SIAM J. NUMERICAL ANALYSIS
Vol. 16, No. 4, August 1979

0036-1429/79/1604-0001 $01.00/0

THEORY OF MULTIVARIATE SECANT METHODS*

JANINA JANKOWSKA†

Abstract. We study multivariate secant methods for the solution of systems of nonlinear equations in complex n-dimensional space ($1 \leqq n < \infty$). We prove a quantitative theorem about the convergence of such methods and consider the dependence of the order of convergence on the position of the iteration points by using the concept of *a set of admissible approximations*. We also find the maximal order of the so-called "standard information" $\mathcal{N} = \{\vec{f}(\vec{x}_j): j = n, n-1, \cdots, 0\}$ with respect to the sets of admissible approximations.

1. Introduction. This paper deals with multivariate secant methods, here called *MS methods*, for the solution of a system of nonlinear equations. We are mostly interested in the dependence of the order of convergence of MS methods on the configuration of the constructed points.

In § 3 we deal with one step of the MS method and prove a basic estimate for the error in the MS method. We then prove a general convergence theorem and briefly consider special cases of the MS method, such as regula falsi, the discrete Newton method and the MS method based on generalized divided differences.

In § 4 we consider the one-point MS method with memory in detail. This is the most interesting case from the practical point of view. In Theorem 4.1 we show how the order of convergence depends on the configuration of the constructed points (compare with Bittner [2] and Barnes [1]).

Section 5 deals with the optimal properties of iterations which use the so-called standard information

$$\mathcal{N} = \{\vec{f}(\vec{x}_j): j = 1, \cdots, n\}.$$

We show that the best possible configuration is closely connected with the admissible approximation set defined by

$$\mathcal{M} = \left\{ (\vec{x}_0, \cdots, \vec{x}_n): \left| \det \left[\frac{\vec{x}_1 - \vec{x}_0}{\|\vec{x}_1 - \vec{x}_0\|}, \cdots, \frac{\vec{x}_n - \vec{x}_{n-1}}{\|\vec{x}_n - \vec{x}_{n-1}\|} \right] \right| \geqq c > 0 \right\}.$$

This means that the $(n+1)$-tuple of iterates $(\vec{x}_k, \vec{x}_{k+1}, \cdots, \vec{x}_{k+n})$ should belong to $\mathcal{M}$ for any k. Moreover the maximal order of any iteration based on $\mathcal{N}$ and $\mathcal{M}$ is achievable by the MS method and is equal to a unique positive zero of the polynomial $t^{n+1} - t^n - 1$.

2. Preliminaries and notation. In this section we introduce the notation and definitions to be used throughout this paper.

Let $\vec{x} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{C}^n$ denote an n-dimensional complex vector and let $A = [\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n]$ be an $n \times n$ complex matrix with columns $\vec{a}_j = [a_{1j}, a_{2j}, \cdots, a_{nj}]^T$, $j = 1, 2, \cdots, n$. By $\operatorname{diag}_{1 \leqq j \leqq n}(b_j)$ we mean an $n \times n$ diagonal matrix with entries b_j. We shall use the Euclidean vector norm $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ and the Frobenius matrix norm $\|A\| = (\sum_{i,j=1}^n |a_{ij}|^2)^{1/2}$.

Consider the problem of solving a system of nonlinear equations,

$$\vec{f}(\vec{x}) = \vec{0}, \tag{2.1}$$

where $\vec{f}: D_f \subset \mathbb{C}^n \to \mathbb{C}^n$ and D_f is an open and convex subset of $\mathbb{C}^n$. We assume throughout this paper that $\vec{f}$ satisfies the following conditions.

* Received by the editors March 17, 1975, and in revised form June 6, 1977.
† Department of Mathematics, University of Warsaw, Warsaw, Poland.

(2.2a) There exists a zero $\vec{\alpha} = \vec{\alpha}(\vec{f}) \in D_f$ of $\vec{f}$ such that $f'(\vec{\alpha})$ is nonsingular. We set $M = \|[f'(\vec{\alpha})]^{-1}\|_2$.

(2.2b) The Fréchet derivative f' of $\vec{f}$ satisfies the Lipschitz condition $\|f'(\vec{x}) - f'(\vec{y})\|_2 \leqq L \cdot \|\vec{x} - \vec{y}\|$ for $\vec{x}, \vec{y} \in D_f$.

We solve (2.1) by the MS method defined as follows. Let $\vec{x}_k, \vec{x}_{k-1}, \cdots, \vec{x}_{k-n} \in D_f$ be approximations to $\vec{\alpha}$ such that the matrices ΔX_{k-n} and ΔF_{k-n} are nonsingular, where

$$\Delta X_{k-n} = [\overrightarrow{\Delta x}_{k-n}, \cdots, \overrightarrow{\Delta x}_{k-1}], \tag{2.3}$$
$$\Delta F_{k-n} = [\overrightarrow{\Delta f}_{k-n}, \cdots, \overrightarrow{\Delta f}_{k-1}],$$

and $\overrightarrow{\Delta x}_j = \vec{x}_{j+1} - \vec{x}_j$, $\overrightarrow{\Delta f}_j = \vec{f}_{j+1} - \vec{f}_j$ with $\vec{f}_j = \vec{f}(\vec{x}_j)$ for $j = k-n, \cdots, k$.

Let $\vec{W}$ be the linear function that interpolates $\vec{f}$ at $\vec{x}_k, \vec{x}_{k-1}, \cdots, \vec{x}_{k-n}$, i.e. $\vec{W}(\vec{x}_{k-j}) = \vec{f}(\vec{x}_{k-j})$ for $j = 0, 1, \cdots, n$. The next approximation $\varphi(\vec{x}_k, \cdots, \vec{x}_{k-n}; \vec{f})$ to $\vec{\alpha}$ in the MS method is a unique zero of $\vec{W}$. It is straightforward to verify that $\vec{W}(\vec{x}) = \vec{f}_k + \Delta F_k \times (\Delta X_k^{-1}) \times (\vec{x} - \vec{x}_k)$ and

$$\varphi(\vec{x}_k, \cdots, \vec{x}_{k-n}; \vec{f}) = \vec{x}_k - \Delta X_k \times (\Delta F_k^{-1}) \times \vec{f}_k. \tag{2.4}$$

For convenience, define

$$d_k = d(\vec{x}_k, \cdots, \vec{x}_{k-n}) = \frac{|\det[\Delta X_{k-n}]|}{\prod_{j-1}^{n} \|\overrightarrow{\Delta x}_{k-j}\|} \tag{2.5}$$

and

$$K(\vec{f}) = M \cdot L \cdot n \cdot \sqrt{n}. \tag{2.6}$$

We are now ready to characterize the convergence of the MS method.

3. Convergence of the MS method. It is well known, that one can get different orders of convergence and different costs of one iterative step in the MS method by taking different points $\vec{x}_{k-1}, \vec{x}_{k-2}, \cdots, \vec{x}_{k-n}$ in (2.3) and (2.4) (e.g. see Woźniakowski [7], Ortega and Rheinboldt [3, IV, § 11.3], Voigt [6]).

However, it turns out that in any case the order of convergence of the sequence $\{\varphi(\vec{x}_k, \cdots, \vec{x}_{k-n}; \vec{f})\}_{k=n}^{\infty}$ mainly depends on the values of d_k given by (2.5), see also Barnes [1], Bittner [2]. To show this we need two lemmas.

