A CONVERGENCE THEOREM FOR NEWTON'S METHOD IN BANACH SPACES (U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER T YAMAMOTO OCT 85 NRC-TSR-2679 DAAG29-80-C-0041
A CONVERGENCE THEOREM FOR NEWTON'S METHOD IN BANACH SPACES

Tetsuro Yamamoto

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

October 1985

(Received September 17, 1985)
A CONVERGENCE THEOREM FOR NEWTON'S METHOD IN BANACH SPACES

Tetsuro Yamamoto

Technical Summary Report #2879
October 1985

ABSTRACT

On the basis of the results obtained in a series of papers [25] - [28], a convergence theorem for Newton's method in Banach spaces is given, which improves the theorems of Kantorovich [4], Lancaster [8] and Ostrowski [16]. The error bounds obtained improve the recent results of Potra [17].

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

Key Words: convergence theorem, Newton's method, Kantorovich's theorem, Lancaster's theorem, Ostrowski's theorem, error estimates, Potra's bounds

Work Unit Number 3 (Numerical Analysis and Scientific Computing)

*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 in Japan.
SIGNIFICANCE AND EXPLANATION

To find sharper error bounds for iterative solution of nonlinear equations under assumptions as weak as possible is of basic importance in numerical analysis. This paper gives a convergence theorem for Newton's method in Banach spaces which improves the theorems of Kantorovich [4], Lancaster [8] and Ostrowski [16]. The error bounds obtained improve the recent results of Potra [17].

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.
A CONVERGENCE THEOREM FOR NEWTON'S METHOD IN BANACH SPACES
Tetsuro Yamamoto

1. Introduction

There is much literature concerning convergence and error estimates for Newton's method in Banach spaces. In a series of papers [25] - [28], we examined the error bounds which have been obtained by many authors (Dennis [1], Tapia [24], Rall-Tapia [20], Ostrowski [15], [16], Gragg-Tapia [3], Miel [9] - [11], Potra-Pták [18], Moret [12]) under the assumptions of the Kantorovich theorem, and compared them with the Kantorovich bounds. As the result, we concluded [28] that their results follow from the Kantorovich theorem so that, under the Kantorovich assumptions, the Kantorovich theorem still gives the best upper bounds for the Newton method.

In this paper, we are interested in improving the assumptions of the Kantorovich theorem and the assertions of the Ostrowski theorem [16; Theorem 38.1]. We shall first state both theorems and several lemmas in §2. Next, in §3, we shall present a convergence theorem which improves both theorems. It will also be shown that results improve the error bounds of Lancaster [8], Kornstad [7] and Potra [17]. Finally, in §4, we shall show that Ostrowski's other theorem [16; Theorem 38.2] can be derived by our approach.

2. Preliminaries

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

*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 in Japan.
F(x) = 0 \quad (2.1)

is defined by

\[ x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) , \quad n \geq 0 , \quad (2.2) \]

provided that \( F'(x_n)^{-1} \in \mathcal{L}(Y,X) \) exists at each step, where \( \mathcal{L}(Y,X) \) denotes the Banach space of bounded linear operators of \( Y \) into \( X \). Sufficient conditions for convergence of the iterates (2.2), error estimates and existence and uniqueness regions of solutions are given by the famous Kantorovich theorem:

Theorem 2.1 (Kantorovich [4], [5] and Kantorovich-Akilov [6]). Let \( F : D \subseteq X + Y \to X \) be Fréchet differentiable in an open convex set \( D_0 \subseteq D \). Assume that for some \( x_0 \in D_0 \), \( F'(x_0) \) is invertible and that

\[
\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq K|x - y|, \quad K > 0, \quad x, y \in D_0 ,
\]

\[
\|F'(x_0)^{-1}F(x_0)\| \leq h, \quad h > 0 ,
\]

\[
h = K \eta \leq \frac{1}{2}
\]

and

\[
S(x_0,t^*) = \{ x \in X \mid \|x - x_0\| < t^* = \frac{1}{1 + \sqrt{1 - 2h}} \} \subseteq D_0 .
\]

Then:

(i) The iterates (2.2) are well-defined, lie in the open ball \( S(x_0,t^*) = \{ x \in X \mid \|x - x_0\| < t^* \} \) and converge to a solution \( x^* \) of the equation (2.1).

(ii) The solution \( x^* \) is unique in \( S(x_0,t^{**}) \cap D_0 \) if \( 2h < 1 \) and in \( \tilde{S}(x_0,t^{**}) \) if \( 2h = 1 \), where \( t^{**} = (1 + \sqrt{1 - 2h})/K \).

(iii) Error estimates

\[
\|x^* - x_n\| \leq \frac{2^n \eta_n}{1 + \sqrt{1 - 2h_n}} \leq 2^{1 - n(2h)^{2^n - 1} n} , \quad n \geq 0 , \quad (2.3)
\]

hold, where \( \eta_n \) and \( h_n \) are defined by the recurrence relations
B_0 = 1, \quad n_0 = n, \quad h_0 = h = Kn,

B_n = \frac{B_{n-1}}{1 - h_{n-1}}, \quad n_n = \frac{h_{n-1} h_{n-1}}{2(1 - h_{n-1})}, \quad h = \frac{Kn}{n}, \quad n \geq 1. \quad (2.4)

(iii) Put \( f(t) = \frac{1}{2} t^2 - t + n \) and define the sequence \( \{t_n\} \) by

\[ t_0 = 0, \quad t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad n \geq 0. \]

Then

\[ |x_{n+1} - x| \leq \frac{t_n - t_{n-1}}{t_n}, \quad n \geq 0, \]

holds.

