

## The Graeffe Process as Applied to Power Series

Of the many methods which have been proposed for solving algebraic equations the most practical one, where complex roots are concerned, is the well known "root-squaring" method usually referred to as the GRAEFFE<sup>1</sup> process. Nearly all accounts of the method show by examples how all the roots of an equation may be found by first finding all the absolute values of the roots, and later by various devices their arguments. The problem of finding the absolute values of the roots is comparatively simple and can be extended (with certain practical difficulties) to the case of the zeros of entire functions. This step was made early in the history of the problem by EULER<sup>2</sup> in generalizing Bernoulli's method to determine the first few zeros of the Bessel function  $J_0(z)$ . More recent accounts of the problem of finding the absolute values of the zeros of entire functions have been given by OSTROWSKI<sup>1</sup> and PÓLYA.<sup>3</sup> The problem of finding by the Graeffe process the arguments of the zeros of entire functions (when these zeros are complex) apparently has not been attempted. At first this may seem a little strange since the Graeffe process owes its popularity to the fact that it is so successful in finding the complex roots of algebraic equations. The explanation is that the various devices, referred to above, for finding the arguments of the complex roots depend on an equation which has a finite number of roots.

There is however one modification of the Graeffe process for polynomials, given more than a score of years ago by BRODETSKY & SMEAL<sup>4</sup> and apparently overlooked by recent writers, in which a pair of complex roots may be found without having to find any other roots. The purpose of this note is to call attention to the fact that this process may be extended to entire functions and their complex zeros.

As mentioned above, accounts of the Graeffe process are nearly always given in terms of examples. To be sure, the "general ideas" of the process are also set forth, but I have yet to see a writer admit that the Graeffe process fails utterly in such simple cases as the equation

$$x^4 + x^3 + x^2 + x + 1 = 0$$

for the primitive fifth roots of unity. Although the Graeffe process is of considerable practical importance its theoretical justification, including the estimation of errors, is not at all easy. Even the lengthy memoir of Ostrowski, which is concerned largely with the absolute values of the roots, leaves something to be desired. If one is given only the coefficients of an equation with no advance information as to the disposition of its roots almost anything may happen if one applies the Graeffe process. Ostrowski shrewdly points out the distinction between the *a priori* and the *a posteriori* evaluation of the error, i.e. before and after the root-squaring process has been applied. It is clear that comparatively little of practical value may be said of the former. Even the number of steps necessary cannot be easily foreseen. The latter, unfortunately, depends on too many contingencies to afford a

simple treatment. In the brief discussion of the following method we must therefore confine ourselves to a few typical cases, leaving a more detailed and rigorous discussion to a later date and a more general method.

Let

$$(1) \quad f(z) = a_0 + a_1z + a_2z^2 + \cdots \quad (a_i \text{ real})$$

be an entire function whose zeros are required. We may suppose that  $a_0 = 1$ . We proceed to calculate for  $m = 0, 1, 2, \dots$  two sequences of numbers

$$(2) \quad \begin{array}{l} A_0^{(m)}, A_1^{(m)}, A_2^{(m)}, \dots, A_r^{(m)}, \dots \\ B_0^{(m)}, B_1^{(m)}, B_2^{(m)}, \dots, B_r^{(m)}, \dots \end{array}$$

defined as follows: For  $m = 0$

$$A_r^{(0)} = a_r \quad B_r^{(0)} = (r-1)a_{r-1}$$

so that  $A_0^{(0)} = 1, B_0^{(0)} = B_1^{(0)} = 0$ . For  $m = 1$

$$A_r^{(1)} = (-1)^r a_r^2 + 2 \sum_{\nu=0}^{r-1} (-1)^\nu a_\nu a_{2r-\nu}$$

$$B_r^{(1)} = \sum_{\nu=0}^{r-1} (-1)^\nu (2r - 2\nu - 1) a_\nu a_{2r-\nu-1},$$

and for  $m > 1$

$$(3) \quad A_r^{(m)} = (-1)^r (A_r^{(m-1)})^2 + 2 \sum_{\nu=0}^{r-1} (-1)^\nu A_\nu^{(m-1)} A_{2r-\nu}^{(m-1)}$$

