Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Vol. No. 02, 1996, pp. 34–41.
Submitted: June 17, 1996. Published: September 9, 1996.

ON THE CONVERGENCE OF PERTURBED NEWTON-LIKE METHODS IN BANACH SPACE AND APPLICATIONS

Ioannis K. Argyros

Abstract. In this study we use disturbed Newton-like methods to approximate a locally unique solution of a nonlinear equation containing a nondifferentiable term. We choose the linear operators to be inverted at every step of the computation of the Newton iterates to be associated not only with the differentiable part of the equation as it has happened so far but also with the nondifferentiable part. This way the upper error bounds on the distances are smaller than before.

1991 A.M.S. (MOS) Subject Classification Codes. 65J15, 65B05, 65L50, 47H17, 49D15.

Key Words and Phrases. Disturbed Newton-like methods, Banach space, Mysovskii theorem, divided difference of order one, Fréchet-derivative.

I. INTRODUCTION

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

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

where the operators F, Q are defined on an open, convex subset D of a Banach space E_1 with values in a Banach space E_2, and are continuous on D.

We introduce disturbed Newton-like methods of the form

$$x_{n+1} = x_n - A_n^{-1}(F(x_n) + Q(x_n)) - z_n \quad (n \geq 0), \tag{2}$$

where

$$A_n = A(x_{n-1}, x_n) = L_1(x_{n-1}, x_n) + L_2(x_{n-1}, x_n) + L_3(x_{n-1}, x_n) \quad (n \geq 0)$$

for $x_{-1}, x_0 \in D$ and $A, L_1, L_2, L_3 : D \times D \to L(E_1, E_2)$.

Cameron University, Lawton, Oklahoma
E-mail Address: ioannisa@cameron.edu

The points $z_n \in E_1$ $(n \geq 0)$ are determined in such a way that the iteration $\{x_n\}$ $(n \geq 0)$ converges to a solution x^* of equation (1). The importance of studying disturbed Newton-like methods comes from the fact that many commonly used variants of Newton's method can be considered procedures of this type. Indeed, approximation (2) characterizes any iterative process in which the corrections are taken as approximate solutions of the Newton equations. Note that if for example an equation on the real line is solved, $F(x_n) + Q(x_n) \geq 0$, and A_n overestimates the derivative, then $x_n - A_n^{-1}(F(x_n) + Q(x_n))$ $(n \geq 0)$ is always larger than the corresponding Newton iterate. In such cases, a positive z_n $(n \geq 0)$ correction term is appropriate. We assume that $L_1(x,y)$ is used to approximate the difference $F(z) - F(y)$ for $x, y, z \in D$; $L_2(x,y)$ is used to approximate the difference $Q(z) - Q(y)$ for all $x, y, z \in D$ and $L_3(x_{n-1}, x_n)$ is a linear operator depending on F, Q, x_n $(n \geq -1)$ and z_n $(n \geq 0)$. Special choices of operators A_n $(n \geq 0)$ and points z_n $(n \geq 0)$ lead to iterations that have already been examined by many authors. Indeed: (a) for $Q(x) = 0$ $(x \in D)$, $L_1 = F'(\cdot)$, $L_2 = L_3 = 0$ (linear operator), and $z_n = 0$ $(n \geq 0)$ iteration (2) reduces to Newton's method [3], [5], [10]; (b) for $Q(x) = 0$ $(x \in D)$, $L_1(x,y) = [x, y; F]$ (divided difference of order one of F on D at the points $x, y \in D$), $L_2 = L_3 = 0$, and $z_n = 0$ $(n \geq 0)$ iteration (2) reduces to the secant method [3], [5], [6]; (c) for $Q(x) = 0$ $(x \in D)$, $L_1(x,y) = [g_1(x), g_2(y); F]$, $E_1 = E_2 = E$, $g_1, g_2 : D \to E$, $L_2 = L_3 = 0$ and z_n $(n \geq 0)$ we get the Steffensen–Aitken method [3], [8], [9]; (d) for $L_2 = L_3 = 0$, $L_1 = F'(\cdot)$, and $z_n = 0$ $(n \geq 0)$ we obtain the Zincenko iteration [3], [9], [10]; (e) for $E_1 = E_2 = E$, $L_1(x_{n-1}, x_n) = F'(x_n)$ $(n \geq 0)$, $L_2(x_{n-1}, x_n) = [x_{n-1}, x_n; Q]$ $(n \geq 0)$, $L_3 = 0$ and $z_n = 0$ $(n \geq 0)$ we obtain the iteration introduced in [4]; (f) for $L_2 = L_3 = 0$, $L_1'(x,y) = A(x)$, (A approximating $F'(\cdot)$) and $z_n = 0$ $(n \geq 0)$ we get iterations considered in [1], [2], [3], [9].

