# MAXIMAL ORDER OF MULTIPOINT ITERATIONS USING n EVALUATIONS*

H. Woźniakowski
Department of Computer Science
Carnegie-Mellon University
(On leave from University of Warsaw)

ABSTRACT

This paper deals with multipoint iterations without memory for the solution of the nonlinear scalar equation $f^{(m)}(x) = 0$, $m \geq 0$. Let $p_n(m)$ be the maximal order of iterations which use n evaluations of the function or its derivatives per step. We prove the Kung and Traub conjecture $p_n(0) = 2^{n-1}$ for Hermitian information. We show $p_n(m+1) \geq p_n(m)$ and conjecture $p_n(m) \equiv 2^{n-1}$. The problem of the maximal order is connected with Birkhoff interpolation. Under a certain assumption we prove that the Pólya conditions are necessary for maximal order.

## 1. INTRODUCTION

We consider the problem of solving the nonlinear scalar equation $f^{(m)}(x) = 0$ where m is a nonnegative integer. We solve this problem by multipoint iterations without memory which use n evaluations of the function or its derivatives per step. For fixed n we seek an iteration of maximal order of convergence. This problem is connected with Birkhoff interpolation and can be expressed in terms of the incidence matrix $E_n^k = (e_{ij})$ where $e_{ij} = 1$ if $f^{(j)}(z_i)$ is computed and

*This work was supported in part by the Office of Naval Research under Contract N0014-67-0314-0010, NR 044-422 and by the National Science Foundation under Grant GJ32111.

$e_{ij} = 0$ otherwise; $z_i \neq z_j$, and $\sum_{i=1}^{k} \sum_{j=0}^{\infty} e_{ij} = n$. (Note that the problem of Birkhoff interpolation has been open for 70 years, see Sharma [72].)

Let $p_n(m)$ be the maximal order of multipoint iterations. For $m = 0$, Kung and Traub showed that $p_n(0) \geq 2^{n-1}$. We show that $p_n(m+1) \geq p_n(m)$ and conjecture $p_n(m) = 2^{n-1}$. For $m = 0$ we prove the Kung and Traub conjecture for Hermitian information, i.e., if $f^{(j)}(z_i)$ is computed, then $f^{(0)}(z_i), \ldots, f^{(j-1)}(z_i)$ are also computed. Under a certain assumption we prove that the Pólya conditions are necessary for the maximal order, i.e., the total number of $f, f', \ldots, f^{(j)}$ evaluations has to be at least $j+1$, $j = 0,1,\ldots,n-1$. We show also that $p_n(0) \leq n(n+1)^{n-1}$. Some special incidence matrices $E_n^k$ are considered and maximal orders of iterations based on $E_n^k$ are discussed.

## 2. THE n-EVALUATION PROBLEM

We consider the problem of solving the nonlinear scalar equation

$$(2.1) \quad f^{(m)}(x) = 0$$

where $f: D_F \subset \mathbb{C} \to \mathbb{C}$, $\mathbb{C}$ denotes the one dimensional complex space and $m$ is a nonnegative integer. We assume that there exists a simple zero $\alpha$ of $f^{(m)}$, $f^{(m)}(\alpha) = 0 \neq f^{(m+1)}(\alpha)$, and that $f$ is analytic in a neighborhood of $\alpha$. Let $\mathfrak{J}$ denote a class of such functions.

We solve (2.1) by stationary iteration and assume that $x_1$ is a sufficiently close approximation to $\alpha$. To get the next approximation $x_2$ to $\alpha$ we need some information on $f$. We assume that this information $\mathfrak{N} = \mathfrak{N}(x_1; f)$ is given by some values of the function and its derivatives at the points $z_i$ defined as follows. Let

$$
\begin{aligned}
z_1 &: f^{(j_1^1)}(z_1),\dots,f^{(j_{\mu_1}^k)}(z_1),\\
&\vdots\\
z_k &: f^{(j_1^k)}(z_k),\dots,f^{(j_{\mu_k}^k)}(z_k)
\end{aligned}
$$

denote points and numbers of derivatives which are computed where nonnegative integers $\{j_\mu^i\}$ satisfy the relations

$$j_\mu^i < j_{\mu+1}^i \quad \text{for } i=1,2,\dots,k \text{ and } \mu=1,2,\dots,\mu_i-1,$$

$$\mu_1 + \mu_2 + \dots + \mu_k = n.$$

Furthermore,

$$
\begin{aligned}
(2.2)\qquad & z_1 = x_1\\
& z_{i+1} = z_{i+1}(z_1,\dots,z_i,\ f^{(j_1^1)}(z_1),\dots,\ f^{(j_{\mu_1}^1)}(z_1),\dots,\\
&\qquad f^{(j_1^i)}(z_i),\dots,f^{(j_{\mu_i}^i)}(z_i)) \quad \text{for } i = 1,2,\dots,k,\\
& z_i \neq z_j \text{ for } x_1 \neq \alpha \text{ and } i \neq j,\ i,j = 1,2,\dots,k,\\
& x_2 = z_{k+1}.
\end{aligned}
$$

This means that every $z_{i+1}$ is the function of the previous information computed at $z_1,\dots,z_i$ and the next approximation $x_2 = z_{k+1}$ depends on n evaluations. Sometimes we shall use the notation $z_i = z_i(x_1)$ or $z_i = z_i(x_1,f)$ to stress the dependence on $x_1$ and f.

To simplify further notations we define <u>an incidence matrix</u> $E_n^k = (e_{ij})$ <u>of the information</u> $\mathfrak{N}$, $i = 1,2,\dots,k$ and $j = 0,1,\dots,$ as follows. Let

$$(2.3)\quad e_{ij} = \begin{cases} 1 & \text{if we compute } f^{(j)}(z_i) \\ 0 & \text{if we do not compute } f^{(j)}(z_i), \end{cases}$$

where

$$(2.4)\quad \sum_{j=0}^{\infty} e_{ij} > 0 \quad \text{for } i = 2,3,\ldots,k,$$

$$(2.5)\quad |E_n^k| = \sum_{i=1}^{k} \sum_{j=0}^{\infty} e_{ij} = n, \text{ (thus } k \le n+1).$$

The condition (2.4) means that at every point $z_i$, $i \ge 2$, we compute at least one derivative. (We consider f to be the zeroth derivative $f^{(0)}$.) However we do not, at this point, insist on any information being computed at $z_1 = x_1$. We show in Lemma 3.2 that $f^{(m)}$ must be evaluated at $x_1$. The condition (2.5) means that we use exactly n evaluations. Let

$$(2.6)\quad e_n^k = \{(i,j):\ e_{ij} = 1,\ i = 1,2,\ldots,k;\ j = 0,1,\ldots\}$$

Hence the information $\mathfrak{N}$ can then be defined in terms of the incidence matrix $E_n^k$ as follows:

$$(2.7)\quad \mathfrak{N} = \mathfrak{N}(x_1;f) = \{f^{(j)}(z_i) : (i,j) \in e_n^k\}.$$

The concept of an incidence matrix is used in Birkhoff interpolation, see Sharma [72]. We shall show some connections between the n evaluation problem and Birkhoff interpolation.

Having the information $\mathfrak{N}$ we define the next approximation $x_2$, $x_2 = z_{k+1}$, as $x_2 = \varphi(x_1;\ \mathfrak{N}(x_1;f))$ where $\varphi$ is a given function.

We call $\varphi$ <u>an iteration function</u> if for every $f \in \mathfrak{F}$, with $f^{(m)}(\alpha) = 0$ there exists $\delta > 0$ such that for any $x_1$, $|x_1-\alpha| \le \delta$, the sequence

(2.8a) $x_{d+1} = \varphi(x_d;\ \mathfrak{N}(x_d;f))$, $d = 1,2,\ldots$

is well-defined and

(2.8b) $\lim_{d\to\infty} x_d = \alpha$,

(2.8c) $\alpha = \varphi(\alpha,\ \mathfrak{N}(\alpha;f))$.

Such iterations are called k-point iteration without memory since they use exactly n new evaluations at k distinct points. If $k > 1$ they are called <u>multipoint iterations</u> (see Traub [61], [64], and Kung and Traub [74]). Let $\Phi$ be a class of iterations $\varphi$ with $k \geq 1$.

