1974

Generalized information and maximal order of iteration for operator equations

H Wozniakowski

Carnegie Mellon University

Follow this and additional works at: http://repository.cmu.edu/compsci
NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:
The copyright law of the United States (title 17, U.S. Code) governs the making
doing photocopies or other reproductions of copyrighted material. Any copying of this
document without permission of its author may be prohibited by law.
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

This research was supported in part by the National Science Foundation under Grant GJ32111 and the Office of Naval Research under Contract N0014-67-A-0314-0100, NR 044-422.
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 concepts 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 $s \geq 0$, by Rissanen [71] for $n = 1$ and $s = 0$, and by Wozniakowski [73] for any $n, s$ such that $n+s \geq 1$.

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, $\mathcal{M}$, and test sequences (Section 2).

The second is to use "generalized information", $\mathcal{R}$ (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 $\mathcal{M}$ - the set of admissible approximations and $\mathcal{R}$ - 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 $\mathcal{M}$ and $\mathcal{R}$ and seek methods which attain maximal order (Section 6). Specifically we prove that the maximal order is equal to the order of information $\mathcal{R}$ with respect to $\mathcal{M}$. We show that maximal order is achievable by the generalized interpolatory method which is defined in Section 8. The last section contains examples of $\mathcal{M}$, $\mathcal{R}$ 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

\[(2.1) \quad F(x) = 0,\]

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 \(\alpha\). 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 \(\mathcal{M}\) defined as follows.
Definition 2

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

(i) \( \mathcal{M} \subseteq B_{1}^{n+1} = B_{1} \times \ldots \times B_{1} \), \( \quad \text{n+1 times} \)

(ii) \( \forall \alpha \in B_{1}, \forall q \geq 1, \mathcal{M} \{ x_{i} \} \subseteq B_{1} \) such that

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

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

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

for \( j = 0, 1, \ldots, n-1 \).

The sequence \( \{ x_{i} \} \) is called a test sequence of qth order in \( \mathcal{M} \). 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 \( \alpha \) of \( B_{1} \) and for any \( q \geq 1 \), \( \mathcal{M} \) contains at least one test sequence of qth order which converges to \( \alpha \).

Example 2.1

\( \mathcal{M} = B_{1}^{n+1} \).

In this case we do not assume any special position of the successive approximations. Many previous papers on iterative processes deal with this \( \mathcal{M} \).
Example 2.2

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

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

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

Note that from the Hadamard inequality it follows that

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

Thus, $M$ contains points for which the Hadamard inequality holds from below.

It is easy to verify that $M$ 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 $M$ such that

\begin{equation}
(2.2) \ \forall \alpha \in B, \exists \Gamma > 0 \text{ such that } S(\alpha, \Gamma)^{n+1} \subset M.
\end{equation}

This means that any points $(x_d, x_{d-1}, \ldots, x_{d-n})$ close enough to $\alpha$ belong to $M$ and no assumptions of position are required. Therefore, if (2.2) holds, we say that $M$ is an unconditional AA set and otherwise $M$ 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 \((x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\) be approximations to \(\alpha\). 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 \(\mathcal{M}\) is called generalized information or simply information iff

(i) \(\mathcal{M} : D_M \to V\)

where \(D_M \subset B^{n+1}_1 \times \mathcal{F}\) and \(V\) is a certain space, and

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

\[
S(\alpha, \Gamma)^{n+1} \times \{F\} \subset D_M
\]

The second condition implies that for any points \((x_1, \ldots, x_0)\) sufficiently close to \(\alpha\), \(\mathcal{M}(x_1, \ldots, x_0; F)\) is well defined. Usually \(\mathcal{M}(x_1, \ldots, x_0; F)\) is defined to be the set of the values of \(F\) or its derivatives at the points \(x_1, \ldots, x_0\). (See the following example.)

**Example 3.1**

\(\mathcal{M}(x_1, \ldots, x_0; F) = \{F^{(k)}(x_j) : k = 0, 1, \ldots, s; j = 0, 1, \ldots, n\}\)

for a fixed \(s\).