We show under more general and weaker conditions that our iteration (2) is faster than the ones mentioned above.

II. CONVERGENCE ANALYSIS

Let $x_{-1}, x_0 \in D$, $z_n \in E_1$ $(n \geq 0)$, and $a_1, a_2, a_3 \geq 0$, $r \in [0, R]$, $R \geq 0$, $U(x_0, R) = \{x \in E_1 \mid \|x - x_0\| \leq R\}$ with $U(x_0, R) \subseteq D$, $C_1, C_2, C_3 : [0, +\infty) \to [0, +\infty)$ be continuous nondecreasing functions that vanish at the origin, and $A, L_1, L_2, L_3 : D \times D \to L(E_1, E_2)$ such that

$$\begin{aligned} (3) \quad & \|A(y,z)^{-1}[F(z) - F(y) - L_1(x,y)(z-y)]\| \\ & \leq [C_1(\|x - x_0\|, \|y - x_0\|, \|z - x_0\|, \|x - y\|, \|z - y\|) + a_1]\|z - y\|, \end{aligned}$$

$$\begin{aligned} (4) \quad & \|A(y,z)^{-1}[Q(z) - Q(y) - L_2(x,y)(z-y)]\| \\ & \leq [C_2(\|x - x_0\|, \|y - x_0\|, \|z - x_0\|, \|x - y\|, \|z - y\|) + a_2]\|z - y\|, \end{aligned}$$

and

$$\begin{aligned} (5) \quad & \|A(x_{n+1}, x_n)^{-1}[A_{n+1}(z_{n+1}) - A_n(z_n) - L_3(x_{n-1}, x_n)(x_{n+1} - x_n)]\| \\ & \leq [C_3(\|x_{n-1} - x_0\|, \|x_n - x_0\|, \|x_{n+1} - x_0\|, \\ & \qquad \|x_{n-1} - x_n\|, \|x_{n+1} - x_n\|) + a_3]\|x_{n+1} - x_n\| \quad (n \geq 0) \end{aligned}$$

for all $x, y, z \in U(x_0, r)$ where x_n $(n \geq -1)$ are given by (2).

Let a_4, a_5 be constants such that

$$\|x_0 - x_{-1}\| \leq a_4 \quad \text{and} \quad \|A_0^{-1}(F(x_0) + Q(x_0)) - z_0\| \leq a_5. \tag{6}$$

It is convenient to introduce the iterations

$$\begin{aligned}\bar{h}_{n+1} = C(&\|x_{n-1} - x_0\|, \|x_n - x_0\|, \|x_{n+1} - x_0\|,\\ &\|x_{n-1} - x_n\|, \|x_{n+1} - x_n\|)\|x_{n+1} - x_n\| \ (n \geq 0)\end{aligned}$$

where

$$C(s_1, s_2, s_3, s_4, s_5) = (C_1 + C_2 + C_3)(s_1, s_2, s_3, s_4, s_5) + a_1 + a_2 + a_3$$

for all $s_1, s_2, s_3, s_4, s_5 \in [0, +\infty)$,

$$t_{n+1} = t_{n+1} + h_{n+1}, \ t_{-1} = 0 \quad (n \geq 0), \tag{7}$$

where

$$h_{n+1} = C(|t_{n-1} - t_0|, t_n - t_0, t_{n+1} - t_0, t_n - t_{n-1}, t_{n+1} - t_n)(t_{n+1} - t_n) \quad (n \geq 0),$$

and the functions $T, C_4 : [0, +\infty) \to [0, +\infty)$ by

$$T(r) = t_1 + C(r, r, r, r, r)(r - t_1) + C(t_0, 0, t_1 - t_0, t_0, t_1 - t_0)(t_1 - t_0) \tag{8}$$

and

$$C_4(s_1, s_2, s_3, s_4, s_5) = (C_1 + C_2)(s_1, s_2, s_3, s_4, s_5) + a_1 + a_2. \tag{9}$$

We can now state and prove the following Newton–Mysovskii [6] semilocal convergence theorem for the disturbed Newton like method generated by (2).

