# OPTIMAL ORDER OF ONE-POINT AND MULTIPOINT ITERATION

H. T. Kung and J. F. Traub

Computer Science Department
Carnegie-Mellon University
Pittsburgh, Pa.

February 1973

This research was supported in part by the National Science Foundation under Grant GJ-32111 and the Office of Naval Research under Contract N00014-67-A-0314-0010, NR 044-422.

## ABSTRACT

The problem is to calculate a simple zero of a non-linear function f by iteration. We exhibit a family of iterations of order $2^{n-1}$ which use n evaluations of f and no derivative evaluations, as well as a second family of iterations of order $2^{n-1}$ based on n-1 evaluations of f and one of f'. In particular, with four evaluations we construct an iteration of eighth order. The best previous result for four evaluations was fifth order.

We prove that the optimal order of one general class of multipoint iterations is $2^{n-1}$ and that an upper bound on the order of a multipoint iteration based on n evaluations of f (no derivatives) is $2^n$.

CONJECTURE. A multipoint iteration without memory based on n evaluations has optimal order $2^{n-1}$.

TABLE OF CONTENTS

## 1. INTRODUCTION

We deal with iterations for calculating simple zeros of a scalar function f. This problem is a prototype for many non-linear numerical problems (Traub [72]). Newton-Raphson iteration is probably the most widely used algorithm for dealing with such problems. It is of second order and requires the evaluation of f and f', that is, it uses two evaluations. Consider an iteration consisting of two successive Newton-Raphson iterates (composition of iterates). This iteration has fourth order and requires four evaluations, two of f and two of f'. More generally an iteration composed of n Newton iterates is of order $2^n$ and requires n evaluations of f and n evaluations of f', that is, 2n evaluations.

We shall show that an iteration of order $2^{n-1}$ may be constructed from just n evaluations of f. We exhibit a second type of iteration which requires n-1 evaluations of f and one evaluation of f' to achieve order $2^{n-1}$.

In particular, with four evaluations we construct an iteration of eighth order. The best previous result (Traub [64, p. 196]) for four evaluations was fifth order.

Newton-Raphson iteration is an example of a one-point iteration. The basic optimality theorem for one-point iteration states that an analytic one-point iteration which is based on n evaluations is of order at most n. (This theorem was first stated by Traub [61], [64, Section 5.4]; we give an improved proof here.) We conjecture that a multipoint iteration based on n evaluations has optimal order $2^{n-1}$. We prove that the optimal order of one important family of multipoint iterations is $2^{n-1}$ and that an upper bound on the order of multipoint iteration based on n evaluations of f is

$2^n$. This upper bound is close to the conjectured optimal order of $2^{n-1}$.

To compare various algorithms, we must define efficiency measures based on speed of convergence (order), cost of evaluating f and its derivatives (problem cost), and the cost of forming the iteration (algorithm cost). We analyze efficiency in another paper (Kung and Traub [73]). We confine ourselves here to iterations <u>without memory</u> deferring the analysis of iterations with memory to a future paper.

We summarize the results of this paper. The class of problems and algorithms studied in this paper is defined in Section 2. Particular families of iterations are defined in the next three sections. The optimality theorem for one-point iterations is proven in Section 6. An optimal order theorem for one general class of multipoint iterations and an upper bound for the order of a second class are proven in the following section. A general conjecture is stated in Section 8. Section 9 contains a small numerical example.

Appendix I gives pseudo-Algol programs for forming two families of multipoint iterations. The last appendix proves a theorem on where evaluations of an iteration must be taken.

## 2. DEFINITIONS

We define the ensemble of problems and algorithms. Let

$D = \{f | f$ is a real analytic function defined on an open interval $I_f \subset R$ which contains a simple zero $\alpha_f$ of f and f' does not vanish on $I_f.\}$.

Let $\Omega$ denote the set of functions $\phi$ which maps every $f \in D$ to $\phi(f)$ with the following properties:

1. $\phi(f)$ is a function mapping $I_{\phi,f} \subset I_f$ into $I_{\phi,f}$ for some open subinterval $I_{\phi,f}$ containing $\alpha_f$.

2. $\phi(f)(\alpha_f) = \alpha_f$.

3. There exists an open subinterval $I^0_{\phi,f} \subset I_{\phi,f}$ containing $\alpha_f$ such that if $x_{i+1} = \phi(f)(x_i)$ then $\lim_{i\to\infty} x_i = \alpha_f$ whenever $x_0 \in I^0_{\phi,f}$.

4. Let $k, d_0,\ldots,d_{k-1}$ be non-negative integers. For $j=-1,\ldots,k-1$, let $u_{j+1}(y_0;y^0_1,\ldots,y^0_{d_0+1};\ldots;y^j_1,\ldots,y^j_{d_j+1})$ be a function of $1 + \sum_{i=0}^{j}(d_i+1)$ variables. For $j=0,\ldots,k-1$, let

$$(2.1)\qquad \begin{aligned} z_0 &= u_0(x), \\ z_{j+1} &= u_{j+1}(x;f(z_0),\ldots,f^{(d_0)}(z_0);\ldots;f(z_j),\ldots,f^{(d_j)}(z_j)). \end{aligned}$$

$\phi(f)(x)$ is defined by

$$(2.2)\qquad \phi(f)(x) = z_k.$$

The assumption that $f \in D$ is needed for theorems dealing with a class of iterations. Any particular $\phi$ can be applied to f having only a certain number of derivatives.

