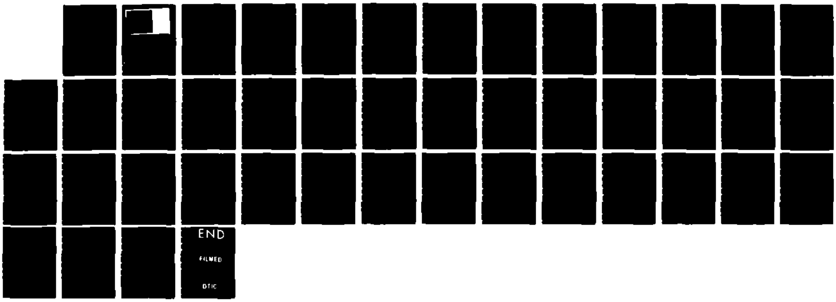
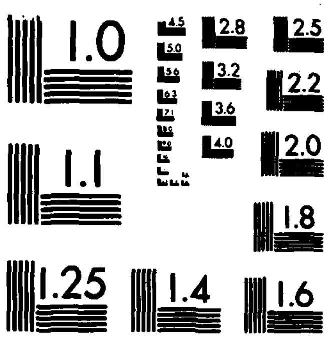
AD-A160 994 ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER KANTOROVICH TYPE ASSUMPTIONS(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER T YAMAMOTO JUL 85 MRC-TSR-2846 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

MRC Technical Summary Report #2846

ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER KANTOROVICH TYPE ASSUMPTIONS

Tetsuro Yamamoto

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

July 1985

(Received July 2, 1985)

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER KANTOROVICH TYPE ASSUMPTIONS

Tetsuro Yamamoto*

Technical Summary Report #2846
July 1985

ABSTRACT

This paper gives a method to derive new *a posteriori* error bounds for Newton-like methods in a Banach space under Kantorovich type assumptions. The bounds found are sharper than those of Miel [10] and include those recently obtained by Moret [12]. The applicability of the author's ~~our~~ method is studied for other types of iterations. Various error bounds for the Newton method under the Kantorovich assumptions are surveyed in the Appendix. Keywords: estimates (operators (mathematics)); convergence.

AMS(MOS) Subject Classifications: 65G99, 65J15

Key Words: Newton-like methods, Newton's method, *a posteriori* error estimates, the Newton-Kantorovich theorem, Kantorovich type assumptions, Dennis' theorem, Rheinboldt's theorem, Miel's bounds, Moret's bounds

Work Unit Number 3 (Numerical Analysis and Scientific Computing)

*Permanent address: Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and by the Ministry of Education, Japan.

SIGNIFICANCE AND EXPLANATION

To find sharper error bounds for iterative solutions of nonlinear equations is one of the important subjects in numerical analysis. This paper gives a simple and powerful technique for improving known error bounds for Newton-like methods in a Banach space under Kantorovich type assumptions.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER KANTOROVICH TYPE ASSUMPTIONS

Tetsuro Yamamoto*

1. INTRODUCTION

Let X and Y be Banach spaces and consider an operator $F : D \subseteq X \to Y$. If F is Fréchet differentiable in an open convex set $D_0 \subseteq D$, then the Newton method for solving the equation

$$F(x) = 0 \tag{1.1}$$

is defined by

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad n = 0, 1, 2, \ldots, \tag{1.2}$$

provided that $F'(x_n)^{-1} \in L(Y,X)$ exists at each step, where $L(Y,X)$ denotes the Banach space of bounded linear operators of Y into X. Since Kantorovich [6] established his famous theorem, called the Kantorovich theorem, which guarantees the convergence of the method and existence and uniqueness of the solution of the equation (1.1), and gave another proof of the theorem with the use of a majorizing sequence, many authors have made efforts to find sharper error bounds for x_n and establish similar convergence theorems for the Newton-like method

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n), \quad n = 0, 1, 2, \ldots, \tag{1.3}$$

where $A(x_n)$ is a linear operator which approximates $F'(x_n)$.

One of the typical generalizations of the Kantorovich theorem is given by Rheinboldt [20] on the basis of his majorant principle, which generalizes Kantorovich's majorant

*Permanent address: Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and by the Ministry of Education, Japan.

technique for the Newton method. A further generalization of Rheinboldt's result is given by Dennis [2], which is stated in an affine invariant form as follows:

THEOREM 1.1. Let $F : D \subseteq X \to Y$ be Fréchet differentiable in an open convex set $D_0 \subseteq D$ and $A : D_0 \to L(X,Y)$. Assume that for a point $x_0 \in D_0$, $A(x_0)^{-1}$ exists and for constants $K > 0$, $L \geq 0$, $M \geq 0$, $\ell \geq 0$, $m \geq 0$, and $\eta > 0$, the following hold:

$$\|A(x_0)^{-1}(F'(x) - F'(y))\| \leq K\|x - y\| , \quad x, y \in D_0 ,$$

$$\|A(x_0)^{-1}(A(x) - A(x_0))\| \leq L \|x - x_0\| + \ell , \quad x \in D_0 , \tag{1.4}$$

$$\|A(x_0)^{-1}(F'(x) - A(x))\| \leq M \|x - x_0\| + m , \quad x \in D_0 ,$$

$$\|A(x_0)^{-1}F(x_0)\| \leq \eta , \quad \ell + m < 1 , \quad \sigma = \max(1, \frac{L + M}{K}) ,$$

$$h = \frac{\sigma K \eta}{(1 - \ell - m)^2} \leq \frac{1}{2} , \tag{1.5}$$

$$t^* = \frac{1 - \sqrt{1 - 2h}}{\sigma K}(1 - \ell - m) ,$$

$$\bar{S}(x_0,t^*) = \{x \in X \mid \|x - x_0\| \leq t^*\} \subseteq D_0 .$$

Then the sequence $\{x_n\}$ generated by (1.3) exists, remains in $\bar{S}(x_0,t^*)$ and converges to a solution x^* of (1.1) which is unique in $D_0 \cap S(x_0,\hat{t})$ where $\hat{t} = (1 - m + \sqrt{(1 - m)^2 - 2K\eta})/K$. Furthermore, error estimates

$$\|x^* - x_n\| \leq t^* - t_n , \quad n = 0, 1, 2, \ldots \tag{1.6}$$

hold, where $\{t_n\}$ is defined by

$$t_0 = 0 , \quad t_{n+1} = t_n + \frac{f(t_n)}{g(t_n)} , \quad n = 0, 1, 2, \ldots , \tag{1.7}$$

with

$$f(t) = \frac{1}{2}\sigma K t^2 - (1 - \ell - m)t + \eta , \quad g(t) = 1 - \ell - Lt . \tag{1.8}$$

Rheinboldt's theorem [20; Theorem 4.3] corresponds to the case $\ell = 0$ in Theorem 1.1. An improved version for the error bounds of Rheinboldt was obtained by Miel [10]. His technique is applicable to Dennis' bounds (1.6), too, and we obtain

$$\|x^* - x_n\| \leq \frac{t^* - t_n}{t_{n+1} - t_n} \|x_{n+1} - x_n\| \leq \frac{t^* - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\| . \tag{1.9}$$

The last bound in (1.9) is of the form found in Miel [10].

Recently, with the use of another type of majorizing sequence, Moret [12] gave a convergence theorem for the iteration (1.3) as well as error bounds, but, under the stronger assumptions than those of Theorem 1.1. In fact, in our notation, he replaced the condition (1.4) by

$$\|A(x_0)^{-1}(A(x) - A(y))\| \leq L\|x - y\| , \quad x, y \in D_0 \tag{1.10}$$

and assumed that $L \leq K$, $M = K - L$ and $2h < 1$. He has shown by numerical experiments that his bounds are sharper than the last bound of (1.9). However, no proof is given.

In this paper, first in §2, we shall present a simple technique to improve the error bounds (1.6) under the same assumptions as in Theorem 1.1. It is shown that the results thus found are sharper than Miel's bounds (1.9) and include what Moret obtained under stronger assumptions. Our technique is simple, but powerful, so that we can improve the error bounds for the other types of iterations which were obtained with the use of majorizing sequences. To show this, in §3, we shall consider the following three types of iterations:

$$x_{n+1} = x_n - A_{\alpha_n}^{-1}F(x_n) , \quad n = 0, 1, 2, \ldots, \tag{1.11}$$

where $\alpha_0 = 0$, $\alpha_n = n$ or $\alpha_n = \alpha_{n-1}$, $n \geq 1$ and $A_{\alpha_n} \in L(X,Y)$;

$$x_{n+1} = x_n - \delta F(x_n,y_n)^{-1}F(x_n) , \quad n = 1, 2, 3, \ldots, \tag{1.12}$$

where $\delta F : D_0 \times D_0 \to L(X,Y)$; and

$$x_{n+1} = x_n - T'(x_n)^{-1}F(x_n) , \quad n = 0, 1, 2, \ldots, \tag{1.13}$$

where $T : D \subseteq X \to Y$ and T is Fréchet differentiable in $D_0 \subsetneq D$, while the differentiability of F is not assumed. Convergence theorems for iterations (1.11) and (1.12) were given by Dennis [3] and Schmidt [21], respectively. The iteration (1.13) was considered by Zincenko [27], Rheinboldt [20] and Moret [12]. We shall show that their error bounds can easily be improved by our method. Furthermore, in §4, we shall specialize our results to the Newton method and show that our bounds improve the basic error bounds [25; Lemma 3] which are obtained from the Kantorovich theorem. Therefore, from the previous results [24] - [26], we can conclude that our bounds for the Newton method which coincide with those of Moret are sharper than those of Miel [11], Potra-Pták [17] and Gragg-Tapia [5], etc. Finally, a more detailed comparison will be made in the Appendix between the various error bounds for the Newton method which have been obtained by many authors under the assumptions of the Kantorovich theorem.