Theorem. *Assume:*

(a) *operators F, Q are defined on some open convex subset D of a Banach space E_1 with values in E_2 and continuous on D;*
(b) *there exist r, R with $r \in [0, R]$ such that $A(x, y) \in L(E_1, E_2)$ for all $x, y \in U(x_0, r)$ is boundedly invertible, $U(x_0, R) \subseteq D$ and conditions* (3)–(5) *are satisfied;*
(c) *sequence $\{z_n\}$ $(n \geq 0) \in E_1$ is null;*
(d) *there exist numbers t_0 and t_1 such that*

$$a_5 \leq t_1 - t_0 \leq \|x_0 - x_{-1}\| \leq a_4 \leq t_0 - t_{-1}; \tag{10}$$

(e) *there exists a minimum nonnegative number r^* satisfying*

$$r^* \leq R \tag{11}$$

and

$$T(r^*) \leq r^* \tag{12}$$

where function T is given by (8);
(f) *moreover r^*, R satisfy*

$$C_4(r^*, r^*, R, 2r^*, r^* + R) < 1 \tag{13}$$

where function C_4 is given by (9).

Then

(i) *Scalar iteration* $\{t_n\}$ $(n \geq -1)$ *generated by* (7) *is monotonically increasing, bounded above by* r^* *and* $\lim_{n\to\infty} t_n = r^*$;

(ii) *disturbed Newton-like method* $\{x_n\}$ $(n \geq -1)$ *generated by* (2) *is well defined, remains in* $U(x_0, r^*)$ *for all* $n \geq -1$, *and converges to a unique solution* x^* *of equation* $F(x) + Q(x) = 0$ *in* $U(x_0, R)$.

Moreover the following error estimates are true for all $n \geq 0$

$$\|x_{n+1} - x_n\| \leq \bar{h}_{n+1} \leq h_{n+1} = t_{n+1} - t_n \tag{14}$$

and

$$\|x_n - x^*\| \leq r^* - t_n. \tag{15}$$

Proof. (i) It follows from (7), (10) and (12) that $t_{-1} \leq t_0 \leq t_1 \leq r^*$. Let us assume that $t_k \leq t_{k+1} \leq r^*$ for $k = -1, 0, 1, 2, \ldots, n$. The C function is nonnegative and nondecreasing, hence $h_{k+1} \geq 0$. That is $t_{k+1} \leq t_{k+2}$ (by (7)). Moreover from (7), (12) and the induction hypothesis we obtain in turn

$$\begin{aligned} t_{k+2} &= t_{k+1} + h_{k+1} \leq t_{k+1} + C(r^*, r^*, r^*, r^*, r^*)(t_{k+1} - t_k) \\ &\leq \cdots \leq t_1 + C(r^*, r^*, r^*, r^*, r^*)(t_{k+1} - t_1) \\ &\qquad + C(t_0, 0, t_1 - t_0, t_0, t_1 - t_0)(t_1 - t_0) \\ &= T(r^*) \leq r^*. \end{aligned}$$

Hence, iteration $\{t_n\}$ $(n \geq -1)$ is also bounded above by r^*. Since r^* is the minimum nonnegative number on $[0, r^*]$ satisfying (12) it follows that $\lim_{n\to\infty} t_n = r^*$.