Since these iterations are stationary and without memory it is sufficient to define how $x_2$ is generated from $x_1$ and to measure the goodness of $\varphi$ by examining some properties of $x_2 - \alpha$ as $x_1$ tends to $\alpha$.

We want to find an iteration for which $x_2$ approximates $\alpha$ as closely as possible, i.e., we seek an iteration with the maximal order. In a previous paper (Wozniakowski [75]) we proved that if a set of iterations $\Phi$ is not empty then the maximal order of iteration is equal to the order of information. This gives us a powerful technique for proving maximal order. Let us briefly recall what we mean by orders of iteration and information.

We shall say <u>$\{\tilde{f}(\cdot;\ x_1)\}$ is equal to f with respect to $\mathfrak{N}$</u> (briefly denoted by $\tilde{f} \underset{\mathfrak{N}}{=} f$) iff

(i) $f,\ \tilde{f}(\cdot;\ x_1) \in \mathfrak{F}$,

(ii) $\tilde{f}^{(m)}(\tilde{\alpha};\ x_1) = 0$ and $f^{(m)}(\alpha) = 0$ where $\tilde{\alpha} = \tilde{\alpha}(x_1)$ and $\lim_{x_1\to\alpha} \tilde{\alpha}(x_1) = \alpha$,

(iii) $\lim_{x_1 \to \alpha} \tilde{f}^{(j)}(\alpha;\ x_1) = g^{(j)}(\alpha)$ where $g(\alpha) = 0$ and $g \in \mathfrak{F}$, $j = 0,1,\ldots$

(iv) $\mathfrak{N}(x_1;\ \tilde{f}) = \mathfrak{N}(x_1;\ f)$, i.e., $\tilde{f}^{(j)}(z_i;x_1) = f^{(j)}(z_i)$ for $(i,j) \in e_n^k$.

The first three conditions mean that $\tilde{f}(x;\ x_1)$ is sufficiently regular with respect to x and tends to a function g, $g \in \mathfrak{F}$, as $x_1$ tends to $\alpha$. The condition (iv) means that $\tilde{f}$ and f have the same information $\mathfrak{N}$ at the point $x_1$. Therefore any iteration $\varphi$ will produce the same approximation $x_2$ for $\tilde{f}$ and f, $\varphi(x_1;\ \mathfrak{N}(x_1;\ \tilde{f})) \equiv \varphi(x_1;\ \mathfrak{N}(x_1;\ f))$. Since we cannot recognize $\tilde{f}$ from f using information (2.7), we should approximate not only the zero $\alpha$ of f, but at the same time, the zero $\tilde{\alpha}$ of $\tilde{f}$. This leads us to the following definitions of orders of iteration and information.

Let A be a set defined by

$$A=\{q \geq 1;\ \forall f \in \mathfrak{F}, f^{(m)}(\alpha)=0,\ \forall \tilde{f} \underset{\mathfrak{N}}{\equiv} f,\ \limsup_{x_1 \to \alpha} \frac{|x_2-\tilde{\alpha}|}{|x_1-\alpha|^{q-\epsilon}} = 0,\ \forall \epsilon>0\}$$

A number $p = p(\varphi)$ is called <u>an order of the iteration $\varphi$</u> iff

$$(2.9)\quad p(\varphi) = \begin{cases} 0 & \text{if A is empty,} \\ \sup A & \text{otherwise.} \end{cases}$$

Using this convention $p(\varphi)$ always exists; however the only interesting cases are for $A \neq \emptyset$. Furthermore, let

$$B = \{q\geq 1;\ \forall f \in \mathfrak{F}, f^{(m)}(\alpha)=0, \forall \tilde{f} \underset{\mathfrak{N}}{\equiv} f, \limsup_{x_1 \to \alpha} \frac{|\alpha-\tilde{\alpha}|}{|x_1-\alpha|^{q-\epsilon}} = 0, \forall \epsilon>0\}.$$

A number $p = p(\mathfrak{N})$ (sometimes denoted $p = p(E_n^k)$) is called <u>an order of the information $\mathfrak{N}$</u> if

$$(2.10)\quad p(\mathfrak{N}) = \begin{cases} 0 & \text{if } B \text{ is empty,} \\ \sup B & \text{otherwise.} \end{cases}$$

We know that if $\Phi \neq \emptyset$ then

$$(2.11)\quad \sup_{\varphi \in \Phi} p(\varphi) = p(\mathfrak{N})$$

and $p(\mathfrak{N}) = p(I_{\mathfrak{N}})$ where $I_{\mathfrak{N}}$ is a generalized interpolatory method. (See Wozniakowski [75].)

We are now in a position to define the n-evaluation problem (see Kung and Traub [73] and [74]). For fixed n and m we wish to find a number $k = k(n,m)$, points $z_i = z_i(x_1)$ for $i = 2,3,\ldots,k$, an incidence matrix $E_n^k$, $|E_n^k| = n$, and an iteration $\varphi$ which uses $E_n^k$ (see (2.8)) such that $p(\varphi)$ is maximal. Due to (2.11) this is equivalent to maximizing the order of information $\mathfrak{N}$, i.e., to find $E{*}_n^k$ such that

$$(2.12)\quad p_n(m) = \sup_{E_n^k} p(E_n^k),$$

$$(2.13)\quad p(E{*}_n^k) = p_n(m).$$

We recall the <u>Kung and Traub conjecture</u> for $m = 0$ (Kung and Traub [74]):

$$(2.14)\quad p_n(0) = 2^{n-1}.$$

They showed two different matrices $E_n^k$, $n \geq 2$, for which the order of iteration is equal to $2^{n-1}$ (see Section 3), so we know that

$$(2.15)\quad p_n(0) \geq 2^{n-1}.$$

We now show a relationship among the $p_n(m)$ for different m.

Lemma 2.1

Let $\varphi = \varphi(\mathfrak{N})$ be an iteration of order p for the problem $f^{(m)}(x) = 0$ which uses n evaluations per step. Then there exists an iteration $\varphi^* = \varphi^*(\mathfrak{N}^*)$ for the problem $f^{(m+1)}(x) = 0$ which also uses n evaluations and has the same order p.

Proof

Let $E_n^k = (e_{ij})$ be the incidence matrix of $\mathfrak{N}$ and $E_n^{*k} = (e_{ij}^*)$ be defined by

$$e_{ij}^* = \begin{cases} 1 & \text{if } e_{i,j-1} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathfrak{N}^*$ be information with the incidence matrix $E_n^{*k}$ based on the points $z_i = z_i(x_1)$, $i = 2,\ldots,k$, from $\mathfrak{N}$. For any $f_1$ from $\mathfrak{F}$, $f_1^{(m+1)}(\alpha) = 0 \neq f_1^{(m+2)}(\alpha)$, define

$$f(x) = f_1'(x).$$

Thus, $f \in \mathfrak{F}$, $f^{(m)}(\alpha) = 0 \neq f^{(m+1)}(\alpha)$, and $f^{(j)}(x) \equiv f_1^{(j+1)}(x)$. Hence

$$\mathfrak{N}^*(x_1;\ f_1) = \mathfrak{N}(x_1;\ f).$$

Let us define $\varphi^*$ by

$$\varphi^*(x_1;\ \mathfrak{N}^*(x_1;\ f_1)) = \varphi(x_1,\ \mathfrak{N}(x_1;\ f)).$$

Since $f_1$ is arbitrary it easily follows that $p(\varphi^*) = p(\varphi)$. ■

From Lemma 2.1 and (2.15) we immediately get

Corollary 2.2

$p_n(m) \geq p_n(m-1) \geq 2^{n-1}$ for any $m \geq 1$. ■

Although Corollary 2.2 states that $p_n(m)$ is at least $p_n(m-1)$ we propose

Conjecture 2.3

$$p_n(m) = 2^{n-1} \qquad \forall\, m \geq 0,\ n \geq 1. \qquad \blacksquare$$

## 3. EXISTENCE OF ITERATIONS

Recall that $\Phi$ is a class of iterations defined by (2.8). In this section we show what we have to assume on the information $\mathfrak{N}$ to be sure that $\Phi$ is not empty. We shall prove that $\Phi = \emptyset$ if any of the following three conditions hold:

(1) If $z_i(x_1)$ does not converge to $\alpha$.