2. IMPROVED ERROR BOUNDS FOR (1.3).

Throughout this section, we keep the notation and assumptions of Theorem 1.1. Furthermore, without loss of generality, we assume that $x_n \neq x_{n+1}$, since, otherwise we have $x_n = x^*$ and $\|x_n - x^*\| = 0$. Then we have

$$x^* - x_{n+1} = x^* - x_n + A(x_n)^{-1}(F(x_n) - F(x^*)) =$$

$$-A(x_n)^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n) + \{F'(x_n) - A(x_n)\}(x^* - x_n)]$$

$$= -A(x_n)^{-1}A(x_0)[\int_0^1 A(x_0)^{-1}\{F'(x_n + t(x^* - x_n)) - F'(x_n)\}(x^* - x_n)dt$$

$$+ A(x_0)^{-1}\{F'(x_n) - A(x_n)\}(x^* - x_n)] ,$$

$$A(x_n) = A(x_0)[I + A(x_0)^{-1}(A(x_n) - A(x_0))]$$

and

$$\|A(x_0)^{-1}(A(x_n) - A(x_0))\| \leq L\|x_n - x_0\| + \ell \leq Lt_n + \ell$$

$$< \sigma K t^* + \ell = 1 - m - (1 - \ell - m)\sqrt{1 - 2h}$$

$$\leq 1 ,$$

where we have used the fact that $t_n < t^*$ if $n > 0$, which is satisfied because of our assumption. Hence, using Banach's lemma, we obtain

$$\|x^* - x_{n+1}\| \leq (1 - \ell - L\Delta_n)^{-1}\{\frac{K}{2}\|x^* - x_n\|^2 + (m + M\Delta_n)\|x^* - x_n\|\}$$

$$\leq (1 - \ell - Lt_n)^{-1}\{\frac{K}{2}\|x^* - x_n\|^2 + (m + Mt_n)\|x^* - x_n\|\} ,$$

where $\Delta_n = \|x_n - x_0\|$. For the sake of simplicity, we put $d_n = \|x_{n+1} - x_n\|$, $a_n = \ell + L\Delta_n$, $b_n = m + M\Delta_n$, $\tilde{a}_n = \ell + Lt_n$, $\tilde{b}_n = m + Mt_n$,

$$\varphi_n(t) = (1 - a_n)^{-1}(\frac{1}{2}Kt^2 + b_n t) \tag{2.1}$$

and

$$\tilde{\varphi}_n(t) = (1 - \tilde{a}_n)^{-1}(\frac{1}{2}Kt^2 + \tilde{b}_n t) . \tag{2.2}$$

Then we have $\varphi_n(t) < \tilde{\varphi}_n(t)$ for all $t > 0$ or $\varphi_n(t) \equiv \tilde{\varphi}_n(t)$ for all $t > 0$. Furthermore, put $\Phi_n(t) = \varphi_n(t) - t + d_n$ and $\tilde{\Phi}_n(t) = \tilde{\varphi}_n(t) - t + d_n$. Then it is clear that if the equation $\tilde{\Phi}_n(t) = 0$ has positive solutions $\tilde{\tau}_n^*$, $\tilde{\tau}_n^{**}$ such that $\tilde{\tau}_n^* \leqq \tilde{\tau}_n^{**}$, then the equation $\Phi_n(t) = 0$ has positive solutions τ_n^*, τ_n^{**} such that $\tau_n^* \leqq \tilde{\tau}_n^* \leqq \tilde{\tau}_n^{**} \leqq \tau_n^{**}$. In particular, we have $0 < \tau_n^* < \tilde{\tau}_n^* \leqq \tilde{\tau}_n^{**} < \tau_n^{**}$ if $\varphi_n(t) < \tilde{\varphi}_n(t)$ for $t > 0$. We first prove that the positive solutions $\tilde{\tau}_n^*$ and $\tilde{\tau}_n^{**}$ do exist.

LEMMA 2.1. The equation $\tilde{\Phi}_n(t) = 0$ has positive solutions so that $\Phi_n(t) = 0$ has positive solutions, too.

<u>Proof</u>. The equation $\tilde{\Phi}_n(t) = 0$ is equivalent to

$$\frac{1}{2} Kt^2 - (1 - \tilde{a}_n - \tilde{b}_n)t + (1 - \tilde{a}_n)d_n = 0.$$

Hence, by noting that $L + M \leqq \sigma K$ and $\sigma K t_n^2 - 2(1 - \ell - m)t_n + 2\eta = 2\nabla t_{n+1} g(t_n) \geqq$ $\geqq 2d_n g(t_n)$, we obtain

$$\begin{aligned}
\tilde{D} &\equiv (1 - \tilde{a}_n - \tilde{b}_n)^2 - 2K(1 - \tilde{a}_n)d_n \\
&\geqq (1 - \ell - m - \sigma K t_n)^2 - 2\sigma K(1 - \ell - Lt_n)d_n \\
&\geqq (1 - \ell - m)^2 + \sigma K(2d_n g(t_n) - 2\eta) - 2\sigma K g(t_n)d_n \\
&= (1 - \ell - m)^2 - 2\sigma K\eta \geqq 0.
\end{aligned}$$

This proves Lemma 2.1. Q.E.D.

LEMMA 2.2. Let τ_n^* be the least solution of the equation $\Phi_n(t) = 0$. Then we have

$$\|x^* - x_n\| \leqq \tau_n^* . \tag{2.3}$$

<u>Proof</u>. By Lemma 2.1, the equations $\Phi_n(t) = 0$ and $\tilde{\Phi}_n(t) = 0$ have positive solutions τ_n^*, τ_n^{**} and $\tilde{\tau}_n^*$, $\tilde{\tau}_n^{**}$ respectively such that $\tau_n^* \leqq \tilde{\tau}_n^* \leqq \tilde{\tau}_n^{**} \leqq \tau_n^{**}$. Let $\tilde{D}$ be defined as in the proof of Lemma 2.1. Then we have

$$\|x^* - x_n\| \leq t^* - t_n = \frac{1 - \sqrt{1 - 2h}}{\sigma K}(1 - \ell - m) - t_n$$

$$\leq \frac{1 - \ell - m - \sigma K t_n}{\sigma K}$$

$$\leq \frac{1 - \tilde{a}_n - \tilde{b}_n}{K}$$

$$\leq \frac{1 - \tilde{a}_n - \tilde{b}_n + \sqrt{\tilde{D}}}{K} = \tilde{\tau}_n^{**} . \tag{2.4}$$

Three cases can occur:

(i) The case where $\Delta_n < t_n$. In this case we have $\varphi_n(t) < \tilde{\varphi}_n(t)$ for all $t > 0$ so that $\tau_n^* < \tilde{\tau}_n^* \leq \tilde{\tau}_n^{**} < \tau_n^{**}$, which, together with (2.4), implies that

$$\|x^* - x_n\| < \tau_n^{**} . \tag{2.5}$$

(ii) The case where $\Delta_n = t_n$ and $\tilde{D} > 0$. In this case, we have $\tilde{\tau}_n^{**} = \tau_n^{**}$ and the inequality $\leq$ in (2.4) is replaced by the strict inequality $<$ so that we again have (2.5).

(iii) The case where $\Delta_n = t_n$ and $\tilde{D} = 0$. In this case, we have $\tau_n^* = \tilde{\tau}_n^* = \tilde{\tau}_n^{**} = \tau_n^{**}$ so that (2.4) means (2.3).

In the cases (i) and (ii), we can also assert (2.3). In fact, we have

$$\|x^* - x_n\| - d_n \leq \|x^* - x_{n+1}\| \leq \varphi_n(\|x^* - x_n\|)$$

or

$$\phi_n(\|x^* - x_n\|) \geq 0 .$$

Solving this inequality yields

$$\|x^* - x_n\| \leq \tau_n^* \quad \text{or} \quad \|x^* - x_n\| \geq \tau_n^{**} .$$

By (2.5), the latter is excluded in the cases (i) and (ii). This proves Lemma 2.2. Q.E.D.

We are now in a position to prove the following theorem:

THEOREM 2.1. Under the assumptions of Theorem 1.1, we have

$$\|x^* - x_n\| \leqq \tau_n^* = \frac{2(1 - a_n)d_n}{1 - a_n - b_n + \sqrt{(1 - a_n - b_n)^2 - 2K(1 - a_n)d_n}} \tag{2.6}$$

$$\leqq \tilde{\tau}_n^* = \frac{2(1 - \tilde{a}_n)d_n}{1 - \tilde{a}_n - \tilde{b}_n + \sqrt{(1 - \tilde{a}_n - \tilde{b}_n)^2 - 2K(1 - \tilde{a}_n)d_n}} \tag{2.7}$$

$$\leqq \frac{t^* - t_n}{\nabla t_{n+1}} d_n \leqq \frac{t^* - t_n}{\nabla t_n} d_{n-1} \leqq t^* - t_n , \tag{2.8}$$

where $d_n = \|x_{n+1} - x_n\|$, $\nabla t_{n+1} = t_{n+1} - t_n$, and a_n, b_n, $\tilde{a}_n$, $\tilde{b}_n$ are defined in (2.1) and (2.2).

Proof. It remains to prove that

$$\tilde{\tau}_n^* \leqq \frac{t^* - t_n}{\nabla t_{n+1}} d_n .$$

From the proof of Lemma 2.1, we already know that

$$\tilde{D} \geqq (1 - \ell - m)^2 - 2\sigma K\eta \geqq 0$$

so that we have

$$\tilde{\tau}_n^* = \frac{2g(t_n)d_n}{1 - \ell - m - (L + M)t_n + \sqrt{\tilde{D}}}$$

$$\leqq \frac{2g(t_n)\nabla t_{n+1}}{1 - \ell - m - \sigma K t_n + \sqrt{(1 - \ell - m)^2 - 2\sigma K\eta}} \cdot \frac{d_n}{\nabla t_{n+1}}$$

$$= \frac{2f(t_n)}{1 - \ell - m - \sigma K t_n + \sqrt{(1 - \ell - m)^2 - 2\sigma K\eta}} \cdot \frac{d_n}{\nabla t_{n+1}}$$

$$= \frac{\sigma K(t_n - t^*)(t_n - t^{**})}{\sigma K(t^{**} - t_n)} \cdot \frac{d_n}{\nabla t_{n+1}}$$

$$= \frac{t^* - t_n}{\nabla t_{n+1}} d_n ,$$

where t^{**} is the largest solution of $f(t) = 0$:

$$t^{**} = \frac{1 - \ell - m + \sqrt{(1 - \ell - m)^2 - 2\sigma Kn}}{\sigma K} . \qquad \text{Q.E.D.}$$

COROLLARY 2.1.1. Under the assumptions of Theorem 1.1, we have

$$\|x^* - x_{n+1}\| \leqq \tau_n^* - d_n \tag{2.9}$$

$$\leqq \frac{t^* - t_{n+1}}{\nabla t_{n+1}} d_n \leqq t^* - t_{n+1} .$$

Proof. We have from Theorem 2.1

$$\|x^* - x_{n+1}\| \leqq \varphi_n(\|x^* - x_n\|) \leqq \varphi_n(\tau_n^*) = \tau_n^* - d_n$$

$$\leqq \frac{t^* - t_n}{\nabla t_{n+1}} d_n - d_n = \frac{t^* - t_{n+1}}{\nabla t_{n+1}} d_n . \qquad \text{Q.E.D.}$$

Remark 2.1. In Theorem 2.1, replace the constants K, L, M and m by G, aG, $(1 - a)G$ and $1 - H$ respectively, where $G > 0$, $0 \leqq a \leqq 1$, $0 < H \leqq 1$. Furthermore, put $\Delta_n = \|x_n - x_0\|$, $Q_n = 1 - aG\Delta_n$, $G_n = G/Q_n$ and $H_n = (H - G\Delta_n)/Q_n$. Then we have $1 - a_n - b_n = H - G\Delta_n$ and

$$\tau_n^* = \frac{H - G\Delta_n}{G} \{1 - \sqrt{1 - \frac{2G(1 - aG\Delta_n)}{(H - G\Delta_n)^2} d_n}\}$$

$$= \frac{H_n}{G_n} \{1 - \sqrt{1 - \frac{2G_n}{H_n^2} d_n}\} . \tag{2.10}$$

This is Moret's bound $(u_n(\|x_{n+1} - x_n\|) - u_n(0)$ in his notation). He obtained (2.10) by replacing (1.4) and (1.5) by the stronger conditions (1.10) and $G\eta/H^2 < 1/2$, respectively. Under his assumptions, the bound (2.9) also reduces to Moret's bound $\beta_n(\|x_{n+1} - x_n\|)$. (See (3.4) in his paper.)

