M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

FLORIAN A. POTRA

On superadditive rates of convergence

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 19, nº 4 (1985), p. 671-685

<http://www.numdam.org/item?id=M2AN_1985__19_4_671_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE
(vol. 19, n° 4, 1985, p. 671 à 685)

ON SUPERADDITIVE RATES OF CONVERGENCE (*)

by Florian A. POTRA [1]

Communicated by Françoise CHATELIN

Abstract. — *In the present paper we prove that a superadditive function is a rate of convergence, in the sense of V. Pták, if and only if its iterates are pointwise convergent to zero. Moreover if the function is continuous from the right then this is equivalent to the inequality* $w(t) < t$. *These results are generalized for rates of convergence of several variables.*

Résumé. — *Dans ce travail nous démontrons qu'une fonction superadditive est un taux de convergence, dans le sens de V. Pták, si et seulement si ses itérées convergent ponctuellement vers zéro. De plus, si la fonction est continue à droite, alors ce fait est équivalent à l'inégalité* $w(t) < t$. *Ces résultats sont généralisés aux taux de convergence de plusieurs variables.*

1. INTRODUCTION

The method of nondiscrete induction was introduced in 1966 by V. Pták [16] in connection with some quantitative refinements of the closed graph theorem. Since then the method has successfully been applied to the study of various iterative constructions in analysis and numerical analysis. The results were published in a series of papers ([1, 2, 4-28, 30]) and they were recently collected in a book [15]. Many of the theorems obtained by using this method are sharp in the sense that neither the convergence conditions nor the error estimates can be improved in the class of problems considered. This is explained by the fact that in the application of the method of nondiscrete induction the classical way of measuring convergence is replaced by a more refined one which makes it possible to obtain estimates sharp not only asymptotically but throughout the whole process. A rate of convergence is defined as a function, not as a number.

(*) Received in December 1984.
(1) Department of Mathematics, University of Iowa, Iowa City, Iowa 52242, U.S.A.

DEFINITION 1.1 : *Let T denote either the positive semiaxis* $]0, \infty[$ *or a half open interval of the form* $]0, b]$. *A function* $w : T \to T$ *is called a rate of convergence on* T *if*

$$\sum_{n=0}^{\infty} w^{(n)}(t) < \infty, \qquad t \in T \tag{1}$$

where $w^{(n)}$ *denotes the* nth *iterate of* w *in the sense of the usual function composition* (i.e., $w^{(0)}(t) = t$, $w^{(n+1)}(t) = w(w^{(n)}(t))$, $n = 0, 1, 2, \ldots$). □

By defining the rate of convergence of an iterative procedure as a function we can retain more information at each step and obtain a sharp final estimate. This departure from tradition in measuring convergence is justified by V. Pták in a paper suggestively entitled « What should be a rate of convergence ? » [25]. In the same paper he stresses the importance of the rates of convergence which satisfy a functional inequality of the form

$$s \circ w \leqslant w \circ s \tag{2}$$

where $s(t)$ denotes the sum of the series (1). This inequality is not implied by Definition 1.1 and V. Pták shows that it is satisfied at least by convex rates of convergence. It is easy to prove that (2) holds for superadditive rates of convergence continuous from the left (see Proposition 3.2). Every convex rate of convergence is superadditive but the converse is not true. For example the rate of convergence of Newton's method

$$w(t) = \frac{t^2}{2\sqrt{t^2 + a^2}} \tag{3}$$

is superadditive on $\mathbb{R}_+$ but it is convex only on the interval $[0, a\sqrt{2}]$ (see definitions in Sections 2.1 and 2.7).

These facts demonstrate that the class of superadditive rates of convergence deserves special attention. In the present paper we intend to show that for superadditive functions $w : T \to T$ condition (1) is equivalent to the simpler condition

$$\lim_{n\to\infty} w^{(n)}(t) = 0, \qquad t \in T. \tag{4}$$

Moreover, if $w : T \to T$ is superadditive and continuous from the right then (1) is equivalent to the inequality

$$w(t) < t, \qquad t \in T. \tag{5}$$

The above-mentioned results will be proved in Section 3 in the more general context of p-dimensional rates of convergence. The notion of a p-dimensional rate of convergence (or rate of convergence of type (p. 1)) was first introduced in [7] in connection with the study of Regula Falsi and it represents a natural generalization of the notion given in Definition 1.1 (see also [9] and [15]).

2. CONVEXITY AND SUPERADDITIVITY

We have mentioned in the introduction that convex rates of convergence are superadditive. This is a consequence of the fact that a convex function $f : \mathbb{R}_+ \to \mathbb{R}$ which vanishes at zero is superadditive. This fact holds for functions of several variables as well, with a proper generalization of the notion of convexity. We will consider the notion of S-convexity, which was introduced (in a more general context) by J. W. Schmidt and H. Leonhardt [29], and will compare it with some more popular generalizations of the notion of a convex function, in order to give the reader a better understanding of its significance.

