GENERALIZED INFORMATION AND MAXIMAL ORDER
OF ITERATION FOR OPERATOR EQUATIONS

H. Wozniakowski

Department of Computer Science
Carnegie-Mellon University

(Visiting from the University of Warsaw)

April 1974

ABSTRACT

We consider stationary iterative methods for the solution of operator

equations which use "generalized information". We seek methods with maximal

order provided the points of iteration are in "good position". The new con-

cepts of order of information and "generalized" interpolatory methods are

introduced. The main result states that the maximal order is equal to the

order of information which depends on generalized information and on the

position of iteration points. We show that the maximal order is achievable

by the generalized interpolatory method.

1. INTRODUCTION

This paper deals with stationary iterative methods for the solution of operator equations. One of the characteristics of an iterative method is the information used at each iterative step. For fixed information we seek methods of maximal order.

Papers in this area have generally assumed that the information is given by the values of the operator and its first s derivatives at n previous iteration points. For this information, previous results are described below.

The problem of maximal order was first posed by Traub [61,64]. Traub's conjecture states that the maximal order for the scalar case is no greater than s+2. This conjecture was proved by Brent, Winograd and Wolfe [73] for a particular class of non-stationary iterations. The maximal order of stationary iterative methods was considered by Kung and Traub [73a], Rissanen [71], Traub [64,72] and Wozniakowski [72,73]. For the scalar case the maximality of the interpolatory methods (see Traub [64], p. 60 and ff.) in certain classes of admissible stationary methods was proved by Traub [64] and Kung and Traub [73a] for n = 0 and s ≥ 0, by Rissanen [71] for n = 1 and s = 0, and by Wozniakowski [73] for any n,s such that n+s ≥ 1.

For the operator case (which includes the multivariate case), Wozniakowski [73] proved that the maximal order does not depend on n and is equal to s+1. Hence the additional information contained at the previous points cannot increase the maximal order. If, however, one assumes a certain position of the successive approximations, then there exist stationary methods with order greater than s+1 (see Barnes [65], Brent [72], Jankowska [73], Ortega and Rheinboldt [70], p. 269 and Wozniakowski [72]).

In this paper we study the subject based on the following two points of view. The first is to assume the suitable position of all approximations. In order to do this we define a set of admissible approximations, $\mathfrak{M}$, and test sequences (Section 2).

The second is to use "generalized information", $\mathfrak{N}$ (Section 3). For many practical problems the information given by the values of the operator and its first derivatives is the most important. But for some other problems, other types of information could be used more effectively and it is interesting to investigate how other types of information can effect the maximal order. For instance, Kacewicz [73] considers iterative methods which use the values of the scalar function, its first s derivatives and one integral per step and proves that the maximal order for this class of iterative methods is equal to s+3. (Note that the maximal order is s+1 if the value of the integral is not used.)

Thus, for fixed $\mathfrak{M}$ - the set of admissible approximations and $\mathfrak{N}$ - the "generalized information", we define the order of a stationary iterative method (Sections 4 and 5). Next we determine the maximal order as a function of $\mathfrak{M}$ and $\mathfrak{N}$ and seek methods which attain maximal order (Section 6). Specifically we prove that the maximal order is equal to the order of information $\mathfrak{N}$ with respect to $\mathfrak{M}$. We show that maximal order is achievable by the generalized interpolatory method which is defined in Section 8. The last section contains examples of $\mathfrak{N}$, $\mathfrak{M}$ and their order of information.

The proof techniques in this paper are based on the concept of order of information and a new definition of the order of iterative methods. We hope that these concepts lead to effective and rather easy proofs of maximal order of one-point stationary iteration with memory. It seems that this technique could be useful for other problems such as the maximal order of multipoint iterations (see Conjecture 8.1 in Kung and Traub [73a],and Kung and Traub [73b]).

2. SET OF ADMISSIBLE APPROXIMATIONS

Let us consider the problem of solving the nonlinear equation

(2.1) $F(x) = 0$,

where $F: D_F \subset B_1 \to B_2$ and B_1, B_2 are real or complex Banach spaces. We solve (2.1) by iteration. Properties of iterative methods depend on the regularity of F and on the multiplicity of the solution. Therefore we assume that F belongs to a class $\mathcal{F}$ defined as follows.

Definition 1

Let $\mathcal{F}$ be a class of F, $F: D_F \subset B_1 \to B_2$ such that

(i) there exists a simple zero $\alpha = \alpha(F) \in D_F$, that is, $F(\alpha) = 0$ and $[F'(\alpha)]^{-1}$ is a linear bounded operator,

(ii) $F(x) = \sum_{k=1}^{\infty} \frac{1}{k!} F^{(k)}(\alpha)(x-\alpha)^k$

for $x \in S(\alpha, \Gamma) = \{x: ||x-\alpha|| \leq \Gamma\}, \Gamma > 0$,

where $F^{(k)}(\alpha)$ denotes the k-th Frechet derivative of F. ∎

For a fixed nonnegative integer n, let $x_d, x_{d-1}, \ldots, x_{d-n}$ be different approximations of the solution α. As we mentioned in the introduction, the crucial point for the operator case is the assumption of suitable positions of $x_d, \ldots, x_{d-n}$ in B_1 space. Hence we assume that these points belong to a set $\mathfrak{M}$ defined as follows.

