MATH. SCAND. 46 (1980), 236–250

NONDISCRETE INDUCTION AND A DOUBLE STEP SECANT METHOD

F.-A. POTRA and VLASTIMIL PTÁK

The application of the method of nondiscrete mathematical induction to the iterative processes treated thus far, in particular to the Newton process, has led to definitive results; not only has it yielded estimates sharp in every step, but also the conditions on the initial data obtained turn out to be optimal. Last but not least it should be mentioned that this method leads to very simple and clear proofs. It is natural that the first processes to be treated were the simple ones; iterative processes involving a larger number of quantities to be computed and estimated at each step make it necessary to consider approximate sets depending on more than one parameter. Recently approximate sets depending on two parameters were applied by F.-A. Potra in his paper [7] on the Regula Falsi; using a two-dimensional analogue of the induction theorem he was able to obtain, for the secant method, results of the same order of sharpness and optimality as those mentioned above.

In the present paper we intend to apply these ideas to the study of a double step secant method.

This method consists in the construction of two sequences $\{x_n\}, \{y_n\}$

$$y_{n+1} = y_n - [y_n, x_n; f]^{-1} f(y_n) \tag{1}$$

$$x_{n+1} = y_{n+1} - [y_n, x_n; f]^{-1} f(y_{n+1}) \tag{2}$$

where f is a mapping from a Banach space E into a Banach space F, and where $[y_n, x_n; f]$ denotes a divided difference of the operator f in the points y_n and x_n (for the definition, see Section 2).

This iterative procedure requires at each step only one inversion of a linear operator, so that it is not much more complicated than the secant method [20], [21]; it is, however, considerably faster than this one. Indeed, from a theorem of J. W. Schmidt and H. Schwetlick [22] it follows that the order of convergence of this method is $1+\sqrt{2}$ (See also the paper [3] of P. Laasonen). In [22] the convergence of the iterative procedure is proved under the assumption that there exists a simple root x^* of the equation $f(x)=0$, the convergence

Received May 11, 1979.

conditions being formulated in terms of $|(f'(x^*))^{-1}|$. In [3] one does not suppose apriori the existence of a root of the equation and the convergence conditions are formulated only in terms of the initial data (i.e. $|f(x_0)|$, $|f(y_0)|$, $|[y_0, x_0;\ f]^{-1}|$). None of the above mentioned papers contains estimates of the distances $|y_n - x^*|$ and $|x_n - x^*|$.

We intend to show in the present paper that sharp estimates, as well as optimal convergence conditions, may be obtained by the method of nondiscrete mathematical induction. The present paper is entirely self-contained and may be read without the knowledge of any article on the method of nondiscrete mathematical induction. Nevertheless, let us mention the Gatliburg lecture [12], or the survey [19], where the reader could learn more about the motivation and the general principles of the application of this method, should he so desire.

SOME NOTATIONS. If (E, d) is a metric space, x_0 an element of E and r a positive number, we denote by $U(x_0, r)$ the closed spherical neighbourhood of x_0 with radius r:

$$U(x_0, r) = \{x \in E\ ;\ d(x, x_0) \leqq r\} .$$

If X is a normed vector space and x an element of X then we shall denote by $|x|$ the norm of this element.

We shall use a measure of invertibility for linear operators in Banach spaces. If A is a linear operator on a Banach space E we define

$$d(A) = \inf\{|Ax|\ ;\ |x| \geqq 1\} .$$

Clearly $d(A) = |A^{-1}|^{-1}$ if A is invertible; also $d(A+B) \geqq d(A) - |B|$ for any perturbation by a bounded linear operator B. If A is invertible and $d(A) - |B| > 0$ then so is $A + B$.

1. The induction theorem.

Let p be a natural number. For each $i = 1, 2, \ldots, p$, let T_i be either the set of all positive numbers, or an open interval $(0, t_i)$ for some $t_i > 0$. Denote by T the cartesian product $T_1 \times \ldots \times T_p$.

DEFINITION. A function ω mapping T into T is said to be a rate of convergence on T if the series $\sum_{n=0}^{\infty} \omega^{(n)}(r)$ is convergent for each $r \in T$.

Here $\omega^0(r) = r$, and the function $\omega^{(n)}$ is the nth iterate of ω, so that $\omega^{(n+1)}(r) = \omega(\omega^{(n)}(r))$.

We shall, sometimes, refer to such functions as rates of convergence of type (p, p).

