SLAC-PUB-172
February 1966

AN IMPROVEMENT TO ITERATIVE METHODS OF POLYNOMIAL FACTORIZATION*

John Bingham
Stanford Linear Accelerator Center, Stanford, California

1. Introduction

Methods of polynomial factorization are essentially of two differnet types, viz:

(a) Those which converge to all zeros (or fairly simply solved functions of them) simultaneously; e.g., the methods of Aitken-Bernoulli [1] and Graeffe [1] and the Q-D algorithm [2] ,

(b) Those which converge to one zero (or a pair of complex conjugate zeros) at a time; e.g., Newton's method and its quadratic equivalent, Bairstow's [1, 3, 4], Laguerre's [1] , D'Alembert's [5] and Lehmer's [6] methods. It is not in general possible by any simple means to find enough approximations which are good enough that each and every factor can be found by iteration by any of these methods on the original polynomial. Therefore it is common to divide each factor as soon as it is found to the maximum accuracy into the dividend polynomial and to continue with the quotient.

Wilkinson [7,8] has shown that if the zeros are found in order of increasing magnitude or if all the zeros are of more or less equal magnitude, then the division introduces very little extra error and all the zeros of the polynomial can be found with nearly the full accuracy permitted by the number of working digits and the original conditioning of the zeros. However, if neither of these conditions is satisfied, then the removal of a relatively large zero (or pair of complex conjugate zeros) may seriously change the remaining smaller zeros.

---

*Work supported by the U. S. Atomic Energy Commission.

Submitted to Communications of the ACM.

Methods of type (a) are generally slower and much more susceptible to the propagation of round-off errors than those of type (b) and if they are used at all it is generally for only a few iteration cycles to find an approximation to each of the factors which can then be used in a method of the type (b).

Of the methods of type (b), those due to Newton and Laguerre are potentially the fastest but their performance is much more sensitive to the first approximation that is used for each factor. Much ingenuity has been used in devising first approximations and criteria for stopping an iteration and starting again if convergence seems unlikely but it would seem that the best that can be expected for all polynomials is that the Newtonian and Laguerre methods will converge fairly quickly to some zero(s) and that very often this zero(s) will have the smallest magnitude.

Kahan* has used a method which is essentially a quadratic version of the Laguerre method; he also uses a fairly elaborate subroutine to calculate upper and lower bounds for the magnitude of the smallest zero(s) at each stage. This is probably the most sophisticated and powerful method yet programmed but, reportedly, even it does not always find all the zeros in order of increasing magnitude.

It would seem that D'Alembert's and Lehmer's methods, albeit slower, are more likely (though still not certain) to find the zeros in order of increasing magnitude and therefore to avoid the deflation error discussed by Wilkinson [6,7].

Wilkinson has suggested that, having found all the factors by successive iteration and deflation, each should be purified in the original polynomial. This will result in accurate zeros in the great majority of cases but it is tedious and not completely foolproof; there are some polynomials for which the errors introduced by removing the zeros in the "wrong" order so perturb remaining small ill-conditioned zeros that two or more of these inaccurate zeros will converge to the same zero when "purified" in the original polynomial.

---

*Private communication.

It is apparent that the weakest step in all of these methods is the division of the polynomial by a factor and that this weakness most adversely affects the Newton and Laguerre methods which are potentially the fastest. This paper will show how this weakness can be eliminated.

2. Division by a Linear Factor

For clarity, division by a linear factor will be considered first in some detail and afterwards it will be shown that the results can be applied almost equally well to quadratic factors.

If a polynomial

$$P(x) \triangleq \sum_{i=0}^{n} p_i \, x^{n-i}$$

is divided by a factor $(x + x_1)$ to form a quotient

$$Q(x) \triangleq \sum_{i=0}^{n-1} q_i \, x^{n-i-1}$$

there are two different sequences of calculations for the coefficients q_i, viz:

$$ {}_d q_o = p_o $$

$$ {}_d q_i = p_i - x_1 \, {}_d q_{(i-1)} \quad \text{for } i = 1 \text{ to } (n-1) \qquad \ldots\ldots (1) $$

by division in order of descending (subscript d) powers of x and

$$ {}_a q_{(n-1)} = p_n / x, $$

$$ {}_a q_i = \left(p_{(i+1)} - {}_a q_{(i+1)} \right) / x_1 \quad \text{for } i = (n-2) \text{ to } 1 \ldots\ldots (2) $$

by division in order of ascending (subscript a) powers.

If $(x + x_1)$ is not an exact factor, then ${}_dq_i \neq {}_aq_i$ and there will be a remainder term.