The bounds (2.3) are of the form found in (4), while the bound (2.5) is found in (5) and (6). We should remark here that \( B_n \) and \( n_n \) are the bounds for \( \|F(x_n)^{-1}F'(x_0)\| \) and \( |x_{n+1} - x| \), respectively. In fact, by induction on \( n \), we have

\[ \|F(x_n)^{-1}F'(x_0)\| \leq \frac{B_{n-1}}{1 - B_{n-1}Kx_n - x_{n-1}} \leq \frac{B_{n-1}}{1 - h_{n-1}} = B_n, \]

\[ x_{n+1} - x_n = -F'(x_n)^{-1}F(x_n) \]

\[ = -F'(x_n)^{-1}[F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})] \]

\[ = -F'(x_n)^{-1} \int_0^1 [F'(x_n + t(x_n - x_{n-1})) - F'(x_{n-1})](x_n - x_{n-1}) dt \quad (2.6) \]

and

\[ |x_{n+1} - x_n| \leq \frac{1}{2} B_n Kx_n - x_{n-1}^2 \leq \frac{1}{2} B_n K^2 n \equiv n_n. \]

On the other hand, Ostrowski [15], [16] proved the convergence of the Newton method under the assumptions which are slightly different from those of Kantorovich.
Theorem 2.2 (Ostrowski [15; Theorems 38.1 and 40.2]). Let $F: D \subseteq X + Y$ and $D^0$ be the interior of $D$. Assume that for some $x_0 \in D^0$, $F'(x_0)$ and $F'(x_0)^{-1}$ exist. Let $\varphi \geq 0$, $\alpha = 1 + \cosh \varphi$, $\rho = e^{-\varphi}x_1 - x_0$ and $\sigma = \alpha F(x_0) - F'(x_0)^{-1}x_1^2$. Consider the line segment $L = \{tx_0 + (1 - t)x_1 | 0 \leq t \leq 1\}$ and the closed ball $\bar{S} = \bar{S}(x_1, \rho)$, and put $C = L \cup \bar{S}$. Assume now that $C \subseteq D^0$, $F$ is Fréchet differentiable on $C$ and
\[
F'(x) - F'(y) \leq \frac{1}{\sigma} \left| x - y \right|, \quad x, y \in L, \quad x, y \in \bar{S}.
\] (2.7)
Then the Newton iterates (2.2) are well-defined, $x_n \in \bar{S}$, $n \geq 1$, and $\{x_n\}$ converges to a solution $x^* \in \bar{S}$ of (2.1), which is unique in $C$. Furthermore, the following inequalities hold:
\[
\left| x_n - x^* \right| \leq e^{-2n-\varphi} \frac{\sinh \varphi}{\sinh 2n-\varphi} \left| x_1 - x_0 \right|, \quad n \geq 0
\] (2.8)
\[
\left| x_n - x^* \right| \leq e^{-n} \left| x_1 - x_0 \right|, \quad n \geq 0.
\] (2.9)

In [28], we derived (2.8) and (2.9) under the assumptions of Theorem 2.1 and showed that they do not improve Gragg-Tapia's bounds. Furthermore, we proved that Moret's bounds, which also follow from Theorem 2.1, are sharper than those of Gragg-Tapia, Potra-Pták and Miel. The argument in [28] also works under the assumptions of an affine invariant version of Theorem 2.2, which are weaker than those of Theorem 2.1 with
\[n = \left| x_1 - x_0 \right|,\] provided that $x_0 \neq x_1$. Therefore, on the basis of results obtained in [28], we can improve Theorems 2.1 and 2.2.

Before giving an improved version of both theorems, we state several lemmas. In the following, without loss of generality, we assume that $F(x_0) \neq 0$. This assumption will be kept throughout this paper.

Lemma 2.1. Let $F: D \subseteq X + Y$ and $D^0$ be the interior of $D$. Assume that for some $x_0 \in D^0$, $F'(x_0)$ and $F'(x_0)^{-1}$ exist and $F(x_0) \neq 0$. Let $\varphi \geq 0$, $\alpha = 1 + \cosh \varphi$, 
-4-
\(\eta = |x_1 - x_0|\) and \(\rho = e^{-\eta}\). Define the sets \(L, S\) and \(C\) as in Theorem 2.2.

Furthermore, assume that \(C \subseteq D^0\), \(F\) is Fréchet differentiable on \(C\) and, for some \(K > 0\),

\[
|F'(x_0)^{-1}(F'(x) - F'(y))| \leq K|x - y|, \quad x, y \in L, \quad x, y \in S, \tag{2.10}
\]

and \(\alpha_h \equiv \alpha K_n \leq 1\).

