ON INFINITELY MANY ALGORITHMS FOR THE SOLUTION OF AN ANALYTIC EQUATION

by

W. J. Kotzé, B. Sc. Hon.

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of Master in Science.

Department of Mathematics,
McGill University,
Montreal.

August 1961.

Contents.

Preface and a few historical notes

This thesis is the result of an interest aroused, especially in the relatively old paper of E. Schröder [1], by Prof. H. Schwerdtfeger whose brilliant lectures on numerical analysis I attended at McGill University during the year 1960 - 61. After some further reading, the temptation was great to approach the subject from a functional analytical point of view as did L. V. Kantorovich in his "Functional analysis and applied mathematics" (Translated from the Russian (1948). Originally printed in Russian in Uspekhi matematicheskikh Nauk, Vol. III, No. 6, 89 - 185, 1948.); L. Collatz (Näherungsverfahren höherer Ordnung für Gleichungen in Banach - Räumen; Archive for Rational Mechanics and Analysis 2, (1958 - 59), 66 - 75); J. Schröder (Über das Newtonsche Verfahren; Arch Rat. Mech. and An. 1, (1957 - 58), 154 - 180.) and others. However, after having read H. Ehrmann's much neglected and relatively unknown paper [2], I decided to deal with the matter in the conventional functional theoretical way Schröder, Bodewig and others did.

The problem of solving equations by means of iterative methods is not a very new one (i. e. in a mathematical sense of speech). Newton was probably the first (1674) who applied this type of method to the equations

$$x - e \operatorname{Sin} x = N$$

and $$e \operatorname{Sinh} x - x = N \qquad \text{(N constant)}$$

(See: "Principia" (1687) Prop. 31, Book 1). These two equations arose from Kepler's problem to find the position of a planet at a given time in an elliptic or hyperbolic orbit of eccentricity e. Substantially the same method is also mentioned by Newton in his "De analysi per aequationes numero terminarum infinitas". The earliest printed account of this appeared in Wallis' "Algebra" (1685) Chap. 94.

In 1690, Joseph Raphson (1648 - 1715), a fellow of the Royal Society of London, published a tract "Analysis aequationum universalis", in which he expressed Newton's method in the now well-known algorithmic form. Raphson's version of the process represents what J. Lagrange (Resolution des equat. num. (1798), Note V, p. 138) recognized as an advance on the scheme of Newton. According to him the method is "plus simple que celle de Newton". We may add here, that the solution of numerical equations was considered geometrically by Thomas Baker (1684) and Edmund Halley (1687), but in 1694 Halley "had a very great desire of doing the same in numbers". The only difference between Halley's and Newton's methods is that Halley solves a quadratic equation at each step, Newton a linear equation. Halley also modified certain algebraic expressions, yielding approximate cube and fifth roots, given in 1692 by Thomas Fantet de Lagny (1660 - 1734).

In the work following, special attention is given to the second and third order algorithms which are variations of the Newton - Raphson method. This attention is especially directed at the convergence of the different modifications and the error estimate of each. In chapter V the construction of two types of higher order algorithms is discussed. These constructions are very often quite laborious, and in practice it was found that in most cases not much is gained in the use of algorithms of order higher than three. However, there does exist a need for a means of choosing the most expedient algorithm for a specific function. In chapter VII as attempt was made to comply with this demand. Chapter IX consists of a short résumé of well-known, and some lesser-known theorems and methods which might be of some assistance in determining the approximate location of the roots of an equation. The knowledge of such locations are of grave importance to the sometimes arduous task of choosing an initial approximation to the desired root.

The major claims to originality (if any) are the following:

(1) The corollaries to T. 5

(2) The construction and discussions of modifications II(a), IV(a), IV(b), V(a), V(b) of the N - R method. These involve the corollary to T. 3, theorem 8 and proposition 9.

(3) Theorems 10 and 11

(4) Applications following theorem 13

(5) Table for constructing higher order algorithms, Fig. 4

(6) Theorems 17, 20, 22, 25

For many of the other theorems, already known, revised forms of the proofs appear.

To comply with the regulations of this university, I herewith wish to declare that no help from persons outside was received in the preparation of this thesis in general. I am very much indebted though, to my director of studies, Prof. H. Schwerdtfeger, whose inspirational lectures, already mentioned, formed the keystones to this humble work. I furthermore wish to express my sincere gratitude towards the Canadian Mathematical Congress and the H. B. Webb - Stipend Trustees (Cape Town, South Africa) for the monetary assistance received during the past year.

McGill University, Montreal (Canada).

Aug. 1961.

Index of Notations and Symbols used.

T.	:	Theorem
P.	:	Proposition
$z_n \xrightarrow[\infty]{n} a$	:	z_n tends to a as n tends to infinity
$a \in D$	:	a is an element of D
sup. ( )	:	supremum, least upper bound (l. u. b.)
inf. ( )	:	infimum, greatest lower bound (g. l. b.)
iff	:	if and only if

I

Introduction

In the following pages we are going to consider the iterative solution of any analytic function (regular in the neighbourhood of the roots) of the complex variable z. In other words we are going to construct and investigate numerical methods for finding any root ξ of $f(z) = 0$, where we only consider those cases for which $f(z)$ is continuous in the vicinity of ξ, and $f(z)$ becomes zero of finite order at ξ (i.e. the order of the root ξ is finite).

From the fundamental theorem of algebra, we have that in case of the n-th degree polynomial, the equation

$$f(z) = a_o z^n + a_1 z^{n-1} + \dots + a_n = 0$$

where $n \geq 1$ and the a_j are constants with $a_o \neq 0$, has at least one root.

The well-known theorem by Rouché says however:

Let $g(z)$ and $h(z)$ be two functions, analytic in a simply connected open region D. Let C be a rectifiable Jordan curve lying in D. Suppose that along C, $g(z)$ is nowhere zero, and $|h(z)| < |g(z)|$. Then the function $g(z) + h(z)$ has the same number of zeros within C as $g(z)$ has.

Now let $f(z) = a_o z^n + \dots + a_n$ with $n \geq 1$ and $a_o \neq 0$. If C' is a circle of large radius, with centre at the origin, we have all along C

$$|a_1 z^{n-1} + \dots + a_n| < |a_o z^n|$$

Hence, $f(z)$ has the same number of zeros within C as $a_o z^n$ does. That number is n. Thus $f(z)$ has n roots, and it is the process of finding these roots by means of numerical iteration that we are interested in. We will however, not confine ourselves to polynomials alone. Indeed, any analytic function regular within a simply connected open region circumscribing ξ will be considered.

At this stage it may not be entirely redundant to reiterate the following well-known fact:

Given $f(z)$ analytic within a simply connected open region D. Then the roots of $f(z)$ lying in D will not have a limit point in D; since if that were the case, we know by a well-known theorem on the zeros of an analytic function, that $f(z)$ will vanish identically throughout D - a trivial case which will obviously be excluded in our following discussions.

II

A few basic theorems and the concept of the order of an algorithm

The problem at hand is thus to find an iteration formula

$$z_n = F(z_{n-1}) = G(z_{n-1}, f(z_{n-1}), f'(z_{n-1}), \ldots, f^{(s)}(z_{n-1}))$$
$$s \geq 1, \quad n = 0, 1, 2, \ldots \qquad (2.1)$$

which gives an approximation z_n for a root ξ of the analytic function $f(z) = 0$ after n applications. This implies that the distance of z_{n+1} from ξ will be smaller than the distance of any of the previous z_j $(j = 1, \ldots, n-1)$ from ξ.

Further we want F to be such that

$$\lim_{n\to\infty} z_n = \xi$$

i. e. $$|z_n - \xi| < |z_{n-1} - \xi| \qquad \text{for all } n$$

and $$|z_N - \xi| < \varepsilon , \qquad N \text{ large enough}$$

Thus we can denote $\xi = \lim_{n\to\infty} F^{(n)}(z)$.

The initial value $z = z_o$ must be an arbitrary point within the so-called domain of convergence of the algorithm.

<u>T. 1:</u> If $\lim_{n\to\infty} z_n = \xi$, then $F(\xi) = \xi$.

<u>Proof:</u> Trivial:

$$\xi = \lim_{n\to\infty} z_n = \lim_{n\to\infty} F(z_{n-1})$$
$$= F(\lim_{n\to\infty} z_{n-1}) = F(\xi)$$

This equality shows that F must be continuous in the vicinity of ξ. Incidentally, from ξ there can of course extend a line of discontinuity of F in the complex plane, i. e. a cut which extends F into more laminae. We will confine ourselves however, to those functions F which are single valued and differentiable (i. e. analytic) at least within a part of the domain of convergence containing ξ.

T. 2: Given $F(z)$ analytic. Then a necessary and sufficient condition for $\lim_{n\to\infty} z_n$ to exist and be equal to ξ, is that $|F'(z)| \leq q < 1$ within the vicinity of ξ (a circle centre ξ.)

Proof:

Necessity: Since $F(z)$ is analytic, we can expand $F(z_o)$ or $F(\xi + \varepsilon)$ in a Taylor series within a circle, centre ξ, as follows:

$$F(z_o) = F(\xi + \varepsilon) = F(\xi) + \varepsilon F'(\xi) + \frac{\varepsilon^2}{2!} F^{(2)}(\xi) + \ldots$$

i.e.
$$z_1 = \xi + \varepsilon F'(\xi) + \frac{\varepsilon^2}{2!} F^{(2)}(\xi) + \ldots$$

From this we see, for $|z_1 - \xi|$ to be smaller than $|z_o - \xi| = |\varepsilon|$, ($\varepsilon$ small) it is necessary for $|F'(\xi)|$ to be smaller than 1.

Sufficiency: We have $F(\xi) = \xi$ and $z_n = F(z_{n-1})$. Then $z_n - \xi = F(z_{n-1}) - F(\xi) = F'(\tilde{z}_{n-1})(z_{n-1} - \xi)$. (Mean Value Theorem. See [16]). Therefore

$$|z_n - \xi| \leq q|z_{n-1} - \xi| \leq q^2|z_{n-2} - \xi| \leq \ldots \leq q^n|z_o - \xi|.$$

As $q^n \underset{\infty}{\overset{n}{\to}} 0$, $|z_n - \xi| \to 0$.

Definition: The order of convergence of an algorithm.

An algorithm of the type (2.1) is said to converge towards a root ξ of $f(z) = 0$ for all initial values $z = z_o$ in a vicinity of ξ of order $k > 0$, when

$$F(z) - \xi = O(|z - \xi|^k), \qquad z \to \xi.$$

[The synbols O and o to be used in this work, are the well-known order symbols, i.e.:

We write $\phi = O(\psi)$; ϕ, ψ functions of z in R if there exists a constant (i.e. a number independent of z) A so that

$$|\phi| \leq A|\psi| \qquad \text{for all } z \text{ in } R;$$

$\phi = O(\psi)$ as $z \to z_o$ if there exists a constant A and a neighbourhood U of z_o so that

$|\phi| \leq A|\psi|$ for all z common to U and R;

and we write $\phi = o(\psi)$ as $z \to z_o$ if for any given $\varepsilon > 0$ there exists a neighbourhood U_ε of z_o so that $|\phi| \leq \varepsilon\, |\psi|$ for all z common to U_ε and R.]

As can be easily seen from the Taylor expansion used above, this condition is satisfied if $F(z)$ has derivatives up to the k-th order in a vicinity of ξ, and the equations

$$F(\xi) = \xi, \quad F'(\xi) = F^{(2)}(\xi) = \dots = F^{(k-1)}(\xi) = 0, \quad k > 0$$
$$\text{and} \quad F^{(k)}(\xi) \neq 0 \quad \text{hold.}$$

(Note further, the smaller $F^{(k)}(\xi)$, the quicker is the k-th order convergence.)

The following propositions follow immediately:

<u>P. 1:</u> Given an algorithm of order $k > 0$, then the algorithm

$$z_n = F_r(z_{n-1}), \qquad n = 0, 1, 2, \dots$$

[where $F_r(z)$ is the "r-fold iterated" function,

i.e. $F_1(z) = F(z)$, $F_2(z) = F(F(z))$, ..., $F_r(z) = F(F_{r-1}(z))$]

has the order k^r.

<u>Proof:</u> Trivial - Since the error made at the j-th approximation (i.e. the deviation of z_j from ξ) is proportional to the k-th power of the error made at the $(j-1)$-th approximation.

Thus, if the first approximation (i.e. $|z_o - \xi|$) is correct to the s-th decimal position, the 2nd. approximation will be correct to the ks-th decimal position ... and the r-th approximation will be correct to the $k^{r-1}s$-th decimal position.

<u>P. 2:</u> If we have the two algorithms $F(z)$ and $G(z)$ with

$$F(z) - \xi = O(|z - \xi|^{k_1}), \qquad z \to \xi$$

and

$$G(z) - \xi = O(|z - \xi|^{k_2}), \qquad z \to \xi, \qquad k_1, k_2 > 0$$

then the algorithms

$$z_n = F(G(z_{n-1})) \quad \text{and} \quad z_n = G(F(z_{n-1}))$$

converge in a vicinity of ξ at least with order $k_1 k_2$.

<u>P. 3:</u> If we have $F(z) = S_k(z)$ as an algorithm of the type (2.1) of order $k > 0$ for finding ξ ($f(\xi) = 0$) then the most general algorithm of the k-th order for this purpose will be of the form

$$F^*(z) = S_k(z) + G(z)$$

where $G(z)$ is a (in a vicinity of ξ singularly defined) function which only has to satisfy the condition

$$G(z) = O(|z - \xi|^k), \qquad z \to \xi. \tag{2.2}$$

<u>P. 4:</u> If $F(z) = S_{k_1}(z)$ is an algorithm which converges towards ξ with order $k_1 > k > 0$, then

$$z_n = S_{k_1}(z_{n-1}) + |f(z_{n-1})|^k \qquad \text{for } f'(\xi) \neq 0$$

$$z_n = S_{k_1}(z_{n-1}) + \left|\frac{f(z_{n-1})}{f'(z_{n-1})}\right|^k \qquad \text{for } f'(\xi) = 0$$

is an algorithm of order k.

<u>Proof:</u> Both $f(z)$ in the case $f'(\xi) \neq 0$, and $\frac{f(z)}{f'(z)}$ have ξ as a simple root. Thus we have

$$|f(z)|^k = O(|z - \xi|^k), \qquad z \to \xi$$

$$\left|\frac{f(z)}{f'(z)}\right|^k = O(|z - \xi|^k), \qquad z \to \xi$$

but not equal to $o(|z - \xi|^k)$, $z \to \xi$. Together with the fact

$$S_{k_1}(z) - \xi = O(|z - \xi|^{k_1}) = o(|z - \xi|^k), \qquad z \to \xi$$

the proposition follows immediately.

It follows from P. 3, that if one already has an algorithm of k-th order e.g. $F(z) = S_k(z)$, infinitely many algorithms of the same order can be obtained by adding a function $G(z)$ satisfying condition (2.2), e.g. $G(z) = (z - \xi)^k$. Since ξ is not known in general, $G(z)$ must be expressed by the function $f(z)$ and its derivatives, e.g.

$$G(z) = [f(z)]^k \quad \text{for } f'(\xi) \neq 0$$

$$G(z) = \left[\frac{f(z)}{f'(z)}\right]^k \quad \text{for } f'(\xi) = 0$$

or more general

$$G(z) = [f(z)]^k \phi(z), \quad f'(\xi) \neq 0$$

$$G(z) = \left[\frac{f(z)}{f'(z)}\right]^k \phi(z), \quad f'(\xi) = 0$$

where $\phi(z)$ stays finite if $z \to \xi$.

Thus:

P. 5: Given an analytic function $f(z)$ with $z = \xi$ as root, $f'(\xi) \neq 0$. Given further an algorithm

$$F(z) = S_k(z), \quad k > 0$$

of the type (2.1) (i.e. an algorithm converging to ξ with k-th order). The most general algorithm of k-th order can then be obtained by putting

$$F^*(z) = S_k(z) + [f(z)]^k \phi(z) \tag{2.3}$$

where $\phi(z)$ is an arbitrary function, finite for $z \to \xi$, and which can still be a function of $f(z)$ and its derivatives. (Note: The unknown quantity ξ does not appear explicitly in this general algorithm. cf. Chap. V later).

Proof: (a) According to the Mean Value Theorem

$$f(z) = (z - \xi) f'(\tilde{z}) \tag{2.4}$$

where $f'(\tilde{z})$ is finite and $\neq 0$ in a vicinity of ξ (given)

From this follows:

$$\begin{aligned} F^*(z) - \xi &= S_k(z) - \xi + [f(z)]^k \phi(z) \\ &= O(|z - \xi|^k) + O(|z - \xi|^k), \quad z \to \xi \\ &= O(|z - \xi|^k), \quad z \to \xi \end{aligned}$$

i.e. the algorithm $F^*(z)$ converges with order $k > 0$.

(b) On the other hand, according to P. 3 the function $F^*(z)$ must be of the type

$$F^*(z) = S_k(z) + G(z)$$

where $$G(z) = O(|z - \xi|^k), \qquad z \to \xi$$

i.e. the quotient $$\frac{|G(z)|}{|z - \xi|^k}$$ must be finite

for $z \to \xi$. Thus we can write

$$G(z) = (z - \xi)^k \psi(z) \tag{2.5}$$

where $\psi(z)$ is finite for $z \to \xi$, but otherwise arbitrary. Since $f'(\tilde{z}) \neq 0$ we can put

$$\phi(z) = \frac{\psi(z)}{[f'(\tilde{z})]^k} \tag{2.6}$$

without any loss of arbitrarity. Again $\phi(z)$ finite for $z \to \xi$. From (2.4), (2.5) and (2.6) we now obtain $G(z)$ in the form

$$G(z) = [f(z)]^k \phi(z).$$

Q.E.D.

<u>Note:</u> If $f'(\xi) = 0$, i.e. if $f(z)$ has ξ as a p-fold root ($f'(\xi) = f^{(2)}(\xi) = \ldots = f^{(p-1)}(\xi) = 0$, $f^{(p)}(\xi) \neq 0$) P. 5 still holds (and the proof is the same) if $f(z)$ is replaced throughout by $\frac{f(z)}{f'(z)}$ or by $f^{(p-1)}(z)$.

Thus P. 4 and 5 show that <u>all</u> (i.e. infinitely many) algorithms of the k-th order for solving the analytic equation $f(z) = 0$ can be obtained if only one such an algorithm is known.

Furthermore by P. 5 an algorithm of the (k+1)-th order can always be found from the general algorithm of the k-th order, by placing restrictions on the arbitrary $\phi(z)$ of equation (2.3). Repeated use will be made of this principle in the following chapters where the construction of such algorithms is discussed.

III.

The Newton-Raphson Method and other algorithms of the second order.

The Newton-Raphson algorithm was derived after the following observations in the real case. (The considerations for the complex case are identical).

If an arbitrary point x_o is chosen "sufficiently near" to the unknown root ξ of $f(x) = o$, a better approximation x_1 of ξ can be found by drawing the ordinate at x_o to cut the curve $y=f(x)$ at $f(x_o)$, and then extend the tangent at $(x_o, f(x_o))$ to cut the x-axis in x_1. Repeat the process. In this way we find a series of successive approximations x_o, x_1, x_2, ... to the root ξ of $f(x) = 0$.

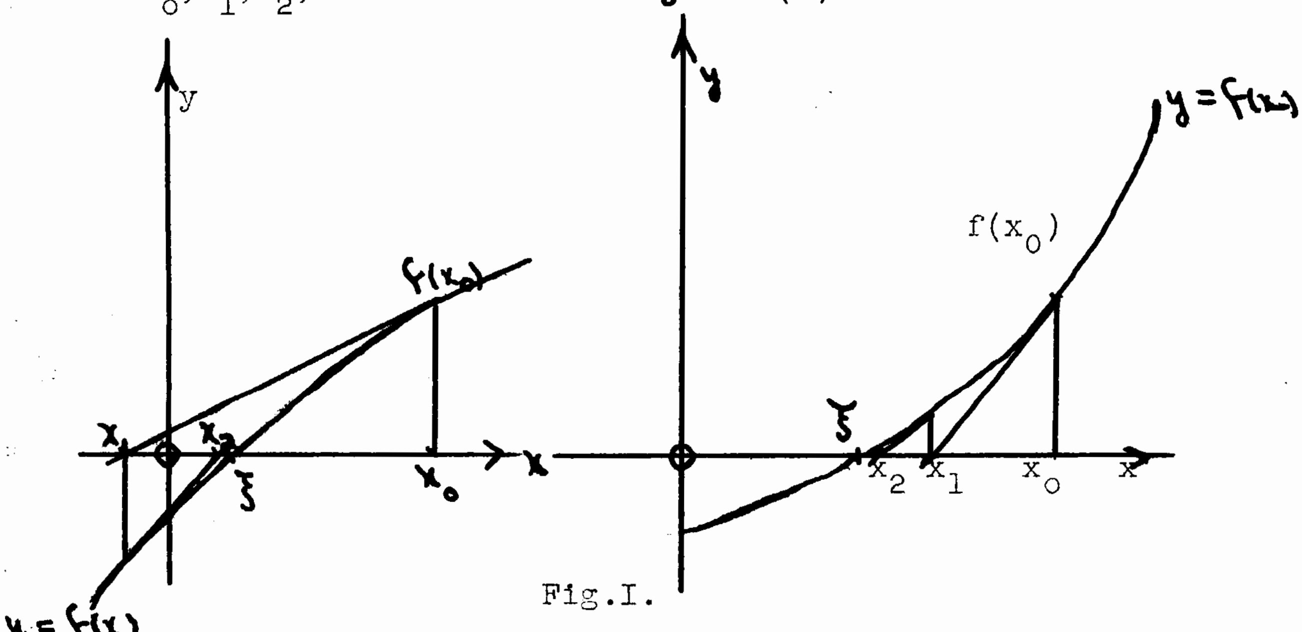

Fig.I.

Analytically, it is obvious that $x_1 = x_0 - \dfrac{f(x_0)}{f^1(x_0)}$

$$x_2 = x_1 - \frac{f(x_1)}{f^1(x_1)}$$

.....

....

Thus we have here an algorithm $x^1 = x - \frac{f(x)}{f^1(x)}$ for finding a root of $f(x) = 0$.

Note: In looking for a simple algorithm of type (2.1) it is by the result of T.1 just natural to choose $F(z)$ as something like

$F(z) = z - cf(z)$. By imposing the further condition of quadratic convergence on this algorithm we have to put $c = \frac{1}{f^1(z)}$, i.e. the Newton-Raphson algorithm.

The similar result is also obtained by answering the question: Find a linear interpolation polynamial in z^1 which will be equal to $f(z)$ for $z^1 = z$ and its first derivative equal to $f^1(z)$ for $z^1 = z$. This polynomial is precisely the Taylor's linear polynomial in $z^1 - z$, i.e.

$$f(z) + (z^1 - z)\, f^1(z).$$

Equate to zero and solve for z^1:

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

This idea was originally due to Newton (1674). In his "Principia", (1687) Prop. 31, Book 1; Newton applied this method first to the equation $x - e \sin x = N$, and next to $e \sinh x - x = N$. The equations arose out of Kepler's problem to find the position of a planet at a given time in an elliptic or hyperbolic orbit of eccentricity e. Raphson however, was the first to express the algorithm in the form

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

or $$z_n = F(z_{n-1}) = z_{n-1} - \frac{f(z_{n-1})}{f^1(z_{n-1})} \qquad (3.1)$$

(See Cajori's "History of Mathematics" p.203).

If ζ is a p-the root of f(z) and if we denote $z-\zeta=\varepsilon$, so that $f=\varepsilon^{p}\psi$

$$f^{1}=\varepsilon^{p-1}\quad(p\psi+\varepsilon\psi^{1})$$

and thus $F(z)=z-\dfrac{\varepsilon\psi}{p\psi+\varepsilon\psi^{1}}$

and $F^{1}(z)=1-\dfrac{\psi}{p\psi+\varepsilon\psi^{1}}-\varepsilon\dfrac{d}{dz}\dfrac{\psi}{p\psi+\varepsilon\psi^{1}}$

Then for $z\cong\zeta$ or $\varepsilon\cong 0$,

$$F^{1}\cong 1-\frac{1}{p}$$

i.e. $|F^{1}|<1$

and $F(\zeta)=\zeta$.

Thus by T.1 and 2 , (3.1) gives an algorithm of the type (2.1).

If $p>1$, $F^{1}(\zeta)\neq 0$, i.e.

$$F(z)-\zeta=0\,(|z-\zeta|)\quad,\quad z\to\zeta\,.$$

Therefore, the algorithm convergences linearly only if f(z) has a multiple root, and the more multiple the root, the larger is $F^{1}(\zeta)$ and subsequently, the slower is the convergence.

If $p=1$, $F^{1}(\zeta)=0$, i.e. $F(z)-\zeta=0\,(|z-\zeta|^{2})$, $z\to\zeta$

$$F^{(2)}(\zeta)\neq 0\quad(\text{in general}).$$

Thus, Newton's algorithm is then of quadratic convergence.

Hence, we can always obtain quadratic convergence in the case of the Newton-Raphson method by simply applying it to $\dfrac{f(z)}{f^{1}(z)}$ instead of f(z).

...

(Since the roots of $\frac{f(z)}{f^1(z)} = 0$ are the same as those of $f(z) = 0$, with the exception that they are all simple).

Then we obtain

<u>Modification I</u>:- of Newton's algorithm, namely:

$$F(z) = z - \frac{f(z)\ f^1(z)}{f^1(z)^2 - f(z).f''(z)} \qquad (3.2)$$

Note: This is a special case of the <u>most general algorithm of the first order</u>, namely

$$z_n = z_{n-1} - \frac{f(z_{n-1})\ \varphi(z_{n-1})}{\varphi(z_{n-1})\ f^1(z_{n-1}) - f(z_{n-1})\ \varphi'(z_{n-1})}$$

$$\text{or} \quad F(z) = z - \frac{1}{\frac{f^1}{f} - \frac{\varphi^1}{\varphi}} \qquad (3.3)$$

This equation (3.3) is obtained by applying the Newton algorithm to $\frac{f(z)}{\varphi(z)} = 0$ instead of $f(z) = 0$, where $\varphi(z)$ does not vanish together with $f(z)$, or at least vanishes with lower order than $f(z)$. Equation (3.2) is obtained from (3.3) by putting $\varphi(z) = f^1(z)$. Incidentally, according to P.5 an equivalent form of (3.3) would be $F(z) = z - \frac{f}{f^1} + f\varphi$ (3.4) where $\varphi(z)$ remains finite for $z \to \zeta$

In future the Newton-Raphson method will consequently always be referred to as an algorithm of second order convegence.

<u>Mod.II:</u> On the other hand, if the degree p of multiplicity of a root is known, then the N-R algorithm can be replaced by

$$F(z) = z - p\ \frac{f(z)}{f^1(z)} \qquad (3.5)$$

since then $F^1(\zeta) = 1 - p \cdot \frac{1}{p} = 0$.

[In the specifie case where $f(z) = (z-\zeta)^p$, i.e. $\psi(z) = 1$, this modifie d algorithm gives for every initial value z_o immediately the correct root since ....

$$z_1 = z_0 - p \cdot \frac{1}{p} (z_0 - \zeta) = \zeta . \quad]$$

Mod.II (a): Another modification which might be used in case of a p-fold root is

$$F(z) = z - \frac{f^{(p-1)}(z)}{f^{(p)}(z)} \qquad \text{since} \quad f^{(p)}(\zeta) \neq 0.$$

Here again $F^1(\zeta) = \dfrac{f^{(p-1)}(\zeta)\, f^{(p+1)}(\zeta)}{[\, f^{(p)}(\zeta)\,]^2} = 0$

We will see later however, that Mod.II is by far the superior of Mod.II (a). As a matter of fact, the convergence speed of Mod.II is $\frac{1}{2}p\,(p+1)$ times that of Mod.II (a).

Before proceeding to the further modifications and improvements of the Newton-Raphson method, we will discuss a few fundamental theorems concerning the convergence of this method as given by (3.1).

T.3: Let f(z) be an analytic function regular within and on a closed contour(rectifiable Jordan curve) C. If $f^1(\zeta) \neq 0$ where ζ is a root within C of $f(z) = 0$ to be obtained by application of the Newton-Raphson algorithm, we have after n applications:

$$|\zeta - z_n| \leq \frac{M}{2m^3} \, |f(z_{n-1})|^2 \quad , \quad n = 1, 2, \ldots .$$

where $|f''(z)| \leq M$,

$|f^1(z)| \geq m > 0$ in the vicinity of ζ, or more precisely, in the domain of convergence of the algorithm.

[In the case where $f^1(\zeta) = 0$, we simply replace f(z) by $\dfrac{f(z)}{f^1(z)}$. Hence, this theorem also holds for Mod.I. See analogous

T.5 for Mod.II.]

