On Calculating the Zeros of Polynomials by the Method of Lucas

Herbert E. Salzer

When $f(x)$ is a polynomial of degree n and x_i, $i=0, 1, \ldots, n$, are any $n+1$ points at which $f(x_i)\neq 0$, the zeros of $f(x)$ are known to be identical with the zeros of [1] $\Sigma a_i/(x-x_i)$, where $a_i=f(x_i)/\Pi'(x_i-x_j)$. Lucas proposed this principle for use in an electric analogue device for finding zeros. The present note evaluates this principle in digital computation for both real and complex zeros when the coefficients of $f(x)$ are given exactly (integral or rational) so that the zeros of $f(x)$ are identical with the zeros of $\Sigma A_i/(x-i)$, A_i integral. The chief advantages are (1) the saving of labor in tabulating $\Sigma A_i/(x-i)$ instead of $f(x)$ in the neighborhood of the zero, especially for complex zeros, and (2) somewhat less work in the inverse interpolation for the zero. Three examples in locating a real root, and one example in locating a complex root were worked out in support of these findings.

In three separate notes F. Lucas[2] describes an electric analogue device for calculating the roots of equations (also mentioned by J. S. Frame).[3] Although the principle is familiar, as far as the writer knows it has not been investigated from the standpoint of digital computation. The present note is intended to call attention to its advantages in finding the roots of polynomial equations with exact coefficients when one has a first approximation as a starting point.

If $f(x)$ denotes a polynomial of degree n, and x_i, $i=0,1, \ldots, n$, denotes any $n+1$ points where $f(x_i)\neq 0$, from

$$f(x)=\sum \frac{\Pi'(x-x_j)}{\Pi'(x_i-x_j)} f(x_i)$$

$$=\Pi(x-x_j)\sum \frac{f(x_i)}{\Pi'(x_i-x_j)(x-x_i)},$$

there follows the well-known result (which is the basis of Lucas's method) that the zeros of $f(x)$ are identical with the zeros of

$$\sum \frac{a_i}{x-x_i}, \text{ where } a_i=\frac{f(x_i)}{\Pi'(x_i-x_j)}.$$

In problems where the coefficients of $f(x)$ are rational, the choice of $x_i=i$ is very convenient. By multiplying through, one obtains the equation in the form $\Sigma A_i/(x-i)=0$, A_i integral, which saves a considerable number of multiplication operations, especially when getting all or even several of the zeros of the same $f(x)$.

By choosing $x_i=x_0+ih$, for any x_0 and h for which $f(x_i)\neq 0$, Lucas's principle can be formulated also as follows:

The zeros of $f(x)$, a polynomial of the nth degree, are identical with the zeros of $\Delta^m\{f(t)/(x-t)\}$, m any integer $\geq n$, where the $f(t)/(x-t)$ is tabulated for any $m+1$ equally spaced values of t.

Of course, in obtaining $\Delta^m\{f(t)/(x-t)\}$, t is the variable, with x as the parameter, and the resulting expression is then regarded as a function of x. (The proof is left for the reader.)

This principle of Lucas, namely, calculating the zeros of $f(x)$ by tabulating $\Sigma a_i/(x-x_i)$ or $\Sigma A_i/(x-x)_i$, instead of $f(x)$ itself, in the neighborhood of a zero, was tried out in several different examples where the zeros had already been obtained from the tabulation of $f(x)$ itself. In all examples, the use of Lucas's method showed a very great saving of computational labor. The two main advantages were (1) much less work in calculating $A_i/(x-x_i)$ instead of the separate terms of $f(x)$, especially those of high degree,[4] and (2) somewhat less work in the inverse interpolation from the tabulated $\Sigma A_i/(x-x_i)$ instead of the tabulated $f(x)$ near the zero, which was apparent from the tendency of Δ^m/Δ to be less in the former case.[5]