If β is a function with values in $\mathbf{R}^p$, we shall denote by β_i its ith component; thus, for instance,

$$\omega^{(n)}(r) = (\omega_1^{(n)}(r),\ldots,\omega_p^{(n)}(r)) .$$

To simplify some of the formulae it will be convenient to introduce the following vector function defined for $r \in T$ by the formula

$$(3) \qquad \psi(r) = (r_1+\ldots+r_p, r_2+\ldots+r_p+\omega_1(r),\ldots, \\ r_p+\omega_1(r)+\ldots+\omega_{p-1}(r)) .$$

Using this function, we can attach to the rate of convergence ω the ($\mathbf{R}^p$ valued) function

$$(4) \qquad \sigma(r) = \sum_{n=0}^{\infty} \psi(\omega^{(n)}(r)) .$$

The functions ω and σ are obviously related by the following functional equation

$$(5) \qquad \sigma(r) = \sigma(\omega(r))+\psi(r), \qquad r \in T .$$

The components of σ satisfy the relations

$$(6) \qquad \sigma_k(r) = \sigma_1(r)-(r_1+\ldots+r_{k-1}), \qquad k=2,\ldots,p .$$

If a and b are two elements of $\mathbf{R}^p$, with coordinates a_i and b_i respectively, we write $a \leqq b$ if $a_i \leqq b_i$ for all i; similarly $a<b$ is taken to mean $a_i<b_i$ for all i.

Now let (E,d) be a complete metric space and let A be a subset of E^p. For each $i=1,2,\ldots,p$ we assign to A the subset A_i of E consisting of those $x \in E$ for which there exists an $a \in A$ whose ith coordinate is x. If $x \in E^p$ we denote by $d^{(p)}(x,A)$ the vector with components $d(x_i,A_i)$, $i=1,2,\ldots,p$. If $r \in \mathbf{R}^p$, $r>0$, we shall denote by $U(A,r)$ a neighbourhood of A of the form

$$U(A,r) = \{x \in E^p \ ; \ d^{(p)}(x,A) \leqq r\} .$$

If $Z(r)$, $r \in T$, is a family of subsets of E^p we denote by $Z(0)$ its limit:

$$Z(0) = \bigcap_{s\in T}\Big(\bigcup_{t\leqq s} Z(t)\Big)^- .$$

Now we can state the following generalisation of the Induction Theorem [12].

THEOREM. *Let T be a p-dimensional interval and let ω be a rate of convergence on T. Suppose that the family $Z(t) \subset E^p$, $t \in T$, satisfies*

$$Z(r) \subset U(Z(\omega(r)), \psi(r)) \tag{7}$$

for each $r \in T$. *Then*

$$Z(r) \subset U(Z(0), \sigma(r)) \tag{8}$$

for each $r \in T$.

PROOF. Given $x_0 \in Z(r)$, successive application of (7) yields a sequence $\{x_n\}$ such that

$$x_{n+1} \in Z(\omega^{(n+1)}(r)) \cap U(x_n, \psi(\omega^{(n)}(r));$$

the sequence $\{x_n\}$ is Cauchy in E^p because the distances of its consecutive terms are majorated by the terms of a convergent series

$$d^{(p)}(x_{n+1}, x_n) \leqq \psi(\omega^{(n)}(r)) .$$

It is easy to prove that the limit x^* of this sequence belongs to $Z(0)$ and that $d^{(p)}(x^*, x_0) \leqq \sigma(r)$. The proof is complete.

Let us add two remarks: Instead of E^p we could have taken the family $Z(\cdot)$ to be subsets of the cartesian product of p complete metric spaces (E_j, d_j). The induction theorem remains true if the function ψ is replaced by any other function ψ' with the property that the series

$$\sigma'(r) = \sum_{n=0}^{\infty} \psi'(\omega^{(n)}(r))$$

is convergent for all $r \in T$.

Let us sketch briefly how the above theorem may be applied to the study of the convergence of iterative procedures of the form

$$x_{n+1} \in F(x_n) , \tag{10}$$

where F is a multivalued mapping of E^p into E^p and x_0 is a given element of E^p. (By a multivalued mapping we mean a mapping which assigns to points $x \in E^p$ subsets of E^p; in many cases this subset $F(x)$ will be just one point so that F will be a mapping from E^p into E^p, but the theorem is designed to take care of errors in computing as well).

