(s_\lambda)(s_\lambda - \rho)^{-1}},$$

(2.13) and (2.14) to be repeated for $\lambda = L - 1, L, \dots$ until s_λ converges to a zero of P .

Remarks. (i) The choice of the derivative polynomial $P'(z)$ to start the process ensures, as Jenkins and Traub [3] point out, that all the polynomials $H^{(\lambda)}$ are linear combinations of the polynomials $\{P_i\}_{i=1}^j$.

(ii) In Stage One, we have given what might be described as the direct generalization of the corresponding Jenkins-Traub process, in that (2.10) is a no-shift process. The effect, in their case, is to accentuate the smaller zeros of P . However, for general ρ , it is not clear that a no-shift process is necessarily best; and we might also consider an infinite-shift process:

$$(2.15) \quad H^{(\lambda+1)}(z) \equiv (z - \rho)H^{(\lambda)}(z) - \frac{h_{n-1}^{(\lambda)}}{a_n} P(z),$$

where $h_{n-1}^{(\lambda)}$ is the coefficient of z^{n-1} in $H^{(\lambda)}$. We shall return to this point in Section 4, but we point out here, for future reference, that Stage One is not necessary as far as the convergence theory is concerned.

(iii) Jenkins and Traub [3] describe a process for choosing the fixed shift s . We do not prescribe such a method, since the choice of s will depend to an extent upon the value of ρ .

(iv) The iteration in Stage Three is normally terminated when the computed value of $P(s_\lambda)$ is less than a "running" error bound (Peters and Wilkinson [7]).

3. Global Convergence. We commence our discussion of the convergence of the class by giving sufficient conditions for the variable-shift iteration to converge. We have already indicated that all the polynomials $\{H^{(\lambda)}\}$ may be written in the form

$$(3.1) \quad H^{(\lambda)}(z) \equiv \sum_{i=1}^j c_i^{(\lambda)} P_i(z).$$

We define the quantities

$$(3.2) \quad \phi = \max_{1 \leq i \leq j} \frac{|\rho_i - \rho|}{|\rho_1 - \rho|}, \quad \alpha = (1 + \phi)^{-1}, \quad R = \min_{2 \leq i \leq j} |\rho_1 - \rho_i|.$$

THEOREM 1. *Assume that*

- (i) $|s_L - \rho_1| < \alpha R$;
- (ii) $c_1^{(L)} \neq 0$;
- (iii) $\beta_L < 1$, where $\beta_\lambda = |s_\lambda - \rho_1|/|\rho - \rho_1|$, $\lambda = L, L + 1, \dots$;
- (iv) $D_L (= \sum_{i=2}^j |c_i^{(L)}|/|c_1^{(L)}|) < \phi(1 - \beta_L)(3 + \beta_L)^{-1}$.

Then $s_\lambda \rightarrow \rho_1$ as $\lambda \rightarrow \infty$.

Proof. The proofs of this theorem and the next are basically the same as those given by Jenkins and Traub [3] for the case $\rho = \infty$ and are therefore omitted.

We now assume that the zeros of P are renumbered, if necessary, so that

$$(3.3) \quad \frac{|\rho_1 - s|}{|\rho_1 - \rho|} = \min_{1 \leq i \leq j} \frac{|\rho_i - s|}{|\rho_i - \rho|},$$

where $s = s_{L-1}$ is the fixed shift of Stage Two. We can now give the first of our two main results on global convergence.

THEOREM 2. *Suppose that the minimum in (3.3) is achieved only at ρ_1 , that is,*

$$(3.4) \quad \frac{|\rho_1 - s|}{|\rho_1 - \rho|} < \frac{|\rho_i - s|}{|\rho_i - \rho|}, \quad i = 2(1)j.$$

Define $\beta = |\rho_1 - s|/|\rho_1 - \rho|$ and assume that

$$(3.5) \quad \beta < 1.$$

Then, for all L sufficiently large and fixed, $s_\lambda \rightarrow \rho_1$.

