PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The Evaluation of Zeros of High-Degree Polynomials

F. W. J. Olver

Phil. Trans. R. Soc. Lond. A 1952 **244**, 385-415
doi: 10.1098/rsta.1952.0010

THE EVALUATION OF ZEROS OF HIGH-DEGREE POLYNOMIALS

By F. W. J. OLVER, *The National Physical Laboratory*

(*Communicated by E. C. Bullard, F.R.S.—Received 6 July 1951*)

CONTENTS

The practical application of the classical numerical methods for solving polynomial equations presents special difficulties in the case of polynomials of high degree. The theoretical and practical aspects of these difficulties are here considered in detail, together with the consequent modifications imposed upon the methods. The treatment of the whole subject is intended to be as comprehensive as possible and comparisons are made of the efficacy and speed of the various processes of solution.

Introduction

The history* of the numerical evaluation of zeros of polynomials of arbitrary degree reveals that the subject received thorough attention from the early mathematicians. With perhaps the exception of Aitken's elegant generalization of Bernoulli's process, comparatively little has been discovered during the last hundred years which is fundamentally new and of practical importance. Indeed, as far as polynomials of degrees greater than six are concerned, nearly every useful known general method which is suitable for use with desk calculating machines is equivalent to one of the following: (i) Newton's rule, (ii) the root-squaring process (including its extensions), or (iii) the Aitken-Bernoulli process. Nevertheless, little research appears to have been directed towards attaining technical perfection in the application of these standard methods. Few writers seem to have carried out the complete numerical solution of a polynomial of degree greater than ten; almost all appear to be unaware of several difficulties arising in practice, which are peculiar to high-degree polynomials.

The main purpose of the present work is to consider the practical application of general methods to polynomials of degree greater than about six. For polynomials of lower degree

* For an account of this see, for example, Whittaker & Robinson (1944, chap. 6).

Vol. 244. A. 885. (Price 7*s.* 6*d.*) 50 [Published 3 April 1952

various special methods exist, the study of which comprises a separate subject and is outside the scope of this paper. The conclusions drawn here are based on experience gained during the past few years at the Mathematics Division, National Physical Laboratory, in the solution of polynomials of degrees ranging from six to twenty-four.

Although the paper has been written primarily for the computer using desk calculating machines, much of it is applicable to the solution of high-degree polynomials using any equipment. This is particularly true, for example, in the sections which deal with the difficulties arising from the presence of zeros which are close together; difficulties that occur more often as the degree is increased and which are bound to manifest themselves whatever equipment is employed. In particular, analogue machines such as isographs can be expected to yield meaningful direct approximations only to the zeros of well-conditioned polynomials (that is, those polynomials whose zeros are not over-sensitive to small changes in the coefficients), whereas in practice ill-conditioned polynomials occur quite frequently. This aspect is further considered in § 3.

The paper is divided into three parts. Parts A and B correspond to the natural classes of the available methods, namely, the direct and the indirect or iterative methods. The latter class needs little explanation; it comprises all processes of successive approximation. Using a direct method some or all of the zeros can be evaluated to any desired accuracy by a *single* application of the process, provided a sufficient number of significant figures is retained during the calculations.

In parts A and B it is supposed that the coefficients of the polynomial are real and known exactly. The latter condition will not, of course, be fulfilled in many of the physical problems in which the polynomials arise, but even then it is convenient to regard the given coefficients as exact during the process of solution. An *a posteriori* examination of the influence upon the zeros of rounding or observational errors in the coefficients can be made with the aid of simple formulae stated in part C. Also included in the third part is a brief discussion of the solution of polynomials with complex coefficients.

Part A. Direct methods

There are three direct methods of importance:

(i) Inverse interpolation.

(ii) The Aitken-Bernoulli processes.

(iii) The root-squaring process (including its extensions for the determination of the phases of the zeros).

Tabulation of the polynomial followed by inverse interpolation is an effective means of evaluating real zeros, especially if a multi-register adding-listing machine (e.g. a National accounting machine) is available. If, however, there are present several complex zeros whose values are also required, a knowledge of the real zeros is seldom of much assistance and it is preferable to employ some other method from the start.

Method (ii), which is described and illustrated in the next section, can sometimes be employed advantageously. It is, however, far exceeded in importance by method (iii), a full discussion of which is given in §§ 2 to 5.

(1) *The Aitken-Bernoulli processes*

A detailed account of these methods has been given by Aitken (1926). The basis of Bernoulli's method is that if the polynomial

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \tag{1·1}$$

has n simple zeros $x_1, x_2, \ldots, x_n$ (supposed for convenience to be arranged in non-ascending order of modulus magnitude), then the difference equation

$$a_0 f_1(t) + a_1 f_1(t+1) + a_2 f_1(t+2) + \ldots + a_n f_1(t+n) = 0 \tag{1·2}$$

has the general solution $$f_1(t) = \omega_1 x_1^t + \omega_2 x_2^t + \ldots + \omega_n x_n^t, \tag{1·3}$$

where $\omega_1, \omega_2, \ldots, \omega_n$ are arbitrary constants. If $|x_1| > |x_2|$ (so that x_1 is necessarily real), then clearly

$$f_1(t) \sim \omega_1 x_1^t \quad \text{and} \quad f_1(t+1)/f_1(t) \to x_1, \quad \text{as } t \to \infty. \tag{1·4}$$

From a sequence of numerical values of $f_1(t)$ constructed by using (1·2) as a recurrence relation, the value of x_1 can usually be obtained to any prescribed degree of accuracy from (1·4), provided the sequence is taken far enough. The factor $x - x_1$ can be removed from $f(x)$ and the process repeated to obtain x_2, provided $|x_2| > |x_3|$. Further applications enable $f(x)$ to be solved completely, provided its zeros are all real and distinct.

Aitken's generalization is designed to determine *all* the zeros from the sequence $\{f_1(t)\}$ and to include the cases of complex and multiple zeros. He constructs further sequences $\{f_s(t)\}$ defined by

$$f_2(t) = \begin{vmatrix} f_1(t) & f_1(t+1) \\ f_1(t-1) & f_1(t) \end{vmatrix}, \quad f_{s+1}(t) = \begin{vmatrix} f_s(t) & f_s(t+1) \\ f_s(t-1) & f_s(t) \end{vmatrix} \div f_{s-1}(t) \quad (s \geqslant 2), \tag{1·5}$$

and proves that if $Z_s(t) \equiv f_s(t+1)/f_s(t)$ and $|x_s| > |x_{s+1}|$, then

$$Z_s(t) \to x_1 x_2 \ldots x_s, \quad \text{as } t \to \infty. \tag{1·6}$$

This result enables us to evaluate numerically the real zeros and the moduli of the complex zeros. To indicate how the phases of the complex zeros may be obtained, suppose that x_1, x_2 are real and distinct, x_3, x_4 are a pair of complex conjugates $r e^{\pm i\theta}$ and that $|x_2| > r > |x_5|$. Then Aitken proves that

$$Z_3(t) \sim x_1 x_2 r \cos\{(t+1)\theta + \alpha\} \sec(t\theta + \alpha), \tag{1·7}$$

where α is an arbitrary constant, and hence that

$$\{Z_3(t+1) Z_3(t) + k^2\}/Z_3(t) \to 2k \cos\theta, \tag{1·8}$$

where $k \equiv x_1 x_2 r$ and is known by previous application of (1·6).

Let us now consider the practical application of Aitken's method to high-degree polynomials. The speed of convergence may be considered first. If $|x_s| > |x_{s+1}|$, then the difference between $Z_s(t)$ and its asymptotic value is of order $|x_{s+1}/x_s|^t$. Each step thus reduces the relative error in the estimate for $x_1 x_2 \ldots x_s$ in the ratio $|x_{s+1}/x_s|$; in effect a fixed number of significant figures is added at each step. On the other hand, it is well known that after t transformations the root-squaring process has a relative error in the estimate for $x_1 x_2 \ldots x_s$ of order $|x_{s+1}/x_s|^T$, where $T = 2^t$, so that the number of accurate figures eventually

doubles with each transformation. Although this simple comparison needs qualifying (since the labour of an Aitken step is considerably less than that of a root-squaring transformation), it does demonstrate the much superior power of separation possessed by the root-squaring process, especially when the modulus ratios are close to unity.

The real drawback to Aitken's process is, however, the severe cancellation that occurs in the numerical calculation of the sequences $\{f_s(t)\}$. Suppose $|x_s|>|x_{s+1}|$, then from (1·6) we have

$$f_s(t+1) = c_s f_s(t)\{1+\epsilon_s(t)\}, \tag{1·9}$$

where $c_s = x_1 x_2 \dots x_s$ and $\epsilon_s(t)\to 0$ as $t\to\infty$. Hence

$$\begin{vmatrix} f_s(t) & f_s(t+1) \\ f_s(t-1) & f_s(t) \end{vmatrix} = \{f_s(t)\}^2 \begin{vmatrix} 1 & 1+\epsilon_s(t) \\ \{1+\epsilon_s(t-1)\}^{-1} & 1 \end{vmatrix}$$
$$= \{f_s(t)\}^2\{\epsilon_s(t-1)-\epsilon_s(t)\} \quad \text{approximately,}$$

demonstrating the cancellation that takes place on forming $f_{s+1}(t)$. In fact, if R leading figures of $Z_s(t)$ and $Z_s(t-1)$ are identical, then R figures are lost 'off the front' in forming $f_{s+1}(t)$. Since R increases with t, this cancellation becomes increasingly severe as the work progresses, and a large number of significant figures must be retained in the early sequences of the process.

Aitken himself is aware of this difficulty and remarks (1926, p. 294): 'If it were really necessary to carry the columns to any great length this [cancellation] would constitute a defect, but we shall show in a later section how, given a few consecutive terms of a slowly approximating sequence Z_m, it is in general possible to derive successively other sequences which give enormously improved results. Again, to find the *smallest* roots, we can make use of the *reciprocal* equation.' Unfortunately, the former of these devices fails to remove the difficulty. The derived sequences to which he refers are denoted by $Z_s^{(m)}(t)$, and if x_s, x_{s+1} and x_{s+2} are real and $|x_s|>|x_{s+1}|>|x_{s+2}|$, then

$$\left.\begin{aligned} Z_s^{(1)}(t) &= \begin{vmatrix} Z_s(t+1) & Z_s(t+2) \\ Z_s(t) & Z_s(t+1) \end{vmatrix} \div \Delta^2 Z_s(t), \\ Z_s^{(2)}(t) &= \begin{vmatrix} Z_s^{(1)}(t+1) & Z_s^{(1)}(t+2) \\ Z_s^{(1)}(t) & Z_s^{(1)}(t+1) \end{vmatrix} \div \Delta^2 Z_s^{(1)}(t), \end{aligned}\right\} \tag{1·10}$$

Δ being the forward difference operator. Although these sequences eventually do converge more rapidly than $Z_s(t)$, they do not give improved accuracy in the estimate for $x_1 x_2 \dots x_s$ until $Z_s(t)$ has begun to settle down to its asymptotic value, and of course it is when this stage is attained that the cancellations start to occur. It is interesting to note that in forming $Z_s^{(1)}(t)$ and $Z_s^{(2)}(t)$ further cancellation takes place, and if a fixed number of figures is retained in $Z_s(t)$, then eventually the derived sequences $Z_s^{(1)}(t)$ and $Z_s^{(2)}(t)$ produce no more accurate results than $Z_s(t)$ itself.

The derived sequences applicable to complex zeros are also considered by Aitken; they are more complicated than (1·10), and have the same limited advantages.