In Appendix II, we show that for any $\phi \in \Omega$, if the function $u_0(x)$ in (2.1) is continuous, then $u_0(x) = x$ for all x. To simplify proofs in this paper we <u>assume</u> that $u_0(x) = x$ for all x.

If $\phi \in \Omega$, $\phi$ is called an <u>iteration without memory</u>, since if the sequence $\{x_i\}$ is generated by $x_{i+1} = \phi(f)(x_i)$, $x_{i+1}$ is computed using information only at the current point $x_i$. In this paper we limit ourselves to iterations without memory.

We classify iterations without memory. If k is the non-negative integer in (2.2), then we say <u>$\phi$ is a k-point iteration</u>. In particular, if k=1, we call $\phi$ a <u>one-point iteration</u> and if $k > 1$ and the value of k is not important, we call $\phi$ a <u>multipoint iteration</u>. (Similar definitions of one-point and multipoint iteration are given in Traub [61], [64, Section 1.22].)

If there exists $p(\phi)$ such that for any $f \in D$,

$$\lim_{x \to \alpha_f} \frac{\phi(f)(x) - \alpha_f}{(x-\alpha_f)^{p(\phi)}} = S(\phi,f)$$

exists for a constant $S(\phi,f)$ and $S(\phi,f) \neq 0$ for at least one $f \in D$, then $\phi$ is said to have <u>order of convergence</u> (order) $p(\phi)$ and asymptotic error constant $S(\phi,f)$.

Let $v_i(\phi)$ denote the number of evaluations of $f^{(i)}$ used to compute $\phi(f)(x)$. Then $v(\phi) = \sum_{i \geq 0} v_i(\phi)$ is the total number of evaluations required by $\phi(f)(x)$ per step.

To simplify notation, we often use $\alpha$, $\phi$, p, $v_i$, v instead of $\alpha_f$, $\phi(f)(x)$, $p(\phi)$, $v_i(\phi)$, $v(\phi)$, if there is no ambiguity.

The following two examples illustrate the definitions.

Example 2.1. (Newton-Raphson Iteration)

$$\phi(f)(x) = x - \frac{f(x)}{f'(x)} .$$

This is a one-point iteration with $v_0=1$, $v_1=1$, $v=2$, and $p=2$.

Example 2.2.

$$z_0 = x,$$

$$z_1 = z_0 - \frac{f(z_0)}{f'(z_0)} ,$$

$$\phi(f)(x) = z_2 = z_1 - \frac{f(z_1)f(z_0)}{[f(z_1)-f(z_0)]^2} \cdot \frac{f(z_0)}{f'(z_0)} .$$

This is a two-point iteration with $v_0=2$, $v_1=1$, $v=3$, and $p=4$. (See Section 5.)

## 3. A FAMILY OF ONE-POINT ITERATIONS

For $f \in D$, let $F$ be the inverse function to $f$. For every $n$, define $\gamma_j(f)\colon I_f \to R$, $j=1,\dots,n$, as follows: $\gamma_1(f)(x) = x$ and for $n > 1$,

$$(3.1)\quad \gamma_{j+1}(f)(x) = \gamma_j(f)(x) + \frac{(-1)^j}{j!} \cdot [f(x)]^j \cdot F^{(j)}(f(x))$$

for $j=1,\dots,n-1$. Note that $F^{(j)}(f(x))$ can be expressed in terms of $f^{(i)}(x)$ for $i=1,2,\dots,j$. It is easy to show that

$$\gamma_1 = x$$
$$\gamma_2 = \gamma_1 - \frac{f(x)}{f'(x)}$$
$$\gamma_3 = \gamma_2 - \frac{f''(x)}{2f'(x)}\left[\frac{f(x)}{f'(x)}\right]^2 .$$

The family $\{\gamma_n\}$ has been thoroughly studied (Traub [64, Section 5.1]). Its essential properties are summarized in

<u>Theorem 3.1.</u>

<u>Let $\gamma_n$ be defined by (3.1). Then for $n > 1$,</u>

1. <u>$\gamma_n \in \Omega$, and $\gamma_n$ is a one-point iteration,</u>
2. <u>$p(\gamma_n) = n$,</u>
3. <u>$v_i(\gamma_n) = 1$, $i=0,\dots,n-1$, $v_i(\gamma_n) = 0$, $i > n-1$. Hence</u>

   <u>$v(\gamma_n) = n$.</u>

Thus $\gamma_n$ requires the evaluation of $f$ and its first $n-1$ derivatives. In Section 6 we shall show that, under a mild smoothness condition on the iteration, <u>every</u> one-point iteration of order $n$ requires the evaluation of at least $f$ and its first $n-1$ derivatives.

## 4. A FAMILY OF MULTIPOINT ITERATIONS

We construct a family of multipoint iterations, $\{\psi_n\}$, which require the evaluation of f at n points, no evaluation of derivatives of f, and for which $p(\psi_n) = 2^{n-1}$.