Definition 2

$\mathfrak{M}$ is called a set of admissible approximations (shortly, $\mathfrak{M}$ is AA set) iff

(i) $\mathfrak{M} \subset B_1^{n+1} = \underbrace{B_1 \times \ldots \times B_1}_{n+1 \text{ times}}$,

(ii) $\forall \alpha \in B_1$, $\forall q \geq 1$, $\exists \{x_i\} \subset B_1$ such that

   (a) $\lim_{i \to \infty} x_i = \alpha$,

   (b) $(x_d, x_{d-1}, \ldots, x_{d-n}) \in \mathfrak{M}$ where $d = i(n+1)$, $i = 0, 1, \ldots$,

   (c) $\lim_{d \to \infty} \dfrac{||x_{d-j} - \alpha||}{||x_{d-n} - \alpha||^{q^{n-j}}} < \begin{cases} + \infty & \text{for } q > 1, \\ 1 & \text{for } q = 1, \end{cases}$

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

The sequence $\{x_i\}$ is called a _test sequence_ of qth order in $\mathfrak{M}$. More briefly we shall say $\{x_i\}$ is of qth order. If the limits in (c) are all non-zero, then we say that $\{x_i\}$ is of exactly qth order. ∎

The second condition of Definition 2 implies that for any point α of B_1 and for any $q \geq 1$, $\mathfrak{M}$ contains at least one test sequence of qth order which converges to α.

Example 2.1

$\mathfrak{M} = B_1^{n+1}$.

In this case we do not assume any special position of the successive approximations. Many previous papers on iterative processes deal with this $\mathfrak{M}$. ∎

Example 2.2

Let $B_1 = B_2 = C^n$ - nth dimensional complex space.

$$\mathfrak{M} = \{(x_n, x_{n-1}, \ldots, x_0) : |\det(\frac{x_n - x_{n-1}}{||x_n - x_{n-1}||}, \ldots, \frac{x_1 - x_0}{||x_1 - x_0||})| \geq c\}$$

for a fixed $c \in (0,1)$.

Note that from the Hadamard inequality it follows that

$$|\det(\frac{x_n - x_{n-1}}{||x_n - x_{n-1}||}, \ldots, \frac{x_1 - x_0}{||x_1 - x_0||})| \leq 1.$$

Thus, $\mathfrak{M}$ contains points for which the Hadamard inequality holds from below. It is easy to verify that $\mathfrak{M}$ is an AA set. This set is especially important for the multivariate secant method (see Barnes [65], Jankowska [73], Ortega and Rheinboldt [70] and the next part of this paper). ∎

Let us consider $\mathfrak{M}$ such that

(2.2) $\forall \alpha \in B_1$, $\exists \Gamma > 0$ such that $S(\alpha, \Gamma)^{n+1} \subset \mathfrak{M}$.

This means that any points $(x_d, x_{d-1}, \ldots, x_{d-n})$ close enough to α belong to $\mathfrak{M}$ and no assumptions of position are required. Therefore, if (2.2) holds, we say that $\mathfrak{M}$ is an _unconditional_ AA set and otherwise $\mathfrak{M}$ is said to be a _conditional_ AA set. Example 2.1 gives an unconditional AA set while Example 2.2 gives a conditional one.

3. INFORMATION

Let the points $(x_d, x_{d-1}, \ldots, x_{d-n}) \in \mathfrak{M}$ be approximations to α. To obtain the next approximation y_d we need certain "generalized information" of F at these points. This generalized information is defined as follows.

Definition 3

An operator $\mathfrak{N}$ is called <u>generalized information</u> or simply <u>information</u> iff

(i) $\mathfrak{N}: D_\mathfrak{N} \to V$

where $D_\mathfrak{N} \subset B_1^{n+1} \times \mathcal{F}$ and V is a certain space, and

(ii) $\forall F \in \mathcal{F}$, $F(\alpha) = 0$, $\exists \Gamma > 0$ such that

$$S(\alpha, \Gamma)^{n+1} \times \{F\} \subset D_\mathfrak{N}$$

The second condition implies that for any points $(x_n, \ldots, x_0)$ sufficiently close to α, $\mathfrak{N}(x_n, \ldots, x_0; F)$ is well defined. Usually $\mathfrak{N}(x_n, \ldots, x_0; F)$ is defined to be the set of the values of F or its derivatives at the points $x_n, \ldots, x_0$. (See the following example.)

Example 3.1

$$\mathfrak{N}(x_n, \ldots, x_0; F) = \{F^{(k)}(x_j): k = 0, 1, \ldots, s; \; j = 0, 1, \ldots, n\}$$

for a fixed s.

More briefly, write $\mathfrak{N}_s$ for $\mathfrak{N}(x_n, \ldots, x_0; F)$. This information is called <u>standard information</u> and was considered by many authors (see, e.g., Traub [64]). The problem of finding maximal methods which use $\mathfrak{N}_s$ with respect to $\mathfrak{M} = B_1^{n+1}$ has been settled (see, e.g., Wozniakowski [73]).

Example 3.2

Let $B_1 = B_2 = C$ and $n = 0$.

$$\Re(x_0;F) = \{F(x_0),F'(x_0),\ldots,F^{(s)}(x_0), \int_{a_0}^{x_0} F(t)dt\}.$$

where $s \geq 1$ and $a_0 = x_0 - \dfrac{s+3}{s+2} \dfrac{F(x_0)}{F'(x_0)}$.

In this case $\Re(x_0;F)$ includes also the value of the integral $\int_{a_0}^{x_0} F(t)dt$.

The particular choice of the a_0 can be shown to be optimal in a certain sense.

For details, see Example 9.3 of this paper and Kacewicz [73]. ■

4. STATIONARY ITERATIVE METHODS

For given $\mathfrak{M}$ and $\mathfrak{N}$ a stationary iterative method is defined as follows.