<u>Lemma:</u> Let $f(z)$ and $g(z)$ be regular within and on a closed contour C. (In the real case we need $f(x)$ and $g(x)$ to be n-times differentiable within C.). Assume there exist n common roots of $f(z)$ and $g(z)$ within C. (If a root is counted with multiplicity p, it must at least have the multiplicity p for both $f(z)$ and $g(z)$.) Assume further that $g^{(n)}(z) \neq 0$. within C. Then for any $z_o \neq z_j$ in C, there exists a ξ within C such that

$$\frac{f(z_o)}{g(z_o)} = \frac{f^{(n)}(\xi)}{g^{(n)}(\xi)} \qquad z_o \neq z_j \text{ , } j=1 \text{ .... } n.$$

<u>Proof of Lemma:</u> Firstly, z_o is not a root of $g(z)$, as if it were the case $g(z)$ would have $n+1$ roots in C, a case which is excluded by the assumption that $g^{(n)}(z) \neq 0$ in C.

$$\text{Let } \lambda = \frac{f(z_o)}{g(z_o)}$$

Consider $F(z) = f(z) - \lambda g(z)$

Evidently $F(z)$ has the $n+1$ roots z_o, z_j ($j = 1$ .... n)

$\therefore$ There exists a ξ in C, such that

$$F^{(n)}(\xi) = f^{(n)}(\xi) - \lambda g^{(n)}(\xi) = 0.$$

$$\therefore \frac{f^{(n)}(\xi)}{g^{(n)}(\xi)} = \lambda = \frac{f(z_o)}{g(z_o)}$$

<u>The Remainder term in general interpolation:</u>

Let $f(z)$, $g(z)$ be regular within and on C (a closed contour). Let $P_n(z)$ be an interpolation polynomial of degree $n-1$ for $f(z)$ with n interpolation points z_1 .... z_n within C.

...

i.e. $f(z_j) = P_n(z_j)$ $(j = 1 \dots n.)$

Let $g(z_j) = 0$ $(j=1 \dots n)$

and $g^{(n)}(z) \neq 0$ for all z in C .

Replace $f(z)$ by $f(z) - P_n(z)$ in the lemma above.

Then for all $z \neq z_j$ $(j=1 \dots n)$, z in C , there exists a ξ in C depending on z such that $\dfrac{f(z) - P_n(z)}{g(z)} = \dfrac{f^{(n)}(\xi)}{g^{(n)}(\xi)}$

Put $g(z) = \prod_{j=1}^{n} (z-z_j)$ and then we have

$$f(z) - P_n(z) = \frac{f^{(n)}(\xi)}{n!} \prod_{j=1}^{n} (z-z_j) \qquad (3.6)$$

When $z_1=z_2 = \dots = z_n=a$, (3.6) is the remainder term of the Taylor series. When the interpolation points are all distinct (3.6) becomes the remainder term of the n-point Lagrangian and also Newton interpolation formula.

<u>Proof of T.3:</u> We have the Newton algorithm (interpolation polynamial for $f(z)$, which will be equal to $f(z_o)$ for $z=z_o$ and its first derivative equal to $f^1(z_o)$ for $z=z_o$ —— see introductory remarks to this chapter:)

$$P_2(z) = f(z_o) + (z-z_o) \; f^1(z_o)$$

Substitute in (3.6) : $f(z) - P_2(z) = \dfrac{f^{11}(\eta)}{2} (z-z_o)^2$

(η an intermediate point lying within C).

Let z_o be the initial approximation to the root ξ of $f(z) = 0$, and put $z= \xi$. Then we obtain:

$$- P_2(\xi) = \frac{f^{11}(\eta)}{2} (\xi - z_o)^2 \quad \dots\dots$$

$$f(z_o) + (\xi - z_o)\, f^{1}(z_o) + \frac{f^{11}(\eta)}{2} (\xi - z_o)^2 = 0$$

$$\frac{f(z_o)}{f^{1}(z_o)} - z_o + \xi = - \frac{f^{11}(\eta)}{2f^{1}(z_o)} (\xi - z_o)^2$$

$$\xi - z_1 = - \frac{f^{11}(\eta)}{2f^{1}(z_o)} (\xi - z_o)^2$$

From this we see, that generally speaking, the approximation will be improved quadratically at each application of the Newton algorithm —— a fact already known since the Newton algorithm is of order 2 for the type of $f(z)$ under consideration.

From the mean value theorem we have:

$$f^{1}(\eta^{1})\ (\xi - z_o) = f(\xi) - f(z_o)$$

$$\therefore\ \xi - z_o = - \frac{f(z_o)}{f^{1}(\eta^{1})}$$

$$\therefore\ \xi - z_1 = - \frac{f^{11}(\eta)\ f(z_o)^2}{2f^{1}(z_o)\ f^{1}(\eta^{1})^2}$$

From this we can immediately obtain an upper bound for the distance of z_1 to ξ (i.e. an error estimate) if we have an upper bound for $| f^{11}(z) |$ and a lower bound for $| f^{1}(z) |$, say

$$|f^{11}(z)| \leq M \ , \quad | f^{1}(z) | \geq m > 0$$

Then $$|\xi - z_1| \leq \frac{M}{2m^3} | f(z_o) |^2 = K .$$

and for further approximations

$$|\xi - z_n| \leq \frac{M}{2m^3} | f(z_{n-1}) |^2 \ , \ n=2 \ \ldots.$$

....

From this inequality it is evident that the Newton algorithm can be applied with great benefit to functions for which m is relatively large, and M relatively small in the vicinity of the roots.

(T.3 was proved by Ostrowski [15] in the real case by considering the inverse of $y=f(x)$).

Corollary: If $f(z)$ is an analytic function, regular within and on a closed contour C, and if ξ is a p-fold root within C of $f(z) = 0$, which is to be obtained by application of Mod.II(a) of the N-R method, we have after n applications

$$|\xi - z_n| \leq \frac{1}{2} \frac{M_{p+1}}{m_p} |\xi - z_{n-1}|^2 \qquad (3.7)$$

where $M_{p+1} = \sup f^{(p+1)}(z)$

$$|\xi - z| < |\xi - z_{n-1}|$$

$$m_p = \inf f^{(p)}(z)$$

$$|\xi - z| < |\xi - z_{n-1}|$$

Proof: Follows immediately. In this case we have

$$P_2(z) = f^{(p-1)}(z_{n-1}) + (z-z_{n-1}) f^{(p)}(z_{n-1})$$

as interpolation polynomial for $f^{(p-1)}(z)$ which will be equal to $f^{(p-1)}(z_{n-1})$ for $z = z_{n-1}$ and its first derivative equal to $f^{(p)}(z_{n-1})$ for $z = z_{n-1}$

Substitute in (3.6):

$$f^{(p-1)}(z) - P_2(z) = \frac{f^{(p+1)}(\eta)}{2} (z-z_{n-1})^2$$

η an intermediate point lying within circle radius $|z-z_{n-1}|$ within C. $\eta \to \xi$ as

$$z_{n-1} \to \xi .$$

Then $$-P_2(\xi) = \frac{f^{(p+1)}(\eta)}{2} (\xi - z_{n-1})^2$$

i.e.
$$f^{(p-1)}(z_{n-1}) + (\xi - z_{n-1}) f^{(p)}(z_{n-1}) + \frac{f^{(p+1)}(\eta)}{2f^{(p)}(z_{n-1})} \times (\xi - z_{n-1})^2 = 0$$

$$\frac{f^{(p-1)}(z_{n-1})}{f^{(p)}(z_{n-1})} - z_{n-1} + \xi = - \frac{f^{(p+1)}(\eta)}{2f^{(p)}(z_{n-1})} (\xi - z_{n-1})^2$$

$$\xi - z_n = - \frac{f^{(p+1)}(\eta)}{2f^{(p)}(z_{n-1})} (\xi - z_{n-1})^2$$

$$\therefore \frac{\xi - z_n}{(\xi - z_{n-1})^2} \to - \frac{1}{2} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$$

In the discussion above we have assumed that a root ξ exists in the neighbourhood of z_o. In practice it very often happens that we must proceed with our computation before this is known. However, we can usually tell after the first steps, whether there is a root in the considered neighbourhood. We will now deal with this in more detail. We have from T.2 that a necessary and sufficient condition for convergence in the case of the N-R algorithm would be that

$$|F^1(z)| = \left| \frac{d}{dz} \left(z - \frac{f(z)}{f^1(z)}\right) \right| = \left| \frac{f(z)\, f^{11}(z)}{f^1(z)^2} \right| < 1$$

for z in the vicinity of the root ξ.

i.e. $| f(z) f^{11}(z) | < | f^1(z) |^2$ ....

We have seen that this condition is always satisfied for functions of the type $f(z) = (z-\xi)^p \psi(z)$. It is however quite another question whether the algorithm will converge for the arbitrary initial value z_o. We will thus try to obtain a condition for convergence in terms of z_o.

T.4: Given f(z) analytic, regular within and on circle C :

$$| z-z_1 | \leq | h_o |$$

where $h_o = -\dfrac{f(z_o)}{f^1(z_o)}$, $f(z_o) f^1(z_o) \neq 0$,

$$z_n = z_{n-1} - \frac{f(z_{n-1})}{f^1(z_{n-1})} \qquad n = 1 \ldots.$$

i.e. $z_1 = z_o + h_o$

and $\max_C | f^{11}(z) | = M$,

$$2 | h_o | M \leq | f^1(z_o) | \tag{3.8}$$

Then all z_n lie in C , and $z_n \xrightarrow[\infty]{n} \xi$

where ξ is the only root in C, and ξ is a simple root unless it lies on C.

Further: $\dfrac{| z_{n+1}-z_n |}{|z_n-z_{n-1}|^2} \leq \dfrac{M}{2|f^1(z_n)|}$ (n=1,2 ....)

$$| \xi - z_{n+1} | \leq \frac{M}{2|f^1(z_n)|} | z_n-z_{n-1} |^2 \qquad (n=1,2 \ldots.)$$

...

[Note: In this theorem no assumptions are made as to the existence of a root],

Proof: Since

$$f^1(z_1) - f^1(z_o) = \int_{z_o}^{z_1} f^{11}(z)\, dz$$

(according to Cauchy for all curves of integration lying in C, connecting z_o and z_1)

i.e. $| f^1(z_1) - f^1(z_o) | \leq | z_1 - z_o | M = |h_o| M \leq \frac{|f^1(z_o)|}{2}$ (3.9)

$$\therefore \quad |f^1(z_1)| \geq | f^1(z_o)| - | f^1(z_1) - f^1(z_o)| \geq |f^1(z_o)| - \frac{|f^1(z_o)|}{2}$$

$$\therefore \quad |f^1(z_1)| \geq \frac{|f^1(z_o)|}{2} \qquad (3.10)$$

Integration by parts gives

$$\int_{z_o}^{z_1} (z_1 - z) f^{11}(z)\, dz = -(z_1 - z_o) f^1(z_o) + f(z_1) - f(z_o)$$

$$= - h_o f^1(z_o) - f(z_o) + f(z_1) = f(z_1)$$

Substitute in this $z = z_o + t h_o$.

i.e. $z_1 - z = z_1 - z_o - t h_o = h_o - t h_o = h_o(1-t)$,

$dz = h_o\, dt$.

Then $f(z_1) = h_o^2 \int_o^1 (1-t) f^{11}(z_o + t h_o) dt$

This will at least hold for all paths of integration lying within the circle $|t| \leq 2$.

[Since the transformed domain of C (described by $|z - z_1| \leq h_o$) in the t-plane will lie within $|t| \leq 2$, as $$th_o = z - z_o$$

$$= z - z_1 - \frac{f(z_o)}{f^1(z_o)}$$

$$\therefore \; |t\,h_o| \leq |z-z_1| + \left| \frac{f(z_o)}{f^1(z_o)} \right|$$

$\therefore$ For z to be within C

$$|t\,h_o| \leq |h_o| + \left| \frac{f(z_o)}{f^1(z_o)} \right|$$

$$\therefore \; |t| \leq 2 \qquad]$$

Thus $$|f(z_1) \leq |h_o|^2 \int_o^1 (1-t)\; |f^{11}(z_o + t\,h_o)|\; dt$$

$$\leq M\; |h_o|^2 \int_o^1 (1-t)\; dt = \frac{|h_o|^2\; M}{2} \qquad (3.11)$$

Since $\max\limits_{|t|\leq 2} |f^{11}(z_o + t\,h_o)\,) = M$.

If we put $h_1 = -\dfrac{f(z_1)}{f^1(z_1)}$, $z_2 = z_1 + h_1$

we have by (3.10) and (3.11)

$$|h_1| = \frac{|f(z_1)|}{|f^1(z_1)|} \leq \frac{|h_o|^2\; M}{|f^1(z_o)|} \qquad (3.12)$$

$$\frac{M|h_1|}{|f^1(z_1)|} \leq \frac{|h_o|^2\; M^2}{|f^1(z_o)|\; |f^1(z_1)|} \leq \frac{|h_o|^2\; M^2}{|f^1(z_o)|\; \frac{1}{2}\; |f^1(z_o)|}$$

i.e. $$\frac{2\;M\;|h_1|}{|f^1(z_1)|} \leq \left(\frac{2\,|h_o|\;M}{|f^1(z_o)|} \right)^2 = \frac{4\;|h_o|^2\;M^2}{|f^1(z_o)|^2} \leq 1$$ by (3.9)

...

$$\therefore \quad 2|h_1| M \leq |f^1(z_1)| \qquad (3.13)$$

Apply (3.12) : $\left|\frac{h_1}{h_o}\right| \leq \frac{1}{2} \left(\frac{2 \; |h_o| \; M}{|f^1(z_o)|} \right) \leq \frac{1}{2}$

$$|h_1| \leq \frac{1}{2} |h_o|$$

and from this we see that the point z_2 will not get beyond the distance $\frac{1}{2} |h_o|$ from z_1, and will remain in C.

From (3.13) it is evident that this will remain true if we replace z_o and h_o by z_1 and h_1 respectively, and thus

$$|h_n| \leq \frac{1}{2} |h_{n-1}| \quad , n=1 \; \ldots .$$

where $h_{n-1} = z_n - z_{n-1}$

and hence all z_n lie in C.

We therefore have a sequence $\{D_n\}$ of rested domains (circles) in the complex plane, with the radius of D_n at most equal to one-half the radius of D_{n-1} . This monotonic sequence is bounded from below, and thus has a limit ζ .

i.e. $z_n \xrightarrow[\infty]{n} \zeta$

ζ is a root of $f(z)$, since

$$z_n = z_{n-1} - \frac{f(z_{n-1})}{f^1(z_{n-1})}$$

i.e. $z_n \cdot f^1(z_{n-1}) = z_{n-1} \cdot f^1(z_{n-1}) - f(z_{n-1})$

let $n \longrightarrow \infty$. $\zeta \, f^1(\zeta) = \zeta \, . \, f^1(\zeta) - f(\zeta)$ ...

$$\therefore \quad f(\xi) = 0.$$

This root ξ is simple <u>within</u> C:

We have for all z in (and on C) - see (3.9)

$$|f^1(z) - f^1(z_o)| \leq |z-z_o| M$$

For z within C the distance between z and z_o is smaller than the diameter of C :.i.e. $2 |h_o|$

$$\therefore \quad |f^1(z) - f^1(z_o)| < 2 |h_o| M \leq |f^1(z_o)|$$

and this implies that $f^1(z) \neq 0$ within C $\therefore$ ξ is a simple root unless it lies on C. By (3.12) and the validity of (3.13) we have

$$|h_n| \leq \frac{M |h_{n-1}|^2}{2|f^1(z_n)|} \qquad , n=1,2 \ldots.$$

and this is equal to:

$$\frac{|z_{n+1} - z_n|}{|z_n-z_{n-1}|^2} \leq \frac{M}{2|f^1(z_n)|} \qquad , n=1,2 \ldots.$$

Further, since ξ lies in a circle centre z_{n+1} and radius

$$|h_n| \quad , \quad [\ |h_n| = |z_{n+1}-z_n| \leq \frac{1}{2^n} |h_o|\ ]$$

$$\therefore \quad |\xi - z_{n+1}| \leq |h_n| \leq \frac{M |h_{n-1}|^2}{2|f^1(z_n)|}$$

$$= \frac{M}{2|f^1(z_n)|} |z_n-z_{n-1}|^2$$

$$n=1,2 \ldots.$$

We therefore know that a root exists if we begin computing by the Newton-Raphson method, and if at the (n+1) - th step, the

....

inequality $2\ |h_n|\ M \leq |f^1(z_n)|$ holds.

It is essential that $f^1(z_o)$ is not zero. It may also be quite hazardous if $f^1(\zeta) = 0$, for if we are sufficiently close to ζ, $|f^1(z_o)|$ will be very small. In this case it will again be advisable to replace $f(z)$ by $\dfrac{f(z)}{f^1(z)}$.

Mod.II : $$z_n = z_{n-1} - p\ \frac{f(z_{n-1})}{f^1(z_{n-1})} \qquad n=1,2\ \ldots.$$

<u>T.5:</u> (Analog to T.3 for Mod.II)

Let f(z) be an analytic function regular within and on a closed contour (rectifiable Jordan curve) C. Let ζ be a root within C of exact multiplicity p to be obtained by Mod.II of the Newton algorithm. Then we have after n applications:

$$|\zeta - z_n| \leq \frac{M_{p+1}}{p(p+1)m_p}\ |\zeta - z_{n-1}|^2$$

where $$\sup_{|\zeta - z| < |\zeta - z_{n-1}|} |f^{(p+1)}(z)| = M_{p+1}$$

$$\inf_{|\zeta - z| < |\zeta - z_{n-1}|} |f^{(p)}(z)| = m_p$$

Note: For $p = 1$ we have the result of T.3.

[The proof of T.3 does not hold for Mod.II , since in this case the $P_2(z)$ cannot be considered as an interpolation polynomial for $f(z)$, which will be equal to $f(z_n)$ for $z=z_n$ and its first derivative equal to $f^1(z_n)$ for $z=z_n$. Incidentally, T.3 can also be proved in a similar way as below.]

....

Proof: Thus we have

$$\xi - z_n = \xi - z_{n-1} + p \frac{f(z_{n-1})}{f^1(z_{n-1})}$$

i.e. $$(\xi - z_n)\, f^1(z_{n-1}) = p\, f(z_{n-1}) + (\xi - z_{n-1})\, f^1(z_{n-1})$$

$$= G(z_{n-1}) \qquad (3.14)$$

We define $G(z)$ as:

$$G(z) = p\, f(z) - (z-\xi)\, f^1(z)$$

Then evidently $G^{(n)}(z) = (p-n)\, f^{(n)}(z) - (z-\xi)\, f^{(n+1)}(z)$

and $$G^{(n)}(\xi) = 0 \qquad (n=o,1,\ \dots .\ p) \qquad (3.15)$$

If we apply Taylor's expansion to $f^1(z)$ we obtain:

$$f^1(z) = \sum_{n=o}^{p-2} \frac{(z-\xi)^n}{n!} f^{(n+1)}(\xi) + R_{p-1} = R_{p-1}$$

(since ξ is a root of p-th order.)

Where $$R_{p-1} = \frac{(z-\xi)^{p-1}}{2\pi i} \int_{C^1} \frac{f^1(t)}{(t-\xi)^p}\, dt$$

(where C^1 is a circle centre ξ inside C such that f(z) is regular in and on C^1.)

$$\therefore\ f^1(z_n) = \frac{(z_n-\xi)^{p-1}}{2\pi i} \int_{C^1} \frac{f^1(t)}{(t-\xi)^p}\, dt = \frac{(z_n-\xi)^{p-1}}{(p-1)!} f^{(p)}(\xi)$$

(Cauchy) (3.16)

From (3.15) it follows further that

$$G(z) = \sum_{n=o}^{p} \frac{(z-\xi)^n}{n!} G^{(n)}(\xi) + R_{p+1} = R_{p+1}$$

$$\therefore \quad G(z) = \frac{(z-\xi)^{p+1}}{2\pi i}\int_{C^1} \frac{G(t)}{(t-\xi)^{p+2}}\,dt$$

$$= \frac{(z-\xi)^{p+1}}{(p+1)!} \qquad G^{(p+1)}(\xi) \qquad \text{(Cauchy)}$$

$$\therefore \quad G(z) = \frac{(z-\xi)^{p+1}}{(p+1)!} \quad [-f^{(p+1)}(\xi)]$$

From (3.14) and (3.16) we obtain

$$\xi - z_{n+1} = \frac{G(z_n)}{f^1(z_n)} = -\frac{(z_n-\xi)^2 f^{(p+1)}(\xi)}{p(p+1)\, f^{(p)}(\xi)}$$

In general $f^{(p+1)}(\xi)$ and $f^{(p)}(\xi)$ is not explicitly known. Thus if we denote the l.u.b of $|f^{(p+1)}(z)|$ in the vicinity of ξ with $M_p{}^1{+1}$ and the g.l.b. of $|f^{(p)}(z)|$ in the vicinity of ξ with $m_p{}^1$, we have

$$|\xi - z_n| \leq \frac{M^1{}_{p+1}}{p(p+1)m^1{}_p} \qquad |\xi - z_{n-1}|^2 \qquad (3.17)$$

As in T.3 we can write for

$$\xi - z_{n-1} = \frac{f(\xi) - f(z_{n-1})}{f^1(\eta)} \qquad \eta \text{ an intermediate point}$$

$$= -\frac{f(z_{n-1})}{f^1(\eta)}$$

$$\therefore \quad |\xi - z_n| \leq \frac{M^1{}_{p+1}}{p(p+1)m^1{}_p\,(m^1{}_1)^2} \qquad |f(z_n-1)|^2$$

<u>Note:</u> The more multiple the root, the quicker the convergence— a fact quite contrary to what we have observed in the application of the ordinary N-R algorithm. (See remark in the beginning of this chapter.)

In many cases it may be difficult to obtain m^1_p . A good estimate of this quantity can be obtained however, if M^1_{p+1} and the p-th derivative at one point, say z_o are known. Then an estimate for m^1_p is derived from

$$| f^{(p)}(z) | \geq | f^{(p)}(z_o) | - M^1_{p+1} | z-z_o | .$$

Corollary:

(a) From T.3:

If in the equation

$$f(z) = (z-\xi_1)(z-\xi_2) \ldots (z-\xi_s) = 0$$

we have

$$|\xi_1-z_o| < |\xi_2-z_o| \leq |\xi_3-z_o| \leq \cdots \leq |\xi_s-z_o|$$

$$\xi_1 \neq \xi_j \quad , \quad j = 2 \ldots\ldots s$$

then the Newton-Raphson algorithm starting with z_o is convergent to the value ξ_1 (i.e. ξ_1 is a so-called attractive fixed point of

$$F = z_n = z_{n-1} - \frac{f(z_{n-1})}{f^1(z_{n-1})} \quad , \quad n = 1 \ldots\ldots)$$

if $|\xi_1 - z_o| < \frac{2m}{M}$, $m = \inf | f^1(z) |$

$$|\xi_1-z| < |\xi_1-z_o|$$

$$M = \sup | f^{11}(z) |$$

$$|\xi_1-z| < |\xi_1-z_o|$$

Proof: We denote

$$\xi_1-z_n = a^1_n \quad , \quad \xi_2-z_n = a^2_n \quad , \ldots\ldots , \quad \xi_s-z_n = a^s_n$$

i.e. we have $|a_o^1| < |a_o^2| \leq |a_o^3| \leq \ldots\ldots \leq |a_o^s|$

and from T.3 it follows that

$$|a_n^1| \le \frac{M}{2m} |a_{n-1}^1|^2 \le \left(\frac{M}{2m}\right)^3 |a_{n-2}^1|^4$$

$$\le \left(\frac{M}{2m}\right)^7 |a_{n-3}^1|^8$$

$$\le \ldots\ldots$$

$$\le \left(\frac{M}{2m}\right)^{2^n-1} |a_o^1|^{2^n} = \frac{2m}{M}\left(\frac{M}{2m}|a_o^1|\right)^{2^n}$$

Then obviously $|a_n^1| = |\xi_1 - z_n| \xrightarrow[\infty]{n} 0$

if $\frac{M}{2m} |a_o^1| = |\xi_1 - z_o| \frac{M}{2m} < 1$

i.e. $|\xi_1 - z_o| < \frac{2m}{M}$

(It is interesting to note here that in the case where m is relatively large, and M relatively small, i.e. where we have a high convergence speed - see remark T.3 — we have the extra benefit of a lesser restriction of the choice of the initial z_o.)

(b) From T.5:

If in the equation

$$f(z) = (z - \xi_1)^p (z - \xi_2) \ldots\ldots (z - \xi_k) = 0$$

we have

$$|\xi_1 - z_o| < |\xi_2 - z_o| \le |\xi_3 - z_o| \le \ldots \le |\xi_k - z_o|$$

$$\xi_1 \ne \xi_j \quad , \quad j = 2 \ldots k$$

then Mod.II of the Newton algorithm will have ξ_1 as an attractive fixed point if

$$|\xi_1 - z_o| < \frac{p(p+1)m_p}{M_{p+1}} \quad , \quad m_p = \text{inf.} |f^{(p)}(z)|)$$

$$|\xi_1 - z| < |\xi_1 - z_o|$$

$$M_{p+1} = \sup |f^{(p+1)}(z)|$$

$$|\xi_1 - z| < |\xi_1 - z_o|$$

(Again a lesser restriction in case of high convergence speed.)

Proof: We have $|a_o^1| < |a_o^2| \leq |a_o^3| \leq \ldots \leq |a_o^k|$ and

$$|a_n^1| \leq \frac{M_{p+1}}{p(p+1)m_p} |a_{n-1}^1|^2 \qquad (\text{ from T.5.})$$

$$\leq \frac{p(p+1)m_p}{M_{p+1}} \left(\frac{M_{p+1}}{p(p+1)m_p} \; |a_o^1| \right)^{2^n}$$

Therefore, for convergence we must have

$$|a_o^1| = |\xi_1 - z_o| < \frac{p(p+1)m_p}{M_{p+1}}$$

(c) Comparing equation (3.17) to(3.7) we observe that Mod.II is by far superior to Mod.II (a), since the speed of convergence in the first case is $\frac{1}{2}$ p(p+1) times that of Mod.II (a). It is thus doubtful whether Mod.II (a) will be of any notable practical significance.

T.6: If in the quadratic equation

$$f(z) = (z - \xi) \; (z - \eta) = 0$$

we have $|\xi - z_o| < |\eta - z_o|$ then the N-R algorithm starting with z_o is convergent to the value ξ, i.e. ξ is an attractive fixed point of $F(z) = z - f(z)/f^1(z)$ (and η is a repulsive fixed point.) The same is true if $\xi = \eta$. On the other hand, if we have $|\xi - z_o| = |\eta - z_o|$, $\xi \neq \eta$ the N-R algorithm starting with z_o is divergent. (i.e. both ξ and η are repulsive fixed points of F.)

Proof: We have by definition of F

$$z_{n+1} = z_n - \frac{(z_n - \xi)(z_n - \eta)}{2z_n - \xi - \eta}$$

$$= z_n + \frac{a_n b_n}{a_n + b_n}$$

where $a_n = \xi - z_n$, $b_n = \eta - z_n$, $a_o = a$, $b_o = b$.

$$\therefore \quad a_{n+1} = a_n - \frac{a_n b_n}{a_n + b_n} = \frac{a_n^2}{a_n + b_n}$$

similarly

$$b_{n+1} = \frac{b_n^2}{a_n + b_n} \qquad (3.18)$$

Claim: For $\xi \neq \eta$:

$$a_n = \frac{a^{2^n}(a-b)}{a^{2^n} - b^{2^n}}$$

$$b_n = \frac{b^{2^n}(a-b)}{a^{2^n} - b^{2^n}} \qquad (3.19)$$

and for $\xi = \eta$, i.e. $a = b$:

$$a_n = b_n = \frac{a}{2^n} \qquad (3.20)$$

We will first consider the case $\xi \neq \eta$: Obviously (3.19) is true for $n = o$. Suppose (3.19) is true for n. Then we have from (3.18):

$$a_{n+1} = \frac{\dfrac{a^{2^{n+1}}(a-b)^2}{(a^{2^n} - b^{2^n})^2}}{\dfrac{(a-b)(a^{2^n} + b^{2^n})}{a^{2^n} - b^{2^n}}} = \frac{a^{2^{n+1}}(a-b)}{a^{2^{n+1}} - b^{2^{n+1}}}$$

Claim (3.19) now follows immediately by induction .
For $\zeta = \eta$, $a_n = b_n$ we have from (3.18)

$$a_{n+1} = b_{n+1} = \frac{a_n}{2}$$

i.e. claim (3.20).

