Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Issue 1, July 2003, pp. 1–9.
Submitted: July 19, 2002. Published: July 1 2003.

ON THE CONVERGENCE OF NEWTON'S METHOD FOR ANALYTIC OPERATORS

Ioannis K. Argyros

Abstract. We approximate a solution of an operator equation using Newton's method. The new local and semilocal convergence theorems are based on the analyticity of the operator involved.

A.M.S. (MOS) Subject Classification Codes. 65B05, 65N30, 47H17, 49M15.

Key Words and Phrases. Newton's method, Banach space, analytic operator, majorizing sequence, Banach lemma on invertible operators.

1. Introduction

In this study we are concerned with the problem of approximating a solution x^* of equation

$$F(x) = 0, \tag{1}$$

where F is an analytic operator defined on an open convex subset of a Banach space X with values in a Banach space Y.

Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (x_0 \in D) \quad (n \geq 0) \tag{2}$$

has been used extensively to generate a sequence approximating x^*. Until 1986 the domain condition in the Kantorovich theorem was used to show convergence of Newton's method to x^*. In 1986 S. Smale [6] and then others [5], [7] replaced the domain condition with the analytic property of F. Here moving in this direction we provide new semilocal and local convergence theorems. It turns out that in the semilocal case under our weaker conditions we provide a more precise information on the location of the solution x^* as well as on the distances $\|x_{n+1}-x_n\|$, $\|x_n-x^*\|$ $(n \geq 0)$ than before [7]. In the local case we can provide a larger radius of convergence than [7]. This observation is very important in computational mathematics and finds applications in projection methods and in the selection of efficient mesh independence strategies [1], [2], [8].

Department of Mathematics, Cameron University, Lawton, OK 73505, U.S.A.
E-mail Address: ioannisa@cameron.edu

2. Semilocal Analysis of Newton's Method

We need the following result on majorizing sequences for Newton's method.

Theorem 1. *Let η, γ_i $(i \geq 2)$ be non-negative parameters. Suppose there exists $\delta \in [0,1]$ such that*

$$h_0 = 2\int_0^1 (1-\theta)\sum_{j=0}^{\infty}(j+1)(j+2)\gamma_{j+2}(\theta\eta)^j\,d\theta + \delta\sum_{j=1}^{\infty}(j+1)\eta^j \leq \delta. \tag{3}$$

Then, iteration $\{t_n\}$ $(n \geq 0)$ given by

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+2} = t_{n+1} + \frac{\int_0^1 (1-\theta)\sum\limits_{j=0}^{\infty}(j+1)(j+2)\gamma_{j+2}(t_n+\theta(t_{n+1}-t_n))^j(t_{n+1}-t_n)d\theta}{1-\sum\limits_{j=1}^{\infty}(j+1)\gamma_{j+1}t_{n+1}^j} \tag{4}$$

is non-decreasing, bounded above by

$$t^{**} = \frac{2\eta}{2-\delta}\,, \tag{5}$$

and converges to some t^* such that

$$0 \leq t^* \leq t^{**}. \tag{6}$$

Moreover, the following error bounds hold for all $n \geq 0$:

$$0 \leq t_{n+2}-t_{n+1} \leq \frac{\delta}{2}(t_{n+1}-t_n) \leq \left(\frac{\delta}{2}\right)^{n+1}\eta. \tag{7}$$

Proof. Set

$$h_n = 2\int_0^1 (1-\theta)\sum_{j=0}^{\infty}(j+1)(j+2)\gamma_{j+2}(t_n+\theta(t_{n+1}-t_n))^j(t_{n+1}-t_n)d\theta + \delta\sum_{j=1}^{\infty}(j+1)\gamma_{j+1}t_{n+1}^j \quad (n \geq 0). \tag{8}$$

We must show using induction on the integer $k \geq 0$

$$h_k \leq \delta, \;\; t_{k+1}-t_k \geq 0, \;\; 1-\sum_{j=1}^{\infty}(j+1)\gamma_{j+1}t_{k+1}^j > 0. \tag{9}$$

Note that $h_0 \le \delta$ (by (3)), $t_1 - t_0 \ge 0$, $1 - \sum_{j=1}^{\infty} (j+1)\gamma_{j+1} t_1^j > 0$ (by (3)). Hence (9) hold for $k = 0$. But then (4) gives

$$0 \le t_2 - t_1 \le \frac{\delta}{2}(t_1 - t_0). \tag{10}$$

Assume that (9) holds for all $k \le n+1$. It can easily be seen that

$$h_k \le h_0 \quad (k \ge 0), \tag{11}$$

which shows the first estimate in (9) for all $k \ge 0$. The second follows from (4).

