Numer. Math. 20, 327—341 (1973)

Optimal Iterative Processes for Root-Finding*

Richard Brent, Shmuel Winograd, and Philip Wolfe

Mathematical Sciences Department, IBM Watson Research Center, Yorktown Heights, New York

Received August 3, 1972

Abstract. Let $f_0(x)$ be a function of one variable with a simple zero at r_0. An iteration scheme is said to be locally convergent if, for some initial approximations $x_1, \ldots, x_s$ near r_0 and all functions f which are sufficiently close (in a certain sense) to f_0, the scheme generates a sequence $\{x_k\}$ which lies near r_0 and converges to a zero r of f. The order of convergence of the scheme is the infimum of the order of convergence of $\{x_k\}$ for all such functions f. We study iteration schemes which are locally convergent and use only evaluations of $f, f', \ldots, f^{[d]}$ at $x_1, \ldots, x_{k-1}$ to determine x_k, and we show that no such scheme has order greater than $d+2$. This bound is the best possible, for it is attained by certain schemes based on polynomial interpolation.

I. Introduction

Many "iterative" methods are known for the numerical solution of the problem of finding a zero r of a function $f(x)$ of a single real variable. The iterative process generates a sequence $\{x_k\}$ of approximations to r, where x_k is determined by the values of f and possibly of some of its derivatives at previous members of the sequence. (The term "iterative" is widely and loosely used; the preceding description seems to cover its use in our subject.) If the process starts at points which are close enough to r, then the sequence $\{x_k\}$ should converge to r. The various methods differ in the amount of information used, the particular way the information is used to generate the next approximation, and consequently the rate at which the sequence $\{x_k\}$ converges to r. The secant method and Newton's method are examples of iterative methods which are much used in practice. Traub's book [1] describes a wide variety of such processes, all fitting the general outline: Given the points $x_{k-1}, \ldots, x_{k-m}$ as well as the values of the function and its first d derivatives at these points, construct the minimal degree interpolating polynomial fitting these data, and choose x_k as a root of this polynomial (or as its value at zero, if it is a polynomial in the dependent variable). The secant method and Newton's method are in this class of iteration methods.

An iterative method does not, however, have to use the root of such a polynomial. For example, the iteration defined by

$$x_k = \frac{f(x_{k-1})\,(x_{k-2}+f(x_{k-1})^2) - f(x_{k-2})\,(x_{k-1}+f(x_{k-1})^2)}{f(x_{k-1}) - f(x_{k-2})},$$

* This work was supported (in part) by the Office of Naval Research under contract numbers N0014-69-C-0023 and N0014-71-C-0112.

which differs slightly from the secant method, is not of this class. It is, of course, trivially true that every iterative method can be said to yield x_{k+1} as a root of the polynomial which interpolates the data *and* has a root at the point x_{k+1}, but this is not the minimum degree interpolating polynomial.

Newton's method is defined by the formula

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}.$$

Let e_k denote $|x_k - r|$. If $f \varepsilon C^2$, $f'(r) \neq 0$, and e_1 is sufficiently small, then $\lim_{k\to\infty} x_k = r$ and

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k^2} = \left| \frac{f''(r)}{2f'(r)} \right|.$$

Hence, we say that the order of convergence of Newton's method is 2 (provided $f'(r) \neq 0$ and $f''(r) \neq 0$.

For the secant method, given by

$$x_k = x_{k-1} - f(x_{k-1}) \frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})},$$

one can show under the same hypotheses that $\lim_{k\to\infty} \frac{e_k}{e_{k-1} e_{k-2}} = \frac{1}{2} \left| \frac{f''(r)}{f'(r)} \right|$, and consequently that $\lim_{k\to\infty} \frac{e_k}{e_{k-1}^{\lambda}} = C$, where $\lambda = \frac{1}{2}(1 + \sqrt{5})$ and $C = \left(\frac{1}{2} \left| \frac{f''(r)}{f'(r)} \right| \right)^{\lambda - 1}$.

We see that the statement about the rate of convergence of the secant method can take at least two forms: that $e_k/(e_{k-1} e_{k-2})$ converges to a finite, nonzero limit, or the same for e_k/e_{k-1}^{λ}. The reason why the second statement is usually preferred in the literature is that it yields an easy comparison with other iteration methods. Since $\frac{1}{2}(1 + \sqrt{5}) < 2$, we see that Newton's method converges faster than the secant method. (This does not imply that Newton's method is more efficient than the secant method, for Newton's method requires the calculation of both f and f' at each step, while only one new value of f is required at one step of the secant method.)

Some methods give a sequence $\{x_k\}$ which converges rapidly to r, but e_k/e_{k-1}^{λ} does not converge to a finite, nonzero limit for any value of λ. Hence, it is not satisfactory for our purposes to define the order to be the number λ such that

$$0 < \lim_{k\to\infty} e_k/e_{k-1}^{\lambda} < \infty. \tag{1.1}$$

Instead, we define the *upper* and *lower orders of convergence* of $\{x_k\}$ to be

$$\bar{\lambda} = \limsup_{k\to\infty} (-\log e_k)^{1/k} \tag{1.2}$$

and

$$\underline{\lambda} = \liminf_{k\to\infty} (-\log e_k)^{1/k} \tag{1.3}$$

respectively. If $\underline{\lambda} = \bar{\lambda}$ we say that the *order of convergence* is $\bar{\lambda}$. It is easy to see that if (1.1) holds then the order of convergence is λ, so our definition gives the usual order if it exists.