Suppose that $(x_d, \ldots, x_{d-n}) \in \mathfrak{M}$ are approximations sufficiently close to the solution α. Hence the information $\mathfrak{N}(x_d, \ldots, x_{d-n}; F)$ is well defined. The next y_d approximation is given by

$$(4.1) \quad y_d = \varphi_{\mathfrak{N}, \mathfrak{M}}(x_d; F)$$

where

$$(4.2) \quad \varphi_{\mathfrak{N}, \mathfrak{M}}(x_d; F) = \varphi(x_d, \ldots, x_{d-n}; \mathfrak{N}(x_d, \ldots, x_{d-n}; F))$$

for an operator φ,

$$(4.3) \quad \varphi: D_\varphi \to B_1, \quad D_\varphi \subset \mathfrak{M} \times V.$$

If $(y_d, x_d, \ldots, x_{d-n+1}) \in \mathfrak{M}$ one can set

$$x_{d+n+1} = y_d, \quad x_{d+n-j} = x_{d-j}, \quad j = 0, 1, \ldots, n-1$$

and perform the next iterative step with $(x_{d+n+1}, \ldots, x_{d+1})$. But if $(y_d, x_d, \ldots, x_{d-n+1}) \notin \mathfrak{M}$ one must define the next approximations $(x_{d+n+1}, \ldots, x_{d+1})$ in a different way and sometimes additional information of F is needed (see e.g. Jankowska [73] and Ortega and Rheinboldt [70], pp. 369-390).

Definition 4

An iterative method defined by (4.1), (4.2) and (4.3) is called a stationary iterative method $\varphi_{\mathfrak{N}, \mathfrak{M}}$ or, more briefly, a $\varphi_{\mathfrak{N}, \mathfrak{M}}$ method. ∎

5. ORDER OF STATIONARY ITERATIVE METHODS

Let $\varphi_{\mathfrak{N},\mathfrak{M}}$ be a stationary iterative method. If $\mathfrak{M}$ is a conditional AA set then $\varphi_{\mathfrak{N},\mathfrak{M}}$ cannot generate, in general, sequences converging to the solution, since then $(y_d, x_d, \ldots, x_{d-n+1})$ does not always belong to $\mathfrak{M}$. definition of the order of $\varphi_{\mathfrak{N},\mathfrak{M}}$ based on properties of generating sequences must fail and we therefore have to define an order of $\varphi_{\mathfrak{N},\mathfrak{M}}$ differently. We shall use test sequences instead of generating sequences. That is, let $\{x_i\}$ be a test sequence of qth order in the sense of Definition 2.

Let

$$(5.1) \quad y_d = \varphi_{\mathfrak{N},\mathfrak{M}}(x_d; F)$$

for any $d = i(n+1)$, $i = 0, 1, \ldots$.

Further, let $\{F_d\} \subset \widetilde{\mathcal{F}}$ be a sequence of operators for which

$$(5.2) \quad F_d(\alpha_d) = 0 \text{ and } \lim_{d \to \infty} \alpha_d = \alpha$$

$$(5.3) \quad \lim_{d \to \infty} F_d^{(k)}(\alpha) = G^{(k)}(\alpha)$$

where $G(\alpha) = 0$ and $G \in \widehat{\mathcal{F}}$.

Next we assume that the information on F and on F_d at the points $x_d, x_{d-1}, \ldots, x_{d-n}$ is the same, that is,

$$(5.4) \quad \mathfrak{N}(x_d, \ldots, x_{d-n}; F_d) = \mathfrak{N}(x_d, \ldots, x_{d-n}; F), \; \forall_d.$$

Hence, $\varphi_{\mathfrak{N},\mathfrak{M}}(x_d; F_d) = \varphi_{\mathfrak{N},\mathfrak{M}}(x_d; F)$ for any stationary method $\varphi_{\mathfrak{N},\mathfrak{M}}$.

Definition 5

If $\{F_d\} \subset \widehat{\mathcal{F}}$ satisfies (5.2), (5.3) and (5.4) then we say that $\{F_d\}$ is equal to F with respect to $\mathfrak{N}$. (Or briefly, $\{F_d\}$ equals F.) ∎

We give an example of $\{F_d\}$.

Example 5.1

Let $B_1 = B_2 = C$, $\mathfrak{M} = C^{n+1}$ and $\mathfrak{N} = \mathfrak{N}_s$ be standard information (see Example 3.1).

We put

$$(5.5) \quad F_d(x) = F(x) + \prod_{j=0}^{n} (x - x_{d-j})^{s+1}.$$

One can verify that there exists $\{\alpha_d\}$ such that

$$F_d(\alpha_d) = 0 \text{ and } \alpha_d - \alpha = O(\prod_{j=0}^{n} (x_{d-j} - \alpha)^{s+1}).$$

Clearly, (5.3) holds for $G(x) = F(x) + (x - \alpha)^{(n+1)(s+1)}$.

From (5.5),

$$F_d^{(k)}(x_{d-j}) = F^{(k)}(x_{d-j}), \quad k = 0,1,\ldots,s; \; j = 0,1,\ldots,n.$$

Hence, $\{F_d\}$ equals F. ∎

By (5.4) the operators F and F_d are same with respect to the information at the points $x_d, x_{d-1}, \ldots, x_{d-n}$. Hence, the next approximation y_d defined by (5.1) is the same for F and F_d. Using information (5.4) we do not know which operator, F or F_d, is considered in the iterative process. Therefore, y_d ought to be an approximation not only to α but also to α_d at the same time. This discussion leads us to the following definition of order.