As a dual of our principle, we have

$$d_n - \|x^* - x_n\| \leq \|x^* - x_{n+1}\| \leq \varphi_n(\|x^* - x_n\|) \leq \tilde{\varphi}_n(\|x^* - x_n\|) ,$$

where $\varphi_n(t)$, $\tilde{\varphi}_n(t)$ are quadratic polynomials defined as in (2.1) and (2.2) respectively. Then $\tilde{\Psi}_n(t) = \tilde{\varphi}_n(t) + t - d_n = 0$ always has only one positive solution $\tilde{\underline{\tau}}_n^*$. Therefore, solving the inequality $\tilde{\Psi}_n(\|x^* - x_n\|) \geq 0$ yields the lower estimates $\|x^* - x_n\| \geq \tilde{\underline{\tau}}_n^*$. This is a technique which was first adapted by Gragg-Tapia [5], and later by Schmidt [22] and Miel [11]. If we denote by $\underline{\tau}_n^*$ the unique positive solution of $\Psi_n(t) = \varphi_n(t) + t - d_n = 0$, then we have $\underline{\tau}_n^* \geq \tilde{\underline{\tau}}_n^*$. Hence, as a dual of Theorem 2.1, we have the following result.

THEOREM 2.2. Under the assumptions of Theorem 2.1, we have

$$\|x^* - x_n\| \geq \underline{\tau}_n^* = \frac{2(1 - a_n)d_n}{1 - a_n + b_n + \sqrt{(1 - a_n + b_n)^2 + 2K(1 - a_n)d_n}} \tag{2.11}$$

$$\geq \tilde{\underline{\tau}}_n^* = \frac{2(1 - \tilde{a}_n)d_n}{1 - \tilde{a}_n + \tilde{b}_n + \sqrt{(1 - \tilde{a}_n + \tilde{b}_n)^2 + 2K(1 - \tilde{a}_n)d_n}} . \tag{2.12}$$

Next, we would like to estimate the ratio d_{n+1}/d_n. For the Newton method, it is well known that $d_{n+1}/d_n \leq 1/2$. Moret [12] obtained under his stronger assumptions that

$$d_{n+1} \leq r_n(d_n) = (1 - aG_nd_n)^{-1}(\tfrac{1}{2} G_nd_n + 1 - H_n)d_n , \tag{2.13}$$

where a, G_n, H_n are defined in (2.10). On the other hand, Miel [10] obtained under the Rheinboldt assumptions that

$$d_{n+1} \leq \frac{\nabla t_{n+2}}{\nabla t_{n+1}} d_n . \tag{2.14}$$

Therefore, it would be interesting to compare (2.13) with (2.14). As remarked in §1 (cf. (1.9)), (2.14) holds true under the assumptions of Theorem 1.1.

We have the following result.

THEOREM 2.3. Under the notation and assumptions of Theorem 2.1, let

$$p_n(t) = \frac{1}{2}(1 - a_{n+1})^{-1}(b_n + \sqrt{b_n^2 + 2K(1 - a_{n+1})t}) ,$$

$$q_n(t) = (1 - a_{n+1})^{-1}(\frac{1}{2} Kt^2 + b_n t)$$

and

$$\bar{r}_n(t) = (1 - a_n - Lt)^{-1}(\frac{1}{2} Kt^2 + b_n t) .$$

Then we have

$$d_{n+1} \leq p_n(d_{n+1})d_n \leq q_n(d_n) \leq \bar{r}_n(d_n) \leq \frac{\nabla t_{n+2}}{\nabla t_{n+1}} d_n . \tag{2.15}$$

<u>Proof</u>. We have

$$x_{n+2} - x_{n+1} = -A(x_{n+1})^{-1}\{F(x_{n+1}) - F(x_n) - F'(x_n)(x_{n+1} - x_n)$$

$$+ (F'(x_n) - A(x_n))(x_{n+1} - x_n)\} ,$$

$$A(x_{n+1}) = A(x_0)\{I + A(x_0)^{-1}(A(x_{n+1}) - A(x_0))\} ,$$

and

$$\|A(x_0)^{-1}(A(x_{n+1}) - A(x_0))\| \leq \ell + L\Delta_{n+1} = a_{n+1} \leq a_n + Ld_n < 1 .$$

Hence we obtain

$$d_{n+1} \leq q_n(d_n) \leq \bar{r}_n(d_n) .$$

The inequality $d_{n+1} \leq q_n(d_n)$ is also equivalent to

$$\frac{1}{2} Kd_n^2 + b_n d_n - (1 - a_{n+1})d_{n+1} \geq 0 . \tag{2.16}$$

Now, to prove the first inequality of (2.15), we may assume that $d_{n+1} \neq 0$. Then $p_n(d_{n+1}) > 0$ and solving (2.16) yields

$$d_n \geq \frac{-b_n + \sqrt{b_n^2 + 2K(1 - a_{n+1})d_{n+1}}}{K} = \frac{d_{n+1}}{p_n(d_{n+1})} ,$$

which implies that $d_{n+1} \leq p_n(d_{n+1})d_n$. Furthermore, let $\alpha = d_{n+1}/p_n(d_{n+1})$. Then $(1 - a_{n+1})d_{n+1} = \frac{1}{2} K\alpha^2 + b_n\alpha$ and

$$q_n(d_n) - p_n(d_{n+1})d_n = \frac{\frac{1}{2} Kd_n^2 + b_n d_n - (1 - a_{n+1})d_{n+1}\alpha^{-1}d_n}{1 - a_{n+1}}$$

$$= \frac{\frac{1}{2} Kd_n^2 + b_n d_n - (\frac{K}{2}\alpha + b_n)d_n}{1 - a_{n+1}}$$

$$= \frac{\frac{1}{2} Kd_n(d_n - \alpha)}{1 - a_{n+1}} \geq 0 ,$$

which proves $p_n(d_{n+1})d_n \leq q_n(d_n)$. Finally, let

$$\psi(s,t) = (1 - \ell - Ls - Lt)^{-1}(\frac{K}{2} t + m + Ms) .$$

Then, with the use of the majorant theory of Rheinboldt and Miel's technique, we have

$$\begin{aligned} \bar{r}_n(d_n) &= \psi(\Delta_n, d_n)d_n \\ &\leq \psi(t_n, \nabla t_{n+1})\nabla t_{n+1} \cdot \frac{d_n}{\nabla t_{n+1}} \\ &= \nabla t_{n+2} \cdot \frac{d_n}{\nabla t_{n+1}} , \end{aligned}$$

which completes the proof of Theorem 2.3. Q.E.D.

<u>Remark</u> 2.2. Under the assumptions of Moret, we have $r_n(d_n) = \bar{r}_n(d_n)$. Therefore, Theorem 2.3 improves the results of Miel [10] and Moret [12].

COROLLARY 2.3.1. Under the assumptions of Theorem 2.3, we have $d_{n+1} < d_n$, provided that $d_n \neq 0$.

Proof. As in the proof of Theorem 2.3, it is easy to see that the inequality

$$\frac{1}{2}\sigma K(\nabla t_{n+1})^2 + \tilde{b}_n \cdot \nabla t_{n+1} - (1 - \tilde{a}_{n+1})\nabla t_{n+2} \geq 0$$

holds. We may again assume that $d_{n+1} \neq 0$. Then it follows that

$$\nabla t_{n+1} \geq \frac{2(1 - \tilde{a}_{n+1})\nabla t_{n+2}}{\tilde{b}_n + \sqrt{\tilde{b}_n^2 + 2\sigma K(1 - \tilde{a}_{n+1})\nabla t_{n+2}}} . \tag{2.17}$$

Observe that the denominator of (2.17) is positive, since $\nabla t_{n+2} \geq d_{n+1} > 0$. Furthermore, by a simple computation, we see that the inequality

$$\frac{2(1 - \tilde{a}_{n+1})}{\tilde{b}_n + \sqrt{\tilde{b}_n^2 + 2\sigma K(1 - \tilde{a}_{n+1})\nabla t_{n+2}}} > 1 \tag{2.18}$$

is equivalent to

$$\sigma K \nabla t_{n+2} + 2(Lt_{n+1} + Mt_n) < 2(1 - \ell - m) . \tag{2.19}$$

However, we have assumed that $\eta > 0$ so that $t^* > t_n$ and

$$\begin{aligned}\sigma K \nabla t_{n+2} + 2(Lt_{n+1} + Mt_n) &\leq \sigma K(\nabla t_{n+2} + 2t_{n+1}) \\ &= \sigma K(t_{n+2} + t_{n+1}) \\ &< 2\sigma K t^* \leq 2(1 - \ell - m) .\end{aligned}$$

Hence the condition (2.18) as well as (2.19) is satisfied, which, together with (2.17), implies that $\nabla t_{n+1} > \nabla t_{n+2}$. Consequently we have

$$d_{n+1} \leq p_n(d_{n+1})d_n \leq q_n(d_n) \leq r_n(d_n) \leq \frac{\nabla t_{n+2}}{\nabla t_{n+1}} d_n < d_n ,$$

provided that $d_n \neq 0$. Q.E.D.