In addition to the two quotients which are obtained by dividing exclusively in order of ascending or descending powers ($Q_o(x)$ and $Q_n(x)$ respectively), n-1 composite quotients can be defined by

$$Q_j(x) \triangleq \sum_{i=0}^{j-1} {}_dq_i \, x^{n-i-1} + \sum_{i=j}^{n-1} {}_aq_i \, x^{n-i-1} \quad \text{for } j = 1 \text{ to } (n-1) \quad \ldots . \qquad (3)$$

and then the remainder term is defined by

$$P(x) \equiv (x + x_1) \; Q_j(x) + r_j \, x^{n-j} \quad .$$

Under these circumstances, the zeros of each $Q_j(x)$ will be different from those of $P(x)$ but they will be the same (except for that at $x = -x_1$) as those of

$$P_j(x) \equiv (x + x_1) \; Q_j(x) = P(x) - r_j \, x^{n-j} \quad .$$

Thus, if there is some value of j such that

$$\left| \frac{r_j}{p_j} \right| < 2^{-t}$$

then the polynomial $P_j(x)$ is almost indistinguishable from $P(x)$ when each coefficient is represented by t binary digits.* In practice this cannot generally be achieved but it has been found that for polynomials with widely separated zeros (those for which the conventional removal of a large zero is most damaging)

*Sometimes, due to rounding errors, the other "reconstituted" coefficients of $P_j(x)$ will not be exactly equal to $P(x)$ but the error can only be of the order of 2^{-t}.

$\left| r_j/p_j \right|$ has a sharp minimum at some value of j and this minimum is of the order of magnitude* of 2^{-t}.

Thus, if it is not possible to find a polynomial $Q_j(x)$ such that $(x + x_1)\, Q_j(x)$ is indistinguishable from $P(x)$, then the least average perturbation of the remaining zeros is probably achieved when the relative difference between the coefficients of $P(x)$ and $P_j(x)$ is minimized. Now,

$$r_j x^{n-j} = P(x) - (x + x_1) \sum_{i=0}^{j-1} d^{q_i} x^{n-i-1} - (x + x_1) \sum_{i=j}^{n} a^{q_i} x^{n-i-1}$$

$$= x^{n-j} \left(p_j - x_1\, d^{q_{(j-1)}} - a^{q_j} \right)$$

so that

$$E_j \triangleq \frac{r_j}{p_j} = 1 - \left[\frac{x_1\, d^{q_{(j-1)}} + a^{q_j}}{p_j} \right] \qquad \ldots . \quad (4)$$

---

*An upper bound for $|r_j/p_j|$ can be found as follows: Wilkinson has shown that if the factor $(x + x_1)$ has been found to the maximum accuracy possible, then the remainder

$$r_j x_1^{n-j} \sim \sum_{i=0}^{n} p_i\, e_i\, x_1^{n-i}$$

where, $|e_i| < 2^{-t}$.

Hence, $$\frac{r_j}{p_j} \sim \sum_{i=0}^{n} e_i \frac{p_i}{p_j} x_1^{j-i}$$

and if j be such that x_1^{j}/p_j is a minimum over all j, then

$$\frac{p_i}{p_j} x_1^{j-i} < 1 \quad \text{for all } i \neq j .$$

Therefore, $$\left| \frac{r_j}{p_j} \right| < \left| \sum_{i=0}^{n} e_i \right| < n 2^{-t} .$$

and it is suggested that j be such that the magnitude of the right hand side of (4) be minimized. If any p_j is zero, then, regardless of the value of $(x_1\, {}_d q_{(j-1)} + {}_a q_j)$, E_j should be treated as a very large number and the search for a minimum continued.

It is possible in some cases, however, for the "best" quotient, $Q_j(x)$,--judged by some external criterion--to correspond to a value of j for which $p_j = 0$; it is apparent that in these cases r_j/p_j fails as an error criterion. However, because, as can be seen from equations (2) and (3), if any p_j is zero then ${}_d q_j$ and ${}_a q_{(j-1)}$ can be found almost as accurately as ${}_d q_{(j-1)}$ and ${}_a q_j$ respectively (no subtractions are involved), $Q_{(j-1)}(x)$ and $Q_{(j+1)}(x)$ are almost equally as good quotients as the "best" $Q_j(x)$.

