CMU-CS-79-139

ANY ITERATION FOR POLYNOMIAL EQUATIONS USING LINEAR INFORMATION HAS INFINITE COMPLEXITY

G. W. Wasilkowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(On leave from the University of Warsaw)

CUCS-145-79

September 1979

This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N00014-76-C-0370, NR044-422.

ABSTRACT

This is the third paper in which we study iterations using linear information for the solution of nonlinear equations. In Wasilkowski [78] and [79] we have considered the existence of globally convergent iterations for the class of analytic functions. Here we study the complexity of such iterations. We prove that even for the class of scalar complex polynomials with simple zeros, any iteration using arbitrary linear information has infinite complexity. More precisely, we show that for any iteration $\bar{\varphi}$ and any integer k, there exists a complex polynomial f with all simple zeros such that the first k approximations produced by $\bar{\varphi}$ do not approximate any solution of $f = 0$ better than a starting approximation x_0. This holds even if the distance between x_0 and the nearest solution of $f = 0$ is arbitrarily small.

1. INTRODUCTION

In this paper we continue the study of iterations using linear information for the solution of nonlinear equations $f = 0$. In Wasilkowski [78] we have proven that no stationary iteration using linear information can be globally convergent for the class of scalar analytic functions with simple zeros. In Wasilkowski [79] we have exhibited nonstationary iterations which are globally convergent for the class of analytic functions with simple zeros even for the abstract case.

In this paper we deal with the complexity of iterations using linear information. We prove the surprising result that any such iteration has infinite complexity even for the class $\mathfrak{Z}$ of scalar complex polynomials with simple zeros. To make this negative result as strong as possible we have chosen a relatively simple class $\mathfrak{Z}$. Furthermore we deal with a very general definition of information and iteration. Namely, any sequence of linear finite dimensional operators is considered as possible information, and any sequence of functionals as an iteration. We also do not specify which zero of f is approximated, and the assumptions concerning the starting points are very weak. Under these assumptions we prove that for any positive L, any integer k, and any iteration $\bar{\varphi}$ using linear information, there exists a complex polynomial f having only simple zeros such that the distance between a starting approximation x_0 and a nearest zero α of f is no larger than L and the first k approximation produced by $\bar{\varphi}$ do not approximate any zero of f better than x_0. Note that L can be arbitrarily small which means that x_0 can be arbitrarily close to α.

2. INFORMATION AND ITERATIONS WITHOUT MEMORY

For the reader's convenience we repeat the very general definition of information and iteration without memory introduced in Wasilkowski [79]. For simplicity, in Sections 2 through 5 we deal only with iterations without memory. The extension to the general case is given in Section 6.

Let H be the class of all complex polynomials and $\mathfrak{J}$ be the subset of H which consists of all polynomials having only simple zeros. Let S(f) denote the set of all zeros of f, $f \in H$. Consider the solution of a nonlinear equation

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

To solve (2.1) iteratively we must know something about f. Let $L_i : H \times \mathbb{C} \to \mathbb{C}$ be a functional which is linear with respect to the first argument, i.e., $L_i(c_1 f_1 + c_2 f_2, x) = c_1 L_i(f_1, x) + c_2 L_i(f_2, x)$, $i = 1,2,\ldots,n$. Then the <u>linear information operator</u> $\mathfrak{N}$, $\mathfrak{N} = [L_1, L_2, \ldots, L_n] : H \times \mathbb{C} \to \mathbb{C}^n$, is defined as

(2.2) $\mathfrak{N}(f,x) = [L_1(f,z_1), L_2(f,z_2), \ldots, L_n(f,z_n)], \ \forall f \in H, \ \forall x \in \mathbb{C}$

where $z_1 = x$ and

$$z_j = \xi_j(z_1; L_1(f,z_1), L_2(f,z_2), \ldots, L_{j-1}(f,z_{j-1}))$$

for some functions ξ_j, $j = 2,3,\ldots,n$. Thus any z_j depends on the previously computed information. For brevity we shall sometimes write $z_j = z_j(f)$. Let ψ_n be the class of all such information operators.

