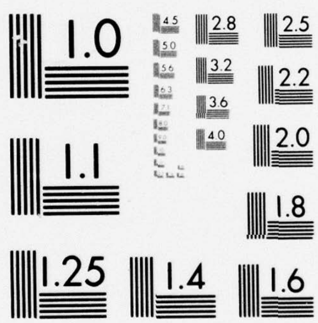
AD-A039 300 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 12/1
CONVERGENCE AND COMPLEXITY OF NEWTON ITERATION FOR OPERATOR EQU--ETC(U)
MAR 77 J F TRAUB, H WOZNIAKOWSKI N00014-76-C-0370

UNCLASSIFIED NL

| OF |

AD
A039300

END
DATE
FILMED
5-77


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6

(12)

CONVERGENCE AND COMPLEXITY OF NEWTON ITERATION FOR OPERATOR EQUATIONS

J. F. Traub
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

H. Woźniakowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(Visiting from the University of Warsaw)

DEPARTMENT
of
COMPUTER SCIENCE

Carnegie-Mellon University

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) CONVERGENCE AND COMPLEXITY OF NEWTON ITERATION FOR OPERATOR EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) J. F. Traub [redacted] H. Wozniakowski		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-76-C-0370, NR 044-422
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NSF-MCS-75-22255
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE (11) Mar [redacted] 77
		13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 26p.		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)

An optimal convergence condition for Newton iteration in a Banach space is established. There exist problems for which the iteration converges but the complexity is unbounded. ~~We show what~~ * stronger condition must be imposed to also assure "good complexity".

* It is shown which

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)

403 081

CONVERGENCE AND COMPLEXITY
OF NEWTON ITERATION FOR OPERATOR EQUATIONS

J. F. Traub
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

H. Woźniakowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(Visiting from the University of Warsaw)

March 1977

This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N0014-76-C-0370, NR 044-422.

ABSTRACT

An optimal convergence condition for Newton iteration in a Banach space is established. There exist problems for which the iteration converges but the complexity is unbounded. We show what stronger condition must be imposed to also assure "good complexity".

1. INTRODUCTION

Numerous papers have analyzed sufficient conditions for the convergence of algorithms for the solution of non-linear problems. In addition to convergence, we consider another fundamental question. What stronger conditions must be imposed to assure "good complexity"? This is clearly one of the crucial issues (in addition to stability) if one is interested in actual computation. We believe it is also a most interesting theoretical question.

We consider Newton iteration for a simple zero of a non-linear operator in a Banach space of finite or infinite dimension. We establish the optimal radius of the ball of convergence with respect to a certain functional. There exist problems where the iteration converges but the complexity increases logarithmically to infinity as the initial iterate approaches the boundary of the ball of convergence. (This phenomenon does not occur in the Kantorovich theory of operator equations; see Section 3.) We establish the optimal radius of the ball of good complexity.

In this paper we limit ourselves to the important case of Newton iteration. In other papers (Traub and Woźniakowski [76b] and [77]) we study optimal convergence and complexity for classes of iterations.

We summarize the results of this paper. Definitions and theorems concerning the optimal ball of convergence are given in Section 2. We conclude this Section by giving conditions under which the radius of the ball of convergence is a constant fraction of the radius of the ball of analyticity of the operator.

Complexity of Newton iteration is studied in Section 3. We show that Newton iteration may converge but have arbitrarily high complexity and conjecture this is a general phenomenon. We establish the radius of the ball of good complexity as well as a lower bound on the complexity of Newton iteration.

