XXV.—On Bernoulli's Numerical Solution of Algebraic Equations.

By A. C. Aitken, D.Sc.

(MS. received April 21, 1926. Read May 24, 1926.)

§ 1. Introductory.

The aim of the present paper is to extend Daniel Bernoulli's method * of approximating to the numerically greatest root of an algebraic equation. On the basis of the extension here given it now becomes possible to make Bernoulli's method a means of evaluating not merely the greatest root, but all the roots of an equation, whether real, complex, or repeated, by an arithmetical process well adapted to mechanical computation, and without any preliminary determination of the nature or position of the roots. In particular, the evaluation of complex roots is extremely simple, whatever the number of pairs of such roots. There is also a way of deriving from a sequence of approximations to a root successive sequences of ever-increasing rapidity of convergence.

Bernoulli's method consists in forming a sequence of solutions of the linear difference equation associated with the given algebraic equation. Thus, if the equation to be solved is

$$a_0z^n + a_1z^{n-1} + a_2z^{n-2} + \ . \ . \ . \ + a_n = 0 \qquad (1.1)$$

then the following difference equation in $f(t)$ is considered:

$$a_0f(t+n) + a_1f(t+n-1) + a_2f(t+n-2) + \ . \ . \ . \ + a_nf(t) = 0 \qquad (1.2)$$

Arbitrary values are assigned to the consecutive n values of f, namely, $f(t+n-1)$, $f(t+n-2)$, . . . $f(t)$, and by (1.2) $f(t+n)$ is found. Next, the n consecutive values $f(t+n)$, $f(t+n-1)$, . . . $f(t+1)$ serve in the same way to determine $f(t+n+1)$, and by repeating this process a sequence of solutions of the difference equation is obtained, with arguments increasing by unity. Then, if $|z_1|$, the modulus of the greatest root of the equation (1.1), is greater than the modulus of any other root, the sequence $f(t)$ will tend to become a geometric sequence with common ratio z_1; in fact, we shall have

$$\lim_{t \to \infty} \frac{f(t+1)}{f(t)} = z_1 \qquad (1.3)$$

* *Commentarii Acad. Sc. Petropol.*, III (1732); *cf.* Euler, *Introductio in Analysin Infinitorum*, I, Cap. XVII; Lagrange, *Résolution des équations numériques*, Note VI.

This is Bernoulli's result. The proof depends on the fact that the solution of the difference equation (1.2), when all the roots of (1.1) are different, is

$$f(t) = w_1 z_1^t + w_2 z_2^t + \ . \ . \ . \ + w_n z_n^t \quad . \quad . \quad . \quad . \quad (1.4)$$

where $z_1, z_2, z_3, \ . \ . \ . \ z_n$ are the roots of (1.1) in descending order of moduli, and $w_1, w_2, \ . \ . \ . \ w_n$ are arbitrary periodic functions of period unity; when the roots are not all different, *e.g.* if z_2, z_3, z_4 coalesce into the one value z_4, then we replace the terms

$$(w_2 z_2^t + w_3 z_3^t + w_4 z_4^t) \quad \text{by} \quad (w_2 z_4^t + t w_3 z_4^t + t^2 w_4 z_4^t);$$

in either case, when t is large, provided w_1 is not zero, the dominating term of $f(t)$ is the term $w_1 z_1^t$. From this Bernoulli's result may be deduced.

§ 2. Extensions of Bernoulli's Formula.

The principal extension we shall give is the following:—

$$\lim_{t\to\infty} \frac{\begin{vmatrix} f(t+1) & f(t+2) & \ . \ . \ . & f(t+m) \\ f(t) & f(t+1) & \ . \ . \ . & f(t+m-1) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ f(t-m+2) & f(t-m+3) & \ . \ . \ . & f(t+1) \end{vmatrix}}{\begin{vmatrix} f(t) & f(t+1) & \ . \ . \ . & f(t+m-1) \\ f(t-1) & f(t) & \ . \ . \ . & f(t+m-2) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ f(t-m+1) & f(t-m+2) & \ . \ . \ . & f(t) \end{vmatrix}} = z_1 z_2 z_3 \ . \ . \ . \ z_m \quad . \quad (2.1)$$

of which Bernoulli's is the simplest case, $m=1$.

Let us first consider the case in which no multiple roots are contained among the m roots in question. As t is increased, the later terms of $f(t)$, provided $|z_m| > |z_{m+1}|$, become of negligible order, and the significant part of $f(t)$ is

$$w_1 z_1^t + w_2 z_2^t + \ . \ . \ . \ + \ w_m z_m^t,$$

assuming none of $w_1, w_2, \ . \ . \ . \ w_m$ to be zero. Hence for the denominator of the quotient on the left of (2.1) we consider the determinant

$$f_m(t) = \begin{vmatrix} \Sigma w_r z_r^t & \Sigma w_r z_r^{t+1} & \ . \ . \ . & \Sigma w_r z_r^{t+m-1} \\ \Sigma w_r z_r^{t-1} & \Sigma w_r z_r^t & \ . \ . \ . & \Sigma w_r z_r^{t+m-2} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \Sigma w_r z_r^{t-m+1} & \Sigma w_r z_r^{t-m+2} & \ . \ . \ . & \Sigma w_r z_r^t \end{vmatrix} \quad . \quad . \quad . \quad (2.2)$$

where Σ denotes summation with respect to r from 1 to m. Since each of the m^2 elements is the sum of m terms, the determinant $f_m(t)$ can be expressed as a sum of m^m determinants with monomial elements, each of the form

$$\begin{vmatrix} w_p z_p^t & w_q z_q^{t+1} & \cdots \\ w_p z_p^{t-1} & w_q z_q^t & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}.$$