3. COMPLEXITY OF ITERATIONS

In this section we define the complexity of an iteration. Let

$$\text{dist}(x,S(f)) = \inf_{\alpha\in S(f)} |x-\alpha|$$

denote the distance between the point x and the set S(f). Let L be a positive number and let $\bar{\varphi}$ be an iteration without memory. For any $f \in \mathfrak{J}$ and x_0 such that

(3.1) $\text{dist}(x_0,S(f)) < L,$

consider the sequence $\{x_i\}$ generated by $\bar{\varphi}$. For any ε, $\varepsilon < 1$, define $N = N(\bar{\varphi},\varepsilon,x_0,f)$ as the minimal integer, if it exists, such that

(3.2) $\text{dist}(x_N,S(f)) \le \varepsilon\ \text{dist}(x_0,S(f)),$

and $N = +\infty$ otherwise. The number N is determined by how many iterative steps are necessary to reduce the starting error by ε.

Let $\text{comp}(\bar{\varphi},\varepsilon,x_0,f)$ be the total cost of computing x_N satisfying (3.2). We do not specify exactly what we mean by the "cost". We merely assume that the cost of the assignment operation is not zero. Since any iterative step performs at least one assignment operation, there exists a positive number c such that

(3.3) $\text{comp}(\bar{\varphi},\varepsilon,x_0,f) \ge cN(\bar{\varphi},\varepsilon,x_0,f),\ \forall\bar{\varphi},\varepsilon,x_0,f.$

In Wasilkowski [79] we showed there exist globally convergent iterations, ie., iterations which for any x_0 and f satisfying (3.1) construct a sequence $\{x_i\}$ such that

$$\mathrm{comp}(\bar{\varphi}, \varepsilon, x_0) = +\infty, \ \forall \varepsilon \in [0,1).$$

This means that the cost of reducing the starting error may be arbitrarily large for some polynomials from $\mathfrak{F}$ even if x_0 is very close to a solution.

(4.4) $\mathfrak{N}(h,x_0) = 0.$

Then there exists β, $\beta \in (0,\frac{L}{2})$, such that $h(x_0+\beta) \neq 0$. For positive σ, define

$$f_\sigma(x) = x-x_0-\beta+\sigma h(x).$$

Let $y_1(\sigma),y_2(\sigma),\ldots,y_r(\sigma)$ be the zeros of f_σ where r is the degree of h. From the theory of algebraic functions (see e.g., Wilkinson [63]) we know that $y_1(\sigma) \neq x_0+\beta$ and $y_1(\sigma) \to x_0+\beta$ as σ tends to zero. It is possible to show that the $y_i(\sigma)$ are simple zeros and $|y_i(\sigma)| \to +\infty$ as σ goes to zero, $i \geq 2$. Thus, for sufficiently small σ, $f_\sigma \in \mathfrak{J}(x_0)$ and $f_\sigma(x_0) \neq 0$. Due to (4.4), $\mathfrak{N}(f_\sigma,x_0) = \mathfrak{N}(x-x_0-\beta,x_0)$ which means that

$$x_i = \varphi_i(x_0,\mathfrak{N}(f_\sigma,x_0)) = \varphi_i(x_0,\mathfrak{N}(x-x_0-\beta,x_0))$$

does not depend on σ, $i = 1,2,\ldots,k$. Note that there exists a small σ_1 such that

(4.5) $\{x_0,x_1,\ldots,x_k\} \cap \{y_1(\sigma_1),y_2(\sigma_1),\ldots,y_r(\sigma_1)\} = \emptyset.$

Indeed, for small σ we have $|y_j(\sigma)| > \max_{0\leq i\leq k} |x_i|$ for $j = 2,3,\ldots,r$. Since $y_1(\sigma)$ takes infinitely many values as σ tends to zero, there exists σ_1 such that $y_1(\sigma_1) \neq x_i$, $i = 1,2,\ldots,k$, which proves (4.5). Taking now $f = f_{\sigma_1}$, we get $f \in \mathfrak{J}(x_0)$ and $x_0,x_1,\ldots,x_k \notin S(f)$. This completes the proof of (4.3) for $n = 1$.