With different functions and/or starting approximations, the same iterative method may generate sequences with different orders of convergence. For example, Newton's method may give a sequence with order greater than two if

$f''(r) = 0$. Hence, we say that the order of a method is the infimum of the orders of convergence of the sequences that it generates when applied to a certain class of functions with a certain class of starting approximations (this will be made precise in Section 4). Similarly for the upper and lower orders of a method.

In discussing orders of convergence we assume that the initial approximations are sufficiently good to ensure convergence to a zero. For Newton's method, it is easily verified that if x_1 is in an interval J (which includes r) such that $0 < A \leqq f'(x) < 2A$ for all x (if $f'(r) > 0$) or $0 > A \geqq f'(x) > 2A$ for all x (if $f'(r) < 0$), then the method converges. (This condition emphasizes the "local" character of the method.) The same condition is also sufficient to guarantee the convergence of the secant method. The condition described above is not often stated in the literature. It is, nevertheless, useful to give quantitative description to statements like "if x_1 is sufficiently close to r".

Both the secant method and Newton's method may be called "stationary" iterative methods. The same function is used to determine x_k in every step of the iteration. Even though "stationary" iterative methods are the ones most often discussed, we need not restrict ourselves to such methods. As an example of a "nonstationary" method, let $P_k (k \geqq 2)$ be the minimal degree polynomial satisfying $P_k(f(x_i)) = x_i$ for $i = 1, 2, \ldots, k$, and define $x_{k+1} = P_k(0)$. (We show below that the order of this method is as great as that of any other iterative method which does not require derivatives.)

The question of the order of convergence of iterative schemes based on minimal-degree interpolating polynomials has been thoroughly discussed in the literature [1–5]; some of it is briefly reviewed in the next section. The question this paper addresses is: Do such schemes make the best possible use of the information obtained? More precisely, what is the highest order of convergence attainable by *any* local iteration scheme?

Our answer is that if a local iteration scheme uses only the values of f and its derivatives through the d-th (evaluated at arbitrarily many previous values of the independent variable), then the order of convergence of the scheme is no higher than $d+2$. We thus answer affirmatively a conjecture of Traub [1, p. 124] inspired by an examination of polynomial iteration schemes. Since the order $d+2$ is asymptotically attained by stationary polynomial schemes, and is attained by nonstationary polynomial schemes, we may say that the polynomial schemes are optimal.

Section 3 gives an imprecise outline of the derivation of the above result. Section 4 is devoted to the precise definition of the concepts our treatment requires, and Sections 5—6 are devoted to the proof of the result. Some possible extensions are mentioned in Section 7.

II. Order of Convergence of Polynomial Schemes

The essence of all the iterative methods that we are studying is the replacement, at each step, of a function f by a function g which agrees with f in a specified way. We therefore require a "remainder" formula for the error incurred when one function is interpolated by another. The most useful for us is the following, due to Cauchy:

Let $f, g \in C^n$, $f^{[i]}(x_j) = g^{[i]}(x_j)$ for all $i = 0, 1, \ldots, b_j - 1$ and $j = 1, \ldots, m$, where $n = b_1 + \cdots + b_m$. Let I be the smallest interval containing $\{x_j\}$. For any x there is ξ, in the smallest interval containing I and x, such that

$$f(x) - g(x) = \frac{1}{n!} [f - g]^{[n]}(\xi) \prod_{j=1}^{m} (x - x_j)^{b_j}. \tag{2.1}$$

The formula (2.1) leads to the derivation of the order of convergence for polynomial interpolation schemes with fixed m, as we can outline here. For simplicity, let all $b_j = d + 1$, so that $n = m(d + 1)$. We assume that $n > 1$ (i.e., either $m > 1$ or $d > 0$). If g_k is the interpolating polynomial agreeing with f at $x_{k-1}, \ldots, x_{k-m}$ through the d-th derivative, then g_k has degree $n - 1$, so (2.1) becomes

$$f(x) - g_k(x) = \frac{1}{n!} f^{[n]}(\xi) \prod_{j=1}^{m} (x - x_{k-j})^{d+1}. \tag{2.2}$$

Letting r be the root sought and $e_l = |r - x_l|$, we have

$$|g_k(r)| = \frac{1}{n!} |f^{[n]}(\xi)| \prod_{j=1}^{m} e_{k-j}^{d+1}. \tag{2.3}$$

Since $g_k(x_k) = 0$, $|g_k(r)| = e_k |g_k'(\eta)|$ for suitable η. Assume that $f'(r) \neq 0$ and $f^{[n]}(r) \neq 0$. Estimation of $|g_k'(\eta)|$ and $|f^{[n]}(\xi)|$ reduces (2.3) to

$$e_k = (C + \delta_k) \prod_{j=1}^{m} e_{k-j}^{d+1}, \tag{2.4}$$

or

$$\log e_k = \log(C + \delta_k) + (d + 1) \sum_{j=1}^{m} \log e_{k-j}, \tag{2.5}$$

where $C = |f^{[n]}(r) / (n! f'(r))|$ and $\delta_k \to 0$. The "difference equation" (2.5) is treated by standard means (noting that $\log e_k \to -\infty$ so that the term $\log(C + \delta_k)$ becomes unimportant), and we find that its solution has the property

$$\lim_{k \to \infty} \frac{\ln e_k}{\ln e_{k-1}} = \lambda, \tag{2.6}$$

where λ is the root of maximum modulus (which is unique and real) of the equation

$$\lambda^m = (d + 1) \sum_{j=0}^{m-1} \lambda^j. \tag{2.7}$$

From (2.6), the order of the method is λ. It is, of course, 2 for Newton's method ($m = 1$, $d = 1$); and $\frac{1}{2}(1 + \sqrt{5})$ for the secant method ($m = 2$, $d = 0$). As $m \to \infty$, the order tends to $d + 2$, which thus constitutes the best order approachable by polynomial interpolation using d derivatives. Moreover, the obvious extension of the nonstationary method mentioned in Section 1 has order exactly $d + 2$ on analytic functions with simple zeros.