$$B_r^{(m)} = \sum_{\nu=0}^{2r} (-1)^\nu A_\nu^{(m-1)} B_{2r-\nu}^{(m-1)}.$$

For all  $m$ :

$$(4) \quad A_0^{(m)} = 1, \quad B_0^{(m)} = 0.$$

Theoretically, the zeros of  $f(z)$  are, in general, limits, as  $m \rightarrow \infty$ , of simple functions of  $A_r^{(m)}$  and  $B_r^{(m)}$ . In what follows it will be convenient to let  $s$  denote  $2^m$ . Let the reciprocals of the zeros of (1) be denoted by  $\alpha_k (k = 1, 2, \dots)$ , the notation being such that

$$(5) \quad |\alpha_1| \geq |\alpha_2| \geq |\alpha_3| \geq \dots.$$

Finally let  $\epsilon$  be an independent variable. Then the function defined recurrently by

$$\varphi_0 = \varphi_0(t) = \varphi_0(t|\epsilon) \equiv \sum_{r=0}^{\infty} a_r (t - \epsilon)^{-r}$$

$$\varphi_m(t) = \varphi_{m-1}(\sqrt{t}) \varphi_{m-1}(-\sqrt{t}),$$

has the expansion

$$(6) \quad \varphi_m(t) = \sum_{r=0}^{\infty} t^{-r} (A_r^{(m)} + \epsilon s B_r^{(m)}) + O(\epsilon^2)$$

where  $A_r^{(m)}, B_r^{(m)}$  are given by (3). Moreover the zeros of this function are at

$$(7) \quad t = \alpha_k^s + \epsilon s \alpha_k^{s-1}.$$

Identifying the elementary symmetric functions of (7) with the coefficients of (6) and recalling the definition of  $\epsilon$  and (4), we have, in the notation of symmetric functions

$$(8) \quad \begin{array}{ll} A_1^{(m)} = - \sum \alpha_1^s & B_1^{(m)} = - \sum \alpha_1^{s-1} \\ A_2^{(m)} = \sum \alpha_1^s \alpha_2^s & B_2^{(m)} = \sum \alpha_1^s \alpha_2^{s-1} \\ A_3^{(m)} = - \sum \alpha_1^s \alpha_2^s \alpha_3^s & B_3^{(m)} = - \sum \alpha_1^s \alpha_2^s \alpha_3^{s-1} \\ \dots & \dots \end{array}$$

It is from these equations that we derive approximations to  $\alpha_k$ . Different cases arise, depending on which of the  $\geq$  signs in (5) may be replaced by  $=$ . The simplest case is that in which

$$(9) \quad |\alpha_1| > |\alpha_2| > |\alpha_3| > \dots$$

Although one might call this the "general" case, a moment's reflection shows that in this case all the zeros of  $f(z)$  are real. This is therefore a special case. By the usual argument, the symmetric functions (8) are, in this case, dominated by their first terms, provided  $s$  is so large that  $A_r^{(m)}$  is sensibly the square of  $A_r^{(m-1)}$ . Hence, approximately

$$\begin{array}{ll} A_1^{(m)} = - \alpha_1^s & B_1^{(m)} = - \alpha_1^{s-1} \\ A_2^{(m)} = \alpha_1^s \alpha_2^s & B_2^{(m)} = \alpha_1^s \alpha_2^{s-1} \\ A_3^{(m)} = - \alpha_1^s \alpha_2^s \alpha_3^s & B_3^{(m)} = - \alpha_1^s \alpha_2^s \alpha_3^{s-1} \\ \dots & \dots \end{array}$$

In general, equations of this type continue as long as  $>$  signs continue to hold in (9). A knowledge of the  $A$ 's alone suffices to determine  $|\alpha_k|$  approximately by the usual relations

$$(10) \quad \begin{array}{l} |\alpha_1| = |A_1^{(m)}|^{1/s} \\ |\alpha_2| = |A_2^{(m)}/A_1^{(m)}|^{1/s} \\ |\alpha_3| = |A_3^{(m)}/A_2^{(m)}|^{1/s} \\ \dots \end{array}$$