Remarks. (i) For the case studied by Jenkins and Traub [3], we see that hypothesis (3.5) is trivially satisfied. In this case, (3.4) is to be interpreted as

$$(3.6) \quad |\rho_1 - s| < |\rho_i - s|, \quad i = 2(1)j;$$

and Theorem 2 then reduces to the global convergence result obtained by Jenkins and Traub. Condition (3.6) requires that s be closest to just one of the zeros of P , a condition that is highly likely to be satisfied in practice. In the same way, it is highly likely in the general case that condition (3.4) will be satisfied.

(ii) Condition (3.5) is a *sufficient* condition for global convergence, but it is not *necessary*. Nevertheless, we can easily guarantee that (3.5) is satisfied by choosing ρ sufficiently large. If ω is an upper bound on the zeros of P , that is,

$$(3.7) \quad |\rho_i| \leq \omega, \quad i = 1(1)j,$$

then condition (3.5) will certainly be satisfied if

$$(3.8) \quad |\rho| > |s| + 2\omega.$$

(iii) The proof of Theorem 1 is accomplished by showing that $|s_{\lambda+1} - \rho_1| \leq \tau |s_\lambda - \rho_1|$, where $\tau < 1$, for $\lambda = L - 1, L, \dots$. This fact, coupled with (3.5), guarantees that $s_\lambda \neq \rho$ for $\lambda = L - 1, L, \dots$, so that the iterations (2.11), (2.13) and (2.14) are well defined.

4. Dual Methods. So far, our development has, to a large degree, followed that of Jenkins and Traub [3]. In particular, we have generalized their proof of global convergence to apply to all cases for which ρ is sufficiently large. We now wish to obtain a corresponding result for small values of ρ ; and, to do so, we adopt a different approach. Essentially, we shall establish an equivalence between a three-stage iteration with a given polynomial and a given value of ρ and another three-stage iteration with the *inverse* of the given polynomial and the *reciprocal* of ρ . In this way, we may demonstrate a duality between pairs of members of the class and, thus, infer global convergence for small values of ρ .

We shall denote the inverse polynomial of a polynomial $Q(z)$ by $Q_I(z)$, so that, if Q has exact degree p , then

$$(4.1) \quad Q_I(z) \equiv z^p Q(z^{-1}).$$

The following three lemmas are all easily established by induction.

LEMMA 1. Consider the following pair of iterations:

$$(4.2) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{z} \left\{ \frac{(z - \rho)}{(-\rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(0)}{P(0)} P(z) \right\};$$

$$(4.3) \quad G^{(\lambda+1)}(z) \equiv (z - \rho^{-1}) G^{(\lambda)}(z) - \frac{G_I^{(\lambda)}(0)}{P(0)} P_I(z);$$

both iterations to be repeated for $\lambda = 0(1)M - 1$. Assume that

$$G^{(0)}(z) \equiv k_1 H_I^{(0)}(z)$$

for some constant k_1 . Then, for $\lambda = 0(1)M$,

$$G^{(\lambda)}(z) \equiv k_1 H_I^{(\lambda)}(z).$$

(The iteration (4.3) will be recognized as the infinite-shift iteration referred to in Section 2.)

LEMMA 2. Consider the following pair of iterations:

$$(4.4) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z - s)} \left\{ \frac{(z - \rho)}{(s - \rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s)}{P(s)} P(z) \right\};$$

$$(4.5) \quad G^{(\lambda+1)}(z) \equiv \frac{1}{(z-t)} \left\{ \frac{(z-\rho^{-1})}{(t-\rho^{-1})} G^{(\lambda)}(z) - \frac{G^{(\lambda)}(t)}{P_I(t)} P_I(z) \right\};$$

both iterations to be repeated for $\lambda = M(1)L - 2$. Assume that $st = 1$ and that

$$(4.6) \quad G^{(M)}(z) \equiv k_2 H_I^{(M)}(z)$$

for some constant k_2 . Then, for $\lambda = M(1)L - 1$,

$$(4.7) \quad G^{(\lambda)}(z) \equiv (-s^2)^{\lambda-M} k_2 H_I^{(\lambda)}(z).$$