In practice inverse polynomial interpolation methods may be preferable to direct methods, for the need to solve a polynomial equation at each iteration is eliminated. However, the orders of convergence of the corresponding direct and inverse methods are the same (see Brent [5] and Traub [1]).

III. Outline of the Method

Recall that we are dealing with iteration schemes which permit use of all previous data. We suppose the quantities $f^{[i]}(x_j)$ for all $j<k$, $i=0,\ldots,b_j-1$ available at the k-th step. For simplicity, suppose here that all $b_j=d+1$.

The error formula (2.1) becomes, with $n=(k-1)(d+1)$,

$$f(x)-g(x)=\frac{1}{n!}[f-g]^{[n]}(\xi)\prod_{j=1}^{k-1}(x-x_j)^{d+1}, \tag{3.1}$$

where f and g agree through the d-th derivative at $x_0,\ldots,x_{k-1}$. This formula indicates that there is an ineradicable uncertainty in extrapolating f to x which is proportional to $\prod_{j<k}(x-x_j)^{d+1}$, quite independent of the nature of extrapolation; the given data simply do not determine f much better than that. It therefore hardly seems likely that any iteration scheme could predict the root of f consistently better than could a polynomial scheme.

The above formula may suggest the quantity

$$T_{k,r}=\left(\frac{|x_k-r|}{\prod_{j<k}|x_j-r|^{d+1}}\right)^{1/((k-1)(d+1))}, \tag{3.2}$$

where r is the root of f, as a measure of the success of a particular iteration scheme which predicts x_k as a root. The discussion of Section 6 shows that if $T_{k,r}$ can be shown to be bounded away from zero, then indeed the order of convergence cannot exceed $d+2$. The bulk of our work below consists in showing that, given any iteration scheme which actually converges for a sensible class of functions, we can construct a function f for which the sequence $T_{k,r}$ obtained in using the scheme on f will be bounded away from zero.

In the next section we set forth the precise definitions we require, and show that our requirements are met by standard iteration schemes. We then choose an arbitrary scheme (from the class that has been defined), a function f_0, and a set $x_1,\ldots,x_s$ of starting values for which the scheme is known to converge.

Applying the scheme to f_0, the next approximation x_{s+1} is obtained. We perturb f_0 to a new function f_{s+1} which agrees (through the d-th derivative) with f_0 on $x_1,\ldots,x_s$. The agreement ensures that the scheme yields the same point x_{s+1} applied to f_0 or to f_{s+1}, but the perturbation can be made in such a way that $|f_{s+1}(x_{s+1})|$ is not too small. Applied to f_{s+1}, the scheme gives a new approximation x_{s+2}. We continue the process by choosing f_{s+2} to agree (through the d-th derivative) with f_{s+1} at $x_1,\ldots,x_{s+1}$ in such a way that $|f_{s+2}(x_{s+2})|$ is not too small, and so on.

A sequence of functions is thus generated, and the sequence has as limit a function f, satisfying the conditions of the scheme, with a root r to which the sequence $\{x_k\}$ converges, but in such a way that $T_{k,r}$ remains bounded away from zero.

The program above is carried out in Section 5, and the conclusions for the overall rate convergence are drawn in Section 6.

IV. Iteration Schemes and Convergence

This section is devoted largely to definitions of the concepts we are dealing with. We have not found them in the literature framed with sufficient precision for our purposes, although we are sure that anyone studying iteration schemes at the level of generality we are aiming at would be led to quite similar formulations.

Definition 1. An iteration scheme is a sequence $\{\phi_k: k > s\}$, where ϕ_k is a function of $\sum_{j=1} (b_j+1)$ real variables [written $\phi_k(x_{k-1}, y_{k-1}^{(0)}, \ldots, y_{k-1}^{(b_{k-1}-1)}, x_{k-2}, \ldots, y_1^{(b_1-1)})$], not necessarily defined everywhere, and $b_1, b_2, \ldots$ are positive integers.

Given an iteration scheme, a sufficiently differentiable real function f, and "starting-values" $x_1, \ldots, x_s$, the associated *iteration sequence* is defined recursively by

$$y_j^{(i)} = f^{[i]}(x_j) \quad \text{for} \quad 0 \leqq i < b_j, \ 1 \leqq j < k,$$

and

$$x_k = \phi_k(\ldots \text{as above} \ldots) \quad \text{for} \quad k > s.$$

If x_k is not defined, the iteration sequence terminates at x_{k-1}.