If we can attach to the pair (F, x_0) a family of sets $Z(r) \subset E^p$, $t \in T$, and a rate of convergence ω on T, so that the following conditions are satisfied:

(11) $\qquad x_0 \in Z(r_0) \quad$ for a certain $r_0 \in T$,

(12) $r \in T$ and $x \in Z(r)$ imply $F(x) \cap Z(\omega(r)) \cap U(x, \psi(r))$ is nonvoid ,

then it follows from the induction theorem that $Z(0)$ is nonvoid. Moreover the sequence $\{x_n\}$ obtained by successive applications of condition (12) converges to a point $x^* \in E^p$ and satisfies the following relations

$$x_n \in Z(\omega^{(n)}(r_0)) \tag{13}$$

$$d^{(p)}(x_n, x_0) \leqq \sigma(r_0) - \sigma(\omega^{(n)}(r_0)) \tag{14}$$

$$d^{(p)}(x_n, x^*) \leqq \sigma(\omega^{(n)}(r_0)) . \tag{15}$$

The last inequality represents an estimate of the distance between the successive approximations x_n and the "solution" x^*. Since this estimate can be given prior performing the iterations we shall call it an *apriori estimate.* On the other hand, having computed $x_1, \ldots, x_n$, we possess, in general, information on which better estimates may be based. Indeed, suppose we find an $r_{n-1} \in T$ such that $x_{n-1} \in Z(r_{n-1})$. Taking x_{n-1} and r_{n-1} for x_0 and r_0 we have, by (15), the estimate

$$d^{(p)}(x_n, x^*) \leqq \sigma(\omega(r_{n-1})) = \sigma(r_{n-1}) - \psi(r_{n-1}) . \tag{17}$$

An r_{n-1} for which $x_{n-1} \in Z(r_{n-1})$ can usually be found as a function of x_{n-1} and x_n, $r_{n-1} = \delta(x_{n-1}, x_n)$. Estimates of the type (17) shall be referred to as *aposteriori estimates* since they can only be given after the elements $x_1, \ldots, x_n$ have been computed.

2. Divided differences.

The iterative procedure (1)–(2) uses divided differences of an operator, a notion introduced by J. Schröder [23]. We shall use this notion in the framework described in [3].

Let E and F be two Banach spaces and let f be a mapping of

$$U = \{x \in E \; ; \; |x - x_0| \leqq m\}$$

into F. If x and y are two points in U, $x \neq y$, a divided difference of f at x, y is a linear operator $A(x, y) \in L(E, F)$ such that

$$A(x, y)(x - y) = f(x) - f(y);$$

of course, this requirement does not determine the operator uniquely except in the case that E has dimension one. Now suppose that for each pair of distinct points $u, v \in U$, an $A(u, v)$ is given which is a divided difference of f at u, v. In the following we shall suppose that the mapping $(u, v) \mapsto A(u, v)$ satisfies a Lipschitz condition.

DEFINITION. A mapping $(u,v) \mapsto [u,v;\, f]$ is said to be a divided difference of f on U satisfying a Lipschitz condition (with constant H) if the following conditions are satisfied:

1°. $[u,v;\, f]$ is a linear and bounded operator from E into F such that

$$[u,v;\, f](u-v) = f(u)-f(v) \tag{18}$$

for each pair $u,v \in U$, $u \neq v$.

2°. If two such pairs (u_1,v_1) and (u_2,v_2) are given then

$$|[u_2,v_2;\, f]-[u_1,v_1;\, f]| \leqq H(|u_2-u_1|+|v_2-v_1|)\,. \tag{19}$$

Suppose now that the above two conditions are satisfied; then the following is easy to prove.

For each $x \in U$ the Fréchet derivative of f exists and equals the limit of $[y,x;\, f]$ as $y \in U$ tends to x, $y \neq x$. We may therefore extend the domain of definition of $[\cdot,\cdot;\, f]$ by setting

$$[x,x;\, f] = f'(x)$$

for each $x \in U$. The Fréchet derivative f' satisfies then a Lipschitz condition

$$|f'(x)-f'(y)| \leqq 2H|x-y|\,.$$

We shall use the following estimate: given three points $x,y,y' \in U$, then

$$|f(y')-f(x)-[y,x;\, f](y'-x)| \leqq H|y'-y|\,|y'-x|\,.$$

If we take $y=x$, this reduces to the familiar formula:

$$|f(y')-f(x)-f'(x)(y'-x)| \leqq H|y'-x|^2\,.$$

The estimate is a consequence of the relation

$$f(y') = f(x)+[y,x;\, f](y'-x)+([y',x;\, f]-[y,x;\, f])(y'-x)\,.$$

3. Convergence of the process.

In this section we intend to prove the main theorem; we shall give sufficient conditions for the convergence and estimates for the distances $|y_n-x^*|$ and $|x_n-x^*|$.

