PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 97, Number 4, August 1986

NEWTON'S METHOD AND THE JENKINS-TRAUB ALGORITHM

R. N. PEDERSON

ABSTRACT. In this paper we propose to show how a multi-increment version of Newton's method can be used to obtain starting points for the Jenkins-Traub algorithm.

1. The Jenkins-Traub algorithm. Let

$$P(z) = \prod_{i=1}^{j} (z - z_j)^{\alpha_j}, \qquad \sum \alpha_j = n, \tag{1.0}$$

where the z_i are distinct roots of multiplicity α_j. The single stage algorithm of Jenkins and Traub starts by creating a sequence of monic polynomials $P^{(\nu)}(z)$ with

$$P^{(1)}(z) = \frac{1}{n}\frac{d}{dz}P(z), \tag{1.1}$$

and for $\nu \geqslant 2$

$$P^{\nu}(z) = \left(P(z) - \frac{P(a)}{P^{\nu-1}(a)}P^{\nu-1}(z)\right) \cdot \frac{1}{(z-a)} \tag{1.2}$$

where a is an arbitrary complex number. They show that if

$$|z_1 - a| < |z_j - a|, \qquad j \geqslant 2, \tag{1.3}$$

then

$$\lim_{\nu \to \infty} \frac{P(z)}{P^{\nu}(z)} = z - z_1. \tag{1.4}$$

However good the probability of an arbitrary complex number satisfying (1.3), it is destined to certain failure when P has real coefficients, a is real and the nearest root is nonreal. Jenkins and Traub overcome this difficulty by constructing an alternative sequence which converges to the quadratic factor

$$(z - z_1)(z - z_2) \tag{1.5}$$

under the condition

$$|z - z_1| = |z - z_2| < |z - z_j|, \qquad j \geqslant 3. \tag{1.6}$$

In this paper we show that the same objective can be realized by saving the data from the two previous iterations.

Received by the editors February 27, 1985 and, in revised form, July 24, 1985.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 30C10, 30-04; Secondary 30C15, 30C80.

THEOREM 1. *If P defined by* (1.0) *satisfies* (1.6), *then with*

$$Q_\nu(z) = \frac{P^{\nu-1}(a)P^\nu(z)P^{\nu-2}(z) - P^{\nu-2}(a)(P_{\nu-1}(z))^2}{P_{\nu-1}(a) - P_{\nu-2}(a)} \tag{1.7}$$

we have

$$\lim_{\nu\to\infty} \frac{[P(z)]^2}{Q_\nu(z)} = (z - z_1)(z - z_2). \tag{1.8}$$

PROOF. Let

$$F^\nu(z) = \left[F^{\nu-1}(z) - \frac{F^{\nu-1}(a)P(z)}{P(a)}\right] \cdot \frac{1}{(z-a)} \tag{1.9}$$

with initial condition

$$F^{(1)}(z) = \frac{1}{n}\frac{d}{dz}(P(z)). \tag{1.10}$$

It follows from (1.1), (1.2), (1.9) and (1.10) and the fact that P is monic that

$$P^\nu(z) = -\frac{P(a)}{F_{\nu-1}(a)}F^\nu(z) \tag{1.11}$$

and, conversely, that

$$F^\nu(z) = -\frac{F^{\nu-1}(a)}{P(a)}P^\nu(z). \tag{1.2}$$

Now, as Jenkins and Traub observe, using the Lagrange interpolation polynomials

$$P_k(z) = P(z)/(z - z_k), \tag{1.13}$$

we have the following expression for the F^ν's:

$$F^\nu(z) = \sum_{i=1}^{j} \frac{\alpha_i P_i(z)}{(z_i - a)^{\nu-1}}, \qquad \nu = 1, 2, \ldots. \tag{1.14}$$

Now let us define

$$\overline{Q}_\nu(z) = F_\nu(z)F_{\nu-2}(z) - [F_{\nu-1}(z)]^2 \tag{1.15}$$

and the corresponding monic polynomial

$$Q_\nu(z) = \frac{\overline{Q}_\nu(z)}{\text{coeff of } z^{2n-2} \text{ in } \overline{Q}_\nu}. \tag{1.16}$$

The function $\overline{Q}_\nu(z)$ is a symmetric quadratic form with zeros on the diagonal. Hence after substituting (1.14) into (1.15) and symmetrizing, we have

$$\overline{Q}_\nu(z) = \sum_{i<j} \frac{\alpha_i\alpha_j P_i(z)P_j(z)}{(a - z_i)^{\nu+1}(a - z_j)^{\nu+1}}(z_i - z_j)^2. \tag{1.17}$$

Since the P_i's are monic, it follows that the coefficient of z^{2n-2} is

$$\sum_{i<j} \frac{\alpha_i\alpha_j(z_i - z_j)^2}{(a - z_i)^{\nu+1}(a - z_j)^{\nu+1}}. \tag{1.18}$$

It now is a consequence of (1.3) that

$$\lim_{z\to\infty} \frac{F^2(z)}{Q_\nu(z)} = (z - z_1)(z - z_2). \tag{1.19}$$

To complete the proof, we have only to use (1.11) and (1.12) to show that Q_ν has the alternative expression (1.7).

