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 \( \{x_n\}, \{y_n\} \)

\[
\begin{align*}
y_{n+1} &= y_n - [y_n, x_n; f]^{-1} f(y_n) \\
x_{n+1} &= y_{n+1} - [y_n, x_n; f]^{-1} f(y_{n+1})
\end{align*}
\]

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 a priori 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 \(r\):

\[
U(x_0, r) = \{ x \in E \ ; \ d(x, x_0) \leq r \} .
\]

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

\[
d(A) = \inf \{ |Ax| \ ; \ |x| \geq 1 \} .
\]

Clearly \(d(A) = |A^{-1}|^{-1}\) if \(A\) is invertible; also \(d(A + B) \geq d(A) - |B|\) 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 \(\omega\) 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 \(n\)th iterate of \(\omega\), 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 $\beta$ is a function with values in $\mathbb{R}^p$, we shall denote by $\beta_i$ its $i$th 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

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

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

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

The functions $\omega$ and $\sigma$ are obviously related by the following functional equation

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

The components of $\sigma$ satisfy the relations

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

If $a$ and $b$ are two elements of $\mathbb{R}^p$, with coordinates $a_i$ and $b_i$ respectively, we write $a \leq b$ if $a_i \leq b_i$ 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 $i$th coordinate is $x$. If $x \in E^p$ we denote by $d^{(p)}(x, A)$ the vector with components $d(x, A_i)$, $i = 1, 2, \ldots, p$. If $r \in \mathbb{R}^p$, $r > 0$, we shall denote by $U(A, r)$ a neighbourhood of $A$ of the form

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

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

$$Z(0) = \bigcap_{s \in T} \left( \bigcup_{t \in S} Z(t) \right)^-.$$

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

**Theorem.** Let $T$ be a $p$-dimensional interval and let $\omega$ be a rate of convergence on $T$. Suppose that the family $Z(t) \subset \mathbb{E}^p$, $t \in T$, satisfies
(7) \[ Z(r) \subset U(Z(\omega(r)), \psi(r)) \]

for each \( r \in T \). Then

(8) \[ Z(r) \subset U(Z(0), \sigma(r)) \]

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 majorized by the terms of a convergent series

\[ d^{(p)}(x_{n+1}, x_n) \leq \psi(\omega^{(n)}(r)) . \]

It is easy to prove that the limit \( x^* \) of this sequence belongs to \( Z(0) \) and that \( d^{(p)}(x^*, x_0) \leq \sigma(r) \). 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 \( (E_p, d_p) \). The induction theorem remains true if the function \( \psi \) is replaced by any other function \( \psi' \) 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

(10) \[ x_{n+1} \in F(x_n) , \]

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 \( \omega \) on \( T \), so that the following conditions are satisfied:

(11) \[ x_0 \in Z(r_0) \quad \text{for a certain } r_0 \in T, \]

(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

\begin{align}
(13) \quad x_n & \in Z(\omega^m(r_0)) \\
(14) \quad d^{(p)}(x_n,x_0) & \leq \sigma(r_0) - \sigma(\omega^m(r_0)) \\
(15) \quad d^{(p)}(x_n,x^*) & \leq \sigma(\omega^m(r_0)).
\end{align}

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 \textit{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

\begin{equation}
(17) \quad d^{(p)}(x_n,x^*) \leq \sigma(\omega(r_{n-1})) = \sigma(r_{n-1}) - \psi(r_{n-1}).
\end{equation}

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
\textit{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

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

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

\[
[u, v; f](u - v) = f(u) - f(v)
\]

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

\[
|[u_2, v_2; f] - [u_1, v_1; f]| \leq H(|u_2 - u_1| + |v_2 - v_1|).
\]

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

\[
|f'(x) - f'(y)| \leq 2H|x - y|.
\]

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

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

If we take \(y = x\), this reduces to the familiar formula:

\[
|f(y') - f(x) - f'(x)(y' - x)| \leq H|y' - x|^2.
\]

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 successive 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

(21) \[ \omega_1(q, r) = \frac{r(q + r)}{q + 2r + 2(r(q + r) + a^2)^{\frac{3}{4}}} \]

(22) \[ \omega_2(q, r) = \omega_1 \frac{q + 2r + \omega_1}{2(r(q + r) + a^2)^{\frac{1}{4}}} - \omega_1 \]

