University of Maryland College Park

Institute for Advanced Computer Studies TR–92–121
Department of Computer Science TR–2990

On Infinitely Many Algorithms for Solving Equations*

E. Schröder

Translated by
G. W. Stewart†

November 1992
(Revised January 1993)

ABSTRACT

This report contains a translation of "Ueber unendlich viele Algorithmen zur Auflösung der Gleichungen," a paper by E. Schröder which appeared in *Mathematische Annalen* in 1870.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742.

On Infinitely Many Algorithms for Solving Equations

Ernst Schröder

Translated by
G. W. Stewart

TRANSLATOR'S INTRODUCTION

Schröder begins his remarkable paper on the solution of a nonlinear equation in a single unknown by thanking a certain H. Eggers for communicating most of the important results in the paper. There is a whiff of mystery here. At some point before Schröder's paper appeared, Dr. Eggers emigrated to America, and in 1876 he published two short papers in *The Analyst* [2, 3] on the solution of nonlinear equations. The mystery is that the papers are inconsequential, and it is difficult to reconcile them with Schröder's sweeping acknowledgement. By way of contrast, Schröder went on to publish important, if somewhat neglected, work in mathematical logic [9]. It would be nice to know if Schröder was simply being overgenerous in his acknowledgement.

In any event, Schröder had a great deal to be generous about. A. S. Householder used to claim you could evaluate a paper on root finding by looking for a citation of Schröder's paper. If it was missing, the author had probably rediscovered something already known to Schröder. This observation was intended as mild hyperbole, since much was done after Schröder; however, it is safe to say that Schröder's paper contains the first systematic, general derivation and analysis of algorithms for solving equations.

Unfortunately, Schröder is only a middling expositor. Line by line he is a considerate writer, giving his reader all the hints necessary to follow his reasoning. But at a higher level he often plows into a thicket of details without bothering to tell one where he is going. For this reason, I will now give a summary of the main ideas in the paper — a sort of road map of the territory.

Schröder's goal is to find the roots of the equation

$$f(z) = 0,$$

where f is analytic about the roots in question. (Although Schröder's mathematical language is not quite ours — for example, he writes $\lim_{\omega=\infty}$ instead of $\lim_{\omega\to\infty}$ — he is rigorous, even by today's standards.) He begins by distinguishing two kinds of methods. The first is typified by Newton's method and consists of the successive substitution of iterates in a fixed formula. Schröder calls such methods *algorithms,* a usage more restricted than ours today. The essence of the second kind of method consists in constructing a sequence of functions $F_\omega(z)$ having the property that $\lim_{\omega\to\infty} F_\omega(z)$ is a root of the equation, the particular root depending on the choice of z. Bernoulli's method can be regarded as a method of the second kind.

At this point Schröder's paper divides naturally into two parts. The first consists of a general treatment of both kinds of methods. The second consists of a systematic way of deriving a two dimensional table of functions (actually two such tables) that can be used to construct algorithms of both kinds. In the second part, Schröder restricts himself to polynomial equations, but as he notes his results are more widely applicable. We will treat each part in turn.

Schröder begins the first part with a careful discussion of the properties of iterations of the form $z' = F(z)$. He derives the now classical result that for such an iteration to converge to a root z_1, the root must be a fixed point of F, and the absolute value of $F'(z_1)$ must be less than one (actually, convergence can occur when $|F'(z)| = 1$, a case missed by Schröder). For the case $0 < |F'(z_1)| < 1$, he calls the rate of convergence linear, as we do today. He then goes on to give the usual conditions for quadratic, cubic, and general ωth order convergence, observing that in the limit each iteration of an ωth order method increases the number of correct digits by a factor of ω.

Schröder now turns to the problem of writing down the most general form of an algorithm with convergence of order ω. He first proceeds by special cases, showing, for example, that the most general quadratically converging algorithm has the form

$$F = z - \frac{f}{f_1} - f^2\varphi_2,$$

where f_1 is the first derivative of f and φ_2 is arbitrary. He then writes down the remarkable formula

$$F = z + \sum_{a=1}^{a=\omega-1} (-1)^a \frac{f^a}{a!} \cdot \left(\frac{1}{f_1}\partial\right)^{a-1} \frac{1}{f_1} - f^\omega \cdot \varphi_\omega.$$

for the general ωth order algorithm and establishes that it has the required convergence rate.

A class of methods of the second kind can be obtained by taking the limit as $\omega \to \infty$ in the above expression, provided the limit exists. Schröder gives a recurrence for the individual terms of the series and an alternative representation. He does not investigate the convergence of the series in general; however, he treats the case of a quadratic equation in tedious detail.

Schröder now moves to the second part of his paper, in which he gives a uniform treatment of a class of algorithms of the first and second kinds. Today the natural approach to these results would be through König's theorem, and it will clarify things if we so describe them.

König's theorem states that if an analytic function has a single, simple pole at the radius of convergence of its power series, then the ratios of the coefficients of its power series converge to that pole. The application to root finding is as follows. The function

$$\frac{1}{f(z-\epsilon)} = \mathcal{A}_0^{(0)}(z) + \mathcal{A}_1^{(0)}(z)\epsilon + \mathcal{A}_2^{(0)}(z)\epsilon^2 + \cdots$$

has a pole at $\epsilon = z - z_1$, where z_1 is a root of $f(z) = 0$ that is nearest z. If z_1 is unique and simple, then by König's theorem

$$\frac{\mathcal{A}_{\omega-1}^{(0)}(z)}{\mathcal{A}_\omega^{(0)}(z)} \to z - z_1.$$

Now consider the related expansion

$$\frac{z-\epsilon}{f(z-\epsilon)} = \mathcal{A}_0^{(1)}(z) + \mathcal{A}_1^{(1)}(z)\epsilon + \mathcal{A}_2^{(1)}(z)\epsilon^2 + \cdots$$

It is easily verified that

$$\mathcal{A}_\omega^{(1)}(z) = z\mathcal{A}_\omega^{(0)}(z) - \mathcal{A}_{\omega-1}^{(0)}(z)$$

(here we assume $\mathcal{A}_{-1}^{(0)}(z) = 0$). Hence if we set

$$F_\omega(z) = \frac{\mathcal{A}_\omega^{(1)}(z)}{\mathcal{A}_\omega^{(0)}(z)} = z - \frac{\mathcal{A}_{\omega-1}^{(0)}(z)}{\mathcal{A}_\omega^{(0)}(z)},$$

it follows that

$$\lim_{\omega\to\infty} F_\omega(z) = z_1,$$

which, in Schröder's terminology, is a method of the second kind. Moreover, it can be shown that the iteration

$$z' = F_\omega(z)$$

is locally convergent with order ω; i.e., it is a method of the first kind.

This is Schröder's development, or might have been if he had known König's theorem. Lacking it, he proceeds indirectly. First, he restricts himself to a rational function f whose roots are z_1, z_2, ..., z_n. He then introduces the functions

$$C_\omega^{(\lambda)}(z) = \sum_{a=1}^{a=n} \frac{z_a^\lambda \chi(z_a)}{(z - z_a)^{\omega+1}},$$

where χ is an essentially arbitrary function. Because he has an explicit formula for the functions C, Schröder is able to define

$$F_\omega(z) = \frac{C_\omega^{(1)}(z)}{C_\omega^{(0)}(z)}$$

and establish the properties mentioned above. He then defines functions $A_\omega^{(\lambda)}$ by taking $\chi(z) = \dfrac{1}{f'(z)}$, and shows that the functions A are essentially the same as the functions $\mathcal{A}$ from the power series. In other words, Schröder passes from certain functions, the A's, whose properties he can establish but which he cannot compute (because he does not actually know the roots) to certain functions, the $\mathcal{A}$'s, whose properties he cannot establish but which he can compute (because he knows the function f and its derivatives). The proof of the identity of the two completes his development.

Incidentally, this description does not do justice to Schröder's virtuosity in finding elegant representations and recurrences for his functions.

The rest of the paper, which is largely devoted to determining convergence regions, is an anticlimax. The task is almost trivial for Schröder's methods of the second kind. For his algorithms the difficulties are almost insuperable, and he is able to obtain results only for the quadratic equation. (Curiously, a few years later Cayley [1] proposed essentially the same problem for Newton's method, commenting that, "The solution is easy and elegant in the case of a quadric equation, but the succeeding case of the cubic equation appears to present considerable difficulty.")

Schröder wrote one more paper on root finding, or rather on iterated functions [8], and then turned his attentions elsewhere. He does not seem to have had much influence on his contemporaries. He is one of a group of people — König [6] and Hadamard [4] among others — who were concerned with extracting the information contained in the coefficients of power series. Yet he is not cited by these people, and there is no evidence that he influenced subsequent developments by Aitkin and Rutishauser that lead to the qd, and ultimately to the QR algorithm (for a survey of this development see [5]). Certainly Schröder deserves credit for the polynomial case of König's theorem. He was also the first to show how by successive origin shifts the ratios of coefficients in a power series could be made to yield algorithms of high order convergence. And the generality of his approach makes him the rediscoverer of some iterative methods but the discover of infinitely many more.

A word on the translation. The page layout and notation is roughly that of the original, as it appeared in *Mathematische Annalen.* Schröder wrote in a style that is convoluted, even by nineteenth century standards, and I have not labored to conceal his shortcomings. On the other hand, I have not attempted to render him in the prose of a century ago. Instead I have looked to the English of our own day, with a little musty elaboration.

Acknowledgement. I am indebted to Thomas Scavo for a detailed reading of the text and many useful suggestions.

College Park 1992

References

[1] A. Cayley. The Newton–Fourier imaginary problem. *American*

Journal of Mathematics, 2:97, 1897.

[2] H. Eggers. Calculation of radicals. *The Analyst*, 3:100–102, 1876.

[3] H. Eggers. A new method of solving numerical equations. *The Analyst*, 3:100–102, 1876.

[4] M. J. Hadamard. Essai sur l'étude des fonctions données par leur développement de taylor. *Journ. de Math. (*4e série*)*, 8:101–186, 1892.

[5] A. S. Householder. *The Numerical Treatment of a Single Non-linear Equation.* McGraw–Hill, New York, 1970.

[6] J. König. Über eine Eigenschaft der Potenzreihen. *Mathematische Annalen*, 23:447–449, 1884.

[7] E. Schröder. Über unendliche viele Algorithmen zur Auflösung der Gleichungen. *Mathematische Annalen*, 2:317–365, 1870. Dated Pforzheim, January 1869.

[8] E. Schröder. Ueber iterirte Functionen. *Mathematische Annalen*, 3:296–322, 1871. Dated Pforzheim, June 1869.

[9] H. Wussing. Schröder, Friedrich Wilhelm Karl Ernst. In C. C. Gillispe, editor, *Dictionary of Scientific Biography, XII.* Charles Scribner's Sons, New York, 1973.

On Infinitely Many Algorithms for Solving Equations

By Dr. E. Schröder at Pforzheim

In this paper the frequently treated problem of solving an equation will be considered from what to my knowledge is a new viewpoint, one which is the common source of the various well-known solution methods and of infinitely many others that have not yet been considered. The investigations are concerned with equations in *one* unknown, not only algebraic equations but also transcendental. I was inspired to this work in 1867 by some communications of Dr. Heinrich Eggers of Meklenburg, who was formerly Professor at the Gymnasium at Schaffhausen and has now emigrated to America. Specifically, I thank him for the knowledge of a great part of the results in §§ 2, 3, 7, 8, 11, 12, and 15, as well as the results I have obviously derived from them.

§ 1.
The Nature of the Solution Methods and the Condition for Their Applicability.

Let $f(z)$ be any single-valued function of the complex argument $z = x + iy$ (which we shall always think of as represented by a point in the complex plane). Then the problem that forms the object of the following deliberations consists of solving the equation

$$f(z) = 0; \tag{1}$$

that is, of finding some number (root) z_1 with the property that

$$f(z_1) = 0. \tag{2}$$

We will only consider those roots of the arbitrary algebraic or transcendental equation (1) at and about which the function $f(z)$ is continuous and for which the zero is of finite order.[1] If we take the symbols z_1 to denote any such root, then the function $f(z)$ will be single-valued, continuous, and finite in some neighborhood T containing the point z_1, and in addition the function will be zero with finite order.

Now it is well known from the theory of functions*) that the degree of multiplicity of the root z_1 — i.e., the order of the zero of $f(z)$ or the pole of the reciprocal function $\dfrac{1}{f(z)}$ — must be a (positive) integer p, so that we may write

$$f(z) = (z - z_1)^p \psi(z), \tag{3}$$

where $\psi(z)$ is single-valued in T and the limit

$$\lim_{z=z_1} \psi(z)$$

is different from 0 and ∞. Further, since the derivative $d_z\psi(z) = \psi^{(1)}(z)$ is itself single-valued and since**)

$$\lim_{z=z_1} (z - z_1)\psi(z) = 0, \tag{4}$$

it follows easily from the relation

$$f^{(1)}(z) = (z - z_1)^{p-1}\{p\psi(z) + (z - z_1)\psi^{(1)}(z)\}, \tag{5}$$

obtained by differentiating (3), that the derivative $f^{(1)}(z)$ of our function has a zero of order $p - 1$ at z_1.

Hence the equation

$$\frac{f(z)}{f^{(1)}(z)} = 0 \tag{6}$$

must have the same roots as equation (1), only these roots are simple. The function f has additional roots at those points in T—and only those points—where f becomes infinite.

[1]Translator's note: Schröder is evidently using the word continuous in its older sense of differentiable or analytic.

*) For example, see B. DURÈGE, Elemente der Theorie der Functionen einer complexen veränderlichen Grösse, Leipzit, 1864, § 29.

**) Ibid. § 24. and 27.

Under these assumptions, one has a choice of different methods for solving the equation (1). Now there are a large number and, as we shall show, even an infinite number of solution methods, all characterized by the fact that one begins calculating in a fixed way with an almost completely arbitrary number z and by a sufficiently extended sequence of operations arrives at a result that comes as near to the root z_1 as desired. The *starting value* z can often be regarded as a first (or zero-th) approximation of the root z_1, from which the algorithm of the solution method generates successively better, more precise approximations. But often it is not necessary to proceed by means of successive approximations. In both cases, which root of the equation $f(z) = 0$ is found depends on choice of the starting value. Otherwise, the starting value appears as a constant, arbitrary within some region, whose influence on the final result diminishes as the calculations proceed. Solution methods of both kinds are the object of the following investigation.

§ 2.
Methods of the First Kind (Algorithms).

Without assuming anything for now about the nature of the function f, we shall first solve the equation (1) by an algorithm whose repeated application, starting from a zero-th approximation, yields successively more precise approximations to the root z. Within certain limits, the zero-th approximation can be chosen or estimated arbitrarily. Our problem then is to find a function F for which the equation

$$z' = F(z) \tag{7}$$

always gives a point z' lying nearer the root z_1 than the original point z.

If we suppose that we already know such a function, then the calculation of

$$z'' = F(z')$$

gives us a new point which is even nearer the desired root z_1 than its two predecessors z and z'. In general, if we define

$$z^{(r)} = F(z^{(r-1)}), \tag{8}$$

then because of the assumptions concerning F the distance of the last point $z^{(r)}$ from the root z_1 will be smaller than that of each of the preceding points

$$z^{(0)} = z, \quad z^{(1)} = z', \quad z^{(2)} = z'', \ldots, \quad z^{(r-1)}.$$

In other words the absolute values of the differences

$$z - z_1,\ z' - z_1,\ z'' - z_1,\ \ldots,\ z^{(r-1)} - z_1,\ z^{(r)} - z_1$$

form a decreasing sequence. We may call these differences the *errors* in the approximations.