LEMMA 3.1. *If $A = [\vec{a}_1, \cdots, \vec{a}_n]$ is nonsingular and $\vec{b}_j^T$ is the j-th row of the matrix A^{-1} then*

$$\|\vec{b}_j\| \leqq \sqrt{n}/(d \cdot a) \quad (j = 1, \cdots, n), \tag{3.1}$$

$$\|A^{-1}\| \leqq n/(d \cdot a), \tag{3.2}$$

where

$$a = \min_{1 \leqq j \leqq n} \|\vec{a}_j\| \quad \textit{and} \quad d = \frac{|\det A|}{\prod_{j=1}^{n} \|\vec{a}_j\|}.$$

Proof. By using Cramer's rule and the Hadamard inequality we get the following inequalities for the elements b_{ij} of the matrix A^{-1}

$$|b_{ij}| \leqq \frac{\prod_{k \neq i} \|\vec{a}_k\|}{d \cdot \prod_{k=1}^{n} \|\vec{a}_k\|} = \frac{1}{d \cdot \|\vec{a}_i\|} \leqq \frac{1}{d \cdot a}.$$

Hence

$$\|\vec{b}_i\| = \sqrt{\sum_{j=1}^{n} |b_{ij}|^2} \leqq \frac{\sqrt{n}}{d \cdot a}$$

and

$$\|A^{-1}\| = \sqrt{\sum_{i=1}^{n} \|\vec{b}_i\|^2} \leqq \frac{n}{d \cdot a},$$

which completes the proof of Lemma 3.1. □

We now deal with one step of the MS method. In the following lemma we give an error estimate for the next approximation $\varphi(\vec{x}_n, \cdots, \vec{x}_0; \vec{f})$ with respect to the starting errors $\|\vec{x}_j = \vec{\alpha}\|$ $j = 0, \cdots, n$.

LEMMA 3.2. *Let*

(3.3) $$\vec{f} \quad \textit{have property} \quad (2.2),$$

(3.4) $$\vec{x}_j \in D_f \quad (j = 0, 1, \cdots, n), \quad \textit{and}$$

(3.5) $$K(\vec{f}) \cdot \max_{1 \leqq j \leqq n} \|\vec{x}_j - \vec{\alpha}\|/d_n < 1.$$

Then

$$\|\varphi(\vec{x}_n, \cdots, \vec{x}_0; \vec{f}) - \vec{\alpha}\| \leqq \frac{3 \cdot K(\vec{f}) \cdot (\max_{1 \leqq j \leqq n} \|\vec{x}_j - \vec{\alpha}\|) \cdot (\min_{1 \leqq j \leqq n} \|\vec{x}_j - \vec{\alpha}\|)}{2 \cdot d_n \cdot \left(1 - \frac{K(\vec{f}) \cdot \max_{1 \leqq j \leqq n} \|\vec{x}_j - \vec{\alpha}\|}{d_n}\right)}$$

where d_n and $K(\vec{f})$ are defined by (2.5) *and* (2.6).

Proof. Without loss of generality we can assume that $\|\vec{x}_n - \vec{\alpha}\| \leqq \|\vec{x}_j - \vec{\alpha}\| \leqq \|\vec{x}_0 - \vec{\alpha}\|$ $(j = 0, \cdots, n)$. *Let* $\vec{z} = \varphi(\vec{x}_n, \cdots, \vec{x}_0; \vec{f})$. From (2.4) we have

(3.6) $$\vec{z} - \vec{\alpha} = \vec{x}_n - \vec{\alpha} - \Delta X_n \times (\Delta F_n^{-1}) \times \vec{f}_n.$$

Define

(3.7) $$\vec{g}(\vec{x}) = \vec{f}(\vec{x}) - f'(\vec{\alpha}) \times (\vec{x} - \vec{\alpha}).$$

Since $\vec{f}$ has property (2.2), we have

(3.8a) $$\|\vec{g}(\vec{x})\| \leqq \tfrac{1}{2} \cdot L \cdot \|\vec{x} - \vec{\alpha}\|^2,$$

(3.8b) $$\|\vec{g}(\vec{x}) - \vec{g}(\vec{y})\| \leqq L \cdot \|\vec{x} - \vec{y}\| \cdot \max(\|\vec{x} - \vec{\alpha}\|, \|\vec{y} - \vec{\alpha}\|),$$

for $\vec{x}, \vec{y} \in D_f$ (see Ortega and Rheinboldt [3, p. 70–73]). From Lemma 3.1 and (3.8) it easily follows that

(3.9) $$\Delta F_n = f'(\vec{\alpha}) \times (I + B) \times \Delta X_n,$$

where

(3.10) $$\|B\| \leqq \frac{M \cdot L \cdot n}{d_n} \cdot \sqrt{\sum_{j=0}^{n-1} \|\vec{x}_j - \vec{\alpha}\|^2} \leqq \frac{K(\vec{f})}{d_n} \cdot \|\vec{x}_0 - \vec{\alpha}\|.$$

This implies that the matrices $(I+B)$ as well as ΔF_n are nonsingular. From (3.9) we get

$$\Delta F_n^{-1} = (\Delta X_n^{-1}) \times (I - C) \times [f'(\vec{\alpha})]^{-1}, \tag{3.11}$$

where $\|C\| \leqq \|B\|/(1-\|B\|)$.

Finally (3.6), (3.10) and (3.11) yield

$$\|\vec{z} - \vec{\alpha}\| \leqq \frac{3 \cdot K(\vec{f}) \cdot \|\vec{x}_0 - \vec{\alpha}\| \cdot \|\vec{x}_n - \vec{\alpha}\|}{2 \cdot d_n \cdot (1 - K(\vec{f}) \cdot (\|\vec{x}_0 - \vec{\alpha}\|/d_n))},$$

which completes the proof of Lemma 3.2. □

Note that in Lemma 3.2 we assume that the starting vectors $\vec{x}_0, \cdots, \vec{x}_n$ are close enough to $\vec{\alpha}$ and additionally that the vectors $\vec{x}_1 - \vec{x}_0, \cdots, \vec{x}_n - \vec{x}_{n-1}$ are sufficiently linearly independent, i.e. d_n is not too small (see also Barnes [1], Bittner [2]).

From Lemma 3.2 it is possible to conclude the convergence of the multivariate regula falsi method (see Ortega and Rheinboldt [3, IV, § 11.3]). Namely we prove

COROLLARY 3.1. *Let*

(3.12) $\vec{f}$ *have property* (2.2),

(3.13) $\vec{y}_j \in D_f$ *for* $j = 0, \cdots, n-1$ *be points such that*

$$s_0 = K(\vec{f}) \cdot \max_{0 \leqq j \leqq n-1} \|\vec{y}_j - \vec{\alpha}\| / d(\vec{\alpha}, \vec{y}_{n-1}, \cdots, \vec{y}_0) < 1.$$

Then, for $s \in (s_0, 1)$,

$$\|\varphi(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0; \vec{f}) - \vec{\alpha}\| \leqq \frac{3 \cdot s}{2 \cdot (1-s)} \cdot \|\vec{x} - \vec{\alpha}\| \quad \text{for} \quad \vec{x} \in T, \tag{3.14}$$

$$\lim_{\substack{\vec{x} \to \vec{\alpha} \\ \vec{x} \in T}} \varphi(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0) = \vec{\alpha}, \tag{3.15}$$

where

$$T = T(s) = \{\vec{x} \in D_f : K(\vec{f}) \cdot \max_{0 \leqq j \leqq n} \|\vec{y}_j - \vec{\alpha}\| / d(\vec{x}, \vec{y}_{n-1} \cdots, \vec{y}_0) \leqq s\}.$$

Proof. Since the function $d(\vec{x}) = d(\vec{x}; \vec{y}_{n-1}, \cdots, \vec{y}_0)$ is continuous with respect to $\vec{x}$, the result follows from (3.13) and Lemma 3.2. □

Assume additionally that

$$\begin{aligned} &s < 0.4, \qquad \vec{x}_0 \in T(s), \\ &\vec{x}_{k+1} = \varphi(\vec{x}_k, \vec{y}_{n-1}, \cdots, \vec{y}_0; \vec{f}) \quad \in T(s) \quad (k = 0, 1, \cdots); \end{aligned} \tag{3.16}$$

then $\|\vec{x}_{k+1} - \vec{\alpha}\| \leqq \rho \cdot \|\vec{x}_k - \vec{\alpha}\|$, where $\rho = ((3 \cdot s)/2 \cdot (1-s)) < 1$. This means that the sequence $\{\vec{x}_k\}_{k=1}^{\infty}$ generated by regula falsi converges to $\vec{\alpha}$. The set $T(s)$ has a simple geometric interpretation. For example, for $n = 2$ we can redefine $T(s)$ as follows

$$T(s) = \{\vec{x} \in D_f : |\sin \sphericalangle (\vec{y}_1 - \vec{y}_0, \vec{x} - \vec{y}_1) \geqq K(\vec{f}) \cdot \|\vec{y}_0 - \vec{\alpha}\|/s\}.$$