(22') \[ = \frac{r(q + r)(q^2 + 5r(q + r)) + 2r(q + r)(q + 2r)(q(q + r) + a^2)^{\frac{1}{4}}}{(q + 2r)(8a^2 + 7r(q + r)) + 2(4a^2 + q^2 + 7r(q + r))(r(q + r) + a^2)^{\frac{3}{4}}} \]

is a rate of convergence on $T$. The corresponding $\sigma$ function is given by

(23) \[ \sigma_1(q, r) = \beta(q, r) + q + r, \quad \sigma_2(q, r) = \beta(q, r) + r \]

where

(24) \[ \beta(q, r) = (r(q + r) + a^2)^{\frac{1}{4}} - a \]

Proof. Consider the real polynomial

(26) \[ f(x) = x^2 - a^2 \]

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

(27) \[ x_0 = r + (r(q + r) + a^2)^{\frac{1}{4}}, \quad y_0 = x_0 + q \]

we have

(28) \[ y_0 - x_0 = q \]

and

(29) \[ x_0 - y_1 = r \]

Taking now

\[ \omega_1(q, r) = y_1 - x_1, \quad \omega_2(q, r) = x_1 - y_2, \quad \sigma_1(q, r) = y_0 - a \]
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:

\begin{align*}
(30) \\ y_n &- x_n = \omega_1^{(n)}(q, r) \\
(31) \\ x_n &- x_{n+1} = \omega_2^{(n)}(q, r) \\
(32) \\ x_n &- x_0 = \sigma_2(q, r) - \sigma_2(\omega^{(n)}(q, r)) \\
(33) \\ y_n &- x_0 = \sigma_2(q, r) - \sigma_1(\omega^{(n)}(q, r)) \\
(34) \\ x_n &- a = \sigma_2(\omega^{(n)}(q, r)) \\
(35) \\ y_n &- a = \sigma_1(\omega^{(n)}(q, r)).
\end{align*}

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

\begin{align*}
(36) \\ a = \frac{1}{2H} (d_0^2 + H^2 q_0^2 - 2d_0 H (q_0 + 2r_0))^\frac{1}{2},
\end{align*}

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 $d([y_0, x_0; f]) \geq d_0$,

2°. $|y_0 - x_0| \leq q_0$,

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

If the initial data satisfy the inequalities

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

5°. $m \geq \sigma_2(q_0, r_0),$

then the iterative procedure

\begin{align*}
&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})
\end{align*}

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:
$6^\circ. \ |x^* - x_0| \leq \sigma_2(q_0, r_0),$

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

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

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

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

where $\omega_1, \omega_2, \beta, \sigma_1, \sigma_2$ 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:

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

where $h$ is a positive function to be determined later.

We intend to prove that

$$(37) \quad (y_0, x_0) \in Z(q_0, r_0)$$

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

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

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

satisfies the inclusion

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

for at suitable rate of convergence $\omega$ on $T$.

The inclusion (40) is equivalent to

$$(41) \quad |y' - x'| \leq \omega_1(q, r)$$

$$(42) \quad d([y', x'; f]) \geq h(\omega(q, r))$$

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

$$(44) \quad |y' - y| \leq q + r$$

$$(45) \quad |x' - x| \leq r + \omega_1(r, q).$$

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

Using this we have
\[
  f(y') = f(x) + [y, x; f](y' - x) + ([y', x; f] - [y, x; f])(y' - x)
  = ([y', x; f] - [y, x; f])(y' - x),
\]
whence
\[ |f(y')| \leq H|y' - y||y' - x|. \]

From the definition of \( Z \) and from (46) we have \(|y' - x| \leq r\), whence
\[ |y' - y| \leq |y' - x| + |x - y| \leq r + q, \]
so that (44) is established. Furthermore
\[ |y' - x'| \leq \frac{|f(y')|}{h(q, r)} \leq \frac{Hr(q + r)}{h(q, r)}; \]
it follows that (41) will be satisfied if we assume that
\[ \frac{Hr(q + r)}{h(q, r)} \leq \omega_1(q, r). \]

Let us do that. Now
\[
  d([y', x'; f]) \geq d([y, x; f]) - [[y', x'; f] - [y, x; f]]
  \geq h(q, r) - H(|y' - y| + |x' - x|).
\]

We already know that \(|y' - x| \leq r\); since
\[ |x' - x| \leq |x' - y'| + |y' - x| \leq \omega_1(q, r) + r \]
we have (45) and
\[ d([y', x'; f]) \geq h(q, r) - H(q + 2r + \omega_1(q, r)). \]

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

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

\[ \omega_1 = \frac{r(q + r)}{k} \]

\[ k\omega = k - (q + 2r + \omega_1) \]

\[ \omega_2 = \omega_1 \frac{q + 2r + \omega_1}{k\omega} \]