We end this section by pointing out that Moret's bounds follow from Theorem 1.1 if the condition (1.4) is replaced by Moret's condition (1.10). In fact, under the same assumptions as in Theorem 1.1, except for replacing (1.4) by (1.10), we have

$$\|A(x_n)^{-1}(F'(x) - F'(y))\| \leq \bar{K}\|x - y\| , \quad x, y \in D_0 ,$$

$$\|A(x_n)^{-1}(A(x) - A(x_n))\| \leq \bar{L} \|x - x_n\| , \quad x \in D_0 ,$$

$$\|A(x_n)^{-1}(F'(x) - A(x))\| \leq \bar{M} \| x - x_n\| + \bar{m} , \quad x \in D_0 ,$$

where

$$\bar{K} = (1 - \bar{a}_n)^{-1}K , \quad \bar{L} = (1 - \bar{a}_n)^{-1}L , \quad \bar{M} = (1 - \bar{a}_n)^{-1}M ,$$

$$\bar{m} = (1 - \bar{a}_n)^{-1}b_n , \quad \bar{a}_n = L\Delta_n , \quad b_n = m + M\Delta_n .$$

Therefore, an application of Theorem 1.1 to x_n leads to

$$\|x^* - x_n\| \leq t^* = \frac{1 - \sqrt{1 - 2\bar{h}}}{\sigma\bar{K}} (1 - \bar{m})$$

$$= \frac{2(1 - \bar{a}_n)d_n}{1 - \bar{a}_n - b_n + \sqrt{(1 - \bar{a}_n - b_n)^2 - 2\sigma K(1 - \bar{a}_n)d_n}} , \tag{2.20}$$

provided that $2\bar{h} \leq 1$, where $\bar{h} = \sigma\bar{K}d_n/(1 - \bar{m})^2$. The condition $2\bar{h} \leq 1$ is indeed satisfied. To show this, we compare the function $\bar{\varphi}_n(t) = (1 - \bar{a}_n)^{-1}(\frac{1}{2}\sigma Kt^2 + b_n t)$ with another function $\hat{\varphi}_n(t) = (1 - \hat{a}_n)^{-1}(\frac{1}{2}\sigma Kt^2 + \tilde{b}_n t)$ where $\hat{a}_n = Lt_n$, $\tilde{b}_n = m + Mt_n$. Then $\bar{\varphi}_n(t) \leq \hat{\varphi}_n(t)$ and $\hat{\Phi}_n(t) = \hat{\varphi}_n(t) - t + d_n = 0$ has positive solutions if $d_n \neq 0$, since we already know from the proof of Lemma 2.1 that

$$\hat{D} \equiv (1 - \hat{a}_n - \tilde{b}_n)^2 - 2\sigma K(1 - \hat{a}_n)d_n$$

$$\geq (1 - m - \sigma Kt_n)^2 - 2\sigma K(1 - Lt_n)d_n$$

$$\geq (1 - m)^2 - 2\sigma K\eta \geq 0 .$$

Hence the equation $\bar{\Phi}_n(t) = \bar{\varphi}_n(t) - t + d_n = 0$ has positive solutions and

$$\bar{D} = (1 - \bar{a}_n - b_n)^2 - 2\sigma K(1 - \bar{a}_n)d_n \geq 0 ,$$

which is equivalent to $2\bar{h} \leq 1$. If $\sigma = 1$, which is also Moret's case, then the bounds (2.20) coincide with (2.6) with $\ell = 0$. Therefore, the bounds (2.20) include those of Moret as a special case. However, if $\sigma > 1$, then the bounds (2.20) are inferior to the corresponding bounds of (2.6) in spite of putting the stronger condition (1.10) in place of (1.4).

Remark 2.3. It also follows from the above argument that $2\bar{h} < 1$ if $2h < 1$. This generalizes Moret's result that $2G_n d_n/H_n^2 < 1$ if $2h < 1$.

3. APPLICATIONS TO OTHER ITERATIONS

The arguments developed in the previous section may be applied to the other type iterations. First we consider the iteration (1.11) considered by Dennis [3]. He assumed that $x_0 \in D_0$, A_0^{-1} exists, $\|A_0^{-1}F(x_0)\| \leq \eta$,

$$\|A_0^{-1}(F'(x) - F'(y))\| \leq K\|x - y\| , \quad x, y \in D_0 ,$$

$$\|A_0^{-1}(F'(x) - A_n)\| \leq \begin{cases} \delta_0 \ (n = 0) \\ \delta_n + \gamma \sum_{j=1}^{n} \|x_j - x_{j-1}\| \ (n \geq 1) , \end{cases}$$

$$\delta_n \leq \delta \ (n \geq 1) , \quad \delta_0 + 2\delta < 1 ,$$

$$h = \frac{(K + 2\gamma)\eta}{(1 - \delta_0 - 2\delta)^2} \leq \frac{1}{2}$$

and

$$\bar{S}(x_0, t^*) \subseteq D_0 .$$

where

$$t^* = \frac{1 - \sqrt{1 - 2h}}{K + 2\gamma} (1 - \delta_0 - 2\delta) .$$

Under these assumptions, he proved that the sequence $\{x_n\}$ generated by (1.11) exists, remains in $\bar{S}(x_0, t^*)$ and converges to a solution x^* of (1.1), which is unique in $D_0 \cap S(x_0, \frac{1 + \sqrt{1 - 2h'}}{K}(1 - \delta_0))$ if $2h' < 1$ where $h' = K\eta/(1 - \delta_0)^2$, and unique in $\bar{S}(x_0, \frac{1 - \delta_0}{K})$ if $2h' = 1$. Furthermore, defining the sequence $\{t_n\}$ by

$$t_0 = 0, \quad t_{n+1} = t_n + \frac{f(t_n)}{g_n} , \quad n = 0, 1, 2, \ldots ,$$

where

$$f(t) = \frac{1}{2}(K + 2\gamma)t^2 - (1 - \delta_0 - 2\delta)t + \eta ,$$

$$g_n = \begin{cases} 1 \ (n = 0) \\ 1 - \delta_0 - \delta_{\alpha_n} - (K + \gamma)t_{\alpha_n} \ (n \geq 1) \end{cases} ,$$

Dennis showed that

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$$

and

$$\|x^* - x_n\| \leq t^* - t_n, \quad n = 0, 1, 2, \ldots . \tag{3.1}$$

We can apply our principle to improve the error bounds (3.1). In fact, we have

$$x^* - x_{n+1} = -A_{\alpha_n}^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)$$

$$+ (F'(x_n) - A_{\alpha_n})(x^* - x_n)] ,$$

$$F'(x_n) - A_{\alpha_n} = \{F'(x_n) - F'(x_{\alpha_n})\} + \{F'(x_{\alpha_n}) - A_{\alpha_n}\} ,$$

$$A_{\alpha_n} = A_0 \{I + A_0^{-1}(A_{\alpha_n} - A_0)\}$$

and

$$A_{\alpha_n} - A_0 = \{A_{\alpha_n} - F'(x_{\alpha_n})\} + \{F'(x_{\alpha_n}) - F'(x_0)\} + \{F'(x_0) - A_0\} .$$

Hence we have

$$\|x^* - x_{n+1}\| \leq \frac{\frac{1}{2} K\|x^* - x_n\|^2 + (\delta_{\alpha_n} + K\|x_n - x_{\alpha_n}\| + \gamma \sum_{j=1}^{\alpha_n} \|x_j - x_{j-1}\|)\|x^* - x_n\|}{1 - \delta_0 - \delta_{\alpha_n} - K\|x_{\alpha_n} - x_0\| - \gamma \sum_{j=1}^{\alpha_n} \|x_j - x_{j-1}\|} ,$$

where we understand that $\sum_{j=1}^{0} = 0$. Therefore, if we again assume that

$d_n = \|x_{n+1} - x_n\| \neq 0$ and put

$$a_n = \delta_0 + \delta_{\alpha_n} + K\|x_{\alpha_n} - x_0\| + \gamma \sum_{j=1}^{\alpha_n} d_{j-1} , \tag{3.2}$$

$$\tilde{a}_n = \delta_0 + \delta_{\alpha_n} + (K + \gamma)t_{\alpha_n} , \tag{3.3}$$

$$b_n = \delta_{\alpha_n} + K\|x_n - x_{\alpha_n}\| + \gamma \sum_{j=1}^{\alpha_n} d_{j-1} , \tag{3.4}$$

$$\tilde{b}_n = \delta_{\alpha_n} + K(t_n - t_{\alpha_n}) + \gamma t_{\alpha_n} , \tag{3.5}$$

Then we have

$$\|x^* - x_{n+1}\| \leq \varphi_n(\|x^* - x_n\|) \leq \tilde{\varphi}_n(\|x^* - x_n\|)$$

where $\varphi_n(t)$ and $\tilde{\varphi}_n(t)$ are of the forms defined in (2.1) and (2.2), respectively. Furthermore, it is easy to see that

$$\begin{aligned}
\tilde{D} &\equiv (1 - \tilde{a}_n - \tilde{b}_n)^2 - 2K(1 - \tilde{a}_n)d_n \\
&\geq (1 - \delta_0 - 2\delta - (K + 2\gamma)t_n)^2 - 2(K + 2\gamma)(1 - \delta_0 - \delta_{\alpha_n} - (K + \gamma)t_{\alpha_n})d_n \\
&\geq (1 - \delta_0 - 2\delta - (K + 2\gamma)t_n)^2 - 2(K + 2\gamma)g_n \nabla t_{n+1} \\
&= (1 - \delta_0 - 2\delta)^2 - 2(K + 2\gamma)\eta \geq 0 \, .
\end{aligned}$$

Therefore, repeating the same arguments as in §2, we again obtain the estimates (2.6) - (2.8) with the a_n, b_n, $\tilde{a}_n$, $\tilde{b}_n$ defined in (3.2) - (3.5).

Next, consider the iteration (1.12) which was discussed by Schmidt [21], [22]. He assumed that F is Fréchet differentiable in D_0 and for some $x_1, y_1 \in D_0$, $\delta F(x_1,y_1)^{-1} \in L(Y,X)$ exists. Furthermore, he assumed that, with *some constants* $a > 0$, $b \geq 0$, $c > 0$, the following hold:

$$\|\delta F(x_1,y_1)^{-1}(F'(u) - F'v))\| \leq 2a\|u - v\|, \quad u, v \in D_0 \, ,$$

$$\|\delta F(x_1,y_1)^{-1}(\delta F(u,v) - F'(x))\| \leq a(\|u - x\| + \|v - x\|), \quad u, v, x \in D_0 \, ,$$

$$\|x_1 - y_1\| \leq b \, , \quad \|\delta F(x_1,y_1)^{-1}F(x_1)\| \leq c \, ,$$

$$h = \frac{2a(b + c)}{(1 + ab)^2} \leq \frac{1}{2} \, , \quad t^* = \frac{1 + ab}{2a}(1 - \sqrt{1 - 2h}) \, ,$$

$$\bar{S} = \bar{S}(x_2, t^* - b - c) \subseteq D_0 \, ,$$

$$y_1 = x_0 \, , \quad y_n = \lambda_n x_n + (1 - \lambda_n)x_{n-1}, \quad \lambda_n \in [0,1], \quad n \geq 2 \, .$$

Then he proved that the sequence $\{x_n\}$ generated by (1.12) exists, remains in $\bar{S}$ and converges to a solution x^* of (1.1). His proof consists in establishing the relations

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \quad \text{and} \quad \|x^* - x_n\| \leq t^* - t_n , \tag{3.6}$$

where the sequence $\{t_n\}$ is defined by

$$t_1 = b , \quad s_1 = t_0 = 0 ,$$

$$t_{n+1} = t_n - \frac{(t_n - s_n)f(t_n)}{f(t_n) - f(s_n)} ,$$

$$s_{n+1} = \lambda_{n+1}t_{n+1} + (1 - \lambda_{n+1})t_n , \quad n = 1, 2, 3, \ldots,$$

with $f(t) = at^2 - (1 + ab)t + b + c$.