For every n, define

$$\psi_j(f): I_{\psi_j,f} \subset I_f \to I_{\psi_j,f}, \quad j=0,\dots,n,$$

as follows: $\psi_0(f)(x) = x$ and for $n > 0$,

$$(4.1)\quad \begin{cases} \psi_1(f)(x) = x + \beta f(x), & \beta \text{ a non-zero constant} \\ \quad\vdots \\ \psi_{j+1}(f)(x) = Q_j(0), \end{cases}$$

for $j=1,\dots,n-1$, where $Q_j(y)$ is the inverse interpolatory polynomial for f at $f(\psi_k(f)(x))$, $k=0,\dots,j$. That is, $Q_j(y)$ is the polynomial of degree at most j such that

$$Q_j(f(\psi_k(f)(x))) = \psi_k(f)(x), \quad k=0,\dots,j.$$

The $\psi_j(f)$, $j=1,\dots,n$, are well defined if

$$(4.2)\quad \psi_j(f)(I_{\psi_j,f}) \subset I_{\psi_j,f}, \quad j=1,\dots,n.$$

That (4.2) holds for $I_{\psi_j,f}$ sufficiently small will be part of the proof of Theorem 4.1..

It is easy to show that

$$\psi_0 = x,$$

$$\psi_1 = \psi_0 + \beta f(\psi_0),$$

$$\psi_2 = \psi_1 - \frac{\beta f(\psi_0)f(\psi_1)}{f(\psi_1) - f(\psi_0)},$$

$$\psi_3 = \psi_2 - \frac{f(\psi_0)f(\psi_1)}{f(\psi_2) - f(\psi_0)}\left(\frac{\psi_1-\psi_0}{f(\psi_1) - f(\psi_0)} - \frac{\psi_2-\psi_1}{f(\psi_2) - f(\psi_1)}\right).$$

A short pseudo-Algol program (Program 1) is given in Appendix I for computing $\psi_n$ for $n \geq 4$.

Our interest in the family of iterations $\{\psi_n\}$ is due to the properties proved in

Theorem 4.1.

Let $\psi_n$ be defined by (4.1). Then for $n > 1$,

1. $\psi_n \in \Omega$, and $\psi_n$ is an n-point iteration,
2. $p(\psi_n) = 2^{n-1}$,
3. $v_0(\psi_n) = n$, $v_i(\psi_n) = 0$, $i > 0$. Hence $v(\psi_n) = n$.

Proof.

We want to show that, for $f \in D$,

$$(4.3) \quad \lim_{x\to\alpha} \frac{\psi_n-\alpha}{(x-\alpha)^{2^{n-1}}} = S(\psi_n,f), \quad n=1,2,\ldots$$

for constants $S(\psi_n,f)$. The proof is by induction on n.

Since

$$\lim_{x\to\alpha} \frac{\psi_1-\alpha}{x-\alpha} = 1 + \beta f'(\alpha),$$

(4.3) holds for n=1. Assume that (4.3) holds for n=1,...,m-1. From general interpolatory iteration theory (Traub [64, Chapter 4]), we know that

$$(4.4)\quad \lim_{x\to\alpha} \frac{\psi_m-\alpha}{\prod_{0\le n<m} (\psi_n-\alpha)} = Y_m(f)$$

where

$$Y_m(f) = \frac{(-1)^{m+1}F^{(m)}(0)}{m![F'(0)]^m}$$

and F is the inverse function of f. From (4.4) and the induction hypothesis,

$$\lim_{x\to\alpha} \frac{\psi_m-\alpha}{(x-\alpha)^{2^{m-1}}} = \lim_{x\to\alpha} \frac{\psi_m-\alpha}{\prod_{0\le n<m} (\psi_n-\alpha)} \cdot \frac{\psi_0-\alpha}{x-\alpha} \cdot \prod_{1\le n<m}\left[\frac{\psi_n-\alpha}{(x-\alpha)^{2^{n-1}}}\right]$$

$$= Y_m(f) \cdot \prod_{1\le n<m} S(\psi_n,f).$$

Hence $S(\psi_m,f) = Y_m(f) \cdot \prod_{1\le n<m} S(\psi_n,f)$ and this completes the induction.

From (4.3) one can easily show that $\psi_j(f)$, $j=2,\ldots,n$, satisfies (4.2) for $I_{\psi_j,f}$ sufficiently small and hence is well defined. Now we prove that $\psi_n \in \Omega$. It follows from (4.3) that $\psi_n$ satisfies properties 1, 2, 3 of Section 2. Define

$$z_0 = u_0(x) = x,$$

$$z_1 = u_1(x,f(x)) = x+\beta f(x),$$

$$z_{j+1} = u_{j+1}(x;f(z_0),f(z_1),\ldots,f(z_j)),$$

for $j=1,\ldots,n-1$, where $u_{j+1}(x;f(z_0),f(z_1),\ldots,f(z_j)) = Q_j(0)$ and $Q_j(y)$ is the inverse interpolatory polynomial for f at $f(z_k)$, $k=0,\ldots,j$. Then by (4.1), $\psi_n(f)(x) = z_n$ for all $f \in D$ and $x \in I_{\psi_n,f}$. Hence $\psi_n(f)$ satisfies property 4 of Section 2. Therefore, $\psi_n \in \Omega$.

