THE STRENGTH OF NONSTATIONARY ITERATION

G. W. Wasilkowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15123
(On leave from the University of Warsaw)

August 1979   CUCS-3-79

ABSTRACT

This is the second of three papers in which we study global convergence of iterations using linear information for the solution of nonlinear equations. In Wasilkowski [78] we proved that for the class of all analytic scalar complex functions having only simple zeros there exists no globally convergent stationary iteration using linear information. Here we exhibit a nonstationary iteration using linear information which is globally convergent even for the multivariate and abstract cases. This demonstrates the strength of nonstationary iteration. In Wasilkowski [79] we shall prove that any globally convergent iteration using linear information has infinite complexity even for the class of scalar complex polynomials having only simple zeros.

1. INTRODUCTION

We deal with the iterative solution of a nonlinear operator equation $F = 0$ where F is an analytic multivariate or abstract function having only simple zeros. Most iterations are only locally convergent, i.e., the sequence $\{x_i\}$ generated by an iteration is convergent to a zero α assuming that the starting points are "sufficiently close" to α. In practice it is very hard to verify this assumption and one therefore wants to use globally convergent iterations. All known globally convergent stationary iterations for the class of analytic operators use nonlinear information. Since most iterations of practical interest use linear information, we would like to know whether there exist globally convergent iterations using linear information. From Wasilkowski [78], we know that no <u>stationary</u> iteration using linear information can be globally convergent even for the scalar case. In this paper we pose and affirmatively answer the following problem:

Do there exist <u>nonstationary</u> iterations using linear information which are globally convergent?

We construct a globally convergent nonstationary iteration which is an interpolatory iteration. The i-th step of this iteration requires the computation of $F(x_0), F'(x_0), \ldots, F^{(i-1)}(x_0)$ and the solution of a polynomial equation of degree i-1. Since, in practice, we cannot solve exactly a polynomial equation this iteration is primarily of theoretical interest. It establishes the power of nonstationary over stationary iteration.

In a forthcoming paper, Wasilkowski [79], we shall prove that any iteration using linear information has complexity equal <u>to infinity</u>. More precisely, we shall prove that for any such iteration there exists a scalar polynomial

having only simple zeros such that the cost of computing a better approximation than the starting one is arbitrarily large. This exhibits the important difference between the concepts of convergence and complexity. The class of linear information operators supplies enough knowledge to find a globally convergent iteration but its cost can be arbitrarily high. Hence from a practical point of view, the class of linear information operators is too "weak" for the solution of nonlinear equations. Therefore we have to use stronger (i.e., some nonlinear) information in order to guarantee global convergence and finite complexity.

We summarize the contents of this paper. For the reader's convenience, in Sections 2 and 3 we deal only with iterations without memory. The extension to the case with memory is given in Section 4. In Section 2 we give a very general definition of information and iteration without memory. We recall the definition of globally convergent iterations and define the constant of global convergence. In Section 3 we prove that for the class $\mathfrak{I}_1$ of all analytic operators having simple zeros, the constant of global convergence is no larger than 1/2 for any iteration. Furthermore we proved that only "one-point" iterations can be globally convergent. We also exhibit an iteration which is globally convergent with the constant of global convergence no less than 1/3, which means this iteration has a "large" domain of convergence. In the Appendix we prove global convergence of all iterations we exhibit in Sections 3 and 4.

2. INFORMATION AND ITERATIONS

In this section we introduce a very general definition of <u>information</u> and <u>iteration</u>. We also discuss very briefly the definition of globally <u>convergent iterations</u> (for more detailed discussion see Wasilkowski [78]) and define the <u>constant of global convergence</u>. For the reader's convenience, Sections 2 and 3 deal with iterations without memory. Iterations with memory are considered in Section 4.

Let B_1, B_2 be two Banach spaces over the complex field $\mathbb{C}$ which have dimension

$$N = \dim B_1 = \dim B_2, \quad 1 \le N < +\infty.$$

Let H be the class of all operators $F : D_F \subset B_1 \to B_2$ analytic in D_F and let $\mathfrak{J}$ be a subset of H which consists of operators having only simple zeros. Let $S(F)$ be the set of all zeros of F. Consider the nonlinear equation

(2.1) $F(x) = 0, \ F \in \mathfrak{J}.$

To motivate our definition of an iteration consider first Newton iteration for a scalar case.

Example 2.1

Let $B_1 = B_2 = \mathbb{C}$. For a given approximation x_0 of a solution of $F(x) = 0$ we construct the sequence of approximations $\{x_i\}$ by the formula

(2.2) $x_{i+1} = \phi(x_i; F(x_i), F'(x_i)) = x_i - F'(x_i)^{-1} F(x_i).$

This means that x_{i+1} requires the information $[F(x_i), F'(x_i)]$. Denote $\mathfrak{N}(f, x) = [F(x), F'(x)]$ and $\phi_F(x) = \phi(x; \mathfrak{N}(F, x))$. Thus x_i depends on x_0,

(2.3) $\quad x_i = \psi_F(x_{i-1}) = \psi_F(\psi_F(x_{i-2})) = \ldots = \psi_F \circ \psi_F \circ \ldots \circ \psi_F(x_0)$,

and on the information

(2.4) $\quad \mathfrak{N}_i(F,x_0) = [F(x_0),F'(x_0),F(x_1),F'(x_1),\ldots,F(x_{i-1}),F'(x_{i-1})]$.

