ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 37, Number 2, 2007

CONVERGENCE BY NONDISCRETE MATHEMATICAL INDUCTION OF A TWO STEP SECANT'S METHOD

S. AMAT, C. BERMÚDEZ, S. BUSQUIER AND J. GRETAY

ABSTRACT. We use the nondiscrete mathematical induction method for the semi-local convergence of a two step secant's iterative scheme on a Banach space. The scheme does not need to evaluate neither any Fréchet derivative nor any bilinear operator, but having a high speed of convergence.

1. Introduction. We consider the problem of approximating locally the unique solution x^* of a nonlinear equation

$$f(x) = 0, \tag{1}$$

where f is a continuous operator defined on the closed convex domain D of a Banach space E_1 with values in a Banach space E_2.

We use the two step Steffensen's method given by

$$\begin{aligned} y_{n+1} &= x_n - [x_n, x_n + \alpha_n(y_n - x_n); f]^{-1} f(x_n), \\ x_{n+1} &= y_{n+1} - [x_n, x_n + \alpha_n(y_n - x_n); f]^{-1} f(y_{n+1}), \quad n \geq 0. \end{aligned} \tag{2}$$

to generate two sequences converging to x^*.

In order to control the stability in practice, the α_n can be computed such that

$$\mathrm{tol}_c << |\alpha_n(y_n - x_n)| \leq \mathrm{tol}_u,$$

where tol_c is related with the computer precision and tol_u is a free parameter.

Here $[x, y; f] \in L(E_1, E_2)$ is a divided difference of order one for the operator f on the points $x, y \in D$. If f is Fréchet differentiable, then

$$[x, y; f] = \int_0^1 f'(x + t(y - x))\, dt, \quad \text{for all} \quad x, y \in D. \tag{3}$$

2000 AMS *Mathematics Subject Classification.* Primary 65F15, 47H17, 49D15.
Key words and phrases. Nonlinear equations, iterative methods, nondiscrete mathematical induction, semi-local convergence.
Received by the editors on September 28, 2004.

In [12], it is proved that these type of divided differences are characterized by the following property

$$[x, y; f] = 2[x, 2y - x; f] - [y, 2y - x; f],$$

for all $x, y \in D$, $x \neq y$ and $2y - x \in D$.

Our iterative method is related with the scheme analyzed in [14] that writes

$$\begin{aligned} y_{n+1} &= x_n - [x_n, y_n; f]^{-1} f(x_n), \\ x_{n+1} &= y_{n+1} - [x_n, y_n; f]^{-1} f(y_{n+1}), \quad n \geq 0. \end{aligned} \tag{4}$$

In practice, as we can see in [2–4], the α_n parameter improves the convergence of the schemes and therefore our scheme becomes more efficient.

In a collection of papers [11, 13, 14] as well as in the monograph [15], the authors present a unified theory of convergence based on the so-called "method of nondiscrete mathematical induction." The theory was introduced in the paper [16] where the classical closed graph theorem and nonlinear mapping are studied.

The basic idea is to define the rate of convergence as a function, not as a number, giving more information than the classical methods, which makes possible to give error estimates that are sharp even throughout the whole process.

In this paper we intend to apply these ideas to the study of the convergence of (2).

2. 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 $T_1 \times \cdots \times T_p$.

Definition 1. A function w mapping T into T is said to be a rate of convergence on T if the series

$$\sum_{n=0}^{\infty} w^{(n)}(r)$$

is convergent for each $r \in T$.

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

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

$$\psi(r) = (r_1 + \cdots, r_p, r_2 + \cdots + r_p + w_1(r), \dots, \\ r_p + w_1(r) + \cdots + w_{p-1}(r)). \tag{5}$$

Using this function, we can attach to the rate of convergence w the function

$$\sigma(r) = \sum_{n=0}^{\infty} \psi(w^{(n)}(r)) \tag{6}$$

The functions w and σ are related by

$$\sigma(r) = \sigma(w(r)) + \psi(r), \quad r \in T \tag{7}$$

If a and b are two vectors, we write $a \le b$, $a < b$, if $a_i \le b_i$, $a_i < b_i$.

Let (E, d) be a complete metric space, and let A be a subset of E^p. For each $i = 1, 2, \dots, p$, 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 $d(x_i, A_i)$, $i = 1, 2, \dots, p$. If $r \in \mathbf{R}^p$, $r > 0$, we shall denote by $U(A, r)$ a subset of A of the form

$$U(A, r) = \{x \in E^p : d^{(p)}(x, A) \le r\}$$

Now we recall the following generalization of the Induction theorem [**14**].

