# SUFFICIENT CONDITIONS FOR THE CONVERGENCE OF NEWTON'S METHOD IN COMPLEX BANACH SPACES[1]

MARVIN L. STEIN

1. **Introduction.** Let $T(y)$ be an operator defined on a Banach space $Y$ into another such space and let $\delta T(y; h)$ be its first variation. The Banach space analogue of Newton's method is given by the iteration formula

$$y_{i+1} = y_i - \delta T^{-1}(y_i; T(y_i)), \qquad i = 0, 1, 2, \cdots, \tag{1.1}$$

where $\delta T^{-1}(y; h)$ denotes the inverse with respect to $h$ of $\delta T(y; h)$. The present writer [**8**][2] has shown that this iteration scheme, henceforth called Newton's method, may be used to solve certain nonlinear systems of second order differential equations subject to two point boundary conditions. Various other applications have been made by Kantorovič [**3**; **4**; **5**] and Mysovskih [**6**; **7**]. In the papers of Kantorovič, conditions under which a sequence $\{y_i\}$ determined by Newton's method will converge to a solution of the equation

$$T(y) = 0 \tag{1.2}$$

are established for the case of a real Banach space. In this paper we complete the picture by presenting a convergence theorem for Newton's method valid for the case of a complex Banach space. This theorem has the virtue of imposing no conditions on the second variation of $T$ as is done by Kantorovič in the real case. This is advantageous in that the second variation plays no role in (1.1) and hence need never be computed.

2. **Convergence theorem.** *Let $Y$ and $Z$ be complex Banach spaces and let $T(y)$ be an operator defined on the sphere*

$$S: \ \|y\| < \rho, \qquad \rho > 0,$$

*into $Z$. Let the following conditions be fulfilled:*

(i) *$T(y)$ is G-differentiable (Hille* [**2**, p. 71]),

(ii) *there exist $y_0$ in $S$, a positive number $a$, and a sufficiently small positive constant $M$ such that*

$$\|T(y)\| \leqq M, \quad \text{if} \quad \|y - y_0\| \leqq a,$$

Presented to the Society, February 23, 1952; received by the editors December 14, 1951 and, in revised form, March 10, 1952.

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research.

[2] The numbers in brackets refer to the list of references at the end of this paper.

(iii) *$\delta T(y_0; h)$ is a one to one mapping of $Y$ onto $Z$. Then the sequence $\{y_i\}$ given by Newton's method is well defined and there exists $d>0$ such that $\{y_i\}$ converges quadratically to a unique element $\bar{y}$ of the $d$-neighborhood of $y_0$. Furthermore $\bar{y}$ is the unique solution of* (1.2) *in this neighborhood.*

It will facilitate the proof to state several intermediate results in the form of lemmas.

LEMMA 2.1. *There exists a positive constant $r<1$ such that for all $y$ in $S$ satisfying $\|y-y_0\|\leqq r$ and any $h$ in $Y$ the inequalities*

$$\|T(y)\| \leqq M, \tag{2.1}$$

$$\|\delta T(y; h)\| \leqq \frac{M}{r}\|h\| \tag{2.2}$$

*hold, while*

$$\|T(y+h) - T(y) - \delta T(y; h)\| \leqq \frac{M\|h\|^2}{r(r - \|h\|)} \tag{2.3}$$

*provided $\|h\|<r$.*

Since we may take the number $a$ of (ii) as small as necessary, (2.1) is merely a restatement of the hypothesis. It follows from (i) and (ii) in consequence of a result of Zorn [9] that $T(y)$ is Fréchet differentiable and hence analytic on $S$. In view of this, inequalities (2.2) and (2.3) with $r\equiv a/2<1$ can be readily established (see Hille [2, proof of theorem 4.5.1]).