We denote (2.3) and (2.4) as

(2.5) $\quad x_i = \varphi_i(x_0; \mathfrak{N}_i(F,x_0))$. ∎

We define an iteration by generalizing the information $\mathfrak{N}_i$ in (2.4) and the function φ_i in (2.5) as follows. Let $L_j : D_{L_j} \subset H \times B_1 \to \mathbb{C}$ be a functional which is linear with respect to the first argument, i.e.,

$L_j(c_1 F_1 + c_2 F_2, x) = c_1 L_j(F_1, x) + c_2 L_j(F_2, x)$ whenever $x \in D_{F_1} \cap D_{F_2}$. We assume that $L_j(F,x)$ is <u>undefined</u> for $x \notin D_F$. Then a <u>linear information operator</u> $\mathfrak{N}$, $\mathfrak{N} = D_{\mathfrak{N}} \subset H \times B_1 \to \mathbb{C}^n$, is defined as

(2.6) $\quad \mathfrak{N}(F,x_0) = [L_1(F,z_1),L_2(F,z_2),\ldots,L_n(F,z_n)]$, $\quad \forall F \in H$, $\forall x_0 \in D_F$,

where $z_1 = x_0$ and $z_{k+1} = \zeta_{k+1}(z_1; L_1(F,z_1),L_2(F,z_2),\ldots,L_k(F,z_k))$ for some functions ζ_j, $j = 1,2,\ldots,n$. Let $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$ be a sequence of linear information operators, $\mathfrak{N}_i : D_{\mathfrak{N}_i} \subset H \times B_1 \to \mathbb{C}^{n_i}$ for $i = 1,2,\ldots$. Let x_0 be an approximation of a solution of (2.1). We construct a sequence of approximations $\{x_i\}$ by the formula

(2.7) $\quad x_i = \varphi_i(x_0; \mathfrak{N}_i(F,x_0))$

where $\varphi_i : D_{\varphi_i} \subset B_1 \times \mathbb{C}^{n_i} \to B_1$ are operators. Then the sequence $\overline{\varphi} = \{\varphi_i\}$ is called an <u>iteration using the information sequence $\overline{\mathfrak{N}}$</u>. Let $\overline{\Phi}(\overline{\mathfrak{N}})$ be the class of all such iterations.

Let $\overline{\pi} = \{\pi_i\}$ be information and let $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}(\overline{\pi})$ be an iteration. We shall say $\overline{\varphi}$ is a __one-point iteration__ iff for any $i = 1,2,\ldots$, the points $z_1, z_2, \ldots, z_{n_i}$ given by (2.6) are equal to x_0. We prove in Theorem 3.2 that any globally convergent iteration is a one-point iteration. We shall say $\overline{\varphi}$ is a __stationary iteration__ iff there exist a linear information operator π and an operator φ such that

$$\pi_i(F, x_0) = \pi(F, x_{i-1}), \quad x_i = \varphi_i(x_0; \pi_i(F, x_0)) = \varphi(x_{i-1}; \pi(f, x_{i-1})), \quad \forall i = 1,2,\ldots,$$

for any $F \in \mathfrak{J}$ and $x_0 \in D_F$. We shall say $\overline{\varphi}$ is __nonstationary__ iff φ is not a stationary iteration.

In most papers iterations are defined in a way that exhibits the dependence of x_i on some previously computed x_j, $j < i$. For example, see Traub [64], Ortega and Rheinboldt [70], and Traub and Woźniakowski [78]. Our definition of an iteration generalizes these definitions. In Wasilkowski [79] we shall establish a negative result for even this very general class of iterations.

We are now ready to define global convergence of an iteration. Let $J(\alpha, R) \overset{df}{=} \{x \in B_1 : \|x - \alpha\| < R\}$ denote the ball of center α and radius R. For any F, $F \in \mathfrak{J}$, and any $\alpha \in S(F)$ define

$$R_F(\alpha, \partial D_F) = \inf_{x \in \partial D_F} \|\gamma - x\|$$

as the distance of α to the boundary ∂D_F of the domain D_F. Let

$$B(b, F) = \bigcup_{\alpha \in S(F)} J(\alpha, bR_F(\alpha))$$

where $b > 0$. Let $\overline{\varphi} \in \overline{\Phi}(\overline{\pi})$ be an iteration. Let A be the set of real numbers a such that for any $F \in \mathfrak{J}$ and any $x_0 \in B(a, F)$ the sequence $\{x_i\}$, $x_i = \varphi_i(x_0; \pi_i(F, x_0))$

is well-defined and $\lim\limits_{i\to\infty} x_i \in S(F)$. We shall say $c = c(\overline{\varphi},\mathfrak{Y})$ is the <u>constant of global convergence of $\overline{\varphi}$ for the class $\mathfrak{Y}$</u> iff

$$c = \begin{cases} \sup A & \text{if } A \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if there exists F from $\mathfrak{Y}$ with finite $R_F(\alpha)$ then $c(\overline{\varphi},\mathfrak{Y}) \in [0,1]$. If $R_F(\alpha) = +\infty$ for any $F \in \mathfrak{Y}$, then $c(\overline{\varphi},\mathfrak{Y})$ is zero or infinity (with the convention $0\cdot\infty = 0$). The set $B(c,F)$ is a convergence domain of $\overline{\varphi}$ for F since taking any starting point from $B(c,F)$ we get convergence of $\{x_i\}$ to a solution of $F = 0$.

Definition 2.1

We shall say that an iteration $\overline{\varphi}$ is <u>globally convergent</u> for the class $\mathfrak{Y}$ iff

$$c(\overline{\varphi},\mathfrak{Y}) > 0. \qquad\qquad \blacksquare$$

3. GLOBALLY CONVERGENT ITERATION

In this section we study global convergence of iterations for the class $\mathfrak{J}_1$ of all analytic operators F from H having only simple zeros. We prove some general properties of globally convergent iterations and we also exhibit an iteration which is globally convergent.

Recall that N, $N < +\infty$, is the dimension of the spaces B_1 and B_2. (See also Remark 3.1.) Let $\{a_{1,1}, a_{1,2}, \ldots, a_{1,N}\}$ be a basis for B_i, $i = 1,2$. Let $t = \sum_{j=1}^{N} t_j a_{1,j}$. We begin with

Theorem 3.1

For any iteration $\bar{\varphi}$

$$c(\bar{\varphi}, \mathfrak{J}_1) \leq 1/2. \qquad \blacksquare$$

Proof

Suppose on the contrary that there exists an $\bar{\pi}$ and a $\bar{\varphi} \in \bar{\Phi}(\bar{\pi})$ with $c = c(\bar{\varphi}, \mathfrak{J}_1) > 1/2$. Let

$$F(t) = (t_1^2 - 1) a_{2,1} + \sum_{j=2}^{N} t_j a_{2,j}.$$

Then F is an entire operator having only simple zeros, $\alpha_1 = a_{1,1}$ and $\alpha_2 = -a_{1,1}$. Consider two operators

$$F_i(t) = \begin{cases} F(t) & \text{for } \| t - \alpha_i \| < 2 \| a_{1,1} \|, \\ \text{undefined} & \text{otherwise,} \end{cases}$$

$i = 1,2$. Then $F_1, F_2 \in \mathfrak{J}_1$, $S(F_i) = \{\alpha_i\}$ and $R_{F_i}(\alpha_i) = 2 \| a_{1,1} \|$, $i = 1,2$.