Theorem 1. *Let (E, d) be a complete metric space. Let T be a p-dimensional interval, and let w be a rate of convergence on T. Suppose that the family $Z(t) \subset E^p$, $t \in T$, satisfies*

$$Z(r) \subset U(Z(w(r)), \psi(r)) \tag{8}$$

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

$$Z(r) \subset U(Z(0), \sigma(r)) \tag{9}$$

for each $r \in T$, *where*

$$Z(0) = \bigcap_{s \in T} \left(\bigcup_{t \leq s} Z(t) \right)$$

Let us sketch 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 multi-valued mapping of E^p into E^p.

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

1) For a given $r_0 \in T$

$$x_0 \in Z(r_0) \tag{11}$$

2) $r \in T$ and $x \in Z(r)$ imply

$$F(x) \bigcap Z(w(r)) \bigcap U(x, \psi(r)) \tag{12}$$

is nonvoid.

Then it follows from the induction theorem that $Z(0)$ is nonvoid. Moreover, the sequence $\{x_n\}$ obtained by successive applications of 2) converges to a point $x^* \in E^p$ and satisfies the following relations

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

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

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

The equation (15) gives us an a priori estimate of the distance between the approximations and the solution. Moreover, if we find $r_{n-1} \in T$

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 following a posteriori estimates,

$$d^{(p)}(x_n, x^*) \le \sigma(w^{(n)}(r_{n-1})) = \sigma(r_{n-1}) - \psi(r_{n-1}). \tag{16}$$

Usually, an $r_{n-1} \in T$ such that $x_{n-1} \in Z(r_{n-1})$ can be found as a function of x_{n-1} and x_n.

3. Convergence analysis. In this section we give sufficient conditions for the convergence of (2).

The iterations to be considered are successive constructions of pairs of points. Thus, we shall work in two dimensions so that $p = 2$. We take for T the positive quadrant; instead of r_1, r_2 we shall write q, r.

Following [**14**], we consider the real polynomial

$$f(x) = x^2 - a^2$$

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

$$\begin{aligned} y_{n+1} &= x_n - \frac{(x_n + \alpha_n(y_n - x_n)) - x_n}{f(x_n + \alpha_n(y_n - x_n)) - f(x_n)} f(x_n) \\ x_{n+1} &= y_{n+1} - \frac{(x_n + \alpha_n(y_n - x_n)) - x_n}{f(x_n + \alpha_n(y_n - x_n)) - f(x_n)} f(y_{n+1}) \end{aligned}$$

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

Choosing $y_0 = x_0 + q$ and x_0 such that $x_0 - y_1 = r$, if we define

$$w_1(q, r) = y_1 - x_1, \quad w_2(q, r) = x_1 - y_2,$$

then, from the convergence of the process, the pair of functions (w_1, w_2) is a rate of convergence on T.

We observe also that the following relations hold:

$$\begin{aligned} y_n - x_n &= w_1^{(n)}(q, r) \\ x_n - y_{n+1} &= w_2^{(n)}(q, r) \\ x_n - x_0 &= \sigma_2(q, r) - \sigma_2(w^{(n)}(q, r)) \end{aligned}$$

$$\begin{aligned} y_n - x_0 &= \sigma_2(q,r) - \sigma_1(w^{(n)}(q,r)) \\ x_n - a &= \sigma_2(q,r) \\ y_n - a &= \sigma_1(q,r). \end{aligned}$$

We shall use the above results in the proof of the following theorem.

Theorem 2. *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:

a) *The operator $[x_0, x_0 + \alpha_0(y_0 - x_0); f]$ is invertible and*

$$d([x_0, x_0 + \alpha_0(y_0 - x_0); f]) \geq d_0,$$

where $d([x_0, x_0 + \alpha_0(y_0 - x_0); f]) = |[x_0, x_0 + \alpha_0(y_0 - x_0); f]^{-1}|^{-1}$;

b) $\max\{|y_0 - x_0|, \mathrm{tol}_u\} \leq q_0$;

c) $|[x_0, x_0 + \alpha_0(y_0 - x_0); f]^{-1} f(x_0)|$;

d) $2r_0(tol_u + r_0) + \mathrm{tol}_u + r_0 \leq d_0/H$;

e) $m \geq \sigma_2(q_0, r_0)$;

then the iterative procedure

$$\begin{aligned} y_{n+1} &= x_n - [x_n, x_n + \alpha_n(y_n - x_n); f]^{-1} f(x_n), \\ x_{n+1} &= y_{n+1} - [x_n, x_n + \alpha_n(y_n - x_n); f]^{-1} f(y_{n+1}), \quad n \geq 0. \end{aligned} \tag{17}$$

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

$$\begin{aligned} |y_n - x^*| &\leq \sigma_1(w^{(n)}(q_0, r_0)) \\ |x_n - x^*| &\leq \sigma_2(w^{(n)}(q_0, r_0)). \end{aligned}$$

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

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

where $\alpha(y,x)$ is a function in $(0,1)$ such that $|\alpha(y,x)(y-x)| \leq \mathrm{tol}_u$ and h is a positive function to be determined later.