We must prove:

$$t_k \le t^{**} \quad (k \ge 0). \tag{12}$$

We get

$$t_0 = 0 \le t^{**}, \quad t_1 = \eta \le t^{**} \quad \text{and} \quad t_2 \le \eta + \frac{\delta}{2}\eta = \frac{2+\delta}{2}\eta \le t^{**}.$$

Assume (12) holds for all $k \le n+1$. It follows from (7) that

$$\begin{aligned} t_{k+2} &\le t_{k+1} + \frac{\delta}{2}(t_{k+1} - t_k) \\ &\le t_k + \frac{\delta}{2}(t_k - t_{k-1}) + \frac{\delta}{2}(t_{k+1} - t_k) \\ &\le \cdots \le t_1 + \frac{\delta}{2}(t_1 - t_0) + \cdots + \frac{\delta}{2}(t_{k-1} - t_k) + \frac{\delta}{2}(t_{k+1} - t_k) \\ &\le n + \frac{\delta}{2}\eta + \left(\frac{\delta}{2}\right)^2 \eta + \cdots + \left(\frac{\delta}{2}\right)^{k+1} \eta \\ &= \frac{1 - \left(\frac{\delta}{2}\right)^{k+2}}{1 - \frac{\delta}{2}}\eta \le \frac{2}{2-\delta}\eta = t^{**}. \end{aligned} \tag{13}$$

Hence, iteration $\{t_n\}$ $(n \ge 0)$ is bounded above by t^{**}. The third estimate in (9) holds by (13) and (3). Moreover it follows from (4) that $\{t_n\}$ $(n \ge 0)$ is also non-decreasing and as such it converges to some t^* satisfying (6).

That completes the proof of Theorem 1.

We show the semilocal convergence theorem for Newton's method and analytic operators.

Theorem 2. *Let $F\colon D \subseteq X \to Y$ be an analytic operator. Assume there exist a point $x_0 \in D$ and parameters $n \ge 0$, $\delta \in [0,1]$ such that*

$$F'(x_0)^{-1} \in L(Y, X) \tag{14}$$

$$\|F'(x_0)^{-1} F(x_0)\| \le \eta, \tag{15}$$

$$h_0 \le \delta, \tag{16}$$

where, h_0, t^* are given by (3), (6) respectively

$$\gamma_j = \left\| \frac{F'(x_0)^{-1}F^{(j)}(x_0)}{j!} \right\| \quad (j \geq 2) \tag{17}$$

and

$$\overline{U}(x_0, t^*) = \{x \in X \mid \|x - x_0\| \leq t^*\} \subseteq D. \tag{18}$$

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ generated by (2) is well defined, remains in $\overline{U}(x_0, t^*)$ for all $n \geq 0$ and converges to a solution $x^* \in \overline{U}(x_0, t^*)$ of equation $F(x) = 0$.

Moreover the following error bounds hold for all $n \geq 0$

$$\|x_{n+2} - x_{n+1}\| \leq t_{n+2} - t_{n+1} \tag{19}$$

and

$$\|x_n - x^*\| \leq t^* - t_n \tag{20}$$

where, iteration $\{t_n\}$ $(n \geq 0)$ is given by (4).

Proof. Let us prove (19) and

$$\overline{U}(x_{n+1}, t^* - t_{n+1}) \leq \overline{U}(x_n, t^* - t_n) \tag{21}$$

hold for all $n \geq 0$. If $z \in \overline{U}(x_1, t^* - t_1)$,

$$\|z - x_0\| \leq \|z - x_1\| + \|x_1 - x_0\| \leq t^* - t_1 + t_1 = t^* - t_0$$

implies $z \in \overline{U}(x_0, t^* - t_0)$. Since also

$$\|x_1 - x_0\| = \|F'(x_0)^{-1}F(x_0)\| \leq \eta = t_1$$

(19) and (21) hold for $n = 0$. Given they hold for $n = 0, 1, \ldots, k+1$, then

$$\|x_{k+1} - x_0\| \leq \sum_{i=1}^{k+1} \|x_i - x_{i-1}\| \leq \sum_{i=1}^{k+1}(t_i - t_{i-1}) = t_{k+1} - t_0 = t_{k+1}$$

and

$$\|x_k + \theta(x_{k+1} - x_k) - x_0\| \leq t_k + \theta(t_{k+1} - t_k) < t^*, \quad \theta \in [0, 1].$$