Since $c > 1/2$, then $x_0 = y_0 = 0 \in B(c,F_1) \cap B(c,F_2)$. Let $x_1 = \varphi_1(x_0; \mathfrak{N}_1(F_1,x_0))$ and $y_1 = \varphi_1(y_0; \mathfrak{N}_1(F_2,y_0))$. Then $\{x_i\}$ and $\{y_i\}$ are well-defined and tend to a_{11} and $-a_{11}$ respectively. Thus, there exists an index i_0 such that $y_i = x_i$ for $i = 0,1,\ldots,i_0$ and $y_{i_0+1} \neq x_{i_0+1}$. Since $\mathfrak{N}_{i_0+1}(F_1,x_0) = \mathfrak{N}_{i_0+1}(F_2,y_0)$ we get $y_{i_0+1} = x_{i_0+1}$ which is a contradiction. ∎

We now give a necessary condition on an iteration to be globally convergent for the class $\mathfrak{I}_1$.

Theorem 3.2

If $\overline{\varphi}$ is a globally convergent iteration for the class $\mathfrak{I}_1$ then $\overline{\varphi}$ is a one-point iteration. ∎

Proof

The proof of Theorem 3.2 is similar to the proof of Theorem 4.1 in Wasilkowski [78]. Therefore we only sketch the proof.

Suppose on the contrary that there exists $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$ and $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}(\overline{\mathfrak{N}})$ which is globally convergent for $\mathfrak{I}_1$ and is not one-point. Let

$$F(t) = F\left(\sum_{s=1}^{N} t_s a_{1,s}\right) = (t_1 - 1)a_{2,1} + \sum_{s=2}^{N} t_s a_{2,s}.$$

Then F is an entire function from $\mathfrak{I}_1$ having only one zero $\alpha = a_{1,1}$. Let $x_0 = 0$. Since $R_F(\alpha) = +\infty$, there exist integers k and j_0, $j_0 \in [1,n_k]$ such that $z_j^i = 0$ for any $i < k$ and $j = 1,\ldots,n_i$, and $z_{j_0}^k \neq 0$. (The points z_j^i are defined by $\mathfrak{N}_i$ and F.) Let $m = N \sum_{i=1}^{k} n_i$. For $\gamma_1,\gamma_2,\ldots,\gamma_{m+1} \in \mathbb{C}$ define

$$W_s(t) = a_{2,s} \sum_{r=1}^{m+1} \gamma_r t_s^r, \quad s = 1,2,\ldots,N.$$

Let $W(t) = \sum_{i=1}^{N} W_s(t)$. Then there exist $\gamma_1, \gamma_2, \ldots, \gamma_{m+1}$, such that $W \not\equiv 0$ and $\mathfrak{N}_i(F+W,0) = \mathfrak{N}_i(F,0)$ for $i = 1,2,\ldots,k$. For $\sigma > 0$, define

$$F_\sigma(t) = \begin{cases} F(t) + \frac{1}{\sigma} W(t) & \text{for } \|t\| < |z_{j_0}^k|, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Of course, $F_\sigma \in \mathfrak{J}_1$ for sufficiently small σ. Furthermore, $F_\sigma(t) = -a_{2,1} + \sum_{s=2}^{N} a_{2,s} [t_s + \frac{1}{\sigma} \sum_{r=1}^{m+1} \gamma_r t_s^r]$ has a zero which tends to zero as σ goes to zero.

Thus, for sufficiently small σ, $x_0 = 0 \in B(c(\overline{\varphi}, \mathfrak{J}_1), F_\sigma)$ which means that $y_i = \varpi_i(0; \mathfrak{N}_i(F_\sigma,0))$ are well-defined for any $i = 1,2,\ldots$. Since $\mathfrak{N}_i(F_\sigma,0) = \mathfrak{N}_i(F,0)$, $z_{j_0}^k \notin D_{F_\sigma}$. This means that y_k is undefined which is a contradiction. ∎

We now exhibit a globally convergent nonstationary iteration for the class $\mathfrak{J}_1$. Let

$$\mathfrak{N}_{n,0}(F,x_0) = [F(x_0), F'(x_0), \ldots, F^{(n-1)}(x_0)].$$

Thus $\mathfrak{N}_n$ is a linear information operator. Let $\overline{\mathfrak{N}}_0 = \{\mathfrak{N}_{n,0}\}$. For given x_0, $x_0 \in D_F$, define

$$(3.1) \quad W_{n,0}(x) = W_{n,0}(x,x_0) = F(x_0) + F'(x_0)(x-x_0) + \ldots + \frac{1}{(n-1)!} F^{(n-1)}(x_0)(x-x_0)^{n-1}.$$

Let $S(W_{n,0})$ be the set of all zeros of $W_{n,0}$. Similarly to Traub and Woźniakowski [76] we define the <u>interpolatory</u> iteration $\overline{I}_0 = \{I_{n,0}\}$, $I_{n,0}(x_0, \mathfrak{N}_{n,0}(F,x_0)) = x_n \in S(W_{n,0})$ with some criterion of the choice of a zero x_n. Thus, for different criteria we obtain different interpolatory iterations. We propose the following criterion for the choice of a zero x_n of $W_{n,0}$. Let

$$(3.2) \quad d_{n,0}(x_0) = \begin{cases} \text{dist}(x_0, S(W_{n,0})) & \text{if } S(W_{n,0}) \neq \emptyset \\ +\infty & \text{otherwise.} \end{cases}$$