Definition 6

A number $p = p(\varphi_{\mathfrak{N},\mathfrak{M}}) \geq 1$ is called an order of $\varphi_{\mathfrak{N},\mathfrak{M}}$ if the following conditions hold:

(i) $\forall F \in \mathcal{F}$, $F(\alpha) = 0$, $\forall \{x_i\}$ of qth order and $\forall \{F_d\}$ equals F, the

values $y_d = \varphi_{\mathfrak{N},\mathfrak{M}}(x_d; F)$ are well defined for large $d = i(n+1)$,

$i = 0,1,\ldots$, and for sufficiently small $\epsilon > 0$,

$$(5.6) \quad \limsup_{d \to \infty} \frac{||y_d - \alpha_d||}{e_d^{\mu}} < +\infty$$

where $\mu = [\min(q + \epsilon, p - \epsilon)]^{n+1}$ and $e_d = ||x_{d-n} - \alpha||$,

(ii) $\exists F \in \mathcal{F}$, $F(\alpha) = 0$, $\{x_i\}$ of pth order and $\{F_d\}$ equals F such that

$$(5.7) \quad \limsup_{d \to \infty} \frac{||y_d - \alpha_d||}{e_d^{(p+\epsilon)^{n+1}}} = +\infty, \quad \forall \epsilon > 0. \qquad \blacksquare$$

In Definition 6 we compare y_d with α_d. But if one sets $F_d = F$ for any

d, then $\alpha_d = \alpha$ and the distance between y_d and α_d is just $||y_d - \alpha||$.

It is easy to verify that if p exists, then p is unique. The new defini-

tion of order is based on test sequences $\{x_i\}$ and on sequences of operators

$\{F_d\}$ equal to F. There are many relations between the order defined in this

paper and the orders used by many other authors, especially when $\mathfrak{M}$ is an un-

conditional AA set. But this subject is not the topic of this paper and will

be deferred to a future paper.

6. MAXIMAL METHODS

Let $\Psi(\mathfrak{N};\mathfrak{M})$ be the class of stationary iterative methods $\varphi_{\mathfrak{N},\mathfrak{M}}$ with well defined order $p(\varphi_{\mathfrak{N},\mathfrak{M}})$.

Definition 7

A $\bar{\varphi}_{\mathfrak{N},\mathfrak{M}}$ method is called maximal if it has order as high as possible, i.e.,

$$p(\bar{\varphi}_{\mathfrak{N},\mathfrak{M}}) = \sup_{\varphi_{\mathfrak{N},\mathfrak{M}} \in \Psi(\mathfrak{N};\mathfrak{M})} p(\varphi_{\mathfrak{N},\mathfrak{M}}). \qquad \blacksquare$$

The next part of the paper deals with maximal methods. For fixed information, we shall define an order of information, $p(\mathfrak{N};\mathfrak{M})$, and then prove that the maximal order is equal to the order of information. Note that the class $\Psi(\mathfrak{N};\mathfrak{M})$ contains all $\varphi_{\mathfrak{N},\mathfrak{M}}$ methods with well defined order. In previous papers the maximality problem was solved for restricted classes of iteration methods. (See Brent, Winograd and Wolfe [73], Theorem 6.1 on the optimal order of one-point iterations in Traub [64] and Kung and Traub [73a] and Wozniakowski [73].) Due to the new definition of order (Definition 6) we are able to solve the maximality problem without additional assumptions on the considered iterative methods.

7. ORDER OF INFORMATION

The order of information is defined by the sequences $\{\alpha_d - \alpha\}$ where α_d and α are the zeros of the operators F_d and F, respectively (see Definition 5).

Definition 7

A number $p = p(\mathfrak{N};\mathfrak{M}) \geq 1$ is called <u>an order of information $\mathfrak{N}$ with respect to $\mathfrak{M}$</u> (shortly <u>an order of information</u>) iff

(i) $\forall F \in \mathcal{F}$, $F(\alpha) = 0$, $\forall\{x_i\}$ of qth order and $\forall\{F_d\}$ equals F,

$$(7.1) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d^{\mu}} < +\infty$$

holds for sufficiently small $\varepsilon > 0$ where

$$\mu = [\min(q + \varepsilon, \; p - \varepsilon)]^{n+1} \quad \text{and} \quad e_d = ||x_{d-n} - \alpha||,$$

(ii) $\exists F \in \mathcal{F}$, $F(\alpha) = 0$, $\{x_i\}$ of pth order and $\{F_d\}$ equals F such that

$$(7.2) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d^{(p+\varepsilon)^{n+1}}} = +\infty, \quad \forall \varepsilon > 0. \quad \blacksquare$$

It can be shown that if the order of information exists, then it is unique. We <u>assume</u> that for the considered $\mathfrak{N}$ and $\mathfrak{M}$, the order of information $p(\mathfrak{N};\mathfrak{M})$ exists.

Lemma 1

Let $p_j \geq 0$, $\sum_{j=0}^{n} p_j t^{n-j} \geq 1$ and let p be the unique positive zero of the polynomial

$$(7.3) \quad t^{n+1} - \sum_{j=0}^{n} p_j t^{n-j}.$$

If the following conditions hold:

(i) $\forall F \in \hat{\mathcal{F}}$, $F(\alpha) = 0$, $\forall \{x_i\}$ of qth order and $\forall \{F_d\}$ equals F,

$$(7.4) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{\prod\limits_{j=0}^{n} ||x_{d-j} - \alpha||^{p_j}} < +\infty$$

holds, and

(ii) $\exists F \in \mathcal{F}$, $F(\alpha) = 0$, $\{x_i\}$ of exactly pth order and $\{F_d\}$ equals F such that

$$(7.5) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{\prod\limits_{j=0}^{n} ||x_{d-j} - \alpha||^{p_j}} > 0,$$

then p is the order of information $\mathfrak{N}$ with respect to $\mathfrak{M}$. ∎

Proof

Since $\{x_i\}$ is of qth order, then

$$(7.6) \quad \prod_{j=0}^{n} ||x_{d-j} - \alpha||^{p_j} = O(e_d^{\sum\limits_{j=0}^{n} p_j q^{n-j}}), \quad e_d = ||x_{d-n} - \alpha||.$$

If $q < p$ then $q^{n+1} < \sum\limits_{j=0}^{n} p_j q^{n-j}$ and for small $\varepsilon > 0$, (7.4) and (7.6) give

$$(7.7) \quad ||\alpha_d - \alpha|| = O(e_d^{(q+\varepsilon)^{n+1}}).$$

Otherwise, if $q \geq p$, then

$$\sum_{j=0}^{n} p_j q^{n-j} \geq \sum_{j=0}^{n} p_j p^{n-j} = p^{n+1}$$

and from (7.4) and (7.6) follows

(7.8) $\quad ||\alpha_d - \alpha|| = 0(e_d^{p^{n+1}})$.