An additional condition that must be imposed on the function F is that the distances do not just decrease monotonically with increasing r but that they actually approach zero, so that

$$\lim_{r=\infty} z^{(r)} = z_1. \tag{9}$$

If this condition along with the previous condition, which can be written analytically in the form

$$\text{mod.}\ \{z^{(r)} - z_1\} < \text{mod.}\ \{z^{(r-1)} - z_1\}, \tag{10}$$

is satisfied for all values of r, or at least from a certain value on to $r = \infty$, then in fact (7) gives an *algorithm* of the kind desired. By successive application of the algorithm, we may determine the root z_1 of equation (1) to arbitrary precision. We shall say simply that the algorithm *converges* to the root z_1 from the starting value z.

The solution of the equation $f(z) = 0$ may be represented symbolically as

$$z_1 = \lim_{r=\infty} F^r(z), \tag{11}$$

where we write

$$\underbrace{F}_{1}(\underbrace{F}_{2}\{\cdots \underbrace{F}_{r-1}[\underbrace{F}_{r}(z)]\cdots\}) = F^r(z) \tag{12}$$

for the r-fold repeated or *iterated* function.

The conditions (9) and (10) that the function F must satisfy may be recast in a more useful way.

Specifically, the starting value is to be any value in a certain region U surrounding the root z_1, a region which we may call the *convergence region* of the algorithm for the root z_1. Thus, if $z = z_1+\epsilon$ represents a starting value, chosen near enough z_1 so that it lies in this convergence region and we calculate the next approximation

$$F(z_1 + \epsilon) = z_1 + \epsilon',$$

then by (10) we always have $|\epsilon'| < |\epsilon|$. Consequently, if ϵ tends toward 0 in any manner, so must ϵ', and we must have

$$\lim_{\epsilon=0} F(z_1 + \epsilon) = z_1.$$

This equation shows first that the function F must be continuous at z_1 and second that it must satisfy the condition

$$F(z_1) = z_1. \tag{13}$$

Of course if the function F were continued in several branches, a line of discontinuity or a cut could proceed from the point z_1. However, we shall restrict our investigations to functions F that are single valued in the convergence region U, or at least in a part of it surrounding the point z_1.

In this case, $F(z)$ or $F(z_1 + \epsilon)$ can be developed in a Taylor series

$$F(z_1 + \epsilon) = F(z_1) + \epsilon F^{(1)}(z_1) + \frac{\epsilon^2}{2} F^{(2)}(z_1) + \cdots$$

inside of some circle with center z_1; i.e., for all sufficiently small ϵ. From (7) and (13) it follows that

$$z' = z_1 + \epsilon F^{(1)}(z_1) + \frac{\epsilon^2}{2} F^{(2)}(z_1) + \cdots. \tag{14}$$

If we now assume that $F^{(1)}(z_1)$ is different from zero, then for sufficiently small ϵ the term $\epsilon F^{(1)}(z_1)$ will dominate all the following terms, and for infinitely small ϵ we can write

$$z' - z_1 = \epsilon F^{(1)}(z_1).$$

According to the requirement (10), the absolute value of this difference is to be smaller than that of $z - z_1$. Thus we obtain a second condition which the function F must satisfy:

$$\text{mod. } F^{(1)}(z_1) < 1. \tag{15}$$

If $F^{(1)}(z_1)$ were zero, then this condition would automatically be satisfied.

Now in general if the condition (15) is satisfied, then the absolute value of the error in the first approximation, i.e. $z' - z_1$, will be only a fraction of the absolute value of the previously hypothesized error, i.e. $z - z_1 = \epsilon$. Since any number can be brought arbitrarily close to zero by repeated multiplication by a constant proper fraction, it is not difficult to prove that with repeated application of the algorithm the error in the approximation actually tends to zero, or in other words that the requirement (9) is satisfied.

Thus if we confine ourselves to a single-valued function F, we may replace the two conditions (9) and (10) by (13) and (15), and, incidentally, state the following theorem.

If $F(z)$ is a function that is single valued about z_1 and satisfies the conditions (13) *and* (15), *then for any number z close enough to z_1 equation* (11) *is satisfied. In other words, for all points z lying in a certain neighborhood about the point z_1, the unboundedly iterated function $F(z)$ tends in the limit to the root z_1 of the equation $F(z) = z$.*

We have found that under the assumptions (13) and (15) equation (7) gives an algorithm of the kind we desire.

If the case $F^{(1)}(z_1) = 0$ is excluded, the convergence of the corresponding algorithm may be called *linear* or *of the first order,* because the error in the approximation is nearly proportional to the first power of the error in the starting value. The smaller the error in the starting value, the stricter the proportionality.

We obtain a much more useful algorithm by choosing the function F so that in addition to the previous conditions we have

$$F^{(1)}(z_1) = 0. \tag{16}$$

In this case, if the next highest derivative $F^{(2)}(z_1)$ is not zero, then for infinitely small ϵ the error in first approximation

$$z' - z_1 = \frac{\epsilon^2}{2} F^{(2)}(z_1)$$

is proportional to the square of the original error. Since the error is of the order of the square, we may call the approximation *quadratic*.

Similarly, if F is defined so that

$$F^{(1)}(z_1) = 0,\ F^{(2)}(z_1) = 0,\ \ldots,\ F^{(\omega-1)}(z_1) = 0, \tag{17}$$

while $F^{(\omega)}(z_1)$ is not zero, we get an algorithm that produces an approximation of the ωth order. For then

$$z' - z_1 = \frac{\epsilon^\omega}{2 \cdot 3 \cdots \omega} F^{(\omega)}(z_1)$$

when ϵ is infinitely small, and the error in the approximation is proportional to the ωth power of the error in the starting value.

In practice if the zero-th approximation or starting value is exact to s places beyond the decimal point in absolute value, then for a quadratically convergent algorithm the following value or first approximation will be exact to $2s$ places, and for an algorithm of the ωth order it will be exact to approximately ωs places. More precisely, the number s can be taken so large that for it and any larger number the assertion is strictly true.

We may now recapitulate our results.

Let z_1 be a root of an equation $f(z) = 0$, and let $F(z)$ be a function that is single valued in a region surrounding z_1 and takes on the value z_1 at the point z_1, so that equation (13),

$$F(z_1) = z_1,$$

is satisfied. Then equation (7),

$$z' = F(z),$$

defines an algorithm that converges to z_1 from any point in some region surrounding z_1, provided that (15),

$$|F^{(1)}(z_1)| < 1,$$

holds. If $|F^{(1)}(z_1)| > 0$, the convergence is only linear, whereas it is of the ωth order when (17),

$$F^{(1)}(z_1) = 0,\ F^{(2)}(z_1) = 0,\ \ldots,\ F^{(\omega-1)}(z_1) = 0,\ |F^{(\omega)}(z_1)| > 0,$$

holds.

§ 3.
Examples of Algorithms.

We will suppose that the assumption mentioned in § 1 that f is single valued holds here.

I. Newton's method is the most famous algorithm of the kind considered here for solving an equation $f(z) = 0$. It is given by the formula

$$z' = z - \frac{f(z)}{f^{(1)}(z)},$$

and its function F is therefore

$$F = z - \frac{f}{f^{(1)}},$$

where the argument z has been omitted for the sake of brevity.

Now if the root z_1 is of multiplicity p and we set

$$z - z_1 = \epsilon,$$

then according to § 1

$$f = \epsilon^p \psi, \qquad f^{(1)} = \epsilon^{p-1}(p\psi + \epsilon\psi^{(1)}),$$

and

$$F = z - \frac{\epsilon\psi}{p\psi + \epsilon\psi^{(1)}},$$

whence

$$F^{(1)} = 1 - \frac{\psi}{p\psi + \epsilon\psi^{(1)}} - \epsilon\partial_z \frac{\psi}{p\psi + \epsilon\psi^{(1)}}.$$

From this we immediately see that when $z = z_1$, i.e., when $\epsilon = 0$,

$$F = z_1, \qquad F^{(1)} = 1 - \frac{1}{p},$$

and consequently mod. $F^{(1)} < 1$. Hence Newton's method satisfies the fundamental conditions we have derived for an algorithm.

When $p > 1$, $F^{(1)}$ is different from zero. Thus the algorithm only converges linearly when it is used to find a multiple root; and the higher the multiplicity, the poorer the convergence. However, if $p = 1$, i.e., the root is simple, then the Newton algorithm is quadratically convergent, since $F^{(1)}(z_1) = 0$ while $F^{(2)}(z_1)$ is in general different from zero.

If we wish to obtain a quadratically convergent algorithm for the first case and the multiplicity of the root is known, we need only set

$$F = z - p\frac{f}{f^{(1)}}.$$

This same algorithm could be used to advantage for p nearly equal roots.

In the special case when $f(z) = (z - z_1)^p$ and hence $\psi(z) = 1$, this last algorithm yields the correct root immediately for any starting value z, since $z' = z - \epsilon = z_1$.

II. Let

$$\frac{\varphi^{(1)}(z)}{\varphi(z)}$$

denote an arbitrary function that is single valued about z_1 and is not infinite at z_1. Then the equation

$$z' = z - \frac{f(z)\varphi(z)}{\varphi(z)f^{(1)}(z) - f(z)\varphi^{(1)}(z)} \quad \text{or} \quad F = z - \frac{1}{\dfrac{f^{(1)}}{f} - \dfrac{\varphi^{(1)}}{\varphi}}$$

yields an algorithm for finding the root z_1 which in general converges linearly when z_1 is a multiple root but converges quadratically when z_1 is a simple root.

The same algorithm results if we construct the Newton algorithm for solving the equation

$$\frac{f(z)}{\varphi(z)} = 0;$$

that is, if in the formulas in I the function f is everywhere replaced by $\dfrac{f}{\varphi}$. For if, as is implicit in the above assumption, f and φ do not vanish simultaneously at z_1, the above equation has the root z_1 and can be solved in place of the equation $f(z) = 0$.

We may also show that this is a legitimate algorithm directly, without deriving it from Newton's method, by substituting $\epsilon^p\psi$ for f to get

$$F = z - \frac{\epsilon}{p + \epsilon\dfrac{\psi^{(1)}}{\psi} - \epsilon\dfrac{\varphi^{(1)}}{\varphi}}$$

and differentiating this expression with respect to z.

Because of the arbitrariness of the function φ, this general algorithm includes infinitely many special algorithms. For example, we get back Newton's method itself by taking

$$\varphi(z) = \text{const}.$$

III. The function φ may be chosen so that the algorithm remains quadratically convergent, even for multiple roots—and this is the most noteworthy special case of the general algorithm. According to the remarks about equation (6), this will happen if we take $\varphi = f^{(1)}$. Here, after an easy reduction, we find that

$$F = z - \epsilon\frac{p + \epsilon\dfrac{\psi^{(1)}}{\psi}}{p + \epsilon^2\left(\dfrac{\psi^{(1)}}{\psi}\right)^2 - \epsilon^2\dfrac{\psi^{(2)}}{\psi}},$$

and for $\epsilon = 0$ it follows immediately that $F^{(1)} = 0$. Hence

$$z' = z - \frac{f(z)f^{(1)}(z)}{f^{(1)}(z)^2 - f(z)f^{(2)}(z)}$$

is an algorithm that always converges quadratically.

§ 4.
The Most General Algorithms with a Given Rate of Convergence.

I will now proceed to show in complete generality how to easily construct algorithms $z' = F(z)$ that converge to a root of the equation $f(z) = 0$ at an arbitrary given rate.

We will let the function f be subject to the condition of § 1 that it be single valued and the function F be subject to the conditions of § 2.

For the sake of brevity we will omit the starting value z (which must be chosen sufficiently near z_1 but is otherwise arbitrary) whenever it appears as an argument of a function. However, to distinguish the case where the special argument z_1 is to be understood in a formula instead of the general argument z, we will write $z = z_1$ nearby—or what is equivalent for our purposes, $f(z) = 0$. Derivatives like

$$\partial_z^a f(z) = f^{(a)}(z)$$

of the function f, which occur especially often, will be denoted by f_a; however, any other differentiation with respect to z will be indicated by the symbols ∂_z or ∂_z^a, or for short ∂ or ∂^a. The scope of these symbols extends to the next $+$ or $-$ sign.

By φ, φ_1, φ_2, ... I will denote arbitrary functions that are single valued about the point z_1 and do not become infinite there.

Finally, without loss of generality we can restrict ourselves to the assumption that the root z_1 of the equation to be solved, i.e. $f = 0$, is simple. For as we have already mentioned, if this equation has multiple roots, the equation $\dfrac{f}{f_1} = 0$ has the same roots, only each is simple. In order to obtain results for the second case that correspond to the first case, one has only to substitute $\dfrac{f}{f_1}$ for f in the first.

Now the first-stated requirement on the function F was that $F = z_1$ for $z = z_1$ or $f = 0$. This requirement will be satisfied with greatest generality if we set

$$F = z - \varphi$$

where the arbitrary function φ vanishes when $f = 0$, in addition to satisfying previous conditions. The present condition will be satisfied, again with greatest generality, if we set

$$\varphi = f \cdot \varphi_1.$$

Hence, as long as φ_1 remains free, the function

$$F = z - f \cdot \varphi_1$$

is of the most general form that satisfies the first requirement.

Now if this algorithm is to have at least quadratic convergence — we are unable to write down a convenient general form for a linearly converging algorithm, which only has to satisfy a single inequality, namely mod. $F < 1$ when $f = 0$; but such algorithms are much less useful in practice — then we have a second requirement, namely $\partial F = 0$ when $f = 0$. This gives

$$f_1\varphi_1 + f\partial\varphi_1 = 1 \quad \text{when} \quad f = 0.$$

Since $\partial\varphi_1$ can no more become infinite at the point z_1 than φ_1 can, we must have

$$f_1\varphi_1 = 1 \quad \text{when} \quad f = 0.$$

And finally, since the assumption of a simple root implies that f_1 cannot vanish along with f, we must have

$$\varphi_1 = \frac{1}{f_1} \quad \text{when} \quad f = 0.$$

If we allow this equation, which has to hold only when $z = z_1$, to hold for arbitrary z and append an arbitrary term that vanishes with f, that is if we set

$$\varphi_1 = \frac{1}{f_1} + f\varphi_2,$$

then the second requirement will be satisfied with greatest generality. Thus the equation

$$F = z - \frac{f}{f_1} - f^2\varphi_2,$$

in which φ_2 remains free, includes all algorithms of the second order or of quadratic convergence.

[In fact, it is easy to determine φ_2 so that the algorithm turns into the equally general algorithm given as an example in II of the previous section. To do this, we need only equate the two expressions for F, and by means of this equation express φ_2 in terms of the function φ of that section.]

If further the algorithm is to converge cubically, the function F must satisfy the additional requirement that $\partial^2 F = 0$ when $f = 0$. If we expand the equation $\partial^2 F = 0$ and set $f = 0$, but do not restrict the argument z to be z_1, we can determine the most general form of the function φ_2 that satisfies the requirement by appending an arbitrary term that vanishes with f. After an easy calculation we obtain

$$\varphi_2 = \frac{f_2}{2f_1^3} + f\varphi_3$$

Consequently the function

$$F = z - \frac{f}{f_1} - \frac{f^2 f_2}{2f_1^3} - f^3\varphi_3.$$

gives the general algorithm with cubic convergence.

If we continue reasoning in this way, we arrive at the following result.