Then the iterates (2.2) are well-defined, \(x_n \in S = S(x_1, \rho)\) (open ball), \(n \geq 1\) and

\[
|\lambda_{n+1} - \lambda_n| \leq \eta_{n+1} - \eta_n, \quad n \geq 0,
\]

where \(\{\eta_n\}\) is the majorizing sequence defined in Theorem 2.1. Therefore, the sequence \(\{\lambda_n\}\) converges to a solution \(\lambda^* \in S\) of (2.1) and

\[
|\lambda^* - \lambda_n| \leq \alpha \eta_n - \eta_n.
\]

Proof. By the assumption \(\alpha h \leq 1\), we have \(2h \leq 1\) since \(\alpha \geq 2\). Therefore, the majorant theory of Kantorovich can be applied to the sequence \(\{\lambda_n\}\), by noting that the condition (2.10) holds and that

\[
|\lambda_{n+1} - \lambda_1| \leq |\lambda_{n+1} - \lambda_n| + |\lambda_n - \lambda_1|
\]

\[
\leq (\eta_{n+1} - \eta_n) + (\eta_n - \eta_1) = \eta_{n+1} - \eta_1
\]

\[
< (t^* - t_1) = \frac{1 - h - \sqrt{1 - 2h}}{K}
\]

\[
< e^{-\eta} \rho = \rho, \tag{2.11}
\]

where equality holds in (2.11) if and only if \(\alpha h = 1\).

Q.E.D.

Lemma 2.2. Under the assumptions of Lemma 2.1, define the sequences \(\{t_n\}, \{n_n\}, \{\eta_n\}\) and \(\{h_n\}\) as in Theorem 2.1. Then

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

\[
t^* - t_n = \frac{1 - \sqrt{1 - 2h}}{h_n} = \frac{2\eta_n}{1 + \sqrt{1 - 2h}}
\]

and

\[-5-\]
That is, \( t^* - t_n \) and \( t^{**} - t_n \) are the solutions of the equation
\[
\frac{1}{2} n K_B n t^2 - t + n = 0.
\]

Proof. The same proof as in [28] works under the assumptions of Lemma 2.1. Q.E.D.

Lemma 2.3. Under the assumptions of Lemma 2.2, we have for \( n \geq 1 \)

(i) \[
S_n = 1 - Kt_n = \sqrt{1 - 2h + (K^n - 1)^2}
\]

(ii) \[
\frac{1}{2} K_B n = \frac{V_{t,n+1}}{(V_t)^2},
\]

(iii) \[
\frac{K_B n}{1 + \sqrt{1 - 2h}} = \frac{t^* - t_n}{(V_t)^2},
\]

where \( V \) denotes the backward difference operator.

Proof. See the proof of Proposition A.3 in [28]. Q.E.D.

Lemma 2.4. Under the assumptions of Lemma 2.2, let \( \theta = t^*/t^{**} = \frac{(1 - \sqrt{1 - 2h})}{(1 + \sqrt{1 - 2h})} \). Then we have for \( n \geq 0 \)

(i) \[
t^* - t_n = \begin{cases} 
 2^{1-n} & (2h = 1), \\
 2^{1-n} \theta^{2^n} & (2h < 1)
\end{cases}
\]

(ii) \[
\frac{t^* - t_{n+1}}{V_{t,n+1}} = \theta^{2^n}.
\]

Proof. See the proofs of Proposition A.1 and Proposition A.4 (ii) in [28]. Q.E.D.
Lemma 2.5. Under the notation and assumptions of Lemma 2.4, we have for \( n \geq 0 \)

\[
\begin{align*}
(1) \quad t^* &= t_n, \\
(2) \quad \phi^2 &\leq \phi^2.
\end{align*}
\]

The equalities hold in (i) and (ii) if and only if \( \phi_n = 1 \).

Proof. Take \( \phi \geq 0 \) such that \( \phi_n = (1 + \cosh \phi^2) > 1 \). Then, by Proposition A.4 in [28], we have that the equalities hold in (i) and (ii). Therefore Lemma 2.5 follows for every \( \phi \in [0, \phi^*] \), since the right hand-sides of (i) and (ii) are monotone decreasing with respect to \( \phi \). Q.E.D.

We end this section by proving the following lemma.

Lemma 2.6. Under the assumptions of Lemma 2.1, define the sequence \( \{B_n\} \) as in

Theorem 2.1. If, for some \( n \), there exists a constant \( M_n > 0 \) such that

\[
1_{x^*} - x_{n+1}^2 \leq \frac{1}{2} M_n x^* - x_n^2
\]

and \( M_n \leq KB_n \). Then

\[
1_{x^*} - x_n^2 \leq \frac{2d_n}{1 + \sqrt{1 - 2d_n}}
\]

where \( d_n = 1_{x^*} - x_n^2 \).

Proof. Without loss of generality, we may assume that \( d_n \neq 0 \), that is \( x^* \neq x_n \).