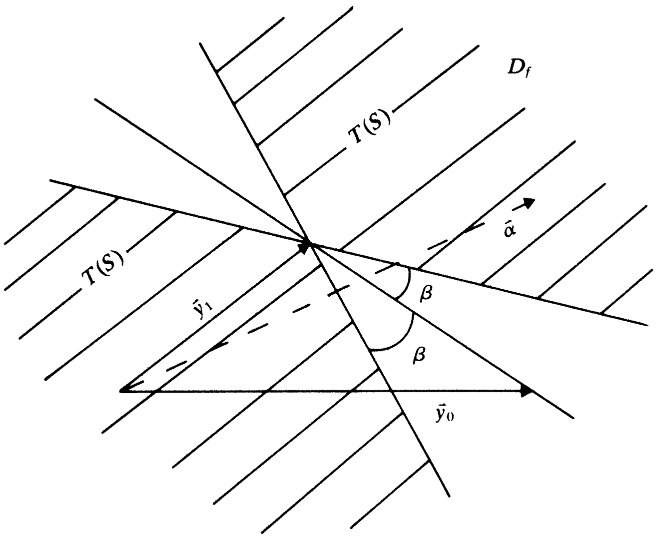

FIG. 1

Let β be an angle such that $\sin(\beta) = K(\vec{f}) \cdot \|\vec{y}_0 - \vec{\alpha}\|/s$, $\beta \in [0, \pi/2]$. Then $T(s)$ is given by Fig. 1. Note that regula falsi requires only one function evaluation per step. However the linear convergence makes this iteration inefficient.

We are now in a position to prove a convergence theorem for the general MS method defined by (2.4).

THEOREM 3.1. *Let*

$$\vec{f} \quad \textit{have property} \quad (2.2), \tag{3.17}$$

(3.18) *the vectors* $\vec{y}_j = \vec{y}_j(\vec{x}) \in D_f$ *for* $\vec{x} \in D_f$, $j = 0, 1, \cdots, n-1$ *satisfy*

$$\tau(\vec{x}) = \frac{K(\vec{f}) \cdot \max_{0 \leqq j \leqq n-1} \|\vec{y}_j - \vec{\alpha}\|}{d(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0)} < 1.$$

Then

$$\|\varphi(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0; \vec{f}) - \vec{\alpha}\| \leqq \frac{3 \cdot \tau(\vec{x})}{2 \cdot (1 - \tau(\vec{x}))} \cdot \|\vec{x} - \vec{\alpha}\|. \tag{3.19}$$

Furthermore if $\tau = \sup_{\vec{x} \in D_f} \tau(\vec{x}) < 0.4$ *then*

$$\|\varphi(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0; \vec{f}) - \vec{\alpha}\| \leqq \rho_1 \cdot \|\vec{x} - \vec{\alpha}\|, \quad \textit{where}$$

$$\rho_1 = \frac{3 \cdot \tau}{2 \cdot (1 - \tau)} < 1, \tag{3.20}$$

$$\lim_{\vec{x} \to \vec{\alpha}} \varphi(\vec{x}, \vec{y}_{n-1}, \cdots, \vec{y}_0; \vec{f}) = \vec{\alpha}. \tag{3.21}$$

Proof. Let $\vec{y}_n = \vec{x}$. From lemma 3.2 we get

$$\|\varphi(\vec{y}_n, \cdots, \vec{y}_0; \vec{f}) - \vec{\alpha}\| \leqq \frac{3 \cdot K(\vec{f}) \cdot \max_{0 \leqq j \leqq n} \|\vec{y}_j - \vec{\alpha}\| \cdot \min_{0 \leqq j \leqq n} \|\vec{y}_j - \vec{\alpha}\|}{2 \cdot d_n \cdot (1 - K(\vec{f})) \cdot (\max_{0 \leqq j \leqq n} \|\vec{y}_j - \vec{\alpha}\|/d_n)}$$

$$\leqq \frac{3 \cdot \tau(\vec{x})}{2 \cdot (1 - \tau(\vec{x}))} \cdot \|\vec{x} - \vec{\alpha}\|,$$

where $d_n = d(\vec{y}_n, \cdots, \vec{y}_0)$, which proves the first part of Theorem 3.1. Moreover, if $\tau = \sup_{\vec{x} \in D_f} \tau(\vec{x}) < 0.4$ then $\rho_1 < 1$, and (3.20) and (3.21) easily follow from (3.19). □

Let us define

$$\vec{x}_{k+1} = \varphi(\vec{x}_k, \vec{y}^{\,k}_{n-1}, \cdots, \vec{y}^{\,k}_0) \quad (k = 0, 1, \cdots), \tag{3.22}$$

where $\vec{x}_0 \in D_f$ and the vectors $\vec{y}^{\,k}_{n-1}, \cdots, y^k_0$ satisfy the assumptions of Theorem 3.1. Then

$$\|\vec{x}_{k+1} - \vec{\alpha}\| \leqq \frac{3 \cdot \tau(\vec{x}_k)}{2 \cdot (1 - \tau(\vec{x}_k))} \cdot \|\vec{x}_k - \vec{\alpha}\|. \tag{3.23}$$

Thus, the convergence of $\{\vec{x}_k\}_{k=0}^{\infty}$ mainly depends on the value $\tau(\vec{x}_k)$. Hence, we want to minimize $\tau(\vec{x}_k)$ by a suitable choice of $\vec{y}^{\,k}_0, \cdots, \vec{y}^{\,k}_{n-1}$. Since in general $\vec{x}_k$ is the best known approximation to $\vec{\alpha}$ at the kth step, it seems reasonable to define the vectors $\vec{y}^{\,k}_j$ $(j = 0, \cdots, n-1)$ so that

$$\|\vec{y}^{\,k}_j - \vec{\alpha}\| = A_j(\vec{f}) \cdot \|\vec{x}_k - \vec{\alpha}\| \tag{3.24}$$

for positive constants $A_j(\vec{f})$ and $d(\vec{x}_k, \vec{y}^{\,k}_{n-1}, \cdots, \vec{y}^{\,k}_0) = 1$. This leads us to the well-known cases of the MS method.

(A) Set $\vec{y}^{\,k}_j = \vec{x}_k + \|\vec{f}_k\| \cdot \sum_{i=j+1}^{n} \vec{e}_i, \qquad j = 0, \cdots, n-1, \quad k = 0, 1, \cdots$

where $\vec{e}_i$ is the ith unit vector. From (3.7) and (3.8) we get

$$\max_{0 \leqq j \leqq n-1} \|\vec{y}^{\,k}_j - \vec{\alpha}\| \leqq \|\vec{x}_k - \vec{\alpha}\| + \|\vec{f}_k\| \cdot \sqrt{n}.$$

Thus, from (3.22), (3.23) and the inequality above we have

$$\|\vec{x}_{k+1} - \vec{\alpha}\| \leqq K_1(\vec{f}) \cdot \|\vec{x}_k - \vec{\alpha}\|^2. \tag{3.25}$$

This proves the quadratic convergence. This version of the MS method is called the *discrete Newton method* or in the scalar case the *Steffensen's method* (see Ortega and Rheinboldt [3, p. 198], Voigt [6] and many others).

(B) Set $\vec{y}^{\,k}_{n-1} = \vec{x}_{k-1}$,

$$\vec{y}^{\,k}_j = \vec{y}^{\,k}_{j+1} - \|\vec{f}_k\| \cdot \vec{z}^{\,k}_j \quad \text{for} \quad j = n-2, \cdots, 0, \quad k = 0, 1, \cdots,$$

where $\{\vec{z}^{\,n}_j\}_{j=0}^{n-1}$ is an orthonormal system with $\vec{z}^{\,k}_{n-1} = \overrightarrow{\Delta x}_{k-1}/\|\overrightarrow{\Delta x}_{k-1}\|$. From (3.7) and (3.8),

$$\max_{0 \leqq j \leqq n-1} \|\vec{y}^{\,k}_j - \vec{\alpha}\| \leqq \|\vec{x}_{k-1} - \vec{\alpha}\| + \|\vec{f}_k\| \cdot \sqrt{n}.$$

Thus, from (3.22), (3.23) and the inequality above we get

$$\|\vec{x}_{k+1} - \vec{\alpha}\| \leqq K_2(\vec{f}) \cdot \|\vec{x}_k - \vec{\alpha}\| \cdot \|\vec{x}_{k-1} - \vec{\alpha}\|. \tag{3.26}$$

The matrices $\Delta X_k \times \Delta F_k^{-1}$ with $\vec{x}_{k-j} = \vec{y}^k_{j+2}$ for $j = 0, 1, \cdots, n-2$ are called *generalized divided differences* (see Ortega and Rheinboldt [3, pp. 203, 213, 363], Ulhm [5], Voigt [6]).