The most general algorithm $z' = F(z)$ *whose convergence is of the* ω*th order is obtained by taking*

$$\begin{aligned} F = z - \frac{f}{1!}\cdot\frac{1}{f_1} + \frac{f^2}{2!}\cdot\frac{1}{f_1}\partial\frac{1}{f_1} - \frac{f^3}{3!}\cdot\left(\frac{1}{f_1}\partial\right)^2\frac{1}{f_1} + \cdots \\ \cdots + (-1)^{\omega-1}\frac{f^{\omega-1}}{(\omega-1)!}\cdot\left(\frac{1}{f_1}\partial\right)^{\omega-2}\frac{1}{f_1} - f^\omega\varphi_\omega, \end{aligned}$$

where φ_ω *is an arbitrary function. In more compact notation*

$$F = z + \sum_{a=1}^{a=\omega-1}(-1)^a\frac{f^a}{a!}\cdot\left(\frac{1}{f_1}\partial\right)^{a-1}\frac{1}{f_1} - f^\omega\cdot\varphi_\omega. \tag{18}$$

By way of explanation I must note that for brevity I have used here (and in the sequel) the symbol $a!$, introduced by SCHLÖMILCH, to denote the factorial $1\cdot 2\cdot 3\cdots(a-1)\cdot a$. Moreover, the expression

$$\left(\frac{1}{f_1}\partial\right)^{a-1}$$

does not represent a quantity but an operator. It directs that the object $\frac{1}{f_1}$ of the operator be successively differentiated and then multiplied by $\frac{1}{f_1}$, the process being repeated $a-1$ times. For example,

$$\left(\frac{1}{f_1}\partial\right)^4\frac{1}{f_1}$$

has the meaning

$$\frac{1}{f_1}\partial\frac{1}{f_1}\partial\frac{1}{f_1}\partial\frac{1}{f_1}\partial\frac{1}{f_1}.$$

In order to establish the truth of the above theorem, one has only to show that all the derivatives of the function F up to and including the $(\omega-1)$th vanish when $f=0$. For this purpose, however, it is unnecessary to calculate all these derivatives; it is sufficient to differentiate the equation (18) a single time. If we do so for all terms, first differentiating the factor before the $\cdot$ sign and then the factor after, but merely indicating the second differentiation, we get

$$\begin{aligned}\partial F = 1 + \sum_{a=1}^{a=\omega-1}(-1)^a\frac{af^{a-1}f_1}{a!}\cdot\left(\frac{1}{f_1}\partial\right)^{a-1}\frac{1}{f_1} - \omega f^{\omega-1}f_1\cdot\varphi_\omega \\ + \sum_{a=1}^{a=\omega-1}(-1)^a\frac{f^a}{a!}\cdot\partial\left(\frac{1}{f_1}\partial\right)^{a-1}\frac{1}{f_1} - f^\omega\cdot\partial\varphi_\omega.\end{aligned}$$

Now in the first sum the general term simplifies to

$$(-1)^a\frac{f^{a-1}}{(a-1)!}\cdot\partial\left(\frac{1}{f_1}\partial\right)^{a-2}\frac{1}{f_1},$$

and one sees right away that the first term of this sum cancels the the first term. Similarly, the following terms of this sum cancel the first $\omega-2$ terms of the second sum, so that only the last term of the second sum remains along with the two terms outside the scope of the summations.

Therefore, we have

$$\partial F = f^{\omega-1}\left\{\frac{(-1)^{\omega-1}}{(\omega-1)!}\partial\left(\frac{1}{f_1}\partial\right)^{\omega-2}\frac{1}{f_1} - \omega f_1\varphi_\omega - f\partial\varphi_\omega\right\}.$$

This function contains the factor $f^{\omega-1}$ and has an $(\omega-1)$-fold zero where f vanishes. Hence, its higher derivatives, up to and including the $(\omega-2)$th vanish when $f=0$, which is what was to be shown.

§ 5.
Solution Methods of the Second Kind as Limiting Cases of Algorithms.

We have now arrived at a general expression (18) for algorithms of order ω, and it is natural to think of taking $\omega=\infty$.

In this case the function F takes the form of the infinite series

$$F = z + \sum_{a=1}^{a=\infty} (-1)^a \frac{f^a}{a!} \cdot \left(\frac{1}{f_1}\partial\right)^{a-1} \frac{1}{f_1}, \tag{19}$$

provided we eliminate the missing term $f^\omega \phi_\omega$ by assuming that $\phi_\omega = 0$. Obviously this function has meaning only to the extent that the series converges. In a large number of cases the series will actually converge in a certain neighborhood of z, since for $f = 0$ it reduces to the starting value $z = z_1$ and since the quantity f, in whose powers the series is expanded, can be made arbitrarily small by taking z sufficiently near z_1. [However, if this neighborhood should reduce to the point z_1, it not infrequently happens that by a suitable choice of ϕ_ω one can replace the divergent series with a limit that remains finite.]

In the case of convergence, the series (19) represents a function whose derivatives all vanish when $f = 0$. Therefore, the value of this function for a starting value z chosen sufficiently near z_1 gives an approximation z' to the root z_1 whose error is proportional to an infinitely high power of the error in the starting value; i.e., the error is zero. Indeed, for this case equation (14) gives F or

$$z' = z_1. \tag{20}$$

The algorithm with infinitely swift convergence therefore gives the true root of the equation immediately as its first approximation. It no longer has the nature of a real algorithm; i.e., a computational method that is to be *repeatedly* applied. Instead it constitutes an algorithm of the second kind mentioned in § 1, in which we do not need to use approximations to solve the equation.

Let us now actually perform the differentiations in the first terms of the series as indicated. The series then reads

$$\begin{aligned} z_1 = F = z - \frac{f}{1!}\cdot\frac{1}{f_1} - \frac{f^2}{2!}\cdot\frac{f_2}{f_1^3} - \frac{f^3}{3!}\cdot\frac{3f_2^2 - f_1 f_3}{f_1^5} \\ -\frac{f^4}{4!}\cdot\frac{15f_2^3 - 10f_1f_2f_3 + f1^2 f_2 f_4}{f_1^7} \\ -\frac{f^5}{5!}\cdot\frac{105f_2^4 - 105f_1f_2^2f_3 + 10f_1^2f_3^2 + 15f_1^2f_4 - f_1^3f_5}{f_1^9} - \cdots. \end{aligned} \tag{21}$$

This form of the series has been derived by THEREMIN*) in a different way and without reference to the algorithms considered here.

If we denote the general term by

$$-\frac{f^a}{a!}\cdot\frac{\chi_a}{f^{2a-1}},$$

the numbers χ_a can be easily computed by the recurrence

$$\chi_{a+1} = (2a-1)f_2\chi_a - f_1\partial\chi_a. \tag{22}$$

The series can also be represented more concisely as follows

$$\begin{aligned} z_1 &= F(z) \\ &= z + \sum_{a=1}^{a=\infty} \frac{(-1)^a f(z)^a}{a!} \lim_{\epsilon=0} \partial_\epsilon^{a-1} \left\{ \frac{f(z+\epsilon)-f(z)}{\epsilon} \right\}^{-a}. \end{aligned} \tag{23}$$

Alternatively,

$$z_1 = \sum_{a=0}^{a=\infty} \frac{(-1)^a f(z)^a}{a!} \partial_{f(z)}^a z, \tag{24}$$

since we have the identity

$$\begin{aligned} \left(\frac{1}{f_1}\partial\right)^{a-1}\frac{1}{f_1} &= \frac{(-1)^{a-1}\chi_a}{f_1^{2a-1}} \\ &= \lim_{\epsilon=0} \partial_\epsilon^{a-1} \left\{ \frac{f(z+\epsilon)-f(z)}{\epsilon} \right\}^{-a} = \partial_{f(z)}^a z, \end{aligned} \tag{25}$$

which can easily be proven from the well-known theorem on the exchange of independent variables.

§ 6.
An Example: the Quadratic Equation.

In order to give an example of the results of the last section, I will apply them to the quadratic equation. This example has already been treated by THEREMIN (loc. cit.), but not with satisfactory rigor and completeness.

*) Crelle's Journal, V. 49, pp. 187–243: Reserches sur la résolution des équations de tous les dégrés.

Let

$$f(z) = (z - z_1)(z - z_2) = z^2 - z(z_1 + z_2) + z_1 z_2 = 0$$

be the equation to be solved in the manner given above. Then

$$\frac{f(z+\epsilon) - f(z)}{\epsilon} = \epsilon + 2z - z_1 - z_2,$$

and

$$\partial_\epsilon^{a-1} \cdot \frac{1}{(\epsilon + 2z - z_1 - z_2)^a} = \frac{(-1)^{a-1}(2a-2)!}{(a-1)!(\epsilon + 2z - z_1 - z_2)^{2a-1}}.$$

If we substitute the value of this expression with $\epsilon = 0$ for the corresponding value in equation (23), it follows that

$$z' = z - \sum_{a=1}^{a=\infty} \frac{(2a-2)!}{(a-1)!a!} \cdot \frac{(z-z_1)^a (z-z_2)^a}{(2z - z_1 - z_2)^{2a-1}}.$$

If we denote the binomial coefficient

$$\frac{s!}{a!(s-a)!}$$

by $(s)_a$, then since

$$\frac{(2a-2)!}{(a-1)!a!} = (-1)^{a-1} 2^{2a-1} \left(\frac{1}{2}\right)_a$$

we have

$$z' = z + \left(z - \frac{z_1 + z_2}{2}\right) \sum_{a=1}^{a=\infty} (-1)^a \left(\frac{1}{2}\right)_a \left\{ \frac{(z-z_1)(z-z_2)}{\left(z - \frac{z_1+z_2}{2}\right)^2} \right\}^a .$$

If we subtract one from the sum on the right, we may take its lower index to be zero. But then the sum represents the binomial series with exponent $\frac{1}{2}$ developed in powers of the quantity

$$t = -\frac{(z-z_1)(z-z_2)}{\left(z - \frac{z_1+z_2}{2}\right)^2}$$

Now ABEL*) has shown that the binomial series

$$\sum_{a=0}^{a=\infty} (s)_a t^a$$

converges and represents one of the values of $(1+t)^s$, whenever mod. $t < 1$. Provided the real part of s is greater than -1, the series even converges when mod. $t = 1$, with the exception of the special value $t = -1$ with real. $s \leq 0$. Since in our case s equals $\frac{1}{2}$, our series converges for mod. $t \leq 1$ and is equal to

$$\begin{aligned} z' &= z + \left(z - \frac{z_1 + z_2}{2}\right)\left\{\pm\sqrt{1+t} - 1\right\} \\ &= z + \left(z - \frac{z_1 + z_2}{2}\right)\left\{\frac{\pm\dfrac{z_1 + z_2}{2}}{z - \dfrac{z_1 + z_2}{2}} - 1\right\}. \end{aligned}$$

Hence

$$z' = \frac{z_1 + z_2}{2} \pm \frac{z_1 - z_2}{2};$$

that is, $z' = z_1$ for the plus sign and $z' = z_2$ for the minus sign. However, we have yet to discuss which of these two cases holds. Equivalently, we need to consider the disposition of the convergence region.

If we set

$$z - z_1 = \rho_1 e^{i\vartheta_1}, \quad z - z_2 = \rho_2 e^{i\vartheta_2}, \quad z - \frac{z_1 + z_2}{2} = \rho e^{i\vartheta},$$

then ρ_1, ρ_2, and ρ are the distances of z from the points z_1, z_2, and $\frac{z_1 + z_2}{2}$, the last lying midway between the first two points. The convergence condition for the series now reads

$$\rho_1 \rho_2 \leq \rho^2.$$

If $2E$ denotes the distance between the two roots z_1 and z_2, then by a well known theorem on the median of a triangle we have

$$\rho^2 = \frac{\rho_1^2 + \rho_2^2}{2} - E^2.$$

*) *Oeuvres Complètes,* T. I. No. VII, Christiania 1839, p. 66.

Our convergence condition thus becomes

$$(\rho_1 - \rho_2)^2 \geq 2E^2 \qquad \text{or} \qquad \pm(\rho_1 - \rho_2) \geq E\sqrt{2}.$$

The convergence region is therefore bounded by an equilateral hyperbola, whose foci are the roots z_1 and z_2 of the quadratic equation. Moreover the region is that part of the plane that contains these foci, and the hyperbola itself belongs to the convergence region.

Abel (loc. cit.) has also shown which value of the infinitely multiple valued function $(1+t)^s$ is given by the binomial series when it converges. However, since the sum is unambiguously known for real arguments, we can obtain this part of Abel's result more easily from the theorem of function theory that says that a power series is a single-valued and continuous function of its argument and that such a function can be continued in the plane—or at least in a sector of the plane—in only one way. As turns out to be suitable in many investigations, the function $\log z$ can be defined to be single valued in the entire complex plane by the stipulation that it be taken real for positive z and that it be continued from the axis of positive numbers to the axis of negative numbers in such a way that the imaginary part of $\log z$ is $+\pi i$ on the negative axis, while infinitely close below the negative axis it is $-\pi i$. Thus the function has a discontinuity along the axis of negative numbers. The sum of the binomial series can then be represented unambiguously by the expression $e^{s\log(1+t)}$. Now for our example

$$1 + t = \left(\frac{Ee^{i\vartheta_0}}{\rho e^{i\vartheta}}\right)^2 = \left(\frac{E}{\rho}\right)^2 e^{2i(\vartheta_0 - \vartheta)},$$

where ϑ_0 denotes the argument of the number

$$\frac{z_1 - z_2}{2} = Ee^{i\vartheta_0}.$$

Hence,

$$e^{\frac{1}{2}\log(1+t)} = \frac{E}{\rho} e^{\frac{1}{2}\log e^{2i(\vartheta_0 - \vartheta)}}.$$

Note that

$$\log e^{iy} = i(y + 2h\pi),$$

where the positive or negative integer h must be chosen so that $y + 2h\pi$ lies between $-\pi$ (exclusive) and π (inclusive). It follows that

$$e^{\frac{1}{2}\log(1+t)} = \frac{E}{\rho} e^{i(\vartheta_0 - \vartheta + h\pi)} = \frac{\dfrac{z_1 + z_2}{2}}{z - \dfrac{z_1 + z_2}{2}} \cdot e^{h\pi i},$$

where h is to be chosen so that $\vartheta_0 - \vartheta + h\pi$ lies between $-\dfrac{\pi}{2}$ (exclusive) and $\dfrac{\pi}{2}$ (inclusive). Now the ray ρ from the point z forms two supplementary angles with the line connecting z_1 and z_2. Let the one lying on the side of the point z_1 be written ω_1 and the other ω_2, and let them be taken between 0 and π. If we take the arguments ϑ_0, ϑ, etc. to be between $-\pi$ and π, then it is easy to express the difference $\vartheta_0 - \vartheta$ in terms of ω_1 and ω_2 using the theorem on the exterior angle of a triangle. We then find that if $\omega_1 < \dfrac{\pi}{2}$ then h is the even number 0 or 2 and consequently $e^{h\pi i} = 1$. On the other hand if $\omega_2 < \dfrac{\pi}{2}$ then h is the odd number ± 1 and consequently $e^{h\pi i} = -1$. In the first case the formula for z has a plus sign, and $z' = z_1$. In the second case it has a minus sign and $z' = z_2$. The series thus yields a sum that is the root of the polynomial that is the focus of the hyperbola lying on the same side of the minor axis as the point z; i.e., the point that z can approach without crossing the curve.

§ 7.
Introduction of the Symmetric Functions A and B of the Roots.

Since the derivation of the most general algorithms for solving the equation $f(z) = 0$ has been completed, I will go on to present the most noteworthy special algorithms. My theme will be *algebraic* equations, and in the sequel I will confine myself to the case where $f(z)$ is a rational function, even though the final results for the most part extend to arbitrary single-valued functions.

Let the roots of the equation, which are different from zero and ∞, be $z_1, z_2, z_3, \ldots, z_n$. Our *sole* concern will be with finding these roots.

The algorithms considered here are based on the properties of certain symmetric functions of all the roots, functions which have the form

$$C_\omega^{(\lambda)}(z) = \sum_{a=1}^{a=n} \frac{z_a^\lambda \chi(z_a)}{(z-z_a)^{\omega+1}}. \tag{26}$$

Here ω and λ denote arbitrary positive integers, and χ is a function of a single variable that does not contain the the arbitrary starting value z and does not become zero or ∞ for any of the roots z_a of the equation.