Now, in the case $\zeta = \eta$, the assertion of our theorem follows immediately from (3.20) , since then we have

$$a_n = \zeta - z_n \xrightarrow[\infty]{n} 0$$

For $\zeta \neq \eta$ we have under the hypothesis

$$| \zeta - z_o | < | \eta - z_o | \text{ , i.e. } | a | < | b |$$

and from (3.19) that $a_n \approx (b-a)\left(\frac{a}{b}\right)^{2^n} \xrightarrow[\infty]{n} 0$

On the other hand ; if $|a| = |b|$, $a \neq b$ we have from (3.19)

$$\left| \frac{a - b}{a^{2^n} - b^{2^n}} \right| = \frac{|a_n|}{|a|^{2^n}} = \frac{|b_n|}{|b|^{2^n}}$$

i.e. $| a_n | = | b_n |$

Thus, in case of convergence both a_n and b_n must tend to zero. It follows from (3.19) however, that

$$a_n - b_n = a - b$$

i.e. if a_n and b_n were both convergent to zero, we would have $a = b$ which is contrary to the hypothesis.

We shall now discuss to what extent , in the case of convergence, the sufficient conditions of T.4 are satisfied. Keeping the notation above, put $p = \frac{b}{a}$, i.e. $p = 1$ or $| p | \neq 1$.

Then we have from above

$$h_n = -\frac{f(z_n)}{f^1(z_n)} = z_{n+1} - z_n = \frac{a_n b_n}{a_n + b_n}$$

$$= \frac{a^{2^n} b^{2^n} (a-b)}{(a^{2^{n+1}} - b^{2^{n+1}})}$$

for $|p| \neq 1$,

or substituting $b = pa$ we have

$$h_n = a\, p^{2^n} \frac{p-1}{p^{2^{n+1}} - 1} \qquad (\,|p| \neq 1\,)$$

and $$h_n = \frac{a}{2^{n+1}} \qquad (\,p = 1\,)$$

Further

$$f^1(z_n) = 2z_n - \xi - \eta = -(a_n + b_n)$$

$$= -(a^{2^n} + b^{2^n}) \frac{a-b}{a^{2^n} - b^{2^n}}$$

$$= -a\,(p-1) \frac{p^{2^n} + 1}{p^{2^n} - 1} \qquad (\,|p| \neq 1)$$

and $$f^1(z_n) = -\frac{a}{2^{n-1}} \qquad (\,p = 1\,)$$

In this case the number M of T.4 is 2 , and

$$\frac{2\,M\,h_n}{f^1(z_n)} = -4 \frac{p^{2^n} (\,p^{2^n} - 1)}{(p^{2^{n+1}} - 1)\,(p^{2^n} + 1)}$$

$$= -4 \frac{p^{2^n}}{(p^{2^n} + 1)^2}$$

$$\therefore \quad \frac{2Mh_n}{f^1(z_n)} = \frac{-4}{(p^{2^{n-1}} + p^{-2^{n-1}})^2} \qquad (|p| \neq 1) \qquad (3.21)$$

$$\frac{2\,M\,h_n}{f^1(z_n)} = -1 \qquad (p = 1)$$

Therefore, for $p = 1$ we have for every n the limiting case $\left|\frac{2\,Mh_n}{f^1(z_n)}\right| = 1$

i.e. convergence.

For $p \neq 1$:

From (3.21) it follows that the modulus of the left-hand expression tends to zero as $n \longrightarrow \infty$. Therefore, the condition of T.4 are satisfied from a certain n onwards.

On the other hand, choosing p conveniently , we can insure that the conditions of T.4 do not hold for $n = 0,1,\ldots,\ N$ where N can be chosen as great as we like. Indeed, if we take

$p = re^{i\alpha}$ we have

$$|p^{2^{n-1}} + p^{-2^{n-1}}|^2 = |(r^{2^{n-1}} + r^{-2^{n-1}}) \operatorname{Cos} 2^{n-1}\alpha$$

$$+ i\,(r^{2^{n-1}} - r^{-2^{n-1}}) \sin\ 2^{n-1}\alpha\ |^2$$

$$= r^{2^n} + r^{-2^n} + 2 \operatorname{Cos} 2^n\alpha$$

Put $\alpha = \frac{\pi}{2^N}$ then we obtain for $n = 0,1,\ \ldots\ N$

$$|p^{2^{n-1}} + p^{-2^{n-1}}|^2 \leq r^{2^n} + r^{-2^n} + 2\,.\operatorname{Cos} \frac{\pi}{2^N}$$

Thus from (3.21)

$$\left|\frac{2\,Mh_n}{f^1(z_n)}\right| \geq \frac{4}{r^{2^n} + r^{-2^n} + 2\,\text{Cos}\,\frac{\pi}{2^N}}$$

The right-hand side of this inequality can be made >1 by taking $r=|p|$ sufficiently near to 1 (but not =1 of course) for then the inequality

$$r^{2^n} + r^{-2^n} + 2\text{Cos}\,\frac{\pi}{2^N} < 4 \;, n= 0,1,\ldots\ldots,N \text{ holds}$$

Note: The modulus of (3.21) can certainly not be equal to 1 for two consecutive values of n:

This follows immediately from the relation

$$q^2 + (1/q^2) = (q + 1/q)^2 - 2,$$

for if both $|q^2 + 1/q^2|$ and $|q + 1/q|$ have the value 2 , this is possible iff we have

$$q^2 + 1/q^2 = 2 \;,\; q^2 = 1 \;,\; q = \pm 1.$$

But in (3.21) $|p| \neq 1$. Contadiction.Claim follows.

Thus, if for a value of n the expression $\left|\frac{2Mh_n}{f^1(z_n)}\right|$ is equal to 1 ,this expression becomes <1 for all greater n , unless our quadratic polynomial has a double root.

Modification III: of the Newton algorithm.

It was suggested that in the N-R algorithm the denominator $f^1(z_n)$ can be replaced by $f^1(z_k)$ as soon as z_k is sufficiently near to ξ .Obviously in this case we will only have linear convergence, and not quadratic convergence characteristic to the N-R method.In the table below, the function $f(x)=x^3 - 2x - 5=0$ is considered.In column I the three values x_1, x_2, x_3 obtained by the N-R formula are given ,whilst in column II the six values x_1 , ... , x_6 obtained by using Mod.III (i.e. replacing $f^1(x_n)$ by $f^1(x_o)$) are given. Comparing the values obtained with the value of ξ, we see that in column II at each step the error is only about 1/10 – th of the preceding error.

$x^3 - 2x - 5 = 0$, $\xi = 2.094\ 551\ 481\ 542\ 326\ 591\ 5$, $x_o = 2$

I	II
$x_1 = 2.1$	$x_1 = 2.1$
$x_2 = 2.094\ 568\ 1$	$x_2 = 2.0939$
$x_3 = 2.094\ 551\ 481\ 72$	$x_3 = 2.094\ 627$
	$x_4 = 2.094\ 542\ 7$
	$x_5 = 2.094\ 552\ 5$
	$x_6 = 2.094\ 551\ 363$

It is therefore doubtful whether this modification will be of great practical significance.

Mod. IV: It may be of some advantage however to compute $f^1(z_n)$ not at every step, but only at every second step , i.e. we have

$$\left.\begin{aligned} z_{n+1} &= z_n - \frac{f(z_n)}{f^1(z_n)} \\ z_{n+2} &= z_{n+1} - \frac{f(z_{n+1})}{f^1(z_n)} \end{aligned}\right\} \qquad (3.22)$$

T.7: (Analog to T.3)

Let $f(z)$ be an analytic function , regular within and on a closed contour C. If $f^1(\xi) \neq 0$, where ξ is a root (within C) of $f(z) = 0$, to be obtained by application of Mod. IV of the Newton algorithm , we have

$$|\xi - z_{n+2}| \leq \frac{M^2}{2m^2} |\xi - z_n|^3$$

[i.e. after n applications of (3.22)

$$|\xi - z_{2n}| \leq \frac{M2}{2m^2} |\xi - z_{2n-2}|^3 \quad]$$

where $M = \sup\limits_{|z-\xi| < |z_n - \xi|} |f^{11}(z)|$

$$m = \inf_{|z-\xi| < |z_n - \xi|} |f^1(z)|$$

Proof: We have from Taylor's expansion:

$$f(z_{n+1}) = (z_{n+1} - \xi)\ f^1(\xi) + 0\,[\,(z_{n+1} - \xi)^2\,]\,,\ z_{n+1} \to \xi \tag{3.23}$$

From (3.22) and (3.23):

$$f^1(z_n)(z_{n+2} - \xi) = f^1(z_n)(z_{n+1} - \xi) - (z_{n+1} - \xi)\,f^1(\xi)$$
$$+ 0\,[\,(z_{n+1} - \xi)^2\,]\,,\ z_{n+1} \to \xi$$

$$\therefore\ f^1(z_n)\ \frac{z_{n+2} - \xi}{z_{n+1} - \xi} = f^1(z_n) - f^1(\xi) + 0\,[\,(z_{n+1} - \xi)\,]\,,$$
$$z_{n+1} \to \xi$$

$$= f^{11}(\eta)\,(z_n - \xi) + 0\,[\,(z_{n+1} - \xi)\,],\ z_{n+1} \to \xi$$

η an intermediate point.

$$\therefore\ f^1(z_n) \cdot \frac{z_{n+2} - \xi}{(z_{n+1} - \xi)(z_n - \xi)} \to f^{11}(\xi)$$

(since $\eta \to \xi$ as $z_n \to \xi$.)

Therefore, as $f^1(\xi) \neq 0$

$$\frac{z_{n+2} - \xi}{(z_{n+1} - \xi)(z_n - \xi)} \longrightarrow \frac{f^{11}(\xi)}{f^{1}(\xi)} \qquad (3.24)$$

From T.3 we have $\dfrac{\xi - z_{n+1}}{(\xi - z_n)^2} \longrightarrow -\dfrac{1}{2} \dfrac{f^{11}(\xi)}{f^{1}(\xi)}$

and this together with (3.24) give:

$$\frac{z_{n+2} - \xi}{(z_n - \xi)^3} \longrightarrow \frac{1}{2} \left(\frac{f^{11}(\xi)}{f^{1}(\xi)} \right)^2$$

$$\therefore \quad | z_{n+2} - \xi | \leq \frac{1}{2} \left(\frac{M}{m} \right)^2 | z_n - \xi |^3 \qquad (3.25)$$

In one application of (3.22), i.e. going from z_n to z_{n+2} we need to compute the three unknowns $f(z_n)$, $f^1(z_n)$, $f(z_{n+1})$. (The work pertained to the other computations involved are generally speaking negligable in comparison.) Thus in going from z_{n-2} to z_{n+2} we have six such quantities to be computed. Raughly speaking, the same amount of work is done after only two applications of the original N-R method. Hence we have in the case of Mod.IV

$$| z_{n+2} - \xi | \leq \frac{1}{16} \left(\frac{M}{m} \right)^8 | z_{n-2} - \xi |^9$$

whilst for the N-R method

$$| z_{n+2} - \xi | \leq \left(\frac{1}{2} \frac{M}{m} \right)^7 | z_{n-1} - \xi |^8$$

It is evident that Mod.IV is definitely an improvement of the N-R method in case of the "smooth" $f(z)$, i.e. where $M < m$. Even in the rather "bad" case $f(x) = x^3 - 2x - 5 = 0$

(considered above) , where $f^{11}(\xi) / f^{1}(\xi) \approx 1.2$,

we have for

$x_o = 2$: $\quad x_1 = 2.1$

$x_2 = 2.0939$

$x_3 = 2.094\ 551\ 72$

$x_4 = 2.094\ 551\ 481\ 367\ 28$

Here the error in x_4 is of the same order of magnitude as that of x_3 in column I above. On the other hand, it must be noticed that in using this modification the values of $f(z_n)$, $f^1(z_n)$ must be calculated to a much higher degree of accuracy. This is due to the factor $| z_n - \xi|^3$ in equation (3.25) .

<u>Mod.IV (a)</u>: We will now investigate the following modification in case of a root ξ of multiplicaty p :

$$\left.\begin{aligned} z_{n+1} &= z_n - p \frac{f(z_n)}{f^1(z_n)} \\ z_{n+2} &= z_{n+1} - p \frac{f(z_{n+1})}{f^1(z_n)} \quad \text{\_\_\_\_\_ (a)} \end{aligned}\right\} \quad (3.26)$$

We have from Taylor

$$f(z_{n+1}) = \frac{(z_{n+1} - \xi)^p}{p!} f^{(p)}(\xi) + \frac{(z_{n+1} - \xi)^{p+1}}{(p+1)!} f^{(p+1)}(\xi)$$

$$+ 0 [(z_{n+1} - \xi)^{p+2}] , z_{n+1} \to \xi$$

and

$$f^1(z_n) = \frac{(z_n - \xi)^{p-1}}{(p-1)!} f^{(p)}(\xi) + \frac{(z_n - \xi)^p}{p!} f^{(p+1)}(\xi)$$

$$+ 0 [(z_n - \xi)^{p+1}] , z_n \to \xi .$$

This together with Mod.IV (a) give,

$$f^{1}(z_n)\,(z_{n+2}-\xi) = f^{1}(z_n)\,(z_{n+1}-\xi) - \frac{(z_{n+1}-\xi)^{p}}{(p-1)!}\, f^{(p)}(\xi)$$

$$- p\,\frac{(z_{n+1}-\xi)^{p+1}}{(p+1)!}\, f^{(p+1)}(\xi) + 0\,[\,(z_{n+1}-\xi)^{p+2}\,]\;,\; n \to \infty$$

or

$$\frac{z_{n+2}-\xi}{z_{n+1}-\xi}\left[\,\frac{(z_n-\xi)^{p-1}}{(p-1)!}\, f^{(p)}(\xi) + \frac{(z_n-\xi)^{p}}{p!}\, f^{(p+1)}(\xi) + 0\,[\,(z_n-\xi)^{p+1}\,]\,\right]$$

$$= \frac{(z_n-\xi)^{p-1}}{(p-1)!}\, f^{(p)}(\xi) + \frac{(z_n-\xi)^{p}}{p!}\, f^{(p+1)}(\xi) - \frac{(z_{n+1}-\xi)^{p-1}}{(p-1)!} \times f^{(p)}(\xi)$$

$$-p\,\frac{(z_{n+1}-\xi)^{p}}{(p+1)!}\, f^{(p+1)}(\xi) + 0\,[\,(z_n-\xi)^{p+1}\,] + 0\,[\,(z_{n+1}-\xi)^{p+1}\,]$$

$$n \to \infty$$

i.e.

$$\frac{z_{n+2}-\xi}{z_{n+1}-\xi}\left[\,\frac{1}{(p-1)!}\, f^{(p)}(\xi) + \frac{(z_n-\xi)}{p!}\, f^{(p+1)}(\xi)\,\right]$$

$$= \frac{1}{(p-1)!}\, f^{(p)}(\xi) + \frac{z_n-\xi}{p!}\, f^{(p+1)}(\xi) - \left(\frac{z_{n+1}-\xi}{z_n-\xi}\right)^{p-1}\frac{1}{(p-1)!} \times f^{(p)}(\xi)$$

$$-\frac{p}{(p+1)!}\,\frac{(z_{n+1}-\xi)^{p}}{(z_n-\xi)^{p-1}}\, f^{(p+1)}(\xi)$$

$$+ 0\left[\frac{(z_{n+2}-\xi)\,(z_n-\xi)}{z_{n+1}-\xi}\right] + 0\,[\,(z_n-\xi)\,] + 0\left[\frac{(z_{n+1}-\xi)^{p+1}}{(z_n-\xi)^{p}}\right]$$

$$n \to \infty$$

$$\therefore \quad \frac{z_{n+2} - \xi}{z_{n+1} - \xi} \xrightarrow[\infty]{n} \frac{(p-1)!}{(p-1)!} \quad \frac{f^{(p)}(\xi)}{f^{(p)}(\xi)} = 1$$

By T.5 we have

$$\frac{z_{n+1} - \xi}{(z_n - \xi)^2} \longrightarrow \frac{f^{(p+1}(\xi)}{p(p+1)\, f^{(p)}\,(\xi)}$$

$$\therefore \quad \frac{z_{n+2} - \xi}{(z_n - \xi)^2} \longrightarrow \frac{f^{(p+1)}(\xi)}{p(p+1)\, f^{(p)}(\xi)}$$

From this it is evident that nothing at all is gained by introducing the intermediate step (a) in the algorithm Mod.II.

Mod. IV (b): Also for a p-fold root ξ.

$$\left.\begin{aligned} z_{n+1} &= z_n - \frac{f^{(p-1)}(z_n)}{f^{(p)}\,(z_n)} \\ z_{n+2} &= z_{n+1} - \frac{f^{(p-1)}(z_{n+1})}{f^{(p)}(z_n)} \end{aligned}\right\}$$

T.8: Given f(a) an analytic function, regular within a closed contour C, and continous within and on C. If ξ is a root of multiplicity p to be obtained by means of Mod.IV (b), we have

$$|\xi - z_{n+2}| \leq \frac{1}{2} \left(\frac{M_{p+1}}{m_p}\right)^2 |\xi - z_n|^3$$

$$M_{p+1} = \sup_{|z - \xi| < |z_n - \xi|} |f^{(p+1)}(\xi)|$$

$$m_p = \inf_{|z - \xi| < |z_n - \xi|} |f^{(p)}(\xi)|$$

Proof: From Taylor:

$$f^{(p)}(z_n) = f^{(p)}(\xi) + f^{(p+1)}(\xi)(z_n - \xi) + o[(z_n - \xi)^2], \quad z_n \to \xi$$

$$f^{(p-1)}(z_{n+1}) = (z_{n+1} - \xi) f^{(p)}(\xi) + \frac{(z_{n+1} - \xi)^2}{2} f^{(p+1)}(\xi)$$

$$+ o[(z_{n+1} - \xi)^3], \quad z_{n+1} \to \xi \ .$$

From Mod.IV (b):

$$f^{(p)}(z_n)(z_{n+2} - \xi) = (z_{n+1} - \xi) f^{(p)}(z_n) - f^{(p-1)}(z_{n+1})$$

or

$$(z_{n+2} - \xi)[f^{(p)}(\xi) + (z_n - \xi) f^{(p+1)}(\xi)] =$$

$$(z_{n+1} - \xi)[f^{(p)}(\xi) + (z_n - \xi) f^{(p+1)}(\xi)]$$

$$- (z_{n+1} - \xi)[f^{(p)}(\xi) + \frac{(z_{n+1} - \xi)}{2} f^{(p+1)}(\xi)]$$

$$+ o[(z_{n+2} - \xi)(z_n - \xi)^2] + o[(z_{n+1} - \xi)(z_n - \xi)^2]$$

$$+ o[(z_{n+1} - \xi)^3], \quad n \to \infty .$$

i.e.

$$\frac{z_{n+2} - \xi}{(z_n - \xi)(z_{n+1} - \xi)}[f^{(p)}(\xi) + (z_n - \xi) f^{(p+1)}(\xi)]$$

$$= f^{(p+1)}(\xi) - \frac{z_{n+1} - \xi}{2(z_n - \xi)} f^{(p+1)}(\xi) + o\left[\frac{(z_n - \xi)(z_{n+2} - \xi)}{z_{n+1} - \xi}\right]$$

$$+ o[(z_n - \xi)] + o\left[\frac{(z_{n+1} - \xi)^2}{z_n - \xi}\right], \quad n \to \infty .$$

We had from Mod.II (a)

$$\frac{z_{n+1} - \xi}{(z_n - \xi)^2} \xrightarrow[\infty]{n} \frac{1}{2} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)} \qquad (3.27)$$

i.e. $$\frac{z_{n+1} - \xi}{z_n - \xi} = 0 [(z_n - \xi)] \quad , \quad n \longrightarrow \infty$$

$$\therefore \quad \frac{z_{n+2} - \xi}{(z_n - \xi)(z_{n+1} - \xi)} \xrightarrow[\infty]{n} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)} \qquad (3.28)$$

(3.27) and (3.28) together give

$$\frac{z_{n+2} - \xi}{(z_n - \xi)^3} \xrightarrow[\infty]{n} \frac{1}{2} \left(\frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)} \right)^2$$

We observe that the order of convergence in this case does not compare very favourably with that of two successive applications of Mod.II, which yield:

$$\frac{z_{n+2} - \xi}{(z_n - \xi)^4} \longrightarrow \left(\frac{1}{p(p+1)} \quad \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)} \right)^3$$

Therefore, Mod. V (b) (see later) will be the better modification by far in the case of a p-fold root.

<u>Mod.V :</u> We can also try to reduce the amount of work done in the N-R formula, by replacing at every second step the denominator $f^1(z_n)$ by a convenient combination of $f(z_n)$ and $f(z_{n-1})$, i.e. we have the modification:

$$\left. \begin{aligned} z_{n+1} &= z_n - \frac{f(z_n)}{f^1(z_n)} \\ z_{n+2} &= z_{n+1} - \frac{f(z_{n+1})(z_{n+1} - z_n)}{2f(z_{n+1}) - f(z_n)} \end{aligned} \right\} \qquad (3.29)$$

<u>T. 9:</u> Let $f(z)$ be an analytic function regular within and on a closed contour C. If $f^1(\xi) \neq 0$ where ξ is a root within C

of $f(z) = 0$ to be obtained by application of Mod. V of the Newton algorithm , we have

$$\frac{\xi - z_{n+2}}{(\xi - z_n)^4} \longrightarrow -\frac{1}{24} \frac{f^{11}(\xi)}{f^1(\xi)^3}\left[3 f^{11}(\xi)^2 - 2f^1(\xi) f^{111}(\xi)\right]$$

Proof: It will be sufficient to show that

$$\frac{\xi - z_2}{(\xi - z_o)^4} \longrightarrow -\frac{1}{24} \frac{f^{11}(\xi)}{f^1(\xi)^3} [3 f^{11}(\xi)^2 - 2 f^1(\xi) f^{111}(\xi)]$$

Put $$z_1 - z_o = -\frac{f(z_o)}{f^1(z_o)} = h \qquad (3.30)$$

$$z_2 - z_1 = -\frac{f(z_1)\, h}{2f(z_1) - f(z_o)} = k \qquad (3.31)$$

Then $$f(z_1) + k f^1(z_1) = f(z_1) \left[1 - \frac{hf^1(z_1)}{2f(z_1) - f(z_o)} \right]$$

$$= f(z_1) (A/B) \qquad (3.32)$$

where $A = 2f(z_1) - f(z_o) - hf^1(z_1)$

$B = 2f(z_1) - f(z_o) = 2f(z_1) + hf^1(z_o)$

From T.3 we have

$$\frac{\xi - z_1}{(\xi - z_o)^2} \longrightarrow -\frac{1}{2} \frac{f^{11}(\xi)}{f^1(\xi)}$$

i.e. $\xi - z_1 = 0 [(\xi - z_o)^2]$, $z_o \longrightarrow \xi$

Since $h = z_1 - z_o = (\xi - z_o) - (\xi - z_1)$

$= \xi - z_o - 0 [(\xi - z_o)^2]$, $z_o \longrightarrow \xi$

$$\therefore \quad h \simeq \xi - z_o \tag{3.33}$$

we can write

$$\frac{\xi - z_1}{h^2} \longrightarrow - \frac{f^{11}(\xi)}{2f^{1}(\xi)}$$

This we can rewrite as

$$\frac{2f(z_1)}{h^2} \longrightarrow f^{11}(\xi) \quad , \; f(z_1) = 0\,(h^2),\; z_o \longrightarrow \xi \tag{3.34}$$

Keeping in mind that

$$f(z_1) \simeq (z_1 - \xi)\; f^{1}(\xi)$$

$$\therefore \quad B = h\, f^{1}(z_o) + 0\,(h^2) \qquad , \; z_o \longrightarrow \xi$$

$$\therefore \quad B \simeq h\, f^{1}(\xi) \tag{3.35}$$

From (3.31) , (3.34), (3.35)

$$\frac{k}{h^2} \longrightarrow - \frac{f^{11}(\xi)}{2f^{1}(\xi)}$$

i.e. $$\frac{k^2}{2h^4} \longrightarrow \frac{f^{11}(\xi)^2}{8f^{1}(\xi)^2} \tag{3.36}$$

Now $A = 2\, f(z_1) - f(z_o) - h\, f^{1}(z_1)$

$$= 2\, f(z_o + h) - h\, [\, f^{1}(z_o + h) - f^{1}(z_o)\,]$$

$$(\text{Since } f(z_o) = -\, h\, f^{1}(z_o)\,)$$

Develop in terms of h up to h^3 , i.e.:

$$A = 2\, f(z_o) + 2\, h\, f^{1}(z_o) + h^2\, f^{11}(z_o) + \frac{h^3}{3}\, f^{111}(\eta_1)$$

$$- [h^2 f^{11}(z_o) + \frac{h^3}{2} f^{111}(\eta_2)]$$

η_1 , η_2 intermediate points.

Since $f(z_o) + h f^1(z_o) = 0$, we finally have

$$A = h^3 [\frac{1}{3} f^{111}(\eta_1) - \frac{1}{2} f^{111}(\eta_2)]$$

But $\eta_1, \eta_2 \longrightarrow \xi$ as $z_o \longrightarrow \xi$

therefore $A/h^3 \longrightarrow - \frac{1}{6} f^{111}(\xi)$ (3.37)

Now from (3.32), (3.34) , (3.35), (3.37) we obtain

$$f(z_1) + k f^1(z_1) = f(z_1) (A/B)$$

$$\longrightarrow - \frac{1}{12} \frac{1}{f^1(\xi)} h^4 f^{11}(\xi) f^{111}(\xi) \quad (3.38)$$

We now develop

$f(z_2) = f(z_1 + k)$ in terms of powers of k up to k^2.

i.e.

$$f(z_2) = [f(z_1) + k f^1(z_1)] + \frac{1}{2} k^2 f^{11}(\eta_3)$$

where $\eta_3 \longrightarrow \xi$ as $z_1 \longrightarrow \xi$.

and by (3.36) and (3.38) this gives

$$\frac{f(z_2)}{h^4} \longrightarrow - \frac{1}{12} \frac{f^{11}(\xi) f^{111}(\xi)}{f^1(\xi)} + \frac{f^{11}(\xi)^3}{8f^1(\xi)^2}$$

Since $z_2 - \xi \simeq \dfrac{f(z_2)}{f^1(\xi)}$

we finally have by (3.33)

$$\frac{z_2 - \xi}{(z_o - \xi)^4} \longrightarrow \frac{1}{24} \frac{f^{11}(\xi)}{f^1(\xi)^3} [3 f^{11}(\xi)^2 - 2 f^1(\xi) f^{111}(\xi)]$$

In the application of Mod. V we again have a similar improvement as in the case of Mod.IV. Again however, the values of $f(z_n)$, $f^1(z_n)$ have to be calculated to a much higher degree of accuracy.

Factually, we must use double the number of decimals as in the case of the N-R method.

In case of a p-fold root ξ the following two modifications (V (a) , (b)) seem to be natural suggestions.

Mod. V (a):

$$\left.\begin{aligned} z_{n+1} &= z_n - p \frac{f(z_n)}{f^1(z_n)} \\ z_{n+2} &= z_{n+1} - p \frac{f(z_{n+1})\,(z_{n+1} - z_n)}{2f(z_{n+1}) - f(z_n)} \end{aligned}\right\} \text{(a)}$$

We will show however, that this is no improvement of Mod.II at all , since

$$\frac{z_{n+2} - \xi}{z_{n+1} - \xi} \;:\!\!\longrightarrow 1$$

and thus by T.5 ,

$$\frac{z_{n+2} - \xi}{(z_n - \xi)^2} \longrightarrow \frac{1}{p(p+1)} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$$

i.e. nothing at all is gained by introducing the intermediate step (a) in the algorithm Mod.II .