It is not difficult to show that $S(\psi_n,f) \neq 0$ for some $f \in D$. Therefore, $P(\psi_n) = 2^{n-1}$. The fact that $v_0(\psi_n) = n$, $v_i(\psi_n) = 0$, $i > 0$ follows from the definition of $\psi_n$. QED

The iteration $\psi_2$ is second order and is based on evaluations of f at x and $x + \beta f(x)$. This iteration is given by Traub [64, Section 8.4]. The iteration $\psi_n$ uses n evaluations of f and is of order $2^{n-1}$. For $n > 2$, no iterations with these properties were previously known.

## 5. A SECOND FAMILY OF MULTIPOINT ITERATIONS

We now construct a second family of multipoint iterations, $\{\omega_n\}$, such that $p(\omega_n) = 2^{n-1}$ and $v(\omega_n) = n$. However, $\omega_n$ requires the evaluation of f at n-1 points and the evaluation of f' at one point.

For every n, define

$$\omega_j(f): \; I_{\omega_j,f} \subset I_f \to I_{\omega_j,f}, \; j=1,\dots,n,$$

as follows: $\omega_1(f)(x) = x$ and for $n > 1$,

$$(5.1) \quad \begin{cases} \omega_2(f)(x) = x - \dfrac{f(x)}{f'(x)}, \\ \quad\vdots \\ \omega_{j+1}(f)(x) = R_j(0), \end{cases}$$

for j=2,...,n-1, where $R_j(y)$ is the inverse Hermite interpolatory polynomial of degree at most j such that

$$(5.2) \quad \begin{aligned} &R_j(f(x)) = x, \\ &R_j'(f(x)) = \frac{1}{f'(x)}, \\ &R_j(f(\omega_k(f)(x))) = \omega_k(x), \; k=2,\dots,j. \end{aligned}$$

One can prove that $\omega_j(f)$, j=2,...,n, are well defined for $I_{\omega_j,f}$ sufficiently small. It is easy to show that

$$\omega_1 = x,$$

$$\omega_2 = \omega_1 - \frac{f(\omega_1)}{f'(\omega_1)},$$

$$\omega_3 = \omega_2 - \frac{f(\omega_1)f(\omega_2)}{[f(\omega_1)-f(\omega_2)]^2} \cdot \frac{f(\omega_1)}{f'(\omega_1)} .$$

A short pseudo-Algol program (Program 2) is given in Appendix I for computing $\omega_4$ for $n \geq 4$.

The basic properties of the family of iterations $\{\omega_n\}$ is stated in the following theorem. The proof is omitted since it is similar to the proof of Theorem 4.1.

Theorem 5.1.

Let $\omega_n$ be defined by (5.1). Then for $n \geq 2$

1. $\omega_n \in \Omega$, and $\omega_n$ is an (n-1)-point iteration,
2. $p(\omega_n) = 2^{n-1}$,
3. $v_0(\omega_n) = n-1$, $v_1(\omega_n) = 1$, $v_i(\omega_n) = 0$, $i > 1$. Hence $v(\omega_n) = n$.

It is straightforward to show that

$$(5.3) \quad \frac{S(\psi_n, f)}{S(\omega_n, f)} = [1 + \beta f'(\alpha)]^{2^{n-2}}.$$

If $\psi_n$ is used, $\beta$ should be chosen so that $1 + \beta f'(\alpha)$ is small.

The iteration $\omega_3$ uses two evaluations of f and one of f' and $p(\omega_3) = 4$. Another iteration with these properties is defined by Ostrowski [66, Appendix G] and a geometrical interpretation is given by Traub [64, Section 8.5]. King [73] gives a family of fourth order methods based on two evaluations of f and one of f'. Jarratt [69] constructs a fourth order iteration based on one evaluation of f and two of f'. The iteration $\omega_n$ uses n-1 evaluations of f and one of f' and $p(\omega_n) = 2^{n-1}$. For $n > 3$, no iterations with these properties were previously known.

## 6. THE OPTIMAL ORDER OF ONE-POINT ITERATIONS

By imposing a mild smoothness condition we can prove that one-point iterations of order n require the evaluation of f and at least its first n-1 derivatives. No such requirement holds for multipoint iterations. For example, the multipoint iteration $\psi_n$ defined in Section 4 has order $2^{n-1}$ but requires no derivative evaluation.

Let $\phi$ be a one-point iteration. Then from (2.1), (2.2)

$$\phi(f)(x) = u_1(x, f(x), \ldots, f^{(d_0)}(x)),$$

where $u_1(y_0, y_1^0, \ldots, y_{d_0+1}^0)$ is a multivariate function of $d_0+2$ variables. In this section we drop the superscript on $y_j$.

The following theorem was first given by Traub [61], [64, Section 5.4]. We regard the proof given here as an improvement of Traub's proof.

Theorem 6.1.

*Let $\phi$ be a one-point iteration of order $p(\phi)$ and let $u_1(y_0, y_1, \ldots, y_{d_0+1})$ be analytic with respect to $y_1$ at $y_1=0$. Then $v_i(\phi) \geq 1$, $i=0,\ldots,p(\phi)-1$, and hence $p(\phi) \leq v(\phi)$.*

Proof.

For $f \in D$, define

$$(6.1) \quad T(f)(x) = \frac{\gamma_p(f)(x) - \phi(f)(x)}{f^p(x)}$$

where $\gamma_p$ is a member of the family of iterations defined in Section 3. To simplify notation, we write f for f(x).