Hence, (7.7) and (7.8) give (7.1).

Let F, $\{x_i\}$ and $\{F_d\}$ be as in condition (ii) of Lemma 1. Now $\{x_i\}$ is of exactly pth order which means that

$$\prod_{j=0}^{n} ||x_{d-j} - \alpha||^{p_j} = C_d \, e_d^{p^{n+1}} \quad \text{where } \liminf_{d \to +\infty} C_d > 0.$$

From (7.5) follows

$$\limsup_{d \to \infty} \frac{||\alpha_d - \alpha||}{e_d^{p^{n+1}}} > 0$$

which proves (7.2) and completes the proof. ∎

Example 7.1

Let $B_1 = B_2 = C$ and $\mathfrak{M} = C^{n+1}$. Let $\mathfrak{N}$ be the standard information $\mathfrak{N}_s$ (see Examples 3.1 and 5.1). Assume that $\{F_d\}$ equals F. Then

$$F_d^{(k)}(x_{d-j}) = F^{(k)}(x_{d-j}), \quad k = 0,1,\ldots,s; \ s = 0,1,\ldots,n.$$

From the remainder formula we get

(7.9) $\quad F(x) - F_d(x) = G_d(x) \prod_{j=0}^{n} (x - x_{d-j})^{s+1}$

where $\quad G_d(x) = \dfrac{\lambda(x)}{(r+1)!} [F - F_d]^{(r+1)}(\xi)$

for $|\lambda(x)| \leq 1$, $r = (n+1)(s+1) - 1$ and $\xi \in \text{conv}(x, x_d, \ldots, x_{d-n})$.

Setting $x = \alpha$ in (7.9), we have

$$(7.10) \quad \alpha_d - \alpha = O(\prod_{j=0}^{n} (x_{d-j} - \alpha)^{s+1})$$

due to (5.3). Moreover (7.10) is sharp. From Lemma 1 follows that the order of standard information is equal to the unique positive zero of the polynomial

$$(7.11) \quad t^{n+1} - (s+1) \sum_{j=0}^{n} t^j. \qquad \blacksquare$$

8. MAIN RESULTS

In this section we prove that the maximal order is equal to the order of information.

Theorem 8.1

If $\varphi_{\mathfrak{N},\mathfrak{M}} \in \Psi(\mathfrak{N};\mathfrak{M})$ then $p(\varphi_{\mathfrak{N},\mathfrak{M}}) \leq p(\mathfrak{N};\mathfrak{M})$. ∎

Proof

Suppose that $\varphi_{\mathfrak{N},\mathfrak{M}}$ has order $\bar{p} > p = p(\mathfrak{N};\mathfrak{M})$. By (7.2) there exist F, $\{x_i\}$ of pth order and $\{F_d\}$ equals F such that

$$(8.1) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d^{\mu}} = +\infty$$

where $e_d = ||x_{d-n} - \alpha||$ and $\mu = (p+\varepsilon)^{n+1}$ for small $\varepsilon > 0$. From the definition of the order of $\varphi_{\mathfrak{N},\mathfrak{M}}$ it follows that for the same F, $\{x_i\}$ and $\{F_d\}$ the next approximations $\{y_d\}$ satisfy

$$(8.2) \quad \limsup_{d \to +\infty} \frac{||y_d - \alpha_d||}{e_d^{\mu}} < +\infty.$$

Setting $F_d \equiv F$ we get

$$(8.3) \quad \limsup_{d \to +\infty} \frac{||y_d - \alpha||}{e_d^{\mu}} < +\infty.$$

Hence, from (8.2) and (8.3),

$$\limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d^{\mu}} \leq \limsup_{d \to \infty} \frac{||y_d - \alpha|| + ||y_d - \alpha_d||}{e_d^{\mu}} < +\infty$$

which contradicts (8.1) and completes the proof. ∎

From Theorem 8.1 it follows that the order of information is an upper

bound on the order of stationary iterative methods. We now shall define a

method whose order achieves this upper bound.

Let F be any operator from $\mathcal{F}$, $F(\alpha) = 0$, and let $\{x_i\}$ be any test sequence

of qth order. For these F and $\{x_i\}$, let $\{F_d^*\}$ be any sequence of operators

equal to F. Note that at least one such sequence exists, for instance, $F_d^* \equiv F$.

We define <u>a generalized interpolatory method</u> $I_{\mathfrak{N},\mathfrak{M}}$ (shortly <u>an interpolatory</u>

<u>method</u> $I_{\mathfrak{N},\mathfrak{M}}$) by

$$(8.4) \quad I_{\mathfrak{N},\mathfrak{M}}(x_d;F) \overset{df}{=} \alpha_d^*$$

where α_d^* is the zero of F_d^*, $\lim_{d \to \infty} \alpha_d^* = \alpha$. The terminology "interpolatory" is

used because F_d^* has the same values as F at the points $x_d, x_{d-1}, \ldots, x_{d-n}$ with

respect to the information $\mathfrak{N}$, i.e.,

$$(8.5) \quad \mathfrak{N}(x_d, \ldots, x_{d-n}; F_d^*) = \mathfrak{N}(x_d, \ldots, x_{d-n}; F).$$

Note that F_d^* is not uniquely defined by (8.5), in general, but we shall prove

that any such F_d^* gives the same maximal order. However, the uniqueness of

F_d^* can be assured for certain $\mathfrak{N}$ and $\mathfrak{M}$. For instance, we can demand that F_d^*

be the minimum degree interpolatory polynomial in case where that is unique.