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

\[ k(q, r) = q + 2r + 2(r(q + r) + a^2)^{\dagger} \]

and take for \( \omega_1, \omega_2 \) 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

\[ a = \frac{1}{2H} (d_0^2 + H^2q_0^2 - 2d_0H(q_0 + 2r_0))^{\dagger} \]

We have to remark that condition 40 of the Theorem implies that

\[ d_0^2 + H^2q_0^2 \geq 2d_0H(q_0 + 2r_0) \]

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 \( |y_n - x_n| \leq \omega^{(n)}_1(q_0, r_0) \), so that both sequences \( x_n \) and \( y_n \) converge to the same limit point \( x^* \). Since by (47)

\[ |f(y_{n+1})| \leq H|y_{n+1} - y_n||y_{n+1} - x_n| \]

and since (19) implies the continuity of \( f \) we have \( f(x^*) = 0 \). The estimates 60, 70, 80 are immediate consequences of (13), (14), (15). The a posteriori estimates 90, 100 are immediate consequences of (17) if we prove that

\[ (y_{n+1}, x_{n-1}) \in Z(|y_{n-1} - x_{n-1}|, |x_{n-1} - y_n|) \]

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

Since we know that \((y_{n-1}, x_{n-1}) \in Z(\omega^{(n-1)}(q_0, r_0))\) we have
\[ d([y_{n-1}, x_{n-1}; f]) \geq h(\omega^{(n-1)}(q_0, r_0)) \]
and the desired inequality is a consequence of the monotonicity of \(h\) and the inequality
\[ (|y_{n-1} - x_{n-1}|, |x_{n-1} - y_n|) \leq \omega^{(n-1)}(q_0, r_0) \, . \]
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
\[ (60) \quad \frac{d_0}{H} \geq q_0 + 2r_0 + 2(r_0(q_0 + r_0) + a^2)^\frac{1}{2} \, . \]

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
\[ ([y_0, x_0; f]^{-1}([y_0, x_0; f] - [z, x^*; f])) \leq \frac{H}{d_0} \, (|y_0 - x_0| + |x_0 - z| + |x_0 - x^*|) \leq \frac{H}{d_0} \, (q_0 + 2\sigma_2(q_0, r_0) + 2a) \]
\[ \frac{H}{d_0} (q_0 + 2r_0 + a + 2(r_0(q_0 + r_0) + a^2)^{\frac{1}{4}}) = 1 , \]

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°. \((r_0^5 + (q_0 + r_0)^4)^2 \leq d_0/H.\)

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 \(4Hr_0 \leq d_0\), 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 \(f: \mathbb{R} \to \mathbb{R}\) and two points \(x_0, y_0 \in \mathbb{R}\) 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 \leq \frac{d_0}{H} \leq 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 \(4^\circ\), there exists a function \(f: \mathbb{R} \to \mathbb{R}\) and two points \(x_0, y_0 \in \mathbb{R}\) which satisfy the hypotheses of Theorem 3.2, and for which all the estimates \(6^\circ, 7^\circ, 8^\circ, 9^\circ, 10^\circ\) are attained.

Proof. As we have already remarked, in the proof of Theorem 3.2, condition \(4^\circ\) implies that \(d_0^2 + H^2q_0^2 \geq 2d_0H(q_0 + 2r_0)\), 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

\[\frac{f(y_0) - f(x_0)}{y_0 - x_0} = d_0, \quad \frac{f(x_0)}{d_0} = r_0 \, ,\]

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

7. F.-A. Potra, \textit{An application of the Induction Method of V. Pták to the study of Regula Falsi,}
11. V. Pták, \textit{A quantitative refinement of the closed graph theorem,} Czechoslovak Math. J. 99
12. V. Pták, \textit{Nondiscrete mathematical induction and iterative existence proofs,} Linear Algebra
and Appl. 13 (1976), 223–238.
Carolinae 16 (1975), 699–705.
15. V. Pták, \textit{A rate of convergence,} Numerical Functional Analysis and Optimization 1 (1979),
255–271.
19. V. Pták, \textit{Nondiscrete Mathematical Induction,} in: General Topology and its Relations to
Modern Analysis and Algebra IV (Lecture Notes in Math. 609), pp. 166–178, Springer-
22. J. W. Schmidt und H. Schwetlick, \textit{Ableitungsfreie Verfahren mit hoherer Konvergenzgeschwindig-
(Basel) 7 (1956), 471–484.
Füüs.-Mat. 16 (1967), 13–26, 146–156.
(1978), 1–12.
27. J. Zemánek, \textit{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. PÁČII 220, 77538 BUCHAREST
ROMANIA

AND

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