The next thing we will consider are the properties of this function (26).

If we set

$$F(z) = \frac{C_\omega^{(\lambda+h)}(z)}{C_\omega^{(\lambda)}(z)}, \tag{27}$$

where h again denotes a natural number, then we can state the following theorem.

I. *If the argument of the function F is a root of the equation $f(z) = 0$, then the value of the function is the hth power of this root; i.e.,*

$$F(z_1) = z_1^h. \tag{28}$$

Here z_1, which is the symbol introduced for the first root, obviously represents an arbitrary root.

To prove this theorem, multiply the numerator and denominator of the fraction (27), so that it does not become ∞ for $z = z_1$. From this we get

$$F(z) = \frac{\sum_{a=1}^{a=n} z_a^{\lambda+h}\chi(z_a)\left(\frac{z-z_1}{z-z_a}\right)^{\omega+1}}{\sum_{a=1}^{a=n} z_a^{\lambda}\chi(z_a)\left(\frac{z-z_1}{z-z_a}\right)^{\omega+1}}. \tag{29}$$

If the root z_1 is p-fold (where p can be one), e.g., if

$$z_1 = z_2 = \cdots = z_p, \tag{30}$$

then for $z = z_1$ all terms in the numerator and denominator which follow the pth vanish (i.e., all terms for which the summation variable

a is greater than p). In the other terms, for which a has the values 1, 2, 3, …, p, the factor

$$\left(\frac{z - z_1}{z - z_a}\right)^{\omega+1}$$

assumes the value one. Since all the terms are equal, their sum comes to p times the first. Hence

$$F(z_1) = \frac{p z_1^{\lambda+h} \chi(z_1)}{p z_1^{\lambda} \chi(z_1)},$$

which amounts to (28).

Under the assumption that $h = 1$, the function F satisfies the first fundamental condition $F(z_1) = z_1$, which, as we have seen, any function that gives an algorithm must obey.

II. *If the point z lies nearer the root z_1 than any of the other roots, i.e., if* mod. $(z - z_1)$ *is the smallest of the distinct moduli of the differences $z - z_1$, $z - z_2$, …, $z - z_n$, so that*

$$\text{(31)} \quad \text{mod.}\ (z - z_1) < \text{mod.}\ (z - z_a), \qquad a = p + 1, p + 2, \ldots, n,$$

then

$$\text{(32)} \qquad \lim_{\omega=\infty} F(z) = z_1^h.$$

For

$$\text{mod.}\ \frac{z - z_1}{z - z_a}$$

is a proper fraction for all $a > p$, and hence

$$\lim_{\omega=\infty} \left(\frac{z - z_1}{z - z_a}\right)^{\omega+1} = 0, \qquad a > p.$$

On the other hand, the same limit is equal to one for $a = 1, 2, 3, \ldots, p$. Hence, in the numerator and denominator of (29) all terms following the pth vanish for $\omega = \infty$, and the rest combine exactly as in the previous theorem, where z was set equal to z_1.

III. *When z becomes equal to a root z_1, the derivative $\partial_z F(z)$ becomes* 0 *with order ω; i.e., we can write*

$$\text{(33)} \qquad \partial_z F(z) = (z - z_1)^{\omega} \cdot \Psi(z),$$

where $\Psi(z)$ is a function that is not infinite for $z = z_1$ and, as long no special relations between the roots are assumed, does not vanish.

For the proof differentiate (27) to get

$$\partial_z F(z) = \frac{C_\omega^{(\lambda)}(z)\partial_z C_\omega^{(\lambda+h)}(z) - C_\omega^{(\lambda+h)}(z)\partial_z C_\omega^{(\lambda)}(z)}{\{C_\omega^{(\lambda)}(z)\}^2}$$

and then substitute the expressions derived from (26) for the functions C. However, to keep things straight replace the summation index a by another letter b in the functions to be differentiated before performing the differentiations. After an easy simplification, it follows that

$$C_\omega^{(\lambda)}(z)^2 \cdot \partial_z F(z) = -(\omega+1)\sum_{a=1}^{a=n}\sum_{b=1}^{b=n} \frac{z_a^\lambda z_b^\lambda \chi(z_a)\chi(z_b)}{(z-z_a)^{\omega+1}(z-z_b)^{\omega+1}} \cdot \frac{z_b^h - z_a^h}{z - z_b}.$$

The expression in the double sum on the right-hand side is unsymmetric with respect to the summation indices a and b because of the factors

$$\frac{z_b^h - z_a^h}{z - z_b}.$$

However, these factors can be cast in symmetric form, if we take into consideration the fact that all the terms in which the summation indices a and b are equal fall out of the double sum owing to the numerators $z_b^h - z_a^h$. Among the remaining factors, for each combination a, b we find a corresponding combination b, a. This allows us to write

$$\frac{1}{2}\left\{\frac{z_b^h - z_a^h}{z - z_b} + \frac{z_a^h - z_b^h}{z - z_a}\right\} = \frac{1}{2}\frac{(z_a^h - z_b^h)(z_a - z_b)}{(z - z_a)(z - z_b)}.$$

for the factors in the double sum. Hence we finally get

$$(34)\ \partial_z F(z) = -\frac{\omega+1}{2} \cdot \frac{\sum_{a=1}^{a=n}\sum_{b=1}^{b=n} \frac{z_a^\lambda z_b^\lambda (z_a^h - z_b^h)(z_a - z_b)}{(z-z_a)^{\omega+2}(z-z_b)^{\omega+2}}\chi(z_a)\chi(z_b)}{\left\{\sum_{c=1}^{c=n} \frac{z_c^\lambda \chi(z_c)}{(z-z_c)^{\omega+1}}\right\}^2},$$

in which we may regard a and b as any two distinct integers from the integers $1, 2, 3, \ldots, n$.

We now will multiply the numerator and denominator of (34) by a power of $z-z_1$ in such a way that they do not become infinite when $z=z_1$. The question is what power of this difference is required.

First of all, if the root z_1 is simple, we obviously must augment the fraction by the factor

$$\{(z-z_1)^{\omega+1}\}^2=(z-z_1)^{2\omega+2}$$

because of the common denominator. However, to obtain the required result in the numerator it would suffice to multiply by $(z-z_1)^{\omega+2}$, since by the distinctness of a and b no higher power of $z-z_1$ can appear in the denominators of the individual terms. Therefore, there remains the factor

$$(z-z_1)^{2\omega+2-(\omega+2)}=(z-z_1)^{\omega}$$

which multiplies a function that no longer becomes ∞ or 0 when $z=z_1$.

In the case of a multiple, say p-fold, root one easily finds that the same factor works by considering that certain terms in the numerator of the formula (34) vanish, while others in the numerator and denominator combine.

Since the derivative $\partial_z F(z)$ vanishes with order ω when $z=z_1$, the higher differential quotients of F, up to an including the ωth, must also vanish. In other words, the function $F(z)$ satisfies the second fundamental condition that a function yielding an algorithm of order $\omega+1$ must obey.

IV. It remains to state the relation

$$C_\omega^{(\lambda)}(z)=\sum_{c=0}^{c=h}(-1)^c(h)_c z^{h-c}C_{\omega-c}^{(\lambda-h)}(z) \tag{35}$$

as a fourth property of the function C. Here $(h)_c$ denotes the binomial coefficient $\dfrac{h!}{c!(h-c)!}$. The proof is easily effected. For if one substitutes the expressions from the definition (26) in the relation, it reduces to the equation

$$z_a^\lambda=\sum_{c=0}^{c=h}(-1)^c(h)_c z^{h-c}(z-z_a)^c,$$

which by the binomial theorem is an identity.

We should mention the most notable special case of the relation (35), which results from setting $h = \lambda$:

$$C_\omega^{(\lambda)}(z) = \sum_{c=0}^{c=\lambda} (-1)^c (\lambda)_c z^{\lambda-c} C_{\omega-c}^{(0)}(z). \tag{36}$$

By means of this equation the functions C with exponent λ can be reduced to the same functions with just the exponent zero.

Finally, if we replace λ by $\lambda + h$, the relation takes a form which we write down for the case $h = 1$, a case that will be especially important in the sequel:

$$C_\omega^{(\lambda+1)}(z) = zC_\omega^{(\lambda)}(z) - C_{\omega-1}^{(\lambda)}(z). \tag{37}$$

Regarding the arbitrary functions χ that enter in the expression (26) for the symmetric function $C_\omega^{(\lambda)}(z)$, two special choices will turn out to be particularly valuable in the sequel. The first choice is

$$\chi(z) = \frac{1}{f^{(1)}(z)},$$

and the second is

$$\chi(z) = 1.$$

If we choose these two expressions for χ, we obtain two functions from (26), which we shall call A and B to distinguish them from the general function C. Specifically,

$$A_\omega^{(\lambda)} = \sum_{a=1}^{a=n} \frac{z_a^\lambda}{(z - z_a)^{\omega+1} f^{(1)}(z_a)}, \tag{38}$$

and

$$B_\omega^{(\lambda)} = \sum_{a=1}^{a=n} \frac{z_a^\lambda}{(z - z_a)^{\omega+1}}. \tag{39}$$

For the function A it is clear that that the case of multiple roots must be excluded. Otherwise, some terms of the sum will become infinite owing to the vanishing of the derivative of f. For the function B this case is permitted. In general, the properties of these two functions exhibit such deep analogies that it is highly advisable to investigate them together.

In fact, the function A becomes the function B when $f(z)$ is replaced by $\dfrac{f(z)}{f^{(1)}(z)}$. Specifically, according to § 1 the equation $\dfrac{f(z)}{f^{(1)}(z)} = 0$ has the same finite roots as the equation $f(z) = 0$, but each of these roots is simple. Therefore, we are allowed to form the function $A_\omega^{(\lambda)}$ for the first equation, which may be done by replacing $\dfrac{1}{\partial_z f(z)}$ with

$$\frac{1}{\partial_z \dfrac{f(z)}{f^{(1)}(z)}} = \frac{f^{(1)}(z)^2}{f^{(1)}(z)^2 - f(z)f^{(2)}(z)}.$$

It is easy to show that this expression takes the value p for $z = z_1$. Hence the sum (38) extending over all the finite roots of the equation $\dfrac{f(z)}{f^{(1)}(z)} = 0$—that is all the distinct roots of the equation $f(z) = 0$— goes over into the sum (39) ranging over all the roots of the equation $f(z) = 0$.

If we think of $f(z)$ as not just a rational function, but an entire function, say

$$\begin{aligned} f(z) &= \gamma_0 z^n + \gamma_1 z^{n-1} + \gamma_2 z^{n-2} + \cdots \gamma_{n-1} z + \gamma_n \\ &= \gamma_0 (z - z_1)(z - z_2) \cdots (z - z_n), \end{aligned} \tag{40}$$

then, as is well known, we can express any symmetric function of all the roots (here presupposed to be finite) in terms of the coefficients of this equation or the derivatives of the polynomial $f(z)$ for $z = 0$. Consequently, we can express our functions A and B in this way. Our immediate problem is to construct these expressions. We have two ways at hand to derive the expressions systematically. On the one hand, starting from the definitions (38) and (39) we can obtain obtain recursions (like the equations (47) and (48) of the following section), from which we see that our expressions are the coefficients of a *recurrent* series, which we then sum. On the other hand, we can decompose the symmetric functions into homogeneous (entire rational) parts and determine the latter according to the method of WARING, GAUSS, and CAUCHY. Better yet, we can seek the generating functions themselves, as explained by BORCHARDT and BETTI.*)

*) Crelle's Journal, V. 53, p. 193 and V. 54, p. 98 ff.

Because these derivations are intricate, I will content myself with simply stating and verifying the results.

§ 8.
Derivation of Related Functions $\mathcal{A}$ and $\mathcal{B}$ from a Generating Function.

We shall now define two brand new functions of z by the following equations:

$$\mathcal{A}_\omega^{(\lambda)}(z) = \left[\frac{(z-\epsilon)^\lambda}{f(z-\epsilon)}\right]_{\epsilon^\omega}, \tag{41}$$

$$\mathcal{B}_\omega^{(\lambda)}(z) = \left[\frac{(z-\epsilon)^\lambda f^{(1)}(z-\epsilon)}{f(z-\epsilon)}\right]_{\epsilon^\omega}. \tag{42}$$

Here (following JACOBI) the symbol $[\Phi(\epsilon)]_{\epsilon^\omega}$ represents the coefficient of ϵ^ω in the development of the bracketed function $\Phi(z)$ in increasing powers of ϵ, a development which is valid for sufficiently small ϵ. Thus the function in brackets would be called by Laplace the generating function *(fonction génératrice)* of the coefficients it defines. To put it otherwise, the above equations are equivalent to the following:

$$\mathcal{A}_\omega^{(\lambda)}(z) = \frac{1}{\omega!}\lim_{\epsilon=0}\partial_\epsilon^\omega \frac{(z-\epsilon)^\lambda}{f(z-\epsilon)}, \tag{43}$$

$$\mathcal{B}_\omega^{(\lambda)}(z) = \frac{1}{\omega!}\lim_{\epsilon=0}\partial_\epsilon^\omega \frac{(z-\epsilon)^\lambda f^{(1)}(z-\epsilon)}{f(z-\epsilon)}. \tag{44}$$

If these definitions are to make sense, the Taylor series

$$\frac{(z-\epsilon)^\lambda}{f(z-\epsilon)} = \sum_{a=0}^{a=\infty} \mathcal{A}_a^{(\lambda)}(z)\cdot\epsilon^a, \tag{45}$$

$$\frac{(z-\epsilon)^\lambda f^{(1)}(z-\epsilon)}{f(z-\epsilon)} = \sum_{a=0}^{a=\infty} \mathcal{B}_a^{(\lambda)}(z)\cdot\epsilon^a \tag{46}$$

must converge for sufficiently small ϵ. In fact it is easy to see that this expansion is always valid for sufficiently small ϵ provided that z is not exactly equal to a root z_1 of the equation $f(z) = 0$.

In this section we will be concerned with recurrent and closed representations of the functions $\mathcal{A}$ and $\mathcal{B}$ in terms of the function $f(z)$ and its derivatives. For the sake of brevity we will once again omit all arguments z and denote the derivatives of $f(z)$ as in § 4.

If we multiply equations (45) and (46) by the expansion of the function $f(z-\epsilon)$ in increasing powers of ϵ (an expansion which always exists) and order both sides by powers of ϵ, then by equating the coefficients of $(-\epsilon)^\omega$ in the right and left-hand sides we get

$$(\lambda)_\omega z^{\lambda-\omega} = \sum_{a=0}^{a=\omega} \frac{(-1)^a f_{\omega-a}}{(\omega-a)!} \mathcal{A}_a^{(\lambda)}, \tag{47}$$

$$\sum_{a=0}^{a=\omega} \frac{(\lambda)_a z^{\lambda-a} f_{\omega-a+1}}{(\omega-a)!} = \sum_{a=0}^{a=\omega} \frac{(-1)^a f_{\omega-a}}{(\omega-a)!} \mathcal{B}_a^{(\lambda)}. \tag{48}$$

By means of these equations the functions $\mathcal{A}$ and $\mathcal{B}$ can be calculated recursively.

Before doing this, it is appropriate to give closed representations of the functions $\mathcal{A}$ and $\mathcal{B}$. We can easily obtain such expressions by writing down the system of equations that result from setting $\omega = 0, 1, 2, 3, \ldots$ in (47) or (48) and then solving the system for the unknowns $(-1)^\omega \mathcal{A}_\omega^{(\lambda)}$ or $(-1)^\omega \mathcal{B}_\omega^{(\lambda)}$.