The definition above seems to cover what is usually meant by an iteration scheme. It could generalized further: for example, b_k could be a (positive integer-valued) function of the same variables as x_k; the theorems below still hold with this generalization. Actually, most iterative procedures that people use are what we would call *stationary*. All the functions ϕ_k are the same, with

$$\phi_k(\ldots) = \phi(x_{k-1}, y_{k-1}^{(0)}, \ldots, y_{k-1}^{(d)}, x_{k-2}, \ldots, y_{k-m}^{(d)})$$

for some fixed m, d, and a function ϕ of $(d+2)m$ variables. (We do not require this specialization for our results.) The "one-point iteration function with memory" of Traub [1, p. 8] seems to be a stationary iteration scheme, while his "multipoint iteration function with memory" falls under our general definition of iteration scheme.

Any discussion of the convergence of the iteration sequence generated by a given scheme requires some conditions on the behavior of the function to which it is applied near the zero of the function. We conceive of associating with any scheme quantities relating to the conditions under which it is to be applied, and call the result a *local* iterative scheme.

If a and e are positive numbers, let $S(s, e)$ be the interval $(a-e, a+e)$, and $H(a, e)$ the set of functions representable in $S(a, e)$ by a convergent Taylor series expansion about a. (Our theorems also apply to iterative methods for finding zeros of analytic functions of a complex variable if $S(a, e)$ is the open disc of radius e about a.)

Definition 2. A *local iteration scheme* is an iteration scheme J such that there exist $r_0 \in R$ (the set of real numbers), $\varrho > 0$, and $f_0 \in H(r_0, \varrho)$ with the following properties:

1. $f_0(r_0) = 0 \neq f_0'(r_0)$.

2. For arbitrary $e > 0$, there exist $\delta > 0$ and starting points $x_1, \ldots, x_s$ in $S(r_0, e)$ such that, for any $h \in H(r_0, \varrho)$ with

$$\sum_{j=0}^{\infty} |h^{[j]}(r_0)| \varrho^j / j! < \delta, \tag{4.1}$$

J applied to $f = f_0 + h$ gives an infinite sequence $\{x_k\}$ which is well-defined and lies in $S(r_0, e)$.

If $x_1, \ldots, x_s$ and f are as above, we say that they *fit* the scheme.

Definition 2 imposes very weak restrictions on an iteration scheme, and all practically useful iteration schemes are in fact local iteration schemes. For example, to show that Newton's method is a local iteration scheme, it is sufficient to take $r_0 = 0$, $\varrho = 1$, $f_0(x) = x$, $e < 1$, $\delta = e/4$, and $x_1 = e/2$. Then, for any h satisfying the conditions of Definition 2, we have

$$|h(0)| < \frac{e}{4}$$

and

$$|h'(x)| < \frac{1}{4} \quad \text{for all} \quad x \in S\left(0, \frac{e}{2}\right),$$

so the Newton iteration converges to the (unique) zero of $f = f_0 + h$ in $S\left(0, \frac{e}{2}\right)$. More generally, the fact that all the direct polynomial interpolation schemes mentioned above (including the nonstationary schemes) are local iteration schemes follows from Lemma 1.

Lemma 1. Let $r_0 \in R$, $\varrho > 0$, $e > 0$, $f_0 \in H(r_0, \varrho)$, $f_0(r_0) = 0 \neq f_0'(r_0)$. There is a positive number δ such that, if $b_1 > 0, \ldots, b_m > 0$, $B = \sum_{j=1}^{m} b_j > 1$, $x_1, \ldots, x_m$ are any distinct points in $S(r_0, e)$, $h \in H(r_0, \varrho)$ satisfies (4.1), $f = f_0 + h$, and g is the minimal-degree polynomial such that

$$g^{[i]}(x_j) = f^{[i]}(x_j)$$

for $i = 0, 1, \ldots, b_j - 1$ and $j = 1, \ldots, m$, then g has a unique zero in $S(r_0, e)$.

Proof. Let

$$c = f_0'(r_0)$$

and

$$A = \sum_{k=0}^{\infty} |f_0^{[k]}(r_0)| \frac{(\varrho/2)^k}{k!}. \tag{4.2}$$

We may suppose e to be so small that

$$e < \varrho/16, \tag{4.3}$$

$$e < \frac{c\varrho^2}{320A} \tag{4.4}$$

and, for all $x \in S(r_0, e)$,

$$|1 - f_0'(x)/c| < \tfrac{1}{8}. \tag{4.5}$$

Take any δ satisfying

$$\delta < \frac{ce}{4}. \tag{4.6}$$

It is sufficient to show that

$$|g(r_0)| < \frac{ce}{2} \tag{4.7}$$

and, for all $x \in S(r_0, e)$,

$$|1 - g'(x)/c| < \tfrac{1}{2}. \tag{4.8}$$

Suppose that $k \geqq 0$ and $x \in S(t_0, e)$. From the Taylor series for $f^{[k]}(x)$ about r_0,

$$|f^{[k]}(x)/k!| \leqq \sum_{j=k}^{\infty} \frac{|f^{[j]}(r_0)|}{k!(j-k)!} e^{j-k} \tag{4.9}$$

$$\leqq \sum_{j=k}^{\infty} \frac{|f^{[j]}(r_0)|}{j!} 2^j e^{j-k} \tag{4.10}$$

$$\leqq \sum_{j=k}^{\infty} \frac{|f^{[j]}(r_0)|}{j!} 2^j (\varrho/4)^{j-k} \tag{4.11}$$

$$\leqq (4/\varrho)^k A. \tag{4.12}$$

Since $f_0(r_0) = 0$, (4.1) gives

$$|f(r_0)| = |h(r_0)| < \delta. \tag{4.13}$$

Also, from (2.1) with $n = B$,

$$|f(r_0) - g(r_0)| \leqq \left|\frac{f^B(\xi)}{B!}\right| e^B \tag{4.14}$$

for some $\xi \in S(r_0, e)$. Thus, from (4.12) and (4.13),

$$|g(r_0)| \leqq \delta + (4e/\varrho)^B A. \tag{4.15}$$

Since $B \geqq 2$, (4.7) follows from (4.3), (4.4), (4.6) and (4.15).