Proof of Claim:

We have from Taylor:

$$f(z_n) = \frac{(z_n - \xi)^p}{p!} f^{(p)}(\xi) + \frac{(z_n - \xi)^{p+1}}{(p+1)!} f^{(p+1)}(\xi)$$

$$+ \, 0 \left[(z_n - \xi)^{p+2} \right] , \quad z_n \longrightarrow \xi \qquad (3.39)$$

Then (3.39) together with Mod.V (a) give

$$(2\, f(z_{n+1}) - f(z_n))\,(z_{n+2} - \xi) = (z_{n+1} - \xi)(2\, f(z_{n+1}) - f(z_n))$$

$$- p\, f(z_{n+1})\,(z_{n+1} - z_n)$$

$$= (z_{n+1} - \xi)\,[\,(2-p)\, f(z_{n+1}) - f(z_n)\,]$$

$$+ p\,(z_n - \xi)\, f(z_{n+1})$$

For large n:

$$\frac{z_{n+2} - \xi}{z_{n+1} - \xi} \Big\{ 1/p\,!\,[\,2(z_{n+1} - \xi)^p - (z_n - \xi)^p\,)\, f^{(p)}(\xi)$$

$$+ 1/(p+1)!\ [\,2\,(z_{n+1} - \xi)^{p+1} - (z_n - \xi)^{p+1}\,]\, f^{(p+1)}(\xi) \Big\} =$$

$$f^{(p)}(\xi) \Big[(2-p) \frac{(z_{n+1} - \xi)^p}{p!} - \frac{(z_n - \xi)^p}{p!} + \frac{(z_n - \xi)\,(z_{n+1} - \xi)^{p-1}}{(p-1)!} \Big]$$

$$+ f^{(p+1)}(\xi) \Big[(2-p) \frac{(z_{n+1} - \xi)^{p+1}}{(p+1)!} - \frac{(z_n - \xi)^{p+1}}{(p+1)!} + \frac{p(z_n - \xi)\,(z_{n+1} - \xi)^p}{(p+1)!} \Big]$$

or

$$\frac{z_{n+2} - \xi}{z_{n+1} - \xi} \Big[\frac{1}{p!} \Big\{ 2 \Big(\frac{z_{n+1} - \xi}{z_n - \xi} \Big)^p - 1 \Big\} f^{(p)}(\xi)$$

$$+ \frac{1}{(p+1)!} \Big\{ 2 \frac{(z_{n+1} - \xi)^{p+1}}{(z_n - \xi)^p} - (z_n - \xi) \Big\} f^{(p+1)}(\xi) \Big]$$

$= \text{L.H.S.}$

L.H.S.

$$= f^{(p)}(\xi)\left[\frac{(2-p)}{p!}\left(\frac{z_{n+1}-\xi}{z_n-\xi}\right)^p - \frac{1}{p!} + \frac{1}{(p-1)!}\left(\frac{z_{n+1}-\xi}{z_n-\xi}\right)^{p-1}\right]$$

$$+ f^{(p+1)}(\xi)\left[\frac{(2-p)}{(p+1)!}\frac{(z_{n+1}-\xi)^{p+1}}{(z_n-\xi)^p} - \frac{(z_n-\xi)}{(p+1)!} + \frac{p}{(p+1)!}\frac{(z_{n+1}-\xi)^p}{(z_n-\xi)^{p-1}}\right]$$

We have from T.5 : $\dfrac{z_{n+1}-\xi}{(z_n-\xi)^2} \longrightarrow \dfrac{1}{p(p+1)}\,\dfrac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$,

i.e. $\left(\dfrac{z_{n+1}-\xi}{z_n-\xi}\right)^p = o\,[\,(z_n-\xi)^p\,],\quad z_n \longrightarrow \xi$

Thus $\dfrac{z_{n+2}-\xi}{z_{n+1}-\xi} \longrightarrow 1$

<u>Mod. V (b):</u> Analog to Mod. II (a) we can introduce the following:

$$\left.\begin{aligned} z_{n+1} &= z_n - \frac{f^{(p-1)}(z_n)}{f^{(p)}(z_n)} \\ z_{n+2} &= z_{n+1} - \frac{f^{(p-1)}(z_{n+1})\,(z_{n+1}-z_n)}{2\,f^{(p-1)}(z_{n+1}) - f^{(p-1)}(z_n)} \end{aligned}\right\}$$

<u>P6 !</u> Let $f(z)$ be an analytic function, regular within and on a closed contour C. If ξ is a root of multiplicity p to be obtained by means of Mod.V (b) , we have

$$\frac{\xi - z_{n+2}}{(\xi - z_n)^4} \longrightarrow - \frac{1}{24} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)^3} \left[3 f^{(p+1)}(\xi)^2 - 2 f^{(p)}(\xi) \; f^{(p+2)}(\xi) \right]$$

Proof: Exactly the same as for T.9 . (Just replace f^1 throughout by $f^{(p)}$, and make use of the Corollary following T.3 instead of T.3 itself.)

This shows, that in contrast with its analog (Mod.II (a), Mod. V (b) is indeed superior (in most cases) to Mod. V (a) (the analog of Mod. II). Factually, two successive applications of Mod. II will roughly give the same degree of approximation as one application of Mod. V (b). In doing this, two values of f(z) and two of $f^1(z)$ must be calculated in the case of Mod.II, whilst in the case of Mod. V (b) we have to calculate (though to twice the degree of accuracy) only two values of f(z) and one of $f^1(z)$. These statements can be verified immediately with the help of the following short synopsis.

Mod.II:
$$\frac{\xi - z_{n+1}}{(\xi - z_n)^2} \longrightarrow - \frac{1}{p(p+1)} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$$

i.e.
$$\frac{\xi - z_{n+2}}{(\xi - z_n)^4} \longrightarrow - \left[\frac{1}{p(p+1)} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)} \right]^3$$

Mod. II (a):
$$\frac{\xi - z_{n+1}}{(\xi - z_n)^2} \longrightarrow - \frac{1}{2} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$$

Mod. V (a):
$$\frac{\xi - z_{n+2}}{(\xi - z_n)^2} \longrightarrow - \frac{1}{p(p+1)} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)}$$

Mod. V (b):

$$\frac{\xi - z_{n+2}}{(\xi - z_n)^4} \longrightarrow -\frac{1}{24} \frac{f^{(p+1)}(\xi)}{f^{(p)}(\xi)^3} \left[3 f^{(p+1)}(\xi)^2 - 2 f^{(p)}(\xi) \cdot f^{(p+2)}(\xi) \right]$$

IV.

"Polynomial Algorithms" and Frame's Modification.

Up till now we have only discussed algorithms of the second order. Before proceeding to the discussion of algorithms of higher orders, it will not be unjustified to mention here another important type of algorithm of the second order which can also be adapted to satisfy the conditions for higher order convergence. These are functions $F(z)$, (for finding the roots of polynomials $f(z)$), which satisfy apart from $F(\zeta)=\zeta$, $F'(\zeta)=0$ also the further restriction, namely that $F(z)$ must be a polynomial. This type of algorithm was i.a. thoroughly discussed by C. Domb, A. S. Householder and H. Schwerdtfeger. These algorithms are constructed in the following way:

We will only consider polynomials $f(z)$ with simple roots. (This is no restriction of course.) Thus, since the g. c. d. $(f(z), f'(z))=1$ two polynomials $h(z)$ and $h_1(z)$ can be found such that

$$h_1(z)f(z)-h(z)f'(z)=1 \qquad (4.1)$$

Then we can choose as algorithm $F(z)=z+f(z)h(z)$ (4.2)

Obviously $F(\zeta)=\zeta$, $F'(\zeta)=0$ and $F(z)$ is a polynomial.

The general solution of (4.1) is:

$$H(z)=h(z)+p(z)f(z) \quad , H_1(z)=h_1(z)+p(z)f'(z)$$

where $p(z)$ is an arbitrary polynomial. Thus $h(z)$ in (4.2) is not uniquely defined. Special solutions $h(z)$ and $h_1(z)$ can always be found by means of the Euclidean algorithm of course. The simplest numerical method to apply here would be the method of unknown coefficients as indicated by H. Schwerdtfeger in his paper [4].

<u>T. 10:</u> Given a polynomial $f(z)$ with a simple root ;then we have after n applications of

$$z_n=z_{n-1}+f(z_{n-1})h(z_{n-1}) \qquad (4.3)$$

$h(z)$ a polynomial,

$$\frac{|z_n - \zeta|}{|z_{n-1} - \zeta|} \leqq 1 + M_h \cdot M_{f'}$$

where $M_h = \underset{|z-\zeta|<|z_{n-1}-\zeta|}{\text{maximum}} |h(z)|$

$M_{f'} = \underset{|z-\zeta|<|z_{n-1}-\zeta|}{\text{maximum}} |f'(z)|$

Proof: Trivial. We have from (4. 3) and Taylors theorem:

$$z_n - \zeta = z_{n-1} - \zeta + [f'(\zeta)(z_{n-1}-\zeta) + O[(z_{n-1}-\zeta)^2]\ ][h(\zeta) + (z_{n-1}-\zeta)h'(\zeta) + O[(z_{n-1}-\zeta)^2]\ ], \quad z_{n-1} \to \zeta.$$

$$\frac{z_n - \zeta}{z_{n-1} - \zeta} \underset{\infty}{\overset{n}{\rightarrow}} 1 + f'(\zeta)h(\zeta)$$

If an algorithm of the third order is required , we can replace (4. 2) by

$$F(z) = z + f(z)H(z) = z + f(z)h(z) + p(z)f(z)^2 \qquad (4.4)$$

p(z) an arbitrary polynomial.

This is according to P. 5 the most general algorithm (of the polynomial type) of the second order. A special algorithm of the third order can now be acquired by choosing p(z) such that $F''(\zeta)=0$.

We have by differentiation:

$$F'(z) = 1 + h(z)f'(z) + f(z)[h'(z) + 2p(z)f'(z) + p'(z)f(z)^2]$$

$$= f(z)[h_1(z) + h'(z) + 2p(z)f'(z) + p'(z)f(z)]$$

Thus, to obtain $F''(\zeta)=0$, we must obviously choose p(z) such that

$$h_1(z) + h'(z) + 2p(z)f'(z) = q(z)f(z)$$

According to (4. 1) this will be satisfied if

$$p(z) = \frac{1}{2}h(z)\ [h'(z) + h_1(z)]$$

$$q(z) = h_1(z)\ [h'(z) + h_1(z)].$$

From this resulting algorithm of the third order, one of the fourth order can be obtained by P. 5 and restrictions on the arbitrary $\psi(z)$ of equation (2. 3). Proceed similarly for higher orders.

Mod. VI:(Frame). J.S.Frame [5], [6] and later also H. S. Wall [7] have suggested the following modification of Newton's algorithm:

$$z_n = z_{n-1} - \frac{2f(z_{n-1})f'(z_{n-1})}{2f'(z_{n-1})^2 - f(z_{n-1})f''(z_{n-1})} \tag{4.5}$$

or $F(z) = z - \dfrac{2f(z)f'(z)}{2f'(z)^2 - f(z)f''(z)}$

(Notice the striking similarity in form with Mod. I)

By substituting $f(z) = (z-\xi)^p \psi(z) = \varepsilon^p \psi(z)$

where $\lim_{z \to \xi} |\psi(z)| < \infty$, $\neq 0$.

we obtain $F'(\xi) = 1 - \frac{2}{p+1}$

We will therefore apply (4.5) only in those cases where $f'(\xi) \neq 0$.

[Mod. VI was arrived at, after the following observations:

The equation of the parabola through the point $(z_{n-1}, f(z_{n-1}))$, having the same first and second derivatives at $z=z_{n-1}$ as $y=f(z)$, is

$$y = f(z_{n-1}) + (z - z_{n-1})f'(z_{n-1}) + \tfrac{1}{2}(z - z_{n-1})^2 f''(z_{n-1})$$

Let z_n be a solution of the equation which results if we put $y=0$.

Then

$$z_n - z_{n-1} = \frac{-f(z_{n-1})}{f'(z_{n-1}) + \frac{1}{2}(z_n - z_{n-1})f''(z_{n-1})}$$

If we take $z_{n-} - z_{n-1} = -f(z_{n-1})/f'(z_{n-1})$ in this formula, we then obtain (4.5)]

In case of multiple roots the following may be used:

Mod. VI(a): $F(z) = z - \dfrac{2f(z)\left[f'(z)^2 - f(z)f''(z)\right]}{2(f'(z))^3 - 3f(z)f'(z)f''(z) + (f(z))^2 f'''(z)}$

(Obtained from Mod. VI by replacing f(z) with $f(z)/f'(z)$.)

Remarkable however is the fact that Mod. VI yields an algorithm of the third order for simple roots (p=1), for then

$$F''(\xi) = -\frac{f''(\xi)}{f'(\xi)} + \frac{2}{f'(\xi)}\frac{1}{2} f''(\xi) = 0.$$

<u>T. 11:</u> Given f(z) regular within a closed contour C. If $f'(\xi) \neq 0$ where ξ is a root of f(z) within C, we have after n applications of Mod. VI:

$$\frac{z_n - \xi}{(z_{n-1} - \xi)^3} \xrightarrow[n \to \infty]{} \frac{1}{12(f'(\xi))^2} \left[3(f''(\xi))^2 - 2f'(\xi)f'''(\xi)\right]$$

<u>Proof:</u> We have from (4.5)

$$(z_n - \xi)\left[2f'(z_{n-1})^2 - f(z_{n-1})f''(z_{n-1})\right] = (z_{n-1} - \xi)\left[2f'(z_{n-1})^2 - f(z_{n-1})f''(z_{n-1})\right] - 2f(z_{n-1})f'(z_{n-1})$$

We also have:

$$\left.\begin{aligned} f(z_{n-1}) &= (z_{n-1} - \xi)f'(\xi) + \frac{(z_{n-1} - \xi)^2}{2} f''(\xi) + \frac{(z_{n-1} - \xi)^3}{6} f'''(\xi) + O[(z_{n-1} - \xi)^4] \\ f'(z_{n-1}) &= f'(\xi) + (z_{n-1} - \xi)f''(\xi) + \frac{(z_{n-1} - \xi)^2}{2} f'''(\xi) + O[(z_{n-1} - \xi)^3] \\ f''(z_{n-1}) &= f''(\xi) + (z_{n-1} - \xi)f'''(\xi) + O[(z_{n-1} - \xi)^2] \end{aligned}\right\} z_{n-1} \to \xi.$$

Then

$$(z_n - \xi)\Big[2(f'(\xi))^2 + 3(z_{n-1} - \xi)f'(\xi)f''(\xi) + \frac{3}{2}(z_{n-1} - \xi)^2(f''(\xi))^2 + \frac{(z_{n-1} - \xi)^4}{3}(f'''(\xi))^2 + (z_{n-1} - \xi)^2 f'(\xi)f'''(\xi) + \frac{4}{3}(z_{n-1} - \xi)^3 f''(\xi)f'''(\xi)\Big]$$

= R:H. S.

$$\text{R.H.S.} = (z_{n-1} - \xi)[2(f'(\xi))^2 + 3(z_{n-1} - \xi)f'(\xi)f''(\xi) + \frac{3}{2}(z_{n-1} - \xi)^2(f''(\xi))^2$$
$$+ \frac{(z_{n-1} - \xi)^4}{3} f'''(\xi)^2 + (z_{n-1} - \xi)^2 f'(\xi)f'''(\xi) + \frac{4}{3}(z_{n-1} - \xi)^3 f''(\xi)f'''(\xi)]$$

$$-2[(z_{n-1} - \xi)f'(\xi)^2 + \frac{3}{2}(z_{n-1} - \xi)^2 f'(\xi)f''(\xi) + \frac{2}{3}(z_{n-1} - \xi)^3 f'(\xi)f'''(\xi)$$
$$+ \frac{(z_{n-1} - \xi)^3}{2} f''(\xi)^2 + \frac{5}{12}(z_{n-1} - \xi)^4 f''(\xi)f'''(\xi) + \frac{(z_{n-1} - \xi)^5}{12} f'''(\xi)^2]$$

$$+ O[(z_{n-1} - \xi)^3(z_n - \xi)] + O[(z_{n-1} - \xi)^4] \quad , \qquad z_{n-1} \to \xi .$$

Therefore:

$$\frac{z_n - \xi}{z_{n-1} - \xi}[2f'(\xi)^2 + 3(z_{n-1} - \xi)f'(\xi)f''(\xi) + \frac{3}{2}(z_{n-1} - \xi)^2 f''(\xi)^2$$
$$+ \frac{(z_{n-1} - \xi)^4}{3} f'''(\xi)^2 + (z_{n-1} - \xi)^2 f'(\xi)f'''(\xi) + \frac{4}{3}(z_{n-1} - \xi)^3 f''(\xi)f'''(\xi)]$$

$$= \frac{1}{2}(z_{n-1} - \xi)^2 f''(\xi)^2 + \frac{(z_{n-1} - \xi)^4}{6} f'''(\xi)^2 + \frac{1}{2}(z_{n-1} - \xi)^3 f''(\xi)f'''(\xi)$$

$$- \frac{1}{3}(z_{n-1} - \xi)^2 f'(\xi)f'''(\xi) + O[(z_{n-1} - \xi)^2(z_n - \xi)] + O[(z_{n-1} - \xi)^3], \quad z_{n-1}, z_n \to \xi$$

Thus:

$$\frac{z_n - \xi}{(z_{n-1} - \xi)^3} \to \frac{1}{2f'(\xi)^2}\left[\frac{1}{2} f''(\xi)^2 - \frac{1}{3} f'(\xi)f'''(\xi)\right]$$

After comparison with the result of T. 3 , it is evident that this algorithm is definitely an improvement of the Newton-Raphson method.

<u>Applications to the equation $f(z)= z^m - a$, $m>1$:</u>

We will first consider the application of the "polynomial" algorithms as given by equations (4. 2) and (4. 4) :

Obviously $h(z)= -\frac{z}{ma}$, $h_1(z) = -\frac{1}{a}$ will be polynomials of lowest degree to solve equation (4. 1). Thus by (4. 2) we obtain as algorithm for approximating the m-th root of a:

$$F(z) = \frac{m+1}{m} z - \frac{1}{ma} z^{m+1} \tag{4. 6}$$

This formula has been given by Hartree and Domb.

D. R. Hartree: Notes on iterative processes ; Proceedings of the Cambridge Philosophical Society 45 (1949), 230-236.

C. Domb: On iterative solutions of algebraic equations ; Ibid. 45 (1949), 237-240.

Using (4. 4) we obtain an improved formula:

$$F(z) = \frac{z}{m} \left[\frac{(2m+1)(m+1)}{2} - \frac{2m+1}{2} z^m + \frac{m+1}{2a^2} z^{2m} \right] \tag{4. 7}$$

If Frame's Mod. VI is applied to $f(z)= z^m -a$, we obtain:

$$F(z)= z \frac{(m-1)z^m + (m+1)a}{(m+1)z^m + (m-1)a} \tag{4. 8}$$

a very handy formula already given by V:A. Bailey in 1941.

[Prodigous Calculation ;Australian Journal of Science 3, No. 4 , (1941) 78-80.]

In using the algorithms given by (4. 6) and (4. 8) , we have from T. 10 and 11 respectively, the following error estimates:

For equation (4. 6): $$\frac{z_n - \zeta}{z_{n-1} - \zeta} \simeq 1 - \zeta^m$$

equation (4. 8): $$\frac{z_n - \zeta}{(z_{n-1} - \zeta)^3} = \frac{1}{2} \left[3(m-1)^2 - 2(m-1)(m-2) \zeta^{-2} \right]$$

The following table gives the approximations to $\sqrt{2}$ obtained from (4. 6) , (4.8) and the N-R algorithm , starting with $z_0 = x_0 = 1$.

	N-R.	(4. 6)	(4. 8)
x_1:	1. 500000000	1. 250000000	1. 400000000
x_2:	1. 416666667	1 . 386718750	1. 41421319797
x_3:	1. 414215686	1. 413416939	1. 4142135623730950487956900 8
x_4:	1. 414213562	1. 414212534	

The value of x_3 found by Newton's formula is correct to four decimal places, while the value of x_3 found by formula (4. 8) is correct to nineteen decimal places. Starting with $x_0=10$, we find that Newton's formula gives for x_4 the value 1. 4442 , which is correct to one decimal place, while formula (4. 8) gives the approximation $x_4 = 1.414213562$; which is correct to nine decimal places.

Another application of (4. 5): The computation of the positive real root of the reduced cubic equation

$$x^3 + bx - c = 0 \ , \ b, c \text{ real}, \ b \neq 0 \ , \ c > 0.$$

Here newton's formula is

$$x_n = \frac{2x_{n-1}^3 + c}{3x_{n-1}^2 + b} \tag{4.9}$$

and formula (4. 5) is now

$$x_n = \frac{3x_{n-1}^5 - bx_{n-1}^3 + 6cx_{n-1}^2 + bc}{6x_{n-1}^4 + 3bx_{n-1}^2 + 3cx_{n-1} + b^2} \tag{4.10}$$

If b=2 , c= 20 and we take $x_0 = 2$, formula (4. 10) gives the approximations

$$x_1 = 2.46 \quad , \quad x_2 = 2.46954551$$

On the other hand (4.9) yields for $x_0 = 2$, the values

$$x_1 = 2.6 \ , \ x_2 = 2.47 \ , \ x_3 = 2.469546$$

The value of the root to nine decimal places is 2.469545649.

V.

Algorithms of order $k > 2$.

On discussing P. 5 in chapter II we have observed a general method for the practical (though laborious) construction of an algorithm of arbitrary order $k<\infty$ if an algorithm of the second (Or even first) order is known. We have already made use of this principle in chapter IV in deriving the cubic algorithm given by(4. 4) fom the quadratic one given by (4. 2). Starting with the N-R algorithm , E. Scröder and later E. Bodewig have obtained in this way as the most general algorithm of the k-th order for f(z) , $f'(\zeta) \neq 0$:

$$F_k(z) = z + \sum_{n=1}^{k-1} (-1)^n \frac{(f(z))^n}{n!} \left[\frac{1}{f'(z)} \frac{d}{dz} \right]^{n-1} \frac{1}{f'(z)} - f(z)^k \varphi_k(z) \qquad k > 2 \tag{5.1}$$

Where $\varphi_k(z)$ is an arbitrary function ,

and $\left[\frac{1}{f'(z)} \frac{d}{dz} \right]^r$ denotes that the operator $\frac{1}{f'(z)} \frac{d}{dz}$ must be applied r-times, i. e.

e. g. $\left[\frac{1}{f'(z)} \frac{d}{dz} \right]^3 g(z) = \frac{1}{f'(z)} \frac{d}{dz} \left[\frac{1}{f'(z)} \frac{d}{dz} \left(\frac{1}{f'(z)} g'(z) \right) \right]$

Again , in the case of a multiple root ζ we just replace f(z) in (5. 1) by $\frac{f(z)}{f'(z)}$.

Formula (5. 1) was given without proof by Scröder. The following two theorems proving its validity are due to H. Schwerdtfeger and D. R. Blaskett.

<u>T. 12:</u> Let w= f(z) be an analytic function regular within and on a closed contour C , and ζ a root of f(z) within C, $f'(\zeta) \neq 0$. Let z_0 be a point within C , "not too far" from ζ . Then, denoting the inverse of f(z) by $z= f^{-1}(w)$ we have

$$\zeta = \sum_{n=0}^{\infty} (-1)^n \frac{f(z_0)^n}{n!} \left(\frac{d^n f^{-1}(w)}{dw^n} \right)_{w= f(z_0)} = \exp. \left[- f(z_0) \frac{d f^{-1}(w)}{dw} \right]_{w= f(z_0)} \tag{5.2}$$

where the exponential function operates symbolically on the differential symbol.

Proof: This theorem follows immediately from the main theorems on the analyticity of the inverse of an analytic function. The requirement $f'(\zeta) \neq 0$ is necessary , since we have for inverse functions: (e. g. Copson p. 121)

If f(z) is an analytic function , regular in a neighbourhood of the point z_0 at which it takes the value w_0 , then the necessary and sufficient condition that the equation f(z)= w should have a unique solution $z = f^{-1}(w)$, regular in a neighbourhood of w_0 , is that $f'(z_0)$ (or $f'(\zeta)$ for z_0 "sufficiently close to" ζ) should not vanish. This unique solution is then given by

$$f^{-1}(w) = f^{-1}(w_0) + \sum_{n=1}^{\infty} \frac{(z-z_0)^n}{n!} \left[\frac{d^{n-1}}{dw^{n-1}} \left\{ \frac{d}{dw}(f^{-1}(w)) \cdot \left\{ \frac{w-w_0}{z-z_0} \right\}^n \right\} \right]_{w=w_0}$$

This is the Lagrange formula (Memoires de l' Acad. Roy. des Sci. -Berlin, 24 (1768) , 251.)

for the reversion of a power series. We can write this as :

$$f^{-1}(w) = f^{-1}(w_0) + \sum_{n=1}^{\infty} \frac{1}{n!} (z-z_0)^n \left[\left(\frac{w-w_0}{z-z_0}\right)^n \frac{d^n}{dw^n} (f^{-1}(w)) + \frac{d}{dw}(f^{-1}(w)) \frac{d^{n-1}}{dw^{n-1}} \left\{ \frac{w-w_0}{z-z_0} \right\}^n \right]_{w=w_0}$$

$$= f^{-1}(w_0) + \sum_{n=1}^{\infty} \frac{1}{n!} (z - z_0)^n f'(z_0)^n \left[\frac{d^n}{dw^n} f^{-1}(w) \right]_{w=w_0}$$

$$= f^{-1}(w_0) + \sum_{n=1}^{\infty} \frac{1}{n!} [f(z) - f(z_0)]^n \left[\frac{d^n}{dw^n} f^{-1}(w) \right]_{w=w_0}$$

and this gives formula (5. 2) for w = 0.

To obtain the Scröder-formula we introduce the operators δ^{μ} ($\mu = 0,1 , 2, \dots$)

as follows:

$$\delta^{0} f(z) = \frac{1}{f'(z)}, \qquad \delta^{1} f(z) = \frac{1}{f'(z)} \frac{d}{dz}\left(\frac{1}{f'(z)}\right),$$

$$\delta^{n} f(z) = \frac{1}{f'(z)} \frac{d}{dz}\left(\delta^{n-1} f(z)\right)$$

Then it can easily be shown by induction that

$$\delta^{1} f(z) = \left[\frac{d^{2} f^{-1}(w)}{dw^{2}}\right]_{w=f(z)}, \qquad \delta^{n} f(z) = \left[\frac{d^{n+1} f^{-1}(w)}{dw^{n+1}}\right]_{w=f(z)}$$

Thus:

$$\zeta = z_0 + \sum_{n=1}^{\infty} (-1)^n \frac{f(z_0)^n}{n!} \left(\delta^{n-1} f(z)\right)_{z=z_0}$$

$$= z_0 - f(z_0)\frac{1}{f'(z_0)} - \frac{f(z_0)^2}{2!} \frac{f''(z_0)}{f'(z_0)^3} + \frac{f(z_0)^3}{3!} \frac{f'(z_0)f'''(z_0) - 3f''(z_0)^2}{f'(z_0)^5}$$

$$+ \ldots\ldots\ldots\ldots \qquad (5.3)$$

Considering

$$F_k(z) = \sum_{n=0}^{k-1} (-1)^n \frac{f(z)^n}{n!} \delta^{n-1} f(z) \qquad (5.4)$$

i.e. a partial sum of (5.3) which will be used as iterative algorithm, we have the following theorem:

<u>T.13:</u> Rewriting (5.4) as

$$z_n = z_{n-1} + \sum_{j=1}^{k-1} (-1)^j \frac{f(z_{n-1})^j}{j!} \left[\delta^{j-1} f(z)\right]_{z=z_{n-1}}$$

we have $z_n \xrightarrow[n\to\infty]{} \zeta$, and

$$F_k(\zeta) = \zeta, \quad F_k'(\zeta) = 0, \ \ldots\ldots\ldots \ F_k^{(k-1)}(\zeta) = 0$$

i.e. $F_k(z)$ is an algorithm of k-th order.

<u>Proof:</u> For $F_k(z)$ to have $F_k'(\zeta) = F_k''(\zeta) = \ldots\ldots = F_k^{(k-1)}(\zeta) = 0$ we must obviously have something of the type:

$$F_k'(z) = f(z)^{k-1} g(z) f'(z) \qquad (5.5)$$

where the undetermined function $g(z)$ is regular at $z = \zeta$.

Now we must choose g(z) such that $F_k(\xi)=\xi$. Then according to T:1 and 2 we have $z_n \xrightarrow[n\to\infty]{} \xi$.

From (5.5) we have :

$$F_k(z) = \int f(z)^{k-1} g(z) f'(z)dz$$

$$= \int w^{k-1} g(f^{-1}(w))dw$$

$$= (k-1)! \sum_{n=0}^{k-1} \frac{(-1)^n}{(k-1-n)!} f(z)^{k-1-n} g_{n+1}(z) \tag{5.6}$$

where $g_{n+1}(z) = \int g_n(f^{-1}(w))\, dw = \ldots\ldots$

For $z = \xi$ this gives

$$F_k(\xi) = (-1)^{k-1}(k-1)!\, g_k(z)$$

Thus, a suitable choice for $g_k(z)$ would be

$$g_k(z) = \frac{(-1)^{k-1}}{(k-1)!} z$$

i.e.

$$g_n(z) = \frac{(-1)^{k-1}}{(k-1)!} \frac{d^{k-n}}{dw^{k-n}} [f^{-1}(w)]$$

Thus from (5.6):

$$F_k(z) = \sum_{n=0}^{k-1} \frac{(-1)^{k-1-n}}{(k-1-n)!} f(z)^{k-1-n} \cdot \delta^{k-n-2}(f(z))$$

which is exactly equation (5.4).

Applications:

I.) The quadratic equation. i.e.