Iterative methods which have been considered in the points (A) and (B) generate quickly convergent sequences. However the cost of one iterative step consists of n function evaluations in the case (A) and $n-1$ function evaluations in the case (B). Hence, they are inefficient if the cost of function evaluation is large.

In order to decrease the number of function evaluations to be computed at every iterative step we pass to *one-point* secant methods *with memory* (see Traub [4, pp. 8–9]).

4. The one-point MS method with memory.

(C) Let us define

$$\vec{y}^k_{n-j} = \vec{x}_{k-j}, \qquad j = 1, \cdots, n, \quad k = 0, 1, \cdots,$$

whenever $\vec{x}_k, \cdots, \vec{x}_{k-n}$ satisfy the assumptions of Theorem 3.1. Otherwise the vectors $\vec{y}^k_{n-j}$ are defined in a different manner (for example the case (A) or (B) can be applied); see Barnes [1], Bittner [2].

Observe that the method (C) used only *one new piece of information* at $\vec{x}_k$ and *reuses* information at $\vec{x}_{k-1}, \cdots, \vec{x}_{k-n}$.

The convergence of this method mainly depends on the behavior of the sequence $\{d_k\}_{k=n}^{\infty}$ (cf. (2.5)). Since we use memory, we can't always guarantee that $d_k = 1$. In some papers [1], [2] it is assumed that

$$d_k \geqq c, \qquad k = n, n+1, \cdots, \tag{4.1}$$

for a given number $c \in (0, 1]$.

As we shall see it is also possible to prove convergence of the method (C) in cases where (4.1) does not hold.

LEMMA 4.1. *Let*

$$\vec{f} \quad \textit{have property} \quad (2.2), \tag{4.2}$$

$$\vec{x}_0, \cdots, \vec{x}_n \in D_f \colon \|\vec{x}_n - \vec{\alpha}\| \leqq \cdots \leqq \|\vec{x}_0 - \vec{\alpha}\|, \tag{4.3}$$

$$d_i \neq 0 \quad \textit{for} \quad i = n, n+1, \cdots. \tag{4.4}$$

If $\sup_{i \geqq n} \tau_i \leqq 0.4$ *and* $\sum_{i=n}^{\infty}(1 - 2.5 \cdot \tau_i) = +\infty$, *where* $\tau_i = K(\vec{f}) \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|/d_i$, *then*

$$\|\vec{x}_{i+1} - \vec{\alpha}\| \leqq \|\vec{x}_i - \vec{\alpha}\|, \tag{4.5}$$

$$\|\vec{x}_{i+1} - \vec{\alpha}\| \leqq \left(\prod_{j=n}^{i} (2.5 \cdot \tau_i)\right) \cdot \|\vec{x}_n - \vec{\alpha}\| \tag{4.6}$$

and

$$\lim_{i \to \infty} \vec{x}_i = \vec{\alpha}. \tag{4.7}$$

Proof. We first prove by induction that $\|\vec{x}_i - \vec{\alpha}\| \leqq \|\vec{x}_{i-1} - \vec{\alpha}\| \cdots \leqq \|\vec{x}_0 - \vec{\alpha}\|$ for $i \geqq n$. By (4.3) this holds for $i = n$. The vectors $\vec{x}_i, \vec{x}_{i-1}, \cdots, \vec{x}_{i-n}$ satisfy the assumptions of Lemma 3.2 since

$$\frac{K(\vec{f}) \cdot \max_{0 \leqq j \leqq n} \|\vec{x}_{i-j} - \vec{\alpha}\|}{d_i} = \frac{K(\vec{f}) \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|}{d_i} = \tau_i \leqq 0.4. \tag{4.8}$$

Hence, by Lemma 3.2 we have

$$\|\vec{x}_{i+1} - \vec{\alpha}\| \leqq \frac{3 \cdot K(\vec{f}) \cdot \|\vec{x}_{i-n} - \vec{\alpha}\| \cdot \|\vec{x}_i - \vec{\alpha}\|}{2 \cdot d_i \cdot (1 - K(\vec{f})) \cdot \|\vec{x}_i - \vec{\alpha}\|/d_i}. \tag{4.9}$$

Next from our assumptions it follows, that

$$\frac{3\cdot K(\vec f)\cdot\|\vec x_{i-n}-\vec\alpha\|}{2\cdot d_i\cdot(1-K(\vec f)\cdot\|\vec x_{i-n}-\vec\alpha\|/d_i}=\frac{3\cdot\tau_i}{2\cdot(1-\tau_i)} \tag{4.10}$$
$$\leqq\frac{3}{5\cdot(1-\frac{2}{5})}=1.$$

The inequalities (4.9) and (4.10) yield,

$$\|\vec x_{i+1}-\vec\alpha\|\leqq 2.5\cdot\tau_i\cdot\|\vec x_i-\vec\alpha\| \tag{4.11}$$
$$\leqq\cdots\leqq\left(\prod_{j=n}^{i}(2.5\cdot\tau_j)\right)\cdot\|\vec x_n-\vec\alpha\|$$

and

$$\|\vec x_{i+1}-\vec\alpha\|\leqq\|\vec x_i-\vec\alpha\|. \tag{4.12}$$

The condition $\sum_{j=n}^{\infty}(1-2.5\cdot\tau_j)=+\infty$ implies $\prod_{j=n}^{\infty}(2.5\cdot\tau_j)=0$, which means that $\lim_{i\to\infty}\vec x_i=\vec\alpha$. This proves Lemma 4.1. □

Note that if $\sup_{i\geqq n}\tau_i<0.4$ then $\sum_{j=n}^{\infty}(1-2.5\cdot\tau_j)_j=+\infty$ and $\{\vec x_k\}_{k=n}^{\infty}$ converges to $\vec\alpha$ at least linearly. However if one wants to get superlinear convergence it is necessary to assure that the sequence $\{d_k\}_{k=n}^{\infty}$ does not tend to zero too quickly.

Define now the family of sets $\mathcal{M}(c,\xi)$ as follows

$$\mathcal{M}(c,\xi)=\{(\vec x_0,\cdots,\vec x_n):\vec x_j\in\mathbb{C}^n,\ \vec x_j\neq\vec x_k,\ j,k=0,\cdots,n,\ j\neq k, \tag{4.13}$$
$$d_n\geqq c\cdot\|\vec x_n-\vec x_0\|^{\xi}\},$$

where $c\in(0,\infty)$ and $\xi\in[0,1)$ (compare with the sets of admissible approximations in Woźniakowski [7]). Note that if $(\vec x_k,\cdots,\vec x_{k-n})\in\mathcal{M}(c,\xi)$ for all k and $\lim_{k\to\infty}\vec x_k=\vec\alpha$ then $1/d_k=O(1/\|\vec x_{k-n}-\vec\alpha\|^{\xi})$.