In section 4 we shall show that the conditions we impose are the weakest possible (a precise meaning of this statement is given in Proposition (4.1)) and that the estimates are sharp.

The iterations to be considered are sucessive constructions of pairs of points; accordingly, we shall work in two dimensions so that $p=2$. We take for T the positive quadrant (i.e. $T=(0,\infty)\times(0,\infty)$); instead of r_1, r_2 we shall write q, r. With these notations we can state

3.1. LEMMA. *Let a be a positive number. The pair of functions*

$$\omega_1(q,r) = \frac{r(q+r)}{q+2r+2(r(q+r)+a^2)^{\frac{1}{2}}} \tag{21}$$

$$\omega_2(q,r) = \omega_1 \frac{q+2r+\omega_1}{2(r(q+r)+a^2)^{\frac{1}{2}}-\omega_1} \tag{22}$$

$$= \frac{r(q+r)(q^2+5r(q+r))+2r(q+r)(q+2r)(r(q+r)+a^2)^{\frac{1}{2}}}{(q+2r)(8a^2+7r(q+r))+2(4a^2+q^2+7r(q+r))(r(q+r)+a^2)^{\frac{1}{2}}} \tag{22'}$$

is a rate of convergence on T. The corresponding σ function is given by

$$\sigma_1(q,r) = \beta(q,r)+q+r,\ \sigma_2(q,r) = \beta(q,r)+r, \tag{23}$$

where

$$\beta(q,r) = (r(q+r)+a^2)^{\frac{1}{2}}-a. \tag{24}$$

PROOF. Consider the real polynomial

$$f(x) = x^2-a^2. \tag{26}$$

If $a<x_0<y_0$, then the algorithm

$$y_{n+1} = y_n - \frac{y_n-x_n}{f(y_n)-f(x_n)} f(y_n)$$

$$x_{n+1} = y_{n+1} - \frac{y_n-x_n}{f(y_n)-f(x_n)} f(y_{n+1})$$

yields two nonincreasing sequences $\{x_n\}$ and $\{y_n\}$ which converge to a. Choosing

$$x_0 = r+(r(q+r)+a^2)^{\frac{1}{2}},\ y_0 = x_0+q \tag{27}$$

we have

$$y_0-x_0 = q \tag{28}$$

and

$$x_0-y_1 = r. \tag{29}$$

Taking now

$$\omega_1(q,r) = y_1-x_1,\ \omega_2(q,r) = x_1-y_2,\ \sigma_1(q,r) = y_0-a,$$

we obtain the formulae (21), (22) and (23). The rest follows from the convergence of the process. We observe also that the following relations hold:

$$y_n - x_n = \omega_1^{(n)}(q,r) \tag{30}$$

$$x_n - y_{n+1} = \omega_2^{(n)}(q,r) \tag{31}$$

$$x_n - x_0 = \sigma_2(q,r) - \sigma_2(\omega^{(n)}(q,r)) \tag{32}$$

$$y_n - x_0 = \sigma_2(q,r) - \sigma_1(\omega^{(n)}(q,r)) \tag{33}$$

$$x_n - a = \sigma_2(\omega^{(n)}(q,r)) \tag{34}$$

$$y_n - a = \sigma_1(\omega^{(n)}(q,r)) . \tag{35}$$

We shall use the above lemma in the proof of the following theorem. There the constant a will be chosen of the form

$$a = \frac{1}{2H}(d_0^2 + H^2 q_0^2 - 2d_0 H(q_0 + 2r_0))^{\frac{1}{2}} , \tag{36}$$

where d_0, H, q_0, r_0 are some constants depending on the initial data.

3.2. THEOREM. *Let E and F be two Banach spaces, let x_0 be a point of E and let f be a mapping from the closed disc $U = U(x_0, m)$ into F. Let a divided difference of f be given which satisfies a Lipschitz condition with constant H. Let y_0 be a given point of U.*

Suppose that the following conditions are satisfied:

1°. *The operator $[y_0, x_0; f]$ is invertible and $d([y_0, x_0; f]) \geqq d_0$,*

2°. $|y_0 - x_0| \leqq q_0$,

3°. $|[y_0, x_0; f]^{-1} f(x_0)| \leqq r_0$.