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

and

$$\frac{dz}{df} = \frac{1}{2z-(z_1+z_2)} = \frac{1}{N}$$

$$\frac{d^n z}{df^n} = (-1)^{n-1} \frac{2^{n-1}.3.5.7.\ldots.(2n-3)}{N^{2n-1}}$$

$$= \left[\begin{matrix}\frac{1}{2}\\ n\end{matrix}\right] \frac{1}{N^{2n-1}} \cdot 2^{2n-1}.n!$$

$$= \delta^{n-1} f(z).$$

Put in (5. 2):

$$\zeta = z + \left(z - \frac{z_1+z_2}{2}\right) \sum_{n=1}^{\infty} (-1)^n \left[\frac{\frac{1}{2}}{n}\right] \cdot \left[\frac{(z-z_1)(z-z_2)}{\left(z - \frac{z_1+z_2}{2}\right)^2}\right]^n \qquad w = f(z)$$

(we write z instead of the initial z_0 to avoid confusion.)

$$= \frac{z_1+z_2}{2} + \left(z - \frac{z_1+z_2}{2}\right) \sum_{n=0}^{\infty} (-1)^n \left[\frac{\frac{1}{2}}{n}\right] \cdot \left[\frac{(z-z_1)(z-z_2)}{\left(z - \frac{z_1+z_2}{2}\right)^2}\right]^n$$

$$= \frac{z_1+z_2}{2} \pm \left(z - \frac{z_1+z_2}{2}\right) (1+t)^{1/2}$$

$$t = -\frac{(z-z_1)(z-z_2)}{\left(z - \frac{z_1+z_2}{2}\right)^2} \qquad \text{i.e. } \underline{\text{if}} \quad |t| \leq 1.$$

Under this condition we then have

$$\zeta = \frac{z_1+z_2}{2} \pm \left(z - \frac{z_1+z_2}{2}\right) \cdot \left[\frac{\frac{z_1-z_2}{2}}{z - \frac{z_1+z_2}{2}}\right]$$

$$= \left.\begin{matrix} z_1 & \text{for } + \\ z_2 & \text{for } - \end{matrix}\right\}$$

The condition for convergence cah be translated as

$$\rho_1 \; \rho_2 \leq \rho^2 \qquad \text{where} \qquad \begin{aligned} \rho_1 e^{i\theta_1} &= z - z_1 \\ \rho_2 e^{i\theta_2} &= z - z_2 \\ \rho \, e^{i\theta} &= z - \frac{z_1+z_2}{2} \end{aligned}$$

Then we know from an elementary theorem on the median of a triangle that

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

where 2E denotes the distance between the two rootpoints z_1 and z_2.

Thus the condition for convergence then changes to

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

Considering for the sake of simplicity the stadard hyperbola $x^2/a^2 - y^2/b^2 = 1$, $b^2 = a^2(e^2 - 1)$ in the real case, we have

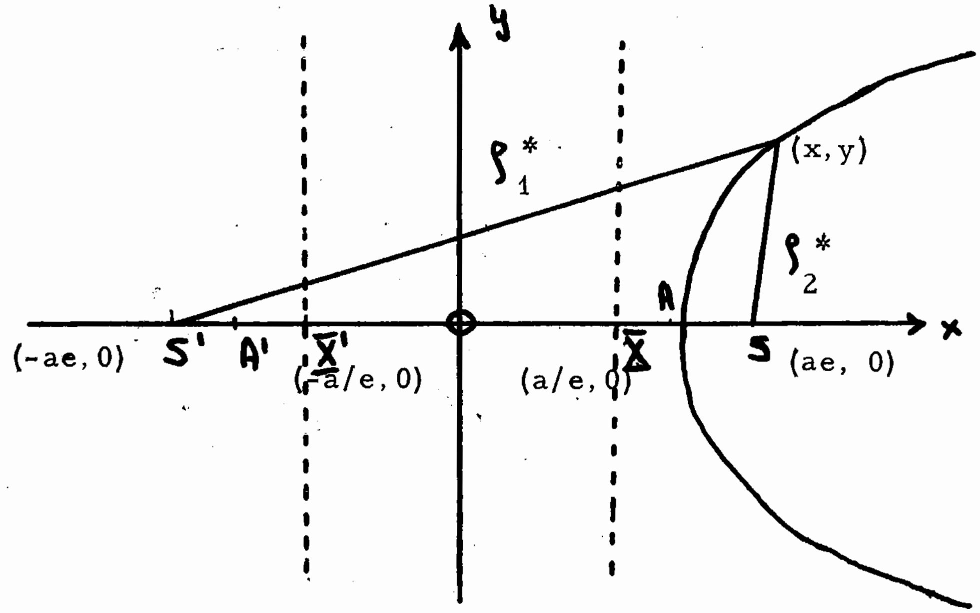

Fig. 2.

$$\rho^{*\,2}_{1} = (x + ae)^2 + y^2 \,, \qquad \rho^{*\,2}_{2} = (x - ae)^2 + y^2$$

$$(\rho^{*}_{1} - \rho^{*}_{2})^2 = \rho^{*\,2}_{1} + \rho^{*\,2}_{2} - 2\,\rho^{*}_{1}\,\rho^{*}_{2}$$

$$= 2x^2 + 2y^2 + 2a^2e^2 - 2\sqrt{(x^2 - a^2e^2)^2 + 2x^2y^2 + 2a^2e^2y^2 + y^4}\,.$$

$$= 2a^2e^2 = 2E^{*\,2} \qquad \text{for} \quad e = \sqrt{2}\,.$$

The domain of convergence is thus bounded by an equi-sided (i. e. $a = \pm b$, $e = \sqrt{2}$) hyperbola , with the roots z_1 and z_2 of the quadratic equation as foci - and incidentally it is that part of the complex number plane in which the foci itselves are situated (the hyperbola itself included.)

The immediate question is now , which one of the two roots is reached by choosing the initial algorithmic approximation z_0 in different areas of the domain of convergence determined above.

We have from above $1+t = (E/\rho)^2\, e^{2i(\theta_0 - \theta)}$

where $\dfrac{z_1 - z_2}{2} = Ee^{i\theta_0}$

Therefore

$$e^{(1/2)\log(1+t)} = (E / \rho)\, e^{(1/2)\log e^{2i(\theta_0 - \theta)}}$$

Considering that for a real number y

$\log e^{iy} = i(\, y + 2h\pi)$ where the integer h must be chosen such that $y + 2h\pi$ lies between $-\pi$ (excl.) and π (incl.) , we find:

$$e^{1/2 \log (1+t)} = (E/\rho)\, e^{i(\theta_0 - \theta + h\pi)} = \frac{\frac{z_1 - z_2}{2}}{z - \frac{z_1+z_2}{2}}\, e^{h\pi i} \tag{5.7}$$

where h must be choseb such that $\theta_0 - \theta + h\pi$ lies between $-\frac{\pi}{2}$ (excl.) and $+\frac{\pi}{2}$ (incl.).

The radius ρ from z to $\frac{z_1+z_2}{2}$ includes two supplementary angles with the line connecting z_1 and z_2. Call the angle on the side of z_1, w_1 and the other w_2.

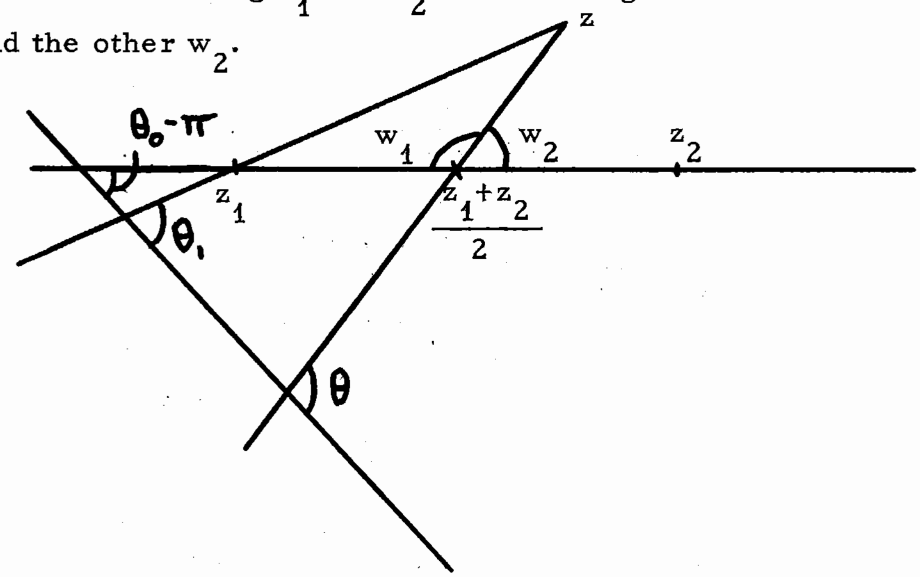

Fig. 3.

Then $\theta = \theta_0 - \pi + w_2$

and $-\frac{\pi}{2} < h\pi + \pi - w_2 \leq \frac{\pi}{2}$

(5.8)

or $-\frac{\pi}{2} < h\pi + w_1 \leq \frac{\pi}{2}$

Obviously for h even it follows from (5.7) that the positive square-root of 1+t is considered, i.e. root z_1 is obtained if z is chosen such that for even h, w_1 and w_2 satisfies (5.8). i.e. $w_1 \leq \frac{\pi}{2}$.

For odd h, $e^{h\pi i} = -1$ and then $w_2 < \frac{\pi}{2}$ in which case root z_2 is obtained from (5.2).

2.) $f(z) = z^m - a = 0.$

$$\frac{dz}{df} = \frac{1}{mz^{m-1}}$$

$$\frac{d^n z}{d\zeta^n} = (-1)^{n+1} \frac{(m-1)(2m-1)(3m-1)\ldots\ldots[(n-1)m-1]}{m^n} \frac{1}{z^{nm-1}}$$

$$= (-1)^{n+1} \frac{(n-1)!\,(m-1)(m-1/2)(m-1/3)\ldots\ldots(m-1/(n-1))}{m^n \cdot z^{nm-1}}$$

Then

$$\zeta = z - \sum_{n=1}^{\infty} \frac{(z^m-a)^n}{n} \frac{(m-1)(m-1/2)(m-1/3)\ldots\ldots(m-1/(n-1))}{m^n} \frac{1}{z^{nm-1}}$$

$$\approx z - \sum_{n=1}^{\infty} \left[\frac{z^m-a}{z^m}\right]^n \frac{1}{nm} z \qquad \text{for large m.}$$

$$= z - \frac{z}{m} \sum_{n=1}^{\infty} (-1)^n \frac{1}{n}\left[\frac{z^m-a}{z^m}\right]^n$$

$$= z - \frac{z}{m}\left[\log\left|1 + \frac{z^m-a}{z^m}\right| + i\theta + 2n\pi i\right], \text{ for } \left|\frac{a-z^m}{z^m}\right| \leqq 1$$

$$\text{except at } \frac{a-z^m}{z^m} = -1$$

Therefore:

$$\zeta = z\left[1 - \frac{1}{m}\log\left|2 - \frac{a}{z^m}\right| - i\frac{\theta}{m} - \frac{2n\pi}{m} i\right]$$

and this holds if $\left|1 - \frac{a}{z^m}\right| \leqq 1$, excluding the case when a=0.

3.) The non- analytic case:

Given f(z) continuous of non-analyticity r in a simply connected region containing the rectifiable Jordan-curve C, given by

$w(t) = u(t) + iv(t)$, $\alpha \leqq t \leqq \beta$.

Que.: Find an approximation for a root ζ of f(z).

Definition of non-analyticity r:

If $f(z) = X(x,y) + iY(x,y)$

is an analytic function of z in a region D , it has at each point of D a unique derivative

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

The derivative will not exist if f(z) is non-analytic. We have however, the following theorem. (See [9])

T. (i): Let $X(x,y)$ and $Y(x,y)$ be continuous and have continuous partial derivatives of the first order near $z_0 = x_0 + iy_0$ and let

$$w(\lambda) = u(\lambda) + iv(\lambda)$$

$$= \lim_{h\to 0}^{\lambda} \frac{f(z_0 + h) - f(z_0)}{h}$$

where $\lim\limits_{h\to 0}^{\lambda}$ denotes that $h\to 0$ along a line of slope λ. Then the point $w(\lambda)$ lies on the circle

$$\left[u - \frac{1}{2}(X_x^0 - Y_y^0) \right]^2 + \left[v - \frac{1}{2}(Y_x^0 + X_y^0) \right]^2 = r^2(z_0) \tag{5.9}$$

where $r(z_0) = \frac{1}{2}\left[(X_x^0 - Y_y^0)^2 + (Y_x^0 + X_y^0)^2 \right]^{1/2}$

and $X_x^0 = \left[\frac{\partial}{\partial x} X(x,y) \right]_{x=x_0,\ y=y_0}$ etc.

We define (5.9) as the derivative circle and its centre the derivative of $f(z)$ at $z=z_0$. We write $f'(z) = \frac{1}{2}(X_x + Y_y) + \frac{1}{2}i(Y_x - X_y)$

It is interesting to note that if $r(z)=0$, the Cauchy-Riemann differential equations are satisfied, and the function $f(z)$ is analytic. Thus we may define $r(z_0)$ as the non-analyticity of $f(z)$ at $z=z_0$, and its least upper bound in D, the non-analyticity of $f(z)$ in D.

We now have the following theorem by Szu-Hoa Min [9]:

T. (ii): IF $f(z)$ is of non-analyticity r in a simply connected region D containing the rectifiable Jordan-curve C, we have for any z inside C:

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(w)dw}{w-z} + kr$$

$$k \leqq 4\sqrt{2}\left(1 + \frac{A}{2\pi}\right)$$ where A is the area enclosed by C.

The following theorem from the analysis of complex numbers is well-known:

T. (iii): If $f(z) = X(x,y) + i\,Y(x,y)$ is continuous on the smooth bounded curve C, which is given by $z(t) = x(t) + i\,y(t)$, $\alpha \leq t \leq \beta$,

Then
$$\int_C f(z)dz = \int_\alpha^\beta f(\,z(t)\,).\,z'(t)dt.$$
$$= \int_\alpha^\beta X(x(t),y(t)\,).\,x'(t)dt - \int_\alpha^\beta Y(x(t),y(t)\,).\,y'(t)dt$$
$$+i\int_\alpha^\beta Y(x(t),y(t)\,).\,x'(t)dt + i\int_\alpha^\beta X(x(t),y(t)\,).\,y'(t)dt.$$

Apply T. (ii) and T. (iii) to the f(z) in question. Then:

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(w)dw}{w-z} + kr$$
$$= \frac{1}{2\pi i}\int_\alpha^\beta \frac{f(w(t)\,).\,w'(t)}{w(t) - (x+iy)}\,dt + kr$$
$$= \frac{1}{2\pi i}\int_\alpha^\beta \frac{Xu' - Yv' + i(\,Yu' + Xv'\,)}{(u-x)^2 + (v+y)^2}\,[\,(u-x) - i(v+y)\,]dt + kr$$
$$= \frac{1}{2\pi i}\int_\alpha^\beta \frac{1}{(u-x)^2+(v+y)^2}\,[\,(Xu' - Yv')(u-x) + (Yu'+Xv')(v+y) + i\{(Yu'+Xv')(u-x) - (Xu'-Yv')(v+y)\}\,]dt$$
$$= \frac{1}{2\pi i}\int_\alpha^\beta [\,U(x,y,t) + i\,V(x,y,t)\,]dt + kr$$

$$\frac{df}{dz} = \frac{1}{2\pi i}\int_\alpha^\beta [\,U_x + iV_x]dt + kr_x$$

$$r_x = \frac{1}{2}\,[(X_x - Y_y)^2 + (Y_x + X_y)^2\,]^{-1/2}.\,[(X_x - Y_y)(X_{xx} - Y_{yx}) + (Y_x + X_y)(Y_{xx} + X_{yx})\,]$$

These values for f(z), f'(z) can now be substituted in (5.3) to obtain the value of ξ.

For practical purposes the general algorithm of the k-th order given by (5.1) becomes very clumsy. It is therefore of grave importance to find iterative algorithms of higher orders ($k > 2$) of which the application is still worth while. The rest of this chapter will therefore be devoted to the construction of such algorithms.

<u>Lemma:</u> Given two functions $G_k(z,\xi)$, $G_{k+1}(z,\xi)$ analytic with reference to both z and ξ.

If $z_{n+1} = G_k(z_n, \xi)$ and $z_{n+1} = G_{k+1}(z_n, \xi)$ represent iteration algorithms with ξ as attractive fixed-point [ξ a root of f(z), analytic] and orders k(> 0 integer) and k+1 respectively;

further

$$\left.\frac{\partial G_{k+1}(z,\xi)}{\partial \xi}\right|_{z=\xi} = 0 \tag{5.10}$$

then $z_{n+1} = G_{k+1}[z_n, G_k(z_n,\xi)]$, $n = 0, 1, 2, \ldots.$

represents an algorithm of (k+1)-th order (at least)

<u>Proof:</u> Put $F(z) = G_{k+1}[z, G_k(z,\xi)]$.

Then $\dfrac{\partial F(z,\xi)}{\partial z} = \dfrac{\partial G_{k+1}}{\partial z} + \dfrac{\partial G_{k+1}}{\partial G_k}\dfrac{\partial G_k}{\partial z}$

$$\frac{\partial^n F(z,\xi)}{\partial z^n} = \frac{\partial^n G_{k+1}}{\partial z^n} + A_1 \frac{\partial G_k}{\partial z} + A_2 \frac{\partial^2 G_k}{\partial z^2} + \ldots\ldots + \frac{\partial G_{k+1}}{\partial G_k}\frac{\partial^n G_k}{\partial z^n}$$

$$= \frac{\partial^n G_{k+1}}{\partial z^n} + \ldots\ldots\ldots + \left[\frac{\partial G_{k+1}}{\partial z}\frac{\partial z}{\partial G_k} + \frac{\partial G_{k+1}}{\partial \xi}\frac{\partial \xi}{\partial G_k}\right]\frac{\partial^n G_k}{\partial z^n}$$

Thus $\left.\dfrac{\partial^n F}{\partial z^n}\right|_{z=\xi} = 0$, $n = 1, 2, \ldots\ldots, k.$

Therefore $F(z,\xi)$ is an algorithm of order k+1 (at least).

We have for the analytic function f(z) (root ζ , $f'(\zeta) \neq 0$), regular within a neighbourhood of ζ .

$$0 = f(\zeta) = f(z) + (\zeta - z)f'(z) + \frac{(\zeta - z)^2}{2} f''(z) + \ldots\ldots$$

Then we can write for a domain D in which this series converges and in which $f'(z) \neq 0$:

$$\zeta = z - \frac{f(z)}{f'(z)} - \frac{1}{f'(z)} \left[\frac{(\zeta - z)^2}{2!} f''(z) + \frac{(\zeta - z)^3}{3!} f'''(z) + \ldots\ldots \right] \qquad (5.11)$$

If we replace ζ on the left-hand side with z_{n+1} , and on the right-hand side z with z_n , we obtain

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} - \frac{1}{f'(z_n)} \left[\frac{(\zeta - z_n)^2}{2!} f''(z_n) + \frac{(\zeta - z_n)^3}{3!} f'''(z_n) + \ldots \right] \qquad (5.12)'$$

which gives an algorithm producing root ζ of f(z) = 0 after one application. ((Due to the presence of the unknown quantity ζ on the right-hand side , equation (5.12)' does not make any sense as an algorithm in practice of course.) We note that a break after the second term in the series (5.12)' gives the Newton-algorithm. A break after the term with $(\zeta - z_n)^k$, $k \geqq 2$ as a factor gives an algorithm of (k+1)-th order.

<u>T. 14:</u> Given f(z) regular in a domain D with ζ , a root of f(z)= 0 as interior point, and $f'(z) \neq 0$ for $z \in D$. Then

$$F(z,\zeta) = z - \frac{f(z)}{f'(z)} - \frac{1}{f'(z)} \left[\frac{(\zeta - z)^2}{2!} f''(z) + \ldots\ldots + \frac{(\zeta - z)^k}{k!} f^{(k)}(z) \right] \qquad k \geqq 2 \qquad (5.12)$$

is an algorithm of (k+1)-th order.

<u>Proof:</u> (5.12) - (5.11): $F(z,\zeta) - \zeta = \frac{1}{f'(z)} \frac{(\zeta - z)^{k+1}}{(k+1)!} f^{(k+1)}(\tilde{z})$

where $\tilde{z}$ is an intermediate point lying within a domain containing z and ζ .

From this follows: $F(z,\zeta) - \zeta = O(|z - \zeta|^{k+1})$, $z \to \zeta$.

Q:E:D:

As we have already remarked, (5.12) is unsuited for practical purposes. However, if we have an algorithm of the k-th order, $z_{n+1} = G_k(z_n)$ in which the quantity ξ does not appear, we can, by the Lemma above obtain a (k+1)-th order algorithm (void of ξ) by replacing in (5.12) ξ with $G_k(z)$. In this recurring way one can obtain algorithms of arbitrary order according to the prescription :

$$\left.\begin{aligned} F_2(z) &= z - \frac{f(z)}{f'(z)} \\ F_{k+1}(z) &= z - \frac{f(z)}{f'(z)} - \frac{1}{f'(z)}\left[\frac{(F_k(z)-z)^2}{2!} f''(z) + \dots\dots \right. \\ &\qquad \left. \dots\dots + \frac{(F_k(z)-z)^k}{k!} f^{(k)}(z) \right], \quad k \geq 2. \end{aligned}\right\} \quad (5.13)$$

For example:

$$F_3(z) = z - \frac{f(z)}{f'(z)} - \frac{f(z)^2 f''(z)}{2f'(z)^3}$$

$$F_4(z) = z - \frac{f(z)}{f'(z)} - \frac{1}{f'(z)}\left[\frac{f''(z)}{2}\left(\frac{f}{f'} + \frac{f^2 f''}{2(f')^3} \right)^2 - \frac{f'''}{6}\left(\frac{f}{f'} + \frac{f^2 f''}{2(f')^3} \right)^3 \right]$$

$$= z - \frac{f}{f'} - \frac{f^2 f''}{2(f')^3} + \frac{f^3}{6(f')^5}\left[f' f''' - 3(f'')^2 \right] + \frac{f^4 f''}{4(f')^6}\left[- \frac{(f'')^2}{2f'} + \frac{f f'' f'''}{2(f')^2} + f''' + \frac{f''' f^2 (f'')^2}{12(f')^4} \right]$$

etc.

In comparing (5.1) with (5.13) it is obvious that not only the derivation of the algorithms is much easier in the latter instance, but also the practical application thereof. Incidentally, the algorithm of (k+1)-th order as given by (5.13) will be extremely expedient for computation by machine.

T. 15: Given f(z) regular in a domain D with ζ as interior point. $f(\zeta) = 0$, $f'(\zeta) \neq 0$. Then the following procedure will give an algorithm of (k+1)-th order (at least):

If z_n is an approximation of ζ , then first compute the values of:

$$\frac{1}{j!}\frac{d^j f(z_n)}{dz^j} = \frac{f_n^{(j)}}{j!} \quad , \quad j = 0, 1, \ldots\ldots, k.$$

and afterwards

$$v_{n,2} = -\frac{f_n}{f'_n} \quad , \quad v_{n,3} = -\frac{1}{f'_n}\left[f_n + v_{n,2}^2 \frac{f''_n}{2!}\right] ,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$v_{n,k+1} = -\frac{1}{f'_n}\left[f_n + v_{n,k}^2 \frac{f''_n}{2!} + v_{n,k}^3 \frac{f'''_n}{3!} + \ldots\ldots\ldots + v_{n,k}^k \frac{f_n^{(k)}}{k!}\right]$$

Then the new approximation is

$$z_{n+1} = z_n + v_{n,k+1} \quad .$$

Proof: Since $f'(\zeta) \neq 0$, $z_{n+1} = z_n + v_{n,2}$ (i. e. the N-R algorithm) will be of order 2.

If $z_{n+1} = z_n + v_{n,r}$ is an algorithm of order r for $2 \leq r \leq k$, then it follows from the Lemma and T. 14 above that

$$z_{n+1} = z_n - \frac{1}{f'_n}\left[f_n + \frac{v_{n,r}^2}{2!} f''_n + \ldots\ldots + \frac{v_{n,r}^r}{r!} f_n^{(r)} \right]$$

is an algorithm of at least order (r+1). (Replace in (5. 12) ζ with $v_{n,r} + z_n$.)By putting succesively r= 2, 3, ... , k the claim is proved.

Note: In the real case the existence of a continuous (k+1)-th derivative of f(z) in the vicinity of ζ is required. The existence of a continuous k-th derivative alone will not necessarily suffice to produce an algorithm of (k+1)-th order. For the evaluation of the coefficients $\frac{f_n^{(j)}}{j!}$ in case of polynomials , Horner' s scheme is proposed.

The following scheme might be useful in computing these higher order algorithms : (v_r denotes $v_{n,r}$)

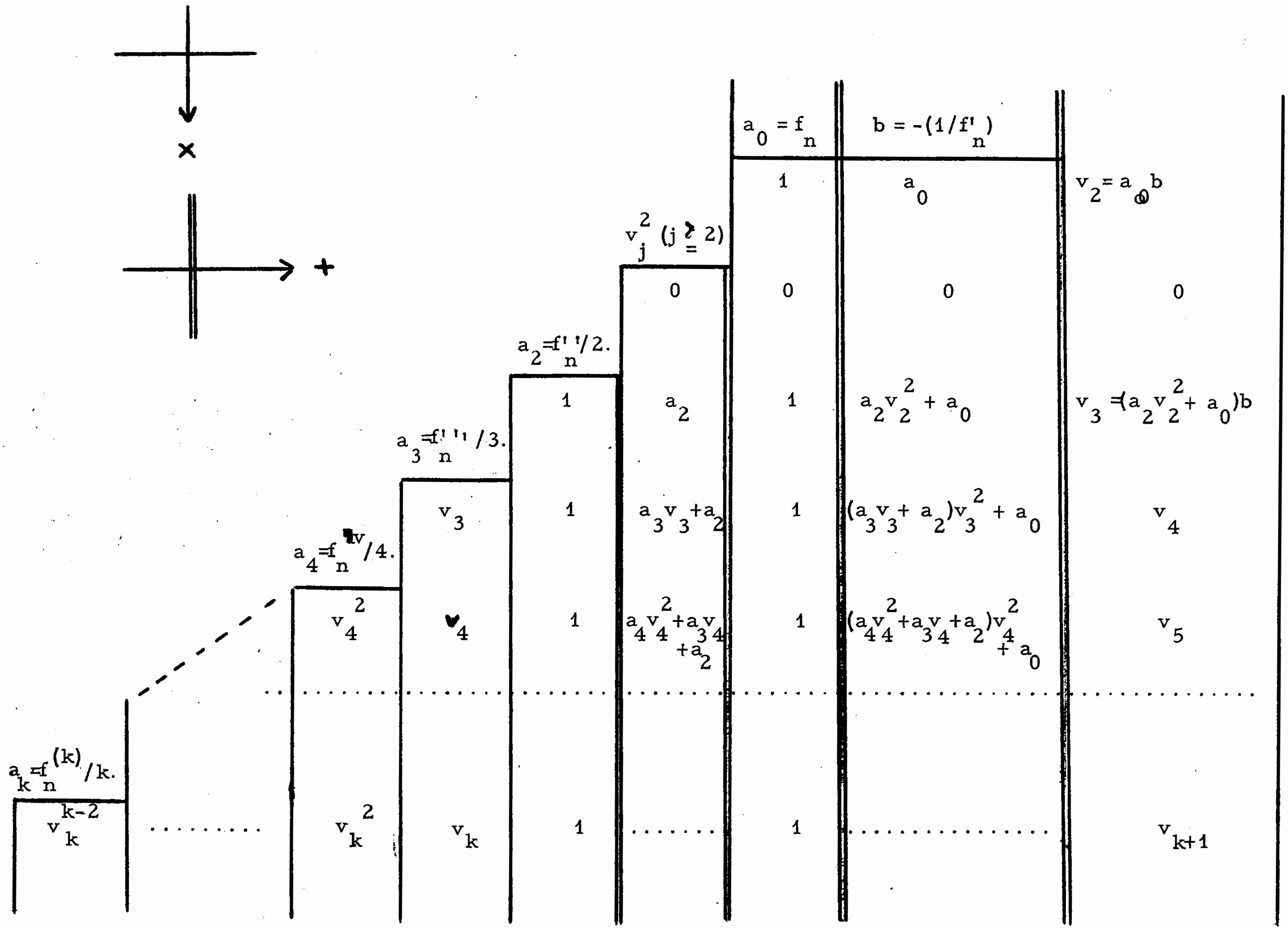
×
+
a0 = fn
b = -(1/f'n)
1
a0
v2 = a0b
v_j^2 (j ≥ 2)
0
0
0
0
a2 = f''n/2.
1
a2
1
a2v2^2 + a0
v3 = (a2v2^2 + a0)b
a3 = f'''n/3.
v3
1
a3v3 + a2
1
(a3v3 + a2)v3^2 + a0
v4
a4 = fn^iv/4.
v4^2
v4
1
a4v4^2 + a3v4 + a2
1
(a4v4^2 + a3v4 + a2)v4^2 + a0
v5
ak = fn^(k)/k.
vk^(k-2)
vk^2
vk
1
1
vk+1

Fig. 4.

(A further column can be introduced at the extreme right for tabulating the algorithms , i. e. $z_{n+1} = z_n + v_r$.)

If f(z) is a polynomial , Horner's scheme for evaluating the a_j can now be fitted in the blank upper left half of this scheme , since the a_j appear in Fig. 4 in the exact positions in which they will appear after application of the Horner-scheme. It should be noted however, that f'_n in the Horner-scheme must be replaced by v_j^2 , j= 2, 3, ... , k.