In what follows we will consider the Euclidean space $\mathbb{R}^p$ endowed with the natural (component-wise) partial ordering « $\leqslant$ » (i.e., for any $u = (u_1, ..., u_p)$, $v = (v_1, ..., v_p)$ from $\mathbb{R}^p$ the relation $u \leqslant v$ is equivalent to $u_i \leqslant v_i$ for $i = 1, 2, ..., p$). The set $\mathbb{R}^p_+$ is called the non-negative cone of $\mathbb{R}^p$. Two elements x, y of $\mathbb{R}^p$ are called comparable if either $x \leqslant y$ or $y \leqslant x$ holds. We will identify the space $(\mathbb{R}^p)^*$ of all linear functionals defined on $\mathbb{R}^p$ with the space $\mathbb{R}^p$ itself. This motivates also the notation uv for the scalar product between the vectors u and v. We will denote by $e_1, e_2, ..., e_p$ the standard basis of $\mathbb{R}^p$ and by e the vector $(1, 1, ..., 1)$. Obviously $e = e_1 + e_2 + \cdots + e_p$.

DEFINITION 2.1 : *Let D be a convex subset of $\mathbb{R}^p$. A function $f : D \to \mathbb{R}$ is called :*

a) convex, *if*

$$f(\lambda x + (1 - \lambda) y) \leqslant \lambda f(x) + (1 - \lambda) f(y) \tag{6}$$

for all $\lambda \in [0, 1]$ and all $x, y \in D$;

b) order convex, *if* (6) *holds for all $\lambda \in [0, 1]$ and all* comparable x, y *from D ;*

c) S-convex, *if there is a mapping*

$$\delta f(., .) : \Delta := \{ (x, y) \in D \times D ; x \neq y \} \to \mathbb{R}^p ,$$

$$\delta f(x, y)(x - y) = f(x) - f(y) , \quad \textit{for all} \quad (x, y) \in \Delta , \tag{7}$$

$$\delta f(x, y) \leqslant \delta f(u, v) , \quad \textit{for all} \quad (x, y), (u, v) \in \Delta \quad \textit{with} \quad x \leqslant u, y \leqslant v . \quad \square \tag{8}$$

For $p = 1$ the above defined notions coincide. If $p > 1$ convexity implies order-convexity, but the reverse is not true. A similar statement holds if we replace convexity by S-convexity.

PROPOSITION 2.2 : *Let D be a convex subset of $\mathbb{R}^p$ and let $f : D \to \mathbb{R}$ be an S-convex function. Then f is also order-convex.*

Proof : Consider two points $x, y \in D$ with $x \leqslant y$ and let λ be a number between 0 and 1. Denote $v = \lambda x + (1 - \lambda) y$. We have obviously $x \leqslant v \leqslant y$, $v - x = (1 - \lambda)(y - x)$, $y - v = \lambda(y - x)$ so that we can write :

$$\begin{aligned}\lambda f(x) + (1 - \lambda) f(y) - f(v) &= \lambda(f(x) - f(v)) + (1 - \lambda)(f(y) - f(v)) \\ &= \lambda \delta f(x, v)(x - v) + (1 - \lambda) \delta f(y, v)(y - v) \\ &= \lambda(1 - \lambda)(\delta f(y, v) - \delta f(x, v))(y - x) \geqslant 0 . \quad \square\end{aligned}$$

We will see in what follows that the reverse of the above proposition is not true. Before that let us give a useful characterization of S-convexity :

PROPOSITION 2.3 : *Let D be a convex subset of $\mathbb{R}^p$. A function $f : D \to \mathbb{R}$ is S-convex if and only if*

$$\frac{1}{s}(f(x + se_i) - f(x)) \leqslant \frac{1}{t}(f(y + te_i) - f(y)) \tag{9}$$

for all $i = 1, 2, \ldots, p$, $x, y \in D$, $s, t \in \mathbb{R} \setminus \{0\}$ satisfying $x \leqslant y$ and

$$x + se_i \leqslant y + te_i . \tag{10}$$

Proof : Suppose (10) is satisfied. If f is S-convex we can write

$$\begin{aligned}\frac{1}{s}(f(x + se_i) - f(x)) &= \frac{1}{s} \delta f(x + se_i, x)\, se_i \\ &= \delta f(x + se_i, x)\, e_i \leqslant \delta f(y + te_i, y)\, e_i \\ &= \frac{1}{t}(f(y + te_i) - f(y))\end{aligned}$$

which proves the necessity of the condition stated in the proposition. Then, taking $x = y$ in (9), it is easy to see that the following limits exist :

$$D_i^+ f(x) = \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x + te_i) - f(x)}{t}$$

$$D_i^- f(x) = \lim_{\substack{t \to 0 \\ t < 0}} \frac{f(x + te_i) - f(x)}{t} .$$