(Compare the scalar case with the standard information $\mathfrak{N}_s$, Traub [64], p.60

and ff.)

<u>Theorem 8.2</u>

The interpolatory method $I_{\mathfrak{N},\mathfrak{M}}$ is maximal, i.e.,

$$p(I_{\mathfrak{N},\mathfrak{M}}) = p(\mathfrak{N};\mathfrak{M})$$

Proof

Recall that $F \in \mathcal{F}$, $F(\alpha) = 0$, $\{x_i\}$ is of qth order and the next approximation y_d in the $I_{\mathfrak{N},\mathfrak{M}}$ method is equal to α_d^*, where $\{\alpha_d^*\}$ are zeros of $\{F_d^*\}$ equals F.

Let $\{F_d\}$ be any sequence equal to F. We want to show (5.6), i.e.,

$$(8.6) \quad \limsup_{d \to \infty} \frac{||y_d - \alpha_d||}{e_d^\mu} < +\infty$$

where as always $e_d = ||x_{d-n} - \alpha||$, $\mu = [\min(q+\epsilon, p-\epsilon)]^{n+1}$ for $p = p(\mathfrak{N};\mathfrak{M})$. From (7.1) follows

$$||\alpha_d^* - \alpha|| = O(e_d^\mu) \text{ and } ||\alpha_d - \alpha|| = O(e_d^\mu).$$

Hence

$$||y_d - \alpha_d|| \leq ||\alpha_d^* - \alpha|| + ||\alpha_d - \alpha|| = O(e_d^\mu)$$

and (8.6) holds.

Now we wish to show (5.7). Let F, $\{x_i\}$ of pth order and $\{F_d\}$ equal to F satisfy (7.2), i.e.,

$$(8.7) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d^\nu} = +\infty \qquad \text{for } \nu = (p+\epsilon)^{n+1}.$$

Let

$$(8.8) \quad C \stackrel{df}{=} \limsup_{d \to \infty} \frac{||y_d - \alpha||}{e_d^\nu}.$$

If $C < +\infty$ then

$$\limsup_{d \to \infty} \frac{||y_d - \alpha_d||}{e_d^\nu} \geq \limsup_{d \to \infty} \frac{||\alpha_d - \alpha||}{e_d^\nu} - \limsup_{d \to \infty} \frac{||y_d - \alpha||}{e_d^\nu} = +\infty$$

and (5.7) holds for the same F, $\{x_i\}$ and $\{F_d\}$ as above. Otherwise, if $C = +\infty$

then (5.7) holds for the same F and $\{x_i\}$ as above and for $F_d \equiv F$.

This completes the proof that the order of information is the order of the interpolatory method $I_{\mathfrak{N},\mathfrak{M}}$. Due to Theorem 8.1 the $I_{\mathfrak{N},\mathfrak{M}}$ method is maximal. ∎

The basic ideas of the interpolatory method are

(i) to find an operator F_d^* which fits the given information,

(ii) and to define the next approximation as the zero of F_d^*.

Note that it is essential that F_d^* tends with all derivatives to an operator G of the class $\mathcal{F}$ (see (5.3)).

Theorems 8.1 and 8.2 lead to

Corollary 8.1

A $\varphi_{\mathfrak{N},\mathfrak{M}}$ method is maximal iff its order is equal to the order of information. ∎

9. EXAMPLES

We shall illustrate the above results for some examples of $\mathfrak{N}$ and $\mathfrak{M}$.

In Examples 9.1 to 9.3 we consider the scalar case, and we assume that $B_1 = B_2 = \mathbb{C}$, and $\mathfrak{M}$ is an arbitrary AA set.

Example 9.1

Let $\mathfrak{N}(x_n, \ldots, x_0; F) = \{F^{(k)}(x_{n-j}): k = 0, 1, \ldots, s; \; j = 0, 1, \ldots, n\}$. From Example 7.1 it follows that the order of information $p_{n,s} = p(\mathfrak{N}; \mathfrak{M})$ does not depend on $\mathfrak{M}$ and is equal to the unique positive zero of the polynomial

$$(9.1) \quad t^{n+1} - (s+1) \sum_{j=0}^{n} t^j.$$

Using Kahan's estimations of $p_{n,s}$ (see Traub [72]) we get

$$s+2 - \frac{s+1}{(s+2)^{n+1} - \frac{(s+1)(n+1)+1}{s+2}} \leq p_{n,s} < s+2 - \frac{s+1}{(s+2)^{n+1} - 1},$$

and

$$\lim_{n \to \infty} p_{n,s} = s+2.$$

Now we can assume that functions $\{F_d^*\}$ which defines the interpolatory method (see (8.4)) are the minimum degree interpolatory polynomials. Thus, F_d^* is given by

$$F_d^{*(k)}(x_{d-j}) = F^{(k)}(x_{d-j}), \quad k = 0, 1, \ldots, s; \; j = 0, 1, \ldots, n,$$

and the degree of F_d is at most $r = (n+1)(s+1) - 1$. The interpolatory method denoted now by $I_{n,s}$ is maximal due to Theorem 8.2. Similar results in certain classes of admissible iterative methods have been given by Traub [64] and Kung and Traub [73a] for $n = 0$ and $s \geq 1$, by Rissanen [71] for $n = 1$ and $s = 0$ by Wozniakowski [73] for any n, s such that $n+s \geq 1$.