2. CONVERGENCE OF NEWTON ITERATION

We consider the solution of the non-linear equation

(2.1) $F(x) = 0$

where $F: D \subset B_1 \to B_2$ and B_1, B_2 are Banach spaces over the real or complex fields of dimension N, $N = \dim(B_1) = \dim(B_2)$, $1 \le N \le +\infty$. We solve (2.1) by Newton iteration which constructs the sequence $\{x_i\}$ as follows

(2.2) $x_{i+1} = x_i - F'(x_i)^{-1}F(x_i)$

where x_0 is an initial guess and $i = 0,1,\dots$. Suppose that F is sufficiently regular and α is a simple zero of (2.1), i.e., $F(\alpha) = 0$ and $F'(\alpha)^{-1}$ exists and is bounded. It is well-known that if x_0 is sufficiently close to α then $\{x_i\}$ converges to α and the order of convergence is two, see, e.g., Kantorovich [48], Ortega and Rheinboldt [70] and Rall [69].

In this paper we are interested in sharp bounds on how close x_0 has to be to α to get quadratic convergence of Newton iteration and how close x_0 has to be to α to guarantee good complexity.

We begin with a theorem which describes the character of convergence of Newton iteration.

Let α be a simple zero of F, $e_i = \|x_i - \alpha\|$ and $J = \{x: \|x-\alpha\| \le \Gamma\}$. Assume $F'(x)$ exists and is a Lipschitz operator for $x \in J$. Define

(2.2) $$A_2 = A_2(\Gamma) = \sup_{x,y \in J} \frac{\|F'(\alpha)^{-1}\{F'(x)-F'(y)\}\|}{2\|x-y\|} .$$

Note that if F'' is continuous for $x \in J$ then

$$A_2(\Gamma) = \sup_{x \in J} \left\|F'(\alpha)^{-1} \frac{F''(x)}{2}\right\|$$

which shows that in this case A_2 is a bound on the "normalized" second derivative.

Let q be any number such that $0 < q < 1$.

Theorem 2.1

If F' is a Lipschitz operator in J,

(2.3) $A_2 \Gamma \le \frac{q}{1+2q}$,

(2.4) $x_0 \in J$,

then Newton iteration is well-defined and

(2.5) $\lim_i x_i = \alpha, \quad e_{i+1} \le q e_i, \quad \forall i,$

(2.6) $e_{i+1} \le C_i e_i^2$

where $C_i = A_2/(1 - 2A_2 e_i)$.

If F is continuously twice-differentiable at α then

(2.7) $x_{i+1} - \alpha = \frac{1}{2} F'(\alpha)^{-1} F''(\alpha)(x_i-\alpha)^2 + o(\|x_i-\alpha\|^2)$.

Proof

The proof is by induction. Let $x_i \in J$ and let

(2.8) $R(x;x_i) = \int_0^1 \{F'(x_i + t(x-x_i)) - F'(x_i)\}(x-x_i)dt.$

Define $w = w(x;x_i)$ as the polynomial of first degree which interpolates $F(x_i)$ and $F'(x_i)$. Then $w(x;x_i) = F(x_i) + F'(x_i)(x-x_i)$ and the next approximation x_{i+1} is the zero of w whenever $F'(x_i)$ is invertible. From Rall [69, p.124] we get the error formula

(2.9) $F(x) - w(x;x_i) = R(x;x_i)$

for $x \in J$ while due to (2.2),

(2.10) $\|F'(\alpha)^{-1}R(x;x_i)\| \le A_2\|x-x_i\|^2.$

From the definition of A_2 we have

$$\|I-F'(\alpha)^{-1}F'(x)\| \le 2A_2\|x-\alpha\| \le 2A_2\Gamma \le \frac{2q}{1+2q} < 1.$$

This means that $F'(x)$ is invertible for any $x \in J$ and x_{i+1} is well-defined. Furthermore

$$\|F'(x)^{-1}F'(\alpha)\| \le 1/(1 - 2A_2\|x-\alpha\|),$$

see Rall [69, p.36]. We can rewrite the polynomial w as $w(x;x_i) = F'(x_i)(x-x_{i+1})$. Set $x = \alpha$ in (2.9). Then $F'(x_i)(x_{i+1}-\alpha) = R(\alpha;x_i)$ which yields

$$\|x_{i+1}-\alpha\| = \|F'(x_i)^{-1}F'(\alpha)F'(\alpha)^{-1}R(\alpha;x_i)\| \le \frac{A_2e_i^2}{1-2A_2e_i} \le \frac{A_2\Gamma}{1-2A_2\Gamma}\, e_i \le qe_i$$

due to (2.3). This proves (2.5) and (2.6).