Since all the elements above the diagonal of the determinant of the system are zero, the determinant is the product of its diagonal elements and takes the value $f^{\omega+1}$. If we write this denominator as a factor on the other side, we get the following formulas:

$$f^{\omega+1}\mathcal{A}_\omega^{(\lambda)} = \left\| (\lambda)_c z^{\lambda-c}, \frac{f_c}{c!}, \frac{f_{c-1}}{(c-1)!}, \ldots \frac{f_1}{1!}, f, 0, 0, \ldots, 0 \right\|, \tag{49}$$

$$f^{\omega+1}\mathcal{B}_\omega^{(\lambda)} = \left\| \sum_{a=0}^{a=c} \frac{(\lambda)_a z^{\lambda-a} f_{c-a+1}}{(c-a)!}, \frac{f_c}{c!}, \frac{f_{c-1}}{(c-1)!}, \ldots \frac{f_1}{1!}, f, 0, 0, \ldots, 0 \right\|. \tag{50}$$

Here I have only written down the $(c+1)$th rows of each determinant, which is of order $(\omega+1)$. The individual rows are obtained by setting $c = 0, 1, 2, \ldots, \omega$, where obviously we take only the first $\omega + 1$ elements from the sequence.

In the especially important case $\lambda = 0$, the first element of the row given for the second equation simplifies to $\frac{f_{c+1}}{c!}$. Moreover, the order of the first equation can be reduced by one, since the initial element is one and the remaining elements in the first column become zero. If we multiply the rows of this simplified determinant in order by 1, f, $f^1, f^2, \ldots, f^\omega$ and divide the columns by the same quantities, we can make the elements above the diagonal (which remains unchanged) equal to one, while any other element $\frac{f_c}{c!}$ becomes $\frac{f^{c-1}f_c}{c!}$.

We shall now show that the functions $\mathcal{A}$ and $\mathcal{B}$ with the exponent λ can be easily expressed in terms of of the corresponding functions for $\lambda = 0$. The most suitable representations of our functions for this purpose are the ones in terms of differential quotients that have already been given in (43) and (44). Specifically, it follows from Leibnitz's theorem on the repeated differentiation of products that

$$\mathcal{A}_\omega^{(\lambda)} = \frac{1}{\omega!}\lim_{\epsilon=0}\sum_{a=0}^{a=\omega}(\omega)_a\partial_\epsilon^a(z-\epsilon)^\lambda \times \partial_\epsilon^{\omega-a}\frac{1}{f(z-\epsilon)},$$

$$\mathcal{B}_\omega^{(\lambda)} = \frac{1}{\omega!}\lim_{\epsilon=0}\sum_{a=0}^{a=\omega}(\omega)_a\partial_\epsilon^a(z-\epsilon)^\lambda \times \partial_\epsilon^{\omega-a}\frac{f^{(1)}(z-\epsilon)}{f(z-\epsilon)}.$$

By the binomial theorem

$$\lim_{\epsilon=0}\partial_\epsilon^a(z-\epsilon)^\lambda = (-1)^a a!(\lambda)_a z^{\lambda-a}.$$

Hence considering the equations (43) and (44) in the case where $\lambda = 0$, we see that

$$\mathcal{A}_\omega^\lambda = \sum_{a=0}^{a=\omega}(-1)^a a!(\lambda)_a z^{\lambda-a}\mathcal{A}_{\omega-a}^{(0)}, \tag{51}$$

$$\mathcal{B}_\omega^\lambda = \sum_{a=0}^{a=\omega}(-1)^a a!(\lambda)_a z^{\lambda-a}\mathcal{B}_{\omega-a}^{(0)}, \tag{52}$$

which is what was to be derived.

In addition, the functions $\mathcal{B}^{(0)}$ can be readily expressed in terms of the functions $\mathcal{A}^{(0)}$. Specifically, proceeding as above we have

$$\mathcal{B}_\omega^{(0)} = \frac{1}{\omega!}\lim_{\epsilon=0}\sum_{a=0}^{a=\omega}(\omega)_a\partial_\epsilon^a f^{(1)}(z-\epsilon) \times \partial_\epsilon^{\omega-a}\frac{1}{f(z-\epsilon)}.$$

Thus since

$$\lim_{\epsilon=0} \partial_\epsilon^a f^{(1)}(z-\epsilon) = (-1)^a f_{a+1},$$

we have

$$\mathcal{B}_\omega^{(0)} = \sum_{a=0}^{a=\omega} \frac{(-1)^a f_{a+1}}{a!} \mathcal{A}_{\omega-a}^{(0)}. \tag{53}$$

More generally, we can write the formula

$$\mathcal{B}_\omega^{(\lambda)} = \sum_{c=0}^{c=\omega} (-1)^c \mathcal{A}_{\omega-c}^{(0)} \sum_{a=0}^{a=c} \frac{(\lambda)_a z^{\lambda-a} f_{c-a+1}}{(c-a)!}, \tag{54}$$

which results from combining (52) and (53).

Thus, it is only necessary to calculate the functions $\mathcal{A}^{(0)}$ by recurrences, since then the functions $\mathcal{A}^{(\lambda)}$ and $\mathcal{B}^{(\lambda)}$ can be easily formed from equations (51) and (54).

In order to actually carry out this calculation, we consider the recursion formula (47) for $\lambda = 0$. In this case the left-hand side vanishes for $\omega > 0$ and takes the value one for $\omega = 0$. Thus the equation splits into two equations, which we write in a fractionless form that seems best suited to our application:

$$\begin{cases} f\mathcal{A}_0^{(0)} = 1, \\ f^{\omega+1}\mathcal{A}_\omega^{(0)} = \sum_{a=1}^{a=\omega} \dfrac{(-1)^{a-1} f^{a-1} f_a}{a!} \cdot f^{\omega-a+1} \mathcal{A}_{\omega-a}^{(0)}, \quad \omega > 0. \end{cases} \tag{55}$$

We can now build a table of the function $f^{\omega+1}\mathcal{A}_\omega^{(0)}$ in the following manner.

First we compute the values (for the argument z) of

$$f, f_1, f_2, f_3, \ldots,$$

from which we obtain

$$f_1, -\frac{1}{2} f f_2, \frac{1}{6} f^2 f_3, -\frac{1}{24} f^3 f_4, \ldots.$$

Then we proceede to multiply horizontally and add vertically according to the following scheme, which requires no further explanation:

$$
\text{(I)}\qquad
\begin{array}{r|l}
f\mathcal{A}_0^{(0)} = 1 & \times f_1 \\ \hline
f^2\mathcal{A}_1^{(0)} = f_1 & \times f_1 \\
1 & \times -\frac{1}{2} f f_2 \\ \hline
f^3\mathcal{A}_2^{(0)} = f_1^2 - \frac{1}{2} f f_2 & \times f_1 \\
f_1 & \times -\frac{1}{2} f f_2 \\
1 & \times \frac{1}{6} f^2 f_3 \\ \hline
f^4\mathcal{A}_3^{(0)} = f_1^3 - f f_1 f_2 + \frac{1}{6} f^2 f_3 & \times f_1 \\
f_1^2 - \frac{1}{2} f f_2 & \times -\frac{1}{2} f f_2 \\
f_1 & \times \frac{1}{6} f^2 f_3 \\
1 & \times -\frac{1}{24} f^3 f_4 \\ \hline
\multicolumn{2}{l}{f^5\mathcal{A}_4^{(0)} = f_1^4 - \frac{3}{2} f f_1^2 f_2 + \frac{1}{3} f^2 f_1 f_3 + \frac{1}{4} f^2 f_2^2 - \frac{1}{24} f^3 f_4} \\
\multicolumn{2}{c}{\text{etc.}}
\end{array}
$$

In order to compute the functions $\mathcal{B}_\omega^{(0)}$ we proceed as above and cast the relation (53) in a more convenient form:

$$
(56)\qquad f^{\omega+1}\mathcal{B}_\omega^{(0)} = \sum_{a=0}^{a=\omega} \frac{(-1)^a f^a f_{a+1}}{a!} \cdot f^{\omega-a+1}\mathcal{A}_{\omega-a}^{(0)}.
$$

Then we need only to combine the expressions we just found for the functions $f^{\omega+1}\mathcal{A}_\omega^{(0)}$ with the appropriate multipliers

$$
f_1,\ -f f_2,\ \frac{1}{2} f^2 f_3,\ -\frac{1}{6} f^3 f_4
$$

to obtain

$$
\text{(II)}\qquad
\begin{aligned}
f\mathcal{B}_0^{(0)} &= f_1, \\
f^2\mathcal{B}_1^{(0)} &= f_1^2 - f f_2, \\
f^3\mathcal{B}_2^{(0)} &= f_1^3 - \tfrac{3}{2} f f_1 f_2 + \tfrac{1}{2} f^2 f_3, \\
f^4\mathcal{B}_3^{(0)} &= f_1^4 - 2 f f_1^2 f_2 + \tfrac{2}{3} f^2 f_1 f_3 + \tfrac{1}{2} f^2 f_2^2 - \tfrac{1}{6} f^3 f_4, \\
&\text{etc.}
\end{aligned}
$$

In addition to the closed representation by determinants that we have already stated for the functions $\mathcal{A}$ and $\mathcal{B}$, we can also derive the following expressions, which I will content myself just to note. Specifically,

$$(57)\quad f^{\omega+1}\mathcal{A}_{\omega}^{(0)} = S\frac{(-1)^a(\omega-a)!}{a_1!a_2!\dots a_\omega!}\Big(\frac{f}{0!}\Big)^a\Big(\frac{f_1}{1!}\Big)^{a_1}\Big(\frac{f_2}{2!}\Big)^{a_2}\cdots\Big(\frac{f_\omega}{\omega!}\Big)^{a_\omega},$$

where the sum extends over all positive integers together with zero that satisfy the pair of simultaneous equations

$$(58)\qquad \begin{cases} a + a_1 + a_2 + \dots + a_\omega = \omega \\ 0\cdot a + 1\cdot a_1 + 2\cdot a_2 + \dots + \omega\cdot a_\omega = \omega. \end{cases}$$

If we multiply the general term in the above sum (57) by $\dfrac{\omega}{\omega-a}$ then the sum represents the value of $f^\omega\mathcal{B}_{\omega-1}^{(0)}$; namely

$$(59)\quad f^{\omega}\mathcal{B}_{\omega-1}^{(0)} = \omega S\frac{(-1)^a(\omega-a-1)!}{a_1!a_2!\dots a_\omega!}\Big(\frac{f}{0!}\Big)^a\Big(\frac{f_1}{1!}\Big)^{a_1}\Big(\frac{f_2}{2!}\Big)^{a_2}\cdots\Big(\frac{f_\omega}{\omega!}\Big)^{a_\omega}.$$

The following relations also hold:

$$(60)\quad \begin{cases} f^{\omega+1}\mathcal{A}_{\omega}^{(0)} = \sum\limits_{a=0}^{a=\omega}\dfrac{(-1)^a f^a}{a!}\lim\limits_{\epsilon=0}\partial_\epsilon^{\omega-a}\left\{\dfrac{f(z+\epsilon)-f(z)}{\epsilon}\right\}^a, \\ f^{\omega}\mathcal{B}_{\omega-1}^{(0)} = \omega\sum\limits_{a=0}^{a=\omega}\dfrac{(-1)^a f^a}{a!(\omega-a)}\lim\limits_{\epsilon=0}\partial_\epsilon^{\omega-a}\left\{\dfrac{f(z+\epsilon)-f(z)}{\epsilon}\right\}^a. \end{cases}$$

§ 9.
The Relation between the Functions A and $\mathcal{A}$.

I will now proceed to show that the functions $A_\omega^{(\lambda)}$ and $\mathcal{A}_\omega^{(\lambda)}$ of the last two sections are equal for any argument z, provided only that the integers ω and λ satisfy a certain inequality.

It is well known that if the equation $f(z) = 0$ has no multiple roots then we have the partial fraction decomposition

$$(61)\qquad \frac{1}{f(z)} = \sum_{a=1}^{a=n}\frac{1}{(z-z_a)f^{(1)}(z_a)}.$$

If z is not equal to a root z_a, we can replace z by $z-\epsilon$ and for sufficiently small ϵ expand each term of the sum on the right in a MacLaurin series in increasing powers of ϵ. To get the generating function of $\mathcal{A}$ on the left we multiply the equation by the development of $(z-\epsilon)^\lambda$ in an (infinite) binomial series. We then order the right-hand side according to powers of ϵ. If we compare the result with equation (45), it follows on equating coefficients of ϵ^ω that

$$\begin{aligned}&\mathcal{A}_\omega^{(\lambda)}(z) = \\ (62)\qquad &\sum_{a=1}^{a=n} \frac{1}{(z-z_a)^{\omega+1} f^{(1)}(z_a)} \sum_{c=0}^{c=\omega} (-1)^c (\lambda)_c z^{\lambda-c} (z-z_a)^c.\end{aligned}$$

This expresses the function $\mathcal{A}_\omega^{(\lambda)}$ as a symmetric function of the roots z_a of the equation $f(z)=0$, just as the function $A_\omega^{(\lambda)}$ has been expressed.

Now if $\omega \geq \lambda$, we can obviously write λ for ω in the upper bound of the second sum on the right, since the binomial coefficient $(\lambda)_c$ vanishes for all c between λ (exclusive) and ω (inclusive). But by the binomial theorem, the sum is then $\{z-(z-z_a)\}^\lambda = z_a^\lambda$. Considering the definition (38) of the function A, we have

$$\mathcal{A}_\omega^{(\lambda)}(z) = A_\omega^{(\lambda)}(z), \qquad \omega \geq \lambda,$$

which is what we were seeking to prove.

However, this relation still holds for $\omega < \lambda$, provided $\lambda < \omega + n$. Specifically, if $\omega < \lambda$ we can decompose the sum

$$\sum_{c=0}^{c=\omega}$$

in the expression (62) for $\mathcal{A}_\omega^{(\lambda)}$ into

$$\sum_{c=0}^{c=\lambda} - \sum_{c=\omega+1}^{c=\lambda}.$$

The sum of all the terms from the first part of this decomposition gives, as above, the function $A_\omega^{(\lambda)}$, so that

$$A_\omega^{(\lambda)}(z) - \mathcal{A}_\omega^{(\lambda)}(z) = \sum_{c=\omega+1}^{c=\lambda} (-1)^c (\lambda)_c z^{\lambda-c} \sum_{a=1}^{a=n} \frac{(z-z_a)^{c-\omega-1}}{f^{(1)}(z_a)}.$$

Now it is well known that

$$\sum_{a=1}^{a=n} \frac{\chi(z_a)}{f^{(1)}(z_a)} = 0,$$

whenever $\chi(z_a)$ is an entire rational function of z_a whose degree does not exceed $n-2$. But the numerator $(z-z_a)^{c-\omega-1}$ in the above double sum can be regarded as such a function χ, provided

$$0 \leq c-\omega-1 \leq n-2$$

for all relevant values of c. Since the first part of this inequality is automatically satisfied and since λ is the largest value that c assumes, the only remaining condition for the vanishing of the double sum is that $\lambda-\omega \leq n-1$ or $\lambda-\omega < n$.