If in any of these determinants the suffixes $p, q, \ldots$ are not all different, then two or more columns are proportional and the determinant vanishes. Thus only those survive in which $p, q, \ldots$ is some permutation of *all* of $1, 2, 3, \ldots m$. Therefore we may remove common factors $w_1, w_2, w_3, \ldots w_m$ and $(z_1 z_2 z_3 \ldots z_m)^{t-m+1}$, and adding together again all the monomial determinants thus reduced, including the vanishing ones, we have

$$f_m(t) = (-)^{\frac{m(m-1)}{2}} w_1 w_2 \ldots w_m (z_1 z_2 \ldots z_m)^{t-m+1} \begin{vmatrix} s_0 & s_1 & \cdots & s_{m-1} \\ s_1 & s_2 & \cdots & s_m \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ s_{m-1} & s_m & \cdots & s_{2m-2} \end{vmatrix},$$

where $$s_r = z_1^r + z_2^r + \ldots + z_m^r,$$

and so
$$f_m(t) = (-)^{\frac{m(m-1)}{2}} \prod_{r=1}^{m} (w_r z_r^{t-m+1}) \cdot \zeta(z_1, z_2, \ldots z_m) \quad . \quad . \quad (2.3)$$

the determinant of s_r being familiar * from the Sturmian theory of equations, and ζ denoting (after Sylvester) a squared difference-product. Hence we have finally the theorem (2.1), namely,

$$\lim_{t\to\infty} \frac{f_m(t+1)}{f_m(t)} = z_1 z_2 \ldots z_m.$$

Now let us consider the case hitherto excluded, in which, included among the m roots in question, there exist one or more sets of multiple roots. Then, as we have seen in § 1, the terms representing the roots which have coalesced must accept as factors, in addition to the arbitrary constants w, powers of t up to the $(p-1)$th, where p denotes the *multiplicity*. Without giving the proof in detail we shall indicate the steps. It is first found that if, *e.g.*, $(w_3 z_3^t + w_4 z_4^t + w_5 z_5^t)$ in the previous determinant be now replaced by $(w_3 z_5^t + t w_4 z_5^t + t^2 w_5 z_5^t)$, then in the monomial determinants all those involving w_3, w_4, or t or t^2 as a *factor* (not as an exponent) vanish. Hence we revert to the whole determinant $f_m(t)$ and simply put $w_3 = w_4 = 0$, and $t = 0$ wherever it occurs as a factor. Some

* *Cf.* Borchardt, *J. für Math.*, **30** (1845), p. 38.

reduction then leads to a result which in the general case may be enunciated as follows:—

In the roots $z_1, z_2, \ldots z_m$ let there be sets of coalescing roots, of multiplicity $p, q, \ldots$ respectively, namely, $z_a, z_{a+1}, \ldots z_{a+p-1}$ coalescing to the value z_{a+p-1}; $z_b, z_{b+1}, \ldots z_{b+q-1}$ coalescing to z_{b+q-1}, etc.; then in the result (2.3) we must replace

$$(w_a w_{a+1} \ldots w_{a+p-1}) \quad \text{by} \quad (w_{a+p-1})^p,$$
$$(w_b w_{b+1} \ldots w_{b+q-1}) \quad \text{by} \quad (w_{b+q-1})^q,$$

and so on, and in the squared difference-product ζ we must replace every squared difference that would otherwise vanish, *e.g.* $(z_a - z_{a+1})^2$, by the square of the particular root *alone*, *e.g.* by z_{a+p-1}^2; finally, we must insert numerical factors $\{(p-1)!\}^p$, $\{(q-1)!\}^q$, . . ., etc.

This modified result may be called the *confluent* form of the result (2.3). In particular, if all of the m roots $z_1, z_2, \ldots z_m$ coalesce to z_m, then (2.3) becomes, provided $|z_m| > |z_{m+1}|$,

$$f_m(t) = (-)^{\frac{m(m-1)}{2}} \{(m-1)!\}^m w_m^m z_m^{m(t+m-1)} \quad . \quad . \quad . \quad (2.4)$$

It is thus seen that the theorem (2.1) holds both in the confluent and non-confluent case. It may even be shown to hold when $z_m = z_{m+1}$, but the convergency of the sequence obtained is too slow * for practical purposes.

In passing it may be observed that a still wider generalisation of (2.1) is possible, for in the foregoing demonstration it is not essential that the argument of $f(t)$ should move by *unit* steps from one row to the next, or from one column to the next. Thus the differences of the arguments from row to row and column to column may be any integers such that no rows or columns of $f_m(t)$ become identical. This extension is not, however, of such practical bearing as the more special result (2.1).

§ 3. The Arithmetical Process of Solution.