We intend to prove that

$$(y_0, x_0) \in Z(q_0, r_0) \tag{18}$$

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

$$\begin{aligned} y' &= x - [x, x+\alpha(y,x)(y-x); f]^{-1} f(x) \\ x' &= y' - [x, x+\alpha(y,x)(y-x); f]^{-1} f(y') \end{aligned}$$

satisfies the inclusion

$$(y', x') \in Z(w(q,r)) \bigcap U((y,x), \psi(q,r)) \tag{19}$$

for a suitable rate of convergence w on T.

The inclusion (19) is equivalent to

$$|y' - x'| \leq w_1(q,r) \tag{20}$$

$$d([x', x' + \alpha(y', x')(y' - x'); f]) \geq h(w(q,r)) \tag{21}$$

$$|[x', x' + \alpha(y', x')(y' - x'); f]^{-1} f(x')| \leq w_2(q,r) \tag{22}$$

$$|y' - y| \leq q + r \tag{23}$$

$$|x' - x| \leq r + w_1(q,r). \tag{24}$$

Suppose $(y,x) \in Z(q,r)$; it follows from the definition of the scheme that

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

Using this, we have

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

whence

$$|f(y)| \leq H\, |y' - (x + \alpha(x,y)(y-x))|\, |y' - x|. \tag{25}$$

From the definition of Z and the scheme, we have $|y' - x| \leq r$; thus,

$$|y' - y| \leq |y' - x| + |x - y| \leq r + q$$

so that (23) follows, and

$$\begin{aligned}|y' - (x + \alpha(y,x)(y-x))| &\leq |y' - x| + |x - (x + \alpha(y,x)(y-x))| \\ &\leq r + tol_u.\end{aligned}$$

Furthermore,

$$\begin{aligned}|y' - x'| &= |-[x, x + \alpha(y,x)(y-x); f]^{-1} f(y')| \\ &\leq \frac{f(y')}{h(q,r)} \\ &\leq \frac{Hr(r + \mathrm{tol}_u)}{h(q,r)}.\end{aligned} \tag{26}$$

It follows that (20) will be satisfied if we assume that

$$\frac{Hr(r + \mathrm{tol}_u)}{h(q,r)} \leq w_1(q,r).$$

Let us do that. Now

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

Since

$$|x' - x| \leq |x' - y'| + |y' - x| \leq w_1(q,r) + r,$$

we have (24), similarly

$$|\alpha(y',x')(y' - x') - \alpha(y,x)(y-x)| \leq w_1(q,r) + 2\mathrm{tol}_u$$

and

$$d([y', x'; f]) \geq h(q,r) - 2H(w_1(q,r) + r + \mathrm{tol}_u). \tag{27}$$

To estimate $f(x')$ we write

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

Thus,

$$\begin{aligned} |[x', x'+&\alpha(y', x')(y' - x'); f]^{-1} f(x')| \\ &\le \frac{H(|x - x'|\ |x + \alpha(y,x)(y-x) - y'|)w_1(q,r)}{h(q,r) - 2H(\mathrm{tol}_u + r + w_1(q,r))} \\ &\le \frac{H(2r + \mathrm{tol}_u + w_1(q,r))w_1(q,r)}{h(q,r) - 2H(\mathrm{tol}_u + r + w_1(q,r))}. \end{aligned} \tag{28}$$

To simplify the formulae, set $h(q,r) = Hk(q,r)$. To satisfy (20)–(22), it will be sufficient, in view of (26)–(28), to consider k such that

$$\begin{aligned} w_1 &= \frac{r(\mathrm{tol}_u + r)}{k} \\ k \circ w &= k - 2(w_1 + r + \mathrm{tol}_u) \\ w_2 &= w_1\,\frac{2r + \mathrm{tol}_u + w_1}{k \circ w}, \end{aligned}$$

where w_1 and w_2 are the functions of the discussion above the present theorem.