From (4.10) with f replaced by h,

$$|h'(x)| \leqq \sum_{j=1}^{\infty} \left|\frac{h^{[j]}(r_0)}{j!}\right| 2^j e^{j-1} \tag{4.16}$$

$$\leqq \sum_{j=1}^{\infty} \left|\frac{h^{[j]}(r_0)}{j!}\right| 2^j (\varrho/2)^{j-1} \tag{4.17}$$

$$\leqq 2\delta/\varrho. \tag{4.18}$$

Hence, from (4.3) and (4.6),

$$|f(x) - f_0'(x)| < c/8 \tag{4.19}$$

for all $x \in S(r, e)$.

By a theorem of Ralston [6],

$$|f'(x) - g'(x)| \leqq B\left|\frac{f^{[B]}(\xi')}{B!}\right| (2e)^{B-1} + \left|\frac{f^{[B+1]}(\eta)}{(B+1)!}\right| (2e)^B, \tag{4.20}$$

where ξ' and η are in $S(r_0, e)$. Thus, from (4.3), (4.4) and (4.12),

$$|f'(x) - g'(x)| \leqq \frac{32Ae}{\varrho^2}\left(\frac{8e}{\varrho}\right)^{B-2}\left(B + \frac{8e}{\varrho}\right) \tag{4.21}$$

$$\leqq \frac{32Ae}{\varrho^2} 2^{2-B}(B + 1/2) \leqq \frac{80Ae}{\varrho^2} \leqq \frac{c}{4} \tag{4.22}$$

for all $x \in S(r_0, e)$. Finally, (4.8) follows from (4.5), (4.19), and (4.22), so the proof is complete.

Lemma 1 shows more than is necessary to establish that direct polynomial interpolation schemes are local iteration schemes. For example, the lemma shows that any function $f_0 \in H(r_0, \varrho)$ with a simple zero at r_0 may be used in Definition 2, but Definition 2 only requires the existence of one such function.

Definition 3. A *locally convergent iteration scheme* is a local iteration scheme J such that the sequence $\{x_k\}$ in Definition 2 converges to a zero of f. The upper (lower) order of convergence of J is the infimum of the upper (lower) order of convergence of such sequences $\{x_k\}$.

A slight modification of Lemma 1 shows that the direct polynomial interpolation schemes are locally convergent (as are the inverse polynomial interpolation schemes), and their order of convergence may be found as indicated in Section 2. However, in the next section we do not need to assume convergence: the results apply to all local iteration schemes.

V. The Main Theorem

Theorem 1. Suppose that J is a local iteration scheme, r_0 and ϱ are as in Definition 2, $R > \varrho$, and $r > 0$. Then there are starting points $x_1, \ldots, x_s$ in $S(r_0, e)$, and $f \in H(r_0, \varrho)$ fitting J, such that f has a simple zero $r \in S(r_0, e)$, J gives a sequence $\{x_k\}$ using evaluations of $f, f', \ldots, f^{[b_k - 1]}$ at x_k (for some $b_k \geqq 1$), and

$$\liminf_{k \to \infty} T_{k,r} \geqq \frac{1}{R}, \tag{5.1}$$

where

$$T_{k,r} = \left(\frac{|x_k - r|}{\prod_{j=1}^{k-1} |x_j - r|^{b_j}} \right)^{1/B_k} \tag{5.2}$$

and

$$B_k = \sum_{j=1}^{k-1} b_j \tag{5.3}$$

for $k > s$.

Proof. Let $f_0 \in H(r_0, \varrho)$ be as in Definition 2,

$$c = |f_0'(r_0)| > 0, \tag{5.4}$$

$$R_0 = (\varrho + R)/2, \tag{5.5}$$

and

$$\lambda = 1 - R_0/R. \tag{5.6}$$

We may suppose e to be so small that

$$e < R_0 - \varrho, \tag{5.7}$$

$$e < \lambda \varrho / 2, \tag{5.8}$$

and, for all $x \in S(r_0, e)$,

$$\left| 1 - \frac{f_0'(x)}{f_0'(r_0)} \right| < \lambda/2. \tag{5.9}$$

Let δ and $x_1, \ldots, x_s$ be as in Definition 2. We may suppose that $f_0(x_i) \neq 0$ for $i \leqq s$ (otherwise replace f_0 by $f_0 + h$, where h is a suitable small perturbation satisfying (4.1)). Choose positive δ_0 and K such that

$$\delta_0 < \delta, \tag{5.10}$$

$$\delta_0 < (1-\lambda)ec, \tag{5.11}$$

$$K < \lambda R_0 c, \tag{5.12}$$

and

$$K < \left(1 - \frac{\varrho + e}{R_0}\right)\delta_0 \tag{5.13}$$

(this is possible by (5.7)).