Suppose $\|x - x_0\| \leq t < t^*$

(22)

$$\begin{aligned} \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| &\leq \sum_{j=1}^{\infty} \frac{\|F'(x_0)^{-1}F^{(j+1)}(x_0)\|}{j!} \|x - x_0\|^j \\ &\leq \sum_{j=1}^{\infty} (j+1)\gamma_{j+1}t^j < 1 \quad \text{(by (3))}. \end{aligned}$$

It follows by the Banach Lemma on invertible operators [4] $F'(x)^{-1}$ exists and

$$\|F'(x)^{-1}F'(x_0)\| \leq \frac{1}{1-\sum\limits_{j=1}^{\infty}(j+1)\gamma_{j+1}t^j}. \tag{23}$$

Moreover from (2), (17) and (19) we get

$$\begin{aligned}
&\|F'(x_0)^{-1}F(x_{k+1})\| \\
&\quad = \|F'(x_0)^{-1}\{F(x_{k+1})-F(x_k)-F'(x_k)(x_{k+1}-x_k)\}\| \\
&\quad = \left\|\int_0^1(1-\theta)F'(x_0)^{-1}F''[x_k+\theta(x_{k+1}-x_k)](x_{k+1}-x_k)^2d\theta\right\| \\
&\quad = \left\|\int_0^1(1-\theta)\sum_{j=0}^{\infty}\frac{F'(x_0)^{-1}F^{(j+2)}(x_0)}{j!}\right. \\
&\qquad \left.\cdot\ [x_k-x_0+\theta(x_{k+1}-x_k)]^j(x_{k+1}-x_k)^2d\theta\right\| \\
&\quad \leq \int_0^1(1-\theta)\sum_{j=0}^{\infty}(j+1)(j+2)\gamma_{j+2}(t_k+\theta(t_{k+1}-t_k))^j(t_{k+1}-t_k)^2d\theta.
\end{aligned} \tag{24}$$

Combining (23) and (24), we get

$$\begin{aligned}
\|x_{n+2}-x_{n+1}\| &= \|[F'(x_{n+1})^{-1}F'(x_0)][F'(x_0)^{-1}F(x_{n+1})]\| \\
&\leq \|F'(x_{n+1})^{-1}F'(x_0)\|\,\|F'(x_0)^{-1}F(x_{n+1})\| \\
&\leq t_{n+2}-t_{n+1}.
\end{aligned} \tag{25}$$

This completes the induction. Theorem 1 implies that $\{t_n\}$ $(n \geq 0)$ is a Cauchy sequence. From (19) and (21) $\{x_n\}$ $(n \geq 0)$ becomes a Cauchy sequence too, and as such it converges to some $x^* \in \overline{U}(x_0,t^*)$ (since $U(x_0,t^*)$ is a closed set) so that (20) holds.

That completes the proof of Theorem 2.

3. Local Convergence Analysis of Newton's Method

We finally provide the local convergence theorem for Newton's method and analytic operators.

Theorem 3. *Let $F\colon D \subseteq X \to Y$ be an analytic operator. Assume there exists a simple zero $x^* \in D$, a parameter $\overline{\delta} \in [0,1]$, and a minimum non-negative parameter r^* satisfying*

$$\overline{h}_0 \leq \overline{\delta}, \tag{26}$$

and

$$\overline{U}(x^*,r^*) \subseteq D, \tag{27}$$

where,

$$\overline{h}_0 = \overline{h}_0(r) = \int_0^1 \sum_{j=0}^{\infty} (j+1)(j+2)\overline{\gamma}_{j+2}[(1-\theta)r]^{j+1} d\theta + \overline{\delta} \sum_{j=1}^{\infty} (j+1)\overline{\gamma}_{j+1} r, \tag{28}$$

$$\overline{\gamma}_j = \left\| \frac{F'(x^*)^{-1}F^{(j)}(x^*)}{j!} \right\| \quad (j \geq 2). \tag{29}$$

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ *generated by (2) is well defined, remains in* $U(x^*, r^*)$ *for all* $n \geq 0$ *and converges to* x^* *provided that* $x_0 \in U(x^*, r^*)$*. Moreover the following error bounds hold for all* $n \geq 0$

$$\|x_{n+1} - x^*\| \leq \frac{\int_0^1 \sum_{j=0}^{\infty} (j+1)(j+2)\overline{\gamma}_{j+2}(1-\theta)^{j+1}\|x_n - x^*\|^{j+2} d\theta}{1 - \sum_{j=1}^{\infty} (j+1)\overline{\gamma}_{j+1}\|x_n - x^*\|^j}. \tag{30}$$

