

# Iteration Methods for Finding all Zeros of a Polynomial Simultaneously

By Oliver Aberth

**Abstract.** Durand and Kerner independently have proposed a quadratically convergent iteration method for finding all zeros of a polynomial simultaneously. Here, a new derivation of their iteration equation is given, and a second, cubically convergent iteration method is proposed. A relatively simple procedure for choosing the initial approximations is described, which is applicable to either method.

**1. Introduction.** Let  $P(z)$  be a monic polynomial of degree  $n$  with real or complex coefficients:

$$P(z) = z^n + c_1 z^{n-1} + \cdots + c_{n-1} z + c_n.$$

E. Durand [1, pp. 277–280] and I. O. Kerner [2] independently have proposed an iteration method for finding all the zeros of  $P(z)$  simultaneously. Starting with  $n$  crude approximations to the zeros, their method proceeds by the repeated refinement of all approximations in a uniform manner, and the order of convergence is quadratic.

Here, we give an alternate way of deriving their iteration equation, and propose a second method for finding all the zeros of  $P(z)$ , with a cubic order of convergence. Both methods were computer-tested on a variety of polynomials, with degrees up to 20. When the initial approximations were chosen as described below, failure of convergence was never observed, although convergence toward a multiple zero was slow. In a later section, it is shown that if the zeros of the polynomial and initial approximations both display certain symmetries, then the methods may fail. However, with proper precautions these cases occur with quite small probability.

**2. The Iteration Method of Durand and Kerner.** Let  $z_i$ ,  $i = 1, \dots, n$ , be approximations to the zeros of  $P(z)$  just before beginning a cycle of iteration, and  $z_i + \Delta z_i$ ,  $i = 1, \dots, n$ , be the approximations after completion of the cycle. We wish to choose  $z_i + \Delta z_i$  closer to the zeros. Ideally, we would like to have

$$(1) \quad \prod_{i=1}^n (z - [z_i + \Delta z_i]) = P(z),$$

for then  $z_i + \Delta z_i$  would equal the zeros. Expanding the left side of (1) in powers of  $\Delta z$ , we obtain

---

Received April 21, 1972.

AMS (MOS) subject classifications (1970). Primary 65H05, 65D99; Secondary 30A06.

Key words and phrases. Polynomial zeros, iteration algorithm.

Copyright © 1973, American Mathematical Society

$$(2) \quad \prod_{k=1}^n (z - z_k) - \sum_{i=1}^n \Delta z_i \prod_{k=1; k \neq i}^n (z - z_k) + \sum_{i,j=1; i < j}^n \Delta z_i \Delta z_j \prod_{k=1; k \neq i, j}^n (z - z_k) - \cdots = P(z).$$

To obtain an iteration equation, we drop all terms in powers of  $\Delta z$  higher than the first:

$$\prod_{k=1}^n (z - z_k) - \sum_{i=1}^n \Delta z_i \prod_{k=1; k \neq i}^n (z - z_k) = P(z).$$

This polynomial equation is easy to solve for  $\Delta z_i$ . Set  $z$  equal successively to  $z_1, \dots, z_n$ , and obtain

$$-\Delta z_i \prod_{k=1; k \neq i}^n (z_i - z_k) = P(z_i), \quad i = 1, \dots, n.$$

Finally, if we define  $Q(z)$  to be the polynomial given by

$$(3) \quad Q(z) = \prod_{k=1}^n (z - z_k),$$

we obtain the equation

$$(4) \quad \Delta z_i = -P(z_i)/Q'(z_i), \quad i = 1, \dots, n.$$

If  $P(z)$  has the zeros  $w_i, i = 1, \dots, n$ , so that

$$(5) \quad P(z) = \prod_{k=1}^n (z - w_k),$$

then we may rewrite (4) in the form

$$(6) \quad \Delta z_i = \prod_{k=1}^n (w_k - z_i) / \prod_{k=1; k \neq i}^n (z_k - z_i).$$

**3. Second Iteration Method.** As a preliminary to the derivation of the second iteration equation, we obtain the formula of Newton's method in an unusual way.