Denote $D_i f(x) = (D_i^- f(x) + D_i^+ f(x))/2$. For any two points $x = (x_1, x_2, \dots, x_p)$, $y = (y_1, y_2, \dots, y_p)$ from $\mathbb{R}^p$ define a vector $\delta f(x, y) \in \mathbb{R}^p$ whose ith component is given by

$$\frac{1}{x_i - y_i}\left[f(x_1, \dots, x_{i-1}, x_i, y_{i+1}, \dots, y_p) - f(x_1, \dots, x_{i-1}, y_i, y_{i+1}, \dots, y_p)\right]$$

if $x_i \neq y_i$; and by

$$D_i f(x_1, \dots, x_{i-1}, x_i, y_{i+1}, \dots, y_p)$$

if $x_i = y_i$. It is easy to check that conditions (7) and (8) are satisfied for the mapping $\delta f(.,.)$ defined as above. The proof is complete. □

In order to clarify the relation between the notions introduced in Definition 2.1, let us suppose that the function f is twice G-differentiable. Then f is convex if and only if

$$f''(x)\, hh \geqslant 0 \quad \text{for all} \quad h \in \mathbb{R}^p, \qquad x \in D. \tag{11}$$

Also, f is order convex if and only if

$$f''(x)\, hh \geqslant 0 \quad \text{for all} \quad h \in \mathbb{R}_+^p, \qquad x \in D. \tag{12}$$

For a proof of the above statements see, for example, [3]. Using Proposition 2.3, we can easily prove that f is S-convex if and only if

$$f''(x)\, hk \geqslant 0 \quad \text{for all} \quad h, k \in \mathbb{R}_+^p, \qquad x \in D. \tag{13}$$

Condition (11) means that the matrix $f''(x)$ is semipositive definite while condition (13) means that the matrix $f''(x)$ has all the entries nonnegative. This observation shows us that there are many functions which are convex but not S-convex as well as functions which are S-convex but not convex. As an example from the first category let us take

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 \tag{14}$$

and as an example from the second category

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x_1, x_2) = x_1^2 + 4\, x_1 x_2 + x_2^2\,. \tag{15}$$

It is known that convex functions are continuous (for a proof, see [3]). For order-convex functions (and even for S-convex functions) we have only continuity from the left and from the right in the sense of the following definition :

DEFINITION 2.4 : *A function* $f : D \subset \mathbb{R}^p \to \mathbb{R}$ *is called continuous from the right (resp. left) at* $x \in D$ *if for any* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that*

$$| f(x) - f(y) | < \varepsilon$$

whenever $y \in D$ *and* $x \leqslant y \leqslant x + \delta e$ (*resp.* $x - \delta e \leqslant y \leqslant x$). $\square$

More precisely we have the following.

THEOREM 2.5 : *Let D be an open convex subset of* $\mathbb{R}^p$ *and let* $f : D \to \mathbb{R}$ *be an order convex function. Then f is continuous from the left and from the right at each point of D.*

The above theorem can be proved by adapting the proof of Theorem 3.4.2 of [3] and using the following lemma.

LEMMA 2.6 : *Let* U_p *be the hypercube*

$$U_p = \left\{ x = (x_1, \ldots, x_p) \in \mathbb{R}^p ; \max_{1 \leqslant j \leqslant p} | x_j | \leqslant 1 \right\}$$

and let $u_1^{(p)}, u_2^{(p)}, \ldots, u_{2^p}^{(p)}$ *be its vertices. If* $f : U_p \to \mathbb{R}$ *is order convex then for any* $x \in U_p$ *we have*

$$f(x) \leqslant \max_{1 \leqslant k \leqslant 2^p} f(u_k^{(p)}) .$$

Proof : Our lemma is trivially verified for $p = 1$. Suppose it holds for a given $p \geqslant 1$. By reordering the vertices of U_{p+1} we may suppose that the last coordinate of

$$u_1^{(p+1)}, u_2^{(p+1)}, \ldots, u_{2^p}^{(p+1)} \tag{16}$$

is equal to -1 while the last coordinate of

$$u_{2^p+1}^{(p+1)}, u_{2^p+2}^{(p+1)}, \ldots, u_{2^{p+1}}^{(p+1)} \tag{17}$$

equals 1. Let us denote by U_p' the convex envelope of the points (16) and by U_p'' the convex envelope of the points (17). From the induction hypothesis we have

$$f(x') \leqslant \max_{1 \leqslant k \leqslant 2^p} f(u_k^{(p+1)}) , \qquad x' \in U_p'$$

and

$$f(x'') \leqslant \max_{2^p + 1 \leqslant k \leqslant 2^{p+1}} f(u_k^{(p+1)}) , \qquad x'' \in U_p'' .$$

Now any $x = (x_1, x_2, \ldots, x_p, x_{p+1}) \in U_{p+1}$ can be written as

$$x = \lambda x' + (1 - \lambda) x'' , \qquad \lambda \in [0, 1]$$

with $x' = (x_1, ..., x_p, -1) \in U'_p$, $x'' = (x_1, ..., x_p, 1) \in U''_p$. We have clearly $x' \leqslant x''$ so that using the order convexity of f we deduce that

$$f(x) \leqslant \lambda f(x') + (1-\lambda) f(x'') \leqslant \max_{1 \leqslant k \leqslant 2^{p+1}} f(u_k^{(p+1)}) .$$