In connection with (1), in adding the separate terms of $\Sigma A_i/(x-x_i)$ considerably more significant figures were lost than in the summation of the separate terms of $f(x)$; but that disadvantage is slight because of the ease in getting any number of places in $A_i/(x-x_i)$ by performing continued division on an ordinary 10-bank desk calculator (the $x-x_i$ is almost certainly an exact number having fewer than 10 significant figures), and the total work in tabulating the $\Sigma A_i/(x-x_i)$ is still much less than that in tabulating the $f(x)$. But if the x_i's are not exact, or the coefficients in $f(x)$ are approximate, this principle of Lucas is severely limited in applicability, even to the point of not yielding a single significant figure. Thus, if in the example below, one were to introduce a relative error of 10^{-10} in those exact coefficients of $f(X)$, it is apparent that the $h(X)$ could not be obtained to even one significant figure. However, the choices of $x_i=i$ or $x_i=i-[n/2]$ seem most suitable for many problems. For example, in the case of the classical

[1] Here and elsewhere the summation is over the range 0 to n; similarly Π indicates a product over the same range, and Π' indicates such a product with the vanishing factor omitted.

[2] Lucas, Compt. rend. Acad. Sci. Paris **106**, 645-48 and 1072-74 (1888); **111**, 965-67 (1890).

[3] J. S. Frame, MTAC **1**, 337-53 (especially 347-50) (1945).

[4] There is the very well-known computational scheme for $a_nx^n+a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+\ldots+a_1x+a_0$ in the form $u_n=a_nx+a_{n-1}$, $u_{n-1}=u_nx+a_{n-2}, \ldots$, etc., until $f(x)=u_1=u_2x+a_0$, which avoids the calculation of powers of x, but which may be less convenient for checking and the retention of significant figures.

[5] The rate of convergence of most inverse interpolation series depends upon the rapidity with which the Δ^m/Δ fall off.

orthogonal polynomials where the coefficients are given exactly (for example, Hermite, Laguerre, Chebyshev, Legendre, etc.) the $A_i/(x-x_i)$ can be easily had to any required number of places.

As an illustration, consider the calculation of the zero of the polynomial

$$f(X)=X^6-156X^5+8580X^4-205920X^3$$
$$+2162160X^2-8648640X+8648640,$$

which is near 67.28. Here $f(X)$ was chosen to be $H_{13}(x)/(2x)$, where $X=4x^2$ and $H_{13}(x)$ is the Hermite polynomial of order 13. Choosing $X_i=i$, one has

$$f(0)=\ 86\,48640 \qquad f(4)=-24\,89408$$
$$f(1)=\ 19\,64665 \qquad f(5)=-13\,89935$$
$$f(2)=-15\,15008 \qquad f(6)=\ 69120$$
$$f(3)=-27\,39879$$

The numbers $\Pi'(X_i-X_j)$ have no more than three digits. After obtaining the a_i in their lowest terms, one sees that the zeros of $f(X)$ are identical with the zeros of

$$g(X)\equiv\frac{12012}{X}-\frac{392933}{24\,(X-1)}-\frac{94688}{3\,(X-2)}+\frac{304431}{4\,(X-3)}$$
$$-\frac{155588}{3\,(X-4)}+\frac{277987}{24\,(X-5)}+\frac{96}{X-6},$$

which, in turn, are identical with the zeros of

$$24g(X)\equiv h(X)\equiv\frac{288288}{X}-\frac{392933}{X-1}-\frac{757504}{X-2}$$
$$+\frac{1826586}{X-3}-\frac{1244704}{X-4}+\frac{277987}{X-5}+\frac{2304}{X-6}.$$

An approximate value of the zero is $X=67.2838$, and $h(X)$ was calculated for $X=67.2838(0.0001)67.2841$. The separate terms of $h(X)$ are given here to show the loss in significant figures upon summation:

	$X=67.2838$	$X=67.2839$	$X=67.2840$	$X=67.2841$
$288288/X=$	4284. 65693 07916 61588 7	4284. 65056 27646 43547 7	4284. 64419 47565 54307 1	4284. 63782 67673 93782 5
$-392933/(X-1)=$	−5928. 03973 21819 20771 0	−5928. 03078 87737 44453 8	−5928. 02184 53925 53255 7	−5928. 01290 20383 47054 6
$-757504/(X-2)=$	−11603. 24613 45693 72493 6	−11603. 22836 10507 33795 0	−11603. 21058 75865 44942 1	−11603. 19281 41768 05684 7
$1826586/(X-3)=$	28414. 40611 78710 65494 0	28414. 36191 64363 08313 6	28414. 31771 51390 70375 2	28414. 27351 39793 51037 0
$-1244704/(X-4)=$	−19668. 60397 13165 13862 9	−19668. 57289 13673 14593 4	−19668. 54181 15163 39043 0	−19668. 51073 17635 86746 1
$277987/(X-5)=$	4463. 23120 93995 54940 5	4463. 22404 34526 41854 5	4463. 21687 75287 39323 1	4463. 20971 16278 47235 5
$2304/(X-6)=$	37. 59557 99085 56584 3	37. 59551 85619 71415 0	37. 59545 72155 86450 0	37. 59539 58694 01688 2
$h(X)=$	−0. 00000 00969 68520 0	+0. 00000 00237 72288 6	+0. 00000 01445 13214 6	+0. 00000 02652 54257 8

That only three values of $h(X)$ at intervals of 0.0001 are required in order to find the zero $\overline{X}$ to the maximum attainable accuracy, follows from these differences:

X	$h(X)$	Δ	Δ^2
67. 2838	−0. (6)09696 85200	+0. (6)12074 08086	+0. (12)1174
67. 2839	+. (6)02377 22886	. (6)12074 09260	. (12)1172
67. 2840	. (6)14451 32146	. (6)12074 10432	
67. 2841	. (6)26525 42578		

When these differences were compared with the corresponding differences in $f(X)$, it was noted that Δ^2/Δ for $h(X)$ was only (1/25) of Δ^2/Δ for $f(X)$, from which one can infer that the inverse interpolation is better for the function $h(X)$. The zero $\overline{X}=67.2839+0.0001p$ was found from the three-point formula[6] $p=r-r^2s$, where

$$r=-2h(67.2839)/\{h(67.2840)-h(67.2838)\}$$
$$s=\Delta^2/\{h(67.2840)-h(67.2838)\}.$$

Thus

$$h(67.2840)-h(67.2838)=0.(6)24148\,17346,$$

$$r=-0.(7)47544\,5772/0.(6)24148\,17346$$
$$=-0.19688\,6846,$$
$$s=0.(12)1174/0.(6)24148=0.(6)486,$$
$$r^2=0.03876,\ r^2s=0.(7)1884,$$
$$p=-0.19688\,6865;\ \overline{X}=67.28388\,03113\,135,$$

which happens to be correct to 13 decimals.

This method was tested upon the calculation of two different roots of the same tenth-degree polynomial in X (the example was to find both the smallest and largest zero of $H_{20}(x)$), the result being that the relative saving of labor was even greater than that for $H_{13}(x)$. In fact, as the degree of the polynomial increases, the proportion of work saved also increases.

Finally, this method was applied to the computation of a complex root of a tenth-degree polynomial, from a rough first approximation. There the relative saving of labor was even greater than that for the examples involving real roots, due in particular to the avoidance of the calculation of high powers of complex numbers. Even when the computation is for complex roots, the choice of $x_i=i$ or $x_i=i-[n/2]$ is still suitable, so that if the coefficients of $f(x)$ are real, the calculation involves the sum of fractions with only real numerators.

WASHINGTON, July 24, 1951.

6 H. E. Salzer, Bul. Am. Math. Soc. **50**, No. 8, 513–16 (1944).