Suppose now by induction, that (4.3) holds for $n \leq n_0$. We want to show that (4.3) also holds for $n = n_0+1$. On the contrary assume that there exist $\mathfrak{N}_n^* \in \Psi_n$,

$$\mathfrak{N}_n^* = [L_1^*,L_2^*,\ldots,L_n^*],$$

Then for any $f \in A_3$, $f(x_0) \neq 0$ and

$$(4.9) \quad L_n^*(\cdot,z_n(f)) \notin \mathrm{lin}\{L_1^*(\cdot;z_1(f)),L_2^*(\cdot,z_2(f)),\ldots,L_{n-1}^*(f))\}.$$

For an information operator $\mathfrak{N}$ and $f \in \mathfrak{J}$, let

$$B(\mathfrak{N}_f) = \{\alpha \in \mathbb{C} : \forall h \in \ker \mathfrak{N}_f,\ h(\alpha) = 0\}$$

where $\mathfrak{N}_f$ is a linear operator defined by (4.1). We need the following lemmas.

<u>Lemma 4.1</u>

If $A_3 \neq \emptyset$ then for any $f \in A_3$,

$$S(f) \cap B(\mathfrak{N}_{n-1,f}^*) \neq \emptyset.$$

<u>Proof</u>

From (4.9) there exists a polynomial ζ, $\zeta = \zeta(f) \in H$, such that $L_n^*(\zeta,z_n(f)) = 1$ and $\zeta \in \ker \mathfrak{N}_{n-1,f}^*$. Define

$$g_\sigma(x) = f(x)+\sigma\zeta(x)$$

for $\sigma > 0$. Since f has only simple zeros, then as in the proof for n = 1, we can conclude that g_σ has only simple zeros which tend to the zeros of f and to infinity (if the degree of f is less than the degree of ζ) as σ goes to zero. Thus, $g_\sigma \in \mathfrak{J}(x_0)$ for sufficiently small σ. Note that $L_j^*(g_\sigma,z_j(f)) = L_j^*(f,z_j(f))$ for $j = 1,2,\ldots,n-1$ which means $z_j(f) = z_j(g_\sigma)$ for $j = 1,2,\ldots,n$. Thus $g_\sigma \notin A_2$. Since $x_0 \notin S(f)$, then x_0 also does not belong to $S(g_\sigma)$ for sufficiently small σ, say $\sigma \in (0,\sigma_0)$. Thus $g_\sigma \in A_3$,

$$\exists i_0 = i_0(\sigma) \in [1,k] : x_{i_0}(g_\sigma) = \varphi_{i_0}(x_0;\ \mathfrak{N}_n^*(g_\sigma,x_0)) \in S(g_\sigma),\quad \forall\sigma \in (0,\sigma_0).$$

$$L^*_{j_s}(\zeta_i, z_{j_s}(f)) = \begin{cases} 1 & \text{if } s = i, \\ 0 & \text{if } s \neq i. \end{cases}$$

We define

$$w_f = \sum_{s=1}^{r} L^*_{j_s}(f, z_{j_s}(f))\zeta_s$$

and

$$A_4 = \{f \in \mathfrak{J}(x_0) : S(w_f) \cap B(\pi^*_{n-1,f}) \neq \emptyset\}.$$

Lemma 4.2

(i) $A_3 \subset A_4$,

(ii) if $A_4 \neq \emptyset$ then for any $f \in A_4$,

$$S(w_f) \cap B(\pi^*_{n-1,f}) \subset S(f).$$

Proof

Without loss of generality we can assume that A_3 is nonempty. Let $f \in A_3$ be arbitrary. Then $h_f \overset{df}{=} f - w_f \in \ker \pi^*_{n-1,f}$ and from Lemma 4.1, there exists $\alpha \in S(f) \cap B(\pi^*_{n-1,f})$. Thus, $w_f(\alpha) = f(\alpha) - h_f(\alpha) = 0$ which means that $S(w_f) \cap B(\pi^*_{n-1,f})$ is nonempty. Thus, $f \in A_4$ which proves that $A_3 \subset A_4$.