Let $h_0(x) = h_1(x) = \cdots = h_s(x) = 0$ and, for $k > s$,

$$h_k(x) = h_{k-1}(x) + s_k K R_0^{-B_k} \prod_{j=1}^{k-1} (x - x_j)^{b_j}, \tag{5.14}$$

where

$$s_k = \begin{Bmatrix} +1 & \text{if} \quad f_{k-1}(x_k) \prod_{j=1}^{k-1} (x_k - x_j)^b \geqq 0 \\ -1 & \text{otherwise} \end{Bmatrix}, \tag{5.15}$$

$$f_k(x) = f_0(x) + h_k(x), \tag{5.16}$$

and the points x_k for $k > s$ will be defined below.

We first show that f_k fits the scheme. The proof goes by induction on $k > s$. Suppose, as inductive hypothesis, that $x_1, \ldots, x_{k-1} \in S(r_0, e)$; J applied to f_{k-1} (with starting points $x_1, \ldots, x_s$) gives $x_1, \ldots, x_{k-1}$; and, for $1 \leqq j < k$, f_j fits the scheme and

$$\sum_{i=0}^{\infty} |h_j^{[i]}(r_0)| \varrho^i / i! \leqq \delta_0. \tag{5.17}$$

(Certainly all this is true for $k = s+1$.)

Let x_k be the k-th point in the sequence produced by J when applied to f_{k-1} (with starting points $x_1, \ldots, x_s$). Since f_{k-1} fits the scheme, x_k is defined and in $S(r_0, e)$. By (5.14) and (5.16),

$$f_k^{[i]}(x_j) = f_{k-1}^{[i]}(x_j) \tag{5.18}$$

for $j = 1, \ldots, k-1$ and $i = 0, \ldots, b_j - 1$. Hence, J applied to either f_k or f_{k-1} (with the same starting points $x_1, \ldots, x_s$) produces the same first k points $x_1, \ldots, x_k$. Suppose $1 \leqq j \leqq k$ and $0 \leqq i \leqq B_j$. From (5.14),

$$\frac{|h_j^{[i]}(r_0) - h_{j-1}^{[i]}(r_0)|}{i!} \leqq \binom{B_j}{i} K R_0^{-B_j} e^{B_j - i} \tag{5.19}$$

where

$$\binom{B_j}{i} = B_j! / (i! [B_j - i]!)$$

(and the left side of (5.19) vanishes for $i > B_j$). Thus

$$\sum_{i=0}^{\infty} |h_k^{[i]}(r_0)| \varrho^i / i! \leqq \sum_{i=0}^{\infty} \sum_{j=1}^{k} |h_j^{[i]}(r_0) - h_{j-1}^{[i]}(r_0)| \varrho^i / i! \tag{5.20}$$

$$\leqq K \sum_{j=1}^{k} \sum_{i=0}^{B_j} \binom{B_j}{i} \left(\frac{e}{R_0}\right)^{B_j - i} \left(\frac{\varrho}{R_0}\right)^i = K \sum_{j=1}^{k} \left(\frac{e + \varrho}{R_0}\right)^{B_j} \leqq \delta_0 \tag{5.21}$$

(from (5.13)). Hence f_k fits the scheme J and, by induction, this is true for all $k > 0$.

We have just shown that

$$\sum_{i=0}^{\infty} \sum_{j=1}^{\infty} |h_j^{[i]}(r_0) - h_{j-1}^{[i]}(r_0)| \varrho^i/i! \leqq \delta_0, \tag{5.22}$$

so there exists

$$h(x) = \lim_{j\to\infty} h_j(x) \tag{5.23}$$

for all $x \in S(r_0, e)$. Moreover, $h \in H(r_0, \varrho)$ and

$$\sum_{i=0}^{\infty} |h^{[i]}(r_0)| \varrho^i/i! \leqq \delta_0. \tag{5.24}$$

Let

$$f(x) = f_0(x) + h(x). \tag{5.25}$$

From (5.24), $f \in H(r_0, \varrho)$ and f fits the scheme J. Also, for $1 \leqq j \leqq k$ and $0 \leqq i < b_j$,

$$f^{[i]}(x_j) = f_k^{[i]}(x_j) \tag{5.26}$$

(since $h_k^{[i]}(x_j) = h_{k+1}^{[i]}(x_j) = \cdots$). Thus, J applied to f (with starting points $x_1, \ldots, x_s$) produces the sequence $\{x_k\} \subseteq S(r_0, e)$.

We shall show that f has a unique, simple zero $r \in S(r_0, e)$. Since $|f(r_0)| = |h(r_0)| \leqq \delta_0$ (from (5.24)), (5.11) gives

$$|f(r_0)| \leqq (1-\lambda) ec. \tag{5.27}$$

Also, for all $x \in S(r_0, e)$,

$$|h'(x)| \leqq \sum_{k=1}^{\infty} |h^{[k]}(r_0)| k e^{k-1}/k! \leqq \frac{1}{\varrho} \sum_{k=1}^{\infty} |h^{[k]}(r_0)| \varrho^k/k! \tag{5.28}$$

(since $k < 2^k$ for $k = 1, 2, \ldots$). Thus, from (5.8), (5.11), and (5.24),

$$|h'(x)| \leqq \lambda c/2 \tag{5.29}$$

and, from (5.9) and (5.25),

$$\left| 1 - \frac{f'(x)}{f_0'(r_0)} \right| < \lambda. \tag{5.30}$$

From (5.27) and (5.30) it is immediate that f has a unique, simple zero r in $S(r_0, e)$.