The proof is complete. □

We note that the above lemma does not hold if we replace U_p by some other convex polyhedron. Indeed if we consider the function f given by (15) and the square of vertices (1, 0), (0, 1), (− 1, 0), (0, − 1) we have

$$f(1,0)=f(0,1)=f(-1,0)=f(0,-1)=1 \quad \text{and} \quad f(-1/2,-1/2)=1+1/2 .$$

Let us state now two very simple results to be used later :

DEFINITION 2.7 : *A function* $f : D \subset \mathbb{R}^p \to \mathbb{R}$ *is called superadditive if*

$$f(x+y) \geqslant f(x) + f(y) \quad \textit{for all} \quad x, y \in D \quad \textit{with} \quad x + y \in D .$$

DEFINITION 2.8 : *A function* $f : D \subset \mathbb{R}^p \to \mathbb{R}$ *is called isotone if*

$$f(x) \leqslant f(y) \quad \textit{whenever} \quad x, y \in D \quad \textit{and} \quad x \leqslant y .$$

PROPOSITION 2.9 : *Let T denote either the whole positive axis $]0, \infty[$ or a half open interval of the form $]0, b]$ and set*

$$D_0 = T^p \cup \{0\}$$

where 0 is the origin of $\mathbb{R}^p$. If $f : D_0 \to \mathbb{R}$ is S-convex and $f(0) = 0$ then f is superadditive.

Proof : $f(x+y) - f(x) - f(y) = f(x+y) - f(x) - (f(y) - f(0)) = (\delta f(x+y, x) - \delta f(y, 0))\, y \geqslant 0$. □

PROPOSITION 2.10 : *If $f : D \subset \mathbb{R}^p \to \mathbb{R}$ is superadditive and nonnegative (i.e., $f(D) \subset \mathbb{R}_+$) then f is isotone.*

Proof : If $x, y \in D$ and $x \leqslant y$ then

$$f(y) = f(x + y - x) \geqslant f(x) + f(y - x) \geqslant 0 . \quad \square$$

It is interesting to remark that Proposition 2.9 does not hold if we replace S-convexity by convexity. For example the function f, given by (14) is convex but if we take $x = (1, 9)$ and $y = (9, 1)$ then we have $f(x+y)=f(10, 10)=100$ and $f(x) + f(y) = 2(81 - 9 + 1) = 146$.

3. p-DIMENSIONAL SUPERADDITIVE RATES OF CONVERGENCE

Let T denote as before either the set of all positive real numbers or a half open interval of the form $]0, b]$. Let w be a mapping of the cartesian product T^p into T and let us consider the « iterates » $w^{(n)}$ of w given for each

$$t = (t_1, t_2, \dots, t_p) \in T^p$$

by the following recurrent scheme :

$$w^{(0)}(t) = t_p, \qquad w^{(n+1)}(t) = w^{(n)}(t_2, \dots, t_p, w(t)), \qquad n = 0, 1, 2, \dots . \tag{18}$$

DEFINITION 3.1 : *A mapping* $w : T^p \to T$ *with the above iteration law is called a p-dimensional rate of convergence (or rate of convergence of type* $(p, 1)$*) on* T *if*

$$\sum_{n=0}^{\infty} w^{(n)}(t) < \infty, \qquad t \in T^p. \tag{19}$$

It is convenient to attach to the mapping $w : T^p \to T$ a mapping $\overline{w} : T^p \to T^p$ defined for every $t = (t_1, t_2, \dots, t_p) \in T^p$ by

$$\overline{w}(t) = (t_2, \dots, t_p, w(t)).$$

It is easily seen that if we denote by $\overline{w}^{(n)}$ the iterates of $\overline{w}$ in the sense of the usual composition of functions (i.e., $\overline{w}^{(0)}(t) = t$, $\overline{w}^{(n+1)}(t) = \overline{w}(\overline{w}^{(n)}(t))$) then from (18) it follows that

$$w^{(n+1)}(t) = w(\overline{w}^{(n)}(t)) = w^{(n)}(\overline{w}(t)), \qquad n = 0, 1, 2, \dots, \tag{20}$$

$$\overline{w}^{(n)}(t) = (w^{(n-p+1)}(t), \dots, w^{(n)}(t)), \qquad n = p-1, p, p+1, \dots . \tag{20'}$$

Let us denote by $s(t)$ the sum of the series (19) and let us consider the mappings $s_i : T^p \to \mathbb{R}$, $i = 1, 2, \dots, p$ given by

$$s_i(t) = s(t) + \sum_{k=i}^{p-1} t_i, \qquad t = (t_1, t_2, \dots, t_p) \in T^p.$$

Finally let us define the mapping

$$\overline{s} : T^p \to T^p, \qquad \overline{s}(t) = (s_1(t), s_2(t), \dots, s_p(t)).$$