LEMMA 3. Consider the following pair of iterations, (I) and (II):

$$(4.8) \quad s_{L-1} = s$$

$$(4.9) \quad H^{(\lambda+1)}(z) \equiv \frac{1}{(z-s_\lambda)} \left\{ \frac{(z-\rho)}{(s_\lambda-\rho)} H^{(\lambda)}(z) - \frac{H^{(\lambda)}(s_\lambda)}{P(s_\lambda)} P(z) \right\}$$

$$(4.10) \quad s_{\lambda+1} = s_\lambda + \frac{H^{(\lambda)}(s_\lambda)}{H^{(\lambda+1)}(s_\lambda) - H^{(\lambda)}(s_\lambda)(s_\lambda - \rho)^{-1}}$$

} (I);

$$(4.11) \quad t_{L-1} = t$$

$$(4.12) \quad G^{(\lambda+1)}(z) \equiv \frac{1}{(z-t_\lambda)} \left\{ \frac{(z-\rho^{-1})}{(t_\lambda-\rho^{-1})} G^{(\lambda)}(z) - \frac{G^{(\lambda)}(t_\lambda)}{P_I(t_\lambda)} P_I(z) \right\}$$

$$(4.13) \quad t_{\lambda+1} = t_\lambda + \frac{G^{(\lambda)}(t_\lambda)}{G^{(\lambda+1)}(t_\lambda) - G^{(\lambda)}(t_\lambda)(t_\lambda - \rho^{-1})^{-1}}$$

} (II);

the iterations (4.9), (4.10) and (4.12), (4.13) to be repeated for $\lambda = L - 1, L, \dots$. Assume that $st = 1$ and, for some constant k_3 ,

$$(4.14) \quad G^{(L-1)}(z) \equiv k_3 H_I^{(L-1)}(z).$$

Then

$$(4.15) \quad G^{(\lambda)}(z) \equiv \left\{ \prod_{i=0}^{\lambda-L} (-s_{L+i-1}^2) \right\} k_3 H_I^{(\lambda)}(z) \quad \text{and} \quad s_\lambda t_\lambda = 1$$

for $\lambda = L - 1, L, \dots$

Taken together, Lemmas 1, 2 and 3 show that, under the appropriate initial conditions, two three-stage iterations using reciprocal values ρ and ρ^{-1} , reciprocal fixed shifts s and s^{-1} , and operating respectively on a polynomial and its inverse are essentially equivalent. We shall regard such a pair of iterations as *dual* methods. In order to complete the duality, we abandon the requirement, for the purposes of this section only, that a Stage One iteration be commenced with the derivative polynomial. Instead, we simply require that the initial polynomial $H^{(0)}(z)$ be a linear combination of $\{P_i(z)\}_{i=1}^j$:

$$(4.16) \quad H^{(0)}(z) \equiv \sum_{i=1}^j c_i^{(0)} P_i(z) \equiv \sum_{i=1}^j c_i^{(0)} P(z)/(z - \rho_i),$$

with the added restriction that

$$(4.17) \quad c_1^{(0)} \neq 0.$$

In consequence of Lemma 1 (taking $k_1 = 1$, for convenience), the dual iteration must commence with the polynomial

$$(4.18) \quad G^{(0)}(z) \equiv H_I^{(0)}(z) \equiv \sum_{i=1}^j \{-c_i^{(0)} \rho_i^{-1}\} P_i(z)/(z - \rho_i^{-1}).$$

We shall denote a three-stage iteration applied to $P(z)$ with initial polynomial $H^{(0)}(z)$, zero or infinite initial shift ϵ , fixed shift s and “ ρ -value” ρ by $(H^{(0)}, \epsilon, s, \rho, P)$. The dual iteration is then $(H_I^{(0)}, \epsilon^{-1}, s^{-1}, \rho^{-1}, P_I)$.