Define

$$(3.3) \quad G_{n,0}(x_0) = \{z \in S(W_{n,0}) : \| z - x_0 \| \leq d_{n,0}(x_0) + \frac{1}{\sqrt{n}}\}.$$

Let $B_1 = \lim\{a_1, a_2, \ldots, a_N\}$. For any $x \in B_1$, $x = \sum_{j=1}^{N} t_j a_j$, $t_j \in \mathbb{C}$, we define $\tilde{x} \in R^{2N}$ as

$$\tilde{x} = [\text{re}(t_1), \text{im}(t_1), \text{re}(t_2), \text{im}(t_2), \ldots, \text{re}(t_N), \text{im}(t_N)].$$

Let $<$ be the lexical order on R^{2N}, i.e., for any b_1, $b_2 \in R^{2N}$, $b_1 \neq b_2$, $b_i = (b_{i,1}, b_{i,2}, \ldots, b_{i,2N})$, we write $b_1 < b_2$ iff there exists an integer $k \in [1, 2N]$ such that $b_{1,k} < b_{2,k}$ and $b_{1,i} = b_{2,i}$ for $i < k$. Then we can define an order on B_1. Namely,

$$(3.4) \quad x_1 < x_2 \text{ iff } \tilde{x}_1 < \tilde{x}_2.$$

If $S(W_n)$ is nonempty then $G_{n,0}(x_0)$ has a minimal element z^*, $z^* = z^*(n, x_0, \mathfrak{R}_{n,0}(F, x_0))$ in the sense of order $<$. We define

$$I_{n,0}(x_0; \mathfrak{R}_{n,0}(F, x_0)) = \begin{cases} z^*(n, x_0, \mathfrak{R}_{n,0}(F, x_0)) & \text{if } S(W_{n,0}) \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\overline{I}_0 = \{I_{n,0}\}$ is a nonstationary interpolatory iteration and $\overline{I}_0 \in \overline{\mathfrak{I}}(\overline{\mathfrak{R}}_0)$.

<u>Theorem 3.3</u>

Iteration $\overline{I}_0$ is globally convergent for the class $\mathfrak{I}_1$ and

$$(3.5) \quad c(\overline{I}_0, \mathfrak{I}_1) \geq \frac{1}{3}.$$

Proof

See Appendix. ∎

We do not know whether (3.5) is sharp. We also do not know what is the maximal constant of global convergence of iterations for the class $\mathfrak{I}_1$. From Theorem 3.1 and (3.5) we can only conclude

Corollary 3.1

$$\frac{1}{3} \le \sup_{\overline{\mathfrak{N}}} \ \sup_{\overline{\varphi} \in \overline{\Phi}(\overline{\mathfrak{N}})} c(\overline{\varphi}, \mathfrak{I}_1) \le \frac{1}{2}. \qquad \blacksquare$$

Remark 3.1

In this section we have assumed that the spaces B_1 and B_2 are finite dimensional. We need this assumption in order to assure that $\mathfrak{N}_{n,0}(F,x) = [F(x), F'(x), \ldots, F^{(n-1)}(x)]$ is representable by a finite number of linear functionals. For the infinite dimension case, $\dim B_1 = \dim B_2 = +\infty$, $\mathfrak{N}_{n,0}(F,x)$ is representable by infinite number of linear functionals. Defining the iteration $\overline{I}_0$ analogously to (3.1)-(3.4), it is possible to verify that Theorem 3.3 still holds. ∎

4. ITERATIONS WITH MEMORY

In this section we extend all previous results to iterations with memory. Since we do not know any globally convergent iteration with memory for the multivariate case we assume in this section that $B_1 = B_2 = \mathbb{C}$, i.e., $N = 1$. We present two globally convergent iterations for different classes. The first of these is the generalization of $\overline{I}_{\varphi}$; it is globally convergent for the class $\mathfrak{I}_1$. The second is based on increasing the size of memory; it is globablly convergent for the class of all entire functions from $\mathfrak{I}_1$.

Let m, $m > 0$, be an integer. Let L_j be a functional defined as in Section 2. A <u>linear information operator with memory</u> $\mathfrak{N}$, $\mathfrak{N} : D_{\mathfrak{N}} \subset H \times \mathbb{C}^{m+1} \to \mathbb{C}^n$, is defined as

$$(4.1) \quad \mathfrak{N}(F,x_0,x_{-1},\ldots,x_{-m}) = [L_1(F,z_1),L_2(F,z_n),\ldots,L_n(F,z_n)], \quad \forall F \in H,$$

$$\forall x_0,x_{-1},\ldots,x_{-m} \in D_F$$

where $z_1 = x_0$, $z_2 = x_{-2},\ldots,z_{m+1} = x_{-m}$ and $z_{k+1} = \zeta_{k+1}(z_1,z_2,\ldots,z_{m+1};$ $L_1(F_1,z_1),L_2(F,z_2),\ldots,L_k(F,z_k))$ for some functions ζ_j, $j = m+2,m+3,\ldots,n$. Let

$$\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$$

be a sequence of linear information operators with memory, $\mathfrak{N}_i : D_{\mathfrak{N}_i} \subset H \times \mathbb{C}^{m+1} \to \mathbb{C}$. Let $x_0,x_{-1},\ldots,x_{-m}$ be distinct approximations of a solution of (2.1). We construct a sequence of approximations $\{x_i\}$ by the formula

$$(4.2) \quad x_i = \varphi_i(x_0,x_{-1},\ldots,x_{-m}; \mathfrak{N}_i(F,x_0,x_{-1},\ldots,x_{-m}))$$

where $\varphi_i : D_{\varphi_i} \subset \mathbb{C}^{m+1+n_i} \to \mathbb{C}$ are functionals. Then $\overline{\varphi} = \{\varphi_i\}$ is called an <u>iteration with memory using information sequence $\overline{\mathfrak{N}}$</u>. Let $\overline{\Phi}_m(\overline{\mathfrak{N}})$ be the class of all such iterations.