Define $\sigma_i$ by

$$\gamma_p = \sum_{i=0}^{p-1} \sigma_i f^i$$

where $\sigma_i$ depends explicitly on $f'(x),\dots,f^{(i)}(x)$ (Traub [64, Section 5.1]). By the analyticity condition on $\phi$,

$$\phi = \sum_{i=0}^{\infty} \lambda_i f^i.$$

Therefore from (6.1),

$$(6.2)\quad T = \sum_{i=0}^{p-1} (\sigma_i - \lambda_i) f^{i-p} - \sum_{i=p}^{\infty} \lambda_i f^{i-p}.$$

Since $\phi$ and $\gamma_p$ are of order p, (6.1) implies that

$$\lim_{x\to\alpha_f} T(f)(x) = \frac{S(\gamma_p, f) - S(\phi, f)}{[f'(\alpha_f)]^p} < \infty$$

for all $f \in D$. Hence it follows from (6.2) that

$$(6.3)\quad \sigma_i = \lambda_i, \; i=0,\dots,p-1, \; \forall f \in D.$$

Consider

$$\sigma_{p-1} = \lambda_{p-1}$$

We know that $\sigma_{p-1}$ depends explicitly on $f'(x),\dots,f^{(p-1)}(x)$ and that the same must be true for $\lambda_{p-1}$. Assume $v_j(\phi) = 0$, for some j, $0 < j \le p-1$. Then $u_1(y_0,y_1,\dots,y_{d_0+1})$ does not depend on $y_{j+1}$. This implies that

$$\frac{\partial^{p-1}}{\partial y_1^{\,p-1}} u_1(y_0,y_1,\dots,y_{d_0+1})$$

is independent of $y_{j+1}$ and that

$$\frac{\partial^{p-1}}{\partial y_1{}^{p-1}} u_1(x,0,f'(x),\ldots,f^{(d_0)}(x))$$

is independent of $f^{(j)}(x)$. Hence

$$\lambda_{p-1} = \frac{1}{(p-1)!} \frac{\partial^{p-1}}{\partial y_1{}^{p-1}} u(x,0,f'(x),\ldots,f^{(d_0)}(x))$$

is independent of $f^{(j)}(x)$, which is a contradiction. Therefore $v_i(\phi) \geq 1$, $i=1,\ldots,p-1$.

Next we show $v_0(\phi) \geq 1$. Suppose this is false. Then $\lambda_i=0$, $i > 0$ and from (6.2)

$$\sigma_i = 0, \ i=1,\ldots,p-1,$$

which is a contradiction. QED

Corollary 6.1.

Let $\gamma_n$ be defined by (3.1). Then $\gamma_n$ achieves the optimal order of any one-point iteration $\phi$ for which $v(\phi) = n$ and which satisfies the analyticity condition of the theorem.

Remark

The analyticity condition is not restrictive. For example, it includes all rational iterations and all iterations defined by simple zeros of polynomials with analytic coefficients.

## 7. TWO OPTIMAL ORDER THEOREMS FOR MULTIPOINT ITERATIONS

We prove an optimal order theorem for one important class of iterations and prove a fairly tight upper bound for the maximal order of a second class of iterations.

Our first class consists of all iterations such that for $j=0,\dots,k-1$, $z_{j+1}$ appearing in (2.1) is given by a Hermite interpolatory iteration based on the points $z_0,\dots,z_j$. If $\phi$ belongs to this family, we say it is a <u>Hermite interpolatory k-point iteration</u>. The order of $\phi$ may be computed as follows. From Traub [64, Section 4.2],

$$z_{j+1} - \alpha = O[(z_j-\alpha)^{d_j+1} \dots (z_0-\alpha)^{d_0+1}]$$

where the $d_i$ are as in (2.1). Hence

$$\phi(f)(x)-\alpha = z_k-\alpha = O[(x-\alpha)^p]$$

where

$$p(\phi) = (d_0+1)\prod_{j=1}^{k-1}(d_j+2).$$

It is easily verified that

$$v(\phi) = \sum_{j=0}^{k-1}(d_j+1).$$

We wish to choose $k, d_0,\dots,d_{k-1}$ such that for $v(\phi)$ fixed, $p(\phi)$ is maximized. The choice of k and the $d_i$ are given by

<u>Theorem 7.1.</u>

<u>Let $d_j \geq 0$, $k \geq 1$ be integers. Let</u>

$$v(\phi) = \sum_{j=0}^{k-1}(d_j+1) = n$$

<u>be fixed. Then</u>

$$p(\emptyset) = (d_0+1) \prod_{j=1}^{k-1} (d_j+2)$$

<u>is maximized exactly when</u>

(7.1) <u>$k=n,\ d_j=0,\ j=0,\dots,n-1$</u>

or

(7.2) <u>$k=n-1,\ d_0=1,\ d_j=0,\ j=1,\dots,n-2.$</u>

<u>Proof</u>.