THEOREM 4.1. *Let*

$$\vec f\quad\textit{have property}\quad(2.2), \tag{4.14}$$

$$\vec x_0,\cdots,\vec x_n\in D_f,\qquad\|\vec x_n-\vec\alpha\|\leqq\cdots\leqq\|\vec x_0-\vec\alpha\|, \tag{4.15}$$

$$(\vec x_i,\cdots,\vec x_{i-n})\in\mathcal{M}(c,\xi)\quad\textit{for } i=n,n+1,\cdots, \tag{4.16}$$
$$\textit{where } c>0 \textit{ if } 0<\zeta<1 \textit{ or } 0<c\leqq1 \textit{ if } \xi=0.$$

If $\tau=\sup_{i\geqq n}\tau_i<0.4$, *where* $\tau_i=K(\vec f)\cdot\|\vec x_{i-n}-\vec\alpha\|/d_i$ *then the sequence* $\{\vec x_i\}_{i=0}^{\infty}$ *satisfies the following conditions*

$$\|\vec x_{i+1}-\vec\alpha\|<\|\vec x_i-\vec\alpha\|, \tag{4.17}$$

$$\lim_{i\to\infty}\vec x_i=\vec\alpha, \tag{4.18}$$

and

$$\|\vec x_{i+1}-\vec\alpha\|\leqq A_i\cdot\|\vec x_{i-n}-\vec\alpha\|^{1-\xi}\cdot\|\vec x_i-\vec\alpha\|,\qquad i=n,n+1,\cdots, \tag{4.19}$$

where

$$A_i=\frac{5\cdot K(\vec f)}{2\cdot c}\cdot\frac{\|\vec x_{i+1}-\vec\alpha\|}{\|\vec x_{i+1}-\vec\alpha\|-\|\vec x_i-\vec\alpha\|}\Bigg)^{\xi}$$
$$\leqq\frac{5\cdot K(\vec f)}{2\cdot c\cdot(1-(5\cdot\tau/2)^n)^{\xi}}\quad\textit{and}\quad\limsup_{i\to\infty}A_i=\frac{5\cdot K(\vec f)}{2\cdot c}.$$

Proof. Observe that the assumptions of Lemma 4.1 are satisfied and hence (4.17) and (4.18) hold. From the definition of τ_i and (4.11) we get

$$\begin{aligned}\|\vec{x}_{i+1}-\vec{\alpha}\| &\leqq \frac{5\cdot K(\vec{f})}{2\cdot c}\cdot\frac{\|\vec{x}_{i-n}-\vec{\alpha}\|\cdot\|\vec{x}_i-\vec{\alpha}\|}{\|\vec{x}_i-\vec{x}_{i-n}\|^{\xi}}\\ &\leqq A_i\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|^{1-\xi}\cdot\|\vec{x}_i-\vec{\alpha}\|.\end{aligned} \tag{4.20}$$

Since

$$\begin{aligned}\|\vec{x}_i-\vec{\alpha}\| &\leqq 2.5\cdot\tau_i\cdot\|\vec{x}_i-\vec{\alpha}\|\\ &\leqq\cdots\leqq\|\vec{x}_{i-n}-\vec{\alpha}\|\prod_{j=i-n}^{i-1}(2.5\cdot\tau_j)\end{aligned}$$

from (4.20) we have

$$A_i\leqq\frac{5\cdot K(\vec{f})}{2\cdot c\cdot(1-(2.5\cdot\tau)^n)^{\xi}}.$$

The inequality (4.20) yields $\|\vec{x}_{i+1}-\vec{\alpha}\|/\|\vec{x}_i-\vec{\alpha}\|\xrightarrow[i\to\infty]{}0$ and $\lim_{i\to\infty}A_i=(5\cdot K(\vec{f}))/(2\cdot c)$, which proves (4.19) and Theorem 4.1. □

In the scalar case ($n=1$) the determinants d_i are equal to unity for arbitrary points $\vec{x}_i, \vec{x}_{i-1}, \vec{x}_i\neq\vec{x}_{i-1}$ and one can apply Theorem 4.1 with $\xi=0$ which yields the known error formula $|x_{i+1}-\alpha|=O(|x_i-\alpha|\cdot|x_{i-1}-\alpha|)$.

From theorem 4.1 it follows that the best convergence is achievable for $\xi=0$ (see Barnes [1], Bittner [2]). In this case

$$\tau_i\leqq\frac{K(\vec{f})}{c}\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|\leqq\frac{K(\vec{f})}{c}\cdot\|\vec{x}_0-\vec{\alpha}\|.$$

One can now rewrite the assumption on τ_i as

$$\|\vec{x}_0-\vec{\alpha}\|<\frac{2\cdot c}{5\cdot K(\vec{f})}. \tag{4.21}$$

The inequality (4.21) defines a ball of convergence for the MS method (C) for $\xi=0$ and $c\in(0,1]$.

5. Optimality of the secant method. The secant method is a stationary iteration which uses at each step the so-called standard information

$$\mathcal{N}=\{\vec{f}(\vec{x}_i),\vec{f}(\vec{x}_{i-1}),\cdots,\vec{f}(\vec{x}_{i-n})\}. \tag{5.1}$$

We shall show that this method makes the optimal use of the information (5.1). If the evaluation cost of $\vec{f}$ is sufficiently large then the most efficient iteration is that of maximal order of convergence. It was shown in Woźniakowski [7] that it is necessary to assume a good position of the points $\vec{x}_i,\cdots,\vec{x}_{i-n}$ to prove superlinear convergence. Therefore we shall assume that the sequence $(\vec{x}_i,\cdots,\vec{x}_{i-n})\in(\mathbb{C}^n)^{n+1}$ belongs for all $i\geqq n$ to a so-called admissible approximation set defined as follows.

DEFINITION 5.1. $\mathcal{M}$ is called *a set of admissible approximations* (shortly, $\mathcal{M}$ is an AA set) if

$$\mathcal{M}\subset B^{n+1}=\mathbb{C}^n\times\cdots\times\mathbb{C}^n, \tag{5.2}$$

and

(5.3) for any $\vec{\alpha} \in \mathbb{C}^n$ and $q \geqq 1$ there exists the sequence $\{\vec{x}_i\}_{i=0}^{\infty} \subset \mathbb{C}^n$ such that

(a) $$\lim_{i\to\infty} \vec{x}_i = \vec{\alpha},$$

(b) $$(\vec{x}_k, \vec{x}_{k-1}, \cdots, \vec{x}_{k-n}) \in \mathcal{M}, \quad \text{where } k = i(n+1),\ i = 0, 1, \cdots,$$

(c) $$\lim_{k\to\infty} \frac{\|\vec{x}_{k-j} - \vec{\alpha}\|}{\|\vec{x}_{k-n} - \vec{\alpha}\|^{q^{n-1}}} < \begin{cases} +\infty & \text{for } q > 1, \\ 1 & \text{for } q = 1, \end{cases}$$

for $j = 0, 1, \cdots, n-1$.

The sequence $\{\vec{x}_i\}$ is called a test sequence of qth order in $\mathcal{M}$ (shortly, $\{\vec{x}_i\}$ is of qth order), (compare with definition 2 in Woźniakowski [7]). If the limits in (c) are all nonzero, then $\{\vec{x}_i\}$ is of exactly qth order.

For fixed $\mathcal{N}$ and $\mathcal{M}$ we want to find an iteration with maximal order of convergence. It was proven in Woźniakowski [7] that the maximal order is equal to an order of information and is achievable by the so-called interpolatory iteration. The definition of the order of information can be found in [7]. A slightly modified proof of Theorem 4.1 points out that the order of the method (C) under assumption (4.16) of Theorem 4.1 as well as the order of information $p(\mathcal{N}, \mathcal{M}(c, \xi))(c > 0, 0 \leqq \xi < 1)$ are equal to the unique positive zero $p(\xi)$ of the polynomial

$$t^{n+1} - t^n - (1-\xi).$$

Since $p(\xi)$ is a decreasing function of $\xi \in [0, 1)$, it follows that

(5.4) $$p = \sup_{\xi\in[0,1)} p(\xi) = p(0).$$

We wish to find the best $\mathcal{M}$ for the standard information $\mathcal{N}$, i.e. that $\mathcal{M}$ which maximizes the order $p(\mathcal{N}, \mathcal{M})$. We shall prove that

$$p(\mathcal{N}, \mathcal{M}(c, 0)) = \sup_{\mathcal{M}\in\Omega} p(\mathcal{N}, \mathcal{M}),$$

where Ω is the class of all AA sets $\mathcal{M}$ such that $\mathcal{M} \subset \mathcal{M}(c, 0)$ or $\mathcal{M}(c, 0) \subset \mathcal{M}$ for some $c \in (0, 1]$. Note that $\mathcal{M}(c, \xi) \in \Omega$ for any positive c and $\xi \in [0, 1)$. We start with the following lemma.