(2) If we do not compute $f^{(m)}(x_1)$, i.e., $e_{1m} = 0$.

(3) If $n = 1$ under the assumption on sufficiently regularity of $\varphi$ as a function of $x_1$.

We prove this in the following Lemmas.

Lemma 3.1

Let $\varphi$ be an iteration which uses the information $\mathfrak{N}$. Then for any $f \in \mathfrak{F}$, $f^{(m)}(\alpha) = 0$,

$$\lim_{x_1 \to \alpha} z_i(x_1;\ f) = \alpha \quad \text{for } i = 1,2,\ldots,k+1.$$

Proof

Suppose on the contrary that there exist $f \in \mathfrak{F}$, $f(\alpha) = 0$, an index $i$, $2 \leq i \leq k$, a number $\varepsilon > 0$ and a sequence $\{x_j\}$ such that

$$\lim_{j\to\infty} x_j = \alpha \quad \text{and} \quad |z_i(x_j)-\alpha| \geq \varepsilon \quad \text{for } j \geq j_0.$$

Let $J = \{x\colon\ |x-\alpha| < \varepsilon\}$. Define $f_1\colon J \to \mathbb{C}$ such that $f_1(x) = f(x)$ for $x \in J$. Since $f_1 \in \mathfrak{F}$ there exists $\delta_1 > 0$ such that any $x_1$, $|x_1-\alpha| \leq \delta_1$ is a good initial approximation.

Setting $x_1 = x_j$, for large j, where $|x_j-\alpha| \le \delta_1$, we get $z_i(x_j) \notin J$ and $\mathfrak{N}(x_1;\ f_1)$ is not well defined which contradicts (2.8a). ■

Lemma 3.2

Let $\mathfrak{N}$ be any information with the incidence matrix $E_n^k$. If $\Phi \neq \emptyset$ then $e_{1m} = 1$, (i.e. we have to compute $f^{(m)}(x_1)$).

Compare Theorem 4.1 in Kung and Traub [73] which proves this result for m = 0.

Proof

Let $\varphi \in \Phi$ and suppose on the contrary that $e_{1m} = 0$. Let f be any function from $\mathfrak{F}$, $f^{(m)}(\alpha) = 0$. Let $x_1$ be sufficiently close approximations to $\alpha$, $x_1 \neq \alpha$. From (2.2) we get $\delta = \min_{2\le i\le k} |z_i(x_1) - x_1| > 0$.

Define

$$f_1(x) = \begin{cases} f(x) - \dfrac{f^{(m)}(x_1)}{m!}(x-x_1)^m & \text{for } |x-x_1| < \delta \\ f(x) & \text{otherwise} \end{cases}$$

Note that $f_1 \in \mathfrak{F}$, $f_1^{(m)}(x_1) = 0$, and

$$f_1^{(j)}(x_1) = f^{(j)}(x_1) \qquad \text{for } j \neq m$$
$$f_1^{(j)}(z_i) = f^{(j)}(z_i) \qquad \text{for any j and } i = 2,\dots,k.$$

Since we do not compute $f^{(m)}(x_1)$ then

$$\mathfrak{N}(x_1;\ f_1) = \mathfrak{N}(x_1;\ f).$$

But $x_1$ is the zero of $f_1$ and due to (2.8c) it follows

$$x_2 = \varphi(x_1;\ \mathfrak{N}(x_1;f)) = \varphi(x_1;\ \mathfrak{N}(x_1;f_1)) = x_1.$$

Thus, $x_d \equiv x_1$ and $\lim_d x_d \neq \alpha$ which contradicts (2.8b). ■

An iteration function $\varphi$ can be treated as a function of x, $\varphi(x) = \varphi(x;\, \mathfrak{N}(x;\, f))$ for x close to $\alpha$. We shall prove that if $\varphi$ is sufficiently regular then the number of evaluations n has to be at least two.

Lemma 3.3

If an iteration $\varphi$ is a sufficiently smooth function of x then $n \geq 2$.

Proof

It is enough to prove Lemma 3.3 for the real case. Assume on the contrary that $n = 1$. From Lemma 3.2 it follows that this unique piece of information is given by $f^{(m)}(x_1)$. Let

$$\varphi(x;\, f^{(m)}(x)) = x + g(x,\, f^{(m)}(x)).$$

From (2.8b) it follows

$$g(\alpha,\, 0) = 0 \qquad \forall\, \alpha \text{ such that } f^{(m)}(\alpha) = 0,\ f \in \mathfrak{F}$$

From this and the regularity of $\varphi$ we can express $g(x, y)$

$$g(x,\, y) = y^k h(x,\, y)$$

for an integer $k \geq 1$ where $h(x, 0) \neq 0$ and $h(x) = h(x, f(x))$ is a continuous function for x close to $\alpha$.

Let $h(\alpha) \neq 0$ and for simplicity we assume that $h(\alpha) > 0$. (If $h(\alpha) < 0$ then the proof is analogous.) Let $f \in \mathfrak{F}$ be a polynomial of degree m+1 and $f^{(m+1)}(x) \equiv 1$, $f(\alpha) = 0$. There exists $\delta = \delta(f) > 0$ such that for any $x_1$, $|x_1 - \alpha| \leq_{\bar{k}} \delta$ the sequence $x_{d+1} = \varphi(x_d,\, f^{(m)}(x_d)) = x_d + \left[f^{(m)}(x_d)\right]^k h(x_d)$ is well defined for any d and converges to $\alpha$ (see (2.8)). For

$e_d = x_d - \alpha$ we get

$$(3.2)\quad e_{d+1} = [1 + e_d^{k-1}\ h(x_d)]\ e_d.$$

If $x_1$ is close but different from $\alpha$ then $e_d \neq 0$ for any $d$. Since $\lim e_d = 0$ then for any $d_1$ there exists $d \geq d_1$ such that $|e_{d+1}|^d < |e_d|$, i.e.

$$(3.3)\quad |1 + e_d^{k-1}\ h(x_d)| < 1.$$

We consider two cases.

Case I. Let $k$ be odd. Then for large $d$ we have

$$e_d^{k-1}\ h(x_d) \cong e_d^{k-1}\ h(\alpha) > 0$$

which contradicts (3.3).

Case II. Let $k$ be even. We prove that $h$ does not change sign for $x \in [\alpha-\delta, \alpha+\delta]$. If so, then by the continuity of $h$ there exists $x^*$ such that $h(x^*) = 0$ and $0 < |x^*-\alpha| < \delta$. Setting $x_1 = x^*$ we get $x_d \equiv x^*$ which contradicts (3.3). Thus $h(x) \geq h_0 > 0$ for $|x-\alpha| \leq \delta$. Define $f_1$: $[\alpha-\delta, \alpha+\delta] \rightarrow \mathfrak{R}$ such that $f_1(x) = f(x)$. Since $f_1$ also belongs to $\mathfrak{J}$, $f_1^{(m)}(\alpha) = 0$, there exists $\delta_1 > 0$ such that $x_{d+1} = \varphi(x_d;\ \mathfrak{N}(x_d;\ f_1))$ is well defined whenever $|x_1-\alpha| \leq \delta_1$. Let $x_1 > \alpha$. Keeping in mind that $\mathfrak{N}(x_d;\ f_1) \equiv \mathfrak{N}(x_d;\ f)$, from (3.2) we get

$$e_{d+1} \geq (1 + e_d^{k-1} h_0) e_d \geq (1 + e_1^{k-1} h_0)^d e_1.$$

Hence, there exists an index $d$ such that $e_{d+1} > \delta$, and since $f_1(x_{d+1})$ is not defined we get a contradiction with (2.8a). ■

## 4. HERMITIAN INFORMATION

In this section we deal with a special case of the n-evaluation problem when the information $\mathfrak{N}$ is hermitian.

Definition 4.1

$\mathfrak{N}$ is called hermitian information if the incidence matrix $E_n^k$ (which is now called hermitian) satisfies

$$e_{ij} = 1 \Rightarrow e_{i0} = e_{i1} = \dots = e_{i,j-1} = 1 \qquad \forall\ (i,j) \in e_n^k \qquad \blacksquare$$

This means that if $f^{(j)}(z_i)$ is computed then $f^{(0)}(z_i),\dots,$ $f^{(j-1)}(z_i)$ are also computed.