VI

Accelerating Iterations with Superlinear Convergence

T16: Given an algorithm of order $k > 1$ which yields after n applications

$$\frac{|z_n - \xi|}{|z_{n-1} - \xi|^k} \xrightarrow[\infty]{n} A \quad , \quad A \neq 0 \, , \quad A \neq \infty$$

We then claim that the approximation to ξ will be improved if z_n is replaced by

$$Z = z_n - \frac{|z_{n-1} - z_n|^{k+1}}{|z_{n-2} - z_{n-1}|^k} \cdot \operatorname{sgn} (z_n - \xi) \tag{6.1}$$

where $\operatorname{sgn} a = \frac{a}{|a|}$ (obviously defined only for $a \neq 0$)

Proof: We can obviously write

$$|z_n - \xi| = A \, |z_{n-1} - \xi|^k (1 + \varepsilon_n) \quad \text{and}$$

$$|z_{n-1} - \xi| = A \, |z_{n-2} - \xi|^k (1 + \varepsilon_{n-1}) \tag{6.2}$$

where $\begin{matrix} \varepsilon_n \\ \varepsilon_{n-1} \end{matrix} \xrightarrow[\infty]{n} 0$

Put $$\Delta_n = \max (|\varepsilon_n| , |\varepsilon_{n-1}|, |z_{n-2} - \xi|^{k-1}) \tag{6.3}$$

We shall now prove that

$$\frac{|z_n - \xi|}{A^{k+1} |z_{n-2} - \xi|^{k^2}} \xrightarrow[\infty]{n} 1 \tag{6.4}$$

and

$$\frac{|Z - \xi|}{A^{k+1} |z_{n-2} - \xi|^{k^2}} = O(\Delta_n) \quad , \quad n \to \infty \tag{6.5}$$

Putting $|z_{n-2} - \zeta| = \delta$

we have by (6.2) and (6.3)

$$\frac{|z_{n-2} - z_{n-1}|}{\delta} = 1 + O\left(\frac{z_{n-1} - \zeta}{\delta}\right) = 1 + O(\delta^{k-1}) = 1 + O(\Delta_n) \text{ , } n \to \infty$$

$$|z_{n-2} - z_{n-1}|^k = \delta^k [1 + O(\Delta_n)] \qquad (6.6)$$

Again by (6.2) and (6.3), as $n \to \infty$:

$$|z_{n-1} - z_n| = |z_{n-1} - \zeta| \left| 1 - \frac{z_n - \zeta}{z_{n-1} - \zeta} \right|$$

$$= A \delta^k [1 + O(|z_{n-1} - \zeta|^{k-1})] (1 + \varepsilon_{n-1})$$

$$= A \delta^k [1 + O(\delta^{k-1})][1 + O(\Delta_n)] \qquad (6.7)$$

From (6.6) and (6.7) we have

$$\frac{|z_n - z_{n-1}|^{k+1}}{|z_{n-1} - z_{n-2}|^k} = A^{k+1} \delta^{k^2} [1 + O(\Delta_n)] \text{ , } \quad n \to \infty \qquad (6.8)$$

On the other hand, after two applications of (6.2), we obtain

$$|z_n - \zeta| = A |z_{n-1} - \zeta|^k [1 + O(\Delta_n)]$$

$$= A^{k+1} \; k^2 [1 + O(\Delta_n)] \text{ , } \quad n \to \infty \qquad (6.9)$$

and from this (6.4) follows immediately.

Put $S_n = \operatorname{sgn}(z_n - \zeta)$

Then from (6.8) and (6.9):

$$S_n \frac{|z_{n-1} - z_n|^{k+1}}{|z_{n-2} - z_{n-1}|^k} = S_n A^{k+1} \delta^{k^2} [1 + O(\Delta_n)]$$

$$z_n - \zeta = S_n A^{k+1} \delta^{k^2} [1 + O(\Delta_n)] \text{ , } \quad n \to \infty$$

Subtract these. Then

$$\frac{|z - \xi|}{A^{k+1}\ \delta^{k^2}} = O(\Delta n) \qquad , \qquad n \to \infty$$

which is (6.5)

In studying the results of the theorems dealing with the error estimates of the algorithms (e.g. T's 3, 5, 7 etc.) we observe that usually

$$\varepsilon_{n-1} = O(\delta^{p}) \qquad p \geq 1$$

[e.g. in T. 3 we had

$$\frac{|\xi - z_{n-1}|}{|\xi - z_{n-2}|^2} = \left| \frac{f''(\eta)}{2\ f'(z_{n-2})} \right| \qquad , \quad \eta \text{ an intermediate point.}$$

Now $''(\eta)$ can be developed in terms of $(\xi - \eta)$, and then we have

$$\frac{|\xi - z_{n-1}|}{|\xi - z_{n-2}|^2} = \left| \frac{f''(\xi)}{2\ f'(z_{n-2})} \right| \left[1 + O(|\xi - \eta|)\right]$$

$$= \frac{|f''(\xi)|}{2m} \left[1 + O(|\xi - z_{n-2}|)\right]$$

$$= A\left[1 + O(\delta)\right] = A\left[1 + \varepsilon_{n-1}\right] \qquad]$$

If we put min. (p, k-1) = d we have obviously

$\Delta_n = O(\delta^{d})$, and (6.5) can be replaced by

$$|z - \xi| = O(\delta^{k^2+d}) \qquad (6.10)$$

We usually have d = 1, and then the use of (6.1) gives an improvement of 25% for k = 2, and

11.1% for k = 3. $\left[\,|z_n - \xi| = O(\delta^{k^2})\,\right]$.

<u>T17</u>: Given an algorithm of order <u>$k>2$</u> which yields after n applications

$$\frac{|z_n - \xi|}{|z_{n-1} - \xi|^k} = A + O(|z_{n-1} - \xi|) \quad , \quad A \neq 0 , \quad A \neq \infty \quad , \quad n \to \infty \qquad (6.11)$$

we then claim that the approximation to ξ will be improved if z_n is replaced by

$$Z* = z_n - A \,|z_{n-1} - z_n|^k \;\mathrm{sgn}\,(z_n - \xi) \qquad (6.12)$$

Further, Z^* will even be a better approximation that Z (see T. 16).

<u>Proof</u>: Since
$$\frac{z_{n-1} - z_n}{z_{n-1} - \xi} = 1 - \frac{z_n - \xi}{z_{n-1} - \xi}$$
$$= 1 - O(|z_{n-1} - \xi|^{k-1}) \quad , \quad n \to \infty$$

$$\therefore \quad \frac{z_{n-1} - \xi}{z_{n-1} - z_n} = 1 + O(|z_{n-1} - \xi|^{k-1}) \quad , \qquad n \to \infty$$

Together with (6.11) this gives

$$|z_n - \xi| = A\,|z_{n-1} - \xi|^k + O(|z_{n-1} - \xi|^{k+1})$$
$$= A\,|z_{n-1} - z_n|^k + O(|z_{n-1} - \xi|^{k^2-k})$$
$$+ O(|z_{n-1} - \xi|^{k+1}) \qquad (6.13)$$
$$= A\,|z_{n-1} - z_n|^k + O(|z_{n-1} - \xi|^{k+1}) \quad , \quad k>2$$

$$\therefore \quad z_n - \xi = A\,|z_{n-1} - z_n|^k \;\mathrm{sgn}\,(z_n - \xi) + O(|z_{n-1} - \xi|^{k+1})$$

Therefore, if we choose Z^* as in (6.12) we have

$$|Z^* - \xi| = O(|z_{n-1} - \xi|^{k+1}) = O(\delta^{k^2+k})$$

Thus Z^* is a better approximation to ξ than z_n since $k^2 + k > k^2$. $(k>0)$

It is also a better approximation than Z since $k^2 + d \leq k^2 + k - 1 < k^2 + k$

Note: In the case of k = 2 , nothing is gained by replacing z_n with Z^*. It is obvious from (6.13) that for k = 2 ,

$$|z^* - \xi| = O(\delta^4)$$

On the other hand $$|z_n - \xi| = O(\delta^4)$$

and $$|z - \xi| = O(\delta^5)$$

VII.

The Choice of a suitable order.

It has often been pointed out that for practical purposes the iterative use of one of the lower order algorithms (discussed in chapters III ; IV) is usually much more expeditious than the application of algorithms of higher order.

[The following problem was solved , first by means of Newton's method and afterwards by application of the 4th.. order algorithm of the Schröder-type (see equations (5. 3) and (5. 4))

Problem: Find the real root of the equations

$x^5 + x^3 = A$ with $A = 1, 2, 3, \ldots\ldots 100.$ correct to 3 decimal places.

An IBM 650 digital computor was used. In the first instance about 5 minutes of computor time was sufficient, whilst in the application of the 4th. order algorithm , the machine required 9 minutes.

(In both instances the same initial approximation for A=1 namely $x_0 = .8$ was used.)

A:

We will first devote our attention to the higher order algorithms as established by T. 15 and will also try to develop a criterium for deciding which order would be the best suited for special cases.

After a glance at Fig. 4 it will be evident that the number of multiplications and divisions together involved in obtaining an algorithm (of the type (5. 13)) of order $k \geq 2$, is equal to $\frac{k(k+1)}{2} - 2$ (i. e. without consideration of the calculation of the a_j). Thus , the "calculation energy " (i. e. the sum of the number of multiplications and divisions) increases in direct proportion with k^2 for increasing k. On the other hand , according to P. 1 this "energy" increases in proportion with log k, if an algorithm of lower order k* is used iteratively to obtain an algorithm of order k.

[$k^{*^r} = k$, Hence $r = \frac{\log k}{\log k^*}$

Therefore, Energy=Const. x r =Const. $\frac{\log k}{\log k^*}$]

(Note. Natural logarithms will be used throughout this chapter.) Further, the number of coefficients $f_n, f'_n, \ldots\ldots, f_n^{(k-1)}/(k-1)!$ to be calculated in the first instance is equal to k. In case of the iterative use of an algorithm order k^*, we have this number as equal to $\frac{k^* \log k}{\log k^*}$.

[In case of the algorithm of order $k \geq 2$ as given by (5.4), the number of calculations is in any case of an order much higher than k. Here computations of derivatives are involved. This usually takes more time than ordinary multiplication or division. Here the number of "coefficients" $f^{(r)}$ to be calculated is also k.]

From these remarks it is evident that there is in the general case no sense at all in arbitrarily increasing the order of an algorithm, since the increase of convergence speed is usually obliterated by the increase in "calculation energy".

<u>T. 18:</u> If the number of multiplications and divisions involved in the computation of each of the values $f(z_n)$, $f'(z_n)$, $f''(z_n)/2!$, ... $f^{(k-1)}(z_n)/(k-1)!$ is approximately equal (let this number be a), then

$$\text{for} \quad a < a^* = \frac{\log \frac{16}{3}}{\log \frac{9}{8}} \cong 14.213 \qquad \text{the order } r=2 \text{, and}$$

$$\text{for} \quad a > a^* \quad , r=3$$

are the best lower order algorithms for iterative use to produce an algorithm of order k. (i. e. considering the algorithms established by T. 15)

<u>Proof:</u> In the iterative use of an r-th. order algorithm of the type under consideration, r such "coefficients" are to be calculated. Thus, after n applications of this algorithm s = anr multiplications and divisions were carried through in the calculation of the

$$f(z_j)\ ,\ f'(z_j)/1\ ,\ \ldots\ldots\ ,\ f^{(r-1)}(z_j)/(r-1)!\quad ,\ j = 0, 1, \ldots\ldots, n-1.$$

Thus, by the remarks above, the total amount of multiplications and divisions will be

$$E = anr + n\left[\frac{r(r+1)}{2} - 2\right]$$

where according to P. 1 $k = r^n$.

$$\text{i.e.}\quad E = \frac{\log k}{2\log r}\left[r^2 + (2a+1)r - 4\right] \tag{7.1}$$

Keeping k constant, and considering E as a continuous function of the real variable r, we have:

$$\frac{dE}{dr} = \frac{\log k}{2r(\log r)^2}\left[(2r^2 + \overline{2a+1}\cdot r)(\log r - 1) + r^2 + 4\right]$$

hence $\quad \dfrac{dE}{dr} > 0$ for $r \geq 3$ (Since $\log r > 1$ for $r \geq 3$)

$$\text{i.e.}\quad E(3) < E(r') \quad \text{for } r' = 4, 5, \ldots\ldots \tag{7.2}$$

Thus the "computation energy" required for the iterative use of algorithms of order higher than 3 is greater than that required for an algorithm of order 3. The cases $r=2$ and $r=3$ remain to be compared. But from (7.1):

$$E(2) = \log k\,\frac{2a+1}{\log 2}\quad ;\quad E(3) = \log k\,\frac{3a+4}{\log 3}$$

Therefore

$$E(2) \lesseqgtr E(3) \quad \text{for} \quad a \lesseqgtr \frac{4\log 2 - \log 3}{2\log 3 - 3\log 2} = \frac{\log\frac{16}{3}}{\log\frac{9}{8}} = a^*$$

Q. E. D.

In the theorem above we have studied the iterative use of an algorithm of order $r \geq 2$ which will produce, after n applications, an algorithm of order $k = r^n$. In the proof above we have, however also allowed non-integral k, i. e. not of the form $k = r^n$. If this might be a cause for anxiety we propose the following theorem:

<u>T. 19:</u> We have the "computation energy" E, (of multiplications and divisions involved) in the construction of an algorithm of order $k = r^n$ by the n-fold use of an iteration formula order $r \geq 2$, as given by

$$E = Q(r)\log k.$$

$Q(r)$ is defined for integral values of $r \geq 2$, for which $Q(r)$ assumes positive values. If it is further given that for a $r_1 \geq 2$, and all $r_j \geq 2$, $r_j \neq r_1$, $Q_1 = Q(r_1) < Q(r_j) = Q_j$,

then it follows from $k_1 = r_1^{n_1} \leq r_j^{n_j} = k_j$

that $E_1 < E_j$.

Further, there exists for every $j \neq 1$ integers n_1 and n_j such that

$$k_1 = r_1^{n_1} > r_j^{n_j} = k_j \quad \text{and} \quad E_1 < E_j .$$

<u>Proof:</u> From $Q_1 < Q_j$ and $k_1 \leq k_j$ we have

$$E_1 = Q_1 \log k_1 < Q_j \log k_j = E_j$$

To prove the second part of T. 19, we choose a rational number $\frac{n_1}{n_j}$ such that

$$\frac{Q_j \log r_j}{Q_1 \log r_1} > \frac{n_1}{n_j} > \frac{Q_1 + Q_j}{2Q_1} \cdot \frac{\log r_j}{\log r_1}$$

(This is possible since $Q_j > Q_1 > 0$, and $r_1, r_j \geq 2$.)

From this follows

$$\frac{n_1 \log r_1}{n_j \log r_j} = \frac{\log k_1}{\log k_j} > \frac{Q_1 + Q_j}{2Q_1} > 1 \quad \text{, i.e.} \quad k_1 > k_j \;.$$

and also

$$\frac{E_j}{E_1} = \frac{Q_j n_j \log r_j}{Q_1 n_1 \log r_1} > 1 \;. \qquad \text{Q. E. D.}$$

Thus far we have only considered cases where the "computation energy" required for every successive derivative $f^{(j)}(z)$ is approximately the same. In the case of polynomials however, this energy decreses for increasing j - a fact which is immediately evident from Horner's scheme. In this case we have:

T. 20: For the iterative solution, according to (5. 13), of an algebraic equation of the n-th degree, we have in consideration of the computation energy, the Newton-algorithm for $n = 1, \ldots\ldots, 10$ and the 3rd. order algorithm for $n > 10$ as the best choices.

Proof: According to Horner's scheme and the scheme in Fig. 4, the number of multiplications and divisions to be performed in applying an algorithm order r of type (5. 13) to an algebraic equation of the n-th degree, will be:

$$E^* = G_1 + G_2 + G_3$$

$$= \left[\frac{r(r+1)}{2} - 2 \right] + \left[\sum_{j=1}^{r} (n+1 - \overline{j-1}) \right] + [r-2]$$

$$= r(n+3) - 4.$$

[G_1 is due to the work done in Fig. 4

G_2 " " " " " " " Horner's scheme for obtaining $f^{(j)}(z_n)$

G_3 " " " " divisions required for obtaining $f^{(j)}(z_n)/j!$]

Thus, after m iterative applications of this r-th order algorithm, we have as total energy for producing a k-th order algorithm:

$$E = mE^* = m[\, r(n+3) - 4 \,],$$ where, according to P:1 $k = r^m$. Hence:

$$E = \frac{\log k}{\log r} [\, r(n+3) - 4] = Q(r) \log k.$$

According to T. 19, the best choice for r will be that integer $r \geq 2$ for which $Q(r)$ has its smallest value. Then, if k is kept constant and $Q(r)$ assumed as being a continuous function of r, we have:

$$\frac{dE}{dr} = \frac{\log k}{r (\log r)^2} [\, r(n+3)(\log r - 1) + 4 \,]$$

Therefore, for $r \geq 3$, $\frac{dE}{dr} > 0$.

Thus , as in T. 18 the only cases left for consideration are $r = 2$ and $r = 3$.

$$E(2) = \log k \; \frac{2n+2}{\log 2} \quad ; \quad E(3) = \log k \; \frac{3n+5}{\log 3}$$. Therefore

$$E(2) \lesseqgtr E(3) \quad \text{for} \quad n \lesseqgtr \frac{5\log 2 - 2\log 3}{2\log 3 - 3\log 2} = \frac{\log \frac{32}{9}}{\log \frac{9}{8}} \cong 10.78$$

There still remain some possibilities of f(z) which are not covered by T.'s 18 and 20. These are the cases where the energy required for the calculation of the $f^{(j)}(z_n)/j!$ differs greatly for different j.
The following theorem might be of some help in deciding on the best order r.

T. 21: Let the "computation energy" (i. e. the number of multiplications and divisions together) for $f(z)$, $f'(z)$, $f''(z)/2$, ... , $f^{(r-1)}(z)/(r-1)!$ respectively be given by $e_o, e_1, e_2, \ldots\ldots e_{r-1}$.

Then the "energy" required for the production of a k-th order algorithm of type (5. 13) by iterative use of one of order $r \geq 2$, will be given by:

$$E = \frac{\log k}{2} \; \frac{r(r+1) + 2\{e_0 + e_1 + \ldots\ldots + e_{r-1}\} - 4}{\log r}$$

The best choice for r will then be that $r \geq 2$ (integral) for which

$$Q = \frac{r(r+1) + 2(e_o + e_1 + \ldots\ldots + e_{r-1}) - 4}{\log r}$$

assumes its smallest value.

Proof: For one application of the algorithm order r, we have

$$E^* = \frac{r(r+1)}{2} - 2 + e_0 + e_1 + \ldots\ldots + e_{r-1}$$, and for n applications

$$E = nE^* = \frac{\log k}{2} \; \frac{r(r+1) + 2(e_0 + e_1 + \ldots\ldots + e_{r-1}) - 4}{\log r}$$

The rest follows from T. 19.

B: Considering algorithms of the type (5.4), we have the following:

T. 22: Let the "computation energy" for $f(z)$, $\delta^0 f(z)$, $\delta^1 f(z), \ldots, \delta^{r-2} f(z)$ respectively be given by $e, e_o, e_1, \ldots, e_{r-2}$.

Then the "energy" required for the production of a k-th order algorithm of type (5.4) by iterative use of one of order $r \geq 2$, will be given by:

$$E = \frac{\log k}{\log r}[2r - 3 + e + e_0 + e_1 + \ldots + e_{r-2}],$$

and the best choice for r will then be that $r \geq 2$ (integral) for which

$$Q = \frac{1}{\log r}[2r - 3 + e + e_0 + e_1 + \ldots + e_{r-2}]$$

assumes its smallest value.

[Similarly as before, e_j denotes the number of multiplications and divisions necessary to obtain $\delta^j f(z)$ from $\delta^{j-1} f(z)$.]

Proof: We have from (5.4)

$$F_r(z) = z + \sum_{j=1}^{r-1} (-1)^j \frac{f(z)^j}{j!} \{ \delta^{j-1} f(z) \}$$

Hence for one application we have

$$E^* = e + r - 2 + r - 1 + e_0 + e_1 + \ldots + e_{r-2}$$

and for n applications:

$$E = nE^* = \frac{\log k}{\log r}[2r - 3 + e + e_o + e_1 + \ldots + e_{r-2}]$$

since $k = r^n$. The rest follows from T. 19.

VIII.

Error Estimates of the higher order algorithms ($k \geqq 3$).

From the definition of the order of an algorithm (see capter II) we have , that for an algorithm of order k , there exists a constant C_k such that

$$\frac{| F(z) - \xi |}{| z - \xi |^k} \leqq C_k$$

holds in the vicinity of the solution ξ of z = F(z) or f(z) = 0. This C_k will obviously be dependent on the derivatives of F(z) , i. e. dependent on the derivatives of f(z). If an explicit form for C_k can be found , it may be very useful as an error estimate for the k-th order algorithm under consideration.

In capters III and IV we have already fully discussed the error estimates of the second (and one 3rd) order algorithms which were mentioned. Our object is now to find error estimates for algorithms with $k \geqq 3$.

A: We will first consider the higher order algorithms as established by T. 15 and Fig. 4 :

From (5. 11) we have for $k \geqq 3$

$$\xi = z_n - \frac{f_n}{f'_n} - \frac{1}{f'_n} \left[\sum_{r=2}^{k-1} \frac{(\xi - z_n)^r}{r!} f_n^{(r)} + R_{k,n} \right] \tag{8. 1}$$

where $R_{k,n} = \frac{(\xi - z_n)^k}{2\pi i} \int_C \frac{f(t)}{(t - \xi)^{k+1}} dt$

$$= \frac{(\xi - z_n)^k}{k!} f^{(k)}(\xi) \tag{8. 2}$$

(C is any circle , centre ξ such that f(z) is analytic within and on C .)

On the other hand we have from (5.13):

$$z_{n+1} = z_n - \frac{f_n}{f'_n} - \frac{1}{f'_n}\left[\sum_{r=2}^{k-1} \cdot \frac{(\tilde{z}_{n+1} - z_n)^r}{r!} \quad f_n^{(r)} \right] \tag{8.3}$$

where $\tilde{z}_{n+1} = F_{k-1}(z_n)$, F_{k-1} being a similar algorithm of order k-1. i.e.

$$\tilde{z}_{n+1} - \xi = g_{k-1}(z_n)\,(z_n - \xi)^{k-1}$$

where $g_{k-1}(z)$ is finite in a vicinity of ξ .

If we write $g_{k-1} = g_{k-1}(z_n)$, we have for $k \geqq 3$

$$\begin{aligned}(\tilde{z}_{n+1} - z_n)^r &= [\,(\tilde{z}_{n+1} - \xi) + (\xi - z_n)\,]^r \\ &= (\xi - z_n)^r\,[\,1 - g_{k-1}\cdot(z_n - \xi)^{k-2}\,]^r \\ &= (\xi - z_n)^r\,[\,1 + \sum_{j=1}^{r} (-1)^j \left[{r \atop j} \right] g_{k-1}^j\,(z_n - \xi)^{j\cdot(k-2)}]\end{aligned} \tag{8.4}$$

Then from (8.1) , (8.3) and (8.4) we obtain for $k \geqq 3$:

$$z_{n+1} - \xi = -\frac{1}{f'_n} \sum_{r=2}^{k-1} \left[\frac{f_n^{(r)}}{r!} \sum_{j=1}^{r} (-1)^{r-j} \left[{r \atop j} \right] g_{k-1}^j\,(z_n - \xi)^{j(k-2)+r} \right]$$

$$+ \quad \frac{R_{k,n}}{f'_n} \tag{8.5}$$

Since $j \geqq 1$, $k \geqq 3$, $r \geqq 2$,

$j(k-2) + r \geqq k.$

Thus (8.2) and (8.5) together give:

$$z_{n+1} - \xi = O(\,|\,z_n - \xi\,|^k\,)$$

Thus equation (8.5) can serve as basis for an error estimate. Let us first consider the most important (according to the previous chapter) case k = 3:

From (8.2) and (8.5) we have then

$$z_{n+1} - \zeta = -\frac{1}{f'_n} \quad \frac{f''_n}{2} [- 2g_2 (z_n - \zeta)^3 + g_2^2 (z_n - \zeta)^4] + \frac{R_{3,n}}{f'_n}$$

$$= \frac{(\zeta - z_n)^3}{2f'_n} [-2f''_n \; g_2 + f''_n \; g_2^2 (z_n - \zeta) + \frac{1}{3} \; f'''(\zeta)]$$

We have from T. 3 :

$$g_2 = - \frac{\tilde{f}''}{2 f'_n}$$

where $\tilde{f}''$ implies that the function $f''(z)$ must be taken at an intermediate value. Then

$$z_{n+1} - \zeta = \frac{(\zeta - z_n)^3}{2f'_n} [\frac{f''_n \tilde{f}''}{f'_n} + \frac{f''_n \tilde{f}''^2}{4 f'^2_n} (z_n - \zeta) + \frac{1}{3} f''(\zeta)] \qquad (8.6)$$

T. 23: Given f(z) regular in a domain D which includes $|z - z_0| \leqq 2c$ where z_0 is the initial approximation to the root ζ of $f(z) = 0$,

$$c \geqq \frac{|f(z_0)|}{m} \quad , \quad m = \inf_{z \in D} |f'(z)| > 0$$

$$C c^2 < 1 ,$$

$$C = \frac{1}{2m^3} [mM_2^2 + \frac{1}{4} M_2^3 c + \frac{1}{3} m^2 M_3]$$

$$M_j = \sup_{z \in D} | f^{(j)}(z) | \qquad , \qquad j = 2, 3.$$

Then all z_n , n = 1, 2, ... obtained by the iterative use of the algorithm (established by T. 15) for k = 3 lie within the sub-domain

$| z - z_0 | \leqq 2c$ of D and $z_n \overset{n}{\underset{\infty}{\rightarrow}} \zeta$

where ζ is the only root of $f(z) = 0$ in D.

Further we have as error estimate:

$$|z_n - \zeta| \leq C^{\frac{1}{2}(3^n - 1)} \cdot c^{3^n} \quad , n = 0, 1, 2, \ldots\ldots$$

(It is interesting to note that here , in contrast with T. 4 we have to assume the existence of ζ within D.)

Proof:

We have $|f'(z)| \geq m > 0$ and $\frac{|f(z_0)|}{m} \leq c$

Since ζ and z_0 are both in D, we also have for an intermediate $\tilde{z}$

$$|f'(\tilde{z})| \geq m.$$

Therefore $\frac{|f(z_0)|}{|f'(\tilde{z})|} \leq c$ or $\frac{|f(z_0)|\,|\zeta - z_0|}{|f(z_0)|} \leq c$

i. e. $|\zeta - z_0| \leq c$

From (8. 6) we have:

$$|z_1 - \zeta| \leq |z_0 - \zeta|^3 \frac{1}{2m}\left[\frac{M_2^2}{m} + \frac{M_2^3}{4m^2} c + \frac{M_3}{3}\right]$$

$$= C\,|z_0 - \zeta|^3 \leq C\,c^3 < c$$

Then obviously

$$|z_n - \zeta| < c \text{ for } n = 0, 1, \ldots\ldots$$

and since $|z_n - z_0| \leq |z_0 - \zeta| + |z_n - \zeta| < 2c$

we conclude that all z_n lie in D, and

$$|z_{n+1} - \zeta| \leq C\,|z_n - \zeta|^3 \quad , \quad n = 0, 1, 2, \ldots\ldots$$

$$\leq C\,c^2\,|z_n - \zeta| < |z_n - \zeta|$$

Therefore $z_n \xrightarrow[\infty]{n} \zeta$. i. e. the algorithm converges.