To illustrate the discussion above, extracts from a numerical example are given in tables 1 and 2. The polynomial being solved is

$$f(x) = 4x^{16}-36x^{15}+139x^{14}-409x^{13}+1056x^{12}-1916x^{11}+3086x^{10}-4618x^9+3912x^8$$
$$-4900x^7+3651x^6+1019x^5+5644x^4+5276x^3+4090x^2+1882x+420, \tag{1·11}$$

the zeros of which are known to be $x_1 = 3\frac{1}{2}$, $x_2 = 3$, $x_3 = 2\frac{1}{2}$, $x_4 = 2$ and twelve irrational complex zeros $x_5, x_6, \ldots, x_{16}$, each of modulus less than 1·94. The initial values of $f_1(t)$ were chosen so that $\omega_1 = \omega_2 = \ldots = \omega_n = 1$, by using Newton's recurrence formulae for the sums of powers of the zeros. In the columns headed $f_1(t)$, $-f_2(t)$ and $-f_3(t)$, the number in italics gives the power of ten by which the given value is to be multiplied. Thus

$$f_3(30) = -1{\cdot}79 \times 10^{38}.$$

TABLE 1. AITKEN'S PROCESS

t	$f_1(t)$	$Z_1(t)$	$-f_2(t)$	$Z_2(t)$	$-f_3(t)$	$Z_3(t)$
30	2·11981 5599 *16*	3·49510 2810	1·05026 8 *29*	10·46019	1·79 *38*	96·6
31	7·40897 3456 *16*	3·49580 2303	1·09860 0 *30*	10·47131	1·73 *40*	17·5
32	2·59003 0647 *17*	3·49640 1787	1·15037 8 *31*	10·47750	3·02 *41*	9·5
33	9·05578 7782 *17*	3·49691 5673	1·20530 9 *32*	10·47939	2·88 *42*	48
34	3·16673 2623 *18*	3·49735 6123	1·26309 0 *33*	10·4822	1·37 *44*	40
35	1·10751 9173 *19*	3·49773 3630	1·32400 *34*	10·4859	5·5 *45*	24
36	3·87380 7057 *19*	3·49805 7227	1·38833 *35*	10·4886	1·3 *47*	17
37	1·35507 9877 *20*	3·49833 4625	1·45616 *36*	10·4901	2·2 *48*	34
38	4·74052 2854 *20*	3·49857 2417	1·52752 *37*	10·4916	7·5 *49*	32
39	1·65850 6250 *21*	3·49877 6257	1·60262 *38*	10·4933	2·4 *51*	21

TABLE 2. AITKEN'S AUXILIARY SEQUENCES

t	$Z_1(t)$	$-Z_1^{(1)}(t)$	$Z_1^{(2)}(t)$	$Z_2(t)$	$-Z_2^{(1)}(t)$	$Z_2^{(2)}(t)$
20	3·47690 1723	3·49940	3·49896	10·67828	10·167	10·114
21	3·48027 7325	3·49832	3·50060	10·28506	10·257	10·447
22	3·48314 6464	3·49991	3·50015	10·19442	10·501	10·372
23	3·48555 9304	3·50039	3·49978	10·39326	10·432	10·405
24	3·48762 4858	3·49989	3·49986	10·46311	10·400	10·05
25	3·48940 2708	3·49980	3·50004	10·40623	10·406	11·06
26	3·49092 2930	3·49997	3·50000	10·40029	10·47	10·46
27	3·49222 0890	3·50002	3·49998	10·44285	10·46	10·46
28	3·49333 2719	3·49998	3·49998	10·46103	10·46	10·46
29	3·49428 6110	3·49998	3·50000	10·45659	10·45	10·48

From these tables it is seen that ten-figure accuracy in the primary sequence $\{f_1(t)\}$ produces only the zeros x_1 and x_2. Even allowing for the fact that the zeros of smallest modulus can be obtained from the reciprocal polynomial, it is clear that something of the order of thirty significant figures would be needed in $f_1(t)$ in order to solve $f(x)$ completely. The same example has been solved by the methods recommended in §§ 2 to 5, keeping a maximum of ten significant figures in the work. Seven root-squaring transformations produced all the moduli of the zeros correct to five or more figures, and the phases were then obtained without difficulty to the same accuracy. Other examples tried both ways have justified the conclusion that in general the labour necessitated by Aitken's method is much greater than that required by the methods described in §§ 2 to 5.

An advantage that is frequently claimed for the Aitken-Bernoulli processes is that checking is unnecessary because isolated errors made in the course of the work will not affect the final results. As far as high-degree polynomials are concerned this is to some extent illusory, since errors are not uncommon and their uncontrolled introduction every few steps may prolong the convergence to a prohibitive extent. Some check is necessary; for example, after every four steps the value of $f_1(t)$ can be verified by direct calculation from the values of $f_1(t-4), f_1(t-5), \ldots, f_1(t-n-3)$.

In conclusion, therefore, from the standpoints of speed and certainty of success the Aitken-Bernoulli processes are much inferior to the processes of §§ 2 to 5. Only if one or two extreme zeros are all that is required, or perhaps if automatic computing equipment is available, can they be considered as useful methods.

(2) *Practical process of root-squaring*

The root-squaring process is essentially a means of calculating the moduli of the zeros. The principles on which it is founded are well known and have the following basis. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the zeros of the polynomial

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \tag{2·1}$$

then a polynomial whose zeros are $-\alpha_1^2, -\alpha_2^2, \ldots, -\alpha_n^2$, is given by

$$g(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n, \tag{2·2}$$

where

$$b_s = a_s^2 - 2a_{s-1}a_{s+1} + 2a_{s-2}a_{s+2} - \ldots \quad (s = 0, 1, 2, \ldots, n). \tag{2·3}$$

The application of this transformation m times in succession yields a polynomial whose zeros are $-\alpha_1^M, -\alpha_2^M, \ldots, -\alpha_n^M$, where $M = 2^m$. For sufficiently large m, the transformed polynomial breaks up and the moduli of the zeros can be computed from the ratios of the arithmetical Mth roots of adjacent coefficients (see, for example, Whittaker & Robinson, 1944, §§ 54 to 58).

One of the major computational difficulties of the process is the checking of each transformation; undetected mistakes can make the subsequent computations useless. As the transformations proceed the coefficients involved become very large or very small numbers, so that algebraical checks based on the relation

$$f(x)\,f(-x) = g(-x^2) \tag{2·4}$$

have no practical value.

The only suitable kind of check seems to be the numerical one of duplication. It is well known that ordinary duplication of complicated calculations is at best a poor check, even if performed by another computer, and this is especially true in root-squaring owing to the inconvenient magnitudes of the numbers involved. Consequently, the computational arrangement and the methods of effecting the transformation and duplication check are of the utmost importance if the work is to be error-free. It is therefore desirable to mention some details of the routine used at the Mathematics Division for carrying out the transformations.

The large and small numbers are conveniently handled when expressed in *standard form*: $N = p.10^q$, where p lies between 1 and 10 and q is an integer. The numbers p and q are recorded in adjacent columns.

Each transformation begins with the coefficients, say $a_0, a_1, \ldots, a_n$, derived from the preceding transformation. The sequence of operations is as follows:

(i) Below the coefficients $a_0, a_1, \ldots, a_n$ we write, in standard form, modified coefficients $A_s = \pm a_s$, where the positive sign is taken if s or $s+1$ is a multiple of four and the negative sign otherwise.

(ii) The following array is formed:

$$\left.\begin{array}{ccccc} A_0^2 & A_1^2 & A_2^2 & A_3^2 & \dots \\ & 2A_0A_2 & 2A_1A_3 & 2A_2A_4 & \dots \\ & & 2A_0A_4 & 2A_1A_5 & \dots \\ & & & 2A_0A_6 & \dots \\ & & & & \dots \end{array}\right\} \tag{2·5}$$

This is most conveniently effected by working along the diagonals using a constant multiplier for each. It is preferable to make a separate task of the determination of the indices and signs, and if this is carried out first the formation of negligible products (of which there is a considerable proportion in the later transformations) can be avoided.

(iii) The columns of the array are summed, taking account of the indices and signs, to give the coefficients $b_0, b_1, \dots, b_n$ of the transformed polynomial.

(iv) The numbers b_s are checked directly from the set a_s by

$$b_s = (\tfrac{1}{2}a_s^2 - a_{s-1}a_{s+1} + a_{s-2}a_{s+2} - \dots) \times 2.$$

In accumulating each b_s on the calculating machine the common index to which the products are formed is taken to be the value recorded at step (iii).

Table 3 gives a numerical example. It shows the transformation from $m = 2$ to $m = 3$ of the polynomial

$$\begin{aligned} &1 + 8x + 32{\cdot}8x^2 + 89{\cdot}6x^3 + 190{\cdot}68x^4 + 304{\cdot}08x^5 + 443{\cdot}576x^6 + 468{\cdot}88x^7 + 524{\cdot}327x^8 \\ &+ 378{\cdot}908x^9 + 345{\cdot}07256x^{10} + 166{\cdot}44768x^{11} + 128{\cdot}218748x^{12} + 37{\cdot}651096x^{13} \\ &+ 25{\cdot}1783048x^{14} + 3{\cdot}4356048x^{15} + 2{\cdot}03253121x^{16}. \end{aligned} \tag{2·6}$$

Two further points should be noted. First, the advantage of writing the coefficients in two staggered rows as in the example is one of numerical convenience; only members of the same row are then multiplied together. Secondly, the principal reason for the introduction of modified coefficients A_s is to enable the signs of the products to be checked independently. It is in the determination of the signs and indices that errors would otherwise be most likely to occur.

By adhering rigidly to the procedure outlined above it has been found that the occurrence of undetected errors is virtually eliminated.

(3) *Special difficulties in the root-squaring of high-degree polynomials*

Common criticisms of the root-squaring process are that it is not self-checking and that end-figure errors rapidly accumulate. The checking described in § 2 shows how the first-named defect may be remedied. The second criticism is superficial, for although it is true that the relative errors of the coefficients are approximately doubled upon each transformation producing relative errors after m transformations of the order $2^m = M$, the final step of taking the arithmetical Mth roots recovers the original precision.

Real difficulties which often occur with the root-squaring of high-degree polynomials are (*a*) the failure of the polynomial to break up after a reasonable number of transforma-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Table 3. Root-squaring

s	0	1	2	3	4	5	6	7	8
$a_{2,s}$	+1·0		+79·75308 2		+ 1·52162 45772 6		+ 0·40363 9028 8		+ 0·00791 08557 10
		− 44·64		− 1·90356 40400 4		+ 0·11538 304414 8		+ 0·00738 29630 10	
$A_{2,s}$	+1·0		− 7·97530 8 3		+ 1·52162 45772 6		− 4·03639 028 7		+ 7·91085 57 7
		+ 4·464 1		− 1·90356 40400 4		− 1·15383 04414 7		+ 7·38296 30 7	
	+1·0	+ 19·92729 6 2	+63·60553 769 6	+ 3·62355 6054 8	+ 2·31534 1354 12	+ 1·33132 46875 14	+16·29244 65 14	+54·50814 3 14	+62·58163 8 14
		− 15·95061 6 3	−16·99501 975 5	− 24·27084 933 9	+ 4·39278 0273 11	− 12·28374 13 13	−17·03737 5 14	−63·86260 2 14	−78·71549 7 14
			+ 3·04324 9154 6	− 10·30139 818 8	+64·38291 14 10	− 28·10788 6 11	+24·07470 5 13	+12·30188 15 14	+19·04834 35 14
				− 8·07278 056 7	+65·91509 4 8	−126·18302 10	+20·29537 11	− 7·18077 9 13	−16·20066 5 13
					+15·82171 7	− 47·594 8	+37·6367 10	−26·728 10	+ 4·29497 12
						− 4·72 7	+62·7 7	−22·51 9	+ 7·19 9
							+ 3 6	−17 6	+26 7
								...	...
$a_{3,s}$	+1·0√		+64·94928 487√ 6		+ 3·40519 8222√ 12		+ 1·68660 7√ 14		+ 1·33744 2√ 14
		−139·57886√ 2		−250·19361 35√ 8		+ 0·06217 630√ 14		+ 2·22644 7√ 14	

s	9	10	11	12	13	14	15	16
$a_{2,s}$		+ 0·00235 957652 10		+ 0·01411 310713 8		+ 1·61136 8060 4		+17·06667 39678
	+ 0·00533 088793 10		+ 0·07020 38378 8		+ 0·18889 5559 6		+ 7·92095 2596 2	
$A_{2,s}$		− 2·35957 652 7		+ 1·41131 0713 6		− 1·61136 8060 4		+ 1·70666 73968 1
	− 5·33088 793 7		+ 7·02038 378 6		− 1·88895 559 5		+ 7·92095 2596 2	
	+ 28·41836 61 14	+ 5·56760 135 14	+ 49·28578 84 12	+ 1·99179 7929 12	+ 3·56815 322 10	+ 2·59650 7025 8	+62·74149 003 4	+ 2·91271 3603 2
	− 37·33253 9 14	−74·84975 83 13	− 6·66019 124 13	−26·52238 64 11	− 4·54828 2011 10	−29·92465 54 7	− 5·50013 8665 5	
	+103·66247 13	+22·32935 1 13	+ 20·13962 11 12	+ 7·60429 248 11	+111·21625 43 8	+ 4·81727 596 7		
	− 11·39320 17 13	−27·89217 8 12	− 25·49460 0 11	−84·45142 12 9	− 8·05402 46 8			
	+ 4·35907 12	+13·00822 11	+116·96020 9	+27·00239 9 8				
	− 4·904 10	−18·279 9	− 13·7776 8					
	− 30 6	+ 5·2 7						
	...							
$a_{3,s}$		+ 0·04946 48√ 14		+ 0·01823 736√ 12		+ 0·08576 908√ 8		+ 2·91217 3603√ 2
	+ 0·35585 4√ 14		+ 0·38961 95√ 12		+ 0·05149 351√ 10		+ 7·74010 338√ 4	

tions, (*b*) the severe loss of significant figures due to cancellation. Although (*b*) is more common and more serious it is convenient to discuss (*a*) first.