3. An Example

The polynomial

$$\begin{aligned} P(x) \equiv\ & x^8 + 1112\,x^7 + 113224\,x^6 + 1225336\,x^5 + 2226446\,x^4 \\ & + 1225336\,x^3 + 113224\,x^2 + 1112\,x + 1 \\ \equiv\ & (x^2 + 2x + 1)(x^2 + 10x + 1)(x^2 + 100x + 1)(x^2 + 1000x + 1) \end{aligned}$$

has two ill-conditioned zeros at x = -1 and very well conditioned zeros near -.001, -.01, -.1, -10, -100 and -1000. Following Wilkinson's example on page 62 [8] and using eight digits, a value of -1.0003333 is acceptable* for one of the two zeros at -1. If the factor (x+ 1.0003333) be divided into P(x) the two quotient polynomials

$$\begin{aligned} Q_8(x) = & x^7 + 1110.9997\,x^6 + 112112.63\,x^5 + 1113186.0\,x^4 \\ & + 1112889.0\,x^3 + 112076.1\,x^2 + 1110.55\,x + 1.0799 \end{aligned}$$

and

$$\begin{aligned} Q_o(x) = & 1.0999334\,x^7 + 1110.8997\,x^6 + 112112.73\,x^5 + 1113185.9\,x^4 \\ & + 1112889.1\,x^2 + 112076.02\,x^2 + 1110.6301\,x + .99966681 \end{aligned}$$

---

*Using eight digits, P(-1.0003333) = -.08.

are obtained. For purposes of comparison the zeros of $Q_8(x)$, $Q_0(x)$ and all the composite polynomials $Q_j(x)$ for $j = 1.....7$ were accurately computed; the values of $|E_j|$ for $j = 0....8$ were also calculated. Table I shows the accurate remaining zeros of $P(x)$ in column 1 and the magnitudes of the relative departures from these of the zeros of $Q_j(x)$ for $j = 0....8$ in columns 2-10.* The last line shows $|E_j|$. The correct value of the zero at -1 is given as -.99966681 because this makes the product of the two zeros near -1 equal to 1.0 but errors of this particular zero in the deflated polynomial are not very meaningful.

It is noteworthy that for $j = 0$ (division exclusively in order of ascending powers of x) the largest zero was in error by 10% and for $j = 8$ (exclusively in order of descending powers) the smallest zero was in error by 8% but for $j = 4$† all the zeros were found to the same degree of accuracy as they could be from the original polynomial. Thus it would appear that, for real zeros at least, this method completely avoids the inaccuracies inherent in removing a medium sized zero first.††

4. Division by a Quadratic Factor

If now $P(x)$ is divided by $(x^2 + bx + c)$ to form the quotient

$$Q(x) \triangleq \sum_{i=0}^{n-2} q_i x^{n-i-2}$$

---

* The symbol (-E) means $\times 10^{-E}$.

† The symmetry of $|r_j/p_j|$ about the minimum at $j = 4$ is of course due to the fact that all the zeros are equally spaced on a log scale and that the zero being removed is the geometric mean of all the zeros; in general $j \neq n/2$. The best value of j is correlated with the ratio of the zero being removed to the geometric mean of the remaining zeros but this correlation seems rather unreliable.

†† Ellenberger [9] had previously suggested that if convergence to the largest zero seemed likely, then the order of the coefficients should be reversed--this is equivalent to division exclusively in order of ascending powers of x.

the two sequences for the coefficients of $Q(x)$ are

$$
\begin{aligned}
{}_{d}q_{o} &= p_{o} \\
{}_{d}q_{1} &= p_{1} - {}_{d}q_{o} \\
{}_{d}q_{i} &= p_{i} - b\ {}_{d}q_{i-1} - c\ {}_{d}q_{i-2} \qquad i = 2\ \ldots\ (n-2)
\end{aligned}
$$

and

$$
\begin{aligned}
{}_{a}q_{n-2} &= p_{n}/c \\
{}_{a}q_{n-3} &= (p_{n-1} - b\ {}_{a}q_{n-2})/c
\end{aligned}
$$

and

$$
{}_{a}q_{i} = (p_{i+2} - b\ {}_{a}q_{i+1} - q_{i+2})/c \qquad i = (n-4)\ \ldots 0 \ .
$$

If now, composite quotients are defined by

$$
Q_{j}(x) \triangleq \sum_{i=0}^{j-1} {}_{d}q_{i}\, x^{n-i-2} + \sum_{i=j}^{n-2} {}_{a}q_{i}\, x^{n-i-2} \qquad j = 1\ldots.\ (n-2)
$$

and $(x^2 + bx + c)$ is not an exact factor, then the remainder terms are defined by

$$
P(x) \equiv (x^2 + bx + c)\ \ Q_{j}(x) \ + \ r_{j} x^{n-j} + s_{j+1}\, x^{n-j-1}
$$

and

$$
P_{j}(x) \equiv (x^2 + bx + c)\ Q_{j}(x) = P(x) - r_{j}x^{n-j} - s_{j+1}x^{n-j-1} \ .
$$

Then the least perturbation of the zeros of $P_{j}(x)$ and hence those of $Q_{j}(x)$ from those of $P(x)$ is probably achieved when some function like

$$
E_{j} \equiv \left| r_{j} / p_{j} \right| + \left| s_{j+1} / p_{j+1} \right|
$$

is minimized. This is not quite such an obvious criterion as in the case of division by a linear factor but as then, for polynomials with widely separated zeros this function has a fairly sharp minimum and the value of j which minimizes the disturbance of the remaining zeros cannot be very different from that which minimizes this function.