LEMMA 5.1. *Let $\{\vec{x}_i\}_{i=1}^{\infty}$ be a test sequence of exactly p-th order converging to $\vec{\alpha}$ such that there exist positive numbers P and Q satisfying*

(5.5) $$Q \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\| \leqq \|\vec{x}_{i+1} - \vec{\alpha}\| \leqq P \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|$$

for $i \geqq n$,

(5.6) $$(\vec{x}_i, \cdots, \vec{x}_{i-n}) \in \mathcal{M}(c, 0) \quad \text{for} \quad i \geqq n.$$

Then for the sequence

$$\vec{\beta}^i = \widehat{\Delta X}_i^{-1} \times \frac{\vec{x}_i - \vec{\alpha}}{\|\vec{x}_i - \vec{\alpha}\|},$$

where

$$\widehat{\Delta X}_i = \Delta X_i \times \text{diag}_{i-n \leqq j \leqq i-1}(1/\|\overrightarrow{\Delta x_j}\|) \quad (i = 1, 2, \cdots)$$

it follows that $\lim_{i\to\infty} \inf |\beta_n^i| \neq 0$.

Proof. First we show that the assumption (5.5) does not contradict the condition (c) in Definition 5.1. From (5.5) it follows that there exist $Q_0, P_0 > 0$ such that

$$Q_0 \cdot \|\vec{x}_i - \vec{\alpha}\|^p \leqq \|\vec{x}_{i+1} - \vec{\alpha}\| \leqq P_0 \cdot \|\vec{x}_i - \vec{\alpha}\|^p, \quad \forall\, i \geqq n,$$

where p was defined by (5.2). Hence

$$Q_0^{l+1} \cdot \|\vec{x}_{i-l} - \vec{\alpha}\|^{p^{l+1}} \leqq \|\vec{x}_{i+1} - \vec{\alpha}\| \leqq p_0^{l+1} \cdot \|\vec{x}_{i-l} - \vec{\alpha}\|^{p^{l+1}}$$

$$\text{for } l = 0, 1, \cdots, n.$$

By replacing $i+1$ with $k-j$ we get

$$Q_0^{n-j} \leqq \frac{\|\vec{x}_{k-j} - \vec{\alpha}\|}{\|\vec{x}_{k-n} - \vec{\alpha}\|^{p^{n-j}}} \leqq P_0^{n-j} \quad \text{for } j = 0, 1, \cdots, n-1,$$

$$k = n, n+1, \cdots.$$

Hence, if the limits $\lim_{k\to\infty}(\|\vec{x}_{k-j} - \vec{\alpha}\| / \|\vec{x}_{k-n} - \vec{\alpha}\|^{p^{n-j}})$ exist $(j = 0, 1, \cdots, n-1)$ then they are positive and finite.

From the definition of $\vec{\beta}^i$ we get

$$\vec{x}_i - \vec{\alpha} = \|\vec{x}_i - \vec{\alpha}\| \cdot \sum_{k=1}^{n} \frac{\overrightarrow{\Delta x_{i-k}}}{\|\overrightarrow{\Delta x_{i-k}}\|} \cdot \beta_k^i, \qquad i = 1, 2, \cdots.$$

Hence, for $i \geqq n$ we have

$$\overrightarrow{\Delta x_i} = \frac{\overrightarrow{\Delta x_i}}{\|\overrightarrow{\Delta x_i}\|} \cdot \|\vec{x}_{i+1} - \vec{\alpha}\| \cdot \beta_1^{i+1} + \sum_{k=1}^{n-1} \frac{\overrightarrow{\Delta x_{i-k}}}{\|\overrightarrow{\Delta x_{i-k}}\|} \cdot (\|\vec{x}_{i+1} - \vec{\alpha}\| \cdot \beta_{k+1}^{i+1}$$

$$- \|\vec{x}_i - \vec{\alpha}\| \cdot \beta_k^i) - \frac{\overrightarrow{\Delta x_{i-n}}}{\|\overrightarrow{\Delta x_{i-n}}\|} \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \beta_n^i.$$

Next dividing by $\|\overrightarrow{\Delta x_i}\|$ we get the equality

$$\sum_{k=0}^{n-1} \frac{\overrightarrow{\Delta x_{i-k}}}{\|\overrightarrow{\Delta x_{i-k}}\|} \cdot \alpha_{k+1}^i = v^i, \tag{5.7}$$

where

$$\alpha_1^i = \frac{\|\vec{x}_{i+1} - \vec{\alpha}\|}{\|\overrightarrow{\Delta x_i}\|} \cdot \beta_1^{i+1} - 1,$$

$$\alpha_{k+1}^i = \frac{\|\vec{x}_{i+1} - \vec{\alpha}\|}{\|\overrightarrow{\Delta x_i}\|} \cdot \beta_{k+1}^{i+1} - \frac{\|\vec{x}_i - \vec{\alpha}\|}{\|\overrightarrow{\Delta x_i}\|} \cdot \beta_k^i$$

$$\text{for } k = 1, 2, \cdots, n-1,$$

$$\vec{v}_i = \frac{\|\vec{x}_i - \vec{\alpha}\|}{\|\overrightarrow{\Delta x_i}\|} \cdot \beta_n^i \cdot \frac{\overrightarrow{\Delta x_{i-n}}}{\|\overrightarrow{\Delta x_{i-n}}\|}.$$

The identity (5.7) means that $\vec{\alpha}^i = \widehat{\Delta X}_i^{-1} \times \vec{v}^i$. Hence

$$\|\vec{v}^i\| \geqq \frac{\|\vec{\alpha}^i\|}{\|\widehat{\Delta X}_i^{-1}\|} \geqq \frac{|\alpha_1^i|}{\|\widehat{\Delta X}_i^{-1}\|}. \tag{5.8}$$

Using Lemma 4.1, we get

$$\|\vec{v}^i\| \geqq \frac{c}{n} \cdot |\alpha_1^i| \tag{5.9}$$

and

$$|\beta_1^{i+1}| \leqq \|\vec{\beta}^{i+1}\| \leqq \|\widehat{\Delta X}^{-1}\| \leqq \frac{n}{c}.$$

The assumption (5.5) and (5.6) imply that for any $\varepsilon \in (0, 1)$ there exists an index i_0 such that

$$\begin{aligned} |\beta_1^{i+1}| \cdot \frac{\|\vec{x}_{i+1} - \vec{\alpha}\|}{\|\overline{\Delta x}_i\|} &\leqq |\beta_1^{i+1}| \cdot \frac{P \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|}{\|\vec{x}_1 - \vec{\alpha}\| - P \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|} \\ &< \frac{n \cdot \varepsilon}{c \cdot (1 - \varepsilon)} \quad \text{for } i \geqq i_0. \end{aligned} \tag{5:10}$$

From (5.9), (5.10) and (5.7) it follows that

$$\|\vec{v}^i\| \geqq \frac{c}{n} \cdot \left(1 - \frac{n \cdot \varepsilon}{c \cdot (1-\varepsilon)}\right) = \frac{c}{n} - \frac{\varepsilon}{1-\varepsilon} \tag{5.11}$$

for sufficiently small $\varepsilon > 0$ and $i \geqq i_0$.

Suppose now that there exists a subsequence $\{\beta_n^{i_s}\}_{i_s}$ of the sequence $\{\beta_n^i\}_{i=1}^{\infty}$ which converges to zero. Then for $i \equiv i_s \geqq i_0$ we have from (5.7) and the assumption (5.5)

$$\begin{aligned} \|\vec{v}^i\| &\leqq |\beta_n^i| \cdot \frac{\|\vec{x}_i - \vec{\alpha}\|}{|\|\vec{x}_i - \vec{\alpha}\| - P \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\||} \\ &\leqq |\beta_n^i| \cdot \frac{1}{1-\varepsilon} \xrightarrow[i \to \infty]{} 0 \end{aligned}$$

which contradicts (5.11). Hence $\liminf_{i \to \infty} |\beta_n^i| \neq 0$ which completes the proof. □

We now prove the theorem about the optimal AA set.