Therefore, under the assumption that either $\omega \geq \lambda$ or $\omega < \lambda < \omega+n$, the functions $\mathcal{A}_\omega^{(\lambda)}$ and $A_\omega^{(\lambda)}$ agree completely. Since these two conditions combine to form one, we have the following theorem:

$$\mathcal{A}_\omega^{(\lambda)}(z) = A_\omega^{(\lambda)}(z), \qquad \lambda-\omega < n. \tag{63}$$

If the side condition is not satisfied, then the above expression for the difference of the two functions takes the place of this theorem. However, since all the terms in the first double sum for which c lies between $\omega+1$ (inclusive) and $\omega+n$ (exclusive) vanish as above, the expression simplifies to

$$\begin{aligned} &A_\omega^{(\lambda)}(z) - \mathcal{A}_\omega^{(\lambda)}(z) = \\ &\sum_{c=\omega+n}^{c=\lambda} (-1)^c (\lambda)_c z^{\lambda-c} \sum_{a=1}^{a=n} \frac{(z-z_a)^{c-\omega-1}}{f^{(1)}(z_a)}, \qquad \lambda-\omega \geq n. \end{aligned} \tag{64}$$

It would not be an uninteresting problem to express this last symmetric function of the roots in terms of the coefficients of the equation $f(z)=0$, just as, in view of § 8, we have now done above for the symmetric function $A_\omega^{(\lambda)}$.

§ 10.
The Relation between the Functions B and $\mathcal{B}$.

Under the conditions for which it holds, the identity of the functions B and $\mathcal{B}$ is easier to recognize than that of A and $\mathcal{A}$. To see this, we will once again seek to develop the generating function (of $\mathcal{B}$) in a series of increasing powers of ϵ whose coefficients are expressed in terms of the roots of the equation $f(z)=0$ rather than its coefficients.

As before, we use a partial-fraction decomposition. Specifically, we have

$$\frac{f^{(1)}(z)}{f(z)} = \sum_{a=1}^{a=n} \frac{1}{z - z_a}, \tag{65}$$

which holds whether or not the equation $f(z) = 0$ has multiple roots. If we replace z in this equation by $z - \epsilon$, develop the right-hand side in increasing powers of ϵ, and multiply the equation by the binomial expansion of $(z - \epsilon)^\lambda$, then on comparing coefficients with those in (46) we obtain as above

$$\mathcal{B}_\omega^{(\lambda)}(z) = \sum_{a=1}^{a=n} \frac{1}{(z - z_a)^{\omega+1}} \sum_{c=0}^{c=\omega} (-1)^c (\lambda)_c z^{\lambda - c} (z - z_a)^c. \tag{66}$$

Now if $\omega \geq \lambda$, the last sum becomes z_a^λ, since λ can be written instead of ω for the upper limit of the sum. Thus

$$B_\omega^{(\lambda)}(z) = \mathcal{B}_\omega^{(\lambda)}(z), \qquad \lambda - \omega \leq 0, \tag{67}$$

which is what we wanted to show. On the other hand, if $\omega < \lambda$, we can decompose the sum in the preceding equation in the form

$$\sum_{c=0}^{c=\omega} = \sum_{c=0}^{c=\lambda} - \sum_{c=\omega+1}^{c=\lambda}.$$

The sum of the terms in the first part turns out to be the same as $\mathcal{B}_\omega^{(\lambda)}(z)$, so that we have

$$B_\omega^{(\lambda)}(z) - \mathcal{B}_\omega^{(\lambda)}(z) = \sum_{c=\omega+1}^{c=\lambda} (-1)^c (\lambda)_c z^{\lambda - c} \sum_{a=1}^{a=n} (z - z_a)^{c - \omega - 1}, \qquad \lambda - \omega > 0. \tag{68}$$

This double sum, which in general is different from zero, can easily be expressed in terms of the coefficients of the equation $f(z) = 0$ instead of its roots.

We have now expressed the symmetric functions A and B — which we introduced in § 7 as the most noteworthy cases of the general function C and have made the object of our study — in terms of the coefficients of the equation $f(z) = 0$. Specifically, they have been represented in the form of simply generated determinants whose elements consist of the the polynomial $f(z)$ and its derivatives — assuming, of course, that ω and λ satisfy the above inequalities (which is always the case when $\omega = \infty$). This understood, we may write A for $\mathcal{A}$ and B for $\mathcal{B}$ (just as earlier we could write A or B for C in the formulas of § 7). Moreover, it is worth noting that several expressions derived for $\mathcal{A}$ and A (or $\mathcal{B}$ and B) in two ways — e.g., the relation (36) of § 7 and (51) and (52) of § 8 — have the same formulas but differ in their upper summation limits, one being λ and the other ω. To the extent that the Latin functions correspond to the script functions, we can obviously choose either of the limits, e.g. the smaller, since the extra terms by which one sum exceeds the other will cancel.

§ 11.
Resulting Solution Methods of the Second Kind.

I will now go on to consider solution methods for higher equations that can be derived from the investigations of §§ 7–10. As we pointed out in § 1, we must distinguish between two kinds of methods.

A solution method of the second kind, as characterized in § 1, follows directly from Theorem (II) of § 7. This theorem states that if we set

$$F(z) = \frac{A_\omega^{(\lambda+h)}}{A_\omega^{(\lambda)}} \qquad \text{or} \qquad F(z) = \frac{B_\omega^{(\lambda+h)}}{B_\omega^{(\lambda)}},$$

then whenever z_1 is a root of the equation $f(z) = 0$ lying nearer the arbitrarily chosen point z than any other root we have

$$\lim_{\omega=\infty} . \, F(z) = z_1^h.$$

In § 8 we gave ways to recursively calculate the functions A and B for ever larger values of ω, as well as to form them independently from the polynomial $f(z)$ and the numbers z, λ, h, and ω. In this way the hth power of the root z_1 (or, if you will, the root itself when $h = 1$) can be determined as precisely as one wishes. And because of the arbitrariness of the numbers listed above, the method — two methods actually, depending on whether the function A or B is chosen — can be applied in an infinite variety of ways.

In the very special case where λ is set to zero and h to one in the function A, our method includes a solution technique recently proposed by FÜRSTENAU,*) a technique he derived from an entirely different point of view. The method is also on the one hand a generalization and on the other a specialization of the method of DANIEL BERNOULLI.

Concerning the starting value z, if one just wants to find some unspecified root — not a particular root — by our method, the starting value may be chosen arbitrarily in the entire complex plane, with the exception of a certain one-dimensional manifold, the *exceptional manifold,* which consists of connected lines, line segments, and rays. If m denotes the number of distinct roots (which must naturally be equal to n for the function A, in which multiple roots are excluded), the exceptional manifold divides the entire plane into m distinct regions, each of which contains a single root, be it simple or multiple. Each region has the property that for any point z chosen within it our algorithm yields the root contained within the region. The boundary of each convergence region is a polygon. The polygon is open to infinity whenever the root lies on a corner of the polygon containing the straight lines connecting the roots to one another. Otherwise, the region is a finite, closed polygon. The sides of the polygon pass at right angles through the middle of the lines connecting two roots. The corners, at which at least three polygons meet, and therefore at which the exceptional manifold splits, are the centers of circles that pass through at least three points but do not contain any other roots. All this follows easily from the requirement that if the point z is not to be an exceptional point it must not be equally removed from the nearest roots. Because of this, there is no difficulty in constructing the exceptional manifold when the roots are given.

First we construct the $\frac{m(m-1)}{2}$ lines that pass at right angles through the middle of the lines connecting two roots. These lines contain all the exceptional points, though in general they will not all belong in their entirety to the exceptional manifold. Specifically, a point on one of these normal lines, which is symmetric to two roots, is an exceptional point only when there is no third root that lies nearer

*)Darstellung der reellen Wurzeln algebraischer Gleichungen durch Determinanten der Coefficienten, Marburg, 1860.

than the other two. These lines are divided by the $\frac{m(m-1)(m-2)}{6}$ centers of the circles that pass through each triplet of roots. Throw out the centers of any circles that contain additional roots and connect the remaining points along the $\frac{m(m-1)}{2}$ lines. Finally starting from the outermost of these points draw infinite rays perpendicular to the polygon containing the roots (and therefore along our lines). The exceptional manifold consists of these rays and connecting lines.

If the roots are not given, the exceptional line will be unknown. In this case, however, if we take the starting value at random, the probability of its falling on the exceptional set is zero. Of course if the equation $f(z) = 0$ has only real coefficients, and therefore complex conjugate roots, we clearly cannot choose z to lie on the real line and hope to find complex roots. Indeed in this case it is clear a priori that we can never arrive at a complex result by a sequence of rational operations involving only real numbers.

If the point z is taken on the exceptional manifold, $F(z)$ will not in general approach a fixed limit with increasing ω, though it will remain bounded.

§ 12.
Resulting Algorithms.

By taking $h = 1$ in Theorems (I) and (III) of § 7, we get solution methods that were called of methods the first kind in § 1; i.e., algorithms.

Specifically, if we set

$$F(z) = \frac{A_\omega^{\lambda+1}(z)}{A_\omega^\lambda(z)} \quad \text{or} \quad F(z) = \frac{B_\omega^{\lambda+1}(z)}{B_\omega^\lambda(z)}, \tag{69}$$

then in § 7 we showed first that $F(z_1) = z_1$ and second that the derivatives

$$\partial_z F(z),\ \partial_z^2 F(z),\ \ldots,\ \partial_z^\omega F(z)$$

are zero for $z = z_1$. According to the results of § 2, these properties imply that the equation

$$z' = F(z)$$

represents an *algorithm of the* $(\omega+1)$*th order,* by means of which we can find any root z_1 of the equation $f(z) = 0$ as accurately as we want from a starting point that has only to be chosen sufficiently near z_1. Moreover, the modulus of each approximation will eventually have $\omega + 1$ times as many accurate decimal places as the modulus of the preceding approximation.

Because the nonnegative integers ω and λ are undetermined and the starting value z is arbitrary this method includes infinitely many algorithms, and it is worth while to write them down for the simplest values of ω and λ. I will label each of the algorithms with the denominators (A_ω^λ) or (B_ω^λ) in (69), which characterize the function F.

Once again recall that if the algorithms (A_ω^λ) are actually to have convergence of order $\omega + 1$, we must exclude multiple roots of the equation $f(z) = 0$. On the other hand, the multiplicity of the roots makes no difference when it comes to the rate of convergence of the algorithms (B_ω^λ).

Finally, if A and B are to be represented by the expressions we have previously given, the function A must satisfy the the inequality $\lambda - \omega < n$, and the function B the inequality $\lambda - \omega \leq 0$.

From equation (37) in Theorem (IV) of § 7, we see that our algorithms can be represented in the form

$$\left.\begin{array}{lll} (A_\omega^\lambda) & \qquad & z' = z - \dfrac{A_{\omega-1}^{(\lambda)}(z)}{A_\omega^{(\lambda)}(z)} = z - f \cdot \dfrac{f^\omega A_{\omega-1}^{(\lambda)}}{f^{\omega+1} A_\omega^{(\lambda)}} \\ (B_\omega^\lambda) & \qquad & z' = z - \dfrac{B_{\omega-1}^{(\lambda)}(z)}{B_\omega^{(\lambda)}(z)} = z - f \cdot \dfrac{f^\omega B_{\omega-1}^{(\lambda)}}{f^{\omega+1} B_\omega^{(\lambda)}} \end{array}\right\} \quad \omega > 0.$$

Here the second term on the right can be regarded as a *correction*, which for each algorithm is to be added to the starting value to form the next approximation. (It is not to be confused with the *error* in the starting value or approximation, that was defined in § 2.)

In these equations we now substitute the values of A and B as we expressed them in terms of $A^{(0)}$ and $B^{(0)}$ in (36) of § 7 and (51) and (52) of § 8. To do this we write these relations in the following fractionless form:

$$(70)\qquad \begin{cases} f^{\omega+1}A_\omega^{(\lambda)} = \sum\limits_{a=0}^{\lambda,\omega}(-1)^a(\lambda)_a z^{\lambda-a}f^a\cdot f^{\omega-a+1}A_{\omega-a}^{(0)}, \\ f^{\omega+1}B_\omega^{(\lambda)} = \sum\limits_{a=0}^{\lambda,\omega}(-1)^a(\lambda)_a z^{\lambda-a}f^a\cdot f^{\omega-a+1}B_{\omega-a}^{(0)}, \end{cases}$$

in which we always choose the smaller of the upper summation limits λ, ω. Hence for $\omega = 0, 1, 2, \ldots$ we have

$$\begin{aligned} f\,A_0^{(\lambda)} &= z^\lambda\cdot f\,A_0^{(0)}, \\ f^2A_1^{(\lambda)} &= z^\lambda\cdot f^2A_1^{(0)} - \lambda z^{\lambda-1}f\;\cdot fA_0^{(0)}, \\ f^3A_2^{(\lambda)} &= z^\lambda\cdot f^3A_2^{(0)} - \lambda z^{\lambda-1}f\cdot f^2A_1^{(0)} + \frac{\lambda(\lambda-1)}{2}z^{\lambda-2}f^2\cdot fA_0^{(0)}, \\ &\text{etc.} \end{aligned}$$

For $\lambda = 1, 2, \ldots$ (the case $\lambda = 0$ is an identity), we have

$$\begin{aligned} f^{\omega+1}A_\omega^{(1)} &= z\cdot f^{\omega+1}A_\omega^{(0)} \;- f\cdot f^\omega A_{\omega-1}^{(0)}, \\ f^{\omega+1}A_\omega^{(2)} &= z^2\cdot f^{\omega+1}A_\omega^{(0)} - 2zf\cdot f^\omega A_{\omega-1}^{(0)} + f^2\cdot f^{\omega-1}A_{\omega-2}^{(0)}, \end{aligned}$$

etc. The corresponding equations for the functions B have exactly the same form and may easily be obtained by writing B for A above.

Since we are unable to express $A_\omega^{(0)}$ or $B_\omega^{(0)}$ generally in a simple manner, we instead obtain simple and easily computable formulas of a general character from the original equations — the ones in which λ is arbitrary — by substituting the functions $A^{(0)}$ and $B^{(0)}$ from formulas (I) and (II) of § 8. Specifically,

$$\begin{aligned} f\,A_0^{(\lambda)} &= z^\lambda, \\ f^2A_1^{(\lambda)} &= z^\lambda f_1 - \lambda z^{\lambda-1}f, \\ f^3A_2^{(\lambda)} &= z^\lambda(f_1^2 - \frac{1}{2}ff_2) - \lambda z^{\lambda-1}ff_1 + \frac{\lambda(\lambda-1)}{2}z^{\lambda-2}f^2, \end{aligned}$$

etc. Moreover,

$$\begin{aligned} f\,B_0^{(\lambda)} &= z^\lambda f_1, \\ f^2B_1^{(\lambda)} &= z^\lambda(f_1^2 - ff_2) - \lambda z^{\lambda-1}ff_1, \end{aligned}$$

etc. In this way it would be easy to construct two-dimensional tables of the functions $A_\omega^{(\lambda)}$ or $B_\omega^{(\lambda)}$.

Finally, we can very easily proceed exhibit the simplest algorithms themselves.

First of all, if we try to take $\omega = 0$, we get the degenerate case

$$(A_0^\lambda) \text{ or } (B_0^\lambda) \qquad z' = z$$

as an algorithm with linear convergence. Since the correction here is zero, the error in the approximation is not only proportional to the first power of the error in the starting value; it is equal to it. thus we might allow the attribute "linear" here, but the designation "algorithm" no longer applies, since $\partial_z F(z)$ is not strictly less than one but equals one.

For $\omega = 1$ we get the most general second order or quadratically converging algorithm that can come from this source: namely,

$$(A_1^\lambda) \qquad z' = z - \frac{zf}{zf_1 - \lambda f},$$

$$(B_1^\lambda) \qquad z' = z - \frac{zff_1}{z(f_1^2 - ff_2) - \lambda ff_1}.$$

For the simplest case $\lambda = 0$ we get Newton's algorithm

$$(A_1^0) \qquad z' = z - \frac{f}{f_1},$$

on the one hand and on the other an equally worthy algorithm

$$(B_1^0) \qquad z' = z - \frac{ff_1}{f_1^2 - ff_2},$$

which to my knowledge has not previously been considered. Besides being almost as simple, this latter algorithm (which we have already mentioned in § 3) has the advantage that it converges quadratically even for multiple roots.