We utilize the identification of complex numbers with vectors in the  $z$ -plane, whereby the vector originating at the point  $z'$  and terminating at  $z''$  is assigned the complex number  $z'' - z'$ . The expression

$$(z - w_0)/|z - w_0|^2 = 1/\overline{(z - w_0)}$$

then defines a vector field such that the vector at the point  $z$  is directed away from the point  $w_0$ , and has magnitude inversely proportional to the distance from  $w_0$ . Because of the obvious analogy from electrostatics, this field may be called the field of a unit plus charge at the point  $w_0$  (cf. [3, pp. 7-9]). If unit plus charges are situated at the  $n$  points  $w_1, w_2, \dots, w_n$ , then the resulting vector field is

$$(7) \quad \sum_{i=1}^n 1/\overline{(z - w_i)} = \overline{(P'(z)/P(z))},$$

where  $P(z)$  is given by (5). When several of the numbers  $w_i$  are identical, say  $w_1 =$

$w_2 = \cdots = w_m$ , so that  $P(z)$  has a zero of multiplicity  $m$  at  $w_1$ , several of the terms on the left side of (7) are equal, and it is as if there are  $m$  units of charge at the point  $w_1$ .

Given a polynomial  $P(z)$ , we may attempt to locate a zero by sampling the field defined by the right side of (7) at some point  $z_i$ , and by then finding the point where a single unit charge would be located if it were causing this field. (This amounts to assuming that there is only one term on the left side of (7), a reasonable approximation when  $z_i$  is near a simple zero.) Sampling at this new point, the cycle then could be repeated. Calling the new point  $z_i + \Delta z_i$ , after taking conjugates, we have the equation

$$1/(z_i - (z_i + \Delta z_i)) = P'(z_i)/P(z_i),$$

which leads to the formula of Newton's method

$$(8) \quad \Delta z_i = -P(z_i)/P'(z_i).$$

Now, if we tried to locate all zeros of  $P(z)$  by simultaneously applying (8) to  $n$  different sampling points  $z_1, \cdots, z_n$ , we would fail, likely as not, since several points  $z_i$  could easily converge to the same (simple) zero of  $P(z)$ . To avoid this, we assign a unit minus charge at each sampling point. The idea here is that when a sampling point  $z_i$  is near a simple zero, the field from the minus charge at  $z_i$  should counteract that from the plus charge at the zero, preventing a second sampling point from converging to this zero. After taking conjugates, our iteration equation for the  $i$ th sampling point now is

$$\frac{1}{z_i - (z_i + \Delta z_i)} = \frac{P'(z_i)}{P(z_i)} + \sum_{k=1; k \neq i}^n \frac{-1}{z_i - z_k},$$

so that

$$(9) \quad \Delta z_i = P(z_i) / \left( P(z_i) \sum_{k=1; k \neq i}^n \frac{1}{z_i - z_k} - P'(z_i) \right).$$

In terms of the polynomial  $Q(z)$  (cf. (3)), the sum appearing in (9) may be written as  $Q''(z_i)/2Q'(z_i)$ , and we obtain

$$(10) \quad \Delta z_i = \frac{P(z_i)Q'(z_i)}{\frac{1}{2}P(z_i)Q''(z_i) - P'(z_i)Q'(z_i)}.$$

A purely algebraic derivation of the iteration equation (9) may be arrived at in the following manner. Referring to the Durand-Kerner formula (4), set  $R_i(z) = -P(z)/\prod_{k \neq i} (z - z_k)$ ,  $i = 1, \cdots, n$ . When Newton's method is applied to the rational function  $R_i(z)$  at the point  $z_i$ ,

$$\Delta z_i = -R_i(z_i)/R'_i(z_i), \quad i = 1, \cdots, n,$$

Eq. (9) is obtained.

**4. Choosing the Initial Approximations.** Two circumstances that interfere with convergence come to mind. The requirement that the approximations be distinct is necessary for either (6) or (9) to be meaningful. However, it is not impossible for two neighboring approximations to move to the same point after a cycle of iteration, especially on the earlier cycles when  $|\Delta z_i|$  is generally larger. Nevertheless, this must

be an unlikely occurrence, and, in any case, provisions to detect and correct an accident like this can be incorporated in the computer program.