If $F''(x)$ is continuous at α then

$$R(\alpha;x_i) = \frac{1}{2}F''(\alpha)(x_i-\alpha)^2 + o(\|x_i-\alpha\|^2),$$
$$F'(x_i) = F'(\alpha) + O(\|x_i-\alpha\|)$$

which implies (2.7). ∎

Remark 2.1

Theorem 2.1 states quadratic convergence of Newton iteration since $e_{i+1} \le C_ie_i^2$ and $C_i \le A_2/(1 - 2A_2e_0)$. It is known (see Ortega and Rheinboldt

[70, p.319]) that if F' is not a Lipschitz operator then Newton iteration does not possess, in general, quadratic convergence. For instance, applying Newton iteration to $F(x) = x + x^a$ where $a = 5/3$ we get

$$e_{i+1} = (a-1)e_i^a/(1 + ae_i^{a-1}).$$ ■

Remark 2.2

Theorem 2.1 states how fast Newton iteration converges. The speed of convergence depends on the value of q. This technique based on q seems to be general and can be applied for any iteration. See Traub and Woźniakowski [77] where the class of interpolatory iterations I_n, $n \geq 3$, is considered. ■

We show that if we want (2.5) to hold then (2.3) is sharp for any value of q, $0 < q < 1$.

Theorem 2.2

There exists an entire F such that

$$A_2\Gamma > \frac{q}{1+2q} \text{ and } e_0 = \Gamma \text{ imply } e_1 > qe_0.$$

for every value of $q \in (0,1)$.

Proof

Let $F(x) = x + x^2$, $x \in \Re$. Then $\alpha = 0$ and $A_2(\Gamma) \equiv 1$. Suppose that $A_2\Gamma = \Gamma > q/(1 + 2q)$ and let $x_0 = -\Gamma$. The next approximation x_1 constructed by Newton iteration is equal to $x_1 = x_0 - F'(x_0)^{-1}F(x_0) = \Gamma^2/(1-2\Gamma) > q\Gamma = -qx_0$. Hence $e_1 > qe_0$. ■

From theorem 2.1 with $q \to 1^-$ it follows that Newton iteration converges whenever

$$A_2\Gamma < 1/3.$$

We show that there exists a problem F such that $A_2\Gamma = 1/3$ and $e_0 = \Gamma$ imply that Newton iteration does not converge.

Theorem 2.3

There exists a problem F where F' is a Lipschitz operator such that if

$$A_2\Gamma = \frac{1}{3} \text{ and } e_0 = \Gamma$$

then Newton iteration is not convergent.

Proof

Define

$$(2.11) \qquad F(x) = \begin{cases} x + x^2 & \text{for } x \le 0 \\ x - x^2 & \text{for } x \ge 0. \end{cases}$$

Then $F(0) = 0$, $F'(0) = 1$ and F' is a Lipschitz function with $A_2(\Gamma) = 1$. Let $\Gamma = \frac{1}{3}$ and $x_0 = -\frac{1}{3}$. Then $x_1 = \frac{1}{3}$ and $x_2 = -\frac{1}{3}$. Hence Newton iteration cycles and is not convergent. ■

Rall [74] showed that assuming the existence of α and using the Kantorovich theorem it is necessary to assume

$$A_2\Gamma \le \frac{2-\sqrt{2}}{4} \doteq 0.15.$$

Thus we have obtained an improvement by a factor of $\frac{4+2\sqrt{2}}{3} \doteq 2.3$.

We define optimal convergence number. Let $X \subset \Re$ be a set of positive numbers r such that for any F with a simple zero α such that F' is a Lipschitz operator in J, $A_2\Gamma < r$ implies that an iteration converges whenever $\|x_0-\alpha\| \le \Gamma$.