(ii) It follows from (10), (12) that $x_{-1}, x_1 \in U(x_0, r^*)$, and that (14) is satisfied for $n = -1, 0$. Assume that $x_k \in U(x_0, r^*)$ and that (14) is satisfied for $k = -1, 0, \ldots, n$. Using (14), the triangle inequality and the induction hypothesis we obtain

$$\|x_{k+1} - x_0\| \leq \sum_{i=0}^{k} \|x_{k+1} - x_i\| \leq \sum_{i=0}^{k} (t_{i+1} - t_i) = t_{k+1} - t_0 \leq r^*.$$

That is $x_{k+1} \in U(x_0, r^*)$. We will show that (14) is true for $k = n + 1$. Using (2)–(5), (7), the induction hypothesis we can obtain the approximation

$$\begin{aligned} F(x_{k+1}) + Q(x_{k+1}) + A_{k+1}(z_{k+1}) &= F(x_{k+1}) + Q(x_{k+1}) \\ &+ A_{k+1}(z_{k+1}) - F(x_k) - Q(x_k) - A_k(z_k) - A_k(x_{k+1} - x_k) \\ &= [F(x_{k+1}) - F(x_k) - L_1(x_{k-1}, x_k)(x_{k+1} - x_k)] \\ &+ [Q(x_{k+1}) - Q(x_k) - L_2(x_{k-1}, x_k)(x_{k+1} - x_k)] \\ &+ [A_{k+1}(z_{k+1}) - A_k(z_k) - L_3(x_{k-1}, x_k)(x_{k+1} - x_k)], \end{aligned} \tag{16}$$

and

(17)
$$
\begin{aligned}
\|x_{k+2}-x_{k+1}\| &= \|A_{k+1}^{-1}[F(x_{k+1})+Q(x_{k+1})-A_{k+1}(z_{k+1})]\| \\
&\leq [C_1(\|x_{k-1}-x_0\|,\|x_k-x_0\|,\|x_{k+1}-x_0\|, \\
&\qquad \|x_{k-1}-x_k\|,\|x_{k+1}-x_k\|)+a_1]\|x_{k+1}-x_k\| \\
&\quad + [C_2(\|x_{k-1}-x_0\|,\|x_k-x_0\|,\|x_{k+1}-x_0\|,\|x_{k-1}-x_k\|, \\
&\qquad \|x_{k+1}-x_k\|)+a_2]\|x_{k+1}-x_k\| \\
&\quad + [C_3(\|x_{k-1}-x_0\|,\|x_k-x_0\|,\|x_{k+1}-x_0\|,\|x_{k-1}-x_k\|, \\
&\qquad \|x_{k+1}-x_k\|)+a_3]\|x_{k+1}-x_k\| \\
&= C(\|x_{k-1}-x_0\|,\|x_k-x_0\|,\|x_{k+1}-x_0\|,\|x_{k-1}-x_k\|, \\
&\qquad \|x_{k+1}-x_k\|)\|x_{k+1}-x_k\| = \bar{h}_{k+1} \\
&\leq C(|t_{k-1}-t_0|, t_k-t_0, t_{k+1}-t_0, t_k-t_{k-1}, t_{k+1}-t_k)(t_{k+1}-t_k) \\
&= h_{k+1} = t_{k+2}-t_{k+1},
\end{aligned}
$$

which shows (14) for all $n \geq 0$.

It now follows from part (i) and (14) that the inexact Newton-like iteration $\{x_n\}$ $(n \geq -1)$ is Cauchy in a Banach space E_1 and as such it converges to some $x^* \in U(x_0, r^*)$ (since $U(x_0, r^*)$ is a closed set). By letting $n \to \infty$ in (16) and using hypothesis (c) we deduce that $F(x^*)+Q(x^*)=0$. That is x^* is a solution of equation (1). Estimate (15) follows immediately from (14) by using standard majorization techniques [3], [5].

To show uniqueness, we reason as follows: under the hypotheses of the theorem iteration (2) for $z_n = 0$ $(n \geq 0)$ becomes

(18)
$$
v_{n+1} = v_n - A_n^{-1}(F(v_n)+Q(v_n)), \; v_{-1} = x_{-1}, v_0 = x_0 \quad (n \geq 0)
$$

and converges to some $x_1^* \in U(x_0, r^*)$. Let us assume that there exists $y^* \in U(x_0, R)$ such that $F(y^*)+Q(y^*)=0$. Using (18), (3), (4) and the approximation

$$
\begin{aligned}
v_{n+1}-y^* &= v_n - A_n^{-1}(F(v_n)+Q(v_n)) - y^* \\
&= A_n^{-1}[F(v_n)-F(y^*)-L_1(v_{n-1},v_n)(v_n-y^*)] \\
&\quad + A_n^{-1}[Q(v_n)-Q(y^*)-L_2(v_{n-1},v_n)(v_n-y^*)]
\end{aligned}
$$