Further $|z_n - \zeta| \leq C\,|z_{n-1} - \zeta|^3 \leq C^4 |z_{n-2} - \zeta|^9 \leq C^{13} |z_{n-3} - \zeta|^{27}$

$$\leq C^{(1/2)(3^n - 1)}\,|z_0 - \zeta|^{3^n}$$

$$\leq C^{(1/2)(3n - 1)} \cdot c^{3n} \qquad n = 0, 1, 2, \ldots\ldots\ldots$$

Q. E. D.

To obtain C_k for $k > 3$ is a little more tedious.
Since $j \geqq 1$, $r \geqq 2$, $k \geqq 3$, $j \leqq r$ we have

$$j(k-2) + r \geqq k-1+j.$$

Thus, if we know that already $|z_n - \zeta| \leqq 1$, we obtain from (8.7)
(8.2) and (8.5) for $k \geqq 3$:

$$|z_{n+1} - \zeta| \leqq \frac{|z_n - \zeta|^{k-1}}{|f'_n|} \left[\sum_{r=2}^{k-1} \frac{|f_n^{(r)}|}{r!} \sum_{j=1}^{r} \left[{r \atop j} \right] |g_{k-1}^j| \, |z_n - \zeta|^j \right]$$

$$+ \frac{|f^{(k)}(\zeta)|}{k! \, |f'_n|} |z_n - \zeta|^k$$

$$= \frac{|z_n - \zeta|^{k-1}}{|f'_n|} \left[\sum_{r=2}^{k-1} \frac{|f_n^{(r)}|}{r!} \left[(1 + |g_{k-1}| \, |z_n - \zeta|)^r - 1 \right] . \right]$$

$$+ \frac{|f^{(k)}(\zeta)|}{k! \, |f'_n|} |z_n - \zeta|^k \tag{8.8}$$

$$\leqq \frac{|z_n - \zeta|^{k-1}}{|f'_n|} \left[(1 + |g_{k-1}| \, |z_n - \zeta|)^{k-1} - 1 \right] \sum_{r=2}^{k-1} \frac{|f_n^{(r)}|}{r!}$$

$$+ \frac{|f^{(k)}(\zeta)|}{k! \, |f'_n|} |z_n - \zeta|^k$$

The sum $S_{k-1} = \sum_{r=2}^{k-1} \frac{|f_n^{(r)}|}{r!}$ is immediately obtainable from the scheme in Fig. 4.

Thus we have:

$$|g_k| = \frac{|z_{n+1} - \zeta|}{|z_n - \zeta|^k} \leqq \frac{1}{|f'_n|} \left[\frac{S_{k-1}}{|z_n - \zeta|} \left\{ (1 + |g_{k-1}||z_n - \zeta|)^{k-1} - 1 \right\} + \frac{|f^{(k)}(\zeta)|}{k!} \right], \quad k \geqq 3$$

If $m = \inf_{D} |f'(z)| > 0$, $M_j = \sup_{D} |f^{(j)}(z)|$ $j = 2, 3, \ldots\ldots, k.$

we obtain C_k in a recursive way from the formula

$$C_k = \frac{1}{m} \left[\frac{S^*_{k-1}}{|z_n - \zeta|} \left\{ (1 + C_{k-1}|z_n - \zeta|)^{k-1} - 1 \right\} + \frac{M_k}{k!} \right] \quad , \ k \geqq 3.$$

where $S^*_{k-1} = \sum_{r=2}^{k-1} \frac{M_r}{r!}$, $C_2 = \frac{M_2}{2m}$ (See T. 3)

Now we can state in general for $k \geqq 3$:

<u>T. 24:</u> Given $f(z)$ regular in the domain D which includes $|z - z_0| \leqq 2c$, where

$$1 > c \geqq \frac{|f(z_0)|}{m}$$

$$m = \inf_{D} |f'(z)| > 0$$

$$C_k c^{k-1} < 1$$

and $C_2 = \frac{M_2}{2m}$

$$C_j = \frac{1}{m} \left\{ \frac{S^*_{j-1}}{c} \left[(1 + C_{j-1} c)^{j-1} - 1 \right] + \frac{M_j}{j!} \right\} \qquad (8.9)$$

$$3 \leqq j \leqq k.$$

$$M_j = \sup_{z \in D} |f^{(j)}(z)| \quad , \ j = 2, 3, \ldots\ldots, k.$$

Then all z_n , $n = 1, 2, \ldots\ldots$ obtained by the iterative use of the k-th order algorithm as established by T. 15 ($k \geqq 3$) , lie within the sub-domain $|z-z_0| \leqq 2c$ of D , and $z_n \xrightarrow{n} \zeta$ where ζ is the only root of $f(z) = 0$ in D.

Further we have as error estimate :

$$|z_n - \zeta| \leqq C_k^{\frac{k^n - 1}{k-1}} \cdot c^{k^n} \quad , \quad n = 0, 1, 2, \ldots\ldots$$

(again, the existence of one root in D is assumed.)

Proof: Completely analog to that of T. 23.

Note:Since $c < 1$ (from (8.7)) , (8.9) can be replaced by the coarser recursion formula

$$C_j = \frac{1}{m}\left\{ S^*_{j-1} \; [(1 + C_{j-1})^{j-1} - 1] + \frac{M_j}{j!} \right\} \qquad 3 \leqq j \leqq k.$$

On the other hand, the "finest" formula would be

$$C_j = \frac{1}{m}\left\{ \frac{1}{c} \sum_{r=2}^{k-1} \frac{M_r}{r!} [(1 + C_{j-1}c)^r - 1] + \frac{M_j}{j!} \right\} \qquad 3 \leqq j \leqq k$$

as is evident from (8.8).

B: Considering the higher order algorithms as established by T. 12 and 13 , we obtain the following analogous results. (These are of course of much less practical importance than those considered above.)

We have from T. 12 :

$$\zeta = z_n - \frac{f_n}{f'_n} + \sum_{r=2}^{k-1} (-1)^r \frac{f(z_n)^r}{r!} \left[\frac{d^r f^{-1}(w)}{dw^r} \right]_{w=f(z_n)} + R_k$$

where

$$R_k = (-1)^k \frac{f(z_n)^k}{2\pi i} \int_C \frac{f^{-1}(t)}{(t - f(z_n))^{k+1}} \, dt \qquad (8.10)$$

(C is any circle , centre $f(z_n)$, such that $f^{-1}(w)$ is analytic within and on C.)

or

$$R_k = (-1)^k \frac{f(z_n)^k}{k!} \left[\frac{d^k}{dw^k} f^{-1}(w) \right]_{w=f(z_n)}$$

On the other hand , from (5.4):

$$z_{n+1} = z_n - \frac{f_n}{f'_n} + \sum_{r=2}^{k-1} (-1)^r \frac{f(z_n)^r}{r!} \left[\frac{d^r f^{-1}(w)}{dw^r} \right]_{w=f(z_n)} \qquad (8.11)$$

Equations (8.10) and (8.11) then give:

$$\zeta - z_{n+1} = (-1)^k \frac{f(z_n)^k}{k!} \left[\frac{d^k}{dw^k} f^{-1}(w) \right]_{w=f(z_n)}$$

If we put $f(z) = (\zeta - z)\psi(z)$ where $\lim_{z \to \zeta} |\psi(z)| < \infty$

we get

$$|g_k| = \frac{|z_{n+1} - \zeta|}{|z_n - \zeta|^k} \leqq \frac{|\psi(z_n)|^k}{k!} \left| \frac{d^k}{dw^k} f^{-1}(w) \right|_{w=f(z_n)}$$

Since $f'(\zeta) = -\psi(\zeta)$

$$|g_k| \leqq \frac{M^k}{k!} L_k \qquad (8.12)$$

where $M = \sup_{z \in D} |f'(z)|$

and $$L_k = \sup_{w \in D' = f(D)} \left| \frac{d^k}{dw^k} f^{-1}(w) \right|$$

T. 25: Given f(z) regular in the domain D which includes

$$|z - z_0| \leqq 2c \text{ , where}$$

$$c \geqq \frac{|f(z_0)|}{m} \quad , \quad m = \inf_{z \in D} |f'(z)| > 0$$

$$C_k c^{k-1} < 1$$

$$C_k = \frac{M^k}{k!} L_k \quad , \quad M = \sup_{z \in D} |f'(z)|$$

$$L_k = \sup_{w \in D' = f(D)} \left| \frac{d^k}{dw^k} f^{-1}(w) \right|$$

Then all z_n , n= 1, 2, ... obtained by the iterative use of the k-th order algorithm as established by (5.4) lie within the sub-domain $|z - z_0| \leqq 2c$ of D and $z_n \underset{\infty}{\overset{n}{\to}} \zeta$ where ζ is the assumed single root of f(z)=0 in D. Further we have as error estimate :

$$|z_n - \zeta| \leqq C_k^{\frac{k^n - 1}{k-1}} \cdot c^{k^n} \quad , \; n = 0, 1, 2, \ldots\ldots$$

Proof: (As before)

From $|f'(z)| \geqq m > 0$, $\dfrac{|f(z_0)|}{m} \leqq c$.

we conclude as before : $|\zeta - z_0| \leqq c$

Then from (8.12): $|z_1 - \zeta| \leqq |z_0 - \zeta|^k . C_k$

$$\leqq C_k c^k < c$$

and also $|z_n - \zeta| < c$

We conclude that all z_n lie in D since

$$|z_n - z_0| \leqq |z_0 - \zeta| + |z_n - \zeta| < 2c.$$

Further we have $|z_{n+1} - \zeta| \leqq C_k c^{k-1} |z_n - \zeta| < |z_n - \zeta|$

Hence $z_n \overset{n}{\underset{\infty}{\rightarrow}} \zeta$,

and $|z_n - \zeta| \leqq C_k |z_{n-1} - \zeta|^k \leqq C_k^{k+1} |z_{n-2} - \zeta|^{k^2} \leqq C_k^{k+1+k^2} |z_{n-3} - \zeta|^{k^3}$

$$\leqq C_k^{\frac{k^n - 1}{k-1}} \cdot c^{k^n} \qquad , n = 0.1, 2, \ldots\ldots.$$

Q. E. D.

IX

The Determination of the Approximate Location of the Roots of $f(z) = 0$.

It is evident that in the iterative application of the algorithms considered above, the choice of the initial approximation z_o to the root ξ of $f(z) = 0$ is a problem in itself. It would therefore be extremely useful if a means can be found of determining the approximate location of the roots of the analytic function under consideration.

[The problem of finding the location of the real roots of the function $f(x) = 0$ of the real variable x is not explicitly discussed in this chapter, and as a result no mention is made of such well-known criteria as the so-called Harriot-Descartes Rule of Signs, Sturm's Theorem etc. For these we wish to refer to most textbooks dealing with the subject, e.g. H. W. Turnbull [20].]

As was already remarked in Chapter I, the zeros of an analytic function are isolated points, i.e.

If a function $f(z)$ is not identically zero, and is analytic in a region including $z = a$, then there is a circle $|z - a| = m$ $(m > o)$ inside which $f(z)$ has no zeros except possibly $z = a$ itself.

A: The problem of determining the number N of zeros of an analytic function which lie in a given region was already solved by Cauchy by means of his theorem:

T (i): If f(z) is analytic within and on a closed contour (rectifiable Jordan curve) C, and $f(z) \neq 0$ on C, then

$$N = \frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)}\, dz$$

where N is the number of zeros inside the contour (a zero of order P being counted P times.)

This result can also be expressed in another way. Since

$$\frac{d}{dz} \log\left[f(z)\right] = \frac{f'(z)}{f(z)}$$

we have
$$\int_C \frac{f'(z)}{f(z)}\, dz = \Delta_C \log\left[f(z)\right]$$

where Δ_C denotes the variation of $\log\left[f(z)\right]$ round the contour C. The value of the logarithm with which we start is clearly indifferent. Also $\log\left[f(z)\right] = \log\left|f(z)\right| + i \arg\left[f(z)\right]$ and log $|f|$ is one-valued. Hence the formula may be written as

$$N = \frac{1}{2\pi} \Delta_C \arg\left[f(z)\right]$$

or better still; if we write $f(z) = r\, e^{2i\pi\theta}$, $r > 0$ on C

i.e. if $f(z) = u(z) + i\, v(z)$ where u(z) and v(z) are real on C,

then $\theta = \frac{1}{2\pi} \arc\tan\left[v(z)/u(z)\right]$

and $N = \Delta_C \arg\left[f(z)\right] = \int_C d\theta.$

Hence N is the amount that θ increases as the point z traverses the curve C in the positive sense.

Keeping Rauché's theorem (see chapter I) in mind might also be of some help for determining N in special cases. The following is an example of the type of problem which can be solved by means of Cauchy's method:

Que.: In which quadrants do the roots of the equation

$$f(z) = z^4 + z^3 + 4z^2 + 2z + 3 = 0 \quad \text{lie?}$$

(a) The equation has no real roots:

Obviously it has no positive root.

Put $z = -x$: $x^4 - x^3 + 4x^2 - 2x + 3 = 0$. For $0 < x < 1$ the first three terms together are positive, and so are the last two. For $x > 1$ the first two terms together are positive, and so are the last three.

Therefore it has no negative roots.

(b) The equation has no purely imaginary roots:

Put $z = iy$: $y^4 - iy^3 - 4y^2 + 2iy + 3 = 0$ and the real and imaginary parts of this do not vanish together.

Now consider $\Delta \arg (z^4 + \ldots + 3)$ taken round the part of the first quadrant bounded by $|z| = R$. The variation along the real axis is zero. On the arc of the circle $z = Re^{i\theta}$. Then for sufficiently large R:

$$\Delta \arg (z^4 + \ldots + 3) = \Delta \arg (R^4 e^{4i\theta}) + \Delta \arg\left[1 + O(R^{-1})\right]$$

$$= 2\pi + O(R^{-1}) \qquad , \qquad R \to \infty$$

On the imaginary axis we have

$$\arg (z^4 + \ldots) = \arc \tan \left(\frac{-y^3 + 2y}{y^4 - 4y^2 + 3} \right)$$

As y varies from $+\infty$ to 0, the expression in brackets varies according to:

to:	$y = \infty$	$\sqrt{3}$	$\sqrt{2}$	1	0
	0	$-\infty$	0	$+\infty$	0

Therefore $\arg (z^4 + \ldots)$ decreases by 2π if y decreases from $+\infty$ to 0.

Thus the total variation of arg $(z^4 + \ldots)$ round the first quadrant is zero, if R is large enough.

Hence there are no zeros in the first quadrant. Since the zeros must occur in conjugate pairs (f(z) has real coefficients), it follows that there are no zeros in the fourth quadrant, and two in each of the second and third quadrants.

Any algebraic equation may be treated in the same way.

The calculations involved in determining this N might be extremely tedious though. Then the following methods might be of some help.

B: The following theorem sometimes gives useful information about the zeros of a function.

T (ii): Let C be a simple closed contour, inside and on which f(z) is analytic. Then if $Re[f(z)]$ vanishes at 2k distinct points on C, f(z) has at most k zeros inside C.

Proof: We have from above that if $f(z) = u + iv$ the number N of zeros of f(z) inside C is given by

$$N = \frac{1}{2\pi} \Delta_C \,(\text{arc tan } v / u)$$
$$= \frac{1}{2\pi} \int_C d\,(\text{arc tan } v / u)$$

Starting at a point where $u \neq 0$, we may take the initial value of arc tan (v/u) ~~tan (v / u)~~ to lie between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$. We can only pass out of this range, say to $(\frac{1}{2}\pi, \frac{3}{2}\pi)$ if u vanishes; and only pass on to $(\frac{3}{2}\pi, \frac{5}{2}\pi)$ if u vanishes again. Thus, if u vanishes twice on C, Δ_C (arc tan v / u) is at most equal to 2π, and N is at most equal to 1. The general result obviously follows from the same argument.

(This theorem was for instance used with great expediency by R. J. Backlund in his "Über die Nullstellen der Riemannschen Zetafunktion" ; Acta Mathematica 41, (1918), 345 - 75).

C: The following theorem is actually a consequence of the maximum-modulus theorem.

T (iii): Let $f(z)$ be regular, and $|f(z)| \leq M$ in the circle $|z - a| \leq R$, and suppose that $f(a) \neq 0$. Then the number of zeros of $f(z)$ in the circle $|z - a| \leq \frac{1}{3}R$ does not exceed $A \log [M / |f(a)|]$

where $$A = \frac{1}{\log 1/h} \quad , \quad 0 < h < 1$$

Proof: For the sake of simplicity, suppose $a = 0$. Let $\xi_1, \xi_2, \ldots \xi_n$ be the zeros of $f(z)$ in $|z| \leq \frac{1}{3}R$ and let

$$g(z) = f(z) / \prod_{j=1}^{n} \left(1 - \frac{z}{\xi_j}\right)$$

$$= \frac{\prod_{j=1}^{n} (z - \xi_j)\, \psi(z)}{\prod_{j=1}^{n} (z - \xi_j) \cdot (-1)^n} \cdot \xi_1 . \xi_2 \cdots \xi_n = (-1)^n \psi(z) \prod_{j=1}^{n} \xi_j$$

where $\lim_{z \to \xi_j} \psi(z) < \infty$, $j = 1 \ldots n$.

Therefore, $g(z)$ is regular for $|z| \leq R$ and on $|z| = R$ we have since

$|\xi_j| \leq \frac{1}{3}R$, $\left|\frac{z}{\xi_j}\right| \geq 3$ for $j = 1, 2, \ldots n$

Thus $$|g(z)| \leq M / \prod_{j=1}^{n} (3 - 1) = 2^{-n} M \qquad (9.1)$$

for $|z| = R$, and by the maximum-modulus theorem also for $|z| < R$.

Since $g(o) = f(o)$ it follows that

$$|f(o)| \leq 2^{-n} M$$

Thus $$n \leq \frac{1}{\log 2} \log \frac{M}{|f(o)|}$$

From (9.1) it is obvious that $\frac{1}{3}$ can be replaced by any number less than $\frac{1}{2}$. A more complete result can be obtained from Jensen's theorem which says:

Let f(z) be analytic for $|z| < R$. Suppose that f(0) is not zero, and let $r_1, r_2, \ldots, r_n, \ldots$ be the moduli of the zeros of f(z) in the circle $|z| < R$, arranged as a non-decreasing sequence. Then if $r_n \leq r \leq r_{n+1}$,

$$\log \frac{r_n |f(0)|}{r_1 r_2 \cdots r_n} = \frac{1}{2\pi} \int_0^{2\pi} \log |f(re^{i\theta})| \, d\theta$$

Now, if the zeros in $|z| \leq R$ have moduli $r_1, r_2, \ldots, r_N$ then, applying Jensen:

$$\log \frac{R^N}{r_1, r_2 \cdots r_N} = \frac{1}{2\pi} \int_0^{2\pi} \log |f(Re^{i\theta})| \, d\theta - \log|f(0)|$$

$$\leq \log M - \log|f(0)|$$

Let the zeros in the circle $|z| \leq hR$, $0<h<1$ have moduli $r_1, r_2, \ldots, r_n$. Then the left-hand side is not less than

$$\log \frac{R^n}{r_1, r_2 \cdots r_n} \geq \log (1/h)^n = n \log(1/h)$$

Thus $$n \leq \frac{1}{\log 1/h} \log \frac{M}{|f(0)|}$$

Q. E. D.

f(z) a Polynomial

For the rest of this chapter we will discuss auxiliary measures which were proposed for finding the approximate location of the roots of a polynomial.

D: Let us begin with a problem of the first category; to find an upper bound for the moduli of all the zeros of a polynomial. A classic solution of such a problem is the result due to Cauchy [Exercises de mathématiques; Oeuvres(2) Vol. 9, (1829) p. 122; Journ. Ecole Poly. Vol. 25 (1837) p. 176] namely:

T (iv): All the zeros of the polynomial $f(z) = a_o + a_1 z + \dots + a_n z^n$, $a_n \neq 0$ lie in the circle $|z| \leq r$, where r is the positive root of the equation

$$|a_o| + |a_1| z + \dots + |a_{n-1}| z^{n-1} - |a_n| z^n = 0 \qquad (9.2)$$

Also

T (v): All the zeros of $f(z) = a_o + a_1 z + \dots + a_n z^n$, $a_n \neq 0$ lie in the circle $|z| < 1 + \max |a_j / a_n|$, $j = 0, 1, \dots, n - 1$

These two theorems formed the basis of the result due to Birkhoff [An elementary double inequality for the roots of an algebraic equation having greatest absolute value; Bull. Amer. Math. Soc. 21 (1914) pp. 494 - 495], Cohn [Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise; Math. Zeit. 14 (1922) pp. 110 - 148] and Berwald [Elementare Sätze über die Abgrenzung der Wurzeln einer algebraische Gleidung; Acta Univ. Szeged 6 (1934) pp. 209 - 221] namely:

T (vi): The zero z_1 of largest modulus of $f(z) = a_o + a_1 z + \dots + a_n z^n$, $a_n \neq o$ satisfies the inequality

$$r \geq |z_1| \geq (2^{1/n} - 1) r$$

where r is the positive root of the equation (9.2).

This again led to the important result of Kuniyeda [Note on the roots of algebraic equations; Tôhaku Math. J. 9 (1916) pp. 167 - 173; Ibid. 10 (1916) pp. 185 - 188.], Montel [Sur la limite supérieure du module des racines d'une équation algébrique; C. R. de la Société des Sciences de Varsovie 24, (1932) pp. 317 - 326; C. R. Acad. Sci. Paris 193, (1931) pp. 974 - 976) and Tôya [Some remarks on Montel's paper concerning upper limit of absolute values of roots of algebraic equations; Science Reports Tokyo Bunrika Daigaku 1, (1933) pp. 275 - 282 which we can state as:

T (vii): For any p and q such that

$$p > 1 \; , \quad q > 1 \; , \quad \frac{1}{p} + \frac{1}{q} = 1$$

the polynomial $f(z) = a_o + a_1 z + \dots + a_n z^n \; , \quad a_n \neq 0$

has all its zeros in the circle.

$$|z| < [\, 1 + (\sum_{j=0}^{n-1} |a_j|^p / \, |a_n|^p \,)^{q/p} \,]^{\frac{1}{q}}$$

$$< (\, 1 + n^{q/p} . M^q \,)^{\frac{1}{q}}$$

where $M = \max |a_j/a_n| \; , \; j = 0, 1, \dots , n - 1$.

An important generalization of Cauchy's T (iv) was published by M. A. Pellet in 1881:

T (viii): If for a polynomial

$$f(z) = a_o + a_1 z + \dots + a_p z^p + \dots + a_n z^n \; , \quad a_p \neq 0$$

the equation

$$F_p(z) = |a_o| + |a_1| z + \dots + |a_{p-1}| z^{p-1} - |a_p| z^p + |a_{p+1}| z^{p+1} + \dots + |a_n| z^n \; .$$

has two positive zeros r and R, $r < R$, then f(z) has exactly p zeros in or on the circle $|z| \leq r$ and no zeros in the "annular" ring $r < |z| < R$.

Let us divide the plane into 2p equal sectors S_k having their common vertex at the origin and having the rays

$$\theta = (\alpha_o + j\pi) / p \qquad j = 1, 2, \ldots, 2p$$

as their bisectors. Let us denote by $G(r_o, r; p, \alpha_o)$ the boundary of the region formed by adding to the circular region $|z| < r_o$ those points of the "annulus" $r_o \leq |z| \leq r$ which lie in the odd numbered sectors $S_1, S_3, \ldots S_{2p-1}$. (see Fig. 5).

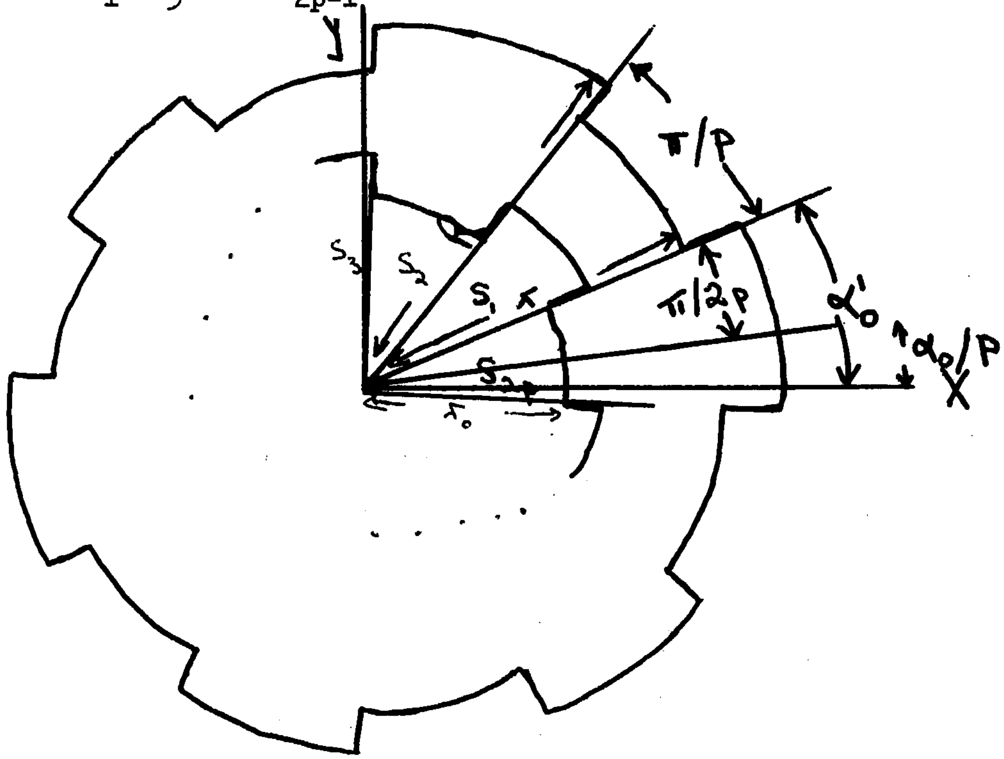

Fig. 5

Then the following refinement of Pellet's Theorem can be proposed. (see for proof M. Marden [23]).

T (ix): If the polynomial

$$f(z) = a_o + a_1 z + \ldots + a_p z^p + \ldots + a_n z^n$$

with $a_o a_1 a_p a_n \neq o$ and $\alpha_o = \arg(a_o / a_p)$ be such that the equation

$$F_p(z) = |a_o| + |a_1| z + \ldots + |a_{p-1}| z^{p-1} - |a_p| z^p + |a_{p+1}| z^{p+1} + \ldots + |a_n| z^n = 0.$$

has two positive zeros r and R, $r < R$; then the equation

$$H_p(z) = |a_1| + |a_2|z + \dots + |a_{p-1}|z^{p-2} - |a_p|z^{p-1} + |a_{p+1}|z^p + \dots + |a_n|z^{n-1} = 0$$

has two positive zeros r_o and R_o with $r_o < r < R < R_o$. Furthermore, the polynomial $f(z)$ has precisely p zeros in or on the curve $G(r_o, r; p, \alpha_o)$ and no zeros in the "annular" region between the curves $G(r_o, r; p, \alpha_o)$ and $G(R, R_o; p, \alpha_o + \pi)$

A generalization of T (v) was established by P. Montel [Sur quelques limites pour les modules des zéros des polynômes; Comm. Math. Helv. 7 (1934 - 35) pp. 178 - 200; C. R. Acad. Sci. Paris 199 (1934) pp. 651 - 653, 760 - 762.]

T (x): At least p zeros of the polynomial

$f(z) = a_o + a_1 z + \dots + a_n z^n$ lie in the circle

$$|z| < 1 + \max. |a_j / a_n|^{1/(n-p+1)} , \qquad j = 0, 1, \dots , p .$$

E: By means of a potential - and function - theoretical approach P. C. Rosenbloom [22] arrived at the following theorem which might be of some help.

T (xi): Let $P(z) = \sum_{n=0}^{k} a_n z^n$ be a polynomial of degree k, and let

$$\lambda = k^{-1} \log (M(1, P)^2 / |a_o||a_k|)$$

where $M(r, P) = \max_{|z| = r} |P(z)|$

Let N(E) be the number of zeros of P in E, divided by k. Then, if E is the set

$$\frac{1}{2} \leq z \leq 2, \quad |\arg z - \alpha| \leq \frac{\tau}{2}$$

where $\alpha = \min_{z \in E} (\arg z)$

then $\left| N(E) - \gamma / 2\pi \right| \leq 36 \lambda^{\frac{1}{4}} \log (2 + 1/\lambda)$

F: H. S. Wall [10] and others have drawn attention to the significance the expansion of polynomials in continued fractions can have in determining the location of the roots of the polynomial.

Let $P(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n$ be a polynomial of degree $n > 0$ with complex coefficients a_j.