Then, by assumptions, we have

\[
1_{x^*} - x_n^2 \leq \frac{1}{2} M_n x^* - x_n^2 \leq \frac{1}{2} KB_n x^* - x_n^2.
\]

Hence, if we put

\[
\phi(t) = \frac{1}{2} M_n t^2 - t + d_n \quad \text{and} \quad \tilde{\phi}(t) = \frac{1}{2} KB_n t^2 - t + d_n,
\]

-7-
then
\[ \tilde{\phi}(t^* - x_n) > \phi(t^* - x_n) > 0 \]
and
\[ \tilde{\phi}(t) > \phi(t) \text{ for } t > 0 \] or \( \tilde{\phi}(t) \geq \phi(t) \).

By Lemmas 2.1 and 2.2, we have
\[ |x^* - x_n| < t^* - t_n \]
and \( t^* - t_n, t^{**} - t_n \) are two solutions of the equation \( \Psi(t) = \frac{1}{2} KB t^2 - t + \eta_n = 0 \).

Furthermore we have \( \Psi(t) = \tilde{\psi}(t) \). Therefore \( \phi(t) \) and \( \tilde{\phi}(t) \) have positive solutions
\[ a_n, \tilde{a}_n \text{ and } a_n^{**}, \tilde{a}_n^{**} \text{ respectively such that} \]
\[ a_n < \tilde{a}_n < t - t_n < t^* - t_n < \tilde{a}_n^{**} < a_n^{**} \text{ if } M_n < KB_n, \]
\[ a_n = \tilde{a}_n < t - t_n < t^* - t_n < \tilde{a}_n^{**} = a_n^{**} \text{ if } M_n = KB_n \text{ and } d_n < \eta_n \]
and
\[ a_n = \tilde{a}_n = t - t_n < t^* - t_n < t - \tilde{a}_n^{**} = a_n^{**} \text{ if } M_n = KB_n \text{ and } d_n = \eta_n \. \]

In any case we have
\[ |x^* - x_n| < a_n, \]

since \( \phi( |x^* - x_n| ) > 0 \) implies \( |x^* - x_n| < a_n \) or \( |x^* - x_n| > a_n^{**} \)
and the latter case can be excluded.

Q.E.D.

3. Results

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

Theorem 3.1. Under the assumptions of Lemma 2.1, the following results hold:

1. The iterates (2.2) are well-defined, the sequence \( \{x_n\} \), \( n \geq 1 \) remains in an open ball \( S = S(x_1, \rho) \) and converges to a solution \( x^* \in S \) of the equation (2.1).
(ii) The solution is unique in $C$. Let

$$K_n = \sup_{x,y \in S} \frac{1}{|x-y|} \left| f'(x_n) - f'(y) \right|, \quad n \geq 1,$$

and put $d_n = |x_{n+1} - x_n|$. Then the following error estimates hold.

(a) A posteriori error estimates:

$$|x^* - x_n| \leq \frac{2d_n}{1 + \sqrt{1 - 2K_n d_n}} \quad (n \geq 0) \quad (3.1)$$

$$t^* = \begin{cases} \frac{2n}{1 + \sqrt{1 - 2h}} & (n = 0) \\ \frac{2d_n}{1 + \sqrt{1 - 2K_n d_n}} & (n \geq 1) \end{cases} \quad (n \geq 0) \quad (3.2)$$

$$\frac{2d_n}{1 + \sqrt{1 - 2K_n d_n}} \quad (n \geq 0) \quad (3.3)$$

$$\frac{d_n}{\sqrt{\kappa_{n+1}}} \quad (n \geq 0) \quad (\text{Miel} [10])$$

$$\frac{d_n^2}{\kappa_n} \quad (n \geq 1) \quad (\text{Potra} - \text{Pták} [18])$$

$$\frac{n^2}{\sqrt{1 - 2h + (2d_{n-1})^2 + \sqrt{1 - 2h}}} \quad (n \geq 1) \quad (\text{Miel} [10])$$
\[
\begin{align*}
&= e^{2n-1} d_{n-1} (n \geq 1) \quad \text{(Gragg-Tapia [3])} \\
&\leq e^{-2n-1} d_{n-1} (n \geq 1) \quad \text{(Ostrowski [16]).} \\
\end{align*}
\]

The equality holds in (3.4) if and only if \( ah = 1 \).

(b) A priori error estimates:

\[
\| x^* - x_n \| \leq \| t^* - t_n \| (n \geq 0) \quad \text{(Kantorovich [5], [6])}
\]

\[
\leq \begin{cases}
\frac{2}{\sqrt{1 - 2h}} \frac{\theta_n^2}{1 - e^{2n}} & (2h < 1) \\
2^{-n} (2h = 1) & (n \geq 0) \quad \text{(Gragg-Tapia [3])}
\end{cases}
\]

\[
\leq \begin{cases}
e^{-2n-1} \frac{\sinh \frac{\theta_n}{2}}{\sinh 2^{-n-1} \frac{\theta_n}{2}} & (\theta > 0) \\
2^{-n} & (\theta = 0) \quad (n \geq 0) \quad \text{(Ostrowski [16])}
\end{cases}
\]

The equality holds in (3.5) if and only if \( ah = 1 \).