we obtain

(19)
$$
\begin{aligned}
\|v_{n+1}-y^*\| &\leq [C_1(\|v_{n-1}-v_0\|,\|v_n-v_0\|,\|y^*-v_0\|,\|v_{n-1}-v_n\|, \\
&\qquad \|v_n-y^*\|)+a_1]\|v_n-y^*\| \\
&\quad + [C_2(\|v_{n-1}-v_0\|,\|v_n-v_0\|,\|y^*-v_0\|,\|v_{n-1}-v_n\|, \\
&\qquad \|v_n-y^*\|)+a_2]\|v_n-y^*\| \\
&\leq C_4(\|v_{n-1}-v_0\|,\|v_n-v_0\|,\|y^*-v_0\|,\|v_{n-1}-v_n\|, \\
&\qquad \|v_n-y^*\|)\|v_n-y^*\|.
\end{aligned}
$$

By (13) there exist $r^*, q \in [0,1]$ such that

$$C(r^*, r^*, R, 2r^*, r^* + R) \leq q < 1. \tag{20}$$

It follows from (19) and (20) that

$$\|v_{n+1} - y^*\| \leq q^{n+1}\|v_0 - y^*\| \leq q^{n+1}R \to 0 \quad \text{as } n \to \infty.$$

That is, $y^* = \lim_{n\to\infty} v_{n+1} = x_1^*$. Hence we deduce $x^* = x_1^* = y^*$.
The proof of the Theorem is now complete.

Remarks. (1) Hypotheses (3)–(5) can be replaced by weaker (3)′–(5)′ if $A(y,z)^{-1}$ is replaced by A_0^{-1} for all $y, z \in U(x_0, r)$. In this case we need to add a hypothesis of the form

$$\|A_0^{-1}(A(x,y) - A_0)\| \leq C_5(\|x - x_{-1}\|, \|y - x_0\|) + a_6, \tag{21}$$

with $C_5 : [0,+\infty)^2 \to [0,+\infty)$ be a continuous nondecreasing function vanishing at the origin and $a_6 \geq 0$. Parameter r^* must satisfy the additional hypothesis

$$C_5(r^*, r^*) + a_6 < 1.$$

It then follows from the Banach Lemma on invertible operators [5] that

$$\|A_{n+1}^{-1}A_0\| \leq \bar{b}_{n+1} \leq b_{n+1} \quad (n \geq 0)$$

where

$$\bar{b}_{n+1} = \{1 - [a_6 + C_5(\|x_n - x_{-1}\|, \|x_{n+1} - x_0\|)]\}^{-1}$$

and

$$b_{n+1} = \{1 - [a_6 + C_5(t_n, t_{n+1} - t_0)]\}^{-1}.$$

The following modifications must also take place: iterations $\bar{h}_{n+1}$, h_{n+1} must be multiplied by $\bar{b}_{n+1}$, b_{n+1} $(n \geq 0)$ respectively, and functions C, C_4 must be multiplied by

$$b(r) = \{1 - [a_6 + C_5(r,r)]\}^{-1}.$$

It can then be easily seen that with the above modifications the conclusions of the Theorem under the above-mentioned weaker assumptions will also hold.

(2) Several choices for the "C" functions appearing in conditions (3)–(5) can be suggested. To compare with conditions suggested by Yamamoto and Chen in [9] under the special choice (f) stated at the introduction of this study, we choose

$$C_1(s_1,s_2,s_3,s_4,s_5) = \int_0^1 [w_1(s_1 + ts_4) - w_2(s_1) + w_3(s_2 + ts_5) - w_4(s_2) + a_1]s_5 dt, \tag{22}$$

$$C_2(s_1,s_2,s_3,s_4,s_5) = [w_5(s_1 + s_4) - w_6(s_1) + w_7(s_2 + s_5) - w_8(s_2) + a_2]s_5 \tag{23}$$

and

(24) $C_3(s_1, s_2, s_3, s_4, s_5) = [w_9(s_1 + s_4) - w_{10}(s_1) + w_{11}(s_2 + s_5) - w_{12}(s_2) + a_3]s_5$