We now extend the definition of global convergence. For any iteration $\bar{\varphi}, \bar{\omega} \in \bar{\Phi}_m(\overline{\mathfrak{M}})$, let A be the set of all real numbers a such that for any $F \in \mathfrak{J}$ and any distinct points $x_0, x_{-1}, \ldots, x_{-m}$ satisfying $x_0 \in B(a, F)$ and $|x_{-j} - x_0| < c \text{ dist}(x_0, S(F))$, the sequence $\{x_i\}$, $x_i = \varphi_i(x_0, x_{-1}, \ldots, x_{-m};$ $\mathfrak{N}_i(F, x_0, x_{-1}, \ldots, x_{-m}))$, is well-defined and $\lim_{i \to \infty} x_i \in S(F)$. We shall say $c = c(\bar{\varphi}, \mathfrak{J})$ is a <u>constant of global convergence of $\bar{\varphi}$ for the class $\mathfrak{J}$</u> iff

$$c = \begin{cases} \sup A & \text{if } A \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

<u>Definition 4.1</u>

We shall say that an iteration $\bar{\varphi}$ is <u>globally convergent</u> for the class $\mathfrak{J}$ iff

$$c(\bar{\varphi}, \mathfrak{J}) > 0. \qquad \blacksquare$$

For this case it can also be shown that for the class $\mathfrak{J}_1$ defined in Section 3,

$$c(\bar{\varphi}, \mathfrak{J}_1) \leq \frac{1}{2}, \quad \forall \bar{\varphi} \in \Phi_m(\overline{\mathfrak{M}}).$$

An example of globally convergent iteration for the class $\mathfrak{J}_1$ is provided by the generalization of $\bar{I}_0$ iteration. Namely, let

$$\mathfrak{N}_{n,m}(F, x_0, x_{-1}, \ldots, x_{-m}) = [F(x_0), F'(x_0), \ldots, F^{(n-1)}(x_0), F(x_{-1}), F'(x_{-1}), \ldots,$$
$$F^{(n-1)}(x_{-1}), \ldots, F(x_{-m}), F'(x_{-m}), \ldots, F^{(n-1)}(x_{-m})]$$
$$(= [\mathfrak{N}_n(f, x_0), \mathfrak{N}_n(F, x_{-1}), \ldots, \mathfrak{N}_n(F, x_{-m})]).$$

Thus $\mathfrak{N}_{n,m} : D_{\mathfrak{N}_{n,m}} \subseteq H \times \mathbb{C}^{m+1} \to \mathbb{C}^{r+1}$ where

(4.3) $\quad r = r(n, m) = n(m+1) - 1.$

Let now $\bar{\mathfrak{N}}_m = \{\mathfrak{N}_{n,m}\}_n$. For distinct $x_0, x_{-1}, \ldots, x_{-m}$ let $W_{n,m} = W_{n,m}(\cdot; x_0, x_{-1}, \ldots, x_{-m})$ be an interpolatory polynomial of degree at most r satisfying

$$(4.4) \quad W_{n,m}^{(k)}(x_{-j}) = F^{(k)}(x_{-j}) \text{ for } k = 0, 1, \ldots, n-1 \text{ and } j = 0, 1, \ldots, m.$$

Let $S(W_{n,m})$ be the set of all zeros of $W_{n,m}$ and let

$$(4.5) \quad d_{n,m}(x_0) = \text{dist}(x_0, S(W_{n,m})).$$

Define

$$(4.6) \quad G_{n,m}(x_0) = \{z \in S(W_{n,m}) : |z - x_0| \le d_{n,m}(x_0) + \frac{1}{\sqrt{r+1}}\}$$

and let $z^* = z^*(n, m, x_0, \mathfrak{N}_{n,m}(F, x_0, x_{-1}, \ldots, x_{-m}))$ be the minimal element from $G_{n,m}(x_0)$ in the sense of the order relation $<$. Then

$$I_{n,m}(x_0, x_{-1}, \ldots, x_{-m}; \mathfrak{N}_{n,m}(F, x_0, x_{-1}, \ldots, x_{-m})) \overset{df}{=} \begin{cases} z^* & \text{if } S(W_{n,m}) \ne \emptyset \\ 0 & \text{otherwise,} \end{cases}$$

is a functional, $I_{n,m} : \mathbb{C}^{m+1+r+1} \to \mathbb{C}$ and

$$\bar{I}_m = \{I_{n,m}\}_n$$

is an iteration with memory, $\bar{I}_m \in \bar{\Phi}_m(\bar{\mathfrak{N}})$.

Theorem 4.1

For any m, $m \ge 1$, iteration $\bar{I}_m$ is globally convergent for the class $\mathfrak{F}_1$ and

$$c(\bar{I}_m, \mathfrak{F}_1) \ge \frac{1}{4}. \qquad \blacksquare$$

Proof

See Appendix. $\qquad \blacksquare$

We now present a globally convergent iteration for the class $\mathfrak{J}_2$ of all entire functions. For fixed n, $n \geq 2$, let

$$\mathfrak{N}_i(F,x_0) = [F(x_0),\ldots,F^{(n-1)}(x_0),F(x_1),\ldots,F^{(n-1)}(x_1),\ldots,F(x_{i-1}),\ldots,F^{(n-1)}(x_{i-1})].$$

Define

$$x_i = \varphi_{i,n}(x_0,\mathfrak{N}_i(F,x_0)) = I_{n,i-1}(x_{i-1},x_{i-2},\ldots,x_0;\ \mathfrak{N}_i(F,x_0)).$$

Then iteration $\overline{\varphi}_n$, $\overline{\varphi}_n = \{\varphi_{i,n}\}_i$, requires the computation of $F(x_{i-1})$, $F'(x_{i-1}),\ldots,F^{(n-1)}(x_{i-1})$ per step.

Theorem 4.2

For any n, $n \geq 2$, the iteration $\overline{\varphi}_n$ is globally convergent for the class $\mathfrak{J}_2$ of all entire functions from $\mathfrak{J}_1$, i.e., $c(\overline{\varphi}_n,\mathfrak{J}_2) = +\infty$. ∎

This theorem can be proven analogously to Theorem 3.3 and therefore its proof is omitted.