2. Starting points. In this section we show that if we take a starting point z outside a circle containing the roots of a polynomial in its interior and multiply the Newton iteration by n, we get into the interior of the circle.

THEOREM 2. *If the roots of P,* (1.0), *satisfy*

$$|z_i - \alpha| < R, \tag{2.1}$$

then with $|z| > R$

$$\beta = z - nf(z)/f'(z) \tag{2.2}$$

satisfies

$$|\beta - \alpha| < R. \tag{2.3}$$

PROOF. We may without loss of generality assume that

$$\alpha = 0 \quad \text{and} \quad R = 1. \tag{2.4}$$

Since

$$\frac{f'(z)}{f(z)} = \sum_{i=1}^{j} \frac{\alpha_i}{z - z_i}, \tag{2.5}$$

with

$$g(z) = z - nf(z)/f'(z), \tag{2.6}$$

we have

$$g(z) = \sum_{i=1}^{n} \frac{\alpha_i z_i}{z - z_i} \Big/ \sum_{i=1}^{j} \frac{\alpha_i}{z - z_i}. \tag{2.7}$$

For a fixed z, $|z| > 1$, $g(z)$ is an analytic function of the complex variables z_i in the polydisc

$$|z_i| < 1, \qquad i = 1, 2, \ldots, n. \tag{2.8}$$

Hence by the maximum principle we may assume that

$$|z_i| = 1, \qquad i = 1, 2, \ldots, n. \tag{2.9}$$

Now, by the theorem of Gauss and Lucas, the zeros of the derivative $P'(z)$ are in the convex hull of

$$z_1, \ldots, z_n \tag{2.10}$$

and so are in a compact subset of the interior of the unit circle. Hence for each fixed set (2.10) satisfying (2.9), g is analytic and bounded for $|z| \geqslant 1$. It therefore assumes its maximum for $|z| = 1$ or at ∞. But

$$\lim_{z\to\infty} g(z) = \sum_{i=1}^{n} \frac{\alpha_i z_i}{n} \tag{2.11}$$

which is a convex combination of the z_i's and so has average less than 1. For $|z| = 1$, we may write the denominator in (2.7) as

$$\sum_{i=1}^{j} \frac{\alpha_i(\bar{z} - \bar{z}_i)}{|z - z_i|^2} = \frac{1}{z} \overline{\sum_{i=1}^{n} \frac{\alpha_j(1 - \bar{z}z)}{|z - z_i|^2}} \tag{2.12}$$

and the numerator as

$$\sum_{i=1}^{j} \frac{\alpha_i z_i(\bar{z} - \bar{z}_i)}{|z - z_i|^2} = -\sum_{i=1}^{n} \frac{\alpha_i(1 - \bar{z}z_i)}{|z - z_i|^2}. \tag{2.13}$$

Hence $|g(z)| = 1$ for $|z| = 1$. This completes the proof.

3. Automatic deflation. It follows from (1.4) that

$$\lim_{\nu \to \infty} P^\nu(z) = \frac{P(z)}{z - z_1} = P_1(z), \tag{3.1}$$

the Lagrange interpolation polynomial corresponding to the factor $z - z_1$. Hence if we use z_1 as starting point and compute a root of P, we may be certain that we do not compute z_1 again when it is simple. If it is multiple, we shall compute z_1 again and by iterating find its multiplicity. Moreover it is clear from (1.14) that if we keep a a reasonable distance from a multiple root, this method of deflation is more accurate than applying the division algorithm. Since all we ask of the initial approximation is that it be closer to the root it approximates than any other root, this procedure appears to have good prospects.

BIBLIOGRAPHY

1. M. A. Jenkins, *Three-stage variable-shift iterations for the solution of polynomial equations with a posteriori error bounds for the zeros*, Dissertation, Stanford Univ., Stanford, Calif., 1969. Available as Rep. CS 138, Computer Science Department, Stanford Univ.

2. M. A. Jenkins and J. F. Traub, *A three-stage variable shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration*, Numer. Math. **14** (1970), 252–261.

3. J. F. Traub, *A class of globally convergent iteration functions for the solution of polynomial equations*, Math. Comp. **20** (1966), 113–138.

4. ______, *Proof of global convergence of an iterative method for calculating complex zeros of a polynomial*, Notices Amer. Math. Soc. **13** (1966), 117.

5. ______, *The calculation of zeros of polynomials and analytic functions*, Mathematical Aspects of Computer Science, Proc. Sympos. Appl. Math., vol. 19, Amer. Math. Soc., Providence, R.I., 1967, pp. 138–152. Also available as Rep. CS 36, Computer Science Department, Stanford Univ., Stanford, Calif., 1966.

6. M. A. Jenkins and J. F. Traub, *A three stage algorithm for real polynomials using quadratic iteration*, Siam J. Numer. Anal. **7** (1970).

DEPARTMENT OF MATHEMATICS, CARNEGIE - MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA 15213