The root-squaring process may converge only slowly when some zeros are of nearly equal modulus. A full understanding of this point necessitates a close examination of the theory of the process.

Suppose that after m root-squaring transformations the polynomial (2·1) becomes

$$f_m(x) = a_{m,0} + a_{m,1}x + a_{m,2}x^2 + \ldots + a_{m,n}x^n. \tag{3·1}$$

If the zeros of $f(x)$, arranged in non-ascending order of modulus magnitude, are $\alpha_1, \alpha_2, \ldots, \alpha_n$, then the zeros of $f_m(x)$ are $-\alpha_1^M, -\alpha_2^M, \ldots, -\alpha_n^M$, where $M = 2^m$, and

$$a_{m,n-s}/a_{m,n} = \Sigma\alpha_1^M\alpha_2^M \ldots \alpha_s^M, \tag{3·2}$$

the summation extending over all the products of $\alpha_1^M, \alpha_2^M, \ldots, \alpha_n^M$ taken s at a time.

Write $\rho_s \equiv |\alpha_{s+1}/\alpha_s|$, so that $\rho_s \leqslant 1$, and let us first suppose that α_s and α_{s+1} are two zeros whose moduli are unequal, so that $\rho_s < 1$. Then from equation (3·2) we see that

$$a_{m,n-s}/a_{m,n} = \alpha_1^M\alpha_2^M \ldots \alpha_s^M(1+\epsilon_m), \quad \text{where} \quad \epsilon_m \to 0 \ \text{as} \ m \to \infty. \tag{3·3}$$

Hence

$$a_{m+1,n-s} \sim a_{m,n-s}^2 \quad \text{as} \quad m \to \infty. \tag{3·4}$$

Thus if m is sufficiently large, further transformations retaining a limited number of significant figures merely have the effect of squaring the coefficient of x^{n-s}. When this state of affairs is reached

$$a_{m,n-s}/a_{m,n} = \alpha_1^M\alpha_2^M \ldots \alpha_s^M, \tag{3·5}$$

to the number of accurate figures in $a_{m,n-s}$, and we shall say that the coefficient of x^{n-s} has *settled*.

Secondly, suppose that α_s and α_{s+1} are two zeros with equal moduli, so that $\rho_s = 1$. Then the asymptotic expression for $a_{m,n-s}/a_{m,n}$ comprises at least two terms of the right-hand side of equation (3·2), and the characteristic feature of the coefficient of x^{n-s} ultimately squaring itself upon each transformation disappears. We shall say, when $\rho_s = 1$, that the coefficient of x^{n-s} has *settled* when $a_{m,n-s}/a_{m,n}$ equals its asymptotic expression to the number of accurate significant figures in $a_{m,n-s}$. For example, if α_s, α_{s+1} are equal zeros and $\rho_{s-1}, \rho_{s+1} < 1$, then

$$a_{m+1,n-s} \sim \tfrac{1}{2}a_{m,n-s}^2, \tag{3·6}$$

and settling occurs when these two quantities are equal, to the number of accurate significant figures they possess. Again, if α_s, α_{s+1} are complex conjugates $r\,e^{\pm i\theta}$ and $\rho_{s-1}, \rho_{s+1} < 1$, then

$$a_{m+1,n-s} \sim \tfrac{1}{2}\cos 2M\theta \sec^2 M\theta\, a_{m,n-s}^2. \tag{3·7}$$

What is the smallest value of m for which the coefficients have settled? Clearly if $\rho_s < 1$, then with rare exceptions $a_{m,n-s}$ has settled when $|\epsilon_m| < \frac{1}{2}10^{-S}$, where S is the number of accurate significant figures in $a_{m,n-s}$. Now from equation (3·2) it is seen that $|\epsilon_m| < k\rho_s^M$, where $k \geqslant 1$ is independent of m. Hence when $\rho_s < 1$, settling occurs as a rule when

$$M > S/\log_{10}(\rho_s^{-1}) \tag{3·8}$$

(since for practical purposes $\log_{10}(2k)$ can be assumed to be negligible compared with S).

When $\rho_s = 1$ and $|\alpha_s| = |\alpha_{s+1}| = |\alpha_{s+2}| = \ldots = |\alpha_{s+\nu}|$, where $\nu \geqslant 1$, then either $s+\nu = n$ and $a_{m,n-s}$ is settled for all m, or $\rho_{s+\nu} < 1$. In the latter event similar reasoning shows that $a_{m,n-s}$ settles when

$$M > S/\log_{10}(\rho_{s+\nu}^{-1}). \tag{3·9}$$

With an *a priori* knowledge of the approximate ratios of the moduli of the zeros, the least value of m for which the coefficients have settled can be predicted from the relations (3·8) and (3·9). Generally, however, no such information is available, and the standard procedure is to continue until most of the coefficients are squaring upon each further transformation. Adjacent pairs of such coefficients correspond to real zeros, and pairs which are separated by a single oscillatory coefficient correspond to pairs of complex zeros. Usually, of course, not all the coefficients begin settling for the same value of m, and if there exists a set of zeros with nearly equal moduli, then the corresponding group of coefficients takes longer to settle and breaks away from the main body. In an effort to separate the moduli such groups may be further root-squared independently; their extreme coefficients square themselves with subsequent transformations. This method may work if the moduli involved are actually unequal and if during the process all the remaining significant figures in the coefficients are not lost by cancellation. The following alternative approach is recommended and believed to be simpler and more general, and it applies also to the case of equal moduli.

Suppose that after several transformations it is found that $a_{m+1,n-s} \neq a_{m,n-s}^2$ to the number of significant figures (S) present in $a_{m,n-s}$. Then either $\rho_s = 1$ or alternatively, if $\rho_s < 1$, then $a_{m,n-s}$ has not settled. In the latter event we deduce from (3·8) that $M\log_{10}(\rho_s^{-1}) < S$ and hence

$$1 - \rho_s < M^{-1} S \log_e 10. \tag{3·10}$$

By taking $S = 1$ we arrive at the following practical rule:

If $a_{m+1,n-s} \neq a_{m,n-s}^2$ even as far as the leading figures are concerned, then $|\alpha_s|$, $|\alpha_{s+1}|$ are equal to within about 2^{8-m} %.

For example, after eight transformations a group of consecutive non-squaring coefficients indicates the presence of zeros whose moduli are equal to within 1 %. The average modulus of such a group of zeros can be deduced approximately from the squaring coefficients bordering the group, and from a knowledge of this quantity to two or three significant figures it is generally possible to determine approximately the associated phases (see § 5 (*c*)) and then complete the numerical evaluation of these zeros by the iterative methods described in part B.

We turn now to the second difficulty mentioned at the beginning of this section, namely, the severe loss of significant figures due to cancellation. When polynomials of degrees exceeding ten are being root-squared it is not uncommon to lose two or three figures 'off the front' in each of the early transformations; by the time three or four transformations have been completed several coefficients may have disappeared entirely and the work must be repeated using more figures.

In order to understand this phenomenon, consider the application of the root-squaring transformation to the polynomial*

$$f(x) = \sum_{s=0}^{n} (-)^s (n\,|\,s)\, a^{n-s} x^s. \tag{3·11}$$

* For typographical reasons $(n\,|\,s)$ is used to denote the binomial coefficient $\binom{n}{s}$.

If the coefficients of the transformed polynomial $g(x)$ are denoted by b_s, then from equation (2·3) we obtain

$$b_s = a^{2n-2s}[(n\,|\,s)^2 - 2(n\,|\,s-1)\,(n\,|\,s+1) + 2(n\,|\,s-2)\,(n\,|\,s+2) - \ldots]. \quad (3{\cdot}12)$$

Now since the zeros of $g(x)$ are all equal to $-a^2$, the sum of the terms in the squared bracket in equation (3·12) must equal $(n\,|\,s)$. If $n-1 \geqslant s \geqslant 1$, some individual terms are considerably greater than $(n\,|\,s)$, showing that if the process were performed numerically partial cancellation would take place; in fact, since $(n\,|\,s)^2 > (n\,|\,s-r)\,(n\,|\,s+r)$ if $r>0$, the number of leading figures which cancel is approximately equal to the number of digits in $(n\,|\,s)$ itself. Further transformations produce the same effect.

Returning to the general polynomial of the nth degree, there is no reason to suppose that systematic cancellation will occur when the zeros are widely separated, and the inference is that it takes place only when some zeros are either equal or close together in the complex plane. This is verified in practice. Even with a few moderately close zeros some cancellation usually occurs during the early transformations. The discussion above indicates the *maximum* amount of cancellation that can happen in the worst cases. For example, not more than three figures should be lost per transformation if $n \leqslant 12$; the corresponding limits for four and five figures being $n \leqslant 15$, 19 respectively. Unless some of the zeros are actually equal these maximum values will be attained, if at all, only in the early transformations, because the zeros of each transformed polynomial are more widely separated than those of its predecessors; eventually there is no further cancellation.

In practice it is important to be able to decide in advance how many significant figures need be retained in the root-squaring computations in order to produce desired accuracy in the final results. With no *a priori* knowledge of the approximate distribution of the zeros the suggested policy is to perform one or two transformations retaining as many figures (usually ten) as can be set conveniently on the calculating machine and to note the cancellation that occurs. If this is severe, it may be assumed that several zeros are close together, and similar cancellations may be expected during the next few transformations. It may then be necessary to repeat the first one or two transformations retaining double the number of figures that can be set directly on the machine. This does not involve any major change in technique in the computations, and the extra labour can be tolerated for a few transformations.

It will be noticed that the slow convergence to a settled state and loss of significant figures owing to cancellation have similar basic causes: the former occurs when several zeros have moduli which are close together, and the latter occurs when, in addition, several zeros are close together in the complex plane. There is more than academic interest in the above discussion of these two main difficulties; they are by no means confined to pathological cases.

As an example* the following polynomials $f(x)$ of degree $2n$ arise in cable unit theory:

$$f(x) = \lambda P(x) + Q(x),$$

where $$P(x) = x(\alpha x^2+1)^{n-1} \sinh n\gamma \operatorname{cosech} \gamma, \quad Q(x) = (\alpha x^2+1)^n \cosh n\gamma,$$

γ being given by $$\cosh \gamma = (\alpha x^2+1)^{-1}\{(\alpha+\tfrac{1}{2})\,x^2+1\}. \quad (3{\cdot}13)$$

* The author is indebted to the British Thomson-Houston Co. for permission to publish this example.

Values of the parameters of physical interest include $n = 6, 8, 10, 12$; $\alpha = 0$ to 1; $\lambda = 0{\cdot}8$ to $1{\cdot}5$. Several of these polynomials have been solved, and with practically all of them the cancellation in the root-squaring process is severe. The polynomial (2·6) comes from the particular case $n = 8$, $\alpha = 0{\cdot}1$, $\lambda = 1$ of this example, and from table 3 it is seen that there is a loss of two significant figures in many of the coefficients in the third transformation. In all, seven transformations were required to separate the zeros to four-figure accuracy, and by that stage nine figures had been lost in about half the coefficients. The zeros of this polynomial to five decimal places are

$$\left.\begin{array}{lll} -0{\cdot}29351 \pm 0{\cdot}14350\mathrm{i}, & -0{\cdot}22447 \pm 0{\cdot}45093\mathrm{i}, & -0{\cdot}14762 \pm 0{\cdot}77176\mathrm{i}, \\ -0{\cdot}09004 \pm 1{\cdot}06119\mathrm{i}, & -0{\cdot}05086 \pm 1{\cdot}29691\mathrm{i}, & -0{\cdot}02567 \pm 1{\cdot}47438\mathrm{i}, \\ -0{\cdot}01049 \pm 1{\cdot}59630\mathrm{i}, & -0{\cdot}00249 \pm 1{\cdot}66712\mathrm{i}, & \end{array}\right\} \quad (3{\cdot}14)$$

and would not, on general grounds, be considered to be particularly close together. If, in fact, the zeros of a given polynomial are much closer than these, their accurate evaluation becomes a really acute computational problem. For example, in the case $n = 8$, $\alpha = 10$, $\lambda = 1$ the zeros of $f(x)$ have been calculated to four decimal places as follows:

$$\left.\begin{array}{lll} -0{\cdot}1324 \pm 0{\cdot}1360\mathrm{i}, & -0{\cdot}0187 \pm 0{\cdot}2530\mathrm{i}, & -0{\cdot}0023 \pm 0{\cdot}2926\mathrm{i}, \\ -0{\cdot}0005 \pm 0{\cdot}3042\mathrm{i}, & -0{\cdot}0001 \pm 0{\cdot}3086\mathrm{i}, & -0{\cdot}0000 \pm 0{\cdot}3107\mathrm{i}, \\ -0{\cdot}0000 \pm 0{\cdot}3117\mathrm{i}, & -0{\cdot}0000 \pm 0{\cdot}3122\mathrm{i}. & \end{array}\right\} \quad (3{\cdot}15)$$

Here the root-squaring process fails to determine all the moduli even if twenty figures are retained in the initial transformations. To solve such a difficult polynomial satisfactorily it must first be transformed into a more convenient form, by methods discussed later in § 9.