THEOREM 5.1. *Let Ω be a class of all* AA *sets $\mathcal{M}$ such that $\mathcal{M} \subset \mathcal{M}(c, 0)$ or $\mathcal{M}(c, 0) \subset \mathcal{M}$ for some $c \in (0, 1]$. Then*

$$\sup_{\mathcal{M} \in \Omega} p(\mathcal{N}, \mathcal{M}) = p(\mathcal{N}, \mathcal{M}(c, 0)).$$

Proof. Let $p = p(\mathcal{N}, \mathcal{M}(c, 0))$. It suffices to show that for any $\mathcal{M} \in \Omega$ there exist

(5.12) a function $\vec{f}$ which has property (2.2),

and a sequence

$\{\vec{x}_i\}_{i=0}^{\infty}$ of pth order in $\mathcal{M}$, for which

$$\limsup_{k \to \infty} \frac{\|\vec{\alpha}_k - \vec{\alpha}\|}{\|\vec{x}_{k-n} - \vec{\alpha}\|^{(p+\varepsilon)^{n+1}}} = +\infty, \quad \forall\, \varepsilon > 0, \tag{5.13}$$

where $\vec{\alpha}_k = \varphi(\vec{x}_k, \cdots, \vec{x}_{k-n}; \vec{f})$.

Let us consider the function

$$\vec{f}(\vec{x}) = \vec{x} - \vec{\alpha} + \vec{a} \cdot \|\vec{x} - \vec{\alpha}\|^2, \quad \vec{a} = [1, \cdots, 1]^T \in \mathbb{C}^n \tag{5.14}$$

and a sequence $\{\vec{x}_i\}_{i=0}^{\infty}$ of exactly pth order in $\mathcal{M}$ for which the assumptions of Lemma 5.1 hold with $\mathcal{M}$ instead of $\mathcal{M}(c, 0)$. From Lemma 5.1 it follows that there exist $K_1, K_2 > 0$ such that

$$K_1 = \frac{\|\vec{x}_{k-j} - \vec{\alpha}\|}{\|\vec{x}_{k-n} - \vec{\alpha}\|^{p^{n-j}}} < K_2 \tag{5.15}$$

for $j = 0, 1, \cdots, n-1$ and sufficiently large k. Let $\vec{w}^i$ be a linear function that interpolates $\vec{f}$ at $\vec{x}_i, \vec{x}_{i-1}, \cdots \vec{x}_{i-n}$, i.e. $\vec{w}^i(\vec{x}_k) = \vec{f}(\vec{x}_k)$ for $k = i, i-1, \cdots, i-n$. It is easy to see that

$$\vec{w}^i(\vec{x}) = \vec{f}(\vec{x}_i) + \Delta F_i \times (\Delta X_i^{-1}) \times (\vec{x} - \vec{x}_i) \tag{5.16}$$

and $\vec{w}^i(\vec{\alpha}_i) = 0$. Let us rewrite (5.10) in the form

$$\vec{w}^i(\vec{x}) = A^i \times \vec{x} + \vec{b}^i,$$

where A^i is a $n \times n$ matrix and $\vec{b}^i \in \mathbb{C}^n$. Hence the rows a_j^i $(j = 1, \cdots, n)$ of A^i satisfy equations

$$\begin{aligned} &(\vec{a}_j^i)^T \times \vec{x}_k + b_j^i - f_{jk} = 0, \quad \text{for } k = i, i-1, \cdots, i-n, \\ &(\vec{a}_j^i)^T \times \vec{\alpha}_i + b_j^i = 0, \end{aligned} \tag{5.17}$$

where $f_{jk} = f_j(\vec{x}_k)$. Note that the system (5.17) has a nonzero solution $[(\vec{a}_j^i)^T, b_j^i, -1]^T$ for any $j \in [1, n]$. Therefore its determinant has to vanish, i.e.,

$$\det \begin{bmatrix} (\vec{x}_{i-n} - \vec{\alpha})^T, & 1, & f_{ji-n} \\ & \cdots & \\ (\vec{x}_i - \vec{\alpha})^T, & 1, & f_{ji} \\ (\vec{\alpha}_i - \vec{\alpha})^T, & 1, & 0 \end{bmatrix} = 0.$$

Hence

$$\sum_{k=1}^{n} H_{jk}^i \times (\vec{\alpha}_k^i - \vec{\alpha}_k) = G_j^i \quad \text{for } j = 1, 2, \cdots, n, \tag{5.18}$$

where

$$H_{jk}^i = (-1)^{n+k} \cdot \det \begin{bmatrix} (\vec{x}_{i-n} - \vec{\alpha})^{kT}, & 1, & f_{ji-n} \\ & \cdots & \\ (\vec{x}_i - \vec{\alpha})^{kT}, & 1, & f_{ji} \end{bmatrix},$$

$$G_j^i = \det \begin{bmatrix} (\vec{x}_{i-n} - \vec{\alpha})^T, & f_{ji-n} \\ \cdots & \\ (\vec{x}_i - \vec{\alpha})^T, & f_{ji} \end{bmatrix},$$

and

$$y^k = [y_1, y_2, \cdots, y_{k+1}, y_{k+1}, \cdots, y_n]^T \quad \in \mathbb{C}^{n-1}.$$

From (5.14) we obtain

$$H_{jk}^i = \delta_{jk} \cdot w + E_{jk}^i, \tag{5.19}$$

where

$$\delta_{jk}=\begin{cases}1 & \text{if } j=k,\\ 0 & \text{if } j\neq k,\end{cases}$$

$$E^i_{jk}=(-1)^{n+k}\det\begin{bmatrix}(\vec{x}_{i-n}-\vec{\alpha})^{kT}, & 1, & \|\vec{x}_{i-n}-\vec{\alpha}\|^2\\ & \cdots & \\ (\vec{x}_i-\vec{\alpha})^{kT}, & 1, & \|\vec{x}_i-\vec{\alpha}\|^2\end{bmatrix},$$

$$w=\det\begin{bmatrix}(\vec{x}_{i-n}-\vec{\alpha})^T, & 1\\ (\vec{x}_i-\vec{\alpha})^T, & 1\end{bmatrix}=\det[\Delta X_i],$$

and

$$G^i_j=\det\begin{bmatrix}(\vec{x}_{i-n}-\vec{\alpha})^T, & \|\vec{x}_{i-n}-\vec{\alpha}\|^2\\ & \cdots \\ (\vec{x}_i-\vec{\alpha})^T, & \|\vec{x}_i-\vec{\alpha}\|^2\end{bmatrix}. \tag{5.20}$$

From this and (5.18) we get

$$(I+E^i/w)\times(\vec{\alpha}_i-\vec{\alpha})=\vec{G}^i/w, \tag{5.21}$$

where $E^i=[E^i_{jk}]^n_{j,k=1}$ and $\vec{G}^i=[G^i_1,\cdots,G^i_n]^T$. Furthermore from (5.19)

$$E^i_{jk}=2\cdot\det\begin{bmatrix}(\overrightarrow{\Delta x}_{i-n})^{kT}, & \dfrac{\|\vec{x}_{i-n+1}-\vec{\alpha}\|^2-\|\vec{x}_{i-n}-\vec{\alpha}\|^2}{2\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|}\\ & \cdots \\ (\overrightarrow{\Delta x}_{i-1})^{kT}, & \dfrac{\|\vec{x}_i-\vec{\alpha}\|^2-\|\vec{x}_{i-1}-\vec{\alpha}\|^2}{2\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|}\end{bmatrix}\cdot\|\vec{x}_{i-n}-\vec{\alpha}$$

and next by the Hadamard inequality we get

$$|E^i_{jk}|<2\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|\prod_{j=i-n+1}^{i}\left(\|\overrightarrow{\Delta x}_{j-1}\|+\frac{\|\vec{x}_j-\vec{\alpha}\|+\|\vec{x}_{j-1}-\vec{\alpha}\|}{2\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|}\cdot\|\Delta X_{j-i}\right|$$

$$\leqq 2^{n+1}\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|\prod_{j=i-n+1}^{i}\|\overrightarrow{\Delta x}_{j-1}\|.$$