More briefly, write \(\mathcal{M}_s\) for \(\mathcal{M}(x_1, \ldots, x_0; F)\). This information is called standard information and was considered by many authors (see, e.g., Traub [64]). The problem of finding maximal methods which use \(\mathcal{M}_s\) with respect to \(\mathcal{M} = B^{n+1}_1\) has been settled (see, e.g., Wozniakowski [73]).
Example 3.2

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

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

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

In this case $\mathcal{R}(x_0; F)$ 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 $\mathcal{M}$ and $\mathcal{N}$ a stationary iterative method is defined as follows. Suppose that $(x_d,...,x_{d-n}) \in \mathcal{M}$ are approximations sufficiently close to the solution $\alpha$. Hence the information $\mathcal{N}(x_d,...,x_{d-n}; F)$ is well defined. The next $y_d$ approximation is given by

\begin{equation}
(4.1) \quad y_d = \varphi_{\mathcal{M}, \mathcal{N}}(x_d; F)
\end{equation}

where

\begin{equation}
(4.2) \quad \varphi_{\mathcal{M}, \mathcal{N}}(x_d; F) = \varphi(x_d,...,x_{d-n}; \mathcal{N}(x_d,...,x_{d-n}; F))
\end{equation}

for an operator $\varphi$,

\begin{equation}
(4.3) \quad \varphi: D \varphi \to B_1, \quad D \varphi \subset \mathcal{M} \times V.
\end{equation}

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

\[ x_{d+n+1} = y_d, \quad x_{d+n-j} = x_{d-j}, \quad j = 0,1,...,n-1 \]

and perform the next iterative step with $(x_{d+n+1},...,x_{d+1})$. But if $(y_d,x_d,...,x_{d-n+1}) \notin \mathcal{M}$ one must define the next approximations $(x_{d+n+1},...,x_{d+1})$ 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 $\varphi_{\mathcal{M}, \mathcal{N}}$ or, more briefly, a $\varphi_{\mathcal{M}, \mathcal{N}}$ method.
5. ORDER OF STATIONARY ITERATIVE METHODS

Let \( \varphi_{\mathcal{M}} \) be a stationary iterative method. If \( \mathcal{M} \) is a conditional AA set then \( \varphi_{\mathcal{M}} \) 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 \( \mathcal{M} \). The definition of the order of \( \varphi_{\mathcal{M}} \) based on properties of generating sequences must fail and we therefore have to define an order of \( \varphi_{\mathcal{M}} \) differently. We shall use test sequences instead of generating sequences. That is, let \( \{x_d\} \) be a test sequence of qth order in the sense of Definition 2.

Let

\[
y_d = \varphi_{\mathcal{M}}(x_d; \mathcal{F})
\]

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

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

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

(5.2)

\[
\lim_{d \to \infty} F_d^{(k)}(\alpha) = G^{(k)}(\alpha)
\]

(5.3)

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

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,

\[
\mathcal{M}(x_d, \ldots, x_{d-n}; F) = \mathcal{M}(x_d, \ldots, x_{d-n}; F), \forall_d.
\]

(5.4)

Hence, \( \varphi_{\mathcal{M}}(x_d; F) = \varphi_{\mathcal{M}}(x_d; F) \) for any stationary method \( \varphi_{\mathcal{M}} \).

**Definition 5**

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

\[\blacksquare\]
We give an example of \( \{ F_d \} \).

**Example 5.1**

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

We put

\[
(5.5) \quad F_d(x) = F(x) + \sum_{j=0}^{n} (x - x_d - j)^{s+1}.
\]

One can verify that there exists \( \{ \alpha_d \} \) such that

\[
F_d(\alpha_d) = 0 \quad \text{and} \quad \alpha_d - \alpha = O(\sum_{j=0}^{n} (x_d - j)^{s+1}).
\]

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

From (5.5),

\[
F_d^{(k)}(x_d - j) = F^{(k)}(x_d - j), \quad k = 0,1,...,s; \quad j = 0,1,...,n.
\]

Hence, \( \{ F_d \} \) equals \( F \).