With the above notation we have :

PROPOSITION 3.2 : *If the p-dimensional rate of convergence* $w : T^p \to T$ *is superadditive and continuous from the left then*

$$s(\overline{w}(t)) \leqslant w(\overline{s}(t)), \quad \textit{for all} \quad t \in T^p. \tag{21}$$

Proof : First, let us observe that

$$s(\overline{w}(t)) = s(t) - w^{(0)}(t) = \sum_{k=1}^{\infty} w^{(k)}(t) .$$

Using the superadditivity of w and (20) we can write

$$\sum_{k=1}^{n} w^{(k)}(t) = \sum_{k=0}^{n-1} w(\overline{w}^{(k)}(t)) \leqslant w\left(\sum_{k=0}^{n-1} \overline{w}^{(k)}(t)\right).$$

Finally, letting n to tend to infinity in the above inequality, we obtain (21). □

The inequality (21) was first proved by V. Pták [25] for unidimensional convex rates of convergence. The heuristic motivation for the importance of such an inequality given in the above-mentioned paper can be generalized to the p-dimensional case as follows : Suppose $\{ x_n \}_{n \geqslant 0}$ is a sequence of points belonging to a complete metric space (X, d) such that

$$d(x_n, x_{n+1}) \leqslant w(d(x_{n-p}, x_{n-p+1}), \ldots, d(x_{n-1}, x_n))$$

where w is an isotone p-dimensional rate of convergence. It is easy to see that under this assumption the sequence $\{ x_n \}_{n \geqslant 0}$ has to be convergent. Moreover if we denote

$$x^* = \lim_{n \to \infty} x_n , \qquad r_0 = (d(x_0, x_1), \ldots, d(x_{p-1}, x_p))$$

then we have the estimates

$$d(x_n, x_{n+1}) \leqslant e_{n,n+1} := w^{(n-p+1)}(r_0)$$
$$d(x_n, x^*) \leqslant e_n := s(\overline{w}^{(n-p+1)}(r_0)) .$$

If the sequence $\{ x_n \}_{n \geqslant 0}$ is obtained *via* an iterative procedure then at a certain stage the distances $d(x_n, x_{n+1})$ are known for, say, all $n \leqslant N$ while the distances $d(x_n, x^*)$ are generally not known. It is then important to note that if (21) holds then the estimates e_n will satisfy a relation similar to the relation satisfied by the distances $d(x_n, x_{n+1})$. Indeed we have :

$$e_{n+1} = s(\overline{w}^{(n-p+2)}(r_0)) = s(\overline{w}(\overline{w}^{(n-p+1)}(r_0)))$$
$$\leqslant w(\overline{s}(\overline{w}^{(n-p+1)}(r_0))) = w(e_{n-p+1}, \ldots, e_n) .$$

In what follows we will show that for superadditive functions $w : T^p \to T$, condition (19), which is generally very difficult to verify, can be replaced by much simpler ones.

THEOREM 3.3 : *Let $w : T^p \to T$ be a superadditive function. Then w is a p-dimensional rate of convergence on T if and only if*

$$\lim_{n\to\infty} w^{(n)}(t) = 0 \quad \textit{for all} \quad t \in T^p . \tag{22}$$

Proof : The necessity of condition (22) is obvious. In order to prove the sufficiency we note first that (22) implies

$$w(\xi, \xi, \ldots, \xi) < \xi , \qquad \xi \in T . \tag{23}$$

Indeed if we suppose that there is an $\eta \in T$ such that $w(\eta, \eta, \ldots, \eta) \geqslant \eta$ then because of the isotony of w (see Proposition 2.10) we have $w^{(n)}(\eta, \eta, \ldots, \eta) \geqslant \eta$ which contradicts (22).

Now, let us fix a point $t \in T^p$ and a positive number $\varepsilon > 0$. Let us denote $\alpha = \varepsilon - w(\varepsilon, \varepsilon, \ldots, \varepsilon)$. From (22) it follows that there is an integer $N = N(t, \varepsilon)$ such that

$$w^{(n)}(t) \leqslant \alpha/p \quad \text{for any} \quad n \geqslant N . \tag{24}$$

We will prove by induction with respect to k that

$$w^{(n)}(t) + w^{(n+1)}(t) + \cdots + w^{(n+k)}(t) \leqslant \varepsilon \quad \text{for any} \quad n \geqslant N . \tag{25}$$

For $k = 0, 1, 2, \ldots, p - 1$ this follows immediately from (24). Suppose (25) holds for a given $k \geqslant p - 1$. Using the superadditivity of w and (20′) we have :

$$\begin{aligned} &w^{(n)}(t) + w^{(n+1)}(t) + \cdots + w^{(n+k+1)}(t) \\ &\qquad \leqslant w^n(t) + \cdots + w^{(n+p-1)}(t) + w(\overline{w}^{n+p-1}(t) + \cdots + \overline{w}^{n+k}(t)) \\ &\qquad \leqslant \alpha + w(\varepsilon, \varepsilon, \ldots, \varepsilon) = \varepsilon . \end{aligned}$$

Thus (25) holds for all $n \geqslant N$ and all $k = 0, 1, 2, \ldots$ Hence the series (19) is convergent. The proof is complete. □

In the proof of the above theorem we have seen that if $w : T^p \to T$ is an isotone function then (22) implies (23). In what follows we show that for right continuous functions the converse is also true.