Let $s_i$ denote the number of evaluations at $z_i$, i.e., $e_{i,s_i-1} = 1$ and $e_{i,s_i} = 0$. Then

(4.1) $\quad s_1 + s_2 + \dots + s_k = n$ where $s_i \geq 1$ for $i = 1,2,\dots,k$.

For given n and k we want to find $s_i$ and $z_i$, $i = 1,2,\dots,k$, to maximize the order of information. Let $p_n(m,H)$ be the maximal order of hermitian information. Note that $p_n(m) \geq p_n(m,H)$.

First we shall discuss a property of hermitian informations for the problem $f(x) = 0$, i.e., $m = 0$.

Theorem 4.1 $\quad (m = 0)$

The order $p(E_n^k)$ of the hermitian information $\mathfrak{N}$ with the incidence matrix $E_n^k$ satisfies

(4.2) $$p(E_n^k) \leq s_1 \prod_{i=2}^{k} (s_i+1). \qquad \blacksquare$$

Proof

It is easy to verify that if $\tilde{f} \underset{\mathfrak{N}}{\equiv} f$ then

$$(4.3)\quad \tilde{f}(x;\ x_1) = f(x) + G(x;\ x_1) \prod_{i=1}^{k} (x-z_i)^{s_i}$$

for an analytic function G. Since $\tilde{f}'(\alpha;\ x_1)$ tends to $g'(\alpha) \neq 0$ then setting $x = \alpha$ in (4.3) we get

$$(4.4)\quad (\alpha-\tilde{\alpha}) = \frac{G(\alpha;\ x_1)}{g'(\alpha)}(1 + o(1)) \prod_{i=1}^{k} (\alpha-z_i)^{s_i}.$$

Define $q_i$ by

$$\frac{\alpha-z_i}{e_1^{q_i-\varepsilon}} \to 0 \quad \text{and} \quad \frac{\alpha-z_i}{e_1^{q_i+\varepsilon}} \to +\infty, \quad \forall\varepsilon > 0$$

where $e_1 \equiv x_1 - \alpha$. Since $z_i = z_i(x_1)$ tends to $\alpha$ (see Lemma 3.1) then $q_i$ exists and $q_i \geq 0$ for $i = 1,2,\ldots,k$. Note that $q_1 = 1$.

Let $p_1 = q_1 = 1$ and

$$(4.5)\quad p_{j+1} = \sum_{i=1}^{j} q_i s_i, \quad j = 1,2,\ldots,k.$$

From (4.4) we get

$$(4.6)\quad \frac{\alpha-\tilde{\alpha}}{e_1^{p_{k+1}-\varepsilon}} = \frac{G(\alpha;\ x_1)}{g'(\alpha)}(1 + o(1)) \prod_{i=1}^{k} \left\{\frac{\alpha-z_i}{e_1^{q_i-\delta}}\right\}^{s_i} \to 0,\ \forall\varepsilon>0,$$

where $\delta = \varepsilon/n$. For $G(\alpha;\ x_1) \equiv \text{const} \neq 0$ we get

$$(4.7)\quad \frac{\alpha-\tilde{\alpha}}{e_1^{p_{k+1}+\varepsilon}} \to \infty,\ \forall\varepsilon > 0.$$

Now we shall prove that there exists a function f such that

(4.8) $q_i \le p_i$ for $i = 1,2,\ldots,k$.

Let $f$ be any function such that $f \in \mathfrak{J}$, $f(\alpha) = 0$ and $f^{(j)}(\alpha) \neq 0$ for $j = 1,2,\ldots$ . Since $p_1 = q_1$, the condition (4.8) holds for $i = 1$. Assume by induction that this holds for $i \le j$. Suppose by the contrary that

$$q_{j+1} > p_{j+1} = \sum_{i=1}^{j} q_i s_i .$$

Define

$$r = \sum_{i=1}^{j} s_i .$$

<u>Case I</u>. Let $r = 1$. This means that $j = 1$, $s_1 = 1$ and $z_2 = z_2(x_1, f(x_1))$ approximates $\alpha$ with order greater than $p_2 = 1$.

Define

$$\text{(4.9)} \quad h(x_1, f(x_1)) = \frac{x_1 - f(x_1) - z_2}{z_2 - x_1} + 1.$$

It is easy to verify that

$$h(x_1, f(x_1)) = f'(\alpha)(1 + o(1)).$$

<u>Case II</u>. Let $r > 1$ and $\tilde{f}$ be the Hermite interpolatory polynomial of degree less than $r$ defined by

$$\tilde{f}^{(1)}(z_i) = f^{(1)}(z_i), \quad i = 1,2,\ldots,j; \; 1 = 0,1,\ldots,s_i-1.$$

Let $\tilde{\alpha}$ be the nearest zero of $\tilde{f}$ to $z_1 = x_1$. Then

$$(4.10)\quad \frac{\tilde{\alpha} - \alpha}{\prod_{i=1}^{j} (\alpha - z_i)^{s_i}} f'(\alpha) = \frac{f^{(r)}(\alpha)}{r!}(1 + o(1)).$$

Note that $\tilde{\alpha}$ is a function of $x_1$ and information $\mathfrak{N}(x_1; f) = \{f^{(l)}(z_i): i = 1,2,\ldots,j;\ l = 0,1,\ldots,s_i-1\}$. Recall that $z_{j+1} = z_{j+1}(x_1, \mathfrak{N}(x_1; f))$ and $z_{j+1} - \alpha = o(e_1^{\,p_{j+1}})$. Define

$$(4.11)\quad h(x_1, \mathfrak{N}(x_1; f)) = \frac{\tilde{\alpha} - z_{j+1}}{\prod_{i=1}^{j} (z_{j+1} - z_i)^{s_i}} \tilde{f}'(z_{j+1}).$$

Thus h is the lefthand side of (4.10) where $\alpha$ is replaced by $z_{j+1}$. Since $z_{j+1}$ is a better approximation to $\alpha$ than $\tilde{\alpha}$, it is straightforward to verify that

$$(4.12)\quad h(x_1, \mathfrak{N}(x_1; f)) = \frac{f^{(r)}(\alpha)}{r!}(1 + o(1)).$$

This means that in both cases using r evaluations of the function and its derivatives given by $\mathfrak{N}$ we can approximate the rth normalized derivative. We prove that this is impossible.

Note that h (see (4.9) or (4.11)) is a continuous function of $x_1$ at $x_1 = \alpha$ and

$$(4.13)\quad h(\alpha, \mathfrak{N}(\alpha; f)) = \frac{f^{(r)}(\alpha)}{r!}.$$

Let $f_1(x) = f(x) + (x-\alpha)^r$ and let us apply h to the function $f_1$. Thus

$$h(\alpha, \mathfrak{N}(\alpha; f)) = h(\alpha, \mathfrak{N}(\alpha; f_1)) = \frac{f^{(r)}(\alpha)}{r!} + 1$$

which contradicts (4.13).

Hence $q_{j+1} \le p_{j+1}$ which proves (4.8). Keeping in mind $p(E_n^k) = p_{k+1}$ and using (4.5), (4.8) we get

$$p(E_n^k) = \sum_{i=1}^{k} q_i s_i \le \sum_{i=1}^{k} p_i s_i = \sum_{i=1}^{k-1} p_i s_i + p_k s_k \le (1+s_k) \sum_{i=1}^{k-1} p_i s_i$$

$$\le s_1 \prod_{i=2}^{k} (s_i+1)$$

which proves Theorem 4.1. ■

We want to show that a bound in (4.2) is sharp, i.e., there exist points $z_2,\ldots,z_k$ such that the order of information is equal to $s_1 \prod_{i=1}^{k} (s_i+1)$.