(iv) If \( F \) is Fréchet differentiable in an open convex set \( D_0 \) such that \( D^0 \supseteq D_0 \supseteq C \) and if \( F'(x) \) satisfies the Lipschitz condition in \( D_0 \) with the Lipschitz constant \( K \), then the solution \( x^* \) is unique if

\[
\tilde{S} = \begin{cases}
S(x_0, t^{**}) \cap D_0 & \text{if } 2h < 1 \\
S(x_0, t^{**}) \cap D_0 & \text{if } 2h = 1
\end{cases}
\]

Furthermore, (3.2) may be replaced by the sharper bound

\[
\| x^* - x_n \| \leq \frac{2d_n}{1 + \sqrt{1 - 2K(1 - \lambda_n)^{-1} d_n}} (n \geq 0) \quad \text{(Moret [12])}
\]

where \( \lambda_n = x_n - x_0 \).

-10-
Proof. (i) was proved in Lemma 2.1. To prove (ii), let $\tilde{x}$ be a solution in $\tilde{S}$.

Then we have

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

$$
= - F'(x_n)^{-1} [F(\tilde{x}) - F(x) - F'(x_n)(\tilde{x} - x_n)]
$$

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

and

$$
x_n + t(\tilde{x} - x_n), \ x_n \in \tilde{S}, \ n \geq 1.
$$

Hence, by (2.10) and Lemma 2.3 (ii) we have

$$
1x^* - x_{n+1}^n \leq \frac{1}{2} b_n K|x^* - x_n|^2 = \frac{\nu_{n+1}}{2} |x^* - x_n|^2
$$

so that

$$
\frac{1}{\nu_{n+1}} \leq \frac{1}{\nu_n} \leq \cdots \leq \frac{1}{\nu_2} \leq \frac{1}{\nu_1} \leq 1.
$$

This implies

$$
1x^* - x_{n+1}^n \leq \nu_{n+1} + 0
$$

as $n \to \infty$. Thus we obtain $\tilde{x}^n = \lim x_n = x^*$. Next, we shall show that there is no solution in $C \setminus \tilde{S}$, provided that the set $C \setminus \tilde{S}$ is not empty. To show this, let $\tilde{x}^n$ be a solution in $C \setminus \tilde{S}$. Then we have

$$
\nu_{n+1} = \rho < |x^* - x_1|^2 \leq \frac{1}{2} K|x^* - x_0|^2 + \frac{1}{2} (x(n - \rho))^2 = \frac{1}{2} K(1 - e^\omega)^2 n^2
$$

so that

$$
1 \leq K (\cosh e - 1) n = K(a - 2) n < 1 - 2h,
$$

which is a contradiction. This proves the uniqueness of the solution in $C$. To prove (iii), we first observe that
\[ x^* - x_{n+1} = -F'(x_n) \int_0^1 \left[ F'(x_n + t(x^* - x_n)) - F'(x_n) \right] (x^* - x_n) \, dt \]

\[ x_n + t(x^* - x_n), \ x_n \in \overline{S}, \ n \geq 1, \]

\[ F'(x_n)^{-1} = \left[ I + F'(x_n)^{-1}(F'(x_n) - F'(x_n)) + F'(x_n)^{-1}(F'(x_n) - F'(x_n)) \right]^{-1} F'(x_n)^{-1} \]

and

\[ x_n, \ x \in \overline{S}, \ n \geq 1, \ x_1, x_0 \in L. \]

Hence we have from Lemma 2.3 (i)

\[ IF'(x_n)^{-1}F'(x_0)I \leq B_n \equiv \frac{1}{1 - K|lx_n - x_0| + d_0} \quad (n \geq 1) \quad (3.7) \]

\[ \frac{1}{1 - Kc_n} = B_n \]

and

\[ |xx^* - x_{n+1}| \leq \frac{1}{2} K_n |lx^* - x_n|^2 \leq \frac{1}{2} KB_n |lx^* - x_n|^2 \leq \frac{1}{2} KB_n |lx^* - x_n|^2. \]

Therefore Lemma 2.6 can be applied to obtain the bounds (3.1) - (3.3) for \( n \geq 1 \). Observe that if \( n = 0 \), they reduce to the Kantorovich bound \( t^* = 2n/(1 + \sqrt{1 - 2h}) \). The other part of (iii) follows from Lemmas 2.2 - 2.5. This proves (iii). Finally we shall prove (iv). Let \( F \) be Fréchet differentiable in an open convex set \( D_0 \) such that

\[ D_0 \supset D_0 \supset C \quad \text{and} \quad F'(x) \text{ satisfy the Lipschitz condition (2.10) in } D_0. \]

Then (3.7) may be replaced by the sharper estimate

\[ IF'(x_n)^{-1}F'(x_0)I \leq \frac{1}{1 - K|lx_n - x_0|}, \ n \geq 0. \]