In the practical application of the extended Bernoulli rule, repeated use is made of an identity familiar in the theory of compound determinants, for example,

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{vmatrix} \equiv \begin{vmatrix} \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} & \begin{vmatrix} a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \\ c_2 & c_3 & c_4 \end{vmatrix} \\ \begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix} & \begin{vmatrix} b_2 & b_3 & b_4 \\ c_2 & c_3 & c_4 \\ d_2 & d_3 & d_4 \end{vmatrix} \end{vmatrix} \div \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}, \quad (3.1)$$

* *Cf.* Euler and Lagrange, *loc. cit.*, with reference to two equal greatest roots.

and its obvious analogues of any order. First of all, the difference equation associated with the algebraic equation to be solved is used to compute a Bernoullian sequence $f(t), f(t+1), f(t+2), \ldots$ in the ordinary way. In general this will enable us to find z_1 as closely as we desire. Next, the first sequence $f(t)$ having been entered in a column to the left of the computing sheet, we form a second column to the right, of entries

$$f_2(t) = \begin{vmatrix} f(t) & f(t+1) \\ f(t-1) & f(t) \end{vmatrix},$$

i.e. we *square the corresponding entry in the first column and subtract the product of the entries above and below.* In this manner we use the first column to obtain a sequence $f_2(t)$ which will in general give $z_1 z_2$ and hence, z_1 being already found, will give z_2. Next we form a third column of entries $f_3(t)$, computed by means of the determinantal identity (3.1) in the form

$$f_3(t) = \begin{vmatrix} f_2(t) & f_2(t+1) \\ f_2(t-1) & f_2(t) \end{vmatrix} \div f(t),$$

and in fact generally we have in this way

$$f_{r+1}(t) = \begin{vmatrix} f_r(t) & f_r(t+1) \\ f_r(t-1) & f_r(t) \end{vmatrix} \div f_{r-1}(t),$$

so that the entries in any column may be computed from those in the *two preceding* columns by the following simple rule :—

Square the entry on the same horizontal line in the preceding column, and subtract the product of the entries immediately above and below it. Divide the result obtained by the corresponding entry in the column preceding that again.

[In exemplifying the above by numerical illustrations we shall adopt the following convention. When dealing with numbers of widely differing magnitudes but known to a certain number of significant digits, we shall represent each number as a decimal fraction between ·1 and 1 multiplied by the requisite power of 10. To indicate the power we prefix a "characteristic," thus :

The number	123·4567	is represented	by	3·1234567,
„	·12345	„	„	0·12345,
„	·00123	„	„	$\bar{2}$·123.

Thus the characteristic, if *positive*, denotes the *number of digits to the left of the decimal point* in the ordinary notation; if *negative*, the *number of zeros to the right.* In this context the characteristic of the *logarithm* of the number to the base 10 is sometimes used, but where computing machines are available the above convention is to be preferred, for by adopting it and remembering that the characteristics obey the index law we can safely leave to the machine the task of correct ranging.]

Example. To solve the equation

$$z^4 - 10z^3 - 92z^2 + 234z + 315 = 0.$$

Putting $f(-2)=0$, $f(-1)=0$, $f(0)=0$, $f(1)=1$, and referring to the associated difference equation, we calculate $f(2)$, $f(3)$, . . . and thence by the rules given construct the following table :—

t.	$f(t)$.	$f_2(t)$.	$f_3(t)$.	$f_4(t)$.
0	0·0	0·0	0·0	0·0
1	1·1	1·1	1·1	1·1
2	2·10	− 2·92	− 3·234	3·315
3	3·192	5·10804	5·83736	
4	4·2606	− 7·109412	− 8·253833	
5	5·41069	9·116902177	10·810292	
6	6·602364	−11·1218756308	−13·254130	
7	7·9131704	13·1283701289	15·801681	
8	9·136303492	−15·1346100347	−18·252406	
9	10·2049261777	17·1414212595		
10	11·3070597564			

Now writing $f_m(t+1)/f_m(t)$ as $Z_m(t)$, we find from the last five entries in the above columns the table of quotients:

Z_1.	Z_2.	Z_3.	Z_4.
15·16	−104·25	−319·22	315
14·93	−105·33	−313·63	
15·03	−104·86	−315·46	
14·98	−105·06	−314·86	

Hence, to three significant digits, $z_1=15{\cdot}0$, $z_1z_2=-105$, $z_1z_2z_3=-315$, $z_1z_2z_3z_4=315$.

Thus the roots are $15{\cdot}0$, $-7{\cdot}00$, $3{\cdot}00$, and $-1{\cdot}00$. (In actual fact they are exactly 15, -7, 3, and -1.)

Some observations must now be made regarding the above and all similar examples. In the first place, the determinants of ascending order $f_2(t)$, $f_3(t)$, . . . are really *difference effects* of increasing degrees of delicacy, and hence the calculation of the entries in any column requires an increasing degree of accuracy in the more and more remote columns to the left. If it were really necessary to carry the columns to any great length this 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.

§ 4. The Evaluation of Conjugate Complex Roots.

In the case of conjugate complex roots of the equation (1.1) the solution $f(t)$ of the difference equation assumes a somewhat modified form. If z_m and z_{m+1} are complex roots, $r_m e^{i\theta}$ and $r_m e^{-i\theta}$, the terms $w_m z_m^t$ and $w_{m+1} z_{m+1}^t$ are replaced by $w_m r_m^t \cos t\theta$ and $w_{m+1} r_m^t \sin t\theta$. The presence of complex roots is therefore indicated (just as in the "root-squaring" method) by *fluctuations in sign and magnitude* in the columns f_m and Z_m.