By (5.4) the operators \( F \) and \( F_d \) are same with respect to the information at the points \( x_d, x_d - 1, ..., x_d - n \). 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 \( \alpha \) but also to \( \alpha_d \) at the same time.

This discussion leads us to the following definition of order.

**Definition 6**

A number \( p = p(\alpha_d, \mathfrak{M}) \geq 1 \) is called an order of \( \alpha_d, \mathfrak{M} \) 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 \( y_d = c_{\min,\max}(x_d; F) \) are well defined for large \( d = i(n+1) \), \( i = 0, 1, \ldots \), and for sufficiently small \( \varepsilon > 0 \),

\[
\limsup_{d \to \infty} \frac{|y_d - \alpha_d|}{e_d} < +\infty
\]

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

\begin{equation}
(5.6)
\end{equation}

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

\[
\limsup_{d \to \infty} \frac{|y_d - \alpha_d|}{e_d^{(p+\varepsilon)n+1}} = +\infty, \forall \varepsilon > 0.
\]

\begin{equation}
(5.7)
\end{equation}

In Definition 6 we compare \( y_d \) with \( \alpha_d \). But if one sets \( F_d = F \) for any \( d \), then \( \alpha_d = \alpha \) and the distance between \( y_d \) and \( \alpha_d \) is just \( ||y_d - \alpha|| \).

It is easy to verify that if \( p \) exists, then \( p \) is unique. The new definition 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 \( \mathcal{M} \) is an unconditional AA set. But this subject is not the topic of this paper and will be deferred to a future paper.
6. MAXIMAL METHODS

Let \( \gamma(\mathcal{R}; \mathcal{M}) \) be the class of stationary iterative methods \( \mathcal{R}, \mathcal{M} \) with well-defined order \( p(\mathcal{R}, \mathcal{M}) \).

**Definition 7**

A \( \mathcal{R}, \mathcal{M} \) method is called maximal if it has order as high as possible, i.e.,

\[
p(\mathcal{R}, \mathcal{M}) = \sup_{\mathcal{R}, \mathcal{M} \in \gamma(\mathcal{R}; \mathcal{M})} p(\mathcal{R}, \mathcal{M}).
\]

The next part of the paper deals with maximal methods. For fixed information, we shall define an order of information, \( p(\mathcal{R}; \mathcal{M}) \), and then prove that the maximal order is equal to the order of information. Note that the class \( \gamma(\mathcal{R}; \mathcal{M}) \) contains all \( \mathcal{R}, \mathcal{M} \) 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 Woźniakowski [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 \( \alpha_d \) and \( \alpha \) are the zeros of the operators \( F_d \) and \( F \), respectively (see Definition 5).

Definition 7

A number \( p = p(\mathcal{R};\mathcal{M}) \geq 1 \) is called an order of information \( \mathcal{M} \) with respect to \( \mathcal{M} \) (shortly an order of information) iff

(i) \( \forall F \in \mathcal{F}, F(\alpha) = 0, \forall \{ x_1 \} \) of \( q \)th order and \( \forall \{ F_d \} \) equals \( F \),

\[
\limsup_{d \to +\infty} \frac{|\alpha_d - \alpha|}{e_d^n} < +\infty
\]

(7.1) holds for sufficiently small \( \varepsilon > 0 \) where

\[
u = \min(q + \varepsilon, p - \varepsilon) \]

(ii) \( \exists F \in \mathcal{F}, F(\alpha) = 0, \{ x_{1_d} \} \) of \( p \)th order and \( \{ F_d \} \) equals \( F \) such that

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

(7.2)

It can be shown that if the order of information exists, then it is unique. We assume that for the considered \( \mathcal{R} \) and \( \mathcal{M} \), the order of information \( p(\mathcal{R};\mathcal{M}) \) 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

\[
t^{n+1} - \sum_{j=0}^{n} p_j t^{n-j}.
\]

(7.3)
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,

\[
\limsup_{d \to +\infty} \sup_{j=0}^{n} \frac{||x_{d-j} - \alpha||}{p_j} < +\infty
\]

(7.4)

holds, and

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

\[
\limsup_{d \to +\infty} \sup_{j=0}^{n} \frac{||x_{d-j} - \alpha||}{p_j} > 0,
\]