It may be remarked that although the polynomials (3·13) are ill-conditioned in the sense that *arbitrary* small changes in the coefficients produce large changes in some of the zeros, the zeros are not unduly sensitive to arbitrary small changes in the parameters α and λ. This example illustrates how the zeros of even very ill-conditioned polynomials can be meaningful physically and demonstrates the importance of studying the solution of such polynomials, quite apart from any intrinsic interest attached to the problem.

(4) *Methods available for evaluating the phases*

Provided an adequate number of significant figures has been retained in the calculations, the root-squaring process furnishes approximations to the moduli of the zeros, including, as demonstrated above, the multiple moduli. The accuracy of the results depends on the extent of the cancellation that has taken place and cannot usually be predicted. The problem discussed here is the determination of the phases or, equivalently, real parts of the zeros using the approximate values of the moduli.

In the case of a modulus which has been 'isolated' by the root-squaring process (that is, one obtained from consecutive squaring coefficients), the corresponding zero is necessarily real, and the phase determination merely amounts to ascertaining the correct sign; this is most speedily effected by trial.

Multiple moduli (including of course the moduli corresponding to conjugate pairs of complex zeros) present a more profound problem unless their total number does not exceed

four, that is, at least all but four of the zeros are real and simple and have been isolated by the root-squaring process. In the latter event the phases can be determined by elementary methods, using the known sums of the zeros and of their reciprocals.* Three methods have been proposed for solving the general problem. We shall consider first their application to isolated conjugate pairs of complex zeros; the case of three or more unseparated moduli is deferred until § 5 (*c*).

I. *Methods based directly on the root-squaring computations*

Using the notation of § 3, suppose that in the final transformed polynomial $f_m(x)$ the triad of terms corresponding to a pair of zeros $r\,e^{\pm i\theta}$ of $f(x)$ is

$$a_{m,s}x^s + a_{m,s+1}x^{s+1} + a_{m,s+2}x^{s+2}.$$

Then from equations (3·2) and (3·5), it follows that

$$r^{2M} = a_{m,s}/a_{m,s+2} \quad \text{and} \quad 2r^M \cos M\theta = a_{m,s+1}/a_{m,s+2}, \tag{4·1}$$

whence r and $\cos M\theta$ can be evaluated. Knowing the latter, we can obtain $\cos\theta$ as one of M possible values. The ambiguity could be resolved by substitution in $f(x)$ but the work involved is excessive. The following two variations of this method, due to Graeffe (1837) and Carvallo (see, for example, Ostrowski 1940, p. 243) respectively, greatly reduce this labour.

(i) Knowing
$$\cos M\theta = \tfrac{1}{2}a_{m,s+1}/(a_{m,s}a_{m,s+2})^{\frac{1}{2}} = c, \quad \text{(say)}, \tag{4·2}$$

we can obtain
$$\cos \tfrac{1}{2}M\theta = \pm\{\tfrac{1}{2}(c+1)\}^{\frac{1}{2}}. \tag{4·3}$$

The determination of the correct sign can be effected by trial, since $-r^{\frac{1}{2}M}\,e^{\frac{1}{2}Mi\theta}$ is a zero of $f_{m-1}(x)$. Similarly, by testing in $f_{m-2}(x)$ the value of $\cos \tfrac{1}{4}M\theta$ can be determined, and by repeating the process m times the value of $\cos\theta$ is obtained. The number of possible ambiguities to be tested is thus reduced by this technique from 2^m to m.

(ii) Write
$$\left.\begin{aligned}\phi_s(x) &\equiv a_{s,0} - a_{s,2}x + a_{s,4}x^2 - \ldots = A_{s,0} + A_{s,2}x + A_{s,4}x^2 + \ldots,\\ \psi_s(x) &\equiv -a_{s,1} + a_{s,3}x - a_{s,5}x^2 + \ldots = A_{s,1} + A_{s,3}x + A_{s,5}x^2 + \ldots,\end{aligned}\right\} \tag{4·4}$$

where A's denote the modified coefficients of § 2. Then

$$f_{m-1}(x) = \phi_{m-1}(-x^2) - x\psi_{m-1}(-x^2). \tag{4·5}$$

Define $x_0 = r\,e^{i\theta}$ and $x_{s+1} = -x_s^2$ if $s \geqslant 0$, so that x_s is a zero of $f_s(x)$. Substituting $x = x_{m-1}$ in (4·5), we obtain

$$x_{m-1} = \frac{\phi_{m-1}(-x_{m-1}^2)}{\psi_{m-1}(-x_{m-1}^2)} = \frac{\phi_{m-1}(x_m)}{\psi_{m-1}(x_m)}. \tag{4·6}$$

Since x_m is known, this formula gives x_{m-1}. Similarly, we can obtain $x_{m-2}, x_{m-3}, \ldots$ and finally x_0.

In practice it is best to combine the two techniques, for although (ii) is speedier, (i) is the more accurate; there is a steady gain in accuracy on taking successive square roots, whereas in (ii) there is gradual accumulation of rounding errors. The suggested combined method is as follows:

* See, for example, Whittaker & Robinson (1944, § 57). These authors remark that 'a similar method ...may be applied to solve equations with more than two pairs of complex roots', but the extension is by no means trivial.

(iii) Starting with $x_m = -r^M e^{Mi\theta}$, computed from (4·1), we form the successive powers $x_m^2, x_m^3, \ldots$, from which, using (4·4), $\phi_{m-1}(x_m)$ and $\psi_{m-1}(x_m)$ can be computed. Then x_{m-1} is obtained from (4·6), and this value is compared with that obtained from

$$x_{m-1} = r^{\frac{1}{2}M} e^{i\theta_{m-1}}, \quad \text{where} \quad \theta_{m-1} = \tfrac{1}{2}\theta_m \pm \tfrac{1}{2}\pi, \tag{4·7}$$

θ_m being the known phase of x_m. The correct sign in equation (4·7) is determined by this comparison. Indeed, equation (4·6) is best regarded merely as a means of resolving the ambiguity in (4·7), and it is usually sufficient to perform the calculations associated with it to three- or four-figure accuracy.

II. *Methods depending on a change of origin of the root-squaring process*

The principle of these methods has been described by Graeffe (1837) and Ostrowski (1940). The polynomials $f(x)$ and $f(x+a)$, where a is a suitably chosen real number, are each root-squared; the zeros of $f(x)$ are then located at the intersections of circles centred at $x = 0$ and $x = a$ with radii equal to the moduli. A practical drawback of this unrefined technique results from the presence of more than one pair of complex zeros; additional work is required to decide which intersections really do correspond to zeros. If there are ν pairs of complex zeros then there are correspondingly ν circles centred at $x = 0$, an equal number at $x = a$, and these may intersect at as many as ν^2 pairs of points. Graeffe pointed out, however, that no ambiguity will arise if a is chosen to be numerically less than half the *least* difference between the distinct moduli. The labour of this process, apparently the most economical of this kind, is equal to that already expended in the root-squaring of $f(x)$.

Another remedy for the defect was suggested by Brodetsky & Smeal (1924). They showed that the ambiguity could also be removed by making an *infinitesimal* change of origin ϵ, retaining ϵ as a symbol in the calculations and ignoring $O(\epsilon^2)$. The additional labour entailed is double that expended in the root-squaring of $f(x)$ and the method is inferior to that of Graeffe.

III. *Highest common factor method*

This also is due in principle to Graeffe (1837). It is supposed* that $f(x)$ is of even degree $2n$. Let $x^2 - qrx + r^2$ be the quadratic factor corresponding to the known modulus r. Formal division of $f(x)$ by this factor, retaining q as a symbol, gives

$$f(x) = (x^2 - qrx + r^2)\, Q(x) + xJ(q) + K(q), \tag{4·8}$$

where $J(q)$, $K(q)$ are polynomials in q of degrees $2n-1$, $2n-2$ respectively, whose coefficients depend on the coefficients of $f(x)$ and on r. Clearly the required value of q must be such that

$$J(q) = K(q) = 0. \tag{4·9}$$

Moreover, since we have supposed there is only one pair of complex zeros of modulus r, there can be only one† real value of q satisfying (4·9) and such that $-2 < q < 2$. Thus q may be obtained from the highest common factor of $J(q)$ and $K(q)$.

It is possible to perform quite formally about half the steps of the usual process for finding the common factor and by these means the following theorem may be established:

* Polynomials of odd degree can be included by multiplying them by x or by dividing them by a linear factor beforehand.

† Even if $f(x)$ has only two zeros of modulus r, it is possible for $J(q)$ and $K(q)$ to possess more than one common zero. They can, however, have only one common zero *in the range* $-2 < q < 2$. See also §5(*a*).

If $x^2-qrx+r^2$ is a quadratic factor of the polynomial

$$f(x) = a_0+a_1x+a_2x^2+\ldots+a_{2n}x^{2n}, \tag{4·10}$$

then q is a common zero of the polynomials

$$U_n(q) = \sum_{s=0}^{n} u_{n,s}\,q^s, \quad U_{n-1}(q) = \sum_{s=0}^{n-1} u_{n-1,s}\,q^s, \tag{4·11}$$

where

$$\left.\begin{aligned} u_{n,s} &= \beta_s-\binom{s+1}{1}\beta_{s+2}+\binom{s+2}{2}\beta_{s+4}-\binom{s+3}{3}\beta_{s+6}+\ldots,\\ u_{n-1,s} &= \gamma_s-\binom{s+1}{1}\gamma_{s+2}+\binom{s+2}{2}\gamma_{s+4}-\binom{s+3}{3}\gamma_{s+6}+\ldots, \end{aligned}\right\} \tag{4·12}$$

and

$$\beta_s = a_{n+s}r^{n+s}-a_{n-s-2}r^{n-s-2}, \quad \gamma_s = a_{n+s+1}r^{n+s+1}-a_{n-s-1}r^{n-s-1}. \tag{4·13}$$

The following points should be noted:

(i) It is supposed in (4·13) that values of a_s are to be replaced by zeros whenever $s<0$ or $s>2n$. Thus

$$\beta_{n-1} = a_{2n-1}r^{2n-1}, \quad \beta_n = a_{2n}r^{2n}, \quad \gamma_n = 0.$$

(ii) It is not possible to continue formally the process for determining the common factor beyond $U_{n-1}(q)$ and still obtain convenient expressions for the coefficients of the resulting polynomials.

(iii) The theorem has not previously appeared in the above form, but it can be deduced from the work of Graeffe and Encke, an account of which is given by Bairstow (1914). In the case we are considering here, a short proof is obtained by making the substitution $q = 2\cos\theta$ in the imaginary parts of each of the two relations

$$\sum_{s=0}^{2n} a_s r^s\,\mathrm{e}^{(s-n)\mathrm{i}\theta} = 0, \quad \sum_{s=0}^{2n} a_s r^s\,\mathrm{e}^{(s-n+1)\mathrm{i}\theta} = 0. \tag{4·14}$$

In the practical application of this result the coefficients $u_{n,s}$, $u_{n-1,s}$ are computed from equations (4·12) and (4·13) and checked by the formulae

$$\left.\begin{aligned} \sum_{s=0}^{n}(\beta_s-\gamma_s) &= a_nr^n+a_{n-1}r^{n-1},\\ \text{and}\quad \sum_{s=0}^{n} u_{n,s} &= \beta_0+\beta_1-\beta_3-\beta_4+\beta_6+\beta_7-\beta_9-\beta_{10}+\ldots,\\ \sum_{s=0}^{n-1} u_{n-1,s} &= \gamma_0+\gamma_1-\gamma_3-\gamma_4+\gamma_6+\gamma_7-\gamma_9-\gamma_{10}+\ldots. \end{aligned}\right\} \tag{4·15}$$

The common factor is then extracted numerically. One of the quickest routines* for doing this is to form the succession of polynomials

$$U_k(q) = \sum_{s=0}^{k} u_{k,s}\,q^s \quad (k = n-2,\ n-3,\ \ldots,\ 0),$$

related by

$$U_k(q) = U_{k+2}(q)+(\mu_kq+\lambda_k)\,U_{k+1}(q), \tag{4·16}$$

where

$$\mu_k = -u_{k+2,k+2}/u_{k+1,k+1}, \quad \lambda_k = -(u_{k+2,k+1}+\mu_ku_{k+1,k})/u_{k+1,k+1}. \tag{4·17}$$

Thus

$$u_{k,s} = u_{k+2,s}+\mu_ku_{k+1,s-1}+\lambda_ku_{k+1,s}. \tag{4·18}$$

* This is equivalent to the school text-book process but saves a certain amount of recording.