For simplicity let us consider a case where the roots in order are z_1, z_2, $re^{i\theta}$, $re^{-i\theta}$, z_5. . . . The Z_1 column will give z_1, the Z_2 will give $z_1 z_2$, the Z_3 column will fluctuate irregularly, thus indicating that z_3 and z_4 are complex, but the Z_4 will give $z_1 z_2 z_3 z_4$, *i.e.* $z_1 z_2 r^2$, and hence r is found. It may be shown that the sequence Z_3 is tending not to $z_1 z_2 r$, but to the following succession:—

$$z_1 z_2 r \frac{\cos(\theta+a)}{\cos(a)}, \quad z_1 z_2 r \frac{\cos(2\theta+a)}{\cos(\theta+a)}, \quad z_1 z_2 r \frac{\cos(3\theta+a)}{\cos(2\theta+a)}, \text{ etc.}, \quad . \quad (4.1)$$

where a is some undetermined angle which we shall eliminate. Since $z_1 z_2 r$ is known, we may divide throughout by it and find a sequence $c(t)$ converging to the cosine-quotients. Eliminating a, we find that a sequence * tending to $\cos\theta$ is

$$\frac{c(t+1)c(t)+1}{2c(t)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.2)$$

and thus θ is determined in an exceedingly simple manner. An exactly similar method can be applied to find any number of different pairs of complex roots. Thus if z_m and z_{m+1} are conjugate complex roots, we find $z_1 z_2 \ldots z_{m-1} r$, and, dividing Z_m by this, obtain $c(t)$ and hence the amplitude, by (4.2).

Example. To solve the equation

$$z^5 - 13z^4 - 121z^3 - 398z^2 + 386z - 520 = 0.$$

Proceeding as before we find the following values:—

$f(t)$.	$f_2(t)$.	$f_3(t)$.
0·0	0·0	0·0
1·1	1·1	1·1
2·13	− 3·121	3·398
3·290	4·9467	6·205110
4·5741	− 6·249109	9·1070826
6·114511	− 8·58382345	11·5574106
7·2294226	11·11940422	14·2898586
8·45860507	−13·1303023237	
9·917298089		

* *Cf.* Euler, *loc. cit.*, §§ 351, 352.

$Z_1(t)$.	$Z_2(t)$.	$Z_3(t)$.
19·946	−26·313	515·35
20·035	234·36	522·07
19·990	−204·52	520·54
20·002	−109·13	520·01

Evidently $z_1 = 20{\cdot}0$, $z_1 z_2 z_3 = 520$. The latter two roots are seen to be complex, $re^{i\theta}$ and $re^{-i\theta}$, and we have $r = \sqrt{26{\cdot}0}$, *i.e.* 5·10, and so $z_1 r = 102{\cdot}0$. Dividing $z_2(t)$ by $z_1 r$ we obtain the cosine-quotients $c(t)$, and hence $\cos\theta$ by the formula (4.2):

$c(t)$.	$\{c(t+1)c(t)+1\} \div 2c(t)$.
− ·2580	−·7892
2·2976	−·7849
−2·0051	−·7843
−1·0699	

Hence $\cos\theta = -{\cdot}784$, and so $\sin\theta = \pm{\cdot}620$, and we have z_2 and z_3 as $-4{\cdot}00 \pm 3{\cdot}16i$. Finally the sum and product of the remaining roots z_4 and z_5 are each found to be 1·00, and thus we find all the roots of the equation as 20·0, $-4{\cdot}00 \pm 3{\cdot}16i$, $\cdot 500 \pm {\cdot}866i$.

(Actually the roots are 20, $-4 \pm i\sqrt{10}$, $\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$.)

§ 5. The Evaluation of Real Multiple Roots.

If $z_m, z_{m+1}, \ldots z_{m+p-1}$ are p equal real roots, but $|z_{m+p-1}| > |z_{m+p}|$, then the columns $Z_m, Z_{m+1}, \ldots Z_{m+p-2}$ do converge to the required products, but very slowly. The sequence Z_{m+p-1} will, however, converge more rapidly, and will give $(z_m)^p$. If we have reason to suspect the presence of equal roots, there is a test which is not difficult to apply so long as the multiplicity p is not too high, as follows:—

(*a*) If $p=2$, the pair of equal roots z_m and z_{m+1} is found to behave as a complex pair with amplitude either 0 or π. In this case we compute the sequence $c(t)$, and apply (4.2) to see whether $\cos\theta$ is ± 1.

Example.

$$z^4 - 22z^3 + 95z^2 + 312z - 144 = 0.$$

In the usual manner we compute several values of f, f_2, Z_1, Z_2:

$f(t)$.	$f_2(t)$.	$Z_1(t)$.	$Z_2(t)$.	$c(t)$.	$\cos\theta$.
0·0	0·0	22·0	95·0	1·242	1·000
1·1	1·1	17·682	167·25	1·195	1·000
2·22	2·95	15·825	138·64	1·163	1·000
3·389	5·15889	14·905	145·40	1·140	1·000
4·6156	7·220286	14·338	143·66	1·123	
5·91757	9·320304	13·957	144·08		
7·1315634	11·460162	13·683	143·98		
8·18362377	13·662997	13·476			
9·251245344	15·954595				
10·3358709653					

The slowness of convergence of $Z_1(t)$ indicates that the roots z_1 and z_2 are either equal or nearly equal, and from $Z_2(t)$ their product is seen to be 144·0. Suspecting them to be 12·00, we divide the last five values of $Z_1(t)$ by 12·00, obtaining $c(t)$, and hence $\cos\theta$ as 1·000.

z_1 and z_2 are therefore 12·00 (bis), and the sum and product of z_3 and z_4 being thus $-2·000$ and 1·000, we find $z_3 = -2·414$, $z_4 = 0·414$.

(The actual roots are 12, 12, $-1 \pm \sqrt{2}$.)