To determine the signs of these real roots it has been customary<sup>5</sup> to substitute  $\pm \alpha_k$  into the equation  $f(z) = 0$ , or to draw a graph of  $f(z)$ . But knowledge of the  $B$ 's provides a rational method of determining the values of  $\alpha_k$  (with their signs). In fact, we have approximately

$$(11) \quad \begin{array}{l} \alpha_1 = A_1^{(m)}/B_1^{(m)} \\ \alpha_2 = A_2^{(m)}/B_2^{(m)} \\ \alpha_3 = A_3^{(m)}/B_3^{(m)} \\ \dots \end{array}$$

The agreement of the two values of  $|\alpha_k|$  obtained in (10) and (11) affords an excellent check on the work. If, however, the  $B$ 's are calculated only roughly, the signs of  $\alpha_k$  can be determined from the signs of  $B_k^{(m)}$ , since  $s - 1$  is odd.

Suppose now that there are zeros of equal absolute value. The most important case of this kind is that in which the  $\alpha$ 's occur in pairs of complex conjugate values, no two pairs having the same absolute value. This has been called the "general case." To discuss it briefly we write

$$\alpha_1 = \rho_1 e^{i\theta_1}, \alpha_2 = \rho_1 e^{-i\theta_1}, \dots, \alpha_{2k-1} = \rho_k e^{i\theta_k}, \alpha_{2k} = \rho_k e^{-i\theta_k}$$

where the absolute values  $\rho_k$  satisfy

$$\rho_1 > \rho_2 > \rho_3 > \dots$$

For large  $s$ , (8) gives in this case, as dominant terms,

$$\begin{aligned} (12) \quad & A_1^{(m)} = -2\rho_1^s \cos s\theta_1 \\ & A_2^{(m)} = \rho_1^{2s} \\ & A_3^{(m)} = -2\rho_1^{2s}\rho_2^s \cos s\theta_2 \\ & A_4^{(m)} = \rho_1^{2s}\rho_2^{2s} \\ & B_1^{(m)} = -2\rho_1^{s-1} \cos (s-1)\theta_1 \\ & B_2^{(m)} = 2\rho_1^{2s-1} \cos \theta_1 \\ & B_3^{(m)} = -(2\rho_1^{2s}\rho_2^{s-1} \cos (s-1)\theta_2 + 4\rho_1^{2s-1}\rho_2^s \cos \theta_1 \cos s\theta_2) \\ & B_4^{(m)} = 2\rho_1^{2s}\rho_2^{2s-1} \cos \theta_2 + 2\rho_1^{2s-1}\rho_2^{2s} \cos \theta_1. \end{aligned}$$

This case is therefore recognized by the oscillatory character of  $A_{2k+1}^{(m)}$  as functions of  $m$ , together with the fact that  $A_{2k}^{(m)}$  approaches the square of  $A_{2k}^{(m-1)}$  ( $k = 0, 1, 2, \dots$ ). It is now clear how we may determine  $(\rho_1, \theta_1), (\rho_2, \theta_2), \dots$  and verify our results. In fact

$$(13) \quad \begin{aligned} & \rho_1 = |A_2^{(m)}|^{1/2s}, \rho_2 = |A_4^{(m)}/A_2^{(m)}|^{1/2s}, \dots \\ & 2 \cos \theta_1 = \rho_1 B_2^{(m)}/A_2^{(m)}, 2 \cos \theta_2 = \rho_2 \left( \frac{B_4^{(m)}}{A_4^{(m)}} - \frac{B_2^{(m)}}{A_2^{(m)}} \right), \dots \end{aligned}$$

The values of  $(\rho_1, \theta_1), (\rho_2, \theta_2), (\rho_3, \theta_3), \dots$  thus found may be verified by using the expressions of  $A_1^{(m)}, A_3^{(m)}, \dots$  and  $B_1^{(m)}, B_3^{(m)}, \dots$ . If the  $B$ 's are calculated only roughly the values of  $\theta_1, \theta_2, \dots$  obtained from them will be sufficiently exact to determine which one of the  $s$  *a priori* values of  $\theta_k$  as defined by

$$2 \cos s\theta_k = \frac{-A_{2k-1}^{(m)}}{\sqrt{|A_{2k}^{(m)}A_{2k-2}^{(m)}|}}$$

is the correct  $\theta_k$  to choose. The case of more than two zeros having equal absolute values breaks into subcases, each recognizable from the behavior of the  $A$ 's and each capable of a similar though more complicated treatment.