Under the assumptions of Schmidt, we can improve the bounds (3.6). In fact, as was shown in his paper [21], $\delta F(x_n,y_n)^{-1}$ exists for each n and we have

$$x^* - x_{n+1} = \delta F(x_n,y_n)^{-1}[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)$$

$$+ \{F'(x_n) - \delta F(x_n,y_n)\}(x^* - x_n)]$$

so that

$$\|x^* - x_{n+1}\| \leq \varphi_n(\|x^* - x_n\|) = (1 - a_n)^{-1}(a\|x^* - x_n\|^2 + b_n\|x^* - x_n\|) ,$$

where

$$a_n = a(\|x_n - y_n\| + 2\|y_n - x_1\| + \|x_1 - y_1\|) , \quad b_n = a\|x_n - y_n\|$$

and

$$\varphi_n(t) = (1 - a_n)^{-1}(at^2 + b_nx) .$$

Therefore, if we put

$$\tilde{a}_n = a(t_n + s_n - t_1) , \quad \tilde{b}_n = a(t_n - s_n) ,$$

and

$$\tilde{\varphi}_n(t) = (1 - \tilde{a}_n)^{-1}(at^2 + \tilde{b}_nt) ,$$

then $\varphi_n(t) \leq \tilde{\varphi}_n(t)$ for $t > 0$ and, by the same argument as in §2, we obtain

$$\tilde{\underline{\tau}}_n^* = \frac{2(1-\tilde{a}_n)d_n}{1-\tilde{a}_n+\tilde{b}_n+\sqrt{(1-\tilde{a}_n+\tilde{b}_n)^2+4a(1-\tilde{a}_n)d_n}}$$

$$\leq \underline{\tau}_n^* = \frac{2(1-a_n)d_n}{1-a_n+b_n+\sqrt{(1-a_n+b_n)^2+4a(1-a_n)d_n}}$$

$$\leq \|x^*-x_n\| \leq \tau_n^* = \frac{2(1-a_n)d_n}{1-a_n-b_n+\sqrt{(1-a_n-b_n)^2-4a(1-a_n)d_n}} \tag{3.7}$$

$$\leq \tilde{\tau}_n^* = \frac{2(1-\tilde{a}_n)d_n}{1-\tilde{a}_n-\tilde{b}_n+\sqrt{(1-\tilde{a}_n-\tilde{b}_n)^2-4a(1-\tilde{a}_n)d_n}}$$

$$\leq \frac{t^*-t_n}{\nabla t_{n+1}} d_n \leq \frac{t^*-t_n}{\nabla t_n} d_{n-1} \leq t^*-t_n .$$

The bound $\tilde{\underline{\tau}}_n^*$, the positive solution of $\tilde{\varphi}_n(t)+t-d_n=0$, is equal to Schmidt's lower bound u_n in his paper [22].

As the third example, we consider the iteration (1.13) which was considered by Zincenko [27], Rheinboldt [20] and Moret [12]. Let F be continuous in D and T be Fréchet differentiable on some open convex set $D_0 \subseteq D$. Assume that for $x_0 \in D_0$, $T'(x_0)^{-1} \in L(Y,X)$ exists and for $K > 0$, $0 \leq \delta < 1$, $\eta > 0$,

$$\|T'(x_0)^{-1}(T'(x)-T'(y))\| \leq K\|x-y\|, \quad x, y \in D_0 ,$$

$$\|T'(x_0)^{-1}((F-T)(x)-(F-T)(y))\| \leq \delta\|x-y\|, \quad x, y \in D_0 ,$$

$$\|T'(x_0)^{-1}F(x_0)\| \leq \eta$$

and

$$h = \frac{K\eta}{(1-\delta)^2} \leq \frac{1}{2} .$$

Define t^* and t^{**} by

$$t^* = \frac{1 - \sqrt{1 - 2h}}{K}(1 - \delta) \quad \text{and} \quad t^{**} = \frac{1 + \sqrt{1 - 2h}}{K}(1 - \delta)$$

respectively and suppose that $\bar{S}(x_0, t^*) \subseteq D_0$. Then it is known [27], [20] that the sequence $\{x_n\}$ defined by (1.13) exists, remains in $\bar{S}(x_0, t^*)$ and converges to the only solution x^* of (1.1) in $D_0 \cap S(x_0, t^{**})$.

Rheinboldt proved this by showing that

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \quad \|x^* - x_n\| \leq t^* - t_n \tag{3.8}$$

where the sequence $\{t_n\}$ is defined by

$$t_0 = 0, \quad t_{n+1} = t_n + \frac{f(t_n)}{g(t_n)}, \quad n = 0, 1, 2, \ldots,$$

with

$$f(t) = \frac{1}{2}Kt^2 - (1 - \delta)t + \eta, \quad g(t) = 1 - Kt.$$

Therefore we can improve the bounds (3.8) on the basis of our principle: For the iteration (1.13), we have

$$x^* - x_{n+1} = -T'(x_n)^{-1}[T(x^*) - T(x_n) - T'(x_n)(x^* - x_n)$$
$$+ (F - T)(x^*) - (F - T)(x_n)]$$

so that

$$\|x^* - x_{n+1}\| \leq \varphi_n(\|x^* - x_n\|) \equiv (1 - K\Delta_n)^{-1}(\frac{1}{2}K\|x^* - x_n\|^2 + \delta\|x^* - x_n\|)$$
$$\leq \tilde{\varphi}_n(\|x^* - x_n\|) \equiv (1 - Kt_n)^{-1}(\frac{1}{2}K\|x^* - x_n\|^2 + \delta\|x^* - x_n\|),$$

where $\Delta_n = \|x_n - x_0\|$. Hence, an application of our technique yields

$$\tilde{\underline{\tau}}^*_n = \frac{2(1 - Kt_n)d_n}{1 + \delta - Kt_n + \sqrt{(1 + \delta - Kt_n)^2 + 2K(1 - Kt_n)d_n}}$$

$$\leq \underline{\tau}_n^* = \frac{2(1 - K\Delta_n)d_n}{1 + \delta - K\Delta_n + \sqrt{(1 + \delta - K\Delta_n)^2 + 2K(1 - K\Delta_n)d_n}}$$

$$\leq \|x^* - x_n\| \leq \tau_n^* = \frac{2(1 - K\Delta_n)d_n}{1 - \delta - K\Delta_n + \sqrt{(1 - \delta - K\Delta_n)^2 - 2K(1 - K\Delta_n)d_n}}$$

$$\leq \tilde{\tau}_n^* = \frac{2(1 - Kt_n)d_n}{1 - \delta - Kt_n + \sqrt{(1 - \delta - Kt_n)^2 - 2K(1 - Kt_n)d_n}}$$

$$\leq \frac{t^* - t_n}{\nabla t_{n+1}} d_n \leq \frac{t^* - t_n}{\nabla t_n} d_{n-1} \leq t^* - t_n .$$

The upper bounds τ_n^* coincide with those of Moret, which he derived under the assumption $2h < 1$.

4. APPLICATION TO NEWTON'S METHOD

As a special case of the iteration (1.3), we consider the Newton method (1.2). We call the assumptions of Theorem 1.1 with $K = L$, $\ell = M = m = 0$ the Kantorovich assumptions. Then we obtain the following result.

THEOREM 4.1. Under the Kantorovich assumptions, we have

$$\underset{\sim}{\tilde{\tau}}_n^* = \frac{2d_n}{1 + \sqrt{1 + 2K(1 - Kt_n)^{-1}d_n}} \tag{4.1}$$

$$= \frac{2d_n}{1 + \sqrt{1 + 2KB_n d_n}} \tag{4.2}$$

$$\leq \underset{\sim}{\tau}_n^* = \frac{2d_n}{1 + \sqrt{1 + 2K(1 - K\Delta_n)^{-1}d_n}} \tag{4.3}$$

$$\leq \|x^* - x_n\| \leq \tau_n^* = \frac{2d_n}{1 + \sqrt{1 - 2K(1 - K\Delta_n)^{-1}d_n}} \tag{4.4}$$

$$\leq \tilde{\tau}_n^* = \frac{2d_n}{1 + \sqrt{1 - 2K(1 - Kt_n)^{-1}d_n}} \tag{4.5}$$

$$= \frac{2d_n}{1 + \sqrt{1 - 2KB_n d_n}} \tag{4.6}$$

$$\leq \frac{t^* - t_n}{\nabla t_{n+1}} d_n \tag{4.7}$$

$$\leq \frac{t^* - t_n}{(\nabla t_n)^2} d_{n-1}^2 \tag{4.8}$$

where $\Delta_n = \|x_n - x_0\|$ and B_n are defined as follows (cf. Kantorovich [6], Rall [18]):

$$B_0 = 1 , \quad \eta_0 = \eta , \quad h_0 = h = K\eta ,$$

$$B_n = \frac{B_{n-1}}{1 - h_{n-1}} , \qquad \eta_n = \frac{h_{n-1}\eta_{n-1}}{2(1 - h_{n-1})} , \quad h_n = KB_n\eta_n , \quad n \geq 1 .$$

Proof. The bounds (4.1), (4.3) - (4.5) and (4.7) follow from Theorems 2.1 and 2.2. The bounds (4.8) are found in Miel [10]. Therefore, it remains to prove that $(1 - Kt_n)^{-1} = B_n$. This fact is implicitly found in Kantorovich-Akilov [8]. However, we can also prove this by using the relations

$$t^* - t_n = \frac{2\eta_n}{1 + \sqrt{1 - 2h_n}} \quad \text{and} \quad t_{n+1} - t_n = \eta_n ,$$

which were proved in the previous paper [26]. (See Proposition A.2 in the Appendix of this paper.) In fact, we have

$$1 - Kt_n = 1 - K(t^* - \frac{2\eta_n}{1 + \sqrt{1 - 2h_n}})$$

$$= \sqrt{1 - 2h} + \frac{1 - \sqrt{1 - 2h_n}}{B_n} = B_n^{-1},$$

since

$$\sqrt{1 - 2h_n} = \sqrt{1 - (\frac{h_{n-1}}{1 - h_{n-1}})^2} = \frac{\sqrt{1 - 2h_{n-1}}}{1 - h_{n-1}} = \ldots$$

$$= \frac{\sqrt{1 - 2h_0}}{(1 - h_{n-1}) \ldots (1 - h_0)} = B_n \sqrt{1 - 2h}. \tag{4.9}$$

This proves (4.2) and (4.6). Q.E.D.