Since $T(y)$ is Fréchet differentiable in $S$, $\delta T(y; h)$ is a linear and bounded operator in $h$ for $y$ in $S$. It then follows from condition (iii) that $\delta T^{-1}(y_0; h)$ exists and is also linear and bounded. Hence, there exists a constant $B$ such that

$$\|\delta T^{-1}(y_0; h)\| \leqq B\|h\|.$$

We now specify that by $M$ sufficiently small we mean

$$BM < \frac{1}{4}\frac{r^2}{1+r}\cdot \tag{2.4}$$

Consequently,

$$\|y_1 - y_0\| = \|\delta T^{-1}(y_0; T(y_0))\| \leqq BM < \frac{r}{8}\cdot \tag{2.5}$$

LEMMA 2.2. *For all $y$ in the neighborhood $\|y-y_0\|\leqq r$, $\delta T^{-1}(y; h)$*

*exists and is expressible in the form*

$$\delta T^{-1}(y; h) = H_0(y; \delta T^{-1}(y_0; h))$$

*where $H_0$ is linear in $h$ for each $y$ and satisfies*

$$\| H_0 \| \leqq \frac{1}{1-q}, \qquad 0 < q < \frac{1}{4}. \tag{2.6}$$

*Hence,*

$$\|\delta T^{-1}(y; h)\| \leqq \frac{B}{1-q} \|h\|. \tag{2.7}$$

Consider the linear operator $Q_0$ defined by

$$Q_0(y; h) \equiv \delta T^{-1}(y_0; \delta T(y; h) - \delta T(y_0; h)).$$

We have by Lemma 2.1

$$\|Q_0(y; h)\| \leqq B\|\delta T(y; h) - \delta T(y_0; h)\| \leqq \frac{2BM}{r}\|h\|$$

provided $\|y-y_0\| \leqq r$. Define $q = 2BM/r$. It then readily follows from (2.4) that $q$ lies on the interval $0<q<1/4$. Consequently, the linear operator $H_0(y; h)$ defined by

$$H_0(y; h) = [I(h) + Q_0(y; h)]^{-1}, \tag{2.8}$$

where $I$ is the identity operator, exists and has a norm satisfying (2.6). This last statement may be verified by consulting Graves and Hildebrandt [1, Lemma 16.1]. That $H_0(y; \delta T^{-1}(y_0; h))$ serves as the inverse with respect to $h$ of $\delta T(y; h)$ follows in a similar way as in the proof of Lemma 16.2 of the last cited reference.

By the lemma just proved we may define an operator

$$Q(y; h) = \delta T^{-1}(y'; \delta T(y; h) - \delta T(y'; h)),$$

provided $\|y'-y_0\| \leqq r$. It follows from Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned} \|Q(y; h)\| &\leqq \frac{B}{1-q}\|\delta T(y; h) - \delta T(y'; h)\| \\ &\leqq \frac{B}{1-q}\,\frac{2M}{r}\|h\| = \frac{q}{1-q}\|h\| \end{aligned}$$

provided $\|y-y_0\| \leqq r$. By (2.6), $0<q/(1-q)<1$. Hence, the operator $H(y; h)$ defined by dropping the subscript in (2.8) exists and has a norm satisfying

$$\|H\| \leqq \frac{1-q}{1-2q}, \tag{2.9}$$

and is such that

$$\delta T^{-1}(y; h) = H(y; \delta T^{-1}(y'; h))$$

for all $y$ and $y'$ satisfying $\|y-y_0\| \leqq r$, $\|y'-y_0\| \leqq r$.