If the initial data satisfy the inequalities

4°. $(r_0^{\frac{1}{2}} + (q_0 + r_0)^{\frac{1}{2}})^2 \leqq d_0/H$,

5°. $m \geqq \sigma_2(q_0, r_0)$,

then the iterative procedure

$$y_{n+1} = y_n - [y_n, x_n; f]^{-1} f(y_n)$$

$$x_{n+1} = y_{n+1} - [y_n, x_n; f]^{-1} f(y_{n+1})$$

is meaningful, the sequences $\{y_n\}$, $\{x_n\}$ obtained by it converge to a root x^* of the equation $f(x) = 0$, and the following estimates hold:

6°. $|x^* - x_0| \leqq \sigma_2(q_0, r_0)$,

7°. $|y_n - x^*| \leqq \sigma_1(\omega^{(n)}(q_0, r_0))$,

8°. $|x_n - x^*| \leqq \sigma_2(\omega^{(n)}(q_0, r_0))$,

9°. $|y_n - x^*| \leqq \beta(|y_{n-1} - x_{n-1}|, |x_{n-1} - y_n|)$,

10°. $|x_n - x^*| \leqq \beta(|y_{n-1} - x_{n-1}|, |x_{n-1} - y_n|) - \omega_1(|y_{n-1} - x_{n-1}|, |x_{n-1} - y_n|)$,

where ω_1, ω_2, β, σ_1, σ_2 are the functions defined in lemma (3.2) *with a given by* (36).

PROOF. We shall consider a family of sets depending on two positive parameters q, r as follows:

$$Z(q,r) = \{(y,x) \in E^2 \ ; \ |y-x| \leqq q, \ [y,x; \ f]^{-1} \text{ exists },$$
$$d([y,x; \ f]) \geqq h(q,r), |[y,x; \ f]^{-1} f(x)| \leqq r\}$$

where h is a positive function to be determined later.

We intend to prove that

$$(y_0, x_0) \in Z(q_0, r_0) \tag{37}$$

and, given $(y,x) \in Z(q,r)$, that the pair (y', x')

$$y' = y - [y,x; \ f]^{-1} f(y) \tag{38}$$

$$x' = y' - [y,x; \ f]^{-1} f(y') \tag{39}$$

satisfies the inclusion

$$(y', x') \in Z(\omega(q,r)) \cap U((y,x), \psi(q,r)) \tag{40}$$

for at suitable rate of convergence ω on T.

The inclusion (40) is equivalent to

$$|y' - x'| \leqq \omega_1(q,r) \tag{41}$$

$$d([y', x'; \ f]) \geqq h(\omega(q,r)) \tag{42}$$

$$|[y', x'; \ f]^{-1} f(x')| \leqq \omega_2(q,r) \tag{43}$$

$$|y' - y| \leqq q + r \tag{44}$$

$$|x' - x| \leqq r + \omega_1(r,q) \ . \tag{45}$$

Suppose $(y,x) \in Z(q,r)$; it follows from (38) that

$$y' - x = y - x - [y,x; \ f]^{-1} f(y) = -[y,x; \ f]^{-1} f(x) \ ,$$

whence

$$y' = x-[y,x;\ f]^{-1}f(x)\ . \tag{46}$$

Using this we have

$$\begin{aligned} f(y') &= f(x)+[y,x;\ f](y'-x)+([y',x;\ f]-[y,x;\ f])(y'-x) \\ &= ([y',x;\ f]-[y,x;\ f])(y'-x)\ , \end{aligned}$$

whence

$$|f(y')| \leqq H|y'-y|\,|y'-x|\ . \tag{47}$$

From the definition of Z and from (46) we have $|y'-x| \leqq r$, whence

$$|y'-y| \leqq |y'-x|+|x-y| \leqq r+q\ ,$$

so that (44) is established. Furthermore

$$|y'-x'| \leqq \frac{|f(y')|}{h(q,r)} \leqq \frac{Hr(q+r)}{h(q,r)}; \tag{48}$$

it follows that (41) will be satisfied if we assume that

$$\frac{Hr(q+r)}{h(q,r)} \leqq \omega_1(q,r)\ .$$

Let us do that. Now

$$\begin{aligned} d([y',x';\ f]) &\geqq d([y,x;\ f])-|[y',x';\ f]-[y,x;\ f]| \\ &\geqq h(q,r)-H(|y'-y|+|x'-x|)\ . \end{aligned}$$