(7.5)

then p is the order of information \( \mathfrak{M} \) with respect to \( \mathfrak{M} \).

\[\blacksquare\]

Proof

Since \( \{x_i\} \) is of qth order, then

\[
\prod_{j=0}^{n} ||x_{d-j} - \alpha||^{p_j} = 0(e_{d}^{j=0}), \quad e_{d} = ||x_{d-n} - \alpha||.
\]

(7.6)

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

\[
||x_d - \alpha|| = 0(e_d^{(q+\varepsilon)^{n+1}}).
\]

(7.7)

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

\[ \|a_d - a\| = 0(e_d^{n+1}). \]

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 \( p \)-th order which means that

\[ \prod_{j=0}^{n} \|x_{d-j} - a\|^{p_j} = c_d e_d^{n+1} \quad \text{where} \quad \lim \inf_{d \to \infty} c_d > 0. \]

From (7.5) follows

\[ \limsup_{d \to \infty} \frac{\|a_d - a\|^{n+1}}{e_d^{n+1}} > 0 \]

which proves (7.2) and completes the proof.

**Example 7.1**

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

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

From the remainder formula we get

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

where

\[ G_d(x) = \frac{\gamma(x)}{(r+1)!} [F - F_d]^{(r+1)}(\xi) \]

for \( |\gamma(x)| \leq 1, r = (n+1)(s+1) - 1 \) and \( \xi \in \text{conv}(x, x_d, \ldots, x_{d-n}) \).
Setting \( x = a \) in (7.9), we have

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

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

\[
(7.11) \quad t^{n+1} - (s+1) \sum_{j=0}^{n} t^j.
\]
8. MAIN RESULTS

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

Theorem 8.1

If \( \varphi_{\mathcal{M}} \in \Psi(\mathcal{M};\mathfrak{m}) \) then \( p(\varphi_{\mathcal{M}}) \leq p(\mathfrak{m};\mathfrak{m}) \).

Proof

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

\[
(8.1) \quad \limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d} = +\infty
\]

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

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

Setting \( F_d = F \) we get

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

Hence, from (8.2) and (8.3),

\[
\limsup_{d \to +\infty} \frac{||\alpha_d - \alpha||}{e_d} \leq \limsup_{d \to +\infty} \frac{||y_d - \alpha|| + ||y_d - \alpha_d||}{e_d} < +\infty
\]

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 $q$th 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, $F_d^* = F$.
We define a generalized interpolatory method $\mathcal{I}_{\mathcal{N},\mathcal{M}}$ (shortly an interpolatory
method $\mathcal{I}_{\mathcal{N},\mathcal{M}}$) by

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

where $\alpha_d^*$ 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 $\mathcal{N}$, i.e.,

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

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 $\mathcal{N}$ and $\mathcal{M}$. 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 $\mathcal{R}$, Traub [64], p.60
and ff.)

**Theorem 8.2**

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

$$
p(I_{\mathcal{N},\mathcal{M}}) = p(\mathcal{N},\mathcal{M})
$$
Proof

Recall that \( F \in F \), \( F(\alpha) = 0 \), \( \{x_1\} \) is of qth order and the next approximation \( y_d \) in the \( I_{p,n} \) method is equal to \( \alpha_d^* \), 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.,

\[
(8.6) \quad \limsup_{d \to \infty} \frac{\|y_d - \alpha_d^*\|}{e_d^{\nu}} < +\infty
\]

where as always \( e_d = \|x_{d-n} - \alpha\| \), \( u = [\min(q+\varepsilon,p-\varepsilon)]^{n+1} \) for \( p = p(\mathbb{N};\mathbb{R}) \).

From (7.1) follows

\[
\|\alpha_d^* - \alpha\| = O(e_d^{\nu}) \quad \text{and} \quad \|\alpha_d - \alpha\| = O(e_d^{\mu}).
\]

Hence

\[
\|y_d - \alpha_d^*\| \leq \|\alpha_d^* - \alpha\| + \|\alpha_d - \alpha\| = O(e_d^{\nu})
\]

and (8.6) holds.