APPENDIX

We prove Theorems 3.3 and 4.1. We begin with the scalar case, i.e., $B_1 = B_2 = \mathbb{C}$, $N = 1$. Let $F \in \mathfrak{I}_1$ and n, m ($m \geq 0$) be integers. Let

(A.1) $\quad R_{n,m}(x) = R_{n,m}(x; x_0, x_{-1}, \ldots, x_{-m}) = F(x) - W_{n,m}(x)$.

Then

(A.2) $\quad R_{n,m}(x) = \prod_{j=0}^{m} (x - x_{-j}) \int_0^1 \int_0^{t_0} \ldots \int_0^{t_{r-1}} F^{(r+1)}(x_0 + t_{n-1}(x_{-1} - x_0) + \ldots + t_r(x - x_{-m})) dt_r dt_{r-}$

$$\ldots dt_0.$$

where

$$r = r(n,m) = n(m+1) - 1.$$

For any $F \in \mathfrak{I}_1$ and $\alpha \in S(F)$ define

(A.3) $\quad A_k(\Gamma) = A_k(\Gamma, \alpha, F) = \sup_{x \in J(\alpha, \Gamma)} \left| F'(\alpha)^{-1} \frac{F^{(k)}(x)}{k!} \right|$ for $k = 2, 3, \ldots$.

Let $q \in (0,1)$ and $x_0, x_{-1}, \ldots, x_{-m} \in J(\alpha, \Gamma)$ be distinct. Following the proof technique of Theorem 2.1 of Traub and Woźniakowski [76] we prove

Lemma A.1

If

(A.4) $\quad A_2(\Gamma) q \Gamma + \dfrac{A_{r+1}(\Gamma)(1+q)^{r+1} \Gamma^r}{q} < 1$

then the polynomial $W_{n,m}$ has a zero $z = z(x_0, x_{-1}, \ldots, x_{-m})$ such that

(A.5) $\quad |z - \alpha| \leq q|x_0 - \alpha|$. ∎

Proof

We can assume $x_0 \neq \alpha$. Since $F(x) = F'(\alpha)(x - \alpha) + R_{2,0}(x; \alpha)$ and $F(x) = W_{n,m}(x) - R_{n,m}(x)$ we get

(A.6) $x - F'(\alpha)^{-1} W_{n,m}(x) = H(x) \stackrel{df}{=} \alpha + F'(\alpha)^{-1} \{ R_{n,m}(x) - R_{2,0}(x; \alpha) \}.$

Thus $W_{n,m}$ has a zero in $J(\alpha, q|x_0 - \alpha|)$ iff there exists a fixed point of

(A.7) $x = H(x)$, $x \in J(\alpha, q|x_0 - \alpha|).$

We first verify that $H(J(\alpha, q|x_0 - \alpha|)) \subset J(\alpha, q|x_0 - \alpha|)$. Let $|x - \alpha| < q|x_0 - \alpha|$.
From (A.2) and (A.3) we get

$$|\alpha - H(x)| \le A_{r+1}(\Gamma) \prod_{j=0}^{m} |x - x_{-j}|^n + A_2(\Gamma)|x - \alpha|^2 \le A_{r+1}(\Gamma) \prod_{j=0}^{m} (|x - \alpha| + |\alpha - x_{-j}|)^n + A_2(\Gamma)|x - \alpha|^2$$

$$< A_{r+1}(\Gamma)(q|x_0 - \alpha| + |x_0 - \alpha|)^n \prod_{j=0}^{m} (q\Gamma + \Gamma)^n + A_2(\Gamma)q|x_0 - \alpha|q\Gamma$$

$$\le q|x_0 - \alpha| \{ A_{r+1}(\Gamma)(1+q)^{mn}\Gamma^{mn} \frac{(1+q)^{n} \Gamma^{n-1}}{q} + A_2(\Gamma)q\Gamma \}$$

$$= q|x_0 - \alpha| \{ \frac{A_{r+1}(\Gamma)(1+q)^{r+1}\Gamma^r}{q} + A_2(\Gamma)q\Gamma \} < q|x_0 - \alpha|.$$

Thus $H(J(\alpha, q|x_0 - \alpha|)) \subset J(\alpha, q|x_0 - \alpha|)$. From the Brouwer fixed point theorem, see e.g., Ortega and Rheinboldt [70, p. 161], it follows that there exists a zero of the polynomial $W_{n,m}$ in $J(\alpha, q|x_0 - \alpha|)$. This proves (A.5). ∎

Let D_F be the domain of F, $F \in \mathfrak{J}$, such that $D_F \ne \mathbb{C}$. Then for any $\alpha \in S(F)$, $R = R_F(\alpha)$ is finite. From Cauchy's formula there exists a constant $M = M(F, \alpha)$ such that

(A.7) $\left| \dfrac{F^{(k)}(\alpha)}{k!} \right| \le \dfrac{M}{R^k}$, $k = 0, 1, \dots$.

Traub and Woźniakowski [76] established

Lemma A.2

Let $\Gamma < \dfrac{R_F(\alpha)}{2}$. Then there exists n_1 such that (A.4) holds for every

$n \geq n_1$ with

$$q = \frac{R_F(\alpha)}{M(\alpha,F)r} \quad \text{and} \quad r = n(m+1)-1. \qquad \blacksquare$$

For fixed L, α, F define

$$(A.8) \quad \gamma(n,m,K) = \sup_{x_0 \in J(\alpha,L)} \ \sup_{\zeta, x_{-1}, \ldots, x_{-m} \in J(x_0,K)} |F(\zeta) - W_{n,m}(\zeta;\ x_0, x_{-1}, \ldots, x_{-m})|.$$

Lemma A.3

    (i)   If $m = 0$ and $L < \dfrac{R_F(\alpha)}{3}$ then

$$\lim_{n \to \infty} \gamma(n,0,\frac{R_F(\alpha)}{3}) = 0.$$

    (ii)   If $m \geq 1$ and $L < \dfrac{R_F(\alpha)}{4}$ then

$$\lim_{n \to \infty} \gamma(n,m,\frac{R_F(\alpha)}{4}) = 0. \qquad \blacksquare$$