Definition 2.1

r_c is called the optimal convergence number (with respect to A_2) if

$$r_c = \sup X. \qquad \blacksquare$$

From theorem 2.1 and theorem 2.3 we have

Corollary 2.1

For Newton iteration

$$r_c = 1/3. \tag{2.11} \qquad \blacksquare$$

Consider the non-linear scalar equation

$$g(\Gamma) \equiv A_2(\Gamma)\Gamma, \; \Gamma \geq 0.$$

Provided F is not linear, $g(\Gamma)$ is strictly increasing, $g(0) = 0$, $g(\infty) = \infty$. Let h denote the inverse function g^{-1}. Then Newton iteration converges provided $\|x_0 - \alpha\| \leq h(r)$, $r < \frac{1}{3}$.

Definition 2.2

R_c is called the optimal radius of the ball of convergence (with respect to A_2) if

$$R_c = h(r_c). \qquad \blacksquare$$

Corollary 2.2

For Newton iteration

$$R_c = h(1/3). \qquad \blacksquare$$

Remark 2.3

Let φ and ψ be any two iterations whose convergence condition is of the form $A_2(\Gamma)\Gamma < r$. Let $r_c(\varphi)$, $r_c(\psi)$ denote the optimal convergence numbers of φ and ψ, and let $R_c(\varphi)$, $R_c(\psi)$ denote the corresponding optimal radii of the ball of convergence. Then

$$r_c(\varphi) \le r_c(\psi)$$

implies

$$R_c(\varphi) \le R_c(\psi).$$

Therefore the optimal radii of convergence of two iterations can be ordered if their convergence numbers are ordered. ∎

In general, Newton iteration converges only locally. We give conditions under which Newton iteration enjoys a "type of global convergence".

Let $F(x) = \sum_{i=1}^{\infty} \frac{1}{i!} F^{(i)}(\alpha)(x_i - \alpha)^i$ be analytic in $D = \{x: \|x-\alpha\| < R\}$ and

$$\frac{\|F'(\alpha)^{-1}F^{(i)}(\alpha)\|}{i!} \le K^{i-1} \tag{2.12}$$

for $i = 2,3,\ldots$, $R \ge \frac{1}{K}$.

One way to find K is to use Cauchy's formula

$$\frac{\|F'(\alpha)^{-1}F^{(i)}(\alpha)\|}{i!} \le \frac{M}{R^i}$$

where $M = \sup_{x \in D} \|F'(\alpha)^{-1}F(x)\|$. Setting $K = \max\left(\frac{1}{R}, \frac{M}{R^2}\right)$ we get $M/R \le KR \le (KR)^{i-1}$ which yields $M/R^i \le K^{i-1}$.

Theorem 2.4

If F satisfies (2.12) then Newton iteration converges in $J = \{x: \|x-\alpha\| \le \Gamma\}$ where

(2.13) $\Gamma = \frac{c_1}{K}$ and $c_1 \cong \frac{1}{6}$.

<u>Proof</u>

Since $\|F'(\alpha)^{-1}F^{(i)}(x)\| \le f^{(i)}(\|x-\alpha\|)$ where $f(x) = x/(1 - Kx)$ then

$$A_2(\Gamma)\Gamma \le \frac{K\Gamma}{(1-K\Gamma)^3} < \frac{1}{3}$$

which yields $\Gamma = \frac{c_1}{K}$ with $c_1 \cong \frac{1}{6}$. This proves (2.13). ■

Note that this result is especially interesting if the domain radius R is approximately equal to 1/K, say $R = c_2/K$. Then $\Gamma = \frac{c}{K} = \frac{c}{c_2} R \simeq \frac{1}{6c_2} R$ and Newton iteration enjoys a "type of global convergence".