For $\omega = 2$ we get the general algorithms of the third order, that is, cubically convergent algorithms; e.g.,

$$(A_2^\lambda) \qquad z' = z - zf\frac{zf_1 - \lambda f}{z^2(f_1^2 - \frac{1}{2}ff_2) - \lambda zff_1 + \frac{\lambda(\lambda-1)}{2}f^2}.$$

For the simplest case $\lambda = 0$ we have

$$(A_2^0) \qquad z' = z - \frac{ff_1}{f_1^2 - \frac{1}{2}ff_2},$$

$$(B_2^0) \qquad z' = z - f\frac{f_1^2 - ff_2}{f_1^3 - \frac{3}{2}ff_1f_2 + \frac{1}{2}f^2f_3}.$$

Of these algorithms the first (A_2^0) is noteworthy because of its similarity to (B_1^0), from which it differs only by the factor $\frac{1}{2}$ in the denominator.

One can easily continue in this way to construct algorithms of quartic and higher convergence. However, proceeding further has little to recommend it, since in practice the disadvantage of having to evaluate a much more complicated expression more than outweighs the advantages of faster convergence.

As an example, for the binomial or pure equation of the nth degree

$$f(z) = z^n - \gamma = 0,$$

we get the following algorithms:

$$(A_1^\lambda) \qquad z' = z \cdot \frac{(n-\lambda-1)z^n + (\lambda+1)\gamma}{(n-\lambda)z^n + \lambda\gamma}, \qquad \lambda < n+1,$$

$$(B_1^\lambda) \qquad z' = z \cdot \frac{(n+\lambda)\gamma - \lambda z^n}{(n+\lambda-1)\gamma - (\lambda-1)z^n}, \qquad \lambda < 2.$$

Here the two B-algorithms are included among the $n+1$ A-algorithms: namely, (B_1^0) is the same as (A_1^{n-1}) and (B_1^1) the same as (A_1^n).

Suppose we want to find, say, the square root of a number to a very large number of decimal places, e.g. 24. Then after after finding a very good approximation (exact to twelve digits) by the usual root extraction by synthetic division, we should use one of the above formulas to determine a subsequent approximation that is exact to 24 digits. An advantage of this method is that one can make a finite number of arbitrary mistakes in the calculation and (provided one does not jump out of the convergence region) still arrive sooner or later at the correct final value. Moreover, one has a reliable estimate of the current precision; the method has attained the twice number of digits to which the current approximation agrees with the previous one.

An answer to the following question would be of interest. For a particular root extraction what value of λ is the most suitable; i.e., gives the fastest convergence?

One could perhaps combine different algorithms to good effect. For example, one might substitute the value obtained from the algorithm (A_1^0) for z in the formula for (A_1^1) in order to find find the second approximation, and substitute this value in the formula for (A_1^2), and the resulting third approximation in the formula for (A_1^3), and so on. Thus instead of an iteration or the repeated execution of substitutions of the same kind, one must perform a given sequence substitutions of different kinds to approximate the desired root.

Finally it would be worth while to investigate the limits attained by these algorithms when λ is not a positive integer.

§ 13.
On the Convergence Regions of These Algorithms.

There remains the problem of finding the convergence regions of the algorithms we have just given; i.e., the problem of determining the boundaries of the regions, at least when the roots of the equation $f(z) = 0$ are given. Though this problem was easily dispatched in § 11 for solution methods of the second kind, it appears to be comparatively difficult for solution methods of the first kind or the algorithms presented in the previous section. I have succeeded in settling the question of the boundaries of the convergence domains only for the simplest cases: namely, for linear equations — or more generally for equations with one root — and for the quadratic equation.

For the case $\lambda = 0$ the following theorem will help with the solution of the problem.

The contours of the convergence domains of the algorithms (A_ω^0) *or* (B_ω^0) *depend only on the mutual (relative) positions of the* m *distinct roots* $z_1, z_2, \ldots, z_m$. *But they do not depend on the positions* *of this system of points with respect to the points* 0 *and* 1*, or more generally with respect to the real and imaginary axes.* In other words: *If one replaces the system of roots with another similar system positioned arbitrarily, then the contours of new convergence domains will be similar to the old and similarly positioned with respect to the new roots.*

Proof. Imagine two complex planes. The roots z_1, z_2, ..., z_m of the equation $f(z) = 0$ are represented as points in the first; the same number of roots ζ_1, ζ_2, ..., ζ_m of another equation $\varphi(\zeta) = 0$ are represented in the second. Moreover, let z in the first plane and ζ in the second be arbitrary starting points. Let the corresponding approximations from our algorithms be

$$z' = z - \frac{C^{(0)}_{\omega-1}(z)}{C^{(0)}_{\omega}(z)},$$

where C is formed for the function f, and

$$\zeta' = \zeta - \frac{C^{(0)}_{\omega-1}(\zeta)}{C^{(0)}_{\omega}(\zeta)},$$

where C is formed for the function φ. Here, as previously, C denotes either the function A or B. If we fix the relation

$$\zeta = \mu z + \nu$$

and also assume that

$$\zeta_a = \mu z_a + \nu, \qquad a = 1, 2, 3, \ldots, m,$$

then it is easy to see that for arbitrary complex numbers μ and ν the system of points ζ, ζ_a in the second plane is *similar* to the system z, z_a in the first, although the second system is positioned arbitrarily with respect to the real and complex axes. For if $\mu = \rho e^{i\vartheta}$, then multiplication of the number z by ρ effects a transformation of the point system into a similar system, similarly situated with respect to the axes, whose homologous dimension is ρ times as large as that of the first. The multiplication by $e^{i\vartheta}$ effects a common rotation of the system of points through the arbitrary angle ϑ. Finally the addition of ν to the product $\mu z = \rho e^{i\vartheta} z$ corresponds to a translation of the entire system in the direction of the number ν and of length equal to the modulus of ν.

The proof of our theorem will now be complete if we can show that the approximations z' and ζ' are homologous points of the two similar systems of points. For this conclusion can easily be extended to all following and preceding approximations, right up the the boundary of the convergence region.

Now the equation whose roots are $\zeta_a = \mu z_a + \nu$ is obviously

$$\varphi(\zeta) = f\left(\frac{\zeta - \nu}{\mu}\right) = 0.$$

Hence

$$\varphi^{(1)}(\zeta) = \frac{1}{\mu} f^{(1)}\left(\frac{\zeta - \nu}{\mu}\right),$$

and more generally

$$\varphi^{(c)}(\zeta) = \frac{1}{\mu^c} f^{(c)}\left(\frac{\zeta - \nu}{\mu}\right),$$

for every natural number c. Since $\dfrac{\zeta - \nu}{\mu} = z$,

$$\varphi(\zeta) = f(z),\ \varphi^{(1)}(\zeta) = \frac{1}{\mu} f^{(1)}(z),\ \ldots,\ \varphi^{(c)}(\zeta) = \frac{1}{\mu^c} f^{(c)}(z).$$

Moreover, in all terms of the equation (57) of § 8 for the function $f^{\omega+1} A_\omega^{(0)}$ and in the corresponding equation for $f^\omega B_{\omega-1}^{(0)}$ the sums of the the products of the exponents and the derivative indices are the same, namely ω. Hence the expressions

$$\varphi(\zeta)^{\omega+1} A_\omega^{(0)}(\zeta) \quad \text{and} \quad \varphi(\zeta)^\omega B_{\omega-1}^{(0)}(\zeta)$$

formed for the function φ are equal to the product of the factor $\dfrac{1}{\mu^\omega}$ with the expressions

$$f(z)^{\omega+1} A_\omega^{(0)}(z) \quad \text{and} \quad f(z)^\omega B_{\omega-1}^{(0)}(z)$$

formed for the function f. From the second of the two equations

$$z' = z - f(z) \frac{f(z)^\omega C_{\omega-1}^{(0)}(z)}{f(z)^{\omega+1} C_\omega^{(0)}(z)}, \qquad [C \text{ formed for } f],$$

and

$$\zeta' = \zeta - \varphi(\zeta)\frac{\varphi(\zeta)^{\omega} C^{(0)}_{\omega-1}(\zeta)}{\varphi(\zeta)^{\omega+1} C^{(0)}_{\omega}(\zeta)}, \qquad [C \text{ formed for } \varphi],$$

it follows that

$$\zeta' = \mu z + \nu - \mu f(z)\frac{f(z)^{\omega} C^{(0)}_{\omega-1}(z)}{f(z)^{\omega+1} C^{(0)}_{\omega}(z)}, \qquad [C \text{ formed for } f].$$

Hence from the first equation,

$$\zeta' = \mu z' + \nu,$$

which is what was to be shown.

The contours of the convergence domains in the similar systems of z and ζ are therefore similar curves with the same ratio ρ for their homologous dimensions and situated similarly with respect to their systems of points. Thus, when it comes to the study of these contours, one can work in the system ζ as well as in the system z. For example, without loss of generality we can take two of the roots of the equation in question to be arbitrary, say one equal to zero and the other equal to one, since it is only a matter of the relative position of the roots. We can also take all the roots to be as near as we like to each other as well as to the origin, since we can imagine the ratio ρ to be arbitrarily small. An so on.

Incidentally, one important consequence of all this is that the exclusion of a root of zero of the equation $f(z) = 0$, which was necessary for many of our theorems, is irrelevant for the conclusions about the algorithms (A^0_ω) and (B^0_ω).

The above theorem does not hold for algorithms (A^λ_ω) or (B^λ_ω), where $\lambda > 0$. If one attempts to carry out the line of argument in the proof of the theorem for this case, it becomes evident that one must take the number ν to be zero for the new approximation to be a homologous point of the system that is similar to the old system. Thus one can change the fundamental unit of the system; i.e., one can replace the system of roots by another similar system that is similarly positioned with respect to the axes. One can also rotate the system about the origin by an arbitrary angle. However, one cannot translate the system. In other words, we have the following theorem.

For the algorithms (A_ω^λ) *and* (B_ω^λ) *the relative position of the contours of the convergence domain with respect to the roots depends only on the ratio of the distance of the center of gravity of all the roots (the midpoint of their mean distances) from the origin to the common distances among these roots. But in general it is independent of the absolute position of the roots and of the fundamental unit.*

So much for the convergence regions in general.

§ 14.
The Principal Algorithms Applied to Very Simple Examples.

In order to get a closer look at the nature of our algorithms, we will now write down the two most useful ones, namely

$$(A_1^0) \quad z' = z - \frac{f}{f_1} \quad \text{and} \quad (B_1^0) \quad z' = z - \frac{ff_1}{f_1^2 - ff_2},$$

for a simple example and arrange them for practical calculation.

To take care of the simplest case, namely where the equation $f(z) = 0$ has only *one* root, so that it is either of the first degree or as a polynomial is the power of a binomial, we note that according to what has gone before we can set $z_1 = 0$ without loss of generality. Thus we can assume that

$$f(z) = z^n.$$

Hence

$$(A_1^0) \quad z' = \left(1 - \frac{1}{n}\right)z \quad \text{and} \quad (B_1^0) \quad z' = 0.$$

The algorithm (B_1^0) thus gives the correct root (0) of the equation right away. On the other hand for the algorithm (A_1^0) it is easily seen that

$$z'' = \left(1 - \frac{1}{n}\right)^2 z', \ \ldots, \ z^{(r)} = \left(1 - \frac{1}{n}\right)^r z,$$

from which it follows that

$$\lim_{r=\infty} z^{(r)} = 0,$$

whatever the starting value. Thus we have the following theorem. *For an equation with a single root the convergence region of the two algorithms* (A_1^0) *and* (B_1^0) *is the entire complex plane.* There are no finite exceptional points.

The next simple case is the quadratic equation. The case of two equal roots is subsumed in the above case and has already been dealt with. Thus we can assume that the roots are distinct. To gain symmetry it is advisable to take the roots to be $z_1 = +1$ and $z_2 = -1$, which, as we have said, we can do without loss of generality. Then

$$f(z) = (z-1)(z+1) = z^2 - 1,$$

and the algorithms become

$$(A_1^0) \quad z' = \frac{1+z^2}{2z} = \frac{z + \dfrac{1}{z}}{2}, \qquad (B_1^0) \quad z' = \frac{2z}{1+z^2} = \frac{2}{z + \dfrac{1}{z}}.$$

Thus we see that for the same starting value the approximation for one algorithm is always the reciprocal of the approximation for the other.

Now if we want to actually compute the approximation $z' = x' + iy'$ corresponding to some complex starting value $z = x + iy$, we must decompose the *complex* algorithm into *two combined real* algorithms. In general we have to do this for any algorithm $z' = F(z)$ whenever the coefficients and roots, not to mention the starting value, are not all real. This algorithmic decomposition, which results from separating and equating real and imaginary parts on both sides of the equation $z' = F(z)$, would be easy to write down for any one of our algorithms after we take the coefficients $\gamma_0, \gamma_1, \ldots, \gamma_n$ of the equation $f(z) = 0$ in the form

$$\gamma_a = \alpha_a + i\beta_a, \qquad a = 0, 1, \ldots, n.$$

Now for our example,

$$(A_1^0) \qquad x' = \frac{x}{2} \cdot \frac{1 + x^2 + y^2}{x^2 + y^2}, \qquad y' = -\frac{y}{2} \cdot \frac{1 - x^2 - y^2}{x^2 + y^2},$$

$$(B_1^0)\qquad \begin{aligned} x' &= 2x\cdot\frac{1+x^2+y^2}{[x^2+(y+1)^2][x^2+(y-1)^2]},\\ y' &= 2y\cdot\frac{1-x^2-y^2}{[x^2+(y+1)^2][x^2+(y-1)^2]}. \end{aligned}$$

from which the first approximation corresponding to any starting value is easily computed. If we want to find the next approximation, we need only replace x and y by x' and y' in the above equations to obtain x'' and y''; i.e., the elements of $z'' = x'' + iy''$. And so on.

Many conclusions can be drawn from these equations. To express them it will be convenient to regard the passage from the starting value to the approximation as a *jump* of the argument z from its position in the complex plane to the position of the approximation. For example, since $y' = 0$ whenever $x^2+y^2 = 1$, we have the following theorem. From the periphery of the unit circle about the origin the argument always jumps to the real axis.

Moreover, since $y' = 0$ for $y = 0$, the approximations for a real starting value are all real: the argument never jumps away from the real axis. Similarly, for a purely imaginary argument all the approximations remain purely imaginary. Therefore, the argument never jumps away from the imaginary axis, and it is impossible for the algorithm to converge to the points ± 1 from such a starting point. The y axis in its entirety belongs to the divergence region; that is, it consists entirely of exceptional points. In addition, by looking for values of z for which the correction vanishes or tends to infinity we can find individual exceptional points for an arbitrary algorithm. Also we can look for values of z for which at some

$$r = 1, 2, 3, \ldots$$

the rth approximation $z^{(r)}$ becomes equal to the starting value, so that the algorithm repeats itself, i.e., is periodic. It would not be difficult to even discover exceptional curves in general — though it would not be easy to determine whether they actually contained all the exceptional points and whether they perhaps fill in an entire region of the plane.

If the sign of x or y or both is changed, the corresponding changes occur in the approximations x' and y'. Thus the algorithm proceeds symmetrically in the four quadrants of the complex plane. The argument will never jump over the imaginary axis, since x' and x always have the same sign. However, the argument can jump the real axis. This happens for the algorithm (A_1^0) when the argument lies within the circle mentioned above, since then y' and y are of opposite signs. If the argument is outside the circle, its jump leaves it on the same side of the x-axis. Exactly the opposite happens for the algorithm B_1^0.