Let $w_\mu$, $\mu = 1,2,\ldots,k$, be the Hermite interpolatory polynomial of degree less than $r_\mu = s_1 + s_2 + \ldots + s_\mu$ defined by

$$(4.14) \quad w_\mu^{(j)}(z_i) = f^{(j)}(z_i), \; i = 1,2,\ldots,\mu; \; j = 0,1,\ldots,s_i-1.$$

Let $\alpha_\mu$ be the nearest zero of $w_\mu$ to $z_1 = x_1$. (If $s_1 = 1$ then $\alpha_1 = x_1 - \beta f(x_1)$ for any nonzero constant $\beta$.)

Define $z_{\mu+1}$ as a point such that

$$(4.15) \quad z_{\mu+1} = \alpha_\mu + O(e_1^{\beta_\mu}), \; \beta_\mu \ge s_1 \prod_{i=2}^{\mu} (s_i+1).$$

From (4.14) it follows

$$(4.16) \quad \alpha_\mu - \alpha = \begin{cases} (\beta f'(\alpha)-1)(\alpha-z_1) + o(\alpha-z_1) & r_\mu = 1 \\ \dfrac{f^{(r_\mu)}(\alpha)}{r_\mu!\,f'(\alpha)} \prod_{i=1}^{\mu}(\alpha-z_i)^{s_i} + o\left(\prod_{i=1}^{\mu} (\alpha-z_i)^{s_i}\right) & \text{if } r_\mu>1. \end{cases}$$

From (4.15) we get

$$(4.17) \quad z_{\mu+1} - \alpha = O(e_1^{q_{\mu+1}}), \quad q_{\mu+1} = s_1 \prod_{i=2}^{\mu} (s_i+1),$$

which proves that the order of information $\mathfrak{N}$ based on the points $z_{\mu+1}$ from (4.15) is equal to $s_1 \prod_{i=2}^{k} (s_i+1)$.

An iteration which uses this information $\mathfrak{N}$ and has the maximal order can be defined as follows.

For $\mu = 1,2,\ldots,k$

(i) construct $w_\mu$ from (4.14) using a divided-difference algorithm,

(ii) apply Newton iteration to the equation $w_\mu(x) = 0$ setting

$$y_0 = z_\mu$$
$$y_{i+1} = y_i - w'_\mu(y_i)^{-1} w_\mu(y_i), \; i = 0,1,\ldots,i_0-1,$$
$$z_{\mu+1} = y_{i_0}$$

where

$$(4.18) \quad i_0 = \lceil \log_2(s_{\mu+1}+1) \rceil.$$

(If $s_1 = 1$ then $z_2 = x_1 - \beta f(x_1)$.)

Then (4.15) holds and

$$(4.19) \quad z_{k+1} - \alpha = O(e_1^{q_{k+1}}), \; q_{k+1} = s_1 \prod_{i=2}^{k} (s_i+1).$$

Furthermore if $\beta_\mu > q_{\mu+1}$ in (4.15) then we can specify the constant which appears in the "O" notation in (4.19). Note that $\beta_\mu > q_{\mu+1}$ if we redefine $i_0$ in (4.18) as the smallest integer such that $i_0 > \log_2(s_{\mu+1}+1)$.

Lemma 4.2

Let $\varphi$ be the iteration defined as above, $z_{k+1} = \varphi(x_1, \mathfrak{N}(x_1; f))$. If $\beta_\mu > q_{\mu+1}$ for $\mu = 1,2,\ldots,k$ then

$$(4.20) \quad \lim_{x_1 \to \alpha} \frac{z_{k+1}(x_1) - \alpha}{(x_1-\alpha)^{q_{k+1}}} = C_{k+1}$$

where

$$C_{\mu+1} = M_{r_\mu} \prod_{j=1}^{\mu-1} M_{r_j}^{s_{j+1}(s_{j+2}+1)\ldots(s_\mu+1)} \quad \text{for } \mu = 1,2,\ldots,k$$

and

$$M_i = \begin{cases} (-1)^i \dfrac{f^{(i)}(\alpha)}{i!f'(\alpha)} & \text{if } i > 1 \\ -\beta f'(\alpha) + 1 & \text{if } i = 1. \end{cases}$$

If

$$(4.21) \quad \underline{K}^{i-1} \le \left|\frac{f^{(i)}(\alpha)}{i!f'(\alpha)}\right| \le \bar{K}^{i-1} \quad \text{for } i = r_1, r_2, \ldots, r_k$$

then

$$(4.22) \quad c \cdot \underline{K}^{q_{k+1}-1} \le \lim_{x_1 \to \alpha} \left|\frac{z_{k+1}(x_1)-\alpha}{(x_1-\alpha)^{q_{k+1}}}\right| \le \bar{K}^{q_{k+1}-1} \cdot c$$

where

$$c = \begin{cases} 1 & \text{if } r_1 > 1 \\ |M_1|^{s_2(s_3+1)\ldots(s_k+1)} & \text{if } r_1 = 1 \text{ and } k \ge 2 \\ |M_1| & \text{if } r_1 = 1 \text{ and } k = 1 \end{cases} \qquad \blacksquare$$

Note that the righthand side of (4.21) follows from the analyticity of f.

Proof

Let $C_i = \lim_{x_1 \to \alpha} (z_i - \alpha)/(x_1 - \alpha)^{q_i}$. Note that $C_1 = 1$. From (4.15), (4.16) and since $\beta_\mu > q_{\mu+1}$ we get

$$z_{\mu+1} - \alpha = \alpha_\mu - \alpha + z_{\mu+1} - \alpha_\mu = M_{r_\mu} \prod_{i=1}^{\mu} (z_i - \alpha)^{s_i} + o(e_1^{q_{\mu+1}}).$$

Thus

$$(4.23) \quad C_{\mu+1} = M_{r_\mu} \prod_{i=1}^{\mu} C_i^{s_i}.$$

Since $C_1 = 1$ we get after some tedious calculations

$$C_{\mu+1} = M_{r_\mu} \prod_{j=1}^{\mu-1} M_{r_j}^{s_{i+1}(s_{i+2}+1)\ldots(s_\mu+1)}$$

which proves the first part of Lemma 4.2.

Let $r_1 > 1$. Assume by induction that $\underline{K}^{q_i-1} \le |C_i| \le \bar{K}^{q_i-1}$. This is true for $i = 1$ since $C_1 = q_1 = 1$. From (4.23) and (4.21) we have

$$|C_{\mu+1}| \le \bar{K}^{r_\mu - 1 + s_1(q_1-1) + \ldots + s_\mu(q_\mu - 1)} = \bar{K}^{q_{\mu+1}-1}$$

and similarly we get a lower bound.

Let $r_1 = 1$. Assume by induction that $c_i \underline{K}^{q_i-1} \le |C_i| \le \bar{K}^{q_i-1} c_i$ where $c_1 = 1$, $c_2 = |M_1|$ and $c_i = |M_1|^{s_2(s_3+1)\ldots(s_{i-1}+1)}$ for $i \ge 3$. This is true for $i = 1$ and $2$ since $C_1 = q_1 = q_2 = 1$ and $C_2 = M_{r_1}$. Then

$$|C_{\mu+1}| \le \bar{K}^{q_{\mu+1}-1} |M_1|^{s_2 + s_2 s_3 + s_4 s_2(s_3+1)\ldots s_\mu s_2(s_3+1)\ldots(s_{\mu-1}+1)} =$$
$$= \bar{K}^{q_{\mu+1}-1} c_{\mu+1}$$

and similarly we get a lower bound. Hence (4.22) holds which completes the proof. ∎

Lemma 4.2 in the case $r_1 > 1$ states that the asymptotic constant $C_{k+1}$ depends exponentially on the order $q_{k+1}$. This property makes an analysis of the complexity of iteration easier (Traub and Wozniakowski will analyze it in a future paper).