Finally, we must establish (5.1). From (5.26) and (5.30),

$$|x_k - r| \geqq \frac{|f(x_k)|}{c(1+\lambda)} = \frac{|f_k(x_k)|}{c(1+\lambda)} \geqq |f_k(x_k)| (1-\lambda)/c. \tag{5.31}$$

From (5.15), for all $k > s$ we have

$$|f_k(x_k)| \geqq K R_0^{-B_k} \prod_{j=1}^{k-1} |x_k - x_j|^{b_j}. \tag{5.32}$$

Thus, for $k > s$,

$$|x_k - r| \geqq \frac{(1-\lambda) K R_0^{-B_k}}{c} \prod_{j=1}^{k-1} |x_k - x_j|^{b_j}. \tag{5.33}$$

We must show that $x_k \neq r$ for all k. If this is not so, let k be the minimal index such that $x_k = r$. From (5.33), we must have $k \leqq s$. Thus, by the definition of f, $f_0(x_k) = f(x_k) = 0$ for some $k \leqq s$, contradicting one of our assumptions. Hence

$$|x_k - r| > 0$$

for all $k \geqq 1$.

We shall show that, for all $k > s$,

$$|x_k - r| \geqq \frac{K R_0^{-B_k}(1-\lambda)^{1+B_k}}{c} \prod_{j=1}^{k-1} |x_j - r|^{b_j}. \tag{5.34}$$

Suppose, if possible, that (5.34) is false, so

$$|x_k - r| < \frac{K R_0^{-B_k}(1-\lambda)^{1+B_k}}{c} \prod_{j=1}^{k-1} |x_j - r|^{b_j} \tag{5.35}$$

for some $k > s$. For $i = 1, 2, \ldots, k-1$.

$$\prod_{j=1}^{k-1} |x_j - r|^{b_j} / |x_i - r| \leqq e^{B_k - 1} \leqq R_0^{B_k - 1} < R_0^{B_k} \lambda c / K \tag{5.36}$$

(from (5.12)), so

$$\lambda |x_i - r| > \frac{K R_0^{-B_k}}{c} \prod_{j=1}^{k-1} |x_j - r|^{b_j}. \tag{5.37}$$

From (5.35) and (5.37),

$$|x_k - r| < \lambda |x_i - r|, \tag{5.38}$$

so

$$|x_k - x_i| > (1-\lambda)|x_i - r|. \tag{5.39}$$

From (5.33) and (5.39) we obtain (5.34), contradicting (5.35). Hence (5.34) must hold. Finally, from (5.2) and (5.34),

$$T_{k,r} \geqq \left[\frac{K(1-\lambda)}{c}\right]^{1/B_k} \left(\frac{1-\lambda}{R_0}\right), \tag{5.40}$$

so

$$\liminf_{k\to\infty} T_{k,r} \geqq \frac{1-\lambda}{R_0}. \tag{5.41}$$

In view of (5.6), the result follows.

VI. Highest Order of Convergence

Our results on order of convergence are direct consequences of Theorem 1. Recall our definitions (1.2) and (1.3) of upper and lower order of convergence of a sequence, and the definition of upper and lower convergence of a locally convergent iteration scheme (Definition 3).

Theorem 2. Let J be a locally convergent iteration scheme which uses evaluations of $f, f', \ldots, f^{[b_k - 1]}$ at x_k (for some $b_k \geqq 1$) to generate the sequence $\{x_k\}$. Then the upper and lower orders of convergence of J are bounded above by

$$\bar{\lambda} \leqq \limsup_{k\to\infty} \left(\prod_{j=1}^{k} (b_j + 1)\right)^{1/k} \tag{6.1}$$

and

$$\underline{\lambda} \leqq \liminf_{k\to\infty} \left(\prod_{j=1}^{k} (b_j + 1)\right)^{1/k}. \tag{6.2}$$

Proof. With the notation of Theorem 1, there is an f fitting J such that

$$\liminf_{k\to\infty} T_{k,r}^{B_k/(B_k-1)} = \liminf_{k\to\infty} T_{k,r} \geqq \frac{1}{R}. \tag{6.3}$$

Also, in the proof of Theorem 1 we showed that (for a suitable f) x_k is never equal to r, so $T_{k,r} > 0$ for all $k \geqq 1$. Hence, there is a positive constant c such that

$$T_{k,r}^{B_k/(B_k-1)} \geqq c \tag{6.4}$$

for all $k \geqq 1$, so

$$|x_k - r| \geqq c^{B_k-1} \prod_{j=1}^{k-1} |x_j - r|^{b_j}. \tag{6.5}$$

Let

$$\mu_k = -\log(c|x_k - r|). \tag{6.6}$$

From (6.5),

$$\mu_k \leqq \sum_{j=1}^{k-1} b_j \mu_j \tag{6.7}$$

for all $k \geqq 1$. Also, we may suppose c to be sufficiently small that $\mu_1 > 0$. Suppose that, for all $k < m$,

$$\mu_k \leqq \mu_1 \prod_{j=1}^{k-1} (b_j + 1). \tag{6.8}$$

(This is certainly true for $k = 1$.) Then, from (6.7),

$$\begin{aligned}\mu_m &\leqq \mu_1 \sum_{j=1}^{m-1} \left(b_j \prod_{i=1}^{j-1} (b_i + 1) \right) && (6.9)\\ &= \mu_1 \sum_{j=1}^{m-1} \left(\prod_{i=1}^{j} (b_i + 1) - \prod_{i=1}^{j-1} (b_i + 1) \right) \\ &= \mu_1 \left(\prod_{i=1}^{m-1} (b_i + 1) - 1 \right) \\ &\leqq \mu_1 \prod_{i=1}^{m-1} (b_i + 1). && (6.10)\end{aligned}$$

Hence, (6.8) holds for all $k \leqq 1$, by induction. From (6.6) and (6.8),

$$-\log|x_k - r| \leqq \log c + \mu_1 \prod_{j=1}^{k-1} (b_j + 1). \tag{6.11}$$

Since $\mu_1 > 0$ and $b_j \geqq 1$, the results (6.1) and (6.2) follow from (6.11).