Hence

$$\begin{aligned}\frac{\|E^i\|}{|w|}&\leqq n\cdot 2^{n+1}\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|\Big(\prod_{j=i-n+1}^{i}\|\overrightarrow{\Delta x}_{i-1}\|\Big)\Big/|w|\\ &\leqq\frac{n\cdot 2^{n+1}}{c}\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|,\end{aligned} \tag{5.22}$$

and from (5.21) it follows

$$\|\vec{\alpha}_i-\vec{\alpha}\|\geqq\frac{\|G^i\|/|w|}{1+(n\cdot 2^{n+1}/c)\cdot\|\vec{x}_{i-n}-\vec{\alpha}\|}=(1+O(\|\vec{x}_{i-n}-\vec{\alpha}\|))\cdot\frac{\|G^i\|}{|w|}. \tag{5.23}$$

Simple transformations of G_j^i yield

$$G_j^i = \|\vec{x}_i - \vec{\alpha}\| \Big(\prod_{k=1}^{n} \|\overrightarrow{\Delta x}_{i-k}\| \Big) \det \begin{bmatrix} \dfrac{\overrightarrow{\Delta x}_{i-n}^T}{\|\overrightarrow{\Delta x}_{i-n}\|}, & \dfrac{\|\vec{x}_{i-n+1} - \vec{\alpha}\|^2 - \|\vec{x}_{i-n} - \vec{\alpha}\|^2}{\|\overrightarrow{\Delta x}_{i-n}\|} \\ & \cdots \\ \dfrac{\overrightarrow{\Delta x}_{i-1}^T}{\|\overrightarrow{\Delta x}_{i-1}\|}, & \dfrac{\|\vec{x}_i - \vec{\alpha}\|^2 - \|\vec{x}_{i-1} - \vec{\alpha}\|^2}{\|\overrightarrow{\Delta x}_{i-1}\|} \\ \dfrac{(\vec{x}_i - \vec{\alpha})^T}{\|\vec{x}_i - \vec{\alpha}\|}, & \|\vec{x}_i - \vec{\alpha}\| \end{bmatrix}$$

$$= \det \left(\Delta X_i \times \operatorname*{diag}_{1 \leqq k \leqq n} \left(\frac{1}{\|\overrightarrow{\Delta x}_{i-k}\|} \right) \right) \cdot \left(\|\vec{x}_i - \vec{\alpha}\| - \sum_{k=1}^{n} \beta_k^i \frac{\|\vec{x}_{i-k+1} - \vec{\alpha}\|^2}{\|\overrightarrow{\Delta x}_{i-k}\|} \frac{-\|\vec{x}_{i-k} - \vec{\alpha}\|^2}{} \right) \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \prod_{k=1}^{n} \|\overrightarrow{\Delta x}_{i-k}\|,$$

where

$$\frac{\vec{x}_i - \vec{\alpha}}{\|\vec{x}_i - \vec{\alpha}\|} = \sum_{k=1}^{n} \beta_k^i \frac{\overrightarrow{\Delta x}_{i-k}}{\|\overrightarrow{\Delta x}_{i-k}\|}.$$

Hence, from Lemma 5.1 follows that $\liminf_{i \to \infty} |\beta_n^i| \neq 0$.

Let us now set,

$$\gamma_i = \frac{\|\vec{x}_i - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|} - \sum_{k=1}^{n} \beta_k^i \cdot \frac{\|\vec{x}_{i-k+1} - \vec{\alpha}\|^2 - \|\vec{x}_{i-k} - \vec{\alpha}\|^2}{\|\vec{x}_{i-n} - \vec{\alpha}\| \cdot \|\overrightarrow{\Delta x}_{i-k}\|}.$$

The inequality (5.15) gives

(a)
$$\frac{\|\vec{x}_i - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|} \leqq K_2 \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|^{p^n - 1} \xrightarrow[i \to \infty]{} 0,$$

(b) for $k = 1, 2, \cdots, n-1$,

$$\begin{aligned} &\frac{|\|\vec{x}_{i-k+1} - \vec{\alpha}\|^2 - \|\vec{x}_{i-k} - \vec{\alpha}\|^2|}{\|\vec{x}_{i-n} - \vec{\alpha}\| \cdot \|\overrightarrow{\Delta x}_{i-k}\|} \\ &\leqq \frac{\|\vec{x}_{i-k+1} - \vec{\alpha}\| + \|\vec{x}_{i-k} - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|} \\ &\leqq K_2 \cdot (\|\vec{x}_{i-n} - \vec{\alpha}\|^{p^{n+k+1} - 1} + \|\vec{x}_{i-n} - \vec{\alpha}\|^{p^{n-k} - 1}) \xrightarrow[i \to \infty]{} 0, \end{aligned}$$

(c) for $k = n$,

$$\begin{aligned} &\frac{|\|\vec{x}_{i-n+1} - \vec{\alpha}\|^2 - \|\vec{x}_{i-n} - \vec{\alpha}\|^2|}{\|\vec{x}_{i-n} - \vec{\alpha}\| \cdot \|\overrightarrow{\Delta x}_{i-n}\|} \\ &\geqq \frac{\|\vec{x}_{i-n} - \vec{\alpha}\| - \|\vec{x}_{i-n+1} - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|} \\ &\geqq 1 - K_2 \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|^{p-1} \xrightarrow[i \to \infty]{} 1. \end{aligned}$$

Hence $|\gamma_i| \geqq |\beta_n^i| \cdot (1 + O(\|\vec{x}_{i-n} - \vec{\alpha}\|^{p-1}))$ and

$$G_j^i \geqq |w_i| \cdot \|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\| \cdot |\beta_n^i| \cdot (1 + O(\|\vec{x}_{i-n} - \vec{\alpha}\|^{p-1})). \tag{5.24}$$

Finally, using the estimates (5.15) and (5.24) we get from (5.23) the lower bound

$$\begin{aligned} &\frac{\|\vec{\alpha}_i - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|^{(p+\varepsilon)n+1}} \\ &\geqq |\beta_n^i| \cdot (1 + O(\|\vec{x}_{i-n} - \vec{\alpha}\|^{p-1})) \frac{\|\vec{x}_i - \vec{\alpha}\| \cdot \|\vec{x}_{i-n} - \vec{\alpha}\|}{\|\vec{x}_{i-n} - \vec{\alpha}\|^{(p+\varepsilon)n+1}} \\ &\geqq \frac{|\beta_n^i| \cdot (1 + O(\|\vec{x}_{i-n} - \vec{\alpha}\|^{p-1})) \cdot K_1}{\|\vec{x}_{i-n} - \vec{\alpha}\|^{(n+1)\cdot p^n \cdot \varepsilon + O(\varepsilon^2)}} \xrightarrow[i \to \infty]{} \infty \quad \text{for } \varepsilon > 0; \end{aligned}$$

this completes the proof of Theorem 4.1. □

COROLLARY 5.1. *The* AA *set* $\mathcal{M}(c, 0)$ *is optimal in the class* Ω *for the standard information* $\mathcal{N}$.

Acknowledgment. I would like to thank H. Woźniakowski and B. Kacewicz for their comments and help during the preparation of this paper.

REFERENCES

[1] J. G. BARNES, *An algorithm for solving nonlinear equations based on the secant method*, Comput. J. 8 (1965), pp. 66–72.

[2] L. BITTNER, *Eine Verallgemeinerung des Sekantenverfahrens zur naherungsweisen Berechung der Nullstellen eines nichtlinearen Gleichungssystems*, Wiss. Z. Techn. Univ. Dresden, 9 (1959/60), pp. 325–329.

[3] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[4] J. F. TRAUB, *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[5] S. ULHM, *Iteration methods with divided differences of the second order*, Dokl. Akad. Nauk SSSR 158, pp. 55–58; Soviet Math. Dokl., 5 (1964), pp. 1187–1190.

[6] R. VOIGT, *Order of convergence for iterative procedures*, this Journal, 8 (1971), pp. 222–243.

[7] H. WOŹNIAKOWSKI, *Generalized information and maximal order of iteration for operator equations*, this Journal, 12 (1975), pp. 121–135.