To prove (ii), let f be an arbitrary polynomial from A_4. There exists $\alpha_1 \in S(w_f) \cap B(\pi^*_{n-1,f})$. Since $h_f \in \ker \pi^*_{n-1,f}$, $f(\alpha_1) = w_f(\alpha_1) + h_f(\alpha_1) = 0$ which means that $\alpha_1 \in S(f)$. Thus, Lemma 4.2 is proven. ∎

Note that knowing $\pi^*_{n-1}(f, x_0)$ we can verify whether f belongs to A_i, $i = 2,4$. Furthermore for any $\tilde{f} \in \mathfrak{J}(x_0)$ with $\pi^*_{n-1}(\tilde{f}, x_0) = \pi^*_{n-1}(f, x_0)$, $\tilde{f} \in A_i$ iff $f \in A_i$, $i = 2,4$. For $i = 1,2,\ldots,k$, define

Proof

From Theorem 4.1, there exists a polynomial g, $g \in \mathfrak{J}(x_0)$, such that $x_0, x_1 = x_1(g), \ldots, x_k = x_k(g) \notin S(g)$. Let $I = \{i \in [1,k] : x_i(g) \neq x_0\}$. If $I = \emptyset$ then for $f = g$ we have

$$0 \neq \operatorname{dist}(x_0, S(f)) = \operatorname{dist}(x_i(f), S(f)), \quad \forall i = 1,2,\ldots,k,$$

which completes the proof.

Suppose therefore that $I \neq \emptyset$. Consider a polynomial w of the form

$$(4.11) \quad w(x) = \prod_{i \in I} (x-x_i)^m (x-x_0) \sum_{j=0}^{n} a_j x^j, \quad m = \max\{3n, \deg g\},$$

satisfying

$$(4.12) \quad \pi_g(w) = 0.$$

Note that (4.12) is equivalent to the following system of n homogeneous linear equations

$$(4.13) \quad \sum_{j=0}^{n} a_j L_s\Big(\prod_{i \in I} (x-x_i)^m (x-x_0) x^j, z_s(g)\Big) = 0 \quad \text{for} \quad s = 1,2,\ldots,n.$$

Since (4.13) has more unknowns than equations, there exists a non-zero polynomial satisfying (4.11) and (4.12). Consider the factorization of w,

$$w(x) = (x-x_0)^{p_0} \prod_{i \in I} (x-x_i)^{p_i} \prod_{j=1}^{r} (x-y_j)^{s_j}$$

for some r, $r \leq n$, $s_1, s_2, \ldots, s_r$ and p_0, p_i for $i \in I$ where $y_j \neq x_i$ for any i and j. Due to (4.11),

$$(4.14) \quad p_0 \leq n+1 \text{ and } p_i \geq 3n \text{ for } i \in I.$$

$$\min_{i=0,1,\ldots,k} \operatorname{dist}(x_i(f), S(f)) = \operatorname{dist}(x_0, S(f)) \neq 0$$

which completes the proof. ∎

$$\text{dist}(x_{i_0}, S(f_0)) \le \varepsilon\ \text{dist}(x_0, S(f_0)) < \text{dist}(x_0, S(f_0))$$

which contradicts (5.4). Hence Theorem 5.1 is proven. ■

From Theorem 5.1 and (3.8) follows

Corollary 5.1

For any positive L, any sequence of linear information operators $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$, any iteration without memory $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}(\overline{\mathfrak{N}})$, and any starting point $x_0 \in \mathbb{C}$,

$$\text{comp}(\overline{\varphi}, \varepsilon, x_0) = +\infty, \quad \forall \varepsilon < 1.$$ ■

(6.4) $\text{dist}(x_N, S(f)) \le \epsilon\, \text{dist}(x_0, S(f))$,

and $N = +\infty$ otherwise. Let $\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$ be the cost of computing x_N. Let L, $L > 0$, be a given constant. Then

(6.5) $$N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \overset{df}{=} \sup_{f \in \mathfrak{J}(x_0)} N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$$

where $\mathfrak{J}(x_0)$ is defined by (3.6). Similarly, let

(6.6) $$\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \overset{df}{=} \sup_{f \in \mathfrak{J}(x_0)} \text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f).$$

As before, there exists a positive c such that

(6.7) $$\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \ge cN(\bar{\varphi}, \epsilon, x_0, x_1, \ldots, x_{-m}), \quad \forall \bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}.$$