We are now in a position to answer the following question. For given n and k, $k \le n$, find nonnegative integers $s_1, s_2, \ldots, s_k$ to maximize the order of information $p_k = \max_{s_1+\ldots+s_k=n} s_1 \prod_{i=2}^{k} (s_i+1)$. Using a standard technique it is easy to verify that

$$(4.24) \quad \left(n + (k-1)\left\lceil\frac{n-1}{k}\right\rceil\right)\left(1 + \left\lceil\frac{n-1}{k}\right\rceil\right)^{k-1} \le p_k \le \left(\frac{n+k-1}{k}\right)^k < 2^{n-1}$$

for $k \le n-2$ and $p_k = 2^{n-1}$ for $k = n-1$ or n. If k is a divisor of n-1 then the optimal $s_i$ are given by

$$s_1 = 1 + \frac{n-1}{k} \text{ and } s_i = \frac{n-1}{k} \text{ for } i = 2,\ldots,k.$$

For $k = n$ the optimal $s_i \equiv 1$. Furthermore from Theorem 7.1 in Kung and Traub [74] it follows that there are exactly two cases which maximize the order of information,

$$k = n-1,\ s_1 = 2,\ s_i = 1 \quad \text{for } i = 2,\ldots,n,\ p_{n-1} = 2^{n-1}$$
$$k = n,\ s_i = 1 \quad \text{for } i = 1,\ldots,n,\ p_n = 2^{n-1}.$$

The first case means that we use f and f' at the first point and f at the other points. The second case states that we use n function evaluations. From Theorem 4.1 and (4.24) we get

Corollary 4.3

The Kung and Traub conjecture holds for hermitian information ($p_n(0,H) = 2^{n-1}$). ■

The next part of this section deals with the general problem $f^{(m)}(x) = 0$, $m \ge 1$. It seems to us that hermitian information is not always relevant for that problem especially for large m. Note that we have to compute $f^{(m)}(x_1)$ and if the information is hermitian then we have to assume $n \ge m+1$. On the other hand if we use $f^{(m)}(z_1),\dots,f^{(m)}(z_n)$ (which is nonhermitian) then the order of information is $2^{n-1}$. However it is interesting to know the optimal order of information for special hermitian cases, e.g., f, f' at $z_1$ followed by n-1 function evaluation at the other points for the problem $f'(x) = 0$, (see Lemma 4.5).

Recall that $p_n(m,H)$ denotes the maximal order of hermitian information. In general we do not know $p_n(m,H)$. We only show some bounds on it.

Lemma 4.4

$$p_n(m,H) \le 2^{n-1}.$$

Proof

If $\tilde{f} \underset{\mathfrak{N}}{\equiv} f$ then

$$(4.25) \quad \tilde{f}^{(m)}(x) - f^{(m)}(x) = [G(x)\prod_{i=1}^{k}(x-z_i)^{s_i}]^{(m)}$$

for an analytic function G. Let $G(x) = \frac{1}{m!}(x-\alpha)^m$. Since $\tilde{f}^{(m+1)}(\alpha)$ tends to $g^{(m+1)}(\alpha) \ne 0$ as $x_1$ tends to $\alpha$ then setting $x = \alpha$ in (4.25) we have

$$\tilde{\alpha}-\alpha = c(\alpha,x_1)\prod_{i=1}^{k}(\alpha-z_i)^{s_i}$$

where $c(\alpha,x_1)$ tends to a nonzero limit (see (4.4)).

The proof of Lemma 4.4 may now be obtained analogously to the proof of Theorem 4.1.

Lemma 4.5

Let $n \geq m+1 \geq 2$. Then

$$p_n(m,H) \geq c\, q(m)^{n-1}$$

where

$$c = c(m) = \frac{2}{(1+2m+\sqrt{t})}, \quad q(m) = \left(\frac{1+\sqrt{t}}{2}\right)^{\frac{1}{m}}$$

and $t = 1 + 4m$.

Proof

Define $s_1 = m+1$ and $s_i = m$ for $i = 2,\ldots,k$. Let $z_2 = x_1 + \beta f^{(m)}(x_1)$ for $\beta \neq 0$ and let $z_\mu$, $\mu \geq 3$, be the nearest zero to $z_{\mu-1}$ of the polynomial $w_\mu^{(m)}$ where

$$w_\mu^{(j)}(z_i) = f^{(j)}(z_i), \qquad i = 1,2,\ldots,\mu-1; \quad j = 0,1,\ldots,m-1,$$

$$w_\mu^{(m)}(z_1) = f^{(m)}(z_1)$$

and $w_\mu$ is of degree $\leq (\mu-1)m$. It is straightforward to verify that

$$z_\mu - \alpha = O((x_1-\alpha)^{q_\mu})$$

where $q_1 = q_2 = 1$ and for $\mu \geq 3$,

$$q_\mu = m(q_1+\ldots+q_{\mu-2}) + q_1 = q_{\mu-1} + mq_{\mu-2}.$$

It is easy to verify that

$$q_{k+1} \geq c\left(\frac{1+\sqrt{t}}{2}\right)^{k+1}$$

where $c = c(m) = 2/(1 + 2m + \sqrt{t})$.

For a given n let $k = \lfloor (n-1)/m \rfloor = \frac{n-1}{m} + \theta$ where $-1 < \theta \le 0$. The total number of evaluation is equal to $km + 1 \le n$. Hence $p_n(m,H) \ge p_{km+1}(m,H) \ge q_{k+1}^{1+\theta} \ge$

$\ge q_{k+1} \ge c\, q(m)^{n-1} \left(\frac{1+\sqrt{t}}{2}\right) \ge c\, q(m)^{n-1}$ which proves Lemma 4.5. ■

Lemma 4.4 and 4.5 state that $p_n(m,H)$ as a function of n is exponentially bounded from below and above. However $\lim_{m\to\infty} q(m) = 1$.

## 5. GENERAL INFORMATION, m = 0

We deal with the n-evaluation problem for m = 0. For small n it is possible to verify the Kung and Traub conjecture and to characterize the information sets for all iterations which have maximal order.

For n = 1 the unique piece of information is given by $f(x_1)$. Since $\tilde{f}(x) = f(x) + (x-x_1)$ has the same information as f then $p_1(0) = 1$. This means that for any $y = y(x_1, f(x_1))$ the distance $\alpha-y$ can be at most of first order in $\alpha-x_1$. However y is not, in general, an iteration function, see Lemma 3.3. Note also that for any m, $p_1(m) = 1$.

For n = 2, Kung and Traub [73] proved that the maximal order of iteration equals two under a certain assumption on the iterations considered. Using our technique we find the order of information for any $\mathfrak{N}$ with n = 2. Note that if $\mathfrak{N}$ is hermitian information then $p(\mathfrak{N}) \le 2$, by Corollary 4.3. Thus it suffices to consider the non-hermitian case. Let us first consider one-point iterations, i.e., k = 1 and $\mathfrak{N} = \{f(x_1), f^{(j)}(x_1)\}$ for $j \ge 2$. Then $\tilde{f}(x) = f(x) + (x-x_1)$ and $p(\mathfrak{N}) = 1$. Let us pass to two-point iterations, i.e., k = 2 and $\mathfrak{N} = \{f(x_1), f^{(j)}(z_2)\}$ where $j \ge 1$ and

$z_2 = z_2(x_1, f(x_1))$. If $j \geq 2$ then $\tilde{f}(x) = f(x) + (x-x_1)$ and $p(\mathfrak{N}) = 1$. Let $j = 1$. Then $\tilde{f}(x) = f(x) + (x-x_1)(x-2z_2+x_1)$. From this we get

$$\tilde{\alpha}-\alpha \cong (\alpha-x_1)(\alpha-y), \quad y = 2z_2-x_1.$$

Since $y = y(x_1, f(x_1))$ then $\alpha-y$ can be at most of first order in $(\alpha-x_1)$. Hence $p(\mathfrak{N}) \leq 2$ and $p(\mathfrak{N}) = 2$ if, for instance, $z_2 = x_1 + \beta f(x_1)$, for any constant $\beta \neq 0$.

It is easy to verify that, in addition, $p_2(m) = 2$ for any m.

For $n = 3$, $p_3(0) = 4$. There are a number of information sets $\mathfrak{N}$ for which $p(\mathfrak{N}) = 4$. A proof and discussion may be found in Meersman [75].

Unfortunately the proof technique used to establish the cases $n = 2, 3$ cannot be used for general n since there are too many sub-cases to investigate.

We now wish to discuss some general properties of the n-evaluation problem.

Recall that $E_n^k = (e_{ij})$ is the incidence matrix of the information $\mathfrak{N}$ and let

$$(5.1) \quad M_r = \sum_{j=0}^{r} \sum_{i=1}^{k} e_{ij}$$

denote the total number of evaluations $f, f', \ldots, f^{(r)}$ at $z_1, \ldots, z_k$, $r = 0, 1, \ldots$ .