Therefore, Lemma 2.6 can again be applied to replace (3.2) by (3.6). To prove the uniqueness of solution in \( \overline{S} \), let \( \overline{x}^* \) be a solution in \( \overline{S} \). Then there exists a nonnegative constant \( r \) such that \( r < 1 \) if \( 2h < 1 \), \( r \leq 1 \) if \( 2h = 1 \) and

\[ |\overline{x}^* - x_0|^2 \leq r^2 |t^* - t_n|, \ n \geq 0. \quad (3.8) \]
In fact, we have under our assumptions
\[ |h_n^* - x_{n+1}^*| \leq \frac{1}{2} \|\mathbb{K}_n\| |h_n^* - x_n^*| \]
\[ \quad \leq \frac{1}{2} \|\mathbb{K}_n\| r^2 |(t^{**} - t_n)|^2 \]
\[ = r^{n+1} (t^{**} - t_n - n_n) \]
\[ = r^{n+1} (t^{**} - t_{n+1}) \]
where we have used the induction hypothesis and Lemma 2.2. This proves (3.8), from which we obtain
\[ |h_n^* - x_n^*| \to 0 \]
as \( n \to \infty \), since \( r^n \to 0 \) if \( 2h < 1 \) and \( t^{**} - t_n = t^* - t_n \to 0 \) if \( 2h = 1 \). Hence we have \( x^* = \lim x_n \to x^* \), which implies the uniqueness of solution in \( S \). Q.E.D.

Assumptions similar to those of Theorem 2.2 or Lemma 2.1 were also adopted by Lancaster [8], and later by Schmidt [22], [23] for the generalized secant method which includes the Newton method as a special case. Lancaster's assumptions correspond to the case \( e = 0 \) in Lemma 2.1, while Schmidt's correspond to the case where \( e \) is chosen so that \( ah = 1 \), in which case we have \( o = e^* n = t^* - n \). In the following we shall improve Lancaster's result. (In the case of the Newton method, Schmidt's upper bound reduces to the Kantorovich bound \( t^* - t_n \). Also see [28].)

Corollary 3.1.1. Let \( F: D \subseteq X \times Y \) and \( D^0 \) be the interior of \( D \). Assume that for some \( x_0 \in D^0 \), \( F'(x_0) \) and \( F'(x_0)^{-1} \) exist, \( F'(x_0) \neq 0 \) and \( F \) is Fréchet differentiable on \( S_0 = \hat{S}(x_0, |x_1 - x_0|) \). Furthermore, put
\[ L_0 = \sup_{x, y \neq x, y \in S_0} \frac{|F'(x_0)^{-1}(F'(x) - F'(y))|}{|x - y|} \]

-13-
If \( 2L_0|x_1 - x_0| \leq 1 \), then the Newton process (2.2) generates a sequence \( \{x_n\} \subset \mathbb{R}_0^* \)
which converges to the unique solution \( x^* \) of (2.1) in \( \mathbb{R}_0^* \). If we put

\[
L_n = \sup_{x,y \in \mathbb{R}_0^*} \frac{1}{1 - \frac{2}{L_n} \left| x - y \right|}, \quad n \geq 1
\]

and

\[
d_n = \left| x_{n+1} - x_n \right|, \quad n \geq 0,
\]

then the following error estimates hold:

\[

\left| x^* - x_n \right| \leq \frac{2d_n}{1 + \sqrt{1 - 2L_n d_n}} \quad (n \geq 0)
\]

\[

\leq \frac{L_n d_n^2}{1 + \sqrt{1 - (L_n d_n)^2}} \quad (n \geq 1)
\]

\[

\leq \frac{L_n d_n^2}{1 - L_n d_n^2} \quad (n > 1)
\]

\[

= \frac{L_n d_n^2}{1 - L_n d_n^2} \quad (n > 1)
\]

**Proof.** Put \( \phi = 0 \) in Theorem 3.1. Then Corollary 3.1.1 follows from Theorem 3.1 by

noting that

\[
d_n \leq \frac{1}{2} L_n d_n^2 \quad \text{and} \quad L_n \leq \frac{1}{1 - L_n d_n^2}, \quad n \geq 1.
\]

Q.E.D.

The bound (3.10) is due to Lancaster [8] and (3.9) is what Potra cited in his recent paper [17] as Kornstaedt's bound [7].
Corollary 3.1.2. Under the notation and assumptions of Corollary 3.1.1, we have

\[ I^* - x^* \leq \delta_n = \frac{2d_n}{1 + \sqrt{1 - 2L_0(n - L_0\delta_n)^{-1}}(n \geq 0)} \]

\[ 2F'(x_0)^{-1}F(x_n) \leq \frac{1}{1 - L_0\delta_n^2 + \sqrt{(1 - L_0\delta_n)^2 - 2L_0F'(x_0)^{-1}F(x_n)}} (n \geq 0) \]  \hspace{1cm} (3.11)

\[ \leq \frac{L_0\delta_n^2}{1 - L_0\delta_n^2 + \sqrt{(1 - L_0\delta_n)^2 - (L_0\delta_n^2)^2}} (n \geq 1) \]  \hspace{1cm} (3.12)

where \( \delta_n = I^* - x^* \).