We already know that $|y'-x| \leqq r$; since

$$|x'-x| \leqq |x'-y'|+|y'-x| \leqq \omega_1(q,r)+r$$

we have (45) and

$$d([y',x';\ f]) \geqq h(q,r)-H(q+2r+\omega_1(q,r))\ . \tag{49}$$

To estimate $f(x')$ we write

$$f(x') = f(y')+[y,x;\ f](x'-y')+([y',x';\ f]-[y,x;\ f])(x'-y')$$

so that, by (39)

$$f(x') = ([y',x';\ f]-[y,x;\ f])(x'-y')\ .$$

Thus

$$(50) \qquad |[y',x';f]^{-1}f(x')| \leqq \frac{H(|y'-y|+|x'-x|)\omega_1(q,r)}{h(q,r)-H(q+2r+\omega_1(q,r))}$$

$$\frac{H(q+2r+\omega_1(q,r))\omega_1(q,r)}{h(q,r)-H(q+2r+\omega_1(q,r))}$$

To simplify the formulae, set $h=Hk$. To satisfy (41), (42), (43) it will be sufficient, in view of (48), (49), (50), to have

$$(51) \qquad \omega_1 = \frac{r(q+r)}{k}$$

$$(52) \qquad k\circ\omega = k-(q+2r+\omega_1)$$

$$(53) \qquad \omega_2 = \omega_1 \frac{q+2r+\omega_1}{k\circ\omega}.$$

It is easy to see that these functional equations are satisfied if we set

$$k(q,r) = q+2r+2(r(q+r)+a^2)^{\frac{1}{2}}$$

and take for ω_1, ω_2 the functions from Lemma 3.1. Here a is still a free parameter; it will be chosen so as to have $(y_0, x_0) \in Z(q_0, r_0)$. For this it suffices to satisfy $h(q_0, r_0) = d_0$, which leads to the choice

$$(36) \qquad a = \frac{1}{2H}(d_0^2+H^2q_0^2-2d_0H(q_0+2r_0))^{\frac{1}{2}}.$$

We have to remark that condition 4° of the Theorem implies that

$$d_0^2+H^2q_0^2 \geqq 2d_0H(q_0+2r_0)$$

so that the formula (36) makes sense and it defines a nonnegative number.

It follows from the Induction Theorem that the sequence (y_n, x_n) is convergent and that $(y_n, x_n) \in Z(\omega^{(n)}(q_0, r_0))$. This however implies that $|y_n - x_n| \leqq \omega_1^{(n)}(q_0, r_0)$, so that both sequences x_n and y_n converge to the same limit point x^*. Since by (47)

$$|f(y_{n+1})| \leqq H|y_{n+1}-y_n||y_{n+1}-x_n|$$

and since (19) implies the continuity of f we have $f(x^*)=0$. The estimates 6°, 7°, 8° are immediate consequences of (13), (14), (15). The aposteriori estimates 9°, 10° are immediate consequences of (17) if we prove that

$$(58) \qquad (y_{n+1}, x_{n-1}) \in Z(|y_{n-1}-x_{n-1}|, |x_{n-1}-y_n|).$$

The first condition in the definition of $Z(\cdot,\cdot)$ is satisfied trivially while the last condition is a consequence of (46). It remains to prove the inequality

$$d([y_{n-1},x_{n-1};\ f]) \geqq h(|y_{n-1},x_{n-1}|,|x_{n-1}-y_n|)\ .$$

Since we know that $(y_{n-1},x_{n-1}) \in Z(\omega^{(n-1)}(q_0,r_0))$ we have

$$d([y_{n-1},x_{n-1};\ f]) \geqq h(\omega^{(n-1)}(q_0,r_0))$$

and the desired inequality is a consequence of the monotonicity of h and the inequality

$$(|y_{n-1}-x_{n-1}|,|x_{n-1}-y_n|) \leqq \omega^{(n-1)}(q_0,r_0)\ .$$

The proof is complete.

Let us turn now to the question of uniqueness. The inequality 6° implies that

$$x^* \in U(x_0,\sigma_2(q_0,r_0))\ ,$$

while from 5° we have

$$U(x_0,\sigma_2(q_0,r_0)) \subset U = U(x_0,m)\ .$$

Let us denote by V the open disk

$$V = \{x \in E\ ;\ |x-x_0| < \sigma_2(q_0,r_0)+2a\}\ .$$

The following result is easy to prove

3.3. Uniqueness Theorem. *If in Theorem* 3.2 *the inequality* 4° *is strict, then* x^* *is the only root of the equation* $f(x)=0$ *in the set* $U\cap V$.

Proof. First let us remark that the inequality 4° is equivalent to

$$\frac{d_0}{H} \geqq q_0+2r_0+2(r_0(q_0+r_0)+a^2)^{\frac{1}{2}}\ . \tag{60}$$

If either 4° or (60) is strict then $a>0$.