As checks at each stage we may verify that $u_{k,k+1}$ and $u_{k,k+2}$ (as given by (4·18)) both vanish, and that

$$\sum_{s=0}^{k} u_{k,s} = \sum_{s=0}^{k+2} u_{k+2,s} + (\mu_k + \lambda_k) \sum_{s=0}^{k+1} u_{k+1,s}. \tag{4·19}$$

Assuming $U_n(q)$, $U_{n-1}(q)$ have only one* common zero, the value of q is given by

$$q = -u_{1,0}/u_{1,1}. \tag{4·20}$$

The identity $u_{0,0} = 0$ provides another check.

An example of the computational arrangement of this method is given in § 5 (*d*).

It may be mentioned here that the common factor theorem given above remains valid whatever the nature of the zeros of $x^2 - qrx + r^2$. They could be real ($q \geqslant 2$ or $q \leqslant -2$), or non-conjugate (q or r or both complex).

(5) *Comparison of methods for evaluating the phases*

(*a*) *Accuracy*

Methods I and II generally produce as many correct decimals in $\cos\theta$ or q as there are significant figures available in r. The accuracy of method III, if used without refinement, cannot exceed this and may be much less owing to loss by cancellation of significant figures in the coefficients of the polynomials $U_k(q)$. Occasionally a devastating loss of figures occurs in all the coefficients of one polynomial of this sequence, a phenomenon which can be quite simply explained.

Suppose, for example, $f(x)$ has a pair of real zeros x_1, x_2 such that their geometric mean equals r. Then in addition to the factor $x^2 - qrx + r^2$ we are seeking, $f(x)$ has another factor of the same form with q replaced by q_1 where q_1 is also a real number but $|q_1| \geqslant 2$. Thus $U_n(q)$ and $U_{n-1}(q)$ have two common zeros and $U_1(q)$ vanishes identically. More commonly, $f(x)$ may have a pair of real zeros with a geometric mean *approximately* equal to r, and the disappearance of $U_1(q)$ will be incomplete, partial cancellation taking place on forming its coefficients. Similarly, if $f(x)$ has two pairs of zeros $r_1 e^{\pm i\theta_1}$, $r_2 e^{\pm i\theta_2}$ which approximately satisfy $\theta_1 = \theta_2$ and $r_1 r_2 = r^2$, then a loss of significant figures will occur on forming the coefficients of $U_2(q)$.

Fortunately, this drawback to method III is not serious, and it is possible to regain the lost accuracy fairly simply. The way this is done depends on the extent of the cancellation that has taken place.

(i) Suppose that three or four decimals are available in the value of q obtained from the common factor process. The corresponding phase θ is then obtained from

$$\theta = \cos^{-1}(\tfrac{1}{2}q). \tag{5·1}$$

Now, using (4·2), we can compute $\cos M\theta \equiv c$ from the root-squaring calculations, and we then have

$$\theta = (\pm\cos^{-1} c + 2\nu\pi)/M, \tag{5·2}$$

where ν has one of the values $0, 1, 2, \ldots, \frac{1}{2}M$. The substitution in this equation of the value of θ obtained from (5·2) usually determines the correct value of ν and the correct sign. It is then possible to obtain $q = 2\cos\theta$ correct to as many decimals as there are figures available in r.

* The case of several common zeros is considered in §§ 5 (*a*) and (*c*).

(ii) If the available value of q is insufficiently accurate to resolve the ambiguities in equation (5·2), then we can revert to the last polynomial of the sequence $\{U_s(q)\}$ in which no considerable cancellation has occurred, and obtain a better value of q by iteration (see § 7), since q must be a zero of this polynomial. Using the more accurate value we can then proceed as in (i).

(iii) Finally, the cancellation in one of the polynomials, say $U_{k-1}(q)$, may be so severe that the common factor process cannot be continued. This apparent failure may in fact be turned to considerable advantage, for if the polynomial $U_k(q)$ (which is of degree less than half that of $f(x)$) is solved numerically, each real or complex zero q will correspond to an approximate quadratic factor $x^2-qrx+r^2$ of $f(x)$, and we can quickly obtain* $2k$ zeros of $f(x)$.

Summarizing, if we regard methods I, II and III merely as means of resolving the ambiguities in the formula (5·2), then they are of equivalent accuracy. As a corollary, it is often possible to apply them keeping comparatively few figures in the calculations and yet still obtain as many decimals in q at the finish as there are significant figures in r.

(*b*) *Speed*

The labour entailed by method II is independent of the number of zeros that happen to be real. With methods I and III, however, the labour involved is proportionate to the number of complex zeros. It is true that method II automatically determines the signs of the real zeros, but this is a relatively small gain since these signs can be quickly resolved by trial. Even in the case most favourable to II when all the zeros are complex, it has been found to be quicker in practice to employ methods I and III, partly because with these methods fewer significant figures need be retained in the bulk of the calculations.

Comparing methods I and III, we note that with the former the evaluation of each value of q requires a time which is roughly proportional to the product of the degree of $f(x)$ and the number of root-squaring transformations required for separation, whereas with III the time varies roughly as the square of the degree of $f(x)$. In practice the former time is generally slightly less than the latter. There is, however, a further consideration completely outweighing this—an advantage of method III not shared by I.

After each quadratic factor $x^2-qrx+r^2$ has been evaluated, and, if necessary, iterated in $f(x)$ (see § 7), it may be divided out of the polynomial. A succession of polynomial quotients is thereby formed, each of degree two less than its immediate predecessor and each of which may be used in place of $f(x)$ in the common factor process. The labour required for the evaluation of each q thus decreases rapidly as the quadratic factors are successively removed. In addition, a further saving of labour may be made by removing beforehand the linear factors corresponding to the real zeros.

One point only of the procedure just described needs special care. Although the polynomial quotients obtained by successive removal of the accurate quadratic factors are sufficiently accurate for use in conjunction with the root-squaring computations for determining q, they must not be used for *iterating* the quadratic factors (using the methods of § 7) if greater accuracy is required. The original polynomial $f(x)$ should always be used for any

* This phenomenon actually occurred in the solution of a sixteenth degree polynomial and it was possible to obtain *three* pairs of complex zeros with *one* particular application of the process.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

necessary iteration of the linear or quadratic factors, otherwise the accumulation of rounding errors may prove very troublesome.

(c) Multiple or nearly equal moduli

Discussion has, so far, been restricted to the evaluation of phases of isolated conjugate pairs of zeros. If there are three or more zeros whose moduli remain unseparated after a reasonable number of root-squaring transformations, an approximation to the average modulus is available (cf. § 3), and we now consider how the associated phases may be found.

Method III can be applied to such problems with little modification of the computational procedure. If the number of zeros involved exceeds three, then the polynomials $U_n(q)$, $U_{n-1}(q)$ have at least two common zeros in the range $-2 \leqslant q \leqslant 2$, and the common factor process terminates at $U_k(q)$ $(k \geqslant 2)$, the zeros of which correspond to quadratic factors of $f(x)$.

Method II is also applicable with little modification, but method I would require substantial alterations in procedure.

(d) Conclusion and example

From the general standpoint of accuracy, speed and simplicity, method III (with the modifications given in (*a*) and (*b*)) is the most efficient of the available methods for the evaluation of the phases of the complex zeros.

TABLE 4. PHASE DETERMINATION ($r^2 = 1{\cdot}6845$)

s	a'_s	$a'_s r_s$	β_s	γ_s	$u_{4,s}$	$u_{3,s}$	$u_{2,s}$	$u_{1,s}$	$u_{0,s}$
0	52·72957	52·72957	+20·220	−4·053	− 0·024	+ 0·124	− 0·023	+0·125	−0·003
1	4·95768	6·4346	+3·2551	−71·255	− 1·2597	+ 1·475	− 1·1886	+1·599	0·000✓
2	95·33607	160·594	+36·609	−4·1772	−12·486	− 4·1772	−11·866	0·000✓	0·000✓
3	6·28617	13·743	+2·2574	−36·365	+ 2·2574	−36·365	0·000✓	0·000✓	
4	63·72316	180·814	+16·365		+16·365		0·000✓		
5	2·63108	9·6897							
6	18·69090	89·339	$\Sigma(\beta_s-\gamma_s) = +194{\cdot}557$✓		$\Sigma = +\ 4{\cdot}853$✓	−38·943✓	−13·078✓	+1·724✓	
7	0·36389	2·2574				$\mu = +\ 0{\cdot}45002$	− 3·0646	+7·421	
8	2·03253	16·365				$\lambda = +\ 0{\cdot}010383$	− 0·04505	+0·163	

Hence $q = -0{\cdot}0782$, $\cos^{-1}(\frac{1}{2}q) = 92{\cdot}24$ degrees.

From root-squaring computations $\begin{cases} \cos 64\theta = -0{\cdot}806, & \theta = 92{\cdot}245 \text{ degrees}, \\ q = 2\cos\theta = -0{\cdot}07835, & p = qr = -0{\cdot}10169. \end{cases}$

An example of the arrangement of the computations using method III is given in table 4. The polynomial is that given by (2·6), and we are here calculating the value of q corresponding to $r^2 = 1{\cdot}6845$. The factors

$$x^2 + 0{\cdot}58701x + 0{\cdot}10674, \quad x^2 + 0{\cdot}44894x + 0{\cdot}25372,$$

$$x^2 + 0{\cdot}29525x + 0{\cdot}61740 \quad \text{and} \quad x^2 + 0{\cdot}18008x + 1{\cdot}13424$$

have been checked by iteration in $f(x)$, and then divided out successively to leave the quotient $a'_0 + a'_1 x + a'_2 x^2 + \ldots + a'_8 x^8$, the coefficients of which are used in (4·13) (taking $n = 4$) in place of those of $f(x)$, in accordance with the procedure suggested in (*b*) above. The accuracy of the value obtained for q has been improved by the device described in (*a*) (i) above.

Part B. Indirect methods

(6) *The use of indirect methods*

We turn now to a discussion of iterative methods. Given some approximation to a zero, an iterative process if successful must produce a better approximation. Except in special cases (such as the well-known procedure for evaluating a square root), convergence is not guaranteed. Both the possibility and the speed of convergence depend in general upon the accuracy of the first approximation.

Iterative methods are therefore most useful when a fair approximation is already known. In rare cases this can be obtained from physical or intuitive reasoning. More commonly, a direct method has first been applied and has produced a first approximation, the accuracy of which may be unknown, or known to fall short of specified requirements. Iteration can then be used to improve the accuracy, and we shall regard such methods as supplementary, rather than as alternative, to the direct methods discussed in part A.

In this approach we differ from other authors, notably Jelinek (1865, see also Bairstow 1914, p. 63), Frazer & Duncan (1929) and Shih-Nge Lin (1943), who have suggested the use of certain iterative processes using arbitrary initial approximations, dispensing entirely with direct methods. In practice it is found that for polynomials of high degree such methods do not always converge, and even when successful they usually entail much more labour than root-squaring followed (if necessary) by iteration.*

A measure of the power of an iterative process can be taken as the *degree of convergence*, defined by Bodewig (1946*b*, 1949) and Hartree† (1949). If α is a simple or multiple zero of $f(x)$, a is an approximation to α and $\eta \equiv \alpha - a$, then an iteration formula of the form

$$\eta = F(a) + O(|\eta|^k) \quad (k \geqslant 1)$$

is said to be *convergent to the kth degree*.

If $k > 1$, the characteristic feature of such a formula is that once the approximation a is sufficiently accurate, each subsequent application increases the number of correct figures in the ratio k to 1. Among formulae of equal degrees of convergence, variations in power occur only through differing values of the constant implied in the error term $O(|\eta|^k)$, and in practice such variations are usually unimportant; the most useful formula of given degree will generally be that with the simplest form of $F(a)$.