THEOREM 3. *Consider the primal iteration $(H^{(0)}, \epsilon, s, \rho, P)$, where $\epsilon = 0$ or ∞ , and (4.16), (4.17) are satisfied. Assume that (3.4) is true and that*

$$(4.19) \quad \beta (= |\rho_1 - s|/|\rho_1 - \rho|) < \max \{1, |s|/|\rho|\}.$$

Then, for all L sufficiently large and fixed, $s_\lambda \rightarrow \rho_1$ and, in the dual iteration, $t_\lambda \rightarrow \rho_1^{-1}$.

Proof. We have remarked previously that the first stage may be ignored as far as convergence is concerned, so that the value of ϵ is immaterial. If $\beta < 1$, the result follows by a proof practically identical to that of Theorem 2. On the other hand, if $\beta \geq 1$, consider the dual iteration. This will converge if

$$|\rho_1^{-1} - s^{-1}|/|\rho_1^{-1} - \rho^{-1}| < 1,$$

that is, if

$$|\rho_1 - s|/|\rho_1 - \rho| < |s|/|\rho|,$$

and the proof is complete.

It is clear from Theorem 3 that we have achieved the aim stated at the beginning of this section, namely to prove convergence of the iteration for small values of ρ . Let ν be a lower bound for the zeros of P :

$$(4.20) \quad \nu \leq |\rho_i|, \quad i = 1(1)j.$$

Then, corresponding to the inequality (3.8), we may easily show that the iteration will converge if

$$(4.21) \quad |\rho| < \nu|s| \{2|s| + \nu\}^{-1}.$$

We note that the bounds ν and ω are easily computed (Marden [5]).

It is interesting to analyze further the behavior of these methods for particular choices of the shift s . If we propose to employ a large value of ρ in the process, we might well choose to follow the lead of Jenkins and Traub in selecting s such that $|s| = \nu$. With such a choice of s , it is not difficult to show that

$$(4.22) \quad |\rho_1| < 7|\rho_i|, \quad i = 2(1)j,$$

so that the zero obtained, ρ_1 , is at worst only a modest multiple of the smallest zero. For the case $\rho = \infty$, (4.22) can be strengthened to

$$(4.23) \quad |\rho_1| < 3|\rho_i|, \quad i = 2(1)j,$$

a result already obtained by Jenkins and Traub [3].

Correspondingly, if we wish to use a small value of ρ , it would seem to be sensible to choose s such that $|s| = \omega$. We may then show that

$$(4.24) \quad |\rho_1| > |\rho_i|/7, \quad i = 2(1)j,$$

so that the zero obtained is not much smaller than the largest zero of P . In particular, for the case $\rho = 0$ (the *dual* Jenkins-Traub method), we may sharpen (4.24) to obtain

$$(4.25) \quad |\rho_1| > |\rho_i|/3, \quad i = 2(1)j.$$

5. Generalized Rayleigh Iteration. As we have already observed, Jenkins and Traub [3] have shown that the final stage of their iteration is equivalent to a generalized Rayleigh iteration applied to the companion matrix of the polynomial P . It is therefore appropriate to ask whether the generalized third-stage represented by the iterations (2.13) and (2.14) may be represented in a similar fashion. The answer to this question is in the affirmative. The schemes (2.13) and (2.14) may be obtained by applying a generalized Rayleigh iteration to the matrix C defined by Broyden and Ford [1] in their work on polynomial deflation. The details of the derivation are beyond the scope of this paper, but they may be found in Ford [2], together with some of the proofs that have been omitted from this paper.

6. Local Convergence. So far, we have concerned ourselves with establishing the global convergence of members of the class. We now turn to consider the local convergence; and we begin by showing, as Jenkins and Traub [3] have done for their method, that the third-stage iteration (2.14) for a zero of P may be regarded as a Newton-Raphson iteration. The proof is again omitted.