Since $d_j+1 \le n$, $k \le n$, there are only finitely many cases and the maximum exists. Let the maximum of p be achieved at $\bar{d}_j$, $j=0,\dots,\bar{k}-1$. We show first that $\bar{d}_j=0$, $j=1,\dots,\bar{k}-1$. Assume that $\bar{d}_r=m$, $m \ge 1$, for some r, $r=1,\dots,\bar{k}-1$. Define $\bar{\bar{d}}_j$, $j=0,\dots,\bar{k}+m-1$ as

$$\bar{\bar{d}}_j = \bar{d}_j,\ j=0,\dots,\bar{k}-1,\ j\neq r,$$
$$\bar{\bar{d}}_r = 0,$$
$$\bar{\bar{d}}_j = 0,\ j=\bar{k},\dots,\bar{k}+m-1.$$

Then we can verify that

$$\sum_{j=0}^{\bar{k}+m-1} (\bar{\bar{d}}_j+1) = \sum_{j=0}^{\bar{k}-1} (d_j+1) = n$$

and

$$(\bar{\bar{d}}_0+1) \prod_{j=1}^{\bar{k}+m-1} (\bar{\bar{d}}_j+2) > (\bar{d}_0+1) \prod_{j=1}^{\bar{k}-1} (\bar{d}_j+2).$$

This contradiction proves that $\bar{d}_j=0$, $j=1,\dots,\bar{k}-1$. A similar argument may be used to prove $\bar{d}_0 \le 1$. If $\bar{d}_0=0$, $\bar{k}=n$ while if $\bar{d}_0=1$, $\bar{k}=n-1$ which completes the proof. QED

Corollary 7.1.

Let $\phi$ be a Hermite interpolatory iteration with $v(\phi) = n$. Then

$$p(\phi) \le 2^{n-1}.$$

Note that $\psi_n$, defined in Section 4, is an instance of (7.1) while $\omega_n$, defined in Section 5, is an instance of (7.2). (Both $\{\psi_n\}$ and $\{\omega_n\}$ are based on inverse interpolation. There are two other families of iterations based on direct interpolation.) Thus we have

Corollary 7.2.

Let $\psi_n$ and $\omega_n$ be defined by (4.1) and (5.1), respectively. Then $\psi_n$ and $\omega_n$ have optimal order for Hermite interpolatory interation with n evaluations.

The second theorem of this section gives an upper bound on the order achievable for any multipoint iteration uses values of f only (and no derivatives).

Theorem 7.2.

Let $\phi$ be a multipoint iteration with $v_0(\phi) = n$, $v_i(\phi) = 0$, $i > 0$. Then $p(\phi) \le 2^n$.

Proof.

For $f \in D$, let $x_0$ be a starting point such that if $x_{i+1} = \varphi(f)(x_i)$ then $\lim_{i\to\infty} x_i = \alpha_f$. From (2.1), for each i, denote

$$z_{in} = x_i,$$
$$z_{in+j+1} = u_{j+1}(z_{in}, f(z_{in}), \ldots, f(z_{in+j}))$$

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

$$(7.3)\quad \lim_{i\to\infty} \frac{z_{(i+1)n}-\alpha}{(z_{in}-\alpha)^p} = S(\phi, f).$$

Suppose that $p > 2^n$. Choose $q$ such that $p > q > 2^n$. Then (7.3) implies

$$\lim_{i\to\infty} \frac{z_{(i+1)n}-\alpha}{(z_{in}-\alpha)^q} = 0.$$

Let $m$ be sufficiently large so that for $i \geq m$

$$\frac{|z_{(i+1)n}-\alpha|}{|z_{in}-\alpha|^q} < 1, \quad |z_{in}-\alpha| < 1.$$

Then

$$(7.4)\quad |z_{(m+j)n}-\alpha| < |z_{mn}-\alpha|^{q^j}, \ \forall j.$$

However, there exists a $f \in D$, a sequence $\{z_i\}$ and a constant $A$, $0 < A < 1$, such that

$$(7.5)\quad |z_t-\alpha| > A^{2^t}, \ \forall t,$$

(see Winograd and Wolfe [71, Theorem 3]). From (7.4) and (7.5),

$$|z_{mn}-\alpha|^{q^j} > |z_{(m+j)n}-\alpha| > A^{2^{(m+j)n}} = (A^{2^{mn}})^{2^{nj}}.$$

Hence

$$q^j|\log|z_{mn}-\alpha|| < 2^{nj}|\log A^{2^{mn}}|,$$

or

$$\left(\frac{q}{2^n}\right)^j < \frac{|\log A^{2^{mn}}|}{|\log|z_{mn}-\alpha||}.$$

Since the left side is unbounded as $j \to \infty$ while the right side is independent of j, we have a contradiction. Hence $p(\phi) \le 2^n$. QED

In Section 4, we constructed an iteration $\psi_n$ such that $v_0(\psi_n) = n$, $v_i(\psi_n) = 0$, $i > 0$ and $p(\psi_n) = 2^{n-1}$. Hence the upper bound of Theorem 7.3 is within a factor of two of the order of that iteration. We conjecture in the next section that $p = 2^{n-1}$ is optimal.

## 8. A CONJECTURE

Conjecture 8.1.

Let $\phi$ be an iteration (with no memory) with $v(\phi) = n$. Then

$$p(\phi) \le 2^{n-1}. \quad (8.1)$$

This extends a conjecture of Traub [72] which states (8.1) for n=2,3.

## 9. NUMERICAL EXAMPLE

Let $f(x) = x^3 + \ln(1+x)$ where ln denotes the logarithm to the natural base. Hence $\alpha=0$. Starting at $x_0=10^{-1}$ and $10^{-2}$, we compute $x_1$ by iterations $\psi_n$ and $\omega_n$, n=3,4,5. For comparison we also use as many steps of the Newton-Raphson iteration as necessary to bring the error to about $10^{-16}$. Calculations were done in double precision arithmetic on a DEC PDP-10 computer. About 16 digits are available in double precision.