Suppose that J is a locally convergent iteration scheme which uses evaluations of $f, f', \ldots, f^{[d]}$ at x_k to generate $\{x_k\}$. The following observations are simple consequences of the above results.

From (6.1), J has upper order of convergence

$$\bar{\lambda} \leqq d + 2. \tag{6.12}$$

The polynomial interpolation schemes show that this result is the best possible.

From (6.11), there is a positive constant A such that

$$|x_k - r| \geqq A^{(d+2)^k} \tag{6.13}$$

for all $k \geqq 1$.

Since

$$|x_1 - r| \prod_{j=1}^{k-1} \frac{|x_{j+1} - r|}{|x_j - r|^{d+2}} = \frac{|x_k - r|}{\prod_{j=1}^{k-1} |x_j - r|^{d+1}}, \tag{6.14}$$

(5.1) shows that

$$\limsup_{k\to\infty} \frac{|x_{k+1} - r|}{|x_k - r|^{d+2}} > 0. \tag{6.15}$$

However, it is possible that

$$\liminf_{k\to\infty} \frac{|x_{k+1} - r|}{|x_k - r|^{d+2}} = 0 \tag{6.16}$$

(for all $x_1, \ldots, x_s$ and f fitting J) if $|x_k - r|$ is occasionally "large".

VII. Some Possible Extensions

While we have given here an upper bound for the order of convergence of iteration schemes which may make use of all previous information, we believe that a similar technique will yield bounds for the more common schemes which use only information from a limited number of previous calues, and that these bounds will be precisely those obtained for the polynomial methods. This subject, which has some complications not shared by the topic of the present paper, will be reported on in the future.

We note that a method may use only some of the values of $f, f', \ldots, f^{[d]}$ at x_k. (For example, methods which use only f and f'' are possible.) Although Theorems 1 and 2 apply to such methods, they are probably not sharp. We do not know of any iterative method using, say, $f^{[c_{k,0}]}(x_k), f^{[c_{k,1}]}(x_k), \ldots, f^{[c_{k,d}]}(x_k)$, where $0 \leqq c_{k,0} < c_{k,1} < \cdots < c_{k,d}$, and having upper order greater than $d+2$.

Iterative methods for finding zeros of f' are of interest in optimization problems. More generally, we may consider methods for finding zeros of $f^{[q]}$, for some fixed $q \geqq 0$. (For some practical methods, see Brent [2].) If condition 1 of Definition 2 is changed to $f_0^{[q]}(r_0) = 0 \neq f_0^{[q+1]}(r_0)$, and the obvious change is made in Definition 3, then Theorems 1 and 2 still hold (in Theorem 1, $f^{[q]}$ has a simple zero r). The proofs are similar to those given above; the main difference is that the term $\prod_{j=1}^{k-1} (x - x_j)^{b_j}$ in (5.14) must be replaced by $(x - x_k)^q \prod_{j=1}^{k-1} (x - x_j)^{b_j}$. However, these results do not appear to be the best possible for $q \geqq 1$. If each $b_j = d+1$ and $0 \leqq q \leqq d+1$, then a plausible conjecture is that the upper order $\bar{\lambda}$ is bounded by

$$\bar{\lambda} \leqq \tfrac{1}{2}(d+2-q) + \sqrt{4q + (d+2-q)^2}, \tag{7.1}$$

where the right side, suggested by consideration of polynomial interpolation schemes, is the positive root of the equation

$$1 = (d+1-q)\, x^{-1} + (d+1) \sum_{k=2}^{\infty} x^{-k}. \tag{7.2}$$

For $q=1$, a partial result in this direction is that (7.1) holds for all locally convergent schemes with the property that the order of convergence of the sequence $\{x_k\}$ in Definition 2 exists (i.e., the upper and lower orders are equal).

Finally, our results apply only to methods for functions of one variable. Not much is known about the maximal order of convergence of iterative methods for functions of several variables, although some plausible conjectures may be made (see Brent [5]).

References

1. Traub, J. F.: Iterative methods for the solution of equations. New Jersey: Prentice-Hall, Inc. 1964.
2. Brent, Richard P.: Algorithms for minimization without derivatives. New Jersey: Prentice-Hall, Inc. 1972.
3. Ostrowski, A. M.: Solution of equations and systems of equations (second edition). New York: Academic Press 1966.
4. Feldstein, M. Alan, Firestone, Roger M.: Hermite interpolatory iteration theory and parallel numerical analysis. Report, Division of Applied Mathematics, Brown University, Providence, Rhode Island (October 1967).
5. Brent, Richard P.: The computational complexity of iterative methods for systems of nonlinear equations. Proc. Complexity Symposium. Yorktown Heights. March 1972. Plenum Press (1972).
6. Ralston, A.: On differentiating error terms. Amer. Math. Monthly **70**, 187–188 (1963).

Philip Wolfe
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights
New York 10598/U.S.A.