THEOREM 4. *The iteration (2.14) is precisely the same as the Newton-Raphson iteration*

$$(6.1) \quad s_{\lambda+1} = s_{\lambda} - W^{(\lambda)}(s_{\lambda})/\{W^{(\lambda)}(s_{\lambda})\}',$$

where

$$(6.2) \quad W^{(\lambda)}(z) \equiv P(z)/H^{(\lambda)}(z).$$

Since we can easily show that $H^{(\lambda)}(z)$ is tending to a multiple of $P_1(z)$, $W^{(\lambda)}(z)$ must tend to a multiple of the linear polynomial $(z - \rho_1)$. This observation affords us some insight into why the method is so effective.

We shall now examine the rate of convergence of these methods. Jenkins and Traub [3] have shown that the convergence of their method is faster than second-order.

We shall show here that all those members of the class which we have proved to be globally convergent in fact converge with R -order at least $\zeta^2 = 1 + \zeta = 2.618\dots$, where $\zeta = (1 + \sqrt{5})/2$, the Golden Section. Let

$$(6.3) \quad \tau = 2D_L / \{\phi(1 - \beta_L) - D_L(1 + \beta_L)\}.$$

It follows from hypotheses (iii) and (iv) of Theorem 1 that $0 \leq \tau < 1$. We define, for any nonnegative integer N ,

$$(6.4) \quad C_N(\lambda) = \frac{|s_{L+\lambda+1} - \rho_1|}{|s_{L+\lambda} - \rho_1| \prod_{t=\lambda-N}^{\lambda} |s_{L+t} - \rho_1|} \quad \text{for } \lambda \geq N.$$

The first of our two results concerning rate of convergence is an improved version of a theorem given by Jenkins and Traub [3] which covers the case $N = 0$. The proof of Theorem 5 may be found in Ford [2].

THEOREM 5. *Let the hypotheses of Theorem 1 hold. Then, for any nonnegative integer N and any integer $\lambda \geq N$,*

$$(6.5) \quad C_N(\lambda) < (\alpha R)^{-N-1} \tau^{(\lambda-N)(\lambda-N-1)/2}.$$

We deduce from Theorem 5 that, for any fixed value of N , $C_N(\lambda) \rightarrow 0$. A lengthy but straightforward argument then yields our main result:

THEOREM 6. *Let the hypotheses of Theorem 1 hold. Then the R -order of the sequence $\{s_{L+\lambda}\}$ is at least $1 + \zeta$.*

(For a treatment of the concept of R -order, see Ortega and Rheinboldt [6].) In order to accomplish the proof of Theorem 6, it is necessary to choose N so large that the unique positive root ξ_N of the indicial equation

$$t^{N+1} = 2t^N + \sum_{i=0}^{N-1} t^i$$

is sufficiently close to $1 + \zeta$. It might be objected that this would require an unreasonably large number of iterations to be performed in order to achieve anything approaching the rate of convergence indicated by Theorem 6. However, a simple calculation shows that $\xi_2 > 2.5$ and $\xi_4 > 2.6$, so that the convergence indicated by Theorem 6 sets in very quickly.

In order to put this result into perspective, it is instructive to compare these methods with, say, the secant method which has order ζ . One iteration of the Jenkins-Traub variable-shift iteration is equivalent to approximately five polynomial evaluations and, hence, we should compare this with five iterations of the secant method, each of which requires one polynomial evaluation. This "five-fold" secant method converges with order $\zeta^5 \simeq 11.1$, which may be compared with the result of Theorem 6. For a general member of the new class of methods, the difference is even greater, since a valid comparison is then with the "six-fold" secant method, which converges with order $\zeta^6 \simeq 17.9$. The redeeming advantages of the Jenkins-Traub-type methods are, first, the guarantee of global convergence (if ρ is chosen suitably) and, second, that Theorem 6

applies irrespective of the multiplicity of the zero ρ_1 . This contrasts with the secant method, which converges only linearly to a multiple zero. A possible modification, therefore, to improve the efficiency of an algorithm based on a Jenkins-Traub-type method is to switch at some stage from the variable-shift iteration to the secant method, if there are good grounds for believing that the current zero being computed, ρ_1 , is a *simple* zero of P .