Proof

    Let $m = 0$. Denote $R = R_F(\alpha)$. From (A.2) we get

$$\gamma(n,0,\frac{R}{3}) \leq (\frac{R}{3})^n \sup_{x_0 \in J(\alpha,L)} \ \sup_{\zeta \in J(x_0,\frac{R}{3})} \left|\frac{F^{(n)}(\zeta)}{n!}\right| \leq (\frac{R}{3})^n \sup_{\zeta \in J(\alpha,L+\frac{R}{3})} \left|\frac{F^{(n)}(\zeta)}{n!}\right|.$$

From Traub and Woźniakowski [76] we know that

$$\left|\frac{F^{(n)}(\zeta)}{n!}\right| \leq \frac{M}{R^n} \frac{1}{\left(1 - \frac{|\zeta - \alpha|}{R}\right)^{n+1}}$$

which implies

$$(A.9) \quad \gamma(n,0,\tfrac{R}{3}) \leq \frac{M}{R^n} \frac{1}{\left(1-\frac{3L+R}{3R}\right)^{n+1}} (\tfrac{R}{3})^n = \frac{3MR}{2R-3L} (\frac{R}{2R-3L})^n.$$

Since $R/(2R-3L) < 1$, the right-hand side of (A.9) tends to zero. Hence (i) is proven.

For $m \geq 1$, it can be similarly shown that $\gamma(n,m,R/4) = 0((\frac{R}{2R-4L})^n)$. Since $R/(2R-4L) < 1$, we get (ii). This completes the proof. ∎

Lemma A.3 states that the polynomials $W_{n,m}$ uniformly approximate F. It is also easy to show that $W'_{n,m}$ tends to F'. Since α is a simple zero of F, then either re $F'(x)$ or im $F'(x)$ is distinct from zero in a ball $J(\alpha,\Gamma^*)$ for some $\Gamma^*, \Gamma^* = \Gamma^*(\alpha,F) > 0$. Without loss of generality we can assume that re $F' > 0$. Then for sufficiently large n, re $W'_{n,m}(x)$ also does not vanish in $J = J(\alpha,\frac{\Gamma^*}{2})$. Thus, re $W_{n,m}(x)$ has at most one solution in J. From this and Lemma A.2 we conclude that $W_{n,m}$ has exactly one zero in y. This is summarized in

Lemma A.4

There exists Γ^*, $\Gamma^* = \Gamma^*(\alpha,F) > 0$, and an integer n_2, $n_2 = n_2(\alpha,F)$, such that for any $n \geq n_2$ the polynomial $W_{n,m}$ has exactly one zero in $J(\alpha,\Gamma^*/2)$. ∎

Proof of Theorem 3.3 for N = 1

Consider first $F \in \mathfrak{I}_1$ with $D_F \neq \mathbb{C}$, i.e., $R_F(\alpha) < +\infty$, $\forall \alpha \in S(F)$. Let $x_0 \in B(\frac{1}{3},F)$. Define $S(F,x_0) = \{\alpha \in S(F) : |x_0-\alpha| = \text{dist}(x_0,S(F))\} = \{\alpha_0,\alpha_1,\ldots,\alpha_s\}$ where α_0 is the minimal element from $S(F,x_0)$ in the sense of (3.4). Of course, $|x_0-\alpha_0| < \frac{R_F(\alpha_0)}{3}$. We prove that the sequence $\{x_n\}$ generated by $\overline{I}_0$ tends to α_0.

Let ϵ be a positive number such that $\epsilon \leq \frac{1}{2} \min\{\Gamma^*(\alpha,F) : \alpha \in S(F,x_0)\}$. From Lemmas A.1, A.2 and A.4 there exists n_1, $n_1 = n_1(\epsilon)$ such that for $n \geq n_1$

$$(1+q_n)|x_0-\alpha_0| < \frac{R_F(\alpha_0)}{3}$$

where

$$q_n \overset{df}{=} \frac{1}{n} \max_{\alpha \in S(F,x_0)} \frac{R_F(\alpha)}{M(\alpha,F)}$$

and the polynomial $W_{n,0}(x; x_0)$ has zeros z_j^n which satisfy

(A.10) $\quad |z_j^n - \alpha_j| < q_n |x_0 - \alpha_0| \leq \epsilon \quad$ for $j = 0, 1, \ldots, s$.

Furthermore z_j^n is the only zero of $W_{n,0}(x; x_0)$ which belongs to $J(\alpha_j; \epsilon)$. Let

$$B(x_0) \overset{df}{=} J(x_0, \frac{R_F(\alpha_0)}{3}) \setminus \bigcup_{\alpha \in S(F,x_0)} J(\alpha, \epsilon)$$

and

$$\sigma = \sigma(\epsilon) = \inf\{|f(x)| : x \in B(x_0)\}.$$

Of course $\sigma > 0$. From Lemma A.3 we get

$$\sup\{|f(x) - W_{n,0}(x; x_0)| : x \in J(x_0, \frac{R_F(x_0)}{3})\} \leq \frac{\sigma}{2}$$

for sufficiently large n, $n \geq n_2 = n_2(\epsilon)$. Note that $|W_{n,0}(x)| \geq |F(x)| - |F(x) - W_{n,0}(x)| \geq \frac{\sigma}{2} > 0$. Thus $W_{n,0}$ does not have a zero in $B(x_0)$. Let $n_3 = \max\{n_1, n_2\}$. For $n \geq n_3$, $x_n \in \bigcup_{\alpha \in S(F)} \overline{J(\alpha, \epsilon)}$. Note that $h \overset{df}{=} \text{dist}(x_0, S(F) \setminus S(F,x_0)) > |x_0 - \alpha_0|$. Then for an arbitrary $\epsilon \in (0, \frac{1}{2}(h - |x_0 - \alpha_0|))$ we get

$$x_n \in \bigcup_{\alpha \in S(F,x_0)} \overline{J(\alpha, \epsilon)}$$

for sufficiently large n. Therefore if $S(F,x_0) = \{\alpha_0\}$ is a singleton set then $|x_n - \alpha_0| \leq \epsilon$ which implies that $\lim_{n \to \infty} x_n = \alpha_0$.