Proof. Let $x \in U(x^*, r^*)$. By the analyticity of F we get

$$\begin{aligned} &\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \\ &= \left\| \sum_{j=1}^{\infty} \frac{F'(x^*)^{-1}F^{(j+1)}(x^*)}{j!}(x - x^*)^j \right\| \\ &\leq \sum_{j=1}^{\infty} (j+1)\overline{\gamma}_{j+1}\|x - x^*\|^j < 1 \quad \text{(by (26))}. \end{aligned} \tag{31}$$

It follows by the Banach Lemma on invertible operators

$$\|F'(x)^{-1}F'(x^*)\| \leq \frac{1}{1 - \sum_{j=1}^{\infty} (j+1)\overline{\gamma}_{j+1}\|x - x^*\|^j}. \tag{32}$$

Using (2) we obtain in turn the approximation

$$\begin{aligned} x_{n+1} - x^* &= x_n - x^* - F'(x_n)^{-1}F(x_n) \\ &= [F'(x_n)^{-1}F'(x^*)]F'(x^*)^{-1} \\ &\quad \cdot \left\{ \int_0^1 [F'(x_n + \theta(x^* - x_n)) - F'(x_n)](x^* - x_n) d\theta \right\} \\ &= [F'(x_n)^{-1}F'(x^*)]F'(x^*)^{-1} \int_0^1 (1-\theta)F''[x_n + \theta(x^* - x_n)] \\ &\quad \cdot (x^* - x_n)^2 d\theta. \end{aligned} \tag{33}$$

Moreover, by (29) we get

(34)
$$\begin{aligned}
&\left\| F'(x^*)^{-1} \int_0^1 (1-\theta) F''[x_n + \theta(x^* - x_n)](x^* - x_n)^2 d\theta \right\| \\
&\quad = \left\| \int_0^1 (1-\theta) \frac{F'(x^*)^{-1} F^{(j+2)}(x^*)}{j!} [x_n - x^* + \theta(x^* - x_n)]^j (x^* - x_n)^2 d\theta \right\| \\
&\quad \leq \int_0^1 \sum_{j=0}^{\infty} (j+1)(j+2)\overline{\gamma}_{j+2} (1-\theta)^{j+1} \|x_n - x^*\|^{j+2} d\theta.
\end{aligned}$$

Estimate (30) follows from (2), (32), (33), and (34). Furthermore, by the choice of r^*

$$\alpha \equiv \frac{\int_0^1 \sum_{j=0}^{\infty} (j+1)(j+2)\overline{\gamma}_{j+2} [(1-\theta)\|x_0 - x^*\|]^{j+1} d\theta}{1 - \sum_{j=1}^{\infty} (j+1)\overline{\gamma}_{j+1} \|x_0 - x^*\|} \in [0,1) \tag{35}$$

and

$$\|x_{n+1} - x^*\| \leq \alpha \|x_n - x^*\| \leq \alpha^{n+1} \|x_0 - x^*\|, \tag{36}$$

which shows $x_n \in U(x^*, r^*)$ and $\lim_{n\to\infty} x_n = x^*$.

That completes the proof of Theorem 3.

The results in [5] and [6] used the assumption that the sequence $\{\gamma_j\}$ $(j \geq 2)$ is bounded by

$$\gamma = \sup_{j\geq 2} \gamma_j^{\frac{1}{j-1}}. \tag{37}$$

This assumption is not true for many interesting types of operators [7]. In [7] they showed a semilocal result where $\{\gamma_j\}$ $(j \geq 2)$ is not considered to be bounded by γ, whereas in the local case they used the bound

$$\gamma^* = \sup_{j\geq 2} \left\| \frac{F'(x^*)^{-1} F^{(j)}(x^*)}{j!} \right\|^{\frac{1}{j-1}}. \tag{38}$$

The convergence radius is given by

$$r_1^* = \frac{1}{2\gamma^*}(3 - 2\sqrt{2}). \tag{39}$$

Note that in our local result (Theorem 3) we do not use $\{\overline{\gamma}_j\}$ $(j \geq 2)$ is bounded by γ^*.