Example 9.2

Let

$$(9.2) \quad \mathfrak{N}(x_n,\ldots,x_0;F) = \{\int_0^{x_j} \frac{(x_j-t)^i}{i!} F(t)dt, \, F^{(k)}(x_j): \quad i = 0,1,\ldots,m-1,$$

$$k = 0,1,\ldots,s-m, \, j = 0,1,\ldots,n\}$$

where $0 \leq m < (n+1)(s+1) - 1$.

We define

$$(9.3) \quad g(x) = \int_0^x \frac{(x-t)^{m-1}}{(m-1)!} F(t)dt.$$

It is easy to verify that

$$g^{(k)}(x) = \int_0^x \frac{(x-t)^{m-1-k}}{(m-1-k)!} F(t)dt \qquad \text{for } k = 0,1,\ldots,m-1$$

$$g^{(m)}(x) = F(x)$$

$$g^{(k)}(x) = F^{(k-m)}(x) \qquad \text{for } k = m+1,\ldots,s.$$

Thus, the problem $F(x) = 0$ is equivalent to the solution of the equation

$$(9.4) \quad g^{(m)}(x) = 0.$$

The information (9.2) means that for (9.4) we use

$$g^{(k)}(x_{n-j}) \qquad \text{for } k = 0,1,\ldots,s; \, j = 0,1,\ldots,n.$$

Note that $F'(\alpha) \neq 0$ means that $g^{(m+1)}(\alpha) \neq 0$.

Let

$$(9.5) \quad m = m_1(s+1) + m_2, \quad 0 \leq m_2 \leq s,$$

for nonnegative integers m_1, m_2.

Using a similar technique as in Example 7.1 one can prove that the order of information $p_{n,s}(m)$ doesn't depend on $\mathfrak{N}$ and is equal to the unique positive zero of the polynomial

(9.6) $\quad t^{n+1} - (s+1-m_2)t^{n-m_1} - (s+1) \displaystyle\sum_{j=0}^{n-m_1-1} t^j$.

It can be shown that

(9.7) $\quad p_{n,s}(m) < p_{n+1,s}(m)$ and $\displaystyle\lim_{n\to\infty} p_{n,s}(m) = p_s(m)$

where $p_s(m)$ is the unique positive zero of the polynomial

(9.8) $\quad t^{m_1+2} - t^{m_1+1} - (s+1-m_2)t - m_2 = 0,$

(see (9.1)). If $m \le s+1$ then

(9.9) $\quad p_s(m) = \dfrac{s+2-m + \sqrt{(s+2-m)^2 + 4m}}{2}$.

Conditions (9.7), (9.8) and (9.9) state, in a certain sense, the affirmative

answer on Brent, Winograd and Wolfe's [73, p. 340] conjecture.

The most important case is $m = 1$ or 2. ($m = 0$ was considered in Example

(9.1).) For $m = 1$, from (9.6) and (9.9) follows

$$\zeta_s \le p_{n,s}(1) < p_s(1) = \frac{s+1 + \sqrt{(s+1)^2 + 4}}{2}$$

where

$$\zeta_s = \begin{cases} p_{2,0}(1) & s = 0, \\ 2 & s = 1, \\ s & s \ge 2. \end{cases}$$

For $m = 2$ orders $p_{n,s}(2)$ lie in the interval (H_s, w_s) where

$$H_s = \begin{cases} p_{3,0}(2) & s = 0, \\ \sqrt{2} & s = 1, \\ (1+\sqrt{13})/2 & s = 2, \\ s-1 & s \ge 3, \end{cases}$$

and

$$w_s = \begin{cases} p_0(2) & s = 0 \\ \dfrac{s+\sqrt{s^2+8}}{2} & s \geq 1. \end{cases}$$

From (9.7) and (9.8) the asymptotic formulas for $p_s(m)$ follow:

1. Let m be fixed. Then

$$p_s(m) < p_{s+1}(m) \quad \text{and} \quad \lim_{s \to \infty} \frac{p_s(m)}{s+2} = 1,$$

2. Let s be fixed. Then

$$p_s(m+1) < p_s(m) \quad \text{and} \quad \lim_{m \to \infty} p_s(m) = 1.$$

The interpolatory method $I_{n,s}(m)$ can be defined as follows. Let g_d be an interpolatory polynomial of degree at most $r = (n+1)(s+1) - 1$ given by

$$g_d^{(k)}(x_{d-j}) = g^{(k)}(x_{d-j}), \quad k = 0,1,\ldots,s; \ j = 0,1,\ldots,n.$$

We put

$$F_d^*(x) = g_d^{(m)}(x)$$

and the next approximation is a zero of F_d^* which converges to α when $d \to +\infty$. (See Brent [73] where the case $s = 0$ and $n = m+1$ was considered in detail.)

Example 9.3

Let $n = 0$ and

$$(9.10) \quad \mathfrak{R}(x_0;F) = \{F(x_0),F'(x_0),\ldots,F^{(s)}(x_0), \int_{a_0}^{x_0} F(t)dt\}$$

where $s \geq 1$ and $a_0 = x_0 - \dfrac{s+3}{s+2}\dfrac{F(x_0)}{F'(x_0)}$.

One can prove that the order of information does not depend on $\mathfrak{M}$ and is equal to s+3. Note that using only the standard information $F^{(k)}(x_0)$, $k = 0,1,\ldots,s$, the order of this information is equal to s+1. Thus, the value of the integral increases the order of information by 2.