The other difficulty arises when the zeros of  $P(z)$  are symmetrically positioned in the complex plane with respect to some line  $L$ , and the initial approximations happen to be chosen so they also are symmetrically placed with respect to this line. In this case, some of the approximations may fail to converge to a zero. For the iteration method of Durand and Kerner, Eq. (6) for  $\Delta z_i$  implies that the successive approximations will preserve their symmetry with respect to the line  $L$ . That this is also true for the second method is apparent from its field definition. Hence, when the number of the approximations on  $L$  exceeds the number of zeros there (counting multiplicities), the surplus approximations cannot leave  $L$  and must oscillate along this line. Similarly, if fewer approximations than zeros of odd multiplicity are on  $L$ , then since approximations not on  $L$  must approach  $L$  symmetrically, there must be some zeros on  $L$  which do not have approximations converge toward them. Of course, here small perturbations of the approximations caused by limited precision, rounding errors and the like, may act to destroy the symmetry and induce convergence.

The most frequently encountered line of symmetry for the zeros of  $P(z)$  is, no doubt, the real axis, since this occurs whenever all coefficients of  $P(z)$  are real. Accordingly, we avoided selecting the initial approximations so they displayed this symmetry. Our procedure for choosing them so they were reasonably close to the zeros of  $P(z)$  was as follows: Set  $z$  equal to  $w - c_1/n$ , where  $c_1$  is the coefficient of  $z^{n-1}$  of  $P(z)$ . Then

$$P(z) = \hat{P}(w) = w^n + c'_2 w^{n-2} + \cdots + c'_{n-1} w + c'_n,$$

and there is no term in  $w^{n-1}$ . Assuming not all coefficients  $c'_k$  are zero, if  $r$  is the positive zero of the polynomial

$$S(w) = w^n - |c'_2| w^{n-2} - \cdots - |c'_{n-1}| w - |c'_n|,$$

then all zeros of  $\hat{P}(w)$  are inside or on the circle  $|w| = r$  [3, pp. 122–123], and so the zeros of  $P(z)$  lie inside or on the circle  $|z + c_1/n| = r$ . The computer was programmed to compute  $r_0$ , a rough estimate of  $r$ , and to take  $n$  equally spaced points on the circle  $|z + c_1/n| = r_0$  as the initial approximations. Thus, our initial approximations were

$$z_k = -c_1/n + r_0 e^{i((2\pi/n)(k-1) + \alpha_0)}, \quad k = 1, \cdots, n,$$

with the constant  $\alpha_0$  taken as  $\pi/2n$  so that symmetry with respect to the real axis cannot occur. Note that  $-c_1/n$  equals the mean of the zeros of  $P(z)$ .

From these starting positions on the bounding circle, the observed general behavior of the successive approximations was about the same for both of the iteration methods described. Typically, the successive approximations moved inward in the direction of a nearby zero. Often, two approximations approached the same simple zero, with one moving off on later cycles toward a “free” zero. A multiple zero eventually attracted as many distinct approximations as its multiplicity, and in the complex plane the approximations tended to assume symmetrical positions around the zero, advancing slowly toward it.

As an example of the computations, suppose  $P(z)$  is the polynomial

$$z^5 - 10z^4 + 43z^3 - 104z^2 + 150z - 100,$$

with zeros  $1 \pm 2i$ ,  $2, 3 \pm i$ . Setting  $z$  equal to  $w + 2$ ,  $\hat{P}(w)$  is then  $w^5 + 3w^3 - 6w^2 + 10w$ . The polynomial  $S(w)$  is  $w^5 - 3w^3 - 6w^2 - 10w$ , and testing the value of this polynomial at the successive positive integers  $1, 2, 3, \dots$ , we find  $S(2) < 0$ ,  $S(3) > 0$ . Taking  $r_0$  as 3, the initial approximations are

$$z_k = 2 + 3e^{i((2\pi/5)(k-1) + \pi/10)}, \quad k = 1, 2, 3, 4, 5.$$

With these initial approximations and high-precision computation, the first iteration method gives approximations whose real and imaginary parts are within  $10^{-10}$  of the corresponding parts of a zero after nine iteration cycles. To attain the same accuracy, the second iteration method requires only six iteration cycles.