PROPOSITION 3.4 : *Let $w : T^p \to T$ be isotone and continuous from the right. Then condition* (22) *is equivalent to*

$$w(\xi, \xi, \ldots, \xi) < \xi , \qquad \xi \in T . \tag{23}$$

Proof : We only have to prove that (23) implies (22). Consider a vector $t = (t_1, t_2, \ldots, t_p) \in T^p$ and set $\eta = \max \{ t_1, t_2, \ldots, t_p \}$. The sequence

$a_n = w^{(n)}(\eta, \eta, \ldots, \eta)$ is a nonincreasing sequence of positive numbers so that it is convergent. Set $a = \lim_{n \to \infty} a_n$ and suppose $a > 0$. Then from the right continuity of w we have

$$a = \lim_{n \to \infty} a_{n+1} = \lim_{n \to \infty} w(a_{n-p+1}, a_{n-p+2}, \ldots, a_n) = w(a, a, \ldots, a)$$

which contradicts (23). Hence $\lim_{n \to \infty} a_n = 0$. From the isotony of w it follows that $w^{(n)}(t) \leqslant a_n$ so that (22) is satisfied. □

Now from Theorem 3.3 and Proposition 3.4 we can immediately deduce the main result of our paper :

COROLLARY 3.5 : *If* $w : T^p \to T$ *is superadditive and continuous from the right then conditions* (19), (22) *and* (23) *are equivalent.* □

In particular, for S-convex functions we have :

COROLLARY 3.6 : *Let* $w : T^p \to T$ *be an* S*-convex function. Then* w *is a* p*-dimensional rate of convergence on* T *if and only if one of the conditions* (22) *or* (23) *is satisfied.*

Proof : By virtue of 2.2, 2.5, 2.6 and 3.5 we only have to prove that if $w : T^p \to T$ is S-convex and if there is a sequence $a_n \in T^p$ such that

$$a_{n+1} \leqslant a_n\,, \qquad n = 0, 1, 2, \ldots; \qquad \lim a_n = 0\,, \qquad \lim w(a_n) = 0$$

then by taking $w(0) = 0$, w extends to an S-convex function on $D_0 = T^p \cup \{0\}$. Let us fix a point $x \in T^p$ and consider the sequence $\{ \delta w(x, a_n) \}_{n \geqslant 0}$. According to 2.1, c) we have

$$\delta w(x, a_{n+1}) \leqslant \delta w(x, a_n)\,, \qquad n = 0, 1, 2, \ldots \tag{26}$$

$$\delta w(x, a_n)\,(x - a_n) = w(x) - w(a_n)\,. \tag{27}$$

From (26) it follows that all the sequences $\{ \delta w(x, a_n)\, e_i \}_{n \geqslant 0}$ $(i = 1, 2, \ldots, p)$ are nonincreasing. Hence $d_i = \lim_{n \to \infty} \delta w(x, a_n)\, e_i$ exists although it might be equal to $-\infty$. From (27) it follows that $\sum_{i=1}^{p} d_i(xe_i) = w(x) > 0$ and taking into account the fact that $xe_i > 0$ for $i = 1, 2, \ldots, p$ it follows that $d_i > -\infty$ for $i = 1, 2, \ldots, p$. Hence $\lim_{n \to \infty} \delta w(x, a_n)$ exists and we take by definition

$$\delta w(x, 0) = \lim_{n \to \infty} \delta w(x, a_n)\,.$$

From (27) it follows that

$$\delta w(x, 0)(x-0)=\delta w(x, 0) x=w(x)=w(x)-w(0) .$$

We have to prove that

$$\delta w(x, 0) \leqslant \delta w(u, v) \quad \text{for all} \quad u, v \in D_0 \quad \text{with} \quad x \leqslant u .$$

If $v=0$ we have

$$\delta w(x, 0)=\lim_{n \rightarrow \infty} \delta w(x, a_n) \leqslant \lim_{n \rightarrow \infty} \delta w(u, a_n)=\delta w(u, 0) .$$

If $v \neq 0$ then there is an integer N such that $a_n \leqslant v$ for $n \geqslant N$. Hence

$$\delta w(x, 0)=\lim_{n \rightarrow \infty} \delta w(x, a_n) \leqslant \delta w(x, v) \leqslant \delta w(u, v) .$$

In a similar way by defining

$$\delta w(0, x)=\lim_{n \rightarrow \infty} \delta w(a_n, x)$$

we can prove that $\delta w(0, x)(0-x)=w(0)-w(x)$ and $\delta w(0, x) \leqslant \delta(u, v)$ for all $u, v \in D_0$ with $x \leqslant v$. The proof is complete. □

From the proof of the above corollary it follows that any S-convex rate of convergence $w : T^p \rightarrow T$ is superadditive. The reciprocal is clearly not true. It is interesting to note that the rate of convergence of Newton's method (3) is convex only on the interval $[0, a\sqrt{2}]$ but it is superadditive on $\mathbb{R}_+$. The proof of this fact is elementary but not very straightforward and it will be given in the following.