3. COMPLEXITY OF NEWTON ITERATION

In the previous section we showed that the sequence of errors $e_i = \|x_i - \alpha\|$ for Newton iteration satisfies

$$(3.1)\quad e_{i+1} \le \frac{A_2 e_i^2}{1-2A_2 e_i} \quad \text{whenever } A_2 e_0 < 1/3.$$

We want to find x_k such that k is the smallest index for which $e_k \le \varepsilon' e_0$ where ε', $0 < \varepsilon' < 1$, is a given number. Define $\varepsilon \le \varepsilon'$ so that

$$(3.2)\quad e_k = \varepsilon e_0.$$

The complexity of Newton iteration, which is the total cost of computing x_k, $k \ge 1$, is given by

$$(3.3)\quad \text{comp} = k \cdot c$$

where c is the cost of one Newton step. (See Traub and Woźniakowski [76a] for a detailed discussion.) Let $c(F^{(j)})$ denote the cost of one evaluation of $F^{(j)}$, $j = 0$ and 1, and let the combinatory cost d be the cost of evaluating $x_{i+1} = x_i - F'(x_i)^{-1}F(x_i)$. Then the cost per step is given by

$$(3.4)\quad c = c(F) + c(F') + d.$$

Note that the dimension of the problem $N \le +\infty$. For multivariate problems, $N < +\infty$, we can assume that each arithmetic operation costs unity and $c(F^{(j)})$ can denote the total number of arithmetic operations needed to evaluate $F^{(j)}$ and $d = d(N)$ is the cost of the solution of the linear system $F'(x_i)(x_{i+1}-x_i) = -F(x_i)$. Hence $d(N) = O(N^\beta)$ with $\beta \le 3$.

For the rest of this section $N \le +\infty$. Define

$$(3.5)\quad f_{i+1} = \frac{A_2 f_i^2}{1-2A_2 f_i}, \quad f_0 = e_0, \; i = 0,1,\dots .$$

Clearly, the sequence $\{f_i\}$ is a majorizing sequence of $\{e_i\}$, i.e., $e_i \le f_i$, $\forall i$. Let $\text{comp}_1 = k_1 c$ where k_1 is the smallest index for which $f_k \le \epsilon e_0$. Of course, $\text{comp} \le \text{comp}_1$. Let $\rho = A_2 f_0 = A_2 e_0$. Note that $\text{comp}_1 = \text{comp}_1(\epsilon,\rho)$. We derive some properties of comp_1 as ρ approaches $1/3$ (which is the necessary and sufficient condition for guaranteeing convergence).

<u>Lemma 3.1</u>

Let n be any fixed integer. If $\rho = \frac{1}{3} - \delta$ with $\delta \le \frac{1}{4^{n+1}}$ then

$$(3.6)\quad f_i = (1 - g_i\delta) f_0 \quad \text{where } 0 \le g_i \le 4^{i+1} - 4$$

for $i = 0,1,\dots, n$.

<u>Proof</u>

Assume by induction that $f_i = (1 - g_i\delta) f_0$ with $0 \le g_i \le 4^{i+1} - 4$. Since $f_{i+1} \le f_0$ then $g_{i+1} \ge 0$ and

$$f_{i+1} = \frac{\rho(1-g_i\delta)^2 f_0}{1-2\rho(1-g_i\delta)}$$

which yields after algebraic manipulations

$$g_{i+1} \le 4g_i + 9 + 3(g_i\delta)^2 \le 4^{i+2} - 4 + 3(g_i\delta)^2 - 3 \le 4^{i+2} - 4.$$

This proves (3.6). ■

We are ready to prove

<u>Lemma 3.2</u>

(3.7) $\frac{c}{4}\ \lg(\frac{1}{\delta})(1 + o(1)) \le \mathrm{comp}_1(\epsilon,\rho) \le c\ \log(\frac{1}{\delta})(1 + o(1))$

where $\rho = \frac{1}{3} - \delta$ with $\delta \to 0^+$ and lg is the logarithm to base 2. ■