where $w_m : [0, +\infty) \to [0, +\infty)$, $m = 1, 2, \ldots, 12$ are continuous functions vanishing at the origin and satisfying

$$0 \leq w_i(d_1 + d_2) - w_{i+1}(d_1) \leq w_i(d_3 + d_4) - w_{i+1}(d_3), i = 1, 3, 5, 7, 9, 11$$

for all $r_1 \leq r_3$ and $r_2 \leq r_4$. Yamamoto–Chen conditions (4), (5) and (6) in [9] are given respectively by (following their notation)

$$\|A(x_0)^{-1}(A(x) - A(x_0))\| \leq w_0(\|x - x_0\|) + b$$

$$\|A(x_0)^{-1}(F'(x + t(y - x)) - A(x))\| \leq w(\|x - x_0\| + (t\|y - x\|) - w_0(\|x - x_0\|) + c$$

and

$$\|A(x_0)^{-1}(Q(x) - Q(y))\| \leq e(r)\|x - y\|.$$

These conditions correspond to our conditions (21), (3) and (4) respectively. It can easily be seen from (22)–(24) that for special choices of the w_i functions our conditions (21) and (4) reduce to Yamamoto–Chen (4) and (6), whereas our condition (4) is weaker than Yamamoto–Chen (5). Moreover since the left-hand side of Yamamoto–Chen condition (5) (times $(y - x)$ inside the norm) is larger than the left-hand side of our condition (3), our upper bounds on the distances $\|x_{n+1} - x_n\|$ and $\|x_n - x^*\|$ $(n \geq 0)$ are smaller than Yamamoto–Chen upper bounds.

(3) Favorable comparisons can also be made with the rest of the cases (a)–(e) mentioned at the introductions of this study. The details are left to the motivated reader. Numerical examples for the usefulness of iteration (2) can also be found in [1]–[4].

(4) Define the residuals $r_n = -A_n(z_n)$ and set $\Delta_n = x_{n+1} - x_n$ $(n \geq 0)$. Since the sequence $\{z_n\}$ $(n \geq 0)$ is null and the linear operators $\{A_n\}$ $(n \geq 0)$ are bounded it follows that the sequence $\{r_n\}$ $(n \geq 0)$ is null also. Consequently the conclusions of the Theorem hold for the system

$$x_{n+1} - x_n = \Delta_n,$$

and

$$A_n\Delta_n = -(F(x_n) + Q(x_n)) + r_n \quad (n \geq 0) \ \ [1], [3].$$

REFERENCES

[1] Argyros, I.K. On the convergence of inexact Newton-like methods, Publ. Math. Debrecen, Vol. 43, 1–2, (1993), 79–85.

[2] Argyros, I.K. A unified approach for constructing fast two-step Newton-like methods, Monatshefte für Mathematia, 119, (1995), 1–22.

[3] Argyros, I.K. and Szidarovszky, F. The theory and applications of iteration methods, C.R.C. Press, Inc. Boca Raton, Florida, 1993.

[4] Cătinas, E. On some iterative methods for solving nonlinear equations, Revue d'analyse Numerique et de theorie de l'approximation, 23, 1, (1994), 47–53.

[5] Kantorovich, L.V. and Akilov, G.P. Functional Analysis in Normed Spaces, Academic Press, New York, 1978.

[6] Mysovskii, I. On the convergence of Newton's method, Trudy Mat. Inst. Steklov, 28 (1949), 145–147. (In Russian).

[7] Păvăloiu, I. Sur une generalisation de la methode de Steffensen, Revue d'analyse Numerique et de theorie de l'approximation, 21, 1, (1992), 59–65.

[8] Păvăloiu, I. Bilateral approximations for the solutions of scalar equations, Revue d'analyse Numerique et de theorie de l'approximation, 23, 1, (1994), 95–100.

[9] Yamamoto, T. and Chen, X. Convergence domains of certain iterative methods for solving nonlinear equations, Numer. Funct. Anal. Optimiz. 10, (1 and 2), (1989), 37–48.

[10] Zabrejko, P.P. and Nguen, D.F. The majorant method in the theory of Newton–Kantorovich approximations and the Ptak error estimates, Numer. Funct. Anal. Optimiz. 9, (5 and 6), (1987), 671–684.