Of course  $f(x) = 0$  may have some real and some complex zeros. In practice this is especially true of polynomials but rather rare in the case of typical entire functions. In this case the above discussions are only slightly modified. Thus if the first three zeros  $\alpha_1, \alpha_2, \alpha_3$  are real and all the other  $\alpha$ 's are complex, equations (10) and (11) hold for  $\alpha_1, \alpha_2$ , and  $\alpha_3$ , while equations (12) and (13) are modified by replacing  $A_k^{(m)}, B_k^{(m)}$  by  $A_{k+3}^{(m)}/A_3^{(m)}, B_{k+3}^{(m)}/A_3^{(m)}$  in accordance with the fundamental equations (8).

We consider now certain practical difficulties which arise when one extends the Graeffe process to entire functions. The most conspicuous of

these is due to the fact that the  $A_r^{(m)}$  and  $B_r^{(m)}$  are each functions of the first  $2^{mr}$  coefficients of  $f(z)$  so that even for moderate values of  $r$  and  $m$  a large number of these coefficients must be known. Of course if  $f(z)$  is merely a polynomial of moderate degree the difficulty does not arise and most of the terms in the calculation of  $A_r^{(m)}$  and  $B_r^{(m)}$  vanish. To put it another way, the Graeffe process for a polynomial  $f$  of degree  $n$  makes considerable use of the fact that the coefficients of  $f$ , after the  $n$ -th, are exactly zero. This should warn the computer to be careful in truncating a power series if he expects the roots of the polynomial thus obtained to agree with the zeros of the series, especially for large zeros. In many typical cases the coefficients  $a_n$  for  $n$  large may be closely approximated not by zero but by some simple asymptotic formula. This fact will reflect itself in the formulas (3) for  $A$  and  $B$ , thus enabling one to estimate the contribution from the far distant coefficients. In many cases also the function  $f(z)$  possesses a simple asymptotic expansion for large  $z$  which enables one to compute (with bounded accuracy) a large zero of  $f(z)$ . In these cases the Graeffe process may be used for finding only those zeros which are comparatively near the origin.

Another source of trouble is loss of accuracy in computing the  $A$ 's and  $B$ 's. This may be due to the accumulation of errors from "rounding off" the coefficients in the given equation. If enough accuracy can be maintained until the process begins to converge, that is, until the expressions for  $A_r^{(m)}$  and  $B_r^{(m)}$  in (3) are dominated by their first few terms, then the crisis is passed and no great loss of accuracy occurs thereafter. This difficulty increases when the roots are nearly equal and is intolerable (as Ostrowski points out) when all the roots coincide, a case which is however of no practical importance.

A more insidious source of trouble is a loss of accuracy due to cancellation of terms in computing  $A_r^{(m)}$  and  $B_r^{(m)}$ . Here the extreme case is that in which  $f(z) = \exp g(z)$ , where  $g$  is an odd function. In this case one may show that  $A_k^{(1)} = 0$  for  $k > 0$ . In particular if  $f(z) = e^{cz}$  the  $B$ 's also vanish with the exception of  $B_1^{(1)} = 2c$ . This puts an end to the Graeffe process. Of course in this case the function  $f(z)$  has no zeros anyway. The converse proposition is also true: If  $A_k^{(1)} = 0$  for  $k > 0$ , and  $B_k^{(1)} = 0$  for  $k > 1$ , then  $f(z) = e^{cz}$  and  $c = \frac{1}{2}B_1^{(1)}$ . Hence in practice we need not expect an outright failure of the process. However, if the coefficients differ but slightly from those of an exponential function one may expect a considerable cancellation of terms in computing  $A_r^{(m)}$  and  $B_r^{(m)}$ , resulting in a devastating loss of accuracy. Hence extreme accuracy is required in the given coefficients  $a_r$ . If the  $a_r$  are defined as rational numbers it may be advisable to use these exact values rather than any approximate decimal equivalents. In this way the  $A$ 's and  $B$ 's appear as exact rational numbers, no accuracy being lost. This device may have to be used for several stages until the process shows signs of converging. If the  $a_r$  are given to within only a certain degree of accuracy and there is this tendency toward cancellation, very little can be done about it and some other method may have to be used. This difficulty is only slightly noticeable when  $f(z)$  is a polynomial of low degree. It is quite pronounced in the following example.