These functional equations are satisfied if we set

$$k(q,r) = \mathrm{tol}_u + r + 2(r(\mathrm{tol}_u + r) + a^2)^{1/2}.$$

Here, the free parameter a will be chosen so as to have $(x_0, y_0) \in Z(q_0, r_0)$. For this it suffices to satisfy $h(q_0, r_0) = d_0$, which leads to the choice

$$a^2 = \frac{1}{2}\left(\frac{d_0}{H} - \mathrm{tol}_u - r_0\right) - r_0(\mathrm{tol}_u + r_0). \tag{29}$$

The condition d) of the theorem implies that

$$\frac{1}{2}\left(\frac{d_0}{H} - \mathrm{tol}_u - r_0\right) - r_0(\mathrm{tol}_u + r_0) \ge 0.$$

Then the formula (29) makes sense. The rest of the proof is a consequence of the induction theorem [**14**]. □

If the condition d) is strict, then x^* is the only root of the equation $f(x) = 0$ in the set

$$U \bigcap \{x \in E : |x - x_0| < \sigma_2(q_0, r_0) + 2a\},$$

see [**14**] for more details.

The application of the method of nondiscrete mathematical induction to any iterative processes carries 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 [**14**].

Acknowledgments. Thanks are due to Professor F. Potra for introducing us to the nondiscrete mathematical induction method. Research supported in part by 00675/PI/04 and MTM2004-07114.

REFERENCES

1. C. Altman, *Iterative methods of higher order*, Bull. Acad. Polon. Sci. Sér. Sci. Math., Astr., Phys. **9** (1961), 62–68.

2. S. Amat and S. Busquier, *On a higher order secant method*, Appl. Math. Comput. **141** (2003), 321–329.

3. ———, *Convergence and numerical analysis of a family of two step Steffensen's methods*, Comput. Math. Appl. **49** (2005), 13–22.

4. S. Amat, S. Busquier and V.F. Candela, *A class of quasi-Newton Generalized Steffensen methods on Banach spaces*, J. Comput. Appl. Math. **149** (2003), 397–406.

5. I.K. Argyros, *An error analysis for the Steffensen method under generalized Zabrejko-Nguen-type assumptions*, Rev. Anal. Numér. Théor. Approx. **25** (1996), 11–22.

6. ———, *A new convergence theorem for Steffensen's method on Banach spaces and applications*, Southwest J. Pure Appl. Math. **1** (1997), 23–29.

7. I.K. Argyros and F. Szidarovszky, *The theory and application of iterative methods*, CRC Press, Boca Raton, Florida, 1993.

8. M.A. Hernández and M.J. Rubio, *Semilocal convergence of the secant method under mild convergence conditions of differentiability*, Comput. Math. Appl. **44** (2002), 277–285.

9. M.A. Hernández, M.J. Rubio and J.A. Ezquerro, *Secant-like methods for solving nonlinear integral equations of the Hammerstein type*, J. Comput. Appl. Math. **115** (2000), 245–254.

10. I. Pavaloiu, *Sur une generalization de la methode de Steffensen*, Rev. Anal. Numér. Théor. Approx. **21** (1992), 59–65.

11. F.A. Potra, *Sharp error bounds for a class of Newton-like methods*, Libertas Math. **5** (1985), 71–84.

12. ———, *A characterisation of the divided differences of an operator which can be represented by Riemann integrals*, Rev. Anal. Numér. Théor. Approx. **9** (1981), 251–253.

13. F.A. Potra and V. Pták, *A generalization of regula falsi*, Numer. Math. **36** (1981), 333–346.

14. ———, *Nondiscrete induction and a double step secant method*, Math. Scand. **46** (1980), 236–250.

15. ———, *Nondiscrete induction and iterative processes*, Research Notes Math. **103**, Pitman, Boston, 1984.

16. V. Pták, *Some metric aspects of the open mapping and closed graph theorems*, Math. Ann. **165** (1966), 95–104.

17. W. Werner, *Newton-like methods for the computation of fixed points*, Comput. Math. Appl. **10** (1984), 77–86.

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Paseo de Alfonso XIII, 52, 30203 Cartagena (Murcia), Spain
E-mail address: sergio.amat@upct.es

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Paseo de Alfonso XIII, 52, 30203 Cartagena (Murcia), Spain

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Paseo de Alfonso XIII, 52, 30203 Cartagena (Murcia), Spain
E-mail address: sonia.busquier@upct.es

Ecole Généraliste d'Ingénieurs de Marseille (EGIM)