The particular choice of the lower limit a_0 is optimal in the following sense. If one considers information (9.10) with

$$a_0 = a_0(F(x_0),\ldots,F^{(s)}(x_0))$$

and tries to maximize the order of information, then it can be shown (Kacewicz [73]) that if a_0 is given by, for instance,

$$a_0 = x_0 - \frac{s+3}{s+2}\frac{F(x_0)}{F'(x_0)},$$

then we have the maximal order of information. ∎

We pass to the multivariate and operator cases.

Example 9.4

Let $\dim(B_1) = \dim(B_2) \geq 2$ and $\mathfrak{M}$ be any <u>unconditional</u> AA set. Let

$$\mathfrak{N}(x_n,\ldots,x_0;F) = \{F^{(k)}(x_{n-j}): \quad k = 0,1,\ldots,s; \; j = 0,1,\ldots,n\}.$$

Using the same proof as for Theorem 5 in Wozniakowski [73] one can show

$$p(\mathfrak{N};\mathfrak{M}) = s+1, \; \forall n.$$

Hence, the order of information does not depend on n. This implies that the additional information contained in the previous points of iteration cannot increase the order of information. ∎

Example 9.5

Now we consider a <u>conditional</u> AA set. Let $B_1 = B_2 = \mathbb{C}^n$, $n \geq 2$, and let

$$\mathfrak{M} = \{(x_n, x_{n-1}, \ldots, x_0): \left| \det\left(\frac{x_n - x_{n-1}}{\|x_n - x_{n-1}\|}, \ldots, \frac{x_1 - x_0}{\|x_1 - x_0\|}\right) \right| \geq c \, \|x_n - x_0\|^\zeta\}$$

where $\zeta \in [0,1]$ and c is a positive constant such that $c \in (0,1]$ if $\zeta = 0$. Let

$$\mathfrak{N}(x_n, \ldots, x_0; F) = \{F(x_{n-j}): j = 0, 1, \ldots, n\}.$$

From Theorem 3 in Jankowska [73] it follows that the order of information $p(\zeta) = p(\mathfrak{N}; \mathfrak{M})$ is equal to the unique positive zero of the polynomial

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

The order $p(\zeta)$ is a decreasing function of ζ. Hence, the best case is when $\zeta = 0$. One can show (Jankowska [73]) that

$$p(0) \in [1 + (\tfrac{\sqrt{5}-1}{2})^n, \; 1 + \tfrac{\sqrt{5}-1}{2}].$$

Note that $p(0) > 1$ although we use only the values of the function. (Note that here $\mathfrak{M}$ is a conditional AA set!)

The considered $\mathfrak{N}$ and $\mathfrak{M}$ are closely related to the multivariate secant method which is defined as follows. (See Barnes [65], Jankowska [73] and Ortega and Rheinboldt [70], p. 360 and ff.) Let

$$F_d(x) = A_d x + b_d$$

be an interpolating polynomial of F given by

$$F_d(x_{d-j}) = F(x_{d-j}), \quad j = 0, 1, \ldots, n.$$

The next approximation is the zero of F_d. The existence of F_d and its zero follow from the assumption that $(x_d, \ldots, x_{d-n}) \in \mathfrak{M}$. Due to Theorem 8.2 the secant method is maximal.

One can find other examples of iterative methods with conditional AA sets in Brent [72] and Wozniakowski [72].   ∎

ACKNOWLEDGMENT

REFERENCES

Barnes [65]   Barnes, J. G., An algorithm for solving non-
                                     linear equations based on the secant method.
                                     Computer Journal 8 (1965), 66-72.

Brent [72]   Brent, R. P., The computational complexity of
                                     iterative methods for systems of nonlinear
                                     equations. Proc. Complexity Symposium, Yorktown
                                     Heights, Plenum Press (1972).

Brent [73]   Brent, R. P., Algorithms for minimization with-
                                     out derivatives. Prentice Hall, 1973.

Brent, Winograd and Wolfe [73]   Brent, R. P., Winograd, S. and Wolfe, P.,
                                     Optimal iterative processes for root-finding.
                                     Numer. Math. 20, 1973, 327-341.

Jankowska [73]   Jankowska, J., Multivariate secant method. To
                                     appear.

Kacewicz [73]   Kacewicz, B., An integral-interpolatory itera-
                                     tive method for the solution of scalar equations.
                                     To appear.

Kung and Traub [73a]   Kung, H. T. and Traub, J. F., Optimal order
                                     of one-point and multipoint iteration. To
                                     appear in JACM. (Also available as a CMU
                                     Computer Science Department report.)

Kung and Traub [73b]   Kung, H. T. and Traub, J. F., Optimal order
                                     and efficiency for iterations with two evalua-
                                     tions. Report, Department of Computer Science,
                                     Carnegie-Mellon University, 1973.

Ortega and Rheinboldt [70]   Ortega, J. M. and Rheinboldt, W. C., Iterative
                                     solution of nonlinear equations in several vari-
                                     ables. Academic Press, 1970.

Rissanen [71]   Rissanen, J., On optimum root-finding algorithms.
                                     J. of Math. Anal. and Appl. 36, 1971, 220-225.

Traub [61]   Traub, J. F., On functional iteration and the
                                     functional iteration and the calculation of
                                     roots. Proceedings 16th National ACM Confer-
                                     ence (1961).

Traub [64]   Traub, J. F. Iterative methods for the solu-
                                     tion of equations, Prentice Hall, 1964.

Traub [72]

Traub, J. F., Computational complexity of iterative processes. SIAM J. Comput. 1, 1972, 167-179.

Wozniakowski [72]

Wozniakowski, H., On nonlinear iterative processes in numerical methods, doctoral thesis, University of Warsaw, 1972 (in Polish).

Wozniakowski [73]

Wozniakowski, H., Maximal stationary iterative methods for the solution of operator equations. To appear in SIAM J. Numer. Anal. (Also available as a CMU Computer Science Department report.)