Formulae for which $k = 1$ are termed linearly convergent and are best avoided; convergence is either very slow or non-existent.

The root-squaring process, although it is not an iterative process, is convergent to the second degree in the sense that the number of accurate figures in the estimates for the moduli eventually doubles with each subsequent transformation, until it equals the number of working figures. Bodewig (1946*a*) has remarked that this process therefore has greatest efficiency if it is carried out to many decimal places. This is often true, but the labour can be very great if cancellation is severe, necessitating 'double-length' arithmetic on the

* A possible exception to this statement is Laguerre's formula applied to polynomials whose zeros are all real (see §7).

† Hartree uses the term 'order' in place of 'degree'.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

calculating machine. In such cases it is best not to exceed the machine capacity if possible, but merely to find an approximation by root-squaring and improve it by the use of iterative formulae of which the most common are also quadratically convergent.

It will not be necessary to consider the region of convergence of formulae of iteration. In nearly all cases the approximations obtained by root-squaring lie well within these regions. In the most unfavourable case of zeros of extraordinarily close proximity, special devices (see § 9) can be used for their separation, and iteration can then complete their evaluation.

(7) *Isolated simple zeros*

In this section we suppose that α is a simple zero (real or complex) of $f(x)$ and a is an approximation to α such that $\eta \equiv \alpha - a$ is small compared with the distances from α of the other zeros of $f(x)$. We shall discuss various formulae of iteration, their uses and relative merits.

(*a*) *Quadratically convergent formulae.* The principal formula of this class is due to Newton:

$$\eta = -\frac{f(a)}{f'(a)} + O(|\eta|^2). \tag{7·1}$$

(*b*) *Cubically convergent formulae.* One of the simplest of such formulae is

$$\eta = -\frac{f(a)}{f'(a)} - \frac{\{f(a)\}^2\{f''(a)\}}{2\{f'(a)\}^3} + O(|\eta|^3). \tag{7·2}$$

Another formula is that of Laguerre:

$$\eta = -\frac{nf(a)}{f'(a) \pm \{H(a)\}^{\frac{1}{2}}} + O(|\eta|^3), \tag{7·3}$$

where n is the degree of $f(x)$ and

$$H(a) = (n-1)^2\{f'(a)\}^2 - n(n-1)\, f(a)\, f''(a).$$

The interesting feature of this formula is that if $f(x)$ is a polynomial whose zeros are all real and distinct, then starting from an *arbitrary* real value a the successive application of (7·3) produces two sequences which converge to the nearest zero greater than a and to the nearest zero less than a. Laguerre's formula and its extensions to multiple real zeros have been fully discussed in papers by Bodewig (1946*b*) and van der Corput (1946).

(*c*) *Formulae of higher degrees of convergence.* Examples of these may be constructed by reverting the Taylor expansion

$$f(a) + \eta f'(a) + \frac{1}{2!}\eta^2 f''(a) + \ldots + \frac{1}{n!}\eta^n f^{(n)}(a) = 0. \tag{7·4}$$

Relative merits of the use of formulae (*a*), (*b*) *and* (*c*)

Formulae (7·1) and (7·2) are very powerful and there is little difference in the labour they involve. Suppose, starting from d correct significant figures in the zero, it is desired to determine Nd correct figures. Then the approximate numbers of steps required using formulae (7·1) and (7·2) are $\log_2 N$ and $\log_3 N$, respectively. Thus the approximate numbers of polynomials to be computed using the two methods are $2\log_2 N$ and $3\log_3 N$, and these quantities have the ratio

$$\frac{2\log_e 3}{3\log_e 2} = 1{\cdot}06\ldots.$$

Formula (7·1) is to be preferred, on the grounds of simplicity and the ease with which errors are detected and corrected. In this respect it is superior also to the formulae of class (*c*), notwithstanding their higher rate of convergence.

The practical application of quadratically convergent iteration formulae to isolated simple real zeros

When a and η are real, the best procedure is the straightforward application of (7·1). The polynomials $f(a)$ and $f'(a)$ are most conveniently computed on a transfer calculating machine using a method well known to computers, consisting of repeating cyclically the operations of multiplication, addition and transference as indicated by the sequence

$$q_n = a_n, \quad q_s = aq_{s+1} + a_s \quad (s = n-1, n-2, \ldots, 0),$$

$a_0, a_1, \ldots, a_n$ being the coefficients of $f(x)$. Then $f(a) = q_0$. The quantities $q_n, q_{n-1}, \ldots, q_1$ are the coefficients of the polynomial quotient which results on dividing $f(x)$ by $x-a$; they appear incidentally on the machine but need not be recorded unless required in another connexion.

One point deserves special mention. It has been frequently stated (see, for example, Whittaker & Robinson 1944, § 48) that after one or two applications of (7·1) it becomes unnecessary to recompute $f'(a)$ since this quantity suffers relatively small changes in the subsequent steps. This practice throws away the advantage of the quadratic nature of the convergence; instead of doubling the number of correct figures with each subsequent step, only a fixed number (equal to the number of leading figures of $f'(a)$ and $f'(\alpha)$ that agree) can be added. To secure the maximum advantage, $f'(a)$ must be recomputed at each step, the only exception being in the *last* step when the full gain associated with the formula may not be required.

The practical application of quadratically convergent iteration formulae to isolated complex zeros

The direct application of (7·1) involves the evaluation of polynomials for complex values of the variable a. The least laborious way of effecting this numerically is to divide the polynomials by the real quadratic factor corresponding to a and its conjugate $\bar{a}$. We shall denote this factor by x^2-px-l, so that

$$p = 2\mathscr{R}a, \quad l = -|a|^2. \tag{7·5}$$

Let the results of dividing $f(x)$ and $f'(x)$ by this factor be written* in the forms

$$\left.\begin{aligned} f(x) &= (x^2-px-l)\,q(x) + q_1 x + q_0', \\ f'(x) &= (x^2-px-l)\,r(x) + r_1 x + r_0'. \end{aligned}\right\} \tag{7·6}$$

Then if $a \equiv b+\mathrm{i}c$, $\eta \equiv \delta b + \mathrm{i}\delta c$, we obtain from (7·1) and (7·6) a correction

$$\delta b + \mathrm{i}\delta c = -\frac{q_1(b+\mathrm{i}c)+q_0'}{r_1(b+\mathrm{i}c)+r_0'}.$$

Hence, using (7·5), we obtain

$$\left.\begin{aligned} A\delta b &= q_1 r_1 l - b(q_0' r_1 + q_1 r_0') - q_0' r_0', \\ A\delta c &= (q_0' r_1 - q_1 r_0')\,c, \end{aligned}\right\} \tag{7·7}$$

where

$$A = r_0'^2 + pr_0' r_1 - lr_1^2.$$

* The primes on q_0', r_0' are introduced here for convenience in a later context. See (7·11) and (7·12).

Complex numbers can be avoided completely if, instead of the pairs of complex zeros, we seek the corresponding quadratic factors, and this is to be preferred. The zeros can, of course, be immediately derived from the factors, but in many problems the quadratics themselves are of predominant interest.

Let δp, δl denote the changes in p and l corresponding to the changes δb, δc. Then neglecting $O(|\eta|^2)$, we find that

$$\delta p = 2\delta b, \quad \delta l = -2b\delta b - 2c\delta c,$$

and therefore

$$\left.\begin{aligned} A\delta p &= (2lr_1 - pr_0')\, q_1 - (pr_1 + 2r_0')\, q_0', \\ A\delta l &= -(pr_1 + 2r_0')\, lq_1 + (2lr_1 + p^2 r_1 + pr_0')\, q_0', \end{aligned}\right\} \tag{7·8}$$

A being given by equation (7·7). For complex zeros this is the most readily computable form of Newton's rule.

A simpler formula was given by Bairstow (1914). Let the results of dividing $f(x)$ twice in succession by $x^2 - px - l$ be denoted by

$$f(x) = (x^2 - px - l)\, q(x) + q_1 x + (q_0 - pq_1), \tag{7·9}$$

$$q(x) = (x^2 - px - l)\, T(x) + T_1 x + (T_0 - pT_1). \tag{7·10}$$

Then Bairstow's formula (in a form slightly modified from the original) is given by

$$\left.\begin{aligned} D\delta p &= T_1 q_0 - T_0 q_1, \quad D\delta l = Mq_1 - T_0 q_0, \\ M &= lT_1 + pT_0, \qquad D = T_0^2 - MT_1. \end{aligned}\right\} \tag{7·11}$$

where

Simple checks on the computation of (7·11) are given by

$$T_1\delta l + T_0\delta p = -q_1, \quad T_0\delta l + M\delta p = -q_0.$$

It is easily verified that (7·11) is equivalent to (7·8) with $O(|\eta|^2)$ neglected, so that Bairstow's formula is quadratically convergent. Bairstow's own method of proof shows that this is still true even if the zeros of $x^2 - px - l$ are not complex conjugates, provided they are distinct. Comparing (7·8) with (7·11), we see that slightly less labour is entailed in the application of the latter, the formula is simpler, there is no necessity to form the coefficients of $f'(x)$, and the degree of $q(x)$ is one less than that of $f'(x)$. On these grounds formula (7·11) has the advantage.

In the application of (7·11) it is important that T_0, T_1, M and D are recomputed at each step (except possibly the last), just as $f'(a)$ must be recomputed when applying Newton's rule to real zeros.

Bairstow's formula does not seem to have acquired quite the popularity it deserves. In our experience it is the best, for complex zeros, of all known quadratically convergent formulae.

The division of polynomials by real quadratic factors

The successful application of Bairstow's formula to high-degree polynomials requires a knowledge of the building-up errors that are incurred in the division of polynomials by quadratic factors. We now consider this question.

Let us write the quotient $q(x)$ of equation (7·9) in the form

$$q(x) \equiv q_n x^{n-2} + q_{n-1} x^{n-3} + \ldots + q_2.$$

The process of division amounts to the application of the recursion formula given by

$$\left.\begin{aligned} q_n = a_n, \quad q_{n-1} = pq_n + a_{n-1} \\ \text{and} \qquad q_s = pq_{s+1} + lq_{s+2} + a_s \quad (s = n-2, n-3, \ldots, 0). \end{aligned}\right\} \tag{7·12}$$

It will be noted that the remainder is not $q_1 x + q_0$, but $q_1 x + q_0'$, where

$$q_0' = q_0 - pq_1. \tag{7·13}$$

For the purpose of iteration, the quantities p, l and the coefficients a_s must be regarded as exact. Errors are introduced through the rounding-off of successive q's calculated from equations (7·12). Suppose that the computed value of q_s is rounded-off to $q_s - \epsilon_s$. The error ϵ_s thus introduced builds up according to the difference equation

$$u_m = pu_{m+1} + lu_{m+2}, \tag{7·14}$$

with the initial values

$$u_s = \epsilon_s, \quad u_{s+1} = 0. \tag{7·15}$$

The contribution of ϵ_s to the total error in q_m $(m \leqslant s)$ is u_m. The solution of (7·14) with the initial values (7·15) is

$$u_m = \epsilon_s r^{s-m} \operatorname{cosec}\theta \sin(s-m+1)\theta, \tag{7·16}$$

where $r\,\mathrm{e}^{\pm \mathrm{i}\theta}$ are the zeros of $x^2 - px - l$. Hence the total error in q_m due to accumulation of the rounding errors in $q_m, q_{m+1}, \ldots, q_n$, is

$$e_m = \operatorname{cosec}\theta \sum_{s=m}^{n} \epsilon_s r^{s-m} \sin(s-m+1)\theta = \operatorname{cosec}\theta \sum_{s=0}^{n-m} \epsilon_{m+s} r^s \sin(s+1)\theta. \tag{7·17}$$

In particular, the error in q_0 is

$$e_0 = \operatorname{cosec}\theta \sum_{s=0}^{n} \epsilon_s r^s \sin(s+1)\theta. \tag{7·18}$$

From the practical point of view the trigonometric terms in (7·17) can be ignored, and we arrive at the following simple rule for ascertaining the number of correct decimals in the coefficients of the quotient $q(x)$:

Rounding errors in the sequence $\{q_s\}$ effectively build up at each step by a factor r.

If the division is performed using a transfer calculating machine, then in the calculation of q_s by (7·12) it is usual to employ a transferred value of q_{s+1}, and this is not so severely rounded as the recorded value. A similar examination of this process reveals that the same rule is valid.