Remark 4.1. The bounds (4.4) follow also from the Kantorovich theorem. In fact, under the Kantorovich assumptions, $F'(x_n)^{-1}$ exist and we have

$$\|F'(x_n)^{-1}(F'(x) - F'(y))\| \leq (1 - K\Delta_n)^{-1}K\|x - y\|, \quad x, y \in D_0,$$

and

$$(1 - K\Delta_n)^{-1} \leq B_n, \quad n = 0, 1, 2, \ldots. \tag{4.10}$$

The inequalities (4.10), which we obtained in the proof of Theorem 4.1, follow also by induction on n:

$$\begin{aligned}(1 - K\Delta_n)^{-1} &\leq (1 - K\Delta_{n-1} - Kd_{n-1})^{-1}\\ &= (1 - K\Delta_{n-1})^{-1}(1 - K(1 - K\Delta_{n-1})^{-1}d_{n-1})^{-1}\\ &\leq B_{n-1}(1 - KB_{n-1}d_{n-1})^{-1}\\ &\leq B_{n-1}(1 - KB_{n-1}\eta_{n-1})^{-1} = B_n.\end{aligned}$$

Hence we have

$$2k(1 - K\Delta_n)^{-1}d_n \leq 2KB_nd_n \leq 2h_n \leq 1$$

so that an application of the Kantorovich theorem to x_n and $\eta = d_n$ leads to the bounds (4.4). As was remarked in [25], [26], the bounds (4.6) also follow from the Kantorovich theorem by replacing x_0 and η in the theorem by x_n and d_n respectively. However, Theorem 4.1 asserts that (4.6) is equal to $\tilde{\tau}_n^*$. Furthermore, we remark that it is shown in a series of papers [24] - [26] that the bounds (4.6) are sharper than those of Gragg-Tapia [5], Potra-Pták [17] and Miel [11]. Therefore, under the Kantorovich assumptions, the Kantorovich theorem still gives us the best upper bounds. We note that the bounds (4.4) also follow from Moret's bounds (2.10), provided that $2h < 1$. Finally, we note that Schmidt's lower bounds [22] for the iteration (1.12) reduce to (4.1) and the bounds (4.2) may be found implicitly in Miel [11].

APPENDIX: ERROR BOUNDS FOR NEWTON'S METHOD UNDER THE KANTOROVICH ASSUMPTIONS.

After Kantorovich gave a proof of his theorem for the Newton method, many authors have made efforts to find sharper error bounds under the same hypotheses. In this appendix, we survey such results and clarify the relationships among them.

In 1948, Kantorovich [6] established his theorem by proving that $\|x_{n+1} - x_n\| \leqq \eta_n$ and

$$\|x^* - x_n\| \leqq \frac{2\eta_n}{1 + \sqrt{1 - 2h_n}} \leqq 2\eta_n \leqq \frac{1}{2^{n-1}} (2h)^{2^n - 1} \eta \,, \tag{A.1}$$

where η_n and h_n are defined in Theorem 4.1. A year later [7], he gave another proof by showing that

$$\|x_{n+1} - x_n\| \leqq t_{n+1} - t_n \,, \quad \|x^* - x_n\| \leqq t^* - t_n \,, \tag{A.2}$$

where the sequence $\{t_n\}$ is defined by (1.7) and (1.8) with $\sigma = 1$, $\ell = m = 0$, $L = K$. Since then, it seems to the author that numerical analysts are convinced that the bounds (A.2) are sharper than (A.1). However, we can prove [26] that $2\eta_n/(1 + \sqrt{1 - 2h_n}) = t^* - t_n$. (Undoubtedly Kantorovich had known this fact.) For the sake of convenience, we give the proof here. We begin by proving the following:

PROPOSITION A.1. Let t^* and t^{**} be the smallest and largest solutions of $f(t) = \frac{1}{2} Kt^2 - t + \eta = 0$, respectively. Furthermore, set $\theta = t^*/t^{**}$ and $\Delta = t^{**} - t^* = 2\sqrt{1 - 2h}/K$. Then

$$t^* - t_n = \begin{cases} \dfrac{\Delta\theta^{2^n}}{1 - \theta^{2^n}} & (2h < 1) \\[2ex] \dfrac{1}{2^n K} & (2h = 1) \end{cases}$$

and

$$t^{**} - t_n = \frac{\Delta}{1 - \theta^{2^n}} \quad (2h < 1) .$$

Proof. This proposition is essentially due to Ostrowski [16; Appendix F]. Let $a_n = t^* - t_n$ and $b_n = t^{**} - t_n$. Then we have

$$a_{n+1} = a_n - \frac{a_n b_n}{a_n + b_n} = \frac{a_n^2}{a_n + b_n} , \quad b_{n+1} = b_n - \frac{a_n b_n}{a_n + b_n} = \frac{b_n^2}{a_n + b_n} . \tag{A.3}$$

Hence

$$\frac{a_n}{b_n} = \left(\frac{a_{n-1}}{b_{n-1}}\right)^2 = \ldots = \left(\frac{a_0}{b_0}\right)^{2^n} = \theta^{2^n}$$

and

$$b_n - a_n = t^{**} - t^* = \Delta ,$$

which lead to

$$a_n = \frac{\Delta\theta^{2^n}}{1 - \theta^{2^n}} , \quad b_n = \frac{\Delta}{1 - \theta^{2^n}}$$

if $\theta < 1$. If $\theta = 1$, then we have $a_n = b_n$ and (A.3) implies that

$$a_n = \frac{1}{2} a_{n-1} = \ldots = \frac{1}{2^n} t^* = \frac{1}{2^n K} . \qquad \text{Q.E.D.}$$

PROPOSITION A.2. We have

$$t_{n+1} - t_n = \eta_n ,$$

$$t^* - t_n = \frac{1 - \sqrt{1 - 2h_n}}{KB_n} = \frac{2\eta_n}{1 + \sqrt{1 - 2h_n}}$$

and

$$t^{**} - t_n = \frac{1 + \sqrt{1 - 2h_n}}{KB_n} , \quad n = 0, 1, 2, \ldots .$$

That is, $t^* - t_n$ and $t^{**} - t_n$ are the solutions of the equations $\frac{1}{2} KB_n t^2 - t + \eta_n = 0$.

Proof. As was shown by Gragg-Tapia [5], we have

$$\theta^{2^n} = \left(\frac{1 - \sqrt{1 - 2h}}{1 + \sqrt{1 - 2h}}\right)^{2^n} = \frac{1 - \sqrt{1 - 2h_n}}{1 + \sqrt{1 - 2h_n}} . \tag{A.4}$$

Hence, if $\theta < 1$, then we have

$$\frac{\theta^{2^n}}{1 - \theta^{2^n}} = \frac{1 - \sqrt{1 - 2h_n}}{2\sqrt{1 - 2h_n}} = \frac{1 - \sqrt{1 - 2h_n}}{2B_n\sqrt{1 - 2h}} \quad \text{(cf. (4.9))} ,$$

and

$$\frac{1}{1 - \theta^{2^n}} = \frac{1 + \sqrt{1 - 2h_n}}{2\sqrt{1 - 2h_n}} = \frac{1 + \sqrt{1 - 2h_n}}{2B_n\sqrt{1 - 2h}} .$$

Consequently, we have from Proposition A.1

$$t^* - t_n = \frac{2\sqrt{1 - 2h}}{K} \cdot \frac{1 - \sqrt{1 - 2h_n}}{2B_n\sqrt{1 - 2h}} = \frac{1 - \sqrt{1 - 2h_n}}{KB_n}$$

and

$$t^{**} - t_n = \frac{2\sqrt{1 - 2h}}{K} \cdot \frac{1 + \sqrt{1 - 2h_n}}{2B_n\sqrt{1 - 2h}} = \frac{1 + \sqrt{1 - 2h_n}}{KB_n} ,$$

if $\theta < 1$. These relations hold true for $\theta = 1$, since $B_n = 2^n$ and $2h_n = 1$ if $\theta = 1$. Furthermore, we have

$$t_{n+1} - t_n = (t^* - t_n) - (t^* - t_{n+1})$$

$$= \frac{1 - \sqrt{1 - 2h_n}}{KB_n} - \frac{1 - \sqrt{1 - 2h_{n+1}}}{KB_{n+1}}$$

$$= \frac{1 - \sqrt{1 - 2h_n}}{KB_n} - \frac{1 - h_n}{KB_n}(1 - \sqrt{1 - 2h_{n+1}})$$

$$= \frac{h_n}{KB_n} = \eta_n ,$$

since $(1 - h_n)\sqrt{1 - 2h_{n+1}} = \sqrt{1 - 2h_n}$. Q.E.D.

PROPOSITION A.3. The following relations hold:

(i) $$B_n = \frac{1}{1 - Kt_n} = \frac{1}{\sqrt{1 - 2h + (K\eta_{n-1})^2}} = \frac{1}{K\eta_n + \sqrt{1 - 2h + (K\eta_n)^2}}$$

$$= \begin{cases} \dfrac{\sqrt{1 - 2h_n}}{\sqrt{1 - 2h}} = \dfrac{2}{\Delta K} \cdot \dfrac{1 - \theta^{2^n}}{1 + \theta^{2^n}} & (2h < 1) \\ 2^n & (2h = 1) . \end{cases} \tag{A.5}$$

(ii) $$\frac{\nabla t_{n+1}}{(\nabla t_n)^2} = \frac{K}{2(1 - Kt_n)} = \frac{1}{2} KB_n .$$

(iii) $$\frac{t^* - t_n}{(\nabla t_n)^2} = \frac{KB_n}{1 + \sqrt{1 - 2h_n}} = \begin{cases} \dfrac{1 - \theta^{2^n}}{\Delta} & (2h < 1) \\ \dfrac{2^{n-1}}{\eta} & (2h = 1) . \end{cases}$$

Proof. (i) In the proof of Theorem 4.1, it was shown that $B_n\sqrt{1 - 2h} = \sqrt{1 - 2h_n}$ and $B_n^{-1} = 1 - Kt_n$, $n \geq 0$. Furthermore, if $n = 1$, then we have $\sqrt{1 - 2h + (K\eta)^2} = 1 - h = B_1^{-1}$. If $n \geq 2$, then we obtain

$$B_n^{-2} = \prod_{i=0}^{n-2} (1 - h_i)^2 = (1 - 2h_{n-1}) \prod_{i=0}^{n-2} (1 - h_i)^2 + \left\{h_{n-1} \prod_{i=0}^{n-2} (1 - h_i)\right\}^2$$

$$= (1 - 2h_{n-1})B_{n-1}^{-2} + (h_{n-1}B_{n-1}^{-1})^2$$

$$= 1 - 2h + (K\eta_{n-1})^2 .$$

This proves the second equality so that we have

$$K\eta_n + \sqrt{1 - 2h_n + (K\eta_n)^2} = K\eta_n + B_{n+1}^{-1} = K\eta_n + (1 - h_n)B_n^{-1} = B_n^{-1} .$$