Now we wish to show (5.7). Let \( F, \{x_1\} \) of pth order and \( \{F_d\} \) equal to \( F \) satisfy (7.2), i.e.,

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

Let

\[
(8.8) \quad C = \lim sup \frac{\|y_d - \alpha\|}{e_d^{\nu}}.
\]

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 \( F, \{x_1\} \) 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 $I_{\mathcal{M},\mathcal{M}'}$. Due to Theorem 8.1 the $I_{\mathcal{M},\mathcal{M}'}$ 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 $I_{\mathcal{M},\mathcal{M}}$ 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 $\mathbb{R}$ and $\mathbb{M}$.

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

Example 9.1

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

\[
\sum_{j=0}^{n} t^{n+1} - (s+1) \sum_{j=0}^{n} t^j.
\]

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

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

and

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

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

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

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

\[ (9.2) \quad \mathcal{H}(x_1, \ldots, x_j; F) = \left\{ \int_0^x \frac{(x-t)^i}{i!} F(t) dt, F^{(k)}(x_j): \ i = 0, 1, \ldots, m-1, \ k = 0, 1, \ldots, s-m, \ j = 0, 1, \ldots, n \right\} \]

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

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

\[ g^{(m)}(x) = F(x) \]

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

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

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

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

Let

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

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 \( m \) and is equal to the unique positive zero of the polynomial
\[(9.6) \quad t^{n+1} \cdot (s+1-m_2) t^{n-m_1} - (s+1) \sum_{j=0}^{n-m_1-1} t^j.\]

It can be shown that

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

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

\[(9.8) \quad t^{m_1+1} - t^{m+1} - (s+1-m_2) t - m_2 = 0,\]

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

\[(9.9) \quad p_s(m) = \frac{s+2-m + \sqrt{(s+2-m)^2 + 4m}}{2}.\]

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

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

where

\[
\zeta_s = \begin{cases} 
2p_{s,0}(1) & s = 0, \\
2 & s = 1, \\
s & s \geq 2.
\end{cases}
\]

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

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

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

1. Let $m$ be fixed. Then

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

2. Let $s$ be fixed. Then

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

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

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

We put

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

and the next approximation is a zero of $F_d^*$ which converges to $\alpha$ 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

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

where $s \geq 1$ and $a_0 = x_0 - \frac{s+3}{s+2} \frac{F(x_0)}{F'(x_0)}.$
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 unconditional AA set. Let

\[
\mathfrak{M}(x_n, \ldots, x_0; F) = \{F^{(k)}(x_{n-j}) \mid k = 0, 1, \ldots, s; j = 0, 1, \ldots, n\}.
\]

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

\[
p(\mathfrak{M}; \mathfrak{M}) = s+1, \quad \forall n.
\]

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 conditional AA set. Let $B_1 = B_2 = \mathbb{C}^n$, $n \geq 2$, and let

$$
\mathcal{M} = \{(x_n, x_{n-1}, \ldots, x_0) : \det \left( \begin{array}{cccc}
\frac{x_n - x_{n-1}}{\|x_n - x_{n-1}\|}, & \ldots, & \frac{x_1 - x_0}{\|x_1 - x_0\|} \end{array} \right) \geq c \|x_n - x_0\|^\zeta \}
$$

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

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

From Theorem 3 in Jankowska [73] it follows that the order of information $p(\zeta) = p(\mathcal{M}; \mathcal{M})$ 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 $\zeta$. Hence, the best case is when $\zeta = 0$. One can show (Jankowska [73]) that

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

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

The considered $\mathcal{M}$ and $\mathcal{M}$ 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

$$
F_d(x_{d-j}) = F(x_{d-j}), \; j = 0, 1, \ldots, n.
$$
The next approximation is the zero of $F_d$. The existence of $F_d$ and its zero follow from the assumption that $(x_d, \ldots, x_{d-n}) \in \mathbb{M}$. 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

I wish to thank H. T. Kung and J. F. Traub for their helpful discussions and assistance during the preparation of these results.
REFERENCES


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.)


Traub [72]


Wozniakowski [72]


Wozniakowski [73]