Suppose now $S(F,x_0) = \{\alpha_0,\alpha_1,\ldots,\alpha_s\}$ is not a singleton. Then for sufficiently large n,

$$G_{n,0}(x_0) \subset S(W_{n,0}) \cap J(x_0,(1+q_n)|x_0-\alpha_0|) = \{z_0^n, z_1^n, \ldots, z_s^n\}$$

where

$$|z_j^n - \alpha_j| \leq q_n|x_0-\alpha_0| \text{ and } z_0^n < z_j^n \quad \text{for } j = 1,2,\ldots,s.$$

Define n_0 as an integer such that

$$(A.11) \quad q_{n_0} < \frac{1}{2|x_0-\alpha_0|\sqrt{n_0}} \; .$$

Since $|x_0-\alpha_j| - |\alpha_j-z_j^n| \leq |z_j^n-x_0| \leq |x_0-\alpha_j| + |z_j^n-\alpha_j|$ for $j = 0,1,\ldots,s$, then

$$|z_0^n-x_0| - d_{n,0}(x_0) \leq 2q_n|x_0-\alpha_0| < \frac{1}{\sqrt{n}}$$

for large n, $n > n_0$. Thus $z_0^n \in G_{n,0}(x_0)$, see (3.3), which means that $x_n = z_0^n$. Hence $\lim_{n\to\infty} x_n = \alpha_0$ which completes the proof for the case $D_F \neq \mathbb{C}$.

Consider now an entire function $F \in \mathfrak{I}_1$, i.e., $D_F = \mathbb{C}$. Let α be an element from $S(f)$. Define

$$\tilde{F}(x) = \begin{cases} F(x) & \text{if } x \in J(\alpha,4|\alpha-x_0|), \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Then $\tilde{F} \in \mathfrak{I}_1, D_{\tilde{F}} \neq \mathbb{C}$, $\alpha \in S(\tilde{F}) \subset S(F)$ and $R_{\tilde{F}}(\alpha) = 4|\alpha-x_0|$. Since $x_0 \in B(\frac{1}{3},\tilde{F})$ then the sequence $\{y_n\}$, $y_n = I_{n,0}(x_0; \mathfrak{N}_{n,0}(\tilde{F},x_0))$, is convergent and $\lim_{n\to\infty} y_n \in S(\tilde{F}) \subset S(F)$. Since $\mathfrak{N}_{n,0}(\tilde{F},x_0) = \mathfrak{N}_{n,0}(F,x_0)$, $y_n = x_n$ and therefore $\lim_{n\to\infty} x_n \in S(F)$. This completes the proof of Theorem 3.3 for the case $N = 1$. ∎

Proof of Theorem 4.1

Following the proof of Theorem 3.3 for $N = 1$ and applying Lemma A.3(ii) instead of (i), we easily get the proof of Theorem 4.1. ∎

To prove Theorem 3.3 for the case $N \neq 1$, consider an analytic operator F, $F \in \mathfrak{I}_1$. Then for

$$(A.12) \quad R_{n,0}(x) \stackrel{df}{=} R_{n,0}(x; x_0) \stackrel{df}{=} F(x) - W_{n,0}(x; x_0)$$

we have

$$(A.13) \quad R_{n,0}(x) = \int_0^1 F^{(n)}(x_0 + t(x-x_0))(x-x_0)^n \frac{(1-t)^{n-1}}{(n-1)!} dt,$$

see e.g., Rall [69, p. 124]. Define

$$A_k(\Gamma) = \sup_{x \in \mathcal{J}(\alpha, \Gamma)} \left\| F'(\alpha)^{-1} \frac{F^{(k)}(x)}{k!} \right\|.$$

It is obvious that Lemmas A.1 to A.4 also hold for this case with the modulus replaced by the norm $\| \cdot \|$. The detailed proofs of Lemmas A.1 and A.2 for $m = 0$ can be found in Traub and Woźniakowski [76].

Following the proof of Theorem 3.3 for $N = 1$ with the modulus replaced by $\| \cdot \|$, we easily get the complete proof of Theorem 3.3.

ACKNOWLEDGMENT

REFERENCES

Ortega and Rheinboldt [70]   Ortega, J. M. and W. C. Rheinboldt, _Iterative Solution of Nonlinear Equations in Several Variables_, Academic Press, New York, 1970.

Rall [69]   Rall, L. B., _Computational Solution of Nonlinear Operator Equations_, John Wiley and Sons, New York, 1969.

Traub [64]   Traub, J. F., _Iterative Methods for the Solution of Equations_, Prentice Hall, 1964.

Traub and Woźniakowski [76]   Traub, J. F. and H. Woźniakowski, "Optimal Radius of Convergence of Interpolatory Iterations for Operator Equations," Department of Computer Science Report, Carnegie-Mellon University, 1976. To appear in _Aequationes Mathematicae_.

Traub and Woźniakowski [78]   Traub, J. F. and H. Woźniakowski, "General Theory of Optimal Error Algorithms and Analytic Complexity, Part B: Iterative Information Model," Department of Computer Science Report, Carnegie-Mellon University, 1978. See also _A General Theory of Optimal Algorithms_, Academic Press, New York, 1980.

Wasilkowski [78]   Wasilkowski, G. W., "Can any Stationary Iteration Using Linear Information be Globally Convergent?," Department of Computer Science Report, Carnegie-Mellon University, 1978. To appear in _J.ACM_.

Wasilkowski [79]   Wasilkowski, G. W., "Any Iteration for Polynomial Equations Using Linear Information has Infinite Complexity," Department of Computer Science Report, Carnegie-Mellon University, 1979.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CMU-CS-79-138	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE STRENGTH OF NONSTATIONARY ITERATION		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G.W. Wasilkowski		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0370
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Department 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 August 1979
		13. NUMBER OF PAGES 27
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)