LEMMA 2.3. *Let $y_k$ and $y_{k-1}$ be elements of the $r$-neighborhood of $y_0$ such that $\|y_k-y_{k-1}\| \equiv \eta_{k-1} \leqq BM$. Then if*

$$y_k = y_{k-1} - \delta T^{-1}(y_{k-1}; T(y_{k-1})), \tag{2.10}$$

*we have*

$$\|\delta T^{-1}(y_{k-1}; T(y_k))\| \leqq \frac{1}{2} \frac{q}{1-q} \eta_{k-1}.$$

Using (2.3), (2.7), and (2.10) one can construct the inequality

$$\begin{aligned}
&\|\delta T^{-1}(y_{k-1}; T(y_k))\| \\
&\quad = \|y_k - \delta T^{-1}(y_{k-1}; T(y_k)) - y_{k-1} + \delta T^{-1}(y_{k-1}; T(y_{k-1})) \\
&\qquad\qquad - y_k + y_{k-1} - \delta T^{-1}(y_{k-1}; T(y_{k-1}))\| \\
&\quad = \|\delta T^{-1}(y_{k-1}; \delta T(y_{k-1}; y_k - y_{k-1}) + T(y_{k-1}) - T(y_k))\| \\
&\quad \leqq \frac{B}{1-q} \|\delta T(y_{k-1}; y_k - y_{k-1}) + T(y_{k-1}) - T(y_k)\| \\
&\quad \leqq \frac{B}{1-q} \frac{M\eta_{k-1}^2}{r(r-\eta_{k-1})}.
\end{aligned}$$

Since $\eta_{k-1}/r-\eta_{k-1} \leqq BM/r-BM < 1$, we may finally write

$$\|\delta T^{-1}(y_{k-1}; T(y_k))\| \leqq \frac{1}{2} \frac{q}{1-q} \eta_{k-1}.$$

We are now ready to complete the proof of the convergence theorem. The essential tool is contained in the statement

$$\eta_i = \|y_{i+1} - y_i\| \leqq \left(\frac{1}{4} \frac{2q}{1-2q}\right)^i \eta_0, \qquad i = 0, 1, 2, \cdots, \tag{2.11}$$

whose proof will be obtained by induction. For $i=1$, it follows from (2.5) and the succeeding lemmas that

$$\begin{aligned}\eta_1 &= \|y_2 - y_1\| = \|\delta T^{-1}(y_1;\, T(y_1))\| \\ &= \|H(y_1;\, \delta T^{-1}(y_0;\, T(y_1)))\| \leqq \frac{1-q}{1-2q}\,\frac{1}{2}\,\frac{q}{1-q}\,\eta_0 \\ &= \frac{1}{4}\,\frac{2q}{1-2q}\,\eta_0.\end{aligned}$$

Assume that (2.11) holds for all $i \leqq k-1$. Then since $2q/(1-2q) < 1$, we have

$$\eta_{k-1} = \|y_k - y_{k-1}\| \leqq \left(\frac{1}{4}\,\frac{2q}{1-2q}\right)^{k-1} \eta_0 < \eta_0 \leqq BM$$

and

$$\begin{aligned}\|y_k - y_0\| &\leqq \sum_{i=0}^{k-1} \eta_i \leqq \frac{2(1-2q)}{2-5q}\,\eta_0 \\ &\leqq \frac{8}{3}\,\eta_0 \leqq \frac{r}{3}.\end{aligned}$$

Similarly, $\|y_{k-1} - y_0\| \leqq r/3$. Thus Lemma 2.2 and Lemma 2.3 may be applied to $\|\delta T^{-1}(y_k;\, T(y_k))\|$. The result is as follows:

$$\begin{aligned}\eta_k &= \|\delta T^{-1}(y_k;\, T(y_k))\| = \|H(y_k;\, \delta T^{-1}(y_{k-1};\, T(y_k)))\| \\ &\leqq \frac{1-q}{1-2q}\,\|\delta T^{-1}(y_{k-1};\, T(y_k))\| \leqq \frac{1}{4}\,\frac{2q}{1-2q}\,\eta_{k-1} \\ &= \left(\frac{1}{4}\,\frac{2q}{1-2q}\right)^{k} \eta_0.\end{aligned}$$

Hence, (2.11) holds for $i = k$ and so for all positive integers $i$.