The last relation (A.5) follows from Gragg-Tapia's relation (A.4). In fact, we have

$$\sqrt{1 - 2h_n} = \frac{1 - \theta^{2^n}}{1 + \theta^{2^n}} .$$

(ii) *The second relation follows from (i). The first relation is well known and is* derived as follows:

$$t_{n+1} - t_n = \frac{f(t_n)}{f'(t_n)}$$

$$= \frac{1}{1 - Kt_n} \left\{f(t_n) - f(t_{n-1}) - f'(t_{n-1})(t_n - t_{n-1})\right\}$$

$$= \frac{1}{1 - Kt_n} \cdot \frac{1}{2} f''(\xi)(t_n - t_{n-1})^2 \qquad (t_{n-1} < \xi < t_n)$$

$$= \frac{K}{2(1 - Kt_n)} (t_n - t_{n-1})^2 .$$

(iii) It follows from Proposition A.2 that, if $\theta < 1$, then

$$\frac{1 - \theta^{2^n}}{\Delta} = \frac{K}{2\sqrt{1 - 2h}} \cdot \frac{2\sqrt{1 - 2h_n}}{1 + \sqrt{1 - 2h_n}}$$

$$= \frac{KB_n}{2\eta_n} \cdot \frac{2\eta_n}{1 + \sqrt{1 - 2h_n}}$$

$$= \frac{1}{\eta_{n-1}^2}(t^* - t_n) = \frac{t^* - t_n}{(t_n - t_{n-1})^2} .$$

If $\theta = 1$, then we have

$$\frac{t^* - t_n}{(t_n - t_{n+1})^2} = \frac{2\eta_n}{\eta_{n-1}^2} = \frac{1}{\eta_{n-1}} = \frac{2^{n-1}}{\eta} . \qquad \text{Q.E.D.}$$

Throughout this appendix, we keep the Kantorovich assumptions. Therefore, according to Ostrowski [15], [16], we can take a constant $\alpha \geq 2$ such that $\alpha h = 1$. Then, there exists a unique constant $\varphi \geq 0$ such that $\alpha = 1 + \cosh\varphi = 1 + 2^{-1}(e^{\varphi} + e^{-\varphi})$. We can prove the following:

PROPOSITION A.4 Let α and φ be defined as above. Then we have

$$\text{(i)} \quad t^* - t_{n+1} = \begin{cases} e^{-2^n\varphi} \dfrac{\sinh\varphi}{\sinh 2^n\varphi}\eta = \dfrac{e^{\varphi} - e^{-\varphi}}{e^{2^n\varphi}(e^{2^n\varphi} - e^{-2^n\varphi})}\eta & (2h < 1) \\[2ex] \dfrac{\eta}{2^n} = \lim\limits_{\varphi\to+0} \left(e^{-2^n\varphi} \dfrac{\sinh\varphi}{\sinh 2^n\varphi}\eta\right) & (2h = 1) , \end{cases}$$

$$\text{(ii)} \quad \frac{t^* - t_{n+1}}{\Delta t_{n+1}} = \theta^{2^n} = e^{-2^n\varphi} .$$

<u>Proof</u>. Define the sequence $\{\alpha_n\}$ by

$$\alpha_0 = \alpha , \quad \alpha_n = 1 + \cosh 2^n\varphi = 2(\cosh 2^{n-1}\varphi)^2 .$$

Then we have $\alpha_{n+1} = 2(\alpha_n - 1)^2$ and $\alpha_n h_n = 1$. In fact, by induction on n, we have

$$\alpha_{n+1}h_{n+1} = 2(\alpha_n - 1)^2 \frac{KB_n}{1 - h_n} \cdot \frac{h_n\eta_n}{2(1 - h_n)} = \left(\frac{\alpha_n h_n - h_n}{1 - h_n}\right)^2 = 1 .$$

Therefore

$$\theta^{2^n} = \frac{1 - \sqrt{1 - 2\alpha_n^{-1}}}{1 + \sqrt{1 - 2\alpha_n^{-1}}} = \frac{\sqrt{\alpha_n} - \sqrt{\alpha_n - 2}}{\sqrt{\alpha_n} + \sqrt{\alpha_n - 2}}$$

$$= \frac{\cosh 2^{n-1}\varphi - \sinh 2^{n-1}\varphi}{\cosh 2^{n-1}\varphi + \sinh 2^{n-1}\varphi} = e^{-2^n\varphi} .$$

Furthermore, we have

$$\frac{t^* - t_{n+1}}{\nabla t_{n+1}} = \frac{1}{\eta_n} \cdot \frac{2\eta_{n+1}}{1 + \sqrt{1 - 2h_{n+1}}} = \frac{h_n}{1 - h_n} \cdot \frac{1}{1 + \sqrt{1 - 2h_{n+1}}}$$

$$= \frac{h_n}{1 - h_n + \sqrt{1 - 2h_n}} = \frac{1 - \sqrt{1 - 2h_n}}{1 + \sqrt{1 - 2h_n}} = \theta^{2^n} ,$$

which proves (ii). To prove (i), we observe that

$$\frac{\Delta}{\eta} = \frac{2}{h}\sqrt{1 - 2h} = 2\sqrt{\alpha(\alpha - 2)} = 2 \sinh \varphi .$$

Hence, we obtain from (ii) and Proposition A.1,

$$t^* - t_{n+1} = (2 \sinh \varphi)\eta \frac{e^{-2^{n+1}\varphi}}{1 - e^{-2^{n+1}\varphi}} = e^{-2^n\varphi} \frac{\sinh \varphi}{\sinh 2^n\varphi} \eta ,$$

provided that $\varphi > 0$. Q.E.D.

<u>Remark</u> A.1. Ostrowski [16] chose a constant $\alpha \geq 2$ such that $\alpha h \leq 1$. Then the above proof implies that $\alpha_n h_n \leq 1$ and $\theta^{2^n} \leq e^{-2^n\varphi}$ where the equalities hold if and only if $\alpha h = 1$. Therefore, the best choice of α is $\alpha = h^{-1}$.

PROPOSITION A.5. We have

$$\frac{1}{KB_n}(1 - \sqrt{1 - 2h_n}) \leq \frac{1}{2^n K}(1 - \sqrt{1 - 2h})^{2^n} , \quad n \geq 0 .$$

Proof. Define the sequences $\{\alpha_n\}$ and $\{\beta_n\}$ by

$$\alpha_0 = \beta_0 = t^* = \frac{1}{K}(1 - \sqrt{1 - 2h}) ,$$

$$\alpha_n = \frac{1}{2^n K}(1 - \sqrt{1 - 2h})^{2^n} , \quad \beta_n = \frac{1}{KB_n}(1 - \sqrt{1 - 2h_n}) , \quad n \geq 1 .$$

Then, they satisfy the recurrence relations

$$\alpha_n = 2^{n-2}K\alpha_{n-1}^2 , \quad \beta_n = \frac{1}{2} KB_{n-1}\beta_{n-1}^2 , \quad n \geq 1 .$$

In fact, we have

$$\beta_n = \frac{1 - h_{n-1}}{KB_{n-1}}\left(1 - \sqrt{1 - \left(\frac{h_{n-1}}{1 - h_{n-1}}\right)^2}\right) = \frac{1}{KB_{n-1}}\left(1 - h_{n-1} - \sqrt{1 - 2h_{n-1}}\right)$$

$$= \frac{KB_{n-1}}{2}\left(\frac{1 - \sqrt{1 - 2h_{n-1}}}{KB_{n-1}}\right)^2 = \frac{1}{2} KB_{n-1}\beta_{n-1}^2 , \quad n \geq 1 .$$

Furthermore, we have $\frac{1}{2} KB_{n-1} \leq 2^{n-2}K$. Hence, by induction on n, we obtain $\alpha_n \geq \beta_n$, $n \geq 0$. Q.E.D.

PROPOSITION A.6. We have

$$Kd_n + \sqrt{1 - 2h + (Kd_n)^2} \leq \sqrt{1 - 2h + (Kd_{n-1})^2} , \quad n \geq 1 ,$$

where $d_n = \|x_{n+1} - x_n\|$.

Proof. As a special case of Theorem 2.3, we have

$$d_n \leq (1 - a_n)^{-1} \cdot \frac{1}{2} Kd_{n-1}^2 \leq (1 - \tilde{a}_n)^{-1} \cdot \frac{1}{2} Kd_{n-1}^2 = \frac{1}{2} KB_n d_{n-1}^2 ,$$

where $a_n = K\|x_n - x_0\|$ and $\tilde{a}_n = Kt_n$. Hence, we obtain from Proposition A.3 (i)

$$d_n \leq \frac{Kd_{n-1}^2}{2\sqrt{1 - 2h + (K\eta_{n-1})^2}} \leq \frac{Kd_{n-1}^2}{2\sqrt{1 - 2h + (Kd_{n-1})^2}} , \tag{A.6}$$

since $d_n \leq \eta_n$. It follows from (A.6) that

$$\left(\sqrt{1 - 2h + (Kd_{n-1})^2} - Kd_n\right)^2 \geq 1 - 2h + (Kd_n)^2 . \tag{A.7}$$

The expression in the parenthesis in the left-hand side is non-negative, since we have $d_n \leq \frac{1}{2} d_{n-1}$ from (A.6). Therefore, (A.7) means that

$$\sqrt{1 - 2h + (Kd_{n-1})^2} - Kd_n \geq \sqrt{1 - 2h + (Kd_n)^2} .$$

This proves Proposition A.6. Q.E.D.