(*b*) We have just seen that if the two roots z_m and z_{m+1} are suspected to be equal, Z_m/Z_{m-1} is divided by the presumed value of the roots, obtained as the square root of Z_{m+1}/Z_{m-1}; if by this means a sequence $c(t)$ is determined, then the test for the equality of z_m and z_{m+1} is that ultimately

$$1 - 2c(t) + c(t)c(t+1) \to 0 \quad . \quad . \quad . \quad . \quad . \quad (5.1)$$

This condition is really an *eliminant*, eliminating the arbitrary constants w_m and tw_{m+1} from three consecutive approximate equations; if it is expressed in determinant form

$$R_2(t) \equiv \begin{vmatrix} 1 & 1 & 1 \\ c(t) & 1 & 2 \\ c(t)c(t+1) & 1 & 3 \end{vmatrix} \to 0 \quad . \quad . \quad . \quad . \quad (5.2)$$

the extension to a multiplicity of order p is indicated, namely,

When the behaviour of the $p-1$ sequences $Z_m, Z_{m+1}, \ldots Z_{m+p-2}$ suggests that the roots $z_m, z_{m+1}, \ldots z_{m+p-1}$ may be all equal, Z_{m+p-2}/Z_{m-1} is divided by the $(p-1)$th power of their presumed value, the latter being obtained as the pth root of Z_{m+p-1}/Z_{m+1}; if this gives a sequence $c(t)$, the test for multiplicity p is that

$$R_p(t) \equiv \begin{vmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ c(t) & 1 & 2 & 2^2 & \cdots & 2^{p-1} \\ c(t)c(t+1) & 1 & 3 & 3^2 & \cdots & 3^{p-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c(t)\ldots c(t+p-1) & 1 & p+1 & (p+1)^2 & \cdots & (p+1)^{p-1} \end{vmatrix} \to 0 \quad . \quad (5.3)$$

On reducing and removing numerical factors, we have the simple extension of (5.1), viz.,

$$1-p\,.\,c(t)+\frac{p(p-1)}{2\,!}c(t)c(t+1)-\;.\;.\;.+(-)^{p}c(t)c(t+1)(\;.\;.\;.\;)c(t+p-1)\rightarrow 0,$$

which may be symbolically expressed as

$$(1-\mathrm{C})^{p}\rightarrow 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.4)$$

where C^r is to mean $c(t)c(t+1)\;.\;.\;.\;c(t+r-1)$.

Example. To solve the equation

$$z^5-19z^4+107z^3-91z^2-392z+686=0.$$

Following the usual routine, we obtain

Z_1.	Z_2.	Z_3.	$c(t)$.
19·0	107·0	91·0	...
13·368	90·84	551·9	...
11·354	78·58	357·9	...
10·312	71·59	324·8	...
9·674	67·43	349·3	...
9·243	64·52	342·1	1·317
8·932	62·41	342·9	1·274
8·697	60·80	343·1	1·241
8·513	59·53	343·0	1·215
8·365	58·51	...	1·194
8·244	...	...	...

Evidently z_1, z_2, z_3 are either equal or very nearly equal, and their product is 343. To test whether each is 7·00, we divide the last five values of Z_2 by 49·0 and obtain values of $c(t)$. The corresponding values of

$$1-3c(t)+3c(t)c(t+1)-c(t)c(t+1)c(t+2)$$

prove to be ·001, ·000, and ·000.

Hence the roots are 7·00, 7·00, 7·00, and the remaining roots are easily ascertained to be $-1·00 \pm 1·00i$.

(It may be noted, and is capable of proof in general, that Z_2/Z_1 is tending fairly rapidly to z_1; *e.g.* 58·51/8·365 = 6·995.)

§ 6. The Evaluation of Multiple Pairs of Complex Roots.

The case of multiple complex roots offers a close analogy with that of multiple real roots. If there are p coincident pairs of conjugate complex roots, z_m, z_{m+1}, . . . z_{m+2p-1}, then Z_m, Z_{m+2}, . . . Z_{m+2p-2} will fluctuate in sign and magnitude, Z_{m+1}, Z_{m+3}, . . . Z_{m+2p-3} will converge very slowly, but Z_{m+2p-1}, provided $|z_{m+2p-1}| > |z_{m+2p}|$, will converge more rapidly, and Z_{m+2p-1}/Z_{m-1} will approximate to r^{2p}, where r is the common modulus.

To find θ, the common amplitude, Z_{m+2p-2}/Z_{m-1} is divided by r^{2p-1} and a sequence $c(t)$ is obtained, the properties of which will now be discussed.

In the case of a single pair of complex roots, θ was found from the fact that

$$1-2c(t)\cos\theta+c(t)c(t+1)\rightarrow 0 \quad . \quad . \quad . \quad . \quad (6.1)$$

This relation was originally obtained as an eliminant

$$\begin{vmatrix} 1 & \cos t\theta & \sin t\theta \\ c(t) & \cos(t+1)\theta & \sin(t+1)\theta \\ c(t)c(t+1) & \cos(t+2)\theta & \sin(t+2)\theta \end{vmatrix} \rightarrow 0 \quad . \quad . \quad . \quad (6.2)$$

and the generalisation to p coincident pairs is found to be

$$\begin{vmatrix} 1 & \cos t\theta & \sin t\theta & 0 & 0 & \ldots 0 \\ c(t) & \cos(t+1)\theta & \sin(t+1)\theta & \cos(t+1)\theta & \sin(t+1)\theta & \ldots \sin(t+1)\theta \\ c(t)c(t+1) & \cos(t+2)\theta & \sin(t+2)\theta & 2\cos(t+2)\theta & 2\sin(t+2)\theta & \ldots 2^{p-1}\sin(t+2)\theta \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c(t)\ldots c(t+2p-1) & \cos(t+2p)\theta & \sin(t+2p)\theta & 2p\cos(t+2p)\theta & 2p\sin(t+2p)\theta & \ldots (2p)^{p-1}\sin(t+2p)\theta \end{vmatrix} \rightarrow 0 \quad . \quad (6.3)$$

This may be reduced to a form which when expressed symbolically is the exact analogue of (5.4), viz.

$$(1-2C\cos\theta+C^2)^p\rightarrow 0 \quad . \quad . \quad . \quad . \quad . \quad (6.4)$$

and on expanding and writing for C^r the product $c(t)\ldots c(t+r-1)$, we have an equation from which $\cos\theta$ may be determined.