Now let us suppose that there exist a $z \in U\cap V$ such that $f(z)=0$. We have then

$$z-x^* = [y_0,x_0;\ f]^{-1}([y_0,x_0;\ f]-[z,x^*;\ f])(z-x^*)$$

and

$$\begin{aligned}&|[y_0,x_0;\ f]^{-1}([y_0,x_0;\ f]-[z,x^*;\ f])|\\ &\quad\leqq \frac{H}{d_0}(|y_0-x_0|+|x_0-z|+|x_0-x^*|)\\ &\quad< \frac{H}{d_0}(q_0+2\sigma_2(q_0,r_0)+2a)\end{aligned}$$

$$= \frac{H}{d_0}(q_0+2r_0+a+2(r_0(q_0+r_0)+a^2)^{\frac{1}{2}}) = 1 ,$$

whence $|z-x^*|=0$.

4. Sharpness and optimality.

In Theorem 3.2, in order to assure the convergence of the iterative process to a root of the equation $f(x)=0$, we have imposed condition 4° to the initial data d_0, q_0, r_0 and the Lipschitz constant H:

4°. $(r_0^{\frac{1}{2}}+(q_0+r_0)^{\frac{1}{2}})^2 \leqq d_0/H$.

Of course, in applications, q_0 can be made small if y_0 is taken close enough to the initial approximation x_0. If q_0 is very small, then condition 4° reduces to $4Hr_0 \leqq d_0$, which means that the initial approximation must be good.

This is a property shared by all Newton-like methods. If we compare condition 4° with the corresponding condition required by Laasonen (cf. [3, condition (2.6)]) we can see—in spite of the fact that the two conditions are not formulated in the same terms—that condition 4° is considerably weaker than the one from Laasonen's paper. In terms of d_0, q_0, r_0 and H, condition 4° turns out to be the weakest possible. This is a consequence of

4.1. PROPOSITION. *For each quadruple of positive numbers* d_0, H, q_0, r_0 which do *not satisfy condition* 4°, *there exists a function* $f\colon \mathbf{R} \to \mathbf{R}$ *and two points* $x_0, y_0 \in \mathbf{R}$ *such that*:

(a) *The divided difference of f satisfies a Lipschitz condition with constant H and conditions* 1°, 2°, 3° *of Theorem* 3.2 *are satisfied.*

(b) *The equation* $f(x)=0$ *has no solution.*

PROOF. Take

$$f(x) = Hx^2+\frac{1}{4H}(2d_0H(q_0+2r_0)-d_0^2-H^2q_0^2),$$

$$x_0 = \frac{d_0-Hq_0}{2H}, \quad y_0 = \frac{d_0+Hq_0}{2H}$$

if

$$q_0+2r_0-2\sqrt{r_0(r_0+q_0)} < \frac{d_0}{H} < q_0+2r_0+2\sqrt{r_0(r_0+q_0)}$$

and

$$f(x) = \frac{d_0}{q_0}x^2 + d_0 r_0, \; x_0 = 0, \; y_0 = q_0$$

if

$$0 < \frac{d_0}{H} \leqq q_0 + 2r_0 - 2\sqrt{r_0(r_0+q_0)} \, .$$

The following result shows that the estimates obtained in Theorem 3.2 are sharp.

4.2. Proposition. *For each quadruple of positive numbers* d_0, H, q_0, r_0 *which satisfy* 4°, *there exists a function* $f\colon \mathrm{R} \to \mathrm{R}$ *and two points* $x_0, y_0 \in \mathrm{R}$ *which satisfy the hypotheses of Theorem* 3.2, *and for which all the estimates* 6°, 7°, 8°, 9°, 10° *are attained.*

Proof. As we have already remarked, in the proof of Theorem 3.2, condition 4° implies that $d_0^2 + H^2 q_0^2 \geqq 2d_0 H(q_0 + 2r_0)$, so that formula (36) makes sense. Set $f(x) = H(x^2 - a^2)$, where a is given by this formula and set

$$x_0 = r_0 + (r_0(q_0+r_0)+a^2)^{\frac{1}{2}}, \; y_0 = x_0 + q_0 \, .$$

It can be easily verified that

$$\frac{f(y_0) - f(x_0)}{y_0 - x_0} = d_0, \qquad \frac{f(x_0)}{d_0} = r_0 \, ,$$

and that the sequences obtained by the procedure (1)–(2) for the function $H(x^2 - a^2)$ are the same as those for the function $x^2 - a^2$. The rest follows from the proof of Lemma 3.1.