Using the rule, we can decide the correct and economical number of decimals to retain in each q_s. There are three cases. First, if r is near to unity, a fixed number of decimals can be retained in all the q_s. Secondly, if $r > 1$, figures must be discarded systematically: if $\sigma \equiv 1/\log_{10} r$, so that $r^\sigma = 10$, one decimal must be discarded approximately every σ steps. Finally, if $r < 1$, decimals can be *added* as the division proceeds: if $\tau \equiv -1/\log_{10} r$, so that $r^{-\tau} = 10$, then one decimal may be added approximately every τ steps.

Tables 5 and 6 give examples of the application of Bairstow's process with the above rule used in the processes of division. The polynomial is that given by equation (2·6). The approximate factors $x^2 + 0{\cdot}4489x + 0{\cdot}2537$ and $x^2 + 0{\cdot}10169x + 1{\cdot}6845$ have been obtained from the root-squaring process, using the method recommended in § 5 for the evaluation of the phases (cf. table 4), and in tables 5 and 6 these factors are determined more accurately. Places where decimals are discarded or added during the divisions are indicated by an asterisk, and they are ascertained before the numerical recurrences are carried out.

TABLE 5. BAIRSTOW'S PROCESS

l		$-0{\cdot}2537$		$-0{\cdot}25372\ 3$
p		$-0{\cdot}4489$		$-0{\cdot}44894\ 0$
s	$f(x)$	$q(x)$	$T(x)$	$q(x)$
16	2·03253 121	+2·033	+2	+2·033
15	3·43560 48	2·523	2	2·523
14	25·17830 48	23·530	22	23·530
13*	37·65109 6	26·4484	15·9	26·4475
12	128·21874 8	110·3765	97·6	110·3753
11	166·44768	110·1897	62·3	110·1854
10	345·07256	267·6059	214·9	267·6012
9*	378·908	230·82458	118·57	230·81457
8	524·327	352·81823	245·07	352·80853
7	468·88	251·93970	111·84	251·92717
6	443·576	240·97028	128·59	240·96018
5*	304·08	131·99133 8	45·894	131·98362 1
4	190·68	70·29492 8	17·070	70·29013 3
3	89·6	24·55840 4	$+5{\cdot}252\ T_1$	24·55666 7
2	32·8	3·94190 9	$-2{\cdot}747\ T_0$	+3·94130 6
1	8	$+0{\cdot}00001\ 0\ q_1$		$-0{\cdot}00000\ 1\ q_1$
0	1	$-0{\cdot}00006\ 7\ q_0$		$+0{\cdot}00000\ 0\ q_0$
M			−0·099	
D			+8·07	
δl			−0·00002 3	+0·00000 0
δp			−0·00004 0	−0·00000 0

TABLE 6. BAIRSTOW'S PROCESS

l	−1·6845		−1·68456 6
p	−0·10169		−0·10172 9
s	$q(x)$	$T(x)$	$q(x)$
16	+2·03253 121	+2·03253	+2·03253 121
15	3·22891 670	3·02223	3·22883 743
14	21·42615 744	17·69503	21·42590 543
13	30·03315 987	23·14281	30·03227 031
12	89·07231 377	56·91164	89·07024 337
11	106·79905 861	62·02765	106·79531 175
10*	184·16985 12	81·9946	184·16367 61
9	180·27675 36	67·4531	180·26946 23
8	195·76054 26	50·7813	195·75250 07
7	145·29691 90	26·5082	145·29048 68
6	99·04212 23	10·8054	99·03773 70
5	49·25574 65	3·5039	49·25357 59
4	18·83472 81	+0·2767	18·83387 85
3	4·71339 15	$-1{\cdot}217\ T_1$	4·71314 90
2	+0·59359 57	$+0{\cdot}251\ T_0$	0·59362 47
1*	$-0{\cdot}00007\ 1\ q_1$		$+0{\cdot}00000\ 1\ q_1$
0	$+0{\cdot}00009\ 5\ q_0$		$-0{\cdot}00000\ 0\ q_0$
M		+2·025	
D		+2·527	
δl		−0·00006 6	+0·00000 1
δp		−0·00003 9	−0·00000 0

(8) *Multiple zeros*

Suppose α is a zero of $f(x)$ of known* multiplicity m, and that a is an approximation to α. Then α is a simple zero of $f^{(m-1)}(x)$, and the iteration methods of § 7 may be employed here in full, replacing $f(x)$ by $f^{(m-1)}(x)$. If $\eta \equiv \alpha - a$, then from (7·1), for example, we obtain

$$\eta = -\frac{f^{(m-1)}(a)}{f^{(m)}(a)} + O(|\eta|^2). \tag{8·1}$$

* Polynomials which are known *a priori* to have multiple zeros seldom arise in practice. The detection and iteration of multiple zeros during the normal process of solution are covered in §9.

Another formula, proposed by Bodewig (1949), is a modified form of Newton's rule:

$$\eta = -\frac{mf(a)}{f'(a)} + O(|\eta|^2). \tag{8·2}$$

In practice this formula is inferior to (8·1). It involves the ratio of two small quantities

$$f(a) = O(|\eta|^m), \quad f'(a) = O(|\eta|^{m-1}),$$

whereas, in (8·1), $f^{(m)}(a) = O(1)$ and $f^{(m-1)}(a) = O(|\eta|)$,

and this difficulty does not there arise.

(9) *Zeros which are very close together*

If a polynomial is particularly ill-conditioned and some of its zeros are extraordinarily close together, their direct separation by root-squaring, notwithstanding the great power of this process, may not be the quickest method of solution. As we have seen in § 3, severe cancellation occurs in such cases, and it is necessary to retain a very large number of significant figures during the early transformations. An example of such a polynomial was given at the end of § 3.

For simplicity we confine our attention here to polynomials with only one group of close zeros, or, alternatively, with one conjugate pair of groups of such zeros (the extension to polynomials with more groups is straightforward and of no special interest). The treatment of the difficulty then falls into four cases, depending on whether or not the number of close zeros is small compared with the degree of the polynomial, and whether or not the zeros cluster in the neighbourhood of a *real* point.

(*a*) *Several zeros near a real point*

Let a be a real point around which there cluster several of the zeros of

$$f(x) \equiv a_0 + a_1 x + \ldots + a_n x^n. \tag{9·1}$$

If no *a priori* knowledge of a is available, a suitable value can be deduced from the settled coefficients bordering the group of close zeros in the root-squaring transformations of $f(x)$.

We perform a root-reducing transformation, producing a polynomial whose zeros are well separated. This could be done by constructing the polynomial

$$f(x+a) = f(a) + xf'(a) + x^2\frac{f''(a)}{2!} + \ldots + x^n\frac{f^{(n)}(a)}{n!}, \tag{9·2}$$

but the labour entailed in computing the coefficients $f^{(s)}(a)/s!$ directly or by Horner's method is not inconsiderable and can be largely averted if, instead, we construct the polynomial

$$F(x) = F_0 + F_1 x + F_2 x^2 + \ldots + F_n x^n, \tag{9·3}$$

where

$$\left.\begin{aligned} F_s &= \alpha_s + \binom{s+1}{1}\alpha_{s+1} + \binom{s+2}{2}\alpha_{s+2} + \ldots + \binom{n}{n-s}\alpha_n, \\ \text{and} \quad \alpha_s &= a^s . a_s. \end{aligned}\right\} \tag{9·4}$$

If the zeros of $f(x)$ are $x_1, x_2, \ldots, x_n$ then the zeros of $F(x)$ can be denoted by $\xi_1, \xi_2, \ldots, \xi_n$, where

$$x_s = a(1+\xi_s) \quad (s = 1, 2, \ldots, n). \tag{9·5}$$

The zeros ξ_s, though small themselves, bear ratios to one another which are not small and, therefore, are well separated from the standpoint of root-squaring.

In practice the coefficients F_s are computed by (9·4), using as checks the reciprocal relations

$$\alpha_s = F_s - \binom{s+1}{1} F_{s+1} + \binom{s+2}{2} F_{s+2} - \ldots + (-)^{n+s} \binom{n}{n-s} F_n. \tag{9·6}$$

For $s = 0$ this becomes

$$\alpha_0 = F_0 - F_1 + F_2 - \ldots + (-)^n F_n, \tag{9·7}$$

and this alone is usually an adequate check, the principal practical use of (9·6) with non-zero values of s being to help locate errors when (9·7) is not satisfied.

The polynomial $F(x)$ can be solved by the standard methods already described. The real zeros x_s of $f(x)$ are then obtained by using (9·5) directly, and the quadratic factors $x^2 - px - l$ (corresponding to the conjugate pairs of complex zeros of $f(x)$) are obtained from the equations

$$p = a(\varpi + 2), \quad l = a^2(\lambda - \varpi - 1), \tag{9·8}$$

where $x^2 - \varpi x - \lambda$ is a factor of $F(x)$.

(*b*) *Several zeros near a conjugate pair of points*

The method of (*a*) remains applicable in theory to this case, the practical modification being that a is now complex and $F(x)$ is a polynomial with complex coefficients, the direct solution of which entails four times the labour of solving a similar real polynomial of equal degree (cf. § 11). The following less laborious procedure, which depends on the solution of a *real* polynomial of the *same* degree, is due in principle to Ostrowski (1940).

Let a common approximation to one of the groups of zeros be denoted by $a \equiv u(1 + iv)$, where u and v are both real.* We then proceed as follows.

(i) Transform $f(x)$ into $F(x)$, as indicated by equations (9·3) and (9·4), taking $a = u$ in (9·4). The zeros of $F(x)$ cluster around the points $x = \pm iv$.

(ii) Perform *one* root-squaring transformation on $F(x)$ to obtain a polynomial $F_1(x)$, the zeros of which cluster around the *single* point $x = v^2$.

(iii) Transform $F(x)$ into a polynomial $\phi(x)$, using formulae analogous to (9·3) and (9·4), taking $a = v^2$ in (9·4). The zeros of $\phi(x)$ cluster around the origin $x = 0$, and this polynomial (which is real and of the same degree as $f(x)$) may be solved by root-squaring since its zeros are, in general, adequately separated for this purpose.

If the zeros of $\phi(x)$ are denoted by ζ_s, then it is easily verified that their relation to the zeros x_s of $f(x)$ is given by

$$(x_s - u)^2 + u^2 v^2 (1 + \zeta_s) = 0. \tag{9·9}$$

This equation may be used directly to compute x_s from real values of ζ_s, the ambiguity (concerning which of two possible values is a genuine zero) being resolved by substitution in $f(x)$. For complex values of ζ_s a more convenient form of the relation is given by the expressions

$$\left.\begin{aligned} p &= 2u + S, & l &= -u(S + u + uv^2 R), \\ \text{where} \quad R &= (1 + P - L)^{\frac{1}{2}}, & S &= \pm uv(2R - P - 2)^{\frac{1}{2}}. \end{aligned}\right\} \tag{9·10}$$

* If a is purely imaginary we cannot write it in this form, but a suitable modification of the given procedure can be made. Stage (i) is omitted and if $a = iv$, stages (ii) and (iii) are performed taking $f(x)$ in place of $F(x)$. The polynomial whose zeros are given by (3·15) was solved with little difficulty using this technique, the value of a being taken as $10^{-\frac{1}{2}}i$.

Here x^2-Px-L is a factor of $\phi(x)$ such that $P^2+4L<0$, and x^2-px-l denotes the corresponding factor of $f(x)$. The sign ambiguity is resolved by dividing $f(x)$ by x^2-px-l.

(c) A few zeros near a real point

The numerical procedure for this case can be made to be a shortened version of that suggested for case (*a*).