The incidence matrix $E_n^k$ satisfies the <u>Pólya conditions</u> if

$$(5.2) \quad M_r \geq r+1 \qquad \text{for } r = 0, 1, \ldots, n-1.$$

(See Sharma [72].) If $E_n^k$ satisfies the Pólya conditions then $e_{ij} = 0$ for any i and $j \geq n$. This means we do not use derivatives of order higher than n-1. Note that hermitian $E_n^k$ satisfies the Pólya conditions. Furthermore all known information sets with maximal order of information have $E_n^k$ which satisfy the Pólya conditions.

Let $j' = j'(E_n^k)$ be a nonnegative integer such that

$$M_r \geq r+1 \text{ for } r = 0,1,\ldots,j' \text{ and } M_{j'+1} < j'+2.$$

Since $j'+1 \leq M_{j'} \leq M_{j'+1} \leq j'+1$ then $e_{i,j'+1} = 0$ which means that we do not use the (j'+1) derivative. We shall call such $j' = j'(E_n^k)$ an index of $E_n^k$. $E_n^k$ satisfies the Pólya conditions if and only if its index is equal to n-1.

We introduce the concept of the polynomial order of information pol($\mathfrak{N}$) defined by

$$(5.3) \quad \mathrm{pol}(\mathfrak{N}) = \begin{cases} 0 & \text{if } B \text{ is empty} \\ \sup B & \text{otherwise} \end{cases}$$

where

$$B = \{q \geq 1: \forall f \in \mathfrak{F},\ f(\alpha) = 0,\ \forall \tilde{f} \underset{\mathfrak{N}}{=} f \text{ and } \tilde{f}-f \in \Pi_n,$$

$$\limsup_{x_1 \to \alpha} \frac{|\alpha-\tilde{\alpha}|}{|x_1-\alpha|^{q-\epsilon}} = 0,\ \forall \epsilon > 0\},$$

.

and $\Pi_n$ denotes a class of polynomials of degree $\leq$ n. Compare with the order of information where is not assumed that $\tilde{f}-f \in \Pi_n$, see (2.10). Thus $p(\mathfrak{N}) \leq \mathrm{pol}(\mathfrak{N})$. Similarly let $\mathrm{pol}(n) = \sup_{\mathfrak{N}} p(\mathfrak{N})$. This gives

$$(5.4) \quad p_n(0) \leq \mathrm{pol}(n).$$

We show some properties of pol(n). From Section 4 it follows that $\mathrm{pol}(n) \geq 2^{n-1}$ and $\mathrm{pol}(n) = 2^{n-1}$ for hermitian information. Furthermore it is possible to show that $pol(n) = 2^{n-1}$ for $n = 1,2,3$ and that $pol(n)$ is an increasing function of $n$.

Lemma 5.1

Let $j'$ be the index of the incidence matrix $E_n^k$ of $\mathfrak{N}$. Then

$$pol(\mathfrak{N}) \le pol(j'+1).$$

Proof (Compare with the proof of the Schoenberg Lemma in Schoenberg [66] and Sharma [72], Lemma 1.)

Let $E_{j'}^k$ denote the first $(j'+1)$ columns of $E_n^k$. Assume $f \in \Pi_{j'+1}$. Then $z_i = z_i(x_1; \mathfrak{N}(x_1; f)) = z_i(x_1; \mathfrak{N}_1(x_1; f))$ where $\mathfrak{N}_1$ is the information based on $E_{j'}^k$. Let $h \in \Pi_{j'+1}$ and

$$(5.5) \quad h^{(j)}(z_i) = 0 \quad \text{for } (i,j) \in e_n^k \text{ and } j \le j'.$$

The total number of homogeneous equations in (5.5) is equal to $M_{j'} = j'+1$ and since we have $j'+2$ unknowns then there exists a nonzero $h$ satisfying (5.5). Furthermore $h^{(j)}(x) \equiv 0$ for $j \ge j'+2$ which means that $h^{(j)}(z_i) = 0$ for all $(i,j) \in e_n^k$. Define $\tilde{f}(x) = f(x) + h(x)$ we get

$$(5.6) \quad \tilde{\alpha}-\alpha = \frac{1}{g'(\alpha)}\,(1 + o(1))h(\alpha).$$

But $h(\alpha)$ depends only on $E_{j'}^k$ and it can be at most of order $pol(j'+1)$. This proves that $pol(\mathfrak{N}) \le pol(j'+1)$. ■

Since $pol(n)$ is an increasing function of $n$ we immediately have

Corollary 5.2

A necessary condition for $\mathfrak{N}$ to have the maximal polynomial order $pol(n)$ is that its incidence matrix $E_n^k$ satisfies the Pólya conditions. ■

We believe that $pol(n) = 2^{n-1}$. However to find even a crude upper bound on pol(n) seems to be hard. We give an upper bound on pol(n) under the following conjecture.

Conjecture 5.3

Let $\varphi_1, \varphi_2, \ldots, \varphi_n$ be any n-point iterations. Then there exists a function $f \in \mathfrak{F}$ such that

$$(5.7) \quad \lim_{x_1 \to \alpha} \left| \frac{\varphi_i(x_1;\ \mathfrak{N}(x_1;\ f)) - \alpha}{e_1^{\,pol(n)+\varepsilon}} \right| = +\infty, \ \forall \varepsilon > 0, \ \forall i \le n. \quad ■$$

Assume for simplicity that $C_i = C_i(f, \varphi_i) = \lim_{x_1 \to \alpha} |\{\varphi_i(x_1;\ \mathfrak{N}(x_1;\ f)) - \alpha\}/e_1^{\,pol(n)}|$ exist for $i = 1,2,\ldots,n$. The conjecture 5.3 states that they are all different from zero for one function. Note that it holds for $n = 1$.

Lemma 5.4

If (5.7) holds then $pol(n) < n!$ for $n \ge 3$.

Proof

Let $E_n^k$ be the incidence matrix of $\mathfrak{N}$. Let $0 \not\equiv h \in \Pi_n$ and $h^{(j)}(z_i) = 0$ for $(i,j) \in e_n^k$. Then

$$h(x;x_1) = a(x_1)(x-h_1)(x-h_2)\ldots(x-h_j)$$

where $1 \le j \le n$ and $a(x_1)$ is chosen in order to ensure that $h(x;x_1)$ tends to an analytic function as $x_1$ tends to $\alpha$. Note that $h_1 = x_1$ and $h_i = h_i(z_1, z_2, \ldots, z_k)$ depends on at most (n-1) evaluations. If $\lim_{x_1 \to \alpha} h_i = \alpha$ then $h_i$ can be treated as an iteration. From (5.7) we get

$$|h_i - \alpha| \ge c|e_1|^{pol(n-1)+1-\varepsilon}, \ c > 0,$$

for any $\varepsilon > 0$. Since it holds for any $\mathfrak{N}$ we have

$$\mathrm{pol}(n) \le (n-1)\,\mathrm{pol}(n-1) + 1 < n\,\mathrm{pol}(n-1) \le n! \quad \blacksquare$$

The next part of this section deals with a restrictive class of n-point iterations. We use n evaluations per step and we assume that an iteration is exact for a function $f \in \Pi_{n-1}$. We shall say that $\varphi \in \Phi_n$ if $\varphi(x_1; \mathfrak{N}(x_1; f)) = \alpha$ whenever $f \in \Pi_{n-1}$ and $x_1$ is close to $\alpha$. Note that all iterations considered in Section 4 belong to $\Phi_n$.

Next we shall say that the problem is locally well-poised for f if for every $h \in \Pi_{n-1}$ such that

$$h^{(j)}(z_i) = 0 \quad \text{for} \quad (i,j) \in e_n^k$$

it follows $h \equiv 0$ for all $x_1$ close to x.

Note that Birkhoff interpolation for $E_n^k$ is well-poised if $\forall(x_1, x_2, \ldots, x_k)$ $h^{(j)}(z_i) = 0$ for $(i,j) \in e_n^k$ and $h \in \Pi_{n-1} \Rightarrow h \equiv 0$ (see Sharma [72]). Thus, if Birkhoff interpolation is well-poised than the problem is locally well-poised but not in general vice versa.