A difficulty arises in applying this criterion if every odd-subscripted p_j is either zero or very small; then, regardless of the accuracy of an accepted quadratic factor, E_j is meaninglessly large for all values of j.

A very simple solution of this problem is to define

$$E_j = \min \left(\left| r_j / p_j \right| , \left| s_{j+1} / p_{j+1} \right| \right)$$

and, as before, to minimize E_j over all j. The author is not able to justify this choice of a criterion theoretically but it has indicated the "best" quotient in all cases (admittedly a small number) tested so far. This test for a minimum E_j can be very easily carried out using the following expressions for r_j and s_{j+1}

$$r_j = p_j - c\, d^{q}_{j-2} - b\, d^{q}_{j-1} - a^{q}_{j}$$

$$s_{j+1} = p_{j+1} - c\, d^{q}_{j-1} - b\, d^{q}_{j} - a^{q}_{j+1} \quad .$$

5. Conclusions

There is no doubt that the tactic of forming a composite quotient from the two idependently generated sequences of quotient coefficients greatly increases the accuracy of division by an approximate factor. Furthermore, it appears that using such a division technique all the zeros of a polynomial can be found in any order and to almost the same accuracy that could be obtained by purifying each zero in the original polynomial (without the normally accompanying worries about convergence). These advantages would appear to amply justify the extra

arithmetic--approximately 7n multiplications and 3n divisions are required for the removal of a quadratic factor as compared to 2n multiplications for the conventional division.

The method of choosing the cross-over point (from one sequence of quotient coefficients to the other) has been well justified theoretically for the case of a linear factor but requires a more detailed analysis for the case of a quadratic factor.

REFERENCES

1. Olver, F. W. J. The evaluation of zeros of high degree polynomials. Phil. Trans. Roy. Soc. 244 (1952), 385-415.
2. Rutishauser, H. Der Quotienten-Differenzer-Algorithmus. Z. Angew. Math. Physik 5 (1954), 233-251.
3. Bairstow, L. Rep. Memor. Adv. Comm. Aero. London 154 (1914), 51-63.
4. Henrici, P. Elements of Numerical Analysis. (John Wiley and Sons, New York, 1964.)
5. Motteler, Z. C. A simple method for finding the zeros of polynomials, based on D'Alembert's lemma. Los Alamos Scientific Laboratory report No. LA-3355. Aug. 1965.
6. Lehmer, D. H. A machine method for solving polynomial equations. J. Assoc. Comp. Mach. 8 (1961), 151-162.
7. Wilkinson, J. H. The evaluation of the zeros of ill-conditioned polynomials. Numerische Mathematik 1 (1959), 150-166.
8. Wilkinson, J. H. Rounding Errors in Algebraic Processes. (Prentice-Hall Inc., New Jersey, 1963.)
9. Ellenberger, K. W. On programming the numerical solution of polynomial equations. Comm. Assoc. Comp. Mach. 3 (1960), 644-647.

TABLE I

Accurate zeros of P(x)	Relative errors of zeros of $Q_j(x)$								
	j=0	1	2	3	4	5	6	7	8
.0010000010	<1(-7)	<1(-7)	<1(-7)	<1(-7)	<1(-7)	1(-7)	2(-7)	9(-5)	9(-2)
.010001000	1(-7)	1(-7)	1(-7)	1(-7)	1(-7)	1(-7)	1(-6)	1(-4)	1(-2)
.10102051	1(-7)	1(-7)	1(-7)	1(-7)	1(-7)	2(-7)	1(-6)	1(-5)	1(-4)
.99966681	6(-7)	4(-7)	3(-7)	2(-7)	5(-8)	1(-7)	2(-7)	3(-7)	4(-7)
9.8989795	2(-4)	1(-5)	1(-6)	1(-7)	2(-8)	< 1(-8)	<1(-8)	<1(-8)	<1(-8)
99.989999	1(-2)	1(-5)	1(-6)	3(-8)	4(-8)	4(-8)	4(-8)	4(-8)	4(-8)
999.99900	1(-1)	1(-4)	1(-7)	4(-8)	4(-8)	4(-8)	4(-8)	4(-8)	4(-8)
$\lvert r_j/p_j \rvert$	1(-1)	1(-4)	9(-7)	8(-8)	5(-8)	8(-8)	8(-7)	8(-5)	8(-2)