<u>Proof</u>

Recall that $\mathrm{comp}_1 = k_1 c$ and we seek bounds on k_1. Since

$$f_i \le \frac{A_2}{1-2A_2 f_0}\, f_{i-1}^2 \le \left(\frac{A_2 f_0}{1-2A_2 f_0}\right)^{2^i-1} f_0 = \left(\frac{1-3\delta}{1+6\delta}\right)^{2^i-1} f_0$$

then

$$k_1 \le \lg\left(1+\lg(\epsilon)/\lg\left(\frac{1-3\delta}{1+6\delta}\right)\right) = \lg\frac{1}{\delta}(1 + o(1)).$$

This proves the righthand side of (3.7).

Let $\delta = 1/4^{n+1}$. From Lemma 3.1 for $i = 1,2,\ldots,\lfloor n/2 \rfloor$ we get

$$f_i \ge (1 - (4^{i+1}-4)/4^{n+1})f_0 \ge (1-2^{-n})f_0 > \epsilon f_0$$

for large n. Thus $k_1 \ge \lfloor n/2 \rfloor = \frac{1}{4}\lg(\frac{1}{\delta})(1 + o(1))$ which proves the lefthand side of (3.7). ■

Lemma 3.2 states the important difference between convergence and complexity of the sequence $\{f_i\}$. $\rho < \frac{1}{3}$ assures convergence of $\{f_i\}$ but with only this condition complexity can be arbitrarily (logarithmically) large.

We believe results similar to Lemma 3.2 also hold for comp (rather than comp_1) of an arbitrary iteration and therefore propose

<u>Conjecture 3.1</u>

For any superlinear iteration Φ there exists a function F with a simple zero α and a number Γ, $0 < \Gamma < +\infty$, such that

(i) the sequence $x_{i+1} = \Phi(x_i, F)$ constructed by the iteration Φ is convergent to α and $\|x_{i+1}-\alpha\| < \|x_i-\alpha\|$, $\forall i$, whenever an *initial approximation* x_0 *is any point of* $J = \{x: \|x-\alpha\| < \Gamma\}$

(ii) the complexity comp, which is now a function of x_0 and ϵ, satisfies

$$\overline{\lim_{x_0: \|x_0-\alpha\| \to \Gamma}} \quad \mathrm{comp}(x_0,\epsilon) = +\infty, \quad \forall\epsilon > 0. \qquad \blacksquare$$

Conjecture 3.1 states that starting from any point of J we get convergence; however, $\mathrm{comp}(x_0;\epsilon)$ can be arbitrarily high when x_0 approaches the boundary of J.

We prove Conjecture 3.1 for Newton iteration.

Theorem 3.1

If Newton iteration is applied to

$$(3.8)\quad F(x) = \begin{cases} x + x^2 & \text{for } x \le 0 \\ x - x^2 & \text{for } x \ge 0 \end{cases}$$

then for ϵ fixed $\mathrm{comp}(x_0,\epsilon)$ tends logarithmically to infinity as $A_2\Gamma$ approaches 1/3 and $x_0 = \Gamma$.

Proof

Note that $\alpha = 0$ and $A_2(\Gamma) \equiv 1$, see (2.11). Applying Newton iteration to (3.8) we get

$$x_{i+1} = \frac{-\mathrm{sign}(x_i)x_i^2}{1-2\mathrm{sign}(x_i)x_i} \text{ and } e_{i+1} = \frac{e_i^2}{1-2e_i}$$

which is equivalent to (3.5) with $A_2 = 1$. Let $x_0 = \Gamma = \frac{1}{3} - \delta$. From Lemma (3.2) we get

$$\frac{1}{4}(1+o(1)) \le \frac{\mathrm{comp}(x_0,\epsilon)}{c\ \lg(\frac{1}{\delta})} \le 1+o(1) \quad \text{as } \delta \to 0.$$

This proves theorem 3.1. ■

This phenomenon does not occur in the Kantorovich theory. If $h = \frac{1}{2}$ (in the usual notation; see Rall [74]), Newton iteration converges linearly and the complexity is always bounded for fixed ϵ. Furthermore the complexity depends on $\lg(1/\epsilon)$ rather than $\lg\lg(1/\epsilon)$ only if α is not a simple zero.