Proof. It is easy to see that

\[ L_n = \frac{L_0}{1 - L_0\delta_n} \]

and

\[ d_n \leq \frac{1}{1 - L_0\delta_n} F'(x_0)^{-1}F(x_n) \]

Furthermore, we have

\[ F'(x_0)^{-1}F(x_n) = F'(x_0)^{-1}[F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})] \]

\[ = F'(x_0)^{-1} \int_0^1 [F'(x_{n-1} + t(x_n - x_{n-1})) - F'(x_{n-1})](x_n - x_{n-1})dt \]

so that

\[ IF'(x_0)^{-1}F(x_n) \leq \frac{1}{2} L_0\delta_n^2, \quad n \geq 1. \]

Therefore, Corollary 3.1.2 follows from Corollary 3.1.1.

Q.E.D.
Remark 3.1. The bounds (3.11) and (3.12) were recently obtained by Potra [17] under the assumptions of Theorem 2.1 as \( B_5(\Delta_n, [F'(x_0)]^{-1}F(x_n)) \) and \( B_5(\Delta_n, d_{n-1}) \) respectively in his notation.

We can further improve the bound \( \delta_n \) obtained in Corollary 3.1.1.

Corollary 3.1.3. Under the assumptions of Corollary 3.1.1, put \( M_0 = L_0 \) and for \( n \geq 1 \)

\[
\bar{S}_n = \bar{S}(x_n, 24F'(x_n)F(x_n)) ,
\]

\[
M_n = \sup_{x,y \in S_n} \frac{\|F'(x) - F'(y)\|}{\|x - y\|} .
\]

Then we have

\[
x^* \in \bar{S}_n \subseteq \bar{S}_{n-1} \subseteq \ldots \subseteq \bar{S}_0.
\]

and

\[
1 \leq n + \frac{2d_n}{1 + \sqrt{1 - 2d_n}} \leq \delta_n, n \geq 0 .
\]

Proof. This immediately follows from Lemma 2.6 by noting that \( 2d_{n+1} \leq d_n \) and

\[
1 \leq n + \frac{2d_n}{1 + \sqrt{1 - 2d_n}} \leq \delta_n, n \geq 0 .
\] Q.E.D.

Remark 3.2. As was remarked in [17], the cost of obtaining \( K_n, L_n \) or \( M_n \) might be very high. Therefore, in practical computation, it would be better to make use of one of (3.2), (3.3), (3.6) and (3.12). However, Theorem 3.1 and its Corollaries assert that the error bounds which have been obtained by many authors with the use of different techniques can be derived from the majorant theory of Kantorovich and Lemma 2.6, in a unified manner.

Theorem 3.2. Under the assumptions of Lemma 2.1, we have

\[
dx_n = \frac{V_{n+1}}{(V_n)^2} d_n \leq \frac{V_{n+1}}{V_n} d_{n-1}
\]

\[
= \frac{1}{2 \cosh 2^{n-1} \delta_{n-1}}, n \geq 1 .
\] (3.13)

The equality holds in (3.13) if and only if \( \delta_n = 1 \).
Proof. It follows from (2.6) and Lemma 2.3 (ii) that

\[
d_n \leq \frac{1}{2} \text{Kn}^2 n_{n-1}^2 \leq \frac{Vt_n+1}{Vt_n}^2 n_{n-1}^2 \leq \frac{Vt_n+1}{Vt_n} d_{n-1}.
\]

Choose \( \psi \geq 0 \) such that \((1 + \cosh \psi)h = 1\). Then, from Lemmas 2.4 and 2.5, we have

\[
Vt_{n+1} = e^{2n\psi} (t^n - t_{n+1}) = \begin{cases} 
\frac{\sinh \psi}{\sinh 2n\psi} (\psi > 0) \\
\frac{1}{2} (\psi = 0)
\end{cases}
\]

Hence

\[
\frac{Vt_{n+1}}{Vt_n} = \begin{cases} 
\frac{\sinh 2n-1\psi}{\sinh 2n\psi} (\psi > 0) \\
\frac{1}{2} (\psi = 0)
\end{cases}
\]

\[
= \frac{1}{2 \cosh 2n-1\psi} \leq \frac{1}{2 \cosh 2n-1\psi}
\]

for every \( \psi \in [0, \psi^*] \). The equality holds in (3.15) if and only if \( \psi = \psi^* \). This, together with (3.14), proves Theorem 3.2.

Q.E.D.

Remark 3.3. The bound (3.13) is of the form found in Ostrowski [16; Theorem 38.2] in.

4. Observation

Under the assumptions of Theorem 2.2, Ostrowski proved that

\[
\frac{1}{2^{2m}} \leq \frac{1}{\text{IF}(x_{n+m})} \leq \left\{ \begin{array}{ll}
\frac{(\sinh 2n-1\psi)^2}{\sinh 2n+1\psi} & (\psi > 0) \\
\frac{1}{2^{2m}} & (\psi = 0)
\end{array} \right.
\]

for every \( n \geq 0, m \geq 0 \).
provided that \( F(x_n) \neq 0 \). By our approach, we can easily derive his estimates. Let

\[
\hat{x} = \sigma^{-1}, \quad \hat{B} = IF'(x_0)^{-1}, \quad \tilde{n} = \hat{B}IF(x_0)I, \quad \tilde{F}(t) = \frac{1}{2} \hat{K}at^2 - t + \tilde{n}.
\]