**5. A Bound on the Error of the Approximation.** In a recent article [4], Brian T. Smith gives a useful bound on the error of the approximations  $z_i$  to the zeros of a polynomial  $P(z)$ . He shows (Corollary 1) that if the circular regions  $\Gamma_i$  are defined by

$$\Gamma_i: |z - z_i| \leq n |P(z_i)/Q'(z_i)|, \quad i = 1, 2, \dots, n,$$

then the union of these regions contains all the zeros of  $P(z)$ . Any connected component consisting of just  $k$  circles contains exactly  $k$  zeros of  $P(z)$ . For the method of Durand and Kerner, a slightly superior definition of the regions  $\Gamma_i$  may be obtained from the first equation on p. 664 of Smith's article:

$$\Gamma_i: |z - (z_i + \Delta z_i)| \leq (n - 1) |\Delta z_i|, \quad i = 1, 2, \dots, n,$$

with  $\Delta z_i$  given by (4).

If  $\epsilon$  equals the maximum of the radii of the circles  $\Gamma_i$ , and these circles do not overlap, then each  $\Gamma_i$  centerpoint is a zero approximation with error  $\leq \epsilon$ . If the overlap of the circles is not investigated, the more modest bound  $(2n - 1)\epsilon$  can be used.

**6. Orders of Convergence.** If we define the point  $Z$  as  $(z_1, z_2, \dots, z_n)$ , both iteration methods may be written as

$$z'_i = z_i + \Delta z_i = F_i(Z), \quad i = 1, 2, \dots, n,$$

for appropriate definitions of the functions  $F_i$ . When all zeros  $w_i$  are distinct, these functions may be expanded in a Taylor series about the point  $W = (w_1, w_2, \dots, w_n)$ , obtaining

$$F_i(Z) = w_i + \sum_{j=1}^n \frac{\partial F_i(W)}{\partial z_j} (z_j - w_j) + \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 F_i(W)}{\partial z_j \partial z_k} (z_j - w_j)(z_k - w_k) + \dots$$

Previously, Kerner showed that for  $F_i$  defined from (4), all first partials are zero, but not all second partials, so the successive iterates converge quadratically for an initial approximation point  $Z$  sufficiently close to  $W$ . Similarly, for  $F_i$  defined from (10), all first and second partials are zero but not all third partials, so, in this case, the successive iterates converge cubically. Moreover, convergence is assured if the point  $Z$  is sufficiently close to the point  $W$ .

To estimate the slower rate of convergence toward a multiple zero, we make the following assumptions:  $P(z)$  has a single multiple zero, which, for convenience, we take to be at the origin. Approximations converging toward a simple zero have

reached their zero, and approximations converging toward the origin have reached positions which are evenly spaced on some circle  $|z| = \rho$ . These assumptions amount to setting  $P(z) = z^m$ , where  $m$  is the multiplicity of the zero at the origin, and  $Q(z) = z^m - \rho^m e^{i\beta}$ . Referring to (4) and (10), we find that the approximations converging toward the origin will move directly toward this zero on each cycle, always reducing their distance away by the fraction  $1/m$  for the first method, and  $2/(m + 1)$  for the second.

**Acknowledgment.** I am grateful to a referee for the purely algebraic derivation of the iteration equation (9). I also wish to express my appreciation to A. S. Householder for supplying the reference to Kerner's article, and to A. H. Stroud for the reference to Durand's book.

Mathematics Department  
Texas A&M University  
College Station, Texas 77843

1. E. DURAND, *Solutions Numériques des Équations Algébriques*. Tome I: *Equations du Type  $F(x) = 0$ ; Racines d'un Polynôme*, Masson, Paris, 1960. MR 22 #12714.
2. I. O. KERNER, "Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen," *Numer. Math.*, v. 8, 1966, pp. 290–294. MR 34 #3778.
3. M. MARDEN, *Geometry of Polynomials*, 2nd ed., Math. Surveys, no. 3, Amer. Math. Soc., Providence, R.I., 1966. MR 37 #1562.
4. B. T. SMITH, "Error bounds for zeros of a polynomial based upon Gerschgorin's theorems," *J. Assoc. Comput. Mach.*, v. 17, 1970, pp. 661–674. MR 43 #5719.