It is also informative to transform the algorithm to polar coordinates, so that the the radius and polar angle of the approximation can be calculated from those of the starting value. If we set

$$x + iy = \rho e^{i\vartheta} \quad \text{and} \quad x' + iy' = \rho' e^{i\vartheta'},$$

then from the equations

$$(A_1^0) \;\; x'^2 + y'^2 = \frac{[x^2 + (y+1)^2][x^2 + (y-1)^2]}{4(x^2 + y^2)}, \qquad \frac{y'}{x'} = -\frac{y}{x} \cdot \frac{1 - x^2 - y^2}{1 + x^2 + y^2}$$

we get the combined algorithms

$$(A_1^0) \qquad \rho' = \frac{1}{2\rho}\sqrt{1 + 2\rho^2 \cos 2\vartheta + \rho^4}, \quad \tan\vartheta' = -\frac{1-\rho^2}{1+\rho^2}\tan\vartheta.$$

In the case of (B_1^0) we take the reciprocal value for ρ' and the value of opposite sign for $\tan\vartheta'$.

The factor $\dfrac{1-\rho^2}{1+\rho^2}$ is always a proper fraction, and hence the numerical value of $\tan\vartheta'$ is always less than the value of $\tan\vartheta$, exclusive of the cases $\tan\vartheta = 0$ or ∞. Since we can confine ourselves to acute angles, the angle ϑ' is itself less than ϑ. In other words: with each jump the radius vector rotates toward the polar axis and as the iteration is continued swings almost like a pendulum toward this equilibrium position.

One can also ask what curve separates the regions whose points jump nearer to or farther from zero, which is the mean of the roots. It is the curve whose equation results from requiring that $\rho'^2 = \rho^2$. In the case of (B_1^0), for example, the equation reads

$$[x^2 + (y+1)^2][x^2 + (y-1)^2] = 4$$

or

$$(x^2 + y^2 + 1)^2 = 4(1 + y^2),$$

from which we get

$$x = \pm\sqrt{2\sqrt{1+y^2} - (1+y^2)} \quad \text{and} \quad y = \pm\sqrt{1 - x^2 + 2\sqrt{1-x^2}}.$$

In polar coordinates

$$\rho^4 + 2\rho^2 \cos 2\vartheta = 3.$$

More interesting than questions about polar coordinates is the question of the distances ρ_1, ρ_2 and ρ_1', ρ_2' of z and z' from the points $+1$ and -1. These distances can be regarded as the radii of the points in a bipolar coordinate system, whose poles are the two roots ± 1. In particular, from the expressions we are seeking we can learn whether the argument jumps nearer a root or not.

Now we have the equations

$$\rho_1^2 = y^2 + (x-1)^2, \quad \rho_2^2 = y^2 + (x+1)^2$$

(and the same for quantities ρ, x, y equipped with primes). Hence by inversion

$$x = \frac{\rho_2^2 - \rho_1^2}{4}, y = \frac{1}{16}(2+\rho_1+\rho_2)(\rho_1+\rho_2-2)(\rho_2+2-\rho_1)(2+\rho_1-\rho_2).$$

If we substitute these values in the equations

$$(A_1^0) \qquad y'^2 + (x'-1)^2 = \frac{[y^2+(x-1)^2]^2}{4(x^2+y^2)}$$

(and the same with $+1$ replaced by -1)

$$(B_1^0) \qquad y'^2 + (x'-1)^2 = \frac{[y^2+(x-1)^2]^2}{[x^2+(y+1)^2][x^2+(y-1)^2]}$$

(and the same with $+1$ replaced by -1)

we get the combined algorithms

$$(A_1^0) \qquad \rho_1'^2 = \frac{\rho_1^4}{2(\rho_1^2+\rho_2^2-2)}, \quad \rho_2'^2 = \frac{\rho_2^4}{2(\rho_1^2+\rho_2^2-2)},$$

$$(B_1^0) \qquad \rho_1'^2 = \frac{2\rho_1^4}{(\rho_1^2-2)^2+(\rho_2^2-2)^2}, \quad \rho_2'^2 = \frac{2\rho_2^4}{(\rho_1^2-2)^2+(\rho_2^2-2)^2}.$$

The equation of the curve that separates the region of points that jump toward the point $+1$ from the region of points that jump away is now determined by the condition that $\rho_1'^2 = \rho_1^2$. For the algorithms (A_1^0) it is $\rho_1^2 + 2\rho_2^2 - 4 = 0$ in bipolar coordinates or $3y^2 + 3x^2 + 2x - 1 = 0$ in rectangular coordinates. In other words, the curve is the circle of radius $\frac{2}{3}$ and center $-\frac{1}{3}$. For the algorithm (B_1^0), the equation of the corresponding curve is of the fourth degree, which is easily solved for y. The curve itself contains all points that merely *rotate* about about the point one with a step of the algorithm.

Finally, in order to compute the starting value backward from the approximation — that is, compute the predecessor of an arbitrary approximation — one has only to solve the following equations for z:

$$(A_1^0) \quad z^2 - 2zz' + 1 = 0, \qquad (B_1^0) \quad z^2 - \frac{2z}{z'} + 1 = 0.$$

Since these equations are of the second degree, corresponding to each approximation z' there are two starting values $[z]_1$ and $[z]_2$ that come together at z' and remain united for all subsequent steps. Thus the points of the convergence region form an infinite family of points, all of which sooner or later jump toward the root and cluster about it with infinite density. It is easy to give formulas by which these operations can be extended backwards or forwards.

In a subsequent treatise on *iterated functions* I will give a proof that for the above algorithms and equations the entire complex plane decomposes into two convergence regions separated by the imaginary axis, which is the only exceptional line. Moreover, I entertain the conjecture that for any algebraic equation the region of exceptional points of these algorithms is only *one* dimensional and reduces to the boundary lines of the convergence domains.

In the same place other questions concerning the above example of the two algorithms will be answered. Nonetheless, the considerations of this section will turn out not to be superfluous.

§ 15.
APPENDIX:
A Theorem on the Function A.

In this appendix I will communicate another theorem — one that formed the original starting point for the development of the algorithms of § 12. It would seem to be of interest because the functions A_ω^λ play a role in it.

As previously, let

$$f(z) = \sum_{a=0}^{a=n} \gamma_a z^{n-a} = \gamma_0(z - z_1)(z - z_2)\cdots(z - z_n), \tag{71}$$

and let the roots z_1, z_2, ..., z_n be distinct from one another and hence simple.

The function $f(z)$ can be divided without remainder by the difference $z - z_a$. From the usual division algorithm, one obtains a result of the form

$$\frac{f(z)}{z - z_a} = \sum_{b=0}^{b=n-1} z^{n-b-1}\mathcal{F}_b(z_a). \tag{72}$$

This equation defines certain entire rational functions $\mathcal{F}_b(z_a)$. On multiplying the equation by $z - z_a$ and equating coefficients we obtain the recursion

$$\begin{cases} \mathcal{F}_0(z_a) = \gamma_0, & \\ \mathcal{F}_b(z_a) = z_a\mathcal{F}_{b-1}(z_a) + \gamma_b, & b = 1, 2, 3, \ldots, n-1, \\ \quad 0 \quad = z_a\mathcal{F}_{n-1}(z_a) + \gamma_n, & \end{cases} \tag{73}$$

by means of which the functions can be computed. The next-to-last of the recursions will also hold for $b = n$ if we define $\mathcal{F}_n(z) = f(z)$. With the help of these recursions we obtain the well-known representation

$$\mathcal{F}_b(z_a) = \sum_{c=0}^{c=b} \gamma_c z_a^{b-c}, \qquad b = 0, 1, \ldots, n. \tag{74}$$

We will allow this representation to hold for arbitrary values of the argument z other than z_a, and eventually for $b > n$, in which case we let the coefficients γ_{n+1}, γ_{n+2}, ... be arbitrary.

If in equation (72) we first substitute z_c for z, where c denotes a number from 1, 2, ..., n that is different from a, and next substitute z_a for z, then from

$$f(z_c) = 0 \qquad \text{and} \qquad \lim_{z=z_a} \frac{f(z)}{z - z_a} = f^{(1)}(z_a)$$

we have the following important relations:

$$(75)\qquad \begin{cases} \sum\limits_{b=0}^{b=n-1} z_c^{n-b-1}\mathcal{F}_b(z_a) = 0, & c \neq a \\ \sum\limits_{b=0}^{b=n-1} z_a^{n-b-1}\mathcal{F}_b(z_a) = f^{(1)}(z_a) & \end{cases}$$

We can now state the following theorem, which is based on these relations.

In the two systems of linear equations

$$(76)\qquad \begin{cases} \sum\limits_{c=0}^{c=n-1} z_a^{n-c-1}X_c = Y_a f^{(1)}(z_1), & a = 1, 2, \dots, n, \\ X_c = \sum\limits_{a=1}^{a=n} \mathcal{F}_c(z_a)Y_a, & c = 0, 1, \dots, n-1 \end{cases}$$

or the two systems

$$(77)\qquad \begin{cases} \sum\limits_{c=0}^{c=n-1} \mathcal{F}_{n-c-1}(z_a)\mathcal{Y}_c = \mathcal{X}_a f^{(1)}(z_1), & a = 1, 2, \dots, n, \\ \mathcal{Y}_c = \sum\limits_{a=1}^{a=n} z_a^c \mathcal{X}_a, & c = 0, 1, \dots, n-1 \end{cases}$$

one is always the solution of the other.

For if we substitute the values from second of these equations into the first, where obviously the summation index a must be replaced by another, say b, then (75) implies that the result is an identity.

If for the converse we substitute the expressions from the from the first equations into the second, the truth of the theorem is not immediately evident from what has gone before. Instead we are led in this way to the relation

$$(78)\qquad \sum_{a=1}^{a=n} \frac{z_a^b \mathcal{F}_c(z_a)}{f^{(1)}(z_a)} = \begin{cases} 0, & \text{when } b + c \neq n-1, \\ 1, & \text{when } b + c = n-1, \end{cases}$$

which can be easily derived from a relation given by CAUCHY.*) The solution of systems (76) and (77) has also be carried out by BALTZER from determinantal considerations.**)

Further we now have the following theorem.

*) Cf. BALTZER, Determ. 2nd Edition, p. 79

**) Ibid., p. 81 ff.

Any positive integer power of a linear function of the quantities

$$\mathcal{F}_0(z_a), \mathcal{F}_1(z_a), \ldots, \mathcal{F}_{n-1}(z_a),$$

say

$$P = \sum_{c=0}^{c=n-1} \mathcal{Y}'_a F_{n-c-1}(z_a),$$

in which the coefficients are arbitrary constants, can be expressed as a linear function of the same n quantities. The coefficients of this linear function are symmetric functions of all the roots and therefore contain only the coefficients $\mathcal{Y}'$ and the coefficients γ of the polynomial $f(z)$, but not the roots z_a.

The simplest way of seeing this is the following. Imagine that the expressions for the functions $\mathcal{F}$ formed from the scheme in equation (74) have been substituted in the expression for $P^{\omega+1}$. It is clear that this expression can be ordered according to powers of the root z_a. By means of the equation $f(z) = 0$ the powers of z_a whose exponent is greater than $n - 1$ can be expressed in terms of the lower powers of z_a. This exhibits $P^{\omega+1}$ as a linear function of the quantities

$$z_a^0, z_a^1, \ldots, z_a^{n-1}.$$

By solving the system of equations (74), we can represent these latter quantities by means of the original quantities

$$\mathcal{F}_0(z_a), \mathcal{F}_1(z_a), \ldots, \mathcal{F}_{n-1}(z_a),$$

and when these representations are substituted we indeed obtain $P^{\omega+1}$ as a linear combination of the quantities $\mathcal{F}$.

However, the following proof may be worth noting.

If we imagine the $(\omega + 1)$th power of the sum P developed according to the binomial theorem, then we get a sum of terms, each of which is a product of factors that are the powers of the individual functions $\mathcal{F}_0(z_a)$, $\mathcal{F}_1(z_a)$, $\ldots$, $\mathcal{F}_{n-1}(z_a)$. Such a product, and hence the entire sum, can be expressed linearly in terms of the function $\mathcal{F}$, provided we are able to solve the same problem for the product of any two of these functions.

Thus we pose ourselves the general problem: for any two natural numbers a and b, express the product

$$\mathcal{F}_a(z)\mathcal{F}_b(z)$$

linearly in terms of the quantities $\mathcal{F}_0(z)$, $\mathcal{F}_1(z)$, ... Here we can dispense with the unchanging argument z of the functions $\mathcal{F}$. If we multiply the equation (74) by $\mathcal{F}_a$ and extract the term multiplied by the zero power of z on the right, we get

$$\mathcal{F}_a\mathcal{F}_b = \gamma_b\mathcal{F}_a + z\mathcal{F}_z \sum_{c=0}^{c=b-1} \gamma_c z^{b-c-1} \quad \text{or} \quad \mathcal{F}_a\mathcal{F}_b = \gamma_b\mathcal{F}_a + z\mathcal{F}_a\mathcal{F}_{b-1}.$$

Moreover, since by (73)

$$z\mathcal{F}_a = \mathcal{F}_{a+1} - \gamma_{a+1},$$

it follows by substitution that

$$\mathcal{F}_a\mathcal{F}_b - \mathcal{F}_{a+1}\mathcal{F}_{b-1} = \gamma_b\mathcal{F}_a - \gamma_{a+1}\mathcal{F}_{b-1}.$$

If we write $a+c$ for a and $b-c$ for b and sum over c from 0 to $b-1$, then, recalling that $\mathcal{F}_0 = \gamma_0$, we get by a suitable reordering of the terms

$$\mathcal{F}_a\mathcal{F}_b = \gamma_b\mathcal{F}_a + \sum_{c=1}^{c=b} \left(\gamma_{b-c}\mathcal{F}_{a+c} - \gamma_{a+c}\mathcal{F}_{b-c}\right). \tag{79}$$

This is the desired representation.

This theorem, which we have proved in two ways, will now be applied to the case where P is the expression (72) itself: namely,

$$\frac{f(z)}{z - z_a} = \sum_{c=0}^{c=n-1} z^{n-c-1}\mathcal{F}_c(z_a),$$

in which $\mathcal{Y}'_c = z^c$. As we have shown, the $(\omega + 1)$th power of this expression may be represented linearly in terms of the functions $\mathcal{F}$. Our problem is to actually construct the representation

$$\left\{\frac{f(z)}{z - z_a}\right\}^{\omega+1} = \sum_{c=0}^{c=n-1} \mathcal{Y}_c\mathcal{F}_c(z_a).$$

This can be done directly with the help of the Theorem (77). Specifically, imagine the last equation has been written down for $a = 1, 2, \ldots, n$, in which we take

$$\frac{f(z)^{\omega+1}}{f^{(1)}(z)(z - z_a)^{\omega+1}}$$

for $\mathcal{X}$. Then considering definition (38) of the function A, we immediately see that the solution of the system of equations is

$$\mathcal{Y}_c = f(z)^{\omega+1} A_c^{(c)}(z).$$

By substituting these values and deleting the common factor $f(z)^{\omega+1}$ we get the identity

$$\frac{1}{(z - z_a)^{\omega+1}} = \sum_{c=0}^{c=n-1} A_\omega^{(c)} \mathcal{F}^{n-c-a}(z_a). \tag{80}$$

Thus we can state the following theorem.

If the $(\omega + 1)$*th power of*

$$\frac{f(z)}{z - z_a}$$

is expressed linearly in terms of the functions $\mathcal{F}_0(z_a), \ldots, \mathcal{F}_{n-1}(z_a)$, *then as* ω *grows unboundedly the ratio of the coefficients of any successive pairs of these functions approaches the root of the equation* $f(z) = 0$ *that lies nearest the arbitrary value* z.

To compute the coefficients $f^{\omega+1} A_\omega^{(c)}$ one can also use the polynomial theorem along with the relation (79).

Pforzheim, January 1869