7. Summary and Conclusions. We have derived a class of methods which may be regarded as a generalization of the Jenkins-Traub process and which have similar properties. We have proved global convergence for many members of the class and have shown that a form of duality exists between pairs of members of the class. An examination of the local convergence properties has shown that many members of the class, including the Jenkins-Traub method, converge with R -order at least 2.618 . . . , a result which improves upon the rate of convergence obtained by Jenkins and Traub [3].

As to the question of which member of the class to use, the evidence would seem to point to the use, in general, of the Jenkins-Traub method or its dual. The only situation which we can envisage at present in which another member of the class might be preferred is that in which an approximate zero of P is known a priori, and we wish to determine the remaining zeros of P accurately. By choosing ρ equal to the approximate zero, we effectively prevent convergence to that zero, in general, and thus facilitate the computation of the other zeros (compare Eq. (3.3)). (Note, however, that the method may fail in this case if condition (3.5) is not satisfied.) However, there seems, at present, to be little justification for any substantial modification along these lines to programs based on the Jenkins-Traub method, with one exception, which we discuss below. We have seen that convergence is guaranteed for the Jenkins-Traub method (except when a very unfortunate choice of s is made) and the process for computing the sequence of polynomials $\{H^{(\lambda)}\}$ is computationally simpler than in the general case. The one modification that might be considered is that of applying the Jenkins-Traub method and its dual alternately, thus locating, in turn, a relatively small and then a relatively large zero. This would help towards alleviating the problem of deflation instability referred to by Jenkins and Traub [4]. Of course, this modification may also be considered as the Jenkins-Traub method applied to P , say, to produce a zero ρ_1 , and then the *same* method applied to the *inverse* polynomial of $P_1(z) \equiv P(z)/(z - \rho_1)$.

A possibility that remains to be considered is that of varying ρ during the process in order to improve the convergence. It is clear from an analysis of the class of methods that the optimal value ρ_{opt} of ρ for iteration $L + \lambda + 1$ is given by the value of ρ which minimizes

$$\Gamma(\rho) = \max_{2 \leq i \leq j} \left\{ \frac{|\rho_i - \rho|}{|\rho_1 - \rho|} \cdot \frac{|s_{L+\lambda} - \rho_1|}{|s_{L+\lambda} - \rho_i|} \right\}.$$

However, in general, we will have little or no information about the zeros $\{\rho_i\}_{i=2}^j$ and only an approximation to the zero ρ_1 . It remains to be seen whether this or another criterion for choosing ρ at each iteration will be effective and, indeed, whether varying ρ in such a manner is worthwhile.

8. Acknowledgements. I am particularly grateful to C. G. Broyden for several helpful discussions and for his comments on an earlier draft of this paper. I also wish to acknowledge several helpful comments by the referee.

Computing Centre
University of Essex
Wivenhoe Park,
Colchester, Essex, United Kingdom

1. C. G. BROYDEN & J. A. FORD, "A new method of polynomial deflation," *J. Inst. Math. Appl.*, v. 16, 1975, pp. 271–281.

2. J. A. FORD, "A generalization of the Jenkins-Traub method," Technical Report CSM-9, Univ. of Essex Computing Centre, June 1975.

3. M. A. JENKINS & J. F. TRAUB, "A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration," *Numer. Math.*, v. 14, 1969/70, pp. 252–263 MR 41 #2918.

4. M. A. JENKINS & J. F. TRAUB, "A three-stage algorithm for real polynomials using quadratic iteration," *SIAM J. Numer. Anal.*, v. 7, 1970, pp. 545–566. MR 43 #5716.

5. M. MARDEN, *The Geometry of the Zeros of a Polynomial in a Complex Variable*, Math. Surveys, no. 3, Amer. Math. Soc., Providence, R. I., 1949. MR 11, 101.

6. J. M. ORTEGA & W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970. MR 42 #8686.

7. G. PETERS & J. H. WILKINSON, "Practical problems arising in the solution of polynomial equations," *J. Inst. Math. Appl.*, v. 8, 1971, pp. 16–35. MR 45 #7980.