Example.

$$z^6-18z^5+134z^4-416z^3-200z^2-1152z-2592=0.$$

The respective sequences are found to be

Z_1.	Z_2.	Z_3.	Z_4.	$c(t)$.	$\cos\theta$.
18·0	134·0	416·0	200·0		
10·556	78·12	351·82	2596·2		
7·495	58·31	262·24	1457·2		
5·239	50·51	180·05	1217·0		
2·380	46·98	74·25	1304·6	·34375	
−8·396	46·77	−410·61	1297·4	−1·90097	
16·831	46·45	554·54	1295·6	2·56731	
8·540	45·03	299·47	1296·0	1·38644	·836
5·942	43·30	209·84	1296·0	·97148	·833
3·725	42·20	129·79	...	·60088	·833
−0·500	41·90				
94·494					

The four greatest roots are evidently two complex pairs of equal or nearly equal moduli. Also their product is 1296·0. Taking the fourth root of this and cubing it we have 216·0, and dividing z_3 by this we have a sequence of $c(t)$. Putting $p=2$ in (6.4) and using the last six values of

$c(t)$, we have three consecutive quadratic equations in $\cos\theta$, and the smaller root in each is respectively ·836, ·833, ·833. Hence $\cos\theta = \cdot 833$, $\sin\theta = \pm\cdot 553$, and the roots are

$$5{\cdot}00 \pm 3{\cdot}32i \text{ (bis).}$$

The remaining roots are then found to be $-1{\cdot}00 \pm 1{\cdot}00i$.

(The actual roots are $5 \pm i\sqrt{11}$ (bis), $-1 \pm i$.)

§ 7. Numerically Equal Roots.

The evaluation of roots of the same modulus but different amplitudes presents difficulties. In such a case it is better to transform the equation by a linear change of variable. There is, however, one case which is easily recognised, that of real roots equal numerically but opposite in sign. It is not difficult to see that the alternation of sign in successive powers of the negative root will produce a sequence Z_m which *oscillates*, the alternate terms tending towards *two different limits*. The values of z_m and z_{m+1} will then be found from the *square root of the product of the two limits with sign reversed*, or more readily from the product $z_m z_{m+1}$ as given by Z_{m+1}/Z_{m-1}.

Example.

$$z^4 - 10z^3 - 12z^2 + 130z - 13 = 0.$$

Z_1.	Z_2.	Z_3.
10·000	− 12·00	−130·00
11·200	−120·33	−128·80
9·911	− 12·108	−128·69
10·051	−106·84	−128·69
9·901	− 12·122	−128·69
9·919	−105·30	−128·69
9·899	− 12·123	−128·69
9·902	−105·10	−128·69
9·899	− 12·124	
9·899		

z_1 is 9·899, and z_2 and z_3 are equal in value but opposite in sign. $z_1 z_2 z_3$ is $-128{\cdot}69$, and so $-z_2 z_3$ is 128·69/9·899, or 13·00. Extracting the square root, we have z_2 and z_3 as $\pm 3{\cdot}606$. z_4 is then found to be ·1010.

(The actual roots are $5 \pm \sqrt{24}$, $\pm\sqrt{13}$.)

§ 8. Derived Sequences of more Rapid Convergence.

So far we have not obtained a very great degree of accuracy in the numerical examples. We shall now proceed to derive from the primary sequences successive sequences of increasing approximative power.

Consider first the sequence of approximations to z_1, namely,

$$Z_1(t)=\frac{f(t+1)}{f(t)},$$

and construct their finite differences of first and second order. It is an immediate consequence of (2.1) that $\Delta Z_1(t)$ tends to become a geometric sequence (provided z_2 is not one of a pair of conjugate complex roots, or of a pair equal in value but opposite in sign) of common ratio $\frac{z_2}{z_1}$. Hence the *deviations* of $Z_1(t)$ from z_1 will also tend to become a geometric sequence with the same common ratio. Thus a further approximate relation is suggested, viz.

$$\frac{z_1-Z_1(t+2)}{z_1-Z_1(t+1)}=\frac{\Delta Z_1(t+1)}{\Delta Z_1(t)} \quad . \quad . \quad . \quad . \quad . \quad (8.1)$$

and solving for z_1 we are led to investigate the *derived* sequence

$$Z_1^{(1)}(t)=\frac{\begin{vmatrix} Z_1(t+1) & Z_1(t+2) \\ Z_1(t) & Z_1(t+1) \end{vmatrix}}{\Delta^2 Z(t)} \quad . \quad . \quad . \quad . \quad . \quad (8.2)$$

Retention of terms of higher order shows that $Z_1^{(1)}(t)$ tends to z_1, but its first differences $\Delta Z_1^{(1)}(t)$ and the deviations $z_1-Z_1^{(1)}(t)$ tend to form geometric sequences * each of ultimate common ratio $\left(\frac{z_2}{z_1}\right)^2$ or $\frac{z_3}{z_1}$, whichever is the greater. Thus the convergency of the derived sequence $Z_1^{(1)}(t)$ is more rapid than that of the primary sequence $z_1(t)$, and it also commences at a more advanced stage of approximation. In the same way we can form a second derived sequence

$$Z_1^{(2)}(t)=\frac{\begin{vmatrix} Z_1^{(1)}(t+1) & Z_1^{(1)}(t+2) \\ Z_1^{(1)}(t) & Z_1^{(1)}(t+1) \end{vmatrix}}{\Delta^2 Z_1^{(1)}(t)} \quad . \quad . \quad . \quad . \quad . \quad (8.3)$$

Retention of terms of still higher order shows that the deviations and differences of the second derived sequence possess similar properties to those of the earlier sequences, and indeed it may be shown generally that when the equation possesses only real roots we can proceed indefinitely forming derived sequences in this way, each converging to z_1 with greater rapidity than the preceding.