Note that F defined in (3.8) is not twice-differentiable at α. It seems to us that comp can tend to infinity as $A_2\Gamma$ approaches 1/3 for twice-differentiable problems. Therefore we propose

Conjecture 3.2

There exists a twice-differentiable problem F with a simple zero α such that the comp = comp($A_2\Gamma$) of Newton iteration satisfies

$$\lim_{A_2\Gamma \to 1/3^-} \mathrm{comp}(A_2\Gamma) = +\infty$$ ■

Theorem 3.1 states that if $A_2\Gamma$ is close to 1/3 then complexity can be arbitrarily large. We seek a bound on $A_2\Gamma$ (stronger than $A_2\Gamma < 1/3$) to be sure that complexity is not too large. Define

(3.9) $t = \max(1, \lg 1/\epsilon)$.

Theorem 3.2

If $A_2\Gamma \le 1/4$ then the complexity of Newton iteration satisfies

(3.10) $\mathrm{comp} \le c \lg(1+t)$.

Proof

Let $h_{i+1} = Kh_i^2$ with $K = A_2/(1 - 2A_2e_0)$ and $h_0 = e_0$. From (3.1), $e_i \le h_i$, $\forall i$. In Traub and Woźniakowski [76a] we proved that if $w = (Ke_0)^{-1} \ge 2$ then the complexity comp(h) of the sequence $\{h_i\}$ is bounded by

$$\text{comp}(h) \le c \lg(1+t).$$

Since comp $\le$ comp(h), (3.10) follows. ■

Remark 3.1

We have chosen the endpoint $w = 2$ in the inequality $w \ge 2$ only for convenience and definiteness. Any value w bounded away from unity can be used to get a good bound on complexity. ■

Theorem 3.2 motivates the definition of the optimal complexity number for a superlinear iteration. Let $Y \subset \Re$ be a set of positive numbers r such that for any F with a simple zero α such that F' is a Lipschitz operator in J, $A_2\Gamma \le r$ implies that an iteration converges and comp $\le c \lg(1+t)$, $\forall t \ge 1$, whenever $\|x_0-\alpha\| \le r$.

Definition 3.1

<u>r_g is called the optimal complexity number [with respect to A_2 and the complexity criterion comp $\le c \lg(1+t)$]</u> for a superlinear iteration if

$$r_g = \sup Y.$$ ■

Compare with the definition of the optimal convergence number r_c. Of course $r_g \le r_c$ and from Theorem 3.2 follows that $r_g \ge 1/4$.

Theorem 3.3

For Newton iteration

(3.11) $r_g = \frac{1}{4}$.

Proof

Let $F(x) = x + x^2$. Then $F(\alpha) = 0$, $F'(\alpha) = 1$ and $A_2\Gamma = \Gamma$. Newton iteration produces $\{x_i\}$ such that

(3.12) $$x_{i+1} = \frac{x_i^2}{1+2x_i^2}.$$

Let $t = 1$. Then comp $\le c$ which means that $e_1 \le \frac{1}{2} e_0$. Set $x_0 = -\Gamma$, $\Gamma < \frac{1}{2}$.

From (3.12) we get

$$e_1 = \frac{\Gamma^2}{1-2\Gamma} \le \frac{1}{2}\Gamma = \frac{1}{2}e_0$$

which is equivalent to $\Gamma \le 1/4$. Hence $A_2\Gamma \le r_g \le \frac{1}{4}$ and from theorem 3.2 we conclude $r_g = \frac{1}{4}$. ■

Remark 3.2

It follows from Remark 3.1 that as long as $A_2\Gamma$ is bounded away from 1/3 we are assured of good complexity. ■

We summarize our results on optimal convergence number and optimal complexity number (Corollary 2.1 and Theorem 3.3):

(i) $A_2e_0 < r_c = 1/3$ assures convergence but complexity can be arbitrarily large

(ii) $A_2e_0 \le r_g = 1/4$ assures convergence and good complexity.