4. APPENDIX

The rate of convergence of Newton's method (3) has been intensely studied (see [12, 15, 20 and 21]). For this function the series (1) is convergent for any $t \in \mathbb{R}_+$ and its sum has the simple explicit expression :

$$s(t)=t-a+\sqrt{t^2+a^2} .$$

Explicit expressions for the finite sums

$$s_n(t)=t+w(t)+\cdots+w^n(t)$$

have also been found (see [20] or [15]). These explicit expressions were used to obtain sharp and elegant a priori and a posteriori estimates for Newton's method for solving nonlinear operator equations in Banach spaces (see [23, 12 and 15]). The second derivative of w is of the form

$$w''(t) = \frac{(2\,a^2 - t^2)\,a^2}{2(t^2 + a^2)^{5/2}}$$

which shows that w is convex on $[0, a\sqrt{2}]$ and concave on $[a\sqrt{2}, +\infty[$. In what follows we will show that w is superadditive on $[0, +\infty[$. We have to prove that

$$\frac{(x + y)^2}{[a^2 + (x + y)^2]^{1/2}} \geqslant \frac{x^2}{(a^2 + x^2)^{1/2}} + \frac{y^2}{(a^2 + y^2)^{1/2}}, \qquad x, y \geqslant 0\,.$$

If $a = 0$ this is trivially satisfied. If $a > 0$ then by dividing both sides of the inequality by a and denoting $u = x/a$, $v = y/a$ this reduces to

$$\frac{(u + v)^2}{[1 + (u + v)^2]^{1/2}} - \frac{u^2}{(1 + u^2)^{1/2}} - \frac{v^2}{(1 + v^2)^{1/2}} \geqslant 0\,, \qquad u, v \geqslant 0\,.$$

Let us fix v and consider the left hand side of the above inequality as a function of u, say $f(u)$. We have obviously $f(0) = 0$. We will prove that $f(u) > 0$ for $u > 0$ by showing that :

1) There is an $M > 0$ such that $f(u) > 0$ for all $u \geqslant M$.

2) The derivative of f has a unique root $u^* > 0$ such that $f'(u) > 0$ for $u < u^*$ and $f'(u) < 0$ for $u > u^*$.

Indeed if 1) and 2) hold then for any $u \leqslant u^*$ we have $f(u) > f(0) = 0$ while if $u > u^*$ then by taking $z > \max\{u, M\}$ we get $f(u) > f(z) > 0$.

In order to prove statement 1) we first perform the change of variable $\varepsilon = 1/u$ and obtain a function of ε,

$$\begin{aligned} g(\varepsilon) &= u^{-1}[1+u^2]^{1/2}\,[1+v^2]^{1/2}\,[1+(u+v)^2]^{1/2}\,f(u) \\ &= \{(1+\varepsilon v)^2\,(1+\varepsilon^2)^{1/2} - [\varepsilon^2+(1+\varepsilon v)^2]^{1/2}\}\,(1+v^2)^{1/2} - \\ &\qquad - \varepsilon v^2(1+\varepsilon^2)^{1/2}\,[\varepsilon^2+(1+\varepsilon v)^2]^{1/2}\,. \end{aligned}$$

We have

$$\begin{aligned} &(1 + \varepsilon v)^2\,(1 + \varepsilon^2)^{1/2} - [\varepsilon^2 + (1 + \varepsilon v)^2]^{1/2} = \\ &\qquad = \varepsilon v(2 + \varepsilon v)\frac{(1 + \varepsilon v)^2 + \varepsilon^2[(1 + \varepsilon v)^2 + 1]}{(1 + \varepsilon v)^2\,(1 + \varepsilon^2)^{1/2} + [\varepsilon^2 + (1 + \varepsilon v)^2]^{1/2}} \end{aligned}$$

so that we can write

$$g(\varepsilon) = \varepsilon v[a(\varepsilon)\,(1 + v^2)^{1/2} - b(\varepsilon)\,v]$$

where $\lim_{\varepsilon \to} a(\varepsilon) = \lim_{\varepsilon \to} b(\varepsilon) = 1$. This shows that $g(\varepsilon) > 0$ for ε sufficiently small, which completes the proof of statement 1).