Define functions φ, ψ by

$$\varphi(t) = \eta - \psi(t), \tag{40}$$

$$\psi(t) = t - \sum_{j=2}^{\infty} \gamma_j t^j \tag{41}$$

and iteration $\{v_n\}$ $(n \geq 0)$ by

$$v_{n+1} = v_n - \frac{\varphi(v_n)}{\varphi'(v_n)} \quad (n \geq 0). \tag{42}$$

If $\varphi(t)$ has a unique zero v^* in $[0, +\infty)$, i.e., if

$$\max_{t>0} \psi(t) \quad \text{exists.} \tag{43}$$

they showed convergence of (2) to x^* and

$$\|x_{n+1} - x_n\| \leq v_{n+1} - v_n, \tag{44}$$

$$\|x_{n+1} - x_n\| \leq v^* - v_n. \tag{45}$$

It can easily be seen that (3) implies (43) but not vice versa. Moreover using induction on n we can easily show that for all $n \geq 0$

$$t_n \leq v_n \tag{46}$$

$$t_{n+1} - t_n \leq v_{n+1} - v_n \tag{47}$$

$$t^* - t_n \leq v^* - v_n \tag{48}$$

and

$$t^* \leq v^*, \tag{49}$$

which justify the claims made at the introduction.

We complete this study with an example to show that our convergence radius r^* can be larger than r_1^*.

Example 1. Let $X = Y = \mathbf{R}$, $D = U(0,1)$, and define function F on D by

$$F(x) = e^x - 1. \tag{50}$$

By (38) we obtain $\gamma^* = .5$. Hence (39) gives

$$r_1^* = 3 - 2\sqrt{2} = .171572875. \tag{41}$$

Using (29) we obtain $\overline{\gamma}_j = \frac{1}{j!}$ $(j \geq 2)$. Moreover for $\delta = 1$ we obtain from (28)

$$\begin{aligned} h_0(r) &= \int_0^1 \sum_{j=0}^{\infty} \frac{(j+1)(j+2)}{(j+2)!}(1-\theta)^{j+1} r^{j+1} d\theta \\ &\quad + \sum_{j=1}^{\infty} (j+1)\frac{1}{(j+1)!} r^j \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} r^{j+1} \int_0^1 (1-\theta)^{j+1} d\theta + \sum_{j=1}^{\infty} \frac{r^j}{j!} \\ &= \sum_{j=0}^{\infty} \frac{r^{j+1}}{(j+1)!} \frac{j+1}{j+2} + \sum_{j=1}^{\infty} \frac{r^j}{j!} \\ &\leq \sum_{j=0}^{\infty} \frac{r^{j+1}}{(j+1)!} + \sum_{j=1}^{\infty} \frac{r^j}{j!} \\ &= 2(e^r - 1). \end{aligned} \tag{52}$$

Solving the inequality

$$2(e^r - 1) \leq 1, \tag{53}$$

we obtain

$$r^* = .405465108. \tag{54}$$

That is,

$$r_1^* < r^*. \tag{55}$$

Hence, our local results provide a wider choice for x_0 than the corresponding ones in [7]. This observation is very important in computational mathematics and finds applications in projection methods and in step-length selection in predictor–corrector continuation procedures [1], [2], [8].

REFERENCES

[1] Argyros, I.K., Local convergence theorems for Newton's method using outer or generalized inverses and m-Fréchet differentiable operators, *Mathematical Sciences Research Hot-Line* **4**, 8 (2000), 47–56.

[2] Argyros, I.K. and Szidarovszky, F., *The Theory and Application of Iteration Methods*, CRC Press, Inc., Boca Raton, Florida, U.S.A., 1993.

[3] Gutiérrez, J., A new semilocal convergence theorem for Newton's method, *J. Comput. Appl. Math.* **79** (1997), 131–145.

[4] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[5] Rheinboldt, W.C., On a theorem of S. Smale about Newton's method for analytic mappings, *Appl. Math. Lett.* **1**, 1 (1988), 69–72.

[6] Smale, S., Newton's method estimates from data at one point. In *The Merging of Disciplines in Pure, Applied and Computational Mathematics*, Springer-Verlag, Berlin, New York, 1986, pp. 185–196.

[7] Wang, D. and Zhao, F., The theory of Smale's point estimation and its applications, *J. Comput. Appl. Math.* **60** (1995), 253–269.

[8] Ypma, T.J., Local convergence of inexact Newton methods, *SIAM J. Numer. Anal.* **21**, 3 (1984), 583–590.