Example. The following equation, proposed to me by F. Wolf, arises in

connection with the flow of a hot gas through a pipe:

$$J_1^2(z) = J_0(z)J_2(z)$$

where  $J_n(z)$  is the ordinary Bessel function of the first kind. It may be shown that

$$z^2\Delta(z) \equiv z^2[J_1^2(z) - J_0(z)J_2(z)] = 4 \sum_{n=2}^{\infty} nJ_n^2(z)$$

from which it follows that except for the trivial double zero at  $z = 0$ ,  $\Delta(z)$  has no real or imaginary zeros. Replacing  $z^2$  by  $-4x$  and expanding  $J_0, J_1$  and  $J_2$  in powers of  $x$  one finds that

$$f(x) \equiv \frac{-2\Delta}{x} = 1 + \frac{2}{3}x + \frac{5}{24}x^2 + \frac{7}{180}x^3 + \dots$$

$$= 2 \sum_{m=0}^{\infty} \binom{2m+2}{m} \frac{x^m}{(m+1)!(m+2)!}$$

If we wish to compute for example  $A_{11}^{(1)}$  and we carry the calculations of the terms to ten significant figures we obtain (in units of the 33rd decimal place)

$-a_{11}^2$	=	-28013	32771
$2a_{10}$	$a_{12}$	=	48133 66627
$-2a_9$	$a_{13}$	=	-30467 02393
$2a_8$	$a_{14}$	=	14129 74452
$-2a_7$	$a_{15}$	=	-4754 97203
$2a_6$	$a_{16}$	=	1144 16515
$-2a_5$	$a_{17}$	=	-192 74739
$2a_4$	$a_{18}$	=	22 06130
$-2a_3$	$a_{19}$	=	1 64295
$2a_2$	$a_{20}$	=	7456
$-2a_1$	$a_{21}$	=	185
$2a_0$	$a_{22}$	=	2

$$A_{11}^{(1)} = \sum_{\nu=0}^{22} (-1)^\nu a_\nu a_{22-\nu} = -404$$

Hence only three significant figures of  $A_{11}^{(1)}$  are obtained from 10-figure data. If however the  $a$ 's are taken at their rational values the rational value of  $A_{11}^{(1)}$  is found to be exactly

$$-19471 \ 33037 / (2^{35} \cdot 3^{19} \cdot 5^8 \cdot 7^5 \cdot 11^4 \cdot 13^2 \cdot 17 \cdot 19 \cdot 23) =$$

$$-4.04020 \ 96279 \ 4 \dots \cdot 10^{-31}.$$

In this example the first 32 coefficients of  $f(x)$  were used to find

$$(14) \quad \begin{cases} A_1^{(4)} = 1.9683 \ 98486 \cdot 10^{-12} & B_1^{(4)} = -1.4572 \ 73376 \cdot 10^{-11} \\ A_2^{(4)} = 1.7561 \ 21066 \cdot 10^{-24} & B_2^{(4)} = -1.5624 \ 51484 \cdot 10^{-23} \\ \begin{cases} A_1^{(5)} = -3.62350470 \cdot 10^{-25} \\ B_1^{(5)} = 1.3060 \ 43223 \cdot 10^{-23} \end{cases} & = 2\rho_1^{32} \cos 32\theta_1 \\ A_1^{(3)} = -8.2527 \ 36680 \cdot 10^{-7} & B_1^{(3)} = -3.3783 \ 23209 \cdot 10^{-6} \end{cases}$$