This property of derived sequences is not peculiar to those for the greatest root z_1, but holds for sequences $Z_m(t)$, $Z_m^{(1)}(t)$, $Z_m^{(2)}(t)$, . . . derived in

* Nägelsbach, in the course of a very detailed investigation of Fürstenau's method of solving equations, obtains the formulæ (8.2) and (8.4), but only incidentally. Cf. *Archiv d. Math. u. Phys.*, **59** (1876), pp. 147–192; **61** (1877), pp. 19–85, and especially pp. 22, 31.

the same way, converging to $z_1 z_2 \ldots z_m$, with the restrictions on z_{m+1}, $z_{m+2}, \ldots$ already mentioned.

Since three consecutive terms of one sequence define a term of the derived sequence, p consecutive terms of a primary sequence determine $p-2$ of a derived sequence, $p-4$ of a second derived sequence, and so on. With tabular or mechanical aids to calculation this method of successive approximation is of considerable power, as we shall show.

Example. To obtain a more accurate value for the greatest root z_1 of the example of § 3.

The table below shows the primary and derived sequences obtained from the last five values of Z_1, to eleven significant digits. Recalling that z_1 was 15, we can compare the sequences and observe the notable improvement in the approximation.

Z_1.	$Z_1^{(1)}$.	$Z_1^{(2)}$.
14·667121186	15·001418373	14·999999987
15·159777145	15·000304169	
14·926402783	15·000065221	
15·034550817		
14·983920543		

A similar improvement can be made in the sequences for the remaining roots.

Where there is evidence that z_{m+1} and z_{m+2} are roots equal in value but opposite in sign, the sequence Z_m will really consist of *two alternating sequences*, each converging to $z_1 z_2 \ldots z_m$, but one converging more rapidly than the other. In such a case we choose the alternate terms which have the smaller first differences, and use them to form derived sequences in the manner described.

When the roots z_{m+1} and z_{m+2} form a complex pair, then the first differences of Z_m, as well as its deviations from $z_1 z_2 \ldots z_m$, would tend to be geometric sequences of common ratio $\frac{r_{m+1}}{z_m}$ but for the presence of cosine factors. Eliminating the latter, we may obtain in such a case the improved sequence

$$Z_m^{(1)}(t) = \frac{\begin{vmatrix} Z_m(t+1) & Z_m(t+2) \\ Z_m(t) & Z_m(t+1) \end{vmatrix} - h_t^2 \begin{vmatrix} Z_m(t+2) & Z_m(t+3) \\ Z_m(t+1) & Z_m(t+2) \end{vmatrix}}{\Delta^2 Z_m(t) - h^2 \Delta^2 Zm(t+1)} \quad . \qquad (8.4)$$

where h_t is the approximate value of $\frac{z_m}{r_{m+1}}$ obtained from $Z_{m-1}(t+3)$, $Z_m(t+3)$, $Z_{m+2}(t+3)$. This formula, while not quite so simple as (8.2) and

(8.3), is equally effective in practice. It may even be proved to possess the reproductive property of the other approximations, *i.e.* it gives a succession of derived approximations of the same form, in which h^2 is replaced by other coefficients, but the complete formulation of these is rather tedious and of no great practical value.

The foregoing remarks suffice to show that in actual computation it will in general be unnecessary to carry the calculation of primary sequences to any great length. It will most often be advantageous to arrest these at an early stage and form derived sequences.

In passing, it is interesting to note that in the case of a quadratic equation the sequence Z_1 consists of the successive convergents of a continued fraction, and each derived sequence is merely the preceding sequence carried to a more advanced stage of approximation and with alternate terms deleted.

§ 9. Connection with Fürstenau's Method.

So far the extension of Bernoulli's method has been developed from the standpoint of mechanical application. Returning to theoretical considerations, we can choose special functions for $f(t)$ and thereby exhibit the relation to other methods.

In the first place, we can put $f(t)=s_t$, the sum of the tth powers of all the roots of the given equation, where s_t is readily calculated for the initial values of t by Newton's recurrence relations. The particular advantage of this choice of $f(t)$ has been pointed out by Lagrange (*loc. cit.*) in commenting on some illustrative examples of Euler, with special reference to equal roots.

Again, if we put * $f(t)=\mathrm{H}_t$, the complete symmetric or "aleph" function of degree t of all the roots, we obtain results first used † by Fürstenau. Fürstenau's principal result is as follows:—

$$\lim_{t\to\infty}\frac{\mathrm{A}(m,\,t+1)}{a_0\mathrm{A}(m,\,t)}=(-)^t z_1 z_2\,.\,.\,.\,z_m \qquad (9.1)$$

where $\mathrm{A}(m,\,t)$ denotes the persymmetric determinant of order t

$$\begin{vmatrix} a_m & a_{m+1} & a_{m+2} & \cdots \\ a_{m-1} & a_m & a_{m+1} & \cdots \\ a_{m-2} & a_{m-1} & a_m & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}_t .$$

* The choice of initial values of $f(t)$, . . . 0, 0, 0, 1, is equivalent to this.