Lemma 5.5

If an iteration $\varphi$ is exact for $f \in \Pi_{n-1}$, $\varphi \in \Phi_n$, then

(i) $E_n^k$ satisfies the Pólya conditions,

(ii) the problem is locally well-poised for $f \in \Pi_{n-1}$,

(iii) $p(\mathfrak{N}) \le n(n+1)^{n-1}$.

Proof

Suppose that the problem is not locally well-poised for $f \in \Pi_{n-1}$. Then there exists a nonzero $h \in \Pi_{n-1}$ such that

$h^{(j)}(z_i) = 0$ for $(i,j) \in e_n^k$. Define $\tilde{f}(x) = f(x) + h(x)$. Since $\tilde{f} \in \Pi_{n-1}$ and $\tilde{f}(\alpha) \neq 0$ then

$$\alpha = \varphi(x_1, \mathfrak{N}(x_1, f)) = \varphi(x_1, \mathfrak{N}(x_1, \tilde{f})) \neq \tilde{\alpha}.$$

This contradicts that $\varphi \in \Phi_n$. Hence (ii) holds. Let j' be the index of $E_n^k$. If $j' < n-1$ then there exists a nonzero $h \in \Pi_{j'+1}$ such that $h^{(j)}(z_i) = 0$ for all $(i,j) \in e_n^k$, see the proof of Lemma (5.1). This contradicts that the problem is locally well-poised. Thus, (i) holds.

To prove (iii) it suffices to note that if

$$E_n^k \le \tilde{E}_{\tilde{n}}^k \quad \text{then} \quad p(E_n^k) \le p(\tilde{E}_{\tilde{n}}^k)$$

for $n \le \tilde{n}$ where by $E_n^k = (e_{ij}) \le \tilde{E}_{\tilde{n}}^k = (\tilde{e}_{ij})$ we mean $e_{ij} \le \tilde{e}_{ij}$ for $(i,j) \in e_n^k$.

Define $\tilde{E}_{\tilde{n}}^k$ as a hermitian matrix where $\tilde{n} = kn$,

$$\tilde{e}_{ij} = 1 \quad \text{for } i = 1,2,\ldots,k \text{ and } j = 0,1,\ldots,n-1.$$

Of course $E_n^k \le \tilde{E}_{\tilde{n}}^k$ and from Theorem 4.1 we get

$$p(\tilde{E}_{\tilde{n}}^k) \le n(n+1)^{n-1}$$

which proves (iii). ∎

## 6. FINAL REMARKS

The problem of the maximal order of n-point iterations is connected with Birkhoff interpolation which has been open almost 70 years. The main difficulty is to estimate the difference between the zeros, $\tilde{\alpha}-\alpha$, of any two functions with the same information, $\tilde{f} \underset{\mathfrak{N}}{\equiv} f$. Note that $\tilde{f}$ can belong to $\Pi_{n-1}$ for all f if the problem is well-poised. However up to now we do not know when Birkhoff interpolation is well-poised. There are many reasons to believe that hermitian information (interpolation without gaps) is optimal. However there also exists nonhermitian information with order $2^{n-1}$.

For nonhermitian information $\mathfrak{N}$ it is hard to find the order $p(\mathfrak{N})$. We know the order of such information only in a few cases. The first one is a Brent iteration based on $\mathfrak{N} = \{f(z_1), f'(z_1), \ldots, f^{(j)}(z_1), f^{(r)}(z_2), f^{(r)}(z_3), \ldots, f^{(r)}(z_k)\}$ for suitable chosen $z_i$ where $0 < r \le j+1$ (see Brent [75]). This information uses $n = j+k$ evaluations and has the order $p(\mathfrak{N}) = j + 2k - 1$, see Meersman [75]. Note that this problem is well-poised. The second example is Abel-Goncarov information given by

$$\mathfrak{N} = \{f(z_1), f'(z_2), \ldots, f^{(n-1)}(z_n)\},$$

see Sharma [72]. Recall that if $z_i = z_1$ for $i = 2, \ldots, n$ then we get one-point information which has the order n (even in the multivariate and abstract cases). For Abel-Goncarov information it is possible to prove

$$n \le p(\mathfrak{N}) \le 2n$$

but we do not know whether this upper bound is sharp. Finally let us mention lacunary information given by

$$\mathfrak{N} = \{f(z_1), f''(z_1), f(z_2), f''(z_2), \ldots, f(z_k), f''(z_k)\}$$

and $n = 2k$, see Sharma [72]. It is possible to verify that

$$\frac{1}{2}\, 2^{n/2} \le p(\mathfrak{N}) \le \frac{3}{4}\, 2^n$$

but the exact value of $p(\mathfrak{N})$ is unknown.

ACKNOWLEDGMENT

I am greatly appreciative to J. F. Traub, A. Sharma, B. Kacewicz and R. Meersman for their helpful comments and assistance during the preparation of this paper.

REFERENCES

Brent [75] Brent, R. P.,"A Class of Optimal Order Zero Finding Methods Using Derivative Evaluation," Department of Computer Science Report, Carnegie-Mellon University, 1975. To appear in Analytic Computational Complexity, edited by J. F. Traub, Academic Press, 1975.

Kung and Traub [73] Kung, H. T. and J. F. Traub, "Optimal Order for Iterations Using Two Evaluations," Department of Computer Science Report, Carnegie-Mellon University, 1973. To appear in SIAM J. Numer. Anal.

Kung and Traub [74] Kung, H. T. and J. F. Traub, "Optimal Order of One-Point and Multipoint Iterations," J. Assoc. Comput. Mach., Vol. 21, No. 4, 1974, 643-651.

Meersman [75] Meersman, R., "Optimal Use of Information in Certain Iterative Processes," Department of Computer Science Report, Carnegie-Mellon University, 1975. To appear in Analytic Computational Complexity, edited by J. F. Traub, Academic Press, 1975.

Traub [61] Traub, J. F., "On Functional Iteration and the Calculation of Roots," Proc. 16th Nat. ACM Conf., 5A-1 (1961), 1-4.

| | |
| --- | --- |
| Traub [64] | Traub, J. F., Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, N. J., 1964. |
| Schoenberg [66] | Schoenberg, I, J., "On Hermite-Birkhoff Interpolation," J. Math. Anal. Appl., 16 (1966), 538-543. |
| Sharma [72] | Sharma, A., "Some Poised and Nonpoised Problems of Interpolation," SIAM Review, Vol. 14, No. 1, 1972, 129-151. |
| Wozniakowski [75] | Wozniakowski, H., "Generalized Information and Maximal Order of Iteration for Operator Equations," SIAM J. Numer. Anal., Vol. 12, No. 1, 1975, 121-135. |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (*and Subtitle*)

MAXIMAL ORDER OF MULTIPOINT ITERATIONS USING n EVALUATIONS

5. TYPE OF REPORT & PERIOD COVERED

Interim

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

H. Wozniakowski

8. CONTRACT OR GRANT NUMBER(s)

N0014-67-0314-0010, NR 044-422; CMU 1-51039

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Carnegie-Mellon University
Computer Science Dept.
Pittsburgh, PA 15213

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research
Arlington, VA 22217

12. REPORT DATE

July 1975

13. NUMBER OF PAGES

34

14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*)

15. SECURITY CLASS (*of this report*)

UNCLASSIFIED

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (*of this Report*)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (*of the abstract entered in Block 20, if different from Report*)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (*Continue on reverse side if necessary and identify by block number*)

20. ABSTRACT (*Continue on reverse side if necessary and identify by block number*)

This paper deals with multipoint iterations without memory for the solution of the nonlinear scalar equation $f^{(m)}(x) = 0$, $m \geq 0$. Let $p_n(m)$ be the maximal order of iterations which use n evaluations of the function or its derivatives per step. We prove the Kung and Traub conjecture $p_n(0) = 2^{n-1}$ for Hermitian information. We show $p_n(m+1) \geq p_n(m)$ and conjecture $p_n(m) \equiv 2^{n-1}$. The problem of the maximal order is connected with Birkhoff interpolation. Under a certain assumption we prove that the Polya conditions are necessary for maximal order.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)