Results are summarized in Examples 1-3. The parameter $\beta$ that appears in $\psi_n$ was chosen $\beta = -.2$ which makes the asymptotic error constant of $\psi_n$ for this problem near unity. The asymptotic error constants of $\omega_n$, n=3,4,5 and the Newton-Raphson iteration are also near unity for this problem. Recall that $p(\psi_n) = p(\omega_n) = 2^{n-1}$ and that for Newton-Raphson iteration, p=2. We expect $x_i \doteq x_0^p$ to hold and this is numerically verified in the examples. From (5.3), we expect

$$(9.1) \quad \frac{\psi_n(x_0)}{\omega_n(x_0)} \doteq (.8)^{2^{n-2}}$$

and (9.1) is numerically verified in the examples for $x_0 = 10^{-2}$.

The examples illustrate the advantage of $\psi_n$ and $\omega_n$ over the repeated use of Newton-Raphson iteration. Starting with $x_0 = 10^{-1}$, $\omega_5(x_0)$ calculates the zero to "full accuracy" at a cost of four evaluations of f and one of f'. Four Newton-Raphson iterations are required with a cost of four evaluations of f and four of f'. The difference is significant when the evaluation of f' is expensive. This observation takes only the cost of f' into account. A more complete analysis based on efficiency measure considerations is given by Kung and Traub [73].

EXAMPLE 1

| $x_0$ | $10^{-1}$ | $10^{-2}$ |
|---|---|---|
| $x_1 = \psi_3(x_0)$ | $.21\times10^{-4}$ | $.27\times10^{-8}$ |
| $x_1 = \psi_4(x_0)$ | $-.80\times10^{-9}$ | $-.47\times10^{-16}$ |
| $x_1 = \psi_5(x_0)$ | $-.27\times10^{-16}$ | |

EXAMPLE 2

| $x_0$ | $10^{-1}$ | $10^{-2}$ |
|---|---|---|
| $x_1 = \omega_3(x_0)$ | $.30\times10^{-4}$ | $.42\times10^{-8}$ |
| $x_1 = \omega_4(x_0)$ | $-.15\times10^{-8}$ | $-.12\times10^{-15}$ |
| $x_1 = \omega_5(x_0)$ | $-.24\times10^{-16}$ | |

EXAMPLE 3

Let $x_{i+1} = \phi(x_i)$, where $\phi$ denotes Newton-Raphson iteration.

| $x_0$ | $10^{-1}$ | $10^{-2}$ |
|---|---|---|
| $x_1$ | $-.26\times10^{-2}$ | $-.48\times10^{-4}$ |
| $x_2$ | $-.33\times10^{-5}$ | $-.11\times10^{-8}$ |
| $x_3$ | $-.54\times10^{-11}$ | $.46\times10^{-17}$ |
| $x_4$ | $-.31\times10^{-16}$ | |

# APPENDIX I

Program 1 and Program 2 which are adapted from a result of Krogh [70] compute $\psi_n(f)(x)$ (with the parameter $\beta$) and $\omega_n(f)(x)$ for $n \geq 4$.

Program 1.

```
v_{0,0} := x
h       := β × f(x)
v_{0,1} := v_{0,0} + h
v_{1,1} := h/(f(v_{0,1}) - f(v_{0,0}))
r       := v_{1,1} × f(v_{0,0})
v_{0,2} := v_{0,0} - r
v_{1,2} := r/(f(v_{0,0}) - f(v_{0,2}))
Π_2     := f(v_{0,0}) × f(v_{0,1})
v_{2,2} := (v_{1,1} - v_{1,2})/(f(v_{0,1}) - f(v_{0,2}))
psi     := v_{0,2} + Π_2 × v_{2,2}
for k=3 step 1 until n-1 do
begin
        v_{0,k} = psi
        for i=0 step 1 until k-1 do
        begin
                v_{i+1,k} := (v_{i,i}-v_{i,k})/(f(v_{0,i}) - f(v_{0,k}))
        end
        Π_k := -f(v_{0,k-1}) × Π_{k-1}
        psi:= psi + Π_k × v_{k,k}
end
```

Program 2.

$v_{0,0}$ := x

$v_{0,1}$ := x

$n_{1,1}$ := $f(v_{0,0})/f'(v_{0,0})$

$v_{0,2}$ := $v_{0,0} - n_{1,1}$

d := $f(v_{0,0}) - f(v_{0,2})$

$v_{1,2}$ := $n_{1,1}/d$

$v_{2,2}$ := $v_{1,2} \times f(v_{0,2})/d$

omega := $v_{0,2} - f(v_{0,0}) \times v_{2,2}$

$v_{1,1}$ := $n_{1,1}/f(v_{0,0})$

$v_{2,2}$ := $v_{2,2}/f(v_{0,0})$

$\Pi_2$ := $f(v_{0,0}) \times f(v_{0,0})$

<u>for</u> k=3 <u>step</u> 1 <u>until</u> n-1 <u>do</u>

<u>begin</u>

  $v_{0,k}$ := omega

  <u>for</u> i=0 <u>step</u> 1 <u>until</u> k-1 <u>do</u>

  <u>begin</u>

    $v_{i+1,k}$ := $(v_{i,i}-v_{i,k})/(f(v_{0,i}) - f(v_{0,k}))$

  <u>end</u>

  $\Pi_k$ := $-f(v_{0,k-1}) \times \Pi_{k-1}$

  omega := omega + $\Pi_k \times v_{k,k}$

<u>end</u>

# APPENDIX II

## Theorem

Let $\phi \in \Omega$. If the function $u_0$ appearing in the definition of the set $\Omega$ (property 4) is continuous, then $u_0(x) = x$, for all x.