REFERENCES

1. M. Balazs and G. Goldner, *On existence of divided differences in linear spaces*, Rev. Anal. Numér. Théorie Approximation 2 (1973), 3–6.
2. J. Křižková and P. Vrbová, *A remark on a factorization theorem*, Comment. Math. Univ. Carolinae 15 (1974), 611–614.
3. P. Laasonen, *Ein überquadratisch konvergenter iterativer Algorithms*, Ann. Acad. Sci. Fenn. Ser. AI 450 (1969), 1–10.
4. T. Popoviciu, *Introduction à la théorie des differences divisées*, Bull. Math. Soc. Sci. Math. R. S. Roumanie 42 (1941), 65–78.
5. F.-A. Potra, *The Rate of convergence of a modified Newton's process*, Preprint INCREST (National Institute of Scientific and Technical Creation, Bucharest) No. 36/1978.
6. F.-A. Potra, *On a modified secant method*, Analyse Numérique et le Théorie de l'Approximation 8 (1979), 203–214.

7. F.-A. Potra, *An application of the Induction Method of V. Pták to the study of Regula Falsi*, Preprint INCREST (National Institute of Scientific and Technical Creation, Bucharest) No. 11/1979.
8. V. Pták, *Some metric aspects of the open mapping theorem*, Math. Ann. 163 (1966), 95–104.
9. V. Pták, *Deux théorèmes de factorization*, C. R. Acad. Sci. Paris Sér. A 278 (1974), 1091–1094.
10. V. Pták, *A theorem of the closed graph type*, Manuscripta Math. 13 (1974), 109–130.
11. V. Pták, *A quantitative refinement of the closed graph theorem*, Czechoslovak Math. J. 99 (1974), 503–506.
12. V. Pták, *Nondiscrete mathematical induction and iterative existence proofs*, Linear Algebra and Appl. 13 (1976), 223–238.
13. V. Pták, *The rate of convergence of Newton's process*, Numer. Math. 25 (1976), 279–285.
14. V. Pták, *Concerning the rate of convergence of Newton's process*, Comment. Math. Univ. Carolinae 16 (1975), 699–705.
15. V. Pták, *A rate of convergence*, Numerical Functional Analysis and Optimization 1 (1979), 255–271.
16. V. Pták, *A modification of Newton's Method*, Časopis Pěst. Mat. 101 (1976), 188–194.
17. V. Pták, *Factorization in Banach Algebras*, Studia Math. 65 (1979), 279–285.
18. V. Pták, *Stability of Exactness*, Comment. Math. (Poznań) 21 (1978), 343–348.
19. V. Pták, *Nondiscrete Mathematical Induction*, in: General Topology and its Relations to Modern Analysis and Algebra IV (Lecture Notes in Math. 609), pp. 166–178, Springer-Verlag, Berlin - Heidelberg - New York, 1977.
20. A. S. Sergeev, *O metode chord*, Sibirsk. Mat. Ž. 2 (1961), 282–289.
21. J. Schmidt, *Eine Übertragung der Regula Falsi auf Gleichungen im Banachraum* I, II, Z. Angew. Math. Mech. 43 (1963), 1–8, 9–110.
22. J. W. Schmidt und H. Schwetlick, *Ableitungsfreie Verfahren mit hoherer Konvergenzgeschwindigkeit*, Computing 3 (1968), 215–226.
23. J. Schröder, *Nichtlineare Majoranten beim Verfahren der schriftweisen Näherung*, Arch. Math. (Basel) 7 (1956), 471–484.
24. S. Ulm, *Ob obobscennych razdelenych raznostjach*, I, II, Eesti NSV Tead. Akad. Toimetised Füüs.-Mat. 16 (1967), 13–26, 146–156.
25. P. Vrbová and H. Petzeltová, *A remark on small divisor problems*, Czechoslovak Math. J. 103 (1978), 1–12.
26. P. Vrbová and H. Petzeltová, *An underrelaxed modification of Newton's method*, Rev. Roumaine Math. Pures Appl. 22 (1977), 959–963.
27. J. Zemánek, *A remark on transitivity of operator algebras*, Časopis Pěst. Mat. 100 (1975), 176–178.

DEPARTMENT OF MATHEMATICS
NATIONAL INSTITUTE FOR
SCIENTIFIC AND TECHNICAL CREATION
BD. PĂCII 220, 77538 BUCHAREST
ROMANIA

AND

INSTITUTE OF MATHEMATICS
CZECHOSLOVAK ACADEMY OF SCIENCES
ZITNA 25, 11567 PRAHA 1
CZECHOSLOVAKIA