Since the sequence $\{y_i\}$ of the theorem is such that $\sum_i \eta_i$ is convergent, it has a unique limit $\bar{y}$. It is clear that $\bar{y}$ must be an element of the $r/3$-neighborhood of $y_0$. It is a consequence of the inequality

$$\begin{aligned}0 &\leqq \lim_{i\to\infty} \|T(y_i)\| = \lim_{i\to\infty} \|\delta T(y_i;\, y_{i+1} - y_i)\| \\ &\leqq \lim_{i\to\infty} \frac{M}{r}\,\|y_{i+1} - y_i\| = 0\end{aligned}$$

and the continuity of $T(y)$ that $T(\bar{y}) = 0$.

We choose $d$ of the theorem equal to $r/3$. To see that $\bar{y}$ is the unique solution of (1.2) in the $d$-neighborhood of $y_0$ let $y^*$ be another element of this neighborhood also satisfying (1.2). Then

$$\begin{aligned}\|\bar{y} - y^*\| &= \|\bar{y} - y^* - \delta T^{-1}(\bar{y}; T(\bar{y})) + \delta T^{-1}(\bar{y}; T(y^*))\| \\ &= \|\delta T^{-1}(\bar{y}; \delta T(\bar{y}; \bar{y} - y^*) - T(\bar{y}) + T(y^*))\|.\end{aligned}$$

Since $\|\bar{y}-y^*\| \leqq 2r/3$, Lemma 2.1 and Lemma 2.2 are applicable. Thus

$$\|\bar{y} - y^*\| \leqq \frac{BM}{1-q} \frac{\|\bar{y} - y^*\|^2}{r(r - \|\bar{y} - y^*\|)}.$$

It is easy to show that

$$\frac{1}{2} \frac{\|\bar{y} - y^*\|}{r - \|\bar{y} - y^*\|} \leqq 1$$

while $q/(1-q) < 1$. Hence,

$$\|y^* - \bar{y}\| < \|y^* - \bar{y}\|$$

i.e., $\bar{y} \equiv y^*$.

The quadratic nature of the convergence can be demonstrated by the following computation:

$$\begin{aligned}\|y_{k+1} - y_k\| &= \|\delta T^{-1}(y_k; T(y_k) - T(y_{k-1}) - \delta T(y_{k-1}; y_k - y_{k-1}))\| \\ &\leqq \frac{1}{2} \frac{q}{1-q} \frac{1}{r - \eta_{k-1}} \eta_{k-1}^2 \leqq \frac{1}{2} \frac{q}{1-q} \frac{1}{r - BM} \eta_{k-1}^2 \\ &= C\|y_k - y_{k-1}\|^2.\end{aligned}$$

## References


1. L. M. Graves and T. H. Hildebrandt, *Implicit functions in general analysis*, Trans. Amer. Math. Soc. vol. 29 (1927) pp. 127–153.

2. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, 1948.

3. L. V. Kantorovič, *On Newton's method for functional equations*, Doklady Akad. Nauk SSSR N.S. vol. 59 (1948) pp. 1237–1240.

4. ———, *Functional analysis and applied mathematics*, Uspehi Matematiceskih Nauk N.S. vol. 3 (1948) pp. 89–185.

5. ———, *On Newton's method*, Trudy Mat. Inst. Steklov vol. 28 (1949) pp. 104–144.

6. I. P. Mysovskih, *On the convergence of Newton's method*, Trudy Mat. Inst. Steklov vol. 28 (1949) pp. 145–147.

7. ———, *On the convergence of the method of L. V. Kantorovič for the solution of functional equations and its applications*, Doklady Akad. Nauk SSSR N.S. vol. 70 (1950) pp. 565–568.

8. M. L. Stein, *On methods for obtaining solutions of fixed end point problems in the calculus of variations*, to appear in J. Research NBS.

9. M. Zorn, *Gateaux differentiability and essential boundedness*, Duke Math. J. vol. 12 (1945) pp. 579–583.


National Bureau of Standards, Los Angeles