† *Darstellung der reellen Würzeln algebraischer Gleichungen durch Determinanten der Coefficienten*, Marburg, 1860.

Now it is known that $H_t = (-)^t A(1,\ t)/a_0^t$. Fürstenau's result will therefore follow directly from (2.1) if it can be shown that

$$A_m(1,\ t) \equiv \begin{vmatrix} A(1,\ t) & A(1,\ t+1) & \dots \\ A(1,\ t-1) & A(1,\ t) & \dots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{vmatrix}_m = A(m,\ t)\ . \qquad . \quad (9.2)$$

First assume (9.2) true for 1, 2, 3, . . . $m-1$. Then, by the extensional theorem in determinants (3.1), applied twice, we have

$$\begin{aligned} A_m(1,\ t) &= \{A_{m-1}(1,\ t)^2 - A_{m-1}(1,\ t+1)A_{m-1}(1,\ t-1)\} \div A_{m-2}(1,\ t) \\ &= \{A(m-1,\ t)^2 - A(m-1,\ t+1)A(m-1,\ t-1)\} \div A(m-2,\ t) \\ &= \{A(m-1,\ t)^2 - [A(m-1,\ t)^2 - A(m,\ t)A(m-2,\ t)]\} \div A(m-2,\ t) \\ &= A(m,\ t). \end{aligned}$$

But when $m=1$ the result (9.2) is a mere identity. Hence in succession it is true for general values of m.

Finally, it may be shown that all the preceding results are special cases of a more general result, involving symmetric functions of assigned degree formed from a given number of the roots of (1.1) in descending order. A quotient of two determinants of order m is considered. The denominator consists, as in § 2, of elements $f(t)$; the differences of the arguments t from column to column are all unity, but from row to row may be any integers. Let this denominator be typified by its first row as $|f(t)f(t+1)f(t+2)\ .\ .\ .\ f(t+m-1)|$. As numerator we take a similar determinant $|f(t+k)f(t+k+p+1)f(t+k+q+2)\ .\ .\ .\ f(t+k+s+m-1)|$. Then the limit of the quotient with increasing t is

$$(z_1z_2\ .\ .\ .\ z_m)^k \begin{vmatrix} H_{m,\,p} & H_{m,\,q+1} & \dots & H_{m,\,s+m-2} \\ H_{m,\,p-1} & H_{m,\,q} & \dots & H_{m,\,s+m-3} \\ H_{m,\,p-2} & H_{m,\,q-1} & \dots & H_{m,\,s+m-4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{m,\,p-m+2} & H_{m,\,q-m+3} & \dots & H_{m,\,s} \end{vmatrix} \qquad . \quad . \quad . \quad (9.3)$$

$$= (z_1z_2\ .\ .\ .\ z_m)^k \begin{vmatrix} c_{m,\,p'} & c_{m,\,q'+1} & \dots & c_{m,\,s'+m-2} \\ c_{m,\,p'-1} & c_{m,\,q'} & \dots & c_{m,\,s'+m-3} \\ c_{m,\,p'-2} & c_{m,\,q'-1} & \dots & c_{m,\,s'+m-4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ c_{m,\,p'-m+2} & c_{m,\,q'-m+3} & \dots & c_{m,\,s'} \end{vmatrix} \qquad . \quad . \quad . \quad (9.4)$$

where $H_{m,\,r}$ denotes the complete symmetric or aleph function of degree r formed from the greatest m roots $z_1,\ z_2,\ .\ .\ .\ z_m$, while $c_{m,\,r}$ denotes the elementary symmetric function of the same degree and constituents, and $(p, q, \ldots s)$, $(p', q', \ldots s')$ are conjugate partitions of the integer $p+q+\ldots+s$.

The proof may be summarised as follows. The general expressions for $f(t)$ are substituted in the determinants as in § 2. It is seen as in § 2 that the common factor $w_1 w_2 \ldots w_m$ may be removed, so that s_t may be written for $f(t)$. But then the determinants of the s's are each expressible as the product of two alternants, one of which, common to numerator and denominator, is cancelled out, leaving finally $(z_1 z_2 \ldots z_m)^k$ multiplied by the quotient of an alternant and a difference-product. Such a quotient was shown by Jacobi* to be equivalent to a determinant of aleph functions of the form given in (9.3), the alternative form in elementary symmetric functions being due† to Nägelsbach.

The simplest cases of (9.3) and (9.4) are the two extremes: (i) when the last r of $k, k+p, \ldots k+s$ are all unity and the first $m-r$ all zero, the limit of the quotient in question is the *elementary symmetric function* of degree r of the m greatest roots, which for $m=r$ becomes the simple product, as we have seen in § 2; (ii) when all of $k, k+p, \ldots k+s$ are zero except the last, which is equal to r, the limit is the *aleph function* of degree r of the m greatest roots, and if $r=1$ we have merely the *sum* of these m roots. Such facts indicate alternative methods of approximating by sequences of quotients of determinants to the roots of equations, but none of the methods seems to admit of an arithmetical process as simple and economical as that described in the earlier sections of this paper.

* De functionibus alternantibus, . . . *J. für Math.*, **22** (1841), pp. 370–371.

† Sch.-Programm, Zweibrücken, 1871. Also "Studien zu Fürstenau's neuer Methode," . . . *Archiv d. Math. u. Phys.*, **59** (1876), pp. 150–151. *Cf.* Muir, *History of the Theory of Determinants*, vol. iii, pp. 144, 154.

(*Issued separately October* 21, 1926.)