## Proof

Consider the functions $u_0,\ldots,u_k$ in (2.1). Let $f \in D$, $x \in I^0_{\phi,f}$. Let $y^j_{i+1}(x)$, $z_j(x)$ be defined as follows:

$$z_0(x) = u_0(x),$$

(A.1) $$y^j_{i+1}(x) = f^{(i)}(z_j(x)),$$

$$z_{j+1}(x) = u_{j+1}(x;\ y^0_1(x),\ldots,y^0_{d_0+1}(x);\ldots;y^j_1(x),\ldots,y^j_{d_j+1}(x)),$$

for $j=0,\ldots,k-1$, $i=0,\ldots,d_j$. Then

(A.2) $$\phi(f)(x) = u_k(x;\ y^0_1(x),\ldots,y^0_{d_0+1}(x);\ldots;y^{k-1}_1(x),\ldots,y^{k-1}_{d_{k-1}+1}(x)).$$

Since $\phi(f)(\alpha_f) = \alpha_f$ by Property 2 (Section 2),

(A.3) $$\alpha_f = u_k(\alpha_f;y^0_1(\alpha_f),\ldots,y^{k-1}_{d_{k-1}+1}(\alpha_f)).$$

Suppose that $u_0(x) \not\equiv x$. Then there exists $\omega_0$ such that $u_0(\omega_0) \neq \omega_0$. By the continuity of $u_0$, $u_0(\omega) \neq \omega$ in an open interval $I_0$ containing $\omega_0$. We shall show that $\phi(f)(\omega) = \omega$ for all $\omega$ in $I_0 \cap I^0_{\phi,f}$. For any fixed $f \in D$ and $\omega \in I_0 \cap I^0_{\phi,f}$ define $y^j_{i+1}(\omega)$, $z_j(\omega)$ by setting $x = \omega$ in (A.1). Then by (A.2)

(A.4) $\phi(f)(\omega) = u_k(\omega;\ y_1^0(\omega),\ldots,y_{d_{k-1}+1}^{k-1}(\omega))$.

Since $u_0(\omega) \neq \omega$, there exists a polynomial q such that

$$q(\omega) = 0,\ q'(\omega) = 1,$$

$$q^{(i)}(z_j(\omega)) = y_{i+1}^j(\omega),\ j=0,\ldots,k-1,\ i=0,\ldots,d_j.$$

Certainly, $q \in D$ and $\alpha_q = \omega$. By (A.3)

(A.5) $\omega = u_k(\omega;\ y_1^0(\omega),\ldots,y_{d_{k-1}+1}^{k-1}(\omega))$.

Equations (A.4) and (A.5) imply that $\phi(f)(\omega) = \omega$ for all $f \in D$ and for all $\omega \in I_0 \cap I_{\phi,f}^0$. There exists a $g \in D$ such that $\alpha_g = \omega_0$. Since $\phi \in \Omega$, there exists an open interval $I_{\phi,g}^0$ containing $\omega_0$ such that if $x_{i+1} = \phi(g)(x_i)$, then $\lim x_i = \omega_0$ whenever $x_0 \in I_{\phi,g}^0$. This is a contradiction since for all $x_0 \in I_0 \cap I_{\phi,g}^0$ $(x_0 \neq \omega_0)$, $\phi(g)(x_0) = x_0$, which does not converge to $\omega_0$. QED

# BIBLIOGRAPHY


| | |
|---|---|
| Jarratt [69] | Jarratt, P., Some Efficient Fourth Order Multipoint Methods for Solving Equations. BIT 9 (1969), 119-124. |
| King [73] | King, R. R., A Family of Fourth-Order Methods for Nonlinear Equations. To appear. |
| Krough [70] | Krough, F. T., Efficient Algorithms for Polynomial Interpolation and Numerical Differentiation. Math. Comp. 24 (1970), 185-190. |
| Kung and Traub [73] | Kung, H. T. and Traub, J. F., Computational Complexity of One-Point and Multipoint Iteration. To appear. |
| Ostrowski [66] | Ostrowski, A. M., Solution of Equations and Systems of Equations, second edition. Academic Press, 1966. |
| Traub [61] | Traub, J. F., On Functional Iteration and the Calculation of Roots. Proceedings 16th National ACM Conference, 5A-1 (1961), 1-4. |
| Traub [64] | Traub, J. F., Iterative Methods for the Solution of Equations. Prentice-Hall, 1964. |
| Traub [72] | Traub, J. F., Computational Complexity of Iterative Processes. SIAM Journal Computing, Vol. 1, No. 2, (1972), 167-179. |
| Winograd and Wolfe [71] | Winograd, S. and Wolfe, P., Optimal Iterative Processes. IBM Report, RC3511 (#15883), 1971. |