Put $R_e (a_j) = p_j$ and $I_m (a_j) = q_j$ so that $a_j = p_j + iq_j$

Then the polynomial

$$Q(z) = p_1 z^{n-1} + iq_2 z^{n-2} + p_3 z^{n-3} + iq_4 z^{n-4} + \dots$$

is called the alternant of $P(z)$. The quotient $Q(z) / P(z)$ has, in general, a so-called J-fraction expansion of the form

$$\frac{Q(z)}{P(z)} = \cfrac{1}{c_1 z + 1 + k_1 + \cfrac{1}{c_2 z + k_2 + \cfrac{1}{c_3 z + k_3 + \cdots + \cfrac{1}{c_n z + k_n}}}} \tag{9.3}$$

called the test-fraction of $P(z)$. Here the c_j are real and different from zero; and the k_j are pure imaginary or zero. This test-fraction for $P(z)$ exists iff the determinants

$$F_j = \begin{vmatrix} p_1 & p_3 & p_5 & \cdots & p_{2j-1} & -q_2 & -q_4 & -q_6 & \cdots & -q_{2j-2} \\ 1 & p_2 & p_4 & \cdots & p_{2j-2} & -q_1 & -q_3 & -q_5 & \cdots & -q_{2j-3} \\ 0 & p_1 & p_3 & \cdots & p_{2j-3} & 0 & -q_2 & -q_4 & \cdots & -q_{2j-4} \\ & & \cdots & & & & & \cdots & & \\ 0 & 0 & 0 & \cdots & p_j & 0 & 0 & 0 & \cdots & -q_{j-1} \\ 0 & q_2 & q_4 & \cdots & q_{2j-2} & p_1 & p_3 & p_5 & \cdots & p_{2j-3} \\ 0 & q_1 & q_3 & \cdots & q_{2j-3} & 1 & p_2 & p_4 & \cdots & p_{2j-4} \\ 0 & 0 & q_2 & \cdots & q_{2j-4} & 0 & p_1 & p_3 & \cdots & p_{2j-5} \\ & & \cdots & & & & & \cdots & & \\ 0 & 0 & \cdots & & q_j & 0 & 0 & \cdots & & p_{j-1} \end{vmatrix}$$

$$j = 2, 3, 4, \ldots, n, \quad p_j = q_j = 0 \quad \text{for } j > n$$

are different from zero. (See E. Frank [11])

The following table is suggested for obtaining expansion (9.3):

$$a_{00} = 1, \qquad a_{01} = iq_1, \qquad a_{02} = p_2, \qquad \ldots$$

$$a_{11} = p_1, \qquad a_{12} = iq_2, \qquad a_{13} = p_3, \qquad \ldots$$

$$c_1 = \frac{a_{00}}{a_{11}}, \quad b_{11} = a_{01} - c_1 a_{12}, \quad b_{12} = a_{02} - c_1 a_{13}, \quad b_{13} = a_{03} - c_1 a_{14}, \ \ldots$$

$$k_1 = \frac{b_{11}}{a_{11}}, \quad a_{22} = b_{12} - k_1 a_{12}, \quad a_{23} = b_{13} - k_1 a_{13}, \quad a_{24} = b_{14} - k_1 a_{14}, \ \ldots$$

$$c_2 = \frac{a_{11}}{a_{22}}, \quad b_{22} = a_{12} - c_2 a_{23}, \quad b_{23} = a_{13} - c_2 a_{24}, \quad b_{24} = a_{14} - c_2 a_{25}, \ \ldots$$

$$k_2 = \frac{b_{22}}{a_{22}}, \quad a_{33} = b_{23} - k_2 a_{23}, \quad a_{34} = b_{24} - k_2 a_{24}, \quad a_{35} = b_{25} - k_2 a_{25}, \ \ldots$$

$$c_3 = \frac{a_{22}}{a_{33}}, \quad b_{33} = a_{23} - c_3 a_{34}, \quad b_{34} = a_{24} - c_3 a_{35}, \quad b_{35} = a_{25} - c_3 a_{36}, \ \ldots$$

$$\cdots\cdots\cdots \qquad (9.4)$$

H. S. Wall succeeded in establishing polygonal bounds for the roots of a polynomial by means of the following theorem:

T (xii): Let P(z) be a polynomial of degree n having a test-fraction written in the form

$$\frac{Q(z)}{P(z)} = \cfrac{c}{b_1 + z - \cfrac{d_1^2}{b_2 + z - \cdots - \cfrac{d_{n-1}^2}{b_n + z}}} \qquad (9.5)$$

Let $\delta_j(\theta) = I_m(d_j e^{i\theta})$, $\beta_j(\theta) = I_m(b_j e^{i\theta})$

and let $Y(\theta)$ be any number such that

$$\beta_j(\theta) + Y(\theta) \geq 0 , \quad j = 1, 2, 3, \ldots , n .$$

$$2\delta_1^2(\theta) \leq [\beta_1(\theta) + Y(\theta)][\beta_2(\theta) + Y(\theta)]$$

$$4\delta_j^2(\theta) \leq [\beta_j(\theta) + Y(\theta)][\beta_{j+1}(\theta) + Y(\theta)] , \quad j = 2, 3, \ldots , n-1 .$$

Then all roots of P(z) are contained in the rectangle

$$y \leq Y(\theta) , \quad x \leq Y(\tfrac{\pi}{2})$$

$$y \geq -Y(\pi) , \quad x \geq -Y(\tfrac{3\pi}{2})$$

In 1945 (see [10]) Wall proved the following theorem for a polynomial with real coefficients.

T (xiii): Let $P(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n$ be a polynomial with real coefficients, and let

$Q(z) = a_1 z^{n-1} + a_3 z^{n-3} + a_5 z^{n-5} + \ldots$ be the alternant of P(z). Then all the zeros of P(z) have negative real parts iff

$$\frac{Q(z)}{P(z)} = \cfrac{1}{c_1 z + 1 + \cfrac{1}{c_2 z + \cfrac{1}{c_3 z + \cdots + \cfrac{1}{c_n z}}}} \tag{9.6}$$

where the coeffieients $c_1, c_2, \ldots, c_n$ are all positive.

Also

T (xiv): In the expansion (9.6) let k of the coefficients c_j be negative and the remaining n - k be positive. Then k of the zeros of P(z) have positive real parts, and n - k have negative real parts.

Example: Let $P(z) = z^5 - 3z^4 - 9z^3 - 27z^2 - 32z - 30$

Then $Q(z) = -3z^4 - 27z^2 - 30.$

The expansion (9.6) may be obtained by dividing P(z) by Q(z) until a remainder is obtained which is of lower degree than Q(z); then Q(z) is divided by this remainder, and so on. (Scheme (9.4) may also be used of course.) If we write only the coefficients, we have:

```
                   -1/3
-3 +0 -27 +0 -30 | 1 -3 -9 -27 -32 -30
                   1 +0 +9 + 0 +10
                     -3-18 -27 -42 -30
                     -3 +0 -27 + 0 -30        1/6
                      -18 + 0 -42 + 0 | -3 +0 -27 +0 -30
                                        -3 +0 - 7 +0
                                             -20 +0 -30
                    9/10
Cont.  -20 +0 -30 | -18 +0 -42 +0
                    -18 +0 -27          4/3
                           -15 +0 | -20 +0 -30
                                    -20 +0        1/2
                                          -30 | -15
                                                -15
                                                  0
```

Hence $c_1 = -1/3$, $c_2 = 1/6$, $c_3 = 9/10$, $c_4 = 4/3$, $c_5 = 1/2$. Thus, there is one zero in the right half-plane and four in the left-half plane.

In 1946 E. Frank [11] proved the following two analogous theorems for the polynomial with complex coefficients:

T (xv): Let $P(z) = z^n + a_1 z^{n-1} + \dots + a_n$ be a polynomial of degree $n > 0$ with complex coefficients $a_j = p_j + iq_j$, $j = 1, 2, \dots, n$. Let $Q(z) = p_1 z^{n-1} + iq_2 z^{n-2} + q_3 z^{n-3} + iq_4 z^{n-4} + \dots$ be the alternant of P(z). All roots of P(z) have negative real parts iff P(z) has a test-fraction of the form

$$\frac{Q(z)}{P(z)} = \cfrac{d_0}{z + d_0 + b_1 + \cfrac{d_1}{z + b_2 + \cfrac{d_2}{z + b_3 + \cdots + \cfrac{d_{n-1}}{z + b_n}}}}$$

in which d_0 , d_1 , ... , d_{n-1} are real and positive, and b_1 , b_2, ... , b_n are pure imaginary or zero.

T (xvi): If P(z) has a test-fraction (9.3) in which k of the c_j are negative and n - k are positive, then k of the roots of P(z) have positive real parts, and n - k have negative real parts.

Approximate computation of the roots of a polynomial

We shall now give a method, based upon this last theorem, for determining the approximate location of the roots of a polynomial. Let

$$P(z) = z^n + (p_1 + iq_1)\, z^{n-1} + (p_2 + iq_2)\, z^{n-2} + \dots + p_n + iq_n$$

be the given polynomial. Put $P_h(z) = P(z + h)$.

Let $Q_h(z)$ be the alternant of $P_h(z)$, and let

$c_j(h) = a_{j-1,\ j-1}(h) / a_{j,\ j}(h)$, $j = 1, 2, 3, \ldots, n$.

(cf. (9.4)) be the coefficients of z in the test-fraction (9.3) for $P_h(z)$. By the theorem above, if $k = k(h)$ of the coefficients $c_j(h)$ are positive for a given real value of h, then $P_h(z)$ has just k(h) roots in the half-plane $R_e(z) < 0$, so that P(z) has just k(h) roots in the half-plane $R_e(z) < h$.

In general, the method for determining the roots of P(z) consists in varying h in such a way that $a_{n,\ n}(h) \to 0$ and k(h) changes by one unit. This means that the last remainder in the division process used in forming the test-fraction $P_h(z)$, which is simply the Euclidean algorithm for the greatest common divisor of $P_h(z) - Q_h(z)$ and $Q_h(z)$, approaches zero. If $z_0(h)$ is the root of the next to the last remainder, $a_{n-1,\ n-1}(h)\, z + a_{n-1,\ n}(h)$, then $h + z_0(h)$ approaches a root of P(z) as $a_{n,\ n}(h)$ approaches zero. If two or more roots ofP(z) have a common real part, the process must be suitably modified. (cf. Example 2, following.)

We shall now show how the computation can be so arranged that the roots of P(z) can be effectively determined by this method. Horner's method will be applied for the formation of the polynomials $P_h(z)$. The Euclidean algorithm can be reduced to the computation in table (9.4).

<u>Example 1</u>: Computation of the roots of

$P(z) = z^3 + (1 + 6i)z^2 - (13 - 5i)z - (7 + 10i)$. First compute the test-fraction for P(z) by means of (9.4):

$$
\begin{array}{llllll}
 & & a_{00} = 1, & a_{01} = 6i, & a_{02} = -13, & a_{03} = -10i \\
-c_1 = -1, & -k_1 = -i, & a_{11} = 1, & a_{12} = 5i, & a_{13} = -7, & \\
 & & b_{11} = i, & b_{12} = -6, & b_{13} = -10i, & \\
-c_2 = 1, & -k_2 = 2i, & a_{22} = -1, & a_{23} = -3i, & & \\
 & & b_{22} = 2i, & b_{23} = -7, & & \\
-c_3 = -1, & -k_3 = -3i, & a_{33} = -1, & & & \\
 & & b_{33} = -3i . & & &
\end{array}
$$

Since $c_1 > 0$, $c_2 < 0$, $c_3 > 0$, there are two roots in $R_e(z) < 0$ and one in $R_e(z) > 0$ (T.(xvi)). By T (xii) we find the roots of P(z) are contained in the rectangle

$y \leq -1, \quad x \leq 2, \quad y \geq -3, \quad x \geq -2, \quad (z = x + iy)$

Now compute the polynomial $P_1(z) = P(z + 1)$ by means of Horner's scheme:

1	1+6i	-13+5i	-7-10i	k = 1
1	1	2+6i	-11+11i	
1	2+6i	-11+11i	-18+ i	
	1	3+6i		
1	3+6i	- 8+17i		
	1			
1	4+6i			

Hence, $P_1(z) = z^3 + (4 + 6i)z^2 + (-8 + 17i)z + (-18 + i)$

Now from the table (9.4) for $P_1(z)$:

$$
\begin{array}{lllll}
 & a_{00} = 1 & a_{01} = 6i & a_{02} = -8 & a_{03} = i \\
c_1 = \frac{1}{4} & a_{11} = 4 & a_{12} = 17i & a_{13} = -18 & \\
 & b_{11} = 1.75i & b_{12} = -3.5 & b_{13} = i &
\end{array}
$$

$$
\begin{array}{lll}
 & a_{22} = 3.9375 & a_{23} = 8.8750i \\
c_2 = \frac{4}{3.9375} & & \\
 & b_{22} = 7.98413i & b_{23} = -18.00
\end{array}
$$

$c_3 = -\frac{3.9375}{.003985}$ $\quad a_{33} = -.003985$

Thus, $P(z)$ has one root in the half-plane $R_e(z) > 1$, and $a_{33}(1) = -.003985$.

We now form

$$P_2(z) = z^3 + (7 + 6i)z^2 + (3 + 29i)z + (-21 + 24i)$$

and find that $c_1(2)$, $c_2(2)$ and $c_3(2)$ are all positive, so that all the roots of $P(z)$ are in the half-plane $R_e(z) < 2$. There is one root in the strip $1 < R_e(z) < 2$. We find that $a_{33}(2) = 8.98$. Since we had $a_{33}(1) = -.003985$, it would appear that this root has real part very nearly equal to 1. If we assume that $a_{33}(h)$ varies linearly with h, we find, by interpolation, that we should have $a_{33}(1.0004) = 0$. In the light of this information, we now form

$$P(z + 1.001) = z^3 + (4.003 + 6i)z^2 + (-7.991997 + 17.012i)z$$
$$+ (-18.007995999 + 1.017006i)$$

and construct table (9.4) for this polynomial. We find $a_{33}(1.001) = -.000076233$ and that

$$a_{22}(1.001)z + a_{23}(1.001) = 3.944596397z + 8.8904426667i$$

On setting the latter equal to zero we find for the imaginary part of the root the approximate value $-2.254i$. We thus have as an approximate value of the root $1.001 - 2.254i$. Now we can apply one of the iterative algorithms discussed above, or:

$$P(d + 1.001 - 2.254i) = d^3 + (4.003 - .7621i)d^2 + (3.814455 - 1.033524i)d$$
$$- (.000253547 + .0006456 98i)$$

If we neglect the terms in d^3 and d^2, and set the linear part equal to zero, we obtain the correction

$d = .0000192 + .0001745i$

Then $d + 1.001 - 2.254i = 1.0010192 - 2.2538255i$ is the value of the root. This is actually correct to the number of places given, since it was found by application of the Newton-Raphson algorithm that the root is $1.0010192259 - 2.2538255167i$. (the last digits 9 and 7 are in doubt.) For the other two roots of $P(z)$ we find the values

$$- 1.520324 - 1.399879l6i \qquad \text{and}$$

$- .480695 - 2.3462953i$ correct to the number of places given. As a check, we find the sum and the product of these values of the roots are $-1-6i$ and $7+10i$ respectively, correct to six decimal places.

<u>Example 2:</u> $P(z) = z^3 + 2z + 20.$

This polynomial has a pair of conjugate imaginary roots. Since the coefficient of z^2 is zero, the test-fraction does not exist. This is of little concern, since the test-fraction exists for $P_h(z) = P(z + h)$ when h is near the real parts of the roots. We have by $\mathcal{T}$ (xvi) applied to $P_1(z)$ and $P_2(z)$, that the imaginary roots are in the strip $1 < R_e(z) < 2$.

In the following table, the numbers m_j are the next to the last remainders obtained in applying the Euclidean algorithm to the polynomials $P_h(z) - Q_h(z)$ and $Q_h(z)$.

1	0	2	20		
	1	1	3	h = 1,	$m_1 = 5 - \frac{23}{3}$
1	1	3	23		
	1	2			= 2.7
1	2	5			
	1				
1	3				
	1	4	9	$\frac{1}{h = 2}$	$m_2 = 14 - \frac{32}{6}$
1	4	9	32		
	1	5			= 8.7
1	5	14			
	1				
1	6				
	-.8	-4.16	-7.872	$\frac{-.8}{h = 1.2}$	$m_{1.2} = -.38$
1	5.2	9.84	24.128		
	-.8	-3.52			
1	4.4	6.32			
	-.8				
1	3.6				
	.1	.37	.669	$\frac{.1}{h = 1.3}$	$m_{1.3} = .71$
1	3.7	6.69	24.797		
	.1	.38			
1	3.8	7.07			
	.1				
1	3.9				
	-.07	-.2681	-.476133	$\frac{-.07}{h = 1.23}$	$m_{1.23} = -.041$
1	3.83	6.8019	24.320867		
	-.07	-.2632			
1	3.76	6.5387			
	-.07				
1	3.69				
	.01	.037	.065757	$\frac{.01}{h = 1.24}$	$m_{1.24} = .073$
1	3.7	6.5757	24.386624		
	.01	.0371			

1	3.71	6.6128		
	.01			
1	3.72			
	-.006	-.022284	-.039543096	$\frac{-.006}{h = 1.234}$ $m_{1.234} = -.0084$
1	3.714	6.590516	24.347080904	
	-.006	-.022248		
1	3.708	6.568268		
	-.006			
1	3.702			
	.001	.003703	.006571971	$\frac{.001}{h = 1.235}$ $m_{1.235} = -.0024$
1	3.703	6.571971	24.353652875	
	.001	.003704		$3.705k^2 + 24.354 = 0,$
1	3.704	6.575675		$k = \pm 2.564i$
	.001			
1	3.705			

$$P(d + 1.235 + 2.564\,i) = d^3 + (3.705 + 7.692\,i)d^2 - (13.146613 - 18.99924\,i)d - (.003572805 - .004048556i).$$

The real part of the imaginary roots has the value h = 1.235 correct to three decimal places. The imaginary parts are $\pm 2.564i$, as indicated above. On equating to zero the linear part of $P(d + 1.235 + 2.564\,i)$ we obtain the correction $d = -.000237 + .000247i$. Thus, the imaginary roots are approximately equal to $1.234773 \pm 2.564247i$. Since the sum of the roots is equal to zero, the real root must be -2.469546. One may readily verify (by Newton's method or otherwise) that this is correct to six decimal places.

Note: (1) If this method for determining the roots of a polynomial is to serve only as a means for finding an initial approximation z_0 for the iterative use of an algorithm of the types discussed in previous chapters, this method is terminated as soon as possible. Its application is usually quite laborious in comparison with the application of the Newton algorithm or one of its modifications - especially if a computing machine is used.

(2) This method of computation of the roots is closely related to the method proposed by F. L. Hitchcock [12].

The considerations discussed in A and also D, led Cauchy [A. L. Cauchy: Calcul des indices des fonctions; Journal de l'Ecole Polytechnique, Vol. 15, 1837, pp. 176 - 229 (OEvres (2), Vol. 1, pp. 416 - 466)] to introduce the notion of the "index" of a rational fraction. He also developed formulas for the computation of the index, and introduced the method to compute the index by means of Sturm's series. (Kronecker extended the notion of index to systems of functions).

We may note here the important Cauchy Index Theorem as presented by Hurwitz [Über die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reelen Theilen besitz; Math. Ann. 46 (1895), 273 - 284; Math. Werke 2, 533 - 545.]

<u>T (xvii)</u>: Let $f(z) = a_0 + a_1 z + \dots + a_{n-1} z^{n-1} + z^n = P_0(x) + P_1(x)$ where $P_0(x)$ and $P_1(x)$ are real polynomials with $P_1(x) \neq 0$. As the point x moves on the real axis from $-\infty$ to $+\infty$, let m be the number of real zeros of $P_0(x)$ at which $g(x) = P_0(x)/P_1(x)$ changes from - to +, and k the number of real zeros of $P_0(x)$ at which g(x) changes from + to -. If f(z) has no real zeros, p zeros in the upper half-plane and q zeros in the lower half-plane, then

$$p = (1/2)\left[n + (k - m)\right], \quad q = (1/2)\left[n - (k - m)\right]$$

(For further information on geometric methods to determine domains containing no zeros or at least one zero, or sometimes all zeros of a given polynomial, we refer to Walsh [24] and Marden [23].)

On the basis of Cauchy's work E. J. Routh [13] derived the following rule for testing a polynomial

$$P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n , \quad a_0 > 0 \quad \text{with } a_j \text{ real.}$$

Consider the array

$$\begin{array}{cccc} a_0 & a_2 & a_4 & \cdots \\ a_1 & a_3 & a_5 & \cdots \\ \dfrac{a_1 a_2 - a_0 a_3}{a_1} , & \dfrac{a_1 a_4 - a_0 a_5}{a_1} & & \\ \cdots & & & \end{array}$$

where the third row is obtained from the first two by cross-multiplication. The next row is obtained from the second and third by the same prodess. Thus the first element in the fourth row is

$$\frac{a_3 \left(\dfrac{a_1 a_2 - a_0 a_3}{a_1} \right) - (a_1 a_4 - a_0 a_5)}{\dfrac{a_1 a_2 - a_0 a_3}{a_1}}$$

Each row has one fewer elements than the preceding row. Then the number of variations in signs in the sequence making up the first column of the array is equal to the number of zeros of P(z) having positive real parts. This was shown by Wall to be essentially T (xiv). (The method of Routh fails however in case division by zero is involved in his algorithm.)

1	5	3	2	4	2	6	4		
	- 0.75	- 3.1875	0.140625	- 1.6054688	- 1.7958985	- 0.1530761	- 4.3851929		
1	4.25	- 0.1875	2.140625	2.3945312	0.2041015	5.8469239	- 0.3851929	- 0.1905556	
	- 0.75	- 2.625	2.109375	- 3.1875000	0.5947266	- 0.5991211	1	- 0.3851929	v_2= 0.0734008
1	3.50	- 2.8125	4.250000	- 0.7929688	0.7988281	(5.2478028)			
	- 0.75	- 2.0625	3.656250	- 5.9296875	5.0419922	0	0	0	
1	2.75	- 4.8750	7.90625	- 6.7226563	5.8408203				
	- 0.75	- 1.5	4.78125	- 9.515625	1	5.8408203	1	- 0.3537244	v_3= 0.0674043
1	2	- 6.375	12.68750	- 16.2382813					
	- 0.75	- 0.9375	5.484375	0.0674043	1	4.7459903	1	- 0.3636303	v_4= 0.0692919
1	1.25	- 7.3125	18.1718750						
	-0.75	- 0.375	0.0048014	0.0692919	1	4.8028952	1	- 0.3621325	v_5= 0.0690065
1	0.5	- 7.6875							
	- 0.75	0.0003286	0.0047619	0.0690065	1	4.8071584	1	- 0.3623154	v_6= 0.0690414
1	- 0.25								
	0.0000227	0.0003291	0.0047667	0.0690414	1	4.8038741	1	- 0.3622946	v_7= 0.0690374
1									
0.0000016	0.0000227	0.0003291	0.0047662	0.0690374	1	4.8038765	1	- 0.3622970	v_8= 0.0690378
0.0000016	0.0000227	0.0003291	0.0047662	0.0690374	1	4.8038765	1	- 0.3622970	v_9= 0.0690378 $=v_8$

Scheme for finding a root of the equation

$$x^7 + 5x^6 + 3x^5 + 2x^4 + 4x^3 + 2x^2 + 6x + 4 = 0$$

by means of an iterative algorithm of the k-th order : $x_1 = x_0 + v_k$. $x_0 = -0.75$.

Fig. 6

Example.

Que.: Find the real root of

$$f(x) = x^7 + 5x^6 + 3x^5 + 2x^4 + 4x^3 + 2x^2 + 6x + 4 = 0.$$

lying between $x=-1$ and $x = -0.5$

[$f(-1) = -1$, $f(-0.5) = +1.1016$]

Choose $x_0 = -0.75$.

According to T. 20 the Newton algorithm will be the most expedient in this case. For the sake of a perceptible representation however, we will apply the scheme given in Fig. 4. Then we obtain Fig. 6 and :

2nd.	order algorithm	(Newton):	$x_1^{(2)} = -0.6765992$
3rd.	"	" :	$x_1^{(3)} = -0.6825957$
4th.	"	" :	$x_1^{(4)} = -0.6807081$
5th.	"	" :	$x_1^{(5)} = -0.6809935$
6th.	"	" :	$x_1^{(6)} = -0.6809586$
7th.	"	" :	$x_1^{(7)} = -0.6809626$
8th.	"	" :	$x_1^{(8)} = -0.6809622$

After j=8 no further changes in the values of the v_j appear up to seven decimal places. After two applications of the Newton algorithm the value $x_2^{(2)} = -0.6809622$ ($x_0^{(2)} = -0.75$, $x_1^{(2)} = -0.6765992$) was found. This is exactly the same as that found after one application of the 8th. order algorithm.

We have $f(-0.6809622) = 4.4 \cdot 10^{-8}$

$$\therefore \quad | f(x_1^{(8)}) | < 5 \cdot 10^{-8}$$

Hence, in the domain D: $-0.681 \leqq x \leqq -0.680$

$$| f'(x) | > 5.8 = m.$$

Therefore

$$\frac{| f(x_1^{(8)}) |}{m} \leqq \frac{5}{5.8} \cdot 10^{-8} < 1 \cdot 10^{-8} = c$$

Hence $\xi = -0.68096220 \pm 10^{-8}$.

In the domain D we have the following upper bounds :

$$\frac{M_2}{2!} = 3 \quad , \quad \frac{M_3}{3!} = 12 \quad , \quad \frac{M_4}{4!} = 16 \quad , \quad \frac{M_5}{5!} = 8$$

$$[f^{(j)}(x) \leqq M_j \quad , \quad x \in D.]$$

From this we obtain after T. 24 (with $c = 10^{-8}$).

$C_3 = 2.61$ $\quad |x_n - \xi| \leqq 0.62 [1.62 \cdot 10^{-8}]^{3^n}$

$C_4 = 23.1$ $\quad |x_n - \xi| \leqq 0.36 [2.85 \cdot 10^{-8}]^{4^n}$

$C_5 = 493.9$ $\quad |x_n - \xi| \leqq 0.22 [4.72 \cdot 10^{-8}]^{5^n}$

and after T. 3 :

$C_2 = 0.5173$ $\quad |x_n - \xi| \leqq 1.93 [0.52 \cdot 10^{-8}]^{2^n}$

where $x_0 = -0.68096220.$

Summary.

It is the hope of the author that the preceding pages might be of some help to the engineer and the physicist who are interested in the practical application of iterative algorithms for the solution of analytic equations. However , since he would like to think of himself more as a "pure analyst" than a practical "numerical analyst" , the author hopes above all that this paper might prove to be of some pure theoretical interest as well.

Bibliography.

[1] Schröder, E. : Über unendlich viele Algorithmen zur Auflösung der Gleichungen ; Math. Ann. 2, (1870), 317-365.

[2] Ehrmann, H. : Konstruktion und Durchführung von Iterationsverfahren höherer Ordnung ; Archive for Rational Mechanics and Analysis 4, (1959-60), 65-88.

[3] Schwerdtfeger, H. + Blaskett, D. R. : A formula for the solution of an arbitrary equation ; Quarterly of Applied Mathematics 3, (1945) 266-268.

[4] Schwerdtfeger, H. : Notes on Numerical Analysis, 1. Polynomial Iteration ; Canadian Math. Bulletin 2, (1959), 97-109.

[5] Frame, J:S. : A variation of Newton's method ; Amer. Math. Monthly 51, (1944), 36-38.

[6] ________ : Remarks on a variation of Newton's method ; Ibid. 52 (1945), 212-214.

[7] Wall, H. S. : A modification of Newton's method ; Ibid. 55, (1948), 90-94.

[8] Bodewig, E. : Konvergenztypen und das Verhalten von Approximationen in der Nähe einer mehrfachen Wurzel einer Gleichung ; Z. angew. Math. u. Mech. 29, (1949), 44-51.

[9] Min, Szu-Hoa : Non- analytic functions ; Amer. Math. Monthly 51, (1944), 510-517.

[10] Wall, H. S. : Polynomials whose zeros have negative real parts ; Ibid. 52, (1945), 308-322.

[11] Frank, E. : On the zeros of polynomials with complex coefficients; Bull. Amer. Math. Soc. 52, (1946), 144-157.

[12] Hitchcock, F:L: : Algebraic equations with complex coefficients ; Jour. of Math. and Physics (M. I. T.) 18, (1939), 202-210.

[13] Routh , E:J. : The advanced part of a treatise on the dynamics of a system of rigid bodies ;Macmillan (1892), 194-202.

[14] Ostrowski , A. M. : Vorlesungen über Differential- und Integralrechnung Vol. 2 , Basel (1951).

[15] ________________ : Solution of Equations and Systems of Equations; Academic Press (1960).

[16] Bieberbach , L. : Lehrbuch der Funftionentheorie Vol. 1 ; B. G. Teubner (1921-31) , 113-114.

[17] Behnke , H. + Sommer , F. : Theorie der analytischen Funktionen einer komplexen Veränderlichen ;Springer-Verlag (1955).

[18] Copson , E:T. : Functions of a Complex Variable ;Oxford, Claredon Press (1935).

[19] Titchmarsh , E:C. : The theory of functions ; 2nd. Ed. (1939), Oxford Univ. Press.

[20] Turnbull , H:W. ; Theory of Equations ; 5th. Ed. (1952), Oliver + Boyd Ltd.

[21] Wall , H. S. : Analytic theory of Continued Fractions ;(1948), D. van Nostrand Co. , Inc.

[22] Rosenbloom , P. C. : Distribution of zeros of polynomials ; Functions of a complex variable , Ed. by W. Kaplan (1955), Univ. of Michigan Press , 265-285.

[23] Marden , M. : The geometry of the zeros of a polynomial in a complex variable ; Mathematical Surveys No. 3, N. Y. (1949).

[24] Walsh , J. L. : The location of the critical points of analytic and harmonic functions ;Amer. Math. Soc. Colloquium Publications No. 34 , N. Y. (1950).

[25] Scarborough , J. B. : Numerical mathematical analysis ;Ox. Univ. Press. (1930).

[26] Cajori , F. : History of Mathematics 2nd. ed. ;N. Y. , Macmillan Co. (1931).