__Definition 3.2__

R_g is called the __optimal radius of the ball of good complexity [with respect to A_2 and comp $\le c\ \lg(1+t)$]__ for a superlinear iteration if

$$R_g = h(r_g). \qquad \blacksquare$$

__Corollary 3.1__

For Newton iteration

$$R_g = h(1/4). \qquad \blacksquare$$

__Remark 3.3__

A remark analogous to Remark 2.3 holds here concerning the ordering of the optimal radii of good complexity for two iterations. ■

The ratio between the optimal radius of the ball of convergence and the optimal radius of the ball of good complexity,

$$\frac{R_c}{R_g} = \frac{h(1/3)}{h(1/4)} .$$

indicates what stronger condition must be imposed to assure good complexity.

We end this paper by deriving a __lower bound__ on the complexity of Newton iteration. Let F be any operator for which

(3.13) $\quad e_{i+1} \ge a_2 e_i^2$

for a positive a_2. We show that the class of such operators is not empty.

Let

$$F(x) = \frac{1}{x} - a, \quad a > 0, \ x \ne 0.$$

Then

$$e_{i+1} = a e_i^2.$$

Recall that $c = c(F) + c(F') + d$ is the cost of one Newton step where d is the cost of solving a linear system of size N. Let c_L and c_U denote lower and upper bounds on c. If $a_2 e_0 \geq 1/t$, then Theorem 3.1 in Traub and Woźniakowski [76a] yields comp $\geq c_L(\text{lgt-lglgt})$.

We summarize the complexity analysis for Newton iteration.

Theorem 3.4

If

$$A_2 e_0 \leq \frac{1}{4} \text{ and } a_2 e_0 \geq \frac{1}{t}$$

then the complexity of Newton iteration satisfies

$$c_L(\text{lgt-lglgt}) \leq \text{comp} \leq c_U \ \text{lg}(1+t),$$

where $t = \max(1, \text{lg } 1/\epsilon)$.

ACKNOWLEDGMENT

We are indebted to H. T. Kung for several important suggestions on this paper.

REFERENCES

Kantorovich [48]	Kantorovich, L. V., "Functional Analysis and Applied Mathematics," Uspeln Mat. Nauk, 3 (1948), 89-185 (Russian). Tr. by C. D. Benster, National Bureau of Standards Report No. 1509, Washington, D.C., 1952.
Ortega and Rheinboldt [70]	Ortega, J. M. and Rheinboldt, W. C., 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.
Rall [74]	Rall, L. B., "A Note on the Convergence of Newton's Method," SIAM J. Numer. Anal., 11 (1974), 34-36.
Traub and Woźniakowski [76a]	Traub, J. F. and Woźniakowski, H.,"Strict Lower and Upper Bounds on Iterative Computational Complexity," in Analytic Computational Complexity, edited by J. F. Traub, Academic Press, 1976, 15-34.
Traub and Woźniakowski [76b]	Traub, J. F., and Woźniakowski, H., "Optimal Convergence of Interpolatory Iterations for Operator Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1976.
Traub and Woźniakowski [77]	Traub, J. F., and Woźniakowski, H., "Convergence and Complexity of Interpolatory-Newton Iteration in a Banach Space," Dept. of Computer Science Report, Carnegie-Mellon University, 1977.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONVERGENCE AND COMPLEXITY OF NEWTON ITERATION FOR OPERATOR EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. F. Traub and H. Wozniakowski		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0370; NR 044-422
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE March 1977 13. NUMBER OF PAGES 22
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)

An optimal convergence condition for Newton iteration in a Banach space is established. There exist problems for which the iteration converges but the complexity is unbounded. We show what stronger condition must be imposed to also assure "good complexity".