Furthermore, we define the sequences \( \{\tilde{B}_n\} \) and \( \{\tilde{n}_n\} \) by

\[
\tilde{B}_0 = \hat{B}, \quad \tilde{n}_0 = \tilde{n}, \quad \tilde{h}_0 = \hat{K}B_0\tilde{n}_0,
\]

\[
\tilde{B}_n = \frac{\tilde{B}_{n-1}}{1 - \tilde{h}_{n-1}}, \quad \tilde{n}_n = \frac{\tilde{B}_{n-1}\tilde{n}_{n-1}}{2(1 - \tilde{h}_{n-1})}, \quad \tilde{h}_n = \hat{K}B_n\tilde{n}_n, \quad n \geq 1.
\]

Then it is easy to see that

\[
IF'(x_n)^{-1} \leq \tilde{B}_n, \quad \|x_{n+1} - x_n\| \leq \tilde{n}_n
\]

and

\[
F(x_{n+1}) = F(x_n) + \int_0^1 F'(x_n + t(x_{n+1} - x_n))(x_{n+1} - x_n)dt
\]

\[
= \int_0^1 [F'(x_n) - F'(x_n + t(x_{n+1} - x_n))]F'(x_n)^{-1}F(x_n)dt.
\]

Hence we have

\[
IF(x_{n+1}) \leq \frac{1}{2} \hat{K}(x_{n+1} - x_n)IF'(x_n)^{-1}IF(x_n)
\]

\[
\leq \frac{1}{2} \hat{K} \tilde{B}_n \tilde{n}_n IF(x_n)I = \left(\frac{\tilde{n}_n}{\tilde{h}_{n-1}}\right)^2 IF(x_n)I \quad (4.2)
\]

Define the sequence \( \{\tilde{c}_n\} \) by

\[
\tilde{c}_0 = 0, \quad \tilde{c}_{n+1} = \tilde{c}_n - \tilde{F}(t_n)/\tilde{F}'(t_n), \quad n \geq 0.
\]

Then, by Lemmas 2.2, 2.4 and 2.5, we have
\[
\begin{align*}
\bar{v}_n &= \bar{v}_{n+1} - \bar{v}_n = e^{2n} \left( \ell^* - \ell_{n+1} \right) \\
&=\begin{cases} 
\frac{\sinh \varphi}{\sinh 2^n \varphi} & (\varphi > 0) \\
\frac{2^n \varphi}{\sinh \varphi} & (\varphi = 0) \\
2^n & (\varphi = 0).
\end{cases} \tag{4.3}
\end{align*}
\]

since \( a_{n_0} = 1 \), where \( \ell^* = (1 - \sqrt{1 - 2R_0})/K \). Therefore, we have from (4.2) and (4.3)

\[
\frac{\|F(x_{n+1})\|^2}{\|F(x_n)\|^2} \leq \begin{cases} 
\frac{\sinh 2^{n-1} \varphi}{\sinh 2^n \varphi} & (\varphi > 0) \\
\frac{1}{2} & (\varphi = 0).
\end{cases}
\]

This leads to the estimate (4.1). Therefore, together with (3.13) which holds in our case, we proved the main part of his theorem [16; Theorem 38.2]. The remaining part also follows from our approach.

Finally we remark that the chart for the lower bounds given in [28] is still true under the assumptions of Lemma 2.1 with a slight modification:

\[
\|x^* - x_n\| \geq \frac{2d_n}{1 + \sqrt{1 + 2Kd_n}} \quad (n \geq 0)
\]

\[
\begin{cases} 
\frac{2d_0}{1 + \sqrt{1 + 2h}} & (n = 0) \\
\frac{2d_n}{1 + \sqrt{1 + 2K(1 - K(x_n - x_1, + d_0)^{-1}d_n^{-1}} & (n \geq 1)
\end{cases}
\tag{4.4}
\]

\[
\begin{cases} 
\frac{2d_n}{1 + \sqrt{1 + 2K(1 - Kx_n^{-1})^{-1}d_n^{-1}} & (n \geq 0) \quad (\text{Miel [11], Schmidt [23]})
\end{cases}
\]
If the assumptions of Theorem 3.1 (iv) are satisfied, then (4.4) may be replaced by the sharper lower bound

\[ \frac{2d_n}{1 + \sqrt{1 + 2h_n} - (n \geq 0)} \text{ (Potra-Ptak [18])} \]

\[ \frac{2d_n}{1 + \sqrt{1 + 2h_n} - (n \geq 0)} \text{ (Gragg-Tapia [3]).} \]

We also have

\[ \frac{2d_n}{1 + \sqrt{1 + 2h_n} - (n \geq 0)} \text{ (Yamamoto [28]).} \]

with the notation and assumptions of Corollary 3.1.3.

Acknowledgements. The author wishes to thank Professor L. B. Rail of the Mathematics Research Center, University of Wisconsin-Madison and Professor R. K. Guy of the University of Calgary for bringing the theorems of Ostrowski and Lancaster respectively to his attention.
References


-21-


On the basis of the results obtained in a series of papers [25] - [28], a convergence theorem for Newton's method in Banach spaces is given, which improves the theorems of Kantorovich [4], Lancaster [8] and Ostrowski [16]. The error bounds obtained improve the recent results of Potra [17].