We first find that

$$[A_2^{(4)}]^{1/32} = 0.18098 \ 48935$$

$$[A_2^{(4)}]^{1/32} B_2^{(4)} / A_2^{(4)} = -1.6102 \ 54104.$$

These are approximations to  $\rho_1$  and  $2 \cos \theta_1$  respectively. Substituting these into (14) we find by successive approximations that (14) has the simultaneous solution

$$\begin{aligned}\rho_1 &= 0.1809\ 84894 \\ \theta_1 &= 2.5066\ 8639.\end{aligned}$$

In this calculation it is not feasible to carry the main root-squaring process any further. One could have wished for more coefficients and higher powers, to separate more zeros. In this case the remaining zeros can be approximated by an asymptotic formula with sufficient accuracy except in the case of  $\alpha_3$  and  $\alpha_4$  which are still fairly close to the origin. The formula gives only

$$\rho_2 = 0.063, \quad \theta_2 = 2.70.$$

Since the larger zeros are well determined and since we now know the first pair of complex zeros, it is a simple matter to compute the sums of the  $k$ th powers of all the zeros except  $\alpha_3$  and  $\alpha_4$  for  $k = 7$  and  $8$ . From the above values of  $A_1^{(3)}$  and  $B_1^{(3)}$  it follows that

$$\begin{aligned}\alpha_3^8 + \alpha_4^8 &= 2\rho_2^8 \cos 8\theta_2 = -5.4461 \cdot 10^{-10} \\ \alpha_3^7 + \alpha_4^7 &= 2\rho_2^7 \cos 7\theta_2 = 9.1125 \cdot 10^{-9}\end{aligned}$$

The approximate values of  $\rho_2$  and  $\theta_2$  given above make it easy to determine by successive approximations the solution

$$\rho_2 = 0.064339, \quad \theta_2 = 2.70071$$

of the above system.

The above example is given in such detail in order to indicate what steps may be necessary in dealing with intractable entire functions. When one is deprived of the use of extremely high powers of the zeros (customarily used in treating polynomials) one is obliged to obtain more information than usual from the lower powers.

D. H. L.

<sup>1</sup> For a typical account of this method, together with some historical remarks, see E. T. WHITTAKER & G. ROBINSON, *The Calculus of Observations*, third ed., London, 1940, p. 106–120. References to other accounts are given in C. A. HUTCHINSON, "On Graeffe's method for the numerical solution of algebraic equations," *Am. Math. Mo.*, v. 42, 1935, p. 149–161, and L. L. CRONVICH, "On the Graeffe method of solution of equations," *Am. Math. Mo.*, v. 46, 1939, p. 185–190. An early exponent of this method, who did much to make it well known, was the astronomer J. F. ENCKE (*J. f. d. reine u. angew. Math.*, v. 22, 1841, p. 193f). He attributed the method to Graeffe. He was also responsible for an utterly useless scheme of changing the signs of all the roots. Many writers on the subject still use this confusing notion of the "Encke roots" of an equation. Graeffe's method of solutions of numerical equations was really due to DANDELIN (1794–1847), *Mem. Acad. Brussels*, v. 3, 1826, p. 46, or to WARING (1734–1798), *Meditationes Analyticae*, 1776, p. 311. It was published by C. H. GRAEFFE (1799–1873) in a prize paper, *Die Auflösung der höheren numerischen Gleichungen*, Zürich, 1837. Other accounts of the history of the method will be found in A. M. OSTROWSKI, "Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent," *Acta Math.*, v. 72, 1940, p. 99–155, and E. CARVALLO, *Méthode pratique pour la Résolution numérique complète des équations algébriques ou transcendentes*, Paris diss., 1890.

<sup>2</sup> L. EULER, Akad. Nauk, S.S.S.R., Leningrad, *Acta*, 1781, part 1, 1784, p. 170f.

<sup>3</sup> G. PÓLYA, "Über das Graeffesche Verfahren," *Z. Math. Phys.*, v. 63, 1915, p. 275–290.

<sup>4</sup> S. BRODETSKY & G. SMEAL, "On Graeffe's method for complex roots of algebraic equations," *Camb. Phil. So., Proc.*, v. 22, 1924, p. 83f.

<sup>5</sup> A. C. AITKEN has proposed a root-cubing procedure in which real roots are determined together with their signs; see *Math. Gazette*, v. 15, 1931, p. 490–491.