If we denote

$$h(t) = \frac{t(2 + t^2)}{(1 + t^2)^{3/2}}$$

then the derivative of f can be written as

$$f'(u) = h(u + v) - h(u) .$$

It is easy to prove that for any given v there is a unique $u > 0$ (more precisely we have $[(\sqrt{5} - 1)/2]^{1/2} < u < \sqrt{2}$) such that $h(u + v) = h(u)$. (Intuitively this can be easily realized by drawing the graph of h.) Hence the equation $f'(u) = 0$ has a unique positive solution u^*. We have obviously $f'(0) = h(v) > 0$ and from the mean value theorem it follows that there is a $w > \sqrt{2}$ such that

$$f'(\sqrt{2}) = vh'(w) = v\frac{2 - w^2}{2(w^2 + 1)^{5/2}} < 0 ,$$

which completes the proof of statement 2).

REFERENCES

[1] I. KOREC, *On a problem of V. Pták*, Čas. pro pėst. mat. 103 (1978), 365-379.

[2] J. KŘÍŽKOVÁ, P. VRBOVÁ, *A remark on a factorization theorem*, Comm. Math. Univ. Carol (CMUC) 15 (1974), 611-614.

[3] J. M. ORTEGA, W. C. RHEINBOLDT, *Iterative solution of nonlinear equations in several variables*, Academic Press, London, 1970.

[4] H. PETZELTOVÁ, P. VRBOVÁ, *An overrelaxed modification of Newton's method*, Revue Roumaine des Mathématiques 22 (1977), 959-963.

[5] H. PETZELTOVÁ, P. VRBOVÁ, *A remark on small divisors problems*, Czech. Math. J. 103 (1978), 1-12.

[6] F. A. POTRA, *On a modified secant method*, Math. Rev. Anal. Numer. Theor. Approximation, Anal. Numer. Theor. Approximation, 8, 2 (1979), 203-214.

[7] F. A. POTRA, *An application of the induction method of V. Pták to the study of Regula Falsi*, Aplikace Matematiky 26 (1981), 111-120.

[8] F. A. POTRA, *The rate of convergence of a modified Newton's process*, Aplikace matematiky 26 (1981), 13-17.

[9] F. A. POTRA, *An error analysis for the secant method*, Numer. Math. 38 (1982), 427-445.

[10] F. A. POTRA, V. PTÁK, *Nondiscrete induction and a double step secant method*, Math. Scand. 46 (1980), 236-250.

[11] F. A. Potra, V. Pták, *On a class of modified Newton processes,* Numer. Funct. Anal. and Optimiz. 2 (1980), 107-120.

[12] F. A. Potra, V. Pták, *Sharp error bounds for Newton's process,* Numer. Math. 34 (1980), 63-72.

[13] F. A. Potra, V. Pták : A generalization of Regula Falsi, Numer. Math. 36 (1981), 333-346.

[14] F. A. Potra, V. Pták : *Nondiscrete induction and an inversion free modification of Newton's method,* Čas. pro pěst. mat. 108, 4 (1983), 333-341.

[15] F. A. Potra, V. Pták : *Nondiscrete induction and iterative processes,* Pitman Advanced Publishing Program, London, 1984.

[16] V. Pták : *Some metric aspects of the open mapping theorem,* Math. Ann. 165 (1966), 95-104.

[17] V. Pták : *A quantitative refinement of the closed graph theorem,* Czech. Math. J. 99 (1974), 503-506.

[18] V. Pták : *A theorem of the closed graph type,* Manuscripta Math. 13 (1974), 109-130.

[19] V. Pták : *Deux théorèmes de factorisation,* Comptes Rendus, Acad. Sci. Paris 278 (1974), 1091-1094.

[20] V. Pták : *Concerning the rate of convergence of Newton's process,* Comm. Math. Univ. Carolinae 16 (1975), 599-705.

[21] V. Pták, *A modification of Newton's method,* Čas. pěst mat. 101 (1976), 188-194.

[22] V. Pták, *Nondiscrete mathematical induction and iterative existence proofs,* Linear Algebra and its Applications 13 (1976), 223-236.

[23] V. Pták, *The rate of convergence of Newton's process,* Numer. Math. 25 (1976), 279-285.

[24] V. Pták, *Nondiscrete mathematical induction,* in : General Topology and its Relations to Modern Analysis and Algebra IV, 166-178, Lecture Notes in Mathematics 609, Springer Verlag, 1977.

[25] V. Pták, *What should be a rate of convergence,* R.A.I.R.O., Analyse Numérique 11 (1977), 279-286.

[26] V. Pták, *Stability of exactness,* Comm. Math. (Poznań) 21 (1978), 343-348.

[27] V. Pták, *A rate of convergence,* Numer. Funct. Anal. and Optimiz. 1 (1979), 255-271.

[28] V. Pták, *Factorization in Banach algebras,* Studia Math. 65 (1979), 279-285.

[29] J. W. Schmidt, H. Leonhardt, *Eeingrenzung von Lösungen mit Hilfe der Regula falsi,* Computing 6 (1970), 318-329.

[30] J. Zemánek, *A remark on transitivity of operator algebras,* Čas. pěst. mat. 100 (1975), 176-178.