By a technique similar to the proof of Theorem 5.1 it is possible to prove

<u>Theorem 6.1</u>

For any positive L, any m, $m > 0$, any sequence of linear information operators with memory $\bar{\mathfrak{N}}$, any iteration $\bar{\varphi}$, $\bar{\varphi} \in \Phi_m(\bar{\mathfrak{N}})$ and any distinct starting points $x_0, x_{-1}, \ldots, x_{-m} \in \mathbb{C}$

$$\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) = N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) = +\infty, \quad \forall \epsilon < 1. \qquad \blacksquare$$

<u>Remark 6.1</u>

In practice one often wants to reduce a residual error, i.e., to find a point x_k such that

(6.8) $|f(x_k)| \le \epsilon |f(x_0)|$

7. OPEN PROBLEMS

In this section we pose a number of open problems which are relevant to the questions studied in this paper.

In Theorem 6.1 we prove that for any $m \ge 0$, any linear information $\bar{\mathfrak{N}} = \{\mathfrak{N}_i\}$, $\mathfrak{N}_i \in \Psi_{n_i,m}$, any iteration $\bar{\varphi} = \{\varphi_i\} \in \bar{\Phi}_m(\bar{\mathfrak{N}})$ and any integer k, there exists a "difficult" polynomial f, $f \in \mathfrak{J}(x_0)$, i.e., a polynomial which requires at least k+1 iterative steps to reduce the starting error $\mathrm{dist}(x_0, S(f))$. Let $P = P(\bar{\mathfrak{N}}, \bar{\varphi}, k)$ be the set of all such difficult polynomials and let $d = d(\bar{\mathfrak{N}}, \bar{\varphi}, k)$ be the minimal degree of such polynomials, i.e.,

$$d \overset{\mathrm{df}}{=} \min_{f \in P} \deg f.$$

<u>Problem 1</u>

Find d as a function of m, k, and $n_1, n_2, \ldots, n_k$. ■

It can be shown that

$$(7.1) \quad d \le (k+2)\left(2 + \sum_{i=1}^{k} n_i\right) + k.$$

In general, this bound is not sharp. For instance, for a <u>stationary iteration</u>,

$$(7.2) \quad d \le (k+1)(n_1+1) + k.$$

By a stationary iteration we mean an iteration which constructs a sequence of approximations by the formula

$$(7.3) \quad x_{i+1} = \varphi_1(x_i, x_{i-1}, \ldots, x_{i-m}; \mathfrak{N}_1(f, x_i, x_{i-1}, \ldots, x_{i-m}))$$

for some $\mathfrak{N}_1 \in \Psi_{n_1,m}$ and $\varphi_1 \in \Phi(\mathfrak{N}_1)$.

Problem 4

(i) For a given nonincreasing function g, $g : [0,1) \to R$, find information $\bar{\mathfrak{N}}$ and an iteration $\bar{\varphi}$ using $\bar{\mathfrak{N}}$ such that the complexity $\mathrm{comp}(\bar{\varphi}, \epsilon, x_0) \le g(\epsilon)$ for any $\epsilon \in [0,1)$.

(ii) Characterize the class of all information for which the complexity of finding x_N is finite. ∎

ACKNOWLEDGMENT

I am extremely thankful to H. Woźniakowski for introducing me to this problem, for his assistance during my struggle and his valuable help. He also convinced me that the complexity of writing a paper is very high. I also thank J. F. Traub for his valuable comments and help during the preparation of this paper.

REFERENCES

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 JACM.

Wasilkowski [79] Wasilkowski, G. W., "The Strength of Nonstationary Iteration," Department of Computer Science Report, Carnegie-Mellon University, 1979.

Wilkinson [63] Wilkinson, J. H., Rounding Errors in Algebraic Processes, Prentice-Hall, Inc., 1963.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CMU-CS-79-139	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ANY ITERATION FOR POLYNOMIAL EQUATIONS USING LINEAR INFORMATION HAS INFINITE COMPLEXITY		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 September 1979
		13. NUMBER OF PAGES 30 Pages
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)

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

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)