On the basis of Propositions A.1 - A.6, we have the following chart of the upper bounds for the errors of the Newton sequence $\{x_n\}$, provided that the Kantorovich assumptions are satisfied:

$$\frac{1}{2^{n-1}} (2h)^{2^n - 1} \eta \quad \text{(Kantorovich [6])}$$

$$\geq \frac{1}{2^n K} (1 - \sqrt{1 - 2h})^{2^n} \quad \text{(Dennis [1], Tapia [23])}$$

$$\geq \frac{2\eta_n}{1 + \sqrt{1 - 2h}} = \frac{1 - \sqrt{1 - 2h_n}}{KB_n} \quad \text{(Kantorovich [6])}$$

$$= t^* - t_n \quad \text{(Kantorovich [7])}$$

$$= \frac{\sigma_n}{2^n K} (1 - \sqrt{1 - 2h})^{2^n} \left(\sigma_0 = 1, \ \sigma_n = \frac{\sigma_{n-1}^2}{2^{n-1}\sqrt{1 - 2h} + \sigma_{n-1}(1 - \sqrt{1 - 2h})^{2^{n-1}}}\right) \quad \text{(Rall - Tapia [19])}$$

$$= \begin{cases} e^{-2^{n-1}\varphi} \dfrac{\sinh \varphi}{\sinh 2^{n-1}\varphi} \eta & (2h < 1) \\ 2^{1-n}\eta \quad (2h = 1) & (h^{-1} = 1 + \cosh \varphi, \ \varphi \geq 0) \end{cases} \quad \text{(Ostrowski [15])}$$

$$= \begin{cases} \dfrac{2}{K} \sqrt{1 - 2h} \ \dfrac{\theta^{2^n}}{1 - \theta^{2^n}} & (2h < 1) \\ 2^{1-n}\eta \quad (2h = 1) & (\theta = t^*/t^{**}) \end{cases} \quad \text{(Gragg-Tapia [5])}$$

$$\geq e^{-2^{n-1}\varphi} d_{n-1} \quad \text{(Ostrowski [16])}$$

$$= \theta^{2^{n-1}} d_{n-1} \quad \text{(Gragg-Tapia [5])}$$

$$= \frac{t^* - t_n}{\nabla t_n} d_{n-1} \quad \text{(Miel [10])}$$

$$= \frac{1}{\eta_{n-1}} \left(\frac{2\eta_n}{1 + \sqrt{1 - 2h_n}}\right) d_{n-1}$$

$$= \frac{K\eta_{n-1}}{B_{n-1}^{-1} + \sqrt{1 - 2h}} d_{n-1}$$

$$= \frac{K\eta_{n-1}}{\sqrt{1 - 2h + (K\eta_{n-1})^2} + \sqrt{1 - 2h}} d_{n-1}$$

$$\geq \frac{Kd_{n-1}^2}{\sqrt{1 - 2h + (Kd_{n-1})^2} + \sqrt{1 - 2h}} \quad \text{(Potra-Pták [17])}$$

$$\geq \frac{Kd_{n-1}^2}{B_n^{-1} + \sqrt{1 - 2h}}$$

$$= \frac{KB_n d_{n-1}^2}{1 + \sqrt{1 - 2h_n}}$$

$$= \begin{cases} \dfrac{1 - \theta^{2^n}}{\Delta} d_{n-1}^2 & (2h < 1) \qquad (\Delta = t^{**} - t^*) \\ \dfrac{2^{n-1}}{\eta} d_{n-1}^2 & (2h = 1) \qquad \text{(Miel [11])} \end{cases}$$

$$= \frac{t^* - t_n}{(\nabla t_n)^2} d_{n-1}^2 \quad \text{(Miel [10])}$$

$$\gtrsim \frac{2d_n}{1+\sqrt{1-2h_n}}$$

$$= \frac{t^* - t_n}{\nabla t_{n+1}} d_n \qquad \text{(Theorem 4.1)}$$

$$\gtrsim \frac{2d_n}{1+\sqrt{1-2KB_n d_n}} \qquad \text{(Yamamoto [25])}$$

$$= \begin{cases} \dfrac{2d_n}{1+\sqrt{1-\dfrac{4}{\Delta}\cdot\dfrac{1-\theta^{2^n}}{1+\theta^{2^n}}\, d_n}} & (2h < 1) \\[2ex] \dfrac{2d_n}{1+\sqrt{1-\dfrac{2^n}{\eta}\, d_n}} & (2h = 1) \end{cases}$$

$$= \tilde{\tau}_n^* = \frac{2d_n}{1+\sqrt{1-2K(1-Kt_n)^{-1}d_n}} \qquad \text{(Theorem 4.1)}$$

$$\gtrsim \tau_n^* = \frac{2d_n}{1+\sqrt{1-2K(1-K\Delta_n)^{-1}d_n}} \qquad \text{(Moret [12], Theorem 4.1)}$$

$$\gtrsim \|x^* - x_n\| .$$

Similarly we have the following chart for the lower bounds:

$$\frac{2d_n}{1+\sqrt{1+2h_n}} \qquad \text{(Gragg-Tapia [5])}$$

$$= \frac{2d_n}{1+\sqrt{1+\dfrac{2K\eta_n}{K\eta_n+\sqrt{1-2h+(K\eta_n)^2}}}}$$

$$\leq \frac{2d_n}{1 + \sqrt{1 + \dfrac{2Kd_n}{Kd_n + \sqrt{1 - 2h + (Kd_n)^2}}}} \qquad \text{(Potra-Pták [17])}$$

$$\leq \frac{2d_n}{1 + \sqrt{1 + \dfrac{2Kd_n}{\sqrt{1 - 2h + (Kd_{n-1})^2}}}}$$

$$\leq \frac{2d_n}{1 + \sqrt{1 + \dfrac{2Kd_n}{\sqrt{1 - 2h + (K\eta_{n-1})^2}}}}$$

$$= \frac{2d_n}{1 + \sqrt{1 + 2KB_n d_n}}$$

$$= \begin{cases} \dfrac{2d_n}{1 + \sqrt{1 + \dfrac{4}{\Delta} \cdot \dfrac{1 - \theta^{2^n}}{1 + \theta^{2^n}} d_n}} & (2h < 1) \\[2ex] \dfrac{2d_n}{1 + \sqrt{1 + \dfrac{2^n}{\eta} d_n}} & (2h = 1) \quad \text{(Miel [11])} \end{cases}$$

$$= \tilde{\tau}_n^{*} = \frac{2d_n}{1 + \sqrt{1 + 2K(1 - Kt_n)^{-1} d_n}} \qquad \text{(Schmidt [22])}$$

$$\leq \tau_n^* = \frac{2d_n}{1 + \sqrt{1 + 2K(1 - K\Delta_n)^{-1} d_n}} \quad \text{(Theorem 4.1)}$$

$$\leq \|x^* - x_n\| .$$

Acknowledgements. The author would like to express his thanks to Professor Louis B. Rall of the Mathematics Research Center, University of Wisconsin, for his helpful comments during the preparation of this paper. The author is also grateful to Professor George J. Miel of University of Nevada-Las Vegas for his comments, which also helped to improve the paper.

REFERENCES

[1] J. E. Dennis, Jr., On the Kantorovich hypotheses for Newton's method, SIAM J. Numer. Anal., 6 (1968), pp. 493-507.

[2] ———, On the convergence of Newton-like methods, in Numerical Methods for Nonlinear Algebraic Equations, edited by P. Rabinowitz, Gordon and Breach, New York, 1970.

[3] ———, Toward a unified convergence theory for Newton-like methods, in Nonlinear Functional Analysis and Applications, edited by L. B. Rall, Academic Press, New York, 1971.

[4] P. Deuflhard and G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods, SIAM J. Numer. Anal., 16 (1979), pp. 1-10.

[5] W. B. Gragg and R. A. Tapia, Optimal error bounds for the Newton-Kantorovich theorem, SIAM J. Numer. Anal., (1974), pp. 10 13.

[6] L. V. Kantorovich, On Newton's method for functional equations, Dokl. Akad. Nauk. SSSR, 59 (1948), pp. 1237-1240.

[7] ———, On Newton's method, Trudy Mat. Inst. Steklov, 28 (1949), pp. 104-144.

[8] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, Oxford, 1964.

[9] G. J. Miel, The Kantorovich theorem with optimal error bounds, Amer. Math. Monthly, 86 (1979), pp. 212-215.

[10] ———, Majorizing sequences and error bounds for iterative methods, Math. Comp., 34 (1980), pp. 185-202.

[11] ———, An updated version of the Kantorovich theorem for Newton's method, Computing, 27 (1981), pp. 237-244.

[12] I. Moret, A note on Newton type iterative methods, Computing, 33 (1984), pp. 65-73.

[13] J. M. Ortega, The Newton-Kantorovich theorem, Amer. Math. Monthly, 75 (1968), pp. 658-660.

[14] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variable, Academic Press, New York, 1970.

[15] A. M. Ostrowski, La méthode de Newton dans les espaces de Banach, C. R. Acad. Sci. Paris, 27 (A)(1971), pp. 1251-1253.

[16] ———, Solution of Equations in Euclidean and Banach Spaces, Academic Press, New York, 1973.

[17] F. A. Potra and V. Pták, Sharp error bounds for Newton's process, Numer. Math., 34 (1980), pp. 63-72.

[18] L. B. Rall, Computational Solution of Nonlinear Operator Equations, Krieger, Huntington, New York, 1979.

[19] L. B. Rall and R. A. Tapia, The Kantorovich theorem and error estimates for Newton's method, MRC Technical Summary Report #1043, University of Wisconsin, 1970.

[20] W. C. Rheinboldt, A unified convergence theory for a class of iterative processes, SIAM J. Numer. Anal., 5 (1968), pp. 42-63.

[21] J. W. Schmidt, Regular-falsi-Verfahren mit konsistenter Steigung und Majoranten-prinzip, Periodica Math. Hungarica, 5 (1974), pp. 187-193.

[22] ———, Untere Fehlerschranken für Regular-falsi-Verfahren, Periodica Math. Hungarica, 9 (1978), pp. 241-247.

[23] R. A. Tapia, The Kantorovich theorem for Newton's method, Amer. Math. Monthly, 78 (1971), pp. 389-392.

[24] T. Yamamoto, Error bounds for Newton's process derived from the Kantorovich theorem, to appear in Japan J. Appl. Math.

[25] ———, Error bounds for Newton's iterates derived from the Kantorovich theorem, MRC Technical Summary Report #2843, University of Wisconsin, 1985.

[26] ———, A unified derivation of several error bounds for Newton's process, J. Comp. Appl. Math., 12 & 13 (1985), pp. 179-191.

[27] A. I. Zincenko, A class of approximate methods for solving operator equations with nondifferentiable operators, Depovidi, Akad. Nauk. Ukrain, RSR (1963), pp. 156-161.

TY/jp

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: #2846	2. GOVT ACCESSION NO.: AD-A160 994
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER KANTOROVICH TYPE ASSUMPTIONS	5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Tetsuro Yamamoto	8. CONTRACT OR GRANT NUMBER(s): DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 3 - Numerical Analysis and Scientific Computing
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709	12. REPORT DATE: July 1985
	13. NUMBER OF PAGES: 39
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)

Newton-like methods, Newton's method, a posteriori error estimates, the Newton-Kantorovich theorem, Kantorovich type assumptions, Dennis' theorem, Rheinboldt's theorem, Miel's bounds, Moret's bounds

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper gives a method to derive new a posteriori error bounds for Newton-like methods in a Banach space under Kantorovich type assumptions. The bounds found are sharper than those of Miel [10] and include those recently obtained by Moret [12]. The applicability of our method is studied for other types of iterations. Various error bounds for the Newton method under the Kantorovich assumptions are surveyed in the Appendix.

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

12-85

DTIC