Let us suppose that the root-squaring of $f(x)$ indicates that there is a group of m unseparated zeros $x_1, x_2, \ldots, x_m$ (say), lying in the neighbourhood of a real point $x=a$, m being small compared with n. Then from (*a*) we see that the values of x_s are given by equation (9·5), where $\{\xi_s\}$ are the zeros of the polynomial (9·3). The m zeros of least modulus must be $\xi_1, \xi_2, \ldots, \xi_m$; let the greatest of these have a modulus equal to ϵ and let the least of the zeros $\xi_{m+1}, \xi_{m+2}, \ldots, \xi_n$ have a modulus equal to E. Consider

$$G(x) \equiv F_0+F_1x+F_2x^2+\ldots+F_mx^m. \tag{9·11}$$

We shall show* that if $\eta_1, \eta_2, \ldots, \eta_m$ are the zeros of this polynomial, suitably enumerated, then

$$\xi_s = \eta_s + O(\epsilon^2/E) \tag{9·12}$$

for every simple zero ξ_s of $F(x)$, such that

$$\xi_s-\xi_r \neq o(\epsilon), \quad \text{when } r \neq s. \tag{9·13}$$

In the first place we note that it is sufficient to take $E=1$; the extension to other values follows by a simple change of scale. Using equation (9·3), we see that if $1 \leqslant s \leqslant m$, then

$$G(\xi_s) = -[F_{m+1}\xi_s^{m+1}+F_{m+2}\xi_s^{m+2}+\ldots+F_n\xi_s^n] = O(\epsilon^{m+1}), \tag{9·14}$$

and also
$$\begin{aligned} G'(\xi_s) &= F'(\xi_s)-\xi_s^m[(m+1)F_{m+1}+(m+2)F_{m+2}\xi_s+\ldots+nF_n\xi_s^{n-m-1}] \\ &= F'(\xi_s)+O(\epsilon^m). \end{aligned} \tag{9·15}$$

The result (9·12) will follow from Newton's rule applied to $G(x)$ for $x=\xi_s$, if we can show that

$$[G'(\xi_s)]^{-1} = O(\epsilon^{1-m}). \tag{9·16}$$

Now writing $F(x)$ in the form

$$F(x) = F_n(x-\xi_1)(x-\xi_2)\ldots(x-\xi_m)(x-\xi_{m+1})\ldots(x-\xi_n), \tag{9·17}$$

we see that if $1 \leqslant s \leqslant m$, then

$$F^{(\nu)}(\xi_s) = O(\epsilon^{m-\nu}) \quad (\nu = 0, 1, 2, \ldots, m). \tag{9·18}$$

Also, we have

$$F(\xi_s+x) = xF'(\xi_s)+\frac{x^2}{2!}F''(\xi_s)+\ldots+\frac{x^n}{n!}F^{(n)}(\xi_s). \tag{9·19}$$

It is now clear that $F'(\xi_s)$ cannot be $o(\epsilon^{m-1})$ for otherwise it would follow from (9·18) and (9·19) that $F(\xi_s+x)$ had at least two zeros of the form $x=o(\epsilon)$, contrary to the hypothesis (9·13). Hence from (9·15), $G'(\xi_s)$ also cannot be $o(\epsilon^{m-1})$ and the results (9·16) and (9·12) under the condition (9·13) are now established when $E=1$, and therefore also for all E.

If ξ_s is such that there is another† zero of $F(x)$ in the neighbourhood $|x-\xi_s| = o(\epsilon)$, then the relation (9·12) is not necessarily true; a further shortened form of root-reducing transformation is needed to separate such zeros.

* The proof which follows is not intended to be mathematically rigorous; a full analytical treatment here might obscure the essential numerical simplicity of the result. In any case, in practical applications the approximate zeros derived in this way would invariably be tested and, if necessary, iterated by recourse to one of the polynomials $F(x)$ or $f(x)$.

† This includes of course the possibility of ξ_s being a multiple zero.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Thus the m smallest zeros of $F(x)$ can, as a first approximation, be taken to be equal to the zeros of $G(x)$; more accurate values can then usually be obtained by iteration in $F(x)$. Since the degree of $G(x)$ is low compared with that of $F(x)$, this procedure saves a considerable amount of labour.

(*d*) *A few zeros near a complex point*

In a similar way we can shorten the procedure described in (*b*). If there are m zeros of $f(x)$ clustered in the neighbourhood of $x = a$, where a is now complex, then good approximations to the small zeros of $\phi(x)$ are obtained by solving the polynomial which comprises the $2m+1$ terms of lowest degree of $\phi(x)$.

For the cases $m = 2$ and 3 it is simpler still to follow the procedure of (*c*) taking a complex value of a. The polynomial $G(x)$ is then a quadratic or cubic with complex coefficients and can be solved explicitly without difficulty.

Part C. Miscellaneous topics

(10) *Rounding errors in the coefficients*

During the numerical processes of solution of the polynomial

$$f(x) = a_0+a_1x+a_2x^2+\ldots+a_nx^n, \tag{10·1}$$

we have, so far, regarded the coefficients a_s as exact quantities. We now consider the effect upon the zeros of small changes in these coefficients.

Let α be a simple zero of $f(x)$. Then if a_m is increased by δa_m, the value of $f(a)$ changes from zero to $\alpha^m\delta a_m$, to the first order of approximation. Hence, by Newton's rule, the corresponding change in the zero is $\delta\alpha \equiv -\alpha^m\delta a_m[f'(\alpha)]^{-1}$. Proceeding to the limit, we obtain

$$\frac{\partial\alpha}{\partial a_m} = -\frac{\alpha^m}{f'(\alpha)}. \tag{10·2}$$

The direct use of this formula enables us to determine quite simply the limitations of accuracy that rounding or observational errors in the coefficients impose upon each real simple zero. The value of $f'(\alpha)$ is usually available in these cases from the iterations or checking of α.

For complex zeros, a more convenient form of the result may be derived. Let us suppose that the quadratic factor x^2-px-l, corresponding to the pair of simple zeros $re^{\pm i\theta}$, has been iterated by the use of Bairstow's formula (7·11). Then by differentiating equation (7·9) and using (7·10), it is seen that

$$\begin{aligned} f'(re^{i\theta}) &= (2re^{i\theta}-p)(T_1re^{i\theta}+T_0-pT_1)+q_1, \\ &= 2ir\sin\theta(T_0-r\cos\theta T_1+ir\sin\theta T_1), \end{aligned} \tag{10·3}$$

since $p = 2r\cos\theta$ and q_1 supposedly vanishes at the close of the iterations. Hence, using equation (10·2), we obtain

$$\frac{\partial(re^{i\theta})}{\partial a_m} = \frac{ir^{m-1}e^{mi\theta}}{2\sin\theta(T_0-r\cos\theta T_1+ir\sin\theta T_1)}. \tag{10·4}$$

For estimating the accuracy of the zero, it is sufficient to know the modulus

$$\left|\frac{\partial(r\,e^{i\theta})}{\partial a_m}\right| = \frac{r^{m-1}}{2\,|\sin\theta|}\,|\,T_0^2 - pT_0T_1 + r^2T_1^2\,|^{-\frac{1}{2}},$$

that is,

$$\left|\frac{\partial(r\,e^{i\theta})}{\partial a_m}\right| = \frac{r^m}{(4r^2-p^2)^{\frac{1}{2}}\,|\,D\,|^{\frac{1}{2}}}, \tag{10·5}$$

where D is given by (7·11) and its numerical value is known from the iterations.

From (10·2) we see that zeros which are close together are the most sensitive to slight perturbations in the coefficients, since at such zeros the derivative is necessarily small. This is the phenomenon of ill-conditioning, not uncommon with high-degree polynomials, to which reference has been made in the Introduction and in § 3.

For multiple zeros, the results corresponding to (10·2) involve higher derivatives than the first. Such zeros no longer retain their multiplicity for arbitrary small changes in the coefficients, but break up into groups of simple zeros, the process being extremely sensitive.

(11) *Complex coefficients*

The numerical solution of high-degree polynomials with complex coefficients appears to have received little attention. The problem is not an intractable one, and we consider here, in outline, some practicable methods. Insufficient experience precludes an exhaustive discussion of the relative merits of such methods.

The general nature of the zeros of such a polynomial is of course different from the nature of those of a polynomial whose coefficients are all real. Real zeros are rare and complex zeros no longer necessarily occur in conjugate pairs, so that repeated moduli are the exception rather than the rule.

The root-squaring process can be readily adapted and it will determine the moduli of the zeros. The coefficients in the array (2·5) are now complex numbers, so that the labour involved is about four times that of the root-squaring of a similar real polynomial of equal degree. The evaluation of the phases of the zeros using the known moduli can be performed by methods given in § 4 for real polynomials. Iteration of zeros, when necessary, can be effected by Newton's rule (7·1), obtaining the values of $f(a)$ and $f'(a)$ by direct calculation.

An alternative approach to the problem is to make it depend on the solution of a polynomial with real coefficients of double the original degree. Write the given complex polynomial in the form

$$f(x) = \sum_{s=0}^{n} (b_s + ic_s)\,x^s, \tag{11·1}$$

and let $\beta_s + i\gamma_s$ $(s = 1, 2, \ldots, n)$ be its zeros, where b_s, c_s, β_s and γ_s are all real. Consider the polynomial

$$\phi(x) = f(x)\,\bar{f}(x) = \sum_{s=0}^{2n} \phi_s x^s, \tag{11·2}$$

where

$$\left.\begin{aligned} \phi_{2s} &= b_s^2 + c_s^2 + 2\sum_{m\geqq 1}(b_{s-m}b_{s+m} + c_{s-m}c_{s+m}), \\ \phi_{2s+1} &= 2\sum_{m\geqq 0}(b_{s-m}b_{s+m+1} + c_{s-m}c_{s+m+1}). \end{aligned}\right\} \tag{11·3}$$

These coefficients are real and $\phi(x)$ can be solved by the standard methods already described. The zeros of $\phi(x)$ are $\{\beta_s \pm i\gamma_s\}$ and the resolution of the ambiguities may be effected by substitution in $f(x)$. The total labour entailed by this method is about eight times that required to solve a similar real polynomial of equal degree.

(12) *Checking*

The iteration processes provide sound individual checks on the zeros, but even so it is advisable to apply further simple group checks based upon the algebraical properties of the zeros. If the polynomial is denoted by (10·1), its real zeros by $\alpha_1, \alpha_2, \ldots, \alpha_m$, and the quadratic factors corresponding to the pairs of complex zeros by $x^2 - p_s x - l_s$ $(s = 1, 2, \ldots, k)$ $(m + 2k = n)$, then four of the most useful of such checks are

$$\Pi\alpha_s \, \Pi l_s = (-)^{n+k} \frac{a_0}{a_n}, \quad \Sigma\alpha_s + \Sigma p_s = -\frac{a_{n-1}}{a_n},$$

$$\Sigma \frac{p_s}{l_s} - \Sigma \frac{1}{\alpha_s} = \frac{a_1}{a_0}, \quad \Sigma\alpha_s^2 + \Sigma p_s^2 + 2\Sigma l_s = \frac{a_{n-1}^2 - 2a_n a_{n-2}}{a_n^2}.$$

(13) *Summary*

The solution of an arbitrary polynomial of high degree is best effected in several steps. First, the root-squaring process is used to obtain the moduli of the zeros, both real and complex. Secondly, the phases of the complex zeros are evaluated by means of the 'highest common factor' process described in §§ 4 and 5. Thirdly, higher accuracy when needed can be achieved by an indirect or iterative procedure. The most suitable iteration formulae are Newton's rule, for real zeros, and Bairstow's formula, for complex zeros.

Except in special cases it is uneconomic to use iterative methods from the start, without a reasonably accurate knowledge of the location of the required zero.

Polynomials of high degree present special difficulties in large-scale cancellation and subsequent loss of significant figures. In §§ 3, 5 and 9 this phenomenon was examined in some detail and remedies were suggested.

The amount of work involved and the time taken to obtain all the zeros depend on several factors:

(i) the cube of the degree of the polynomial,
(ii) the degree of ill-conditioning (§§ 3 and 9),
(iii) the proportion of complex zeros,
(iv) the order of accuracy required.

The sixteenth-degree polynomial given by (2·6), which is moderately ill-conditioned and has no real zeros, was solved by an experienced computer, using a desk calculating machine, in 16 hours, to an accuracy of about five significant figures.

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published with the permission of the Director of the Laboratory.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

References

Aitken, A. C. 1926 *Proc. Roy. Soc. Edinb.* **46**, 289–305.

Bairstow, L. 1914 *Rep. Memor. Adv. Comm. Aero., Lond.*, **154**, 51–63.

Bodewig, E. 1946*a* *Quart. Appl. Math.* **4**, 177–190.

Bodewig, E. 1946*b* *Proc. Acad. Sci. Amst.* **49**, 911–921.

Bodewig, E. 1949 *Quart. Appl. Math.* **7**, 325–333.

Brodetsky, S. & Smeal, G. 1924 *Proc. Camb. Phil. Soc.* **22**, 83–87.

Frazer, R. A. & Duncan, W. J. 1929 *Proc. Roy. Soc.* A, **125**, 68–82.

Graeffe, C. H. 1837 *Die Auflösung der höheren numerischen Gleichungen.* Zurich: Friedrich Schulthess.

Hartree, D. R. 1949 *Proc. Camb. Phil. Soc.* **45**, 230–236.

Jelinek, P. C. 1865 *Die Auflösung der höheren numerischen Gleichungen.* Leipzig.

Ostrowski, A. 1940 *Acta Math.* **72**, 99–257.

Shih-Nge Lin 1943 *J. Math. Phys.* **22**, 60–77.

Van der Corput, J. G. 1946 *Proc. Acad. Sci. Amst.* **49**, 922–929.

Whittaker, E. T. & Robinson, G. 1944 *The calculus of observations.* London: Blackie and Son.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES