UDC 519.6

A Local and Semilocal Convergence of the Continuous Analogy of Newton's Method

T. Zhanlav*, **O. Chuluunbaatar**†

* *Department of Applied Mathematics*
National University of Mongolia Ulan-Bator, Mongolia
† *Laboratory of Information Technologies*
Joint Institute for Nuclear Research
Joliot-Curie 6, 141980 Dubna, Moscow region, Russia

In this paper, a region of convergence of the continuous analogy of Newton's method is defined and an optimal choice of the parameter τ is proposed. For the damped Newton's method a global convergence is proved and error bounds are obtained. The damping strategies allow one to extend the convergence domain of the initial guesses. Several damping strategies were compared. Numerical examples are given and confirm the theoretical results.

Key words and phrases: nonlinear equations in Banach spaces; damped Newton's method; recurrence relations; error bounds.

1. Introduction

One of the modifications of Newton method is the well-known continuous analogy of Newton's method (CANM) or damped Newton's method

$$x_{k+1} = x_k + \tau_k v_k, \quad F'(x_k)v_k = -F(x_k), \quad k = 0, 1, \dots \tag{1}$$

to solve nonlinear equations $F(x) = 0$ in Banach spaces. In this case $F : \Omega \subseteq X \to Y$ is an operator defined in an open convex subset Ω of the Banach space X into the Banach space Y. If $\tau_k \equiv 1$ the CANM leads to the Newton method. In our paper [1] we have shown that the iteration (1) is convergent if $\tau_k \in (0; 2)$. In the present paper, we give a τ-region of convergence of CANM and propose an optimal choice for the parameter τ_k in (1). We also prove the semilocal convergence theorem for the iteration (1) by using a technique consisting of a new system of recurrence relations [2].

2. The τ–region of Convergence of the CANM

We study the influence of the parameter τ on the convergence of method (1). Consider the open level set $L_\alpha = \{x \in \Omega; \|F(x)\| < \eta/\beta = \alpha\}$ and let $L \subset \Omega$ be bounded. In that follows we assume that

(c_1) $\|F'(y) - F'(x)\| \leqslant L\,\|y - x\|, \quad x, y \in \Omega,$
(c_2) $\|(F'(x_0))^{-1}\| \leqslant \beta,$
(c_3) $\|(F'(x_0))^{-1}F(x_0)\| \leqslant \eta, \quad a_0 = L\beta\eta,$
(c_4) $L\,\|(F'(x_k))^{-1}\|\,\|(F'(x_k))^{-1}F(x_k)\| \leqslant a_k, \quad k = 0, 1, \dots$

Received 27$^{\text{th}}$ September, 2011.
O. Chuluunbaatar acknowledges a financial support from the RFBR Grant No. 11-01-00523, and the theme 09-6-1060-2005/2013 "Mathematical support of experimental and theoretical studies conducted by JINR"..

Theorem 1. *Let us assume that the conditions* $(c_1)-(c_4)$ *are satisfied and* $\tau_k \in I_k$*, where*

$$I_k = \begin{cases} \left(0, \dfrac{-1+\sqrt{1+8a_k}}{2a_k}\right) \subseteq (0,2), & for \quad 0 < a_k \leqslant 1, \\ \left(0, \dfrac{1}{a_k}\right), & for \quad 1 < a_k < M < \infty. \end{cases} \tag{2}$$

Then the sequence $\{x_k\}$ *obtained by* (1) *is well defined and remains in* L_α *and converges to some* x^* *with* $F(x^*) = 0$.

Proof. We notice first of all that

$$\tau_k a_k < 1, \quad k = 0, 1, \ldots,$$

as $\tau_k \in I_k$. By induction we prove that there exists $\Gamma_{k+1} = (F'(x_{k+1}))^{-1}$, such that

$$\|\Gamma_{k+1}\| \leqslant \frac{\|\Gamma_k\|}{1 - a_k\tau_k}. \tag{3}$$

Now from the conditions (c_1) and (c_4) we have

$$\|I - \Gamma_k F'(x_{k+1})\| \leqslant \|\Gamma_k\| \, \|F'(x_k) - F'(x_{k+1})\| \leqslant L \, \|\Gamma_k\| \, \|x_{k+1} - x_k\| = \\ = \tau_k L \, \|\Gamma_k\| \, \|\Gamma_k F(x_k)\| \leqslant \tau_k a_k < 1, \quad k = 0, 1, \ldots$$

So, by the Banach perturbation lemma there should exist Γ_{k+1} satisfying (3). Thus the sequence $\{x_k\}$ obtained from the damped Newton iteration is well defined. The Taylor expansion of $F(x)$ about x_k gives

$$F(x_{k+1}) = F(x_k) + F'(\xi_k)(x_{k+1} - x_k) = \\ = (1 - \tau_k)F(x_k) + (F'(\xi_k) - F'(x_k))(x_{k+1} - x_k), \tag{4}$$

where $\xi_k = \theta x_k + (1-\theta)x_{k+1}$, $\theta \in (0,1)$. Using the conditions (c_1) and (c_4) in (4) we have

$$\|F(x_{k+1})\| \leqslant \psi(\tau_k)\|F(x_k)\|, \tag{5}$$

$$\psi(\tau_k) = |1 - \tau_k| + a_k\tau_k^2. \tag{6}$$

It is easy to show that

$$\psi(\tau_k) < 1 \tag{7}$$

under the condition $\tau_k \in I_k$. Therefore, from (5) we deduce the following

$$\|F(x_k)\| \leqslant \psi(\tau_{k-1})\psi(\tau_{k-2}) \cdots \psi(\tau_0)\|F(x_0)\|. \tag{8}$$

According to (7) we have

$$\|F(x_k)\| < \|F(x_0)\| < \frac{\eta}{\beta}, \quad k = 1, 2, \ldots,$$

i.e. $x_k \in L_\alpha$ for all $k = 1, 2, \ldots$ From (8) we obtaine

$$\|F(x_k)\| \to 0 \quad \text{at} \quad k \to \infty.$$

Since L_α is bounded, there exists an accumulation point x^* of the sequence $\{x_k\}$ with $F(x^*) = 0$. □

We call the interval I_k by the τ–region of convergence, as well as call the value τ_k, such that

$$\psi(\tau_k) = \min_{\tau_k \in (0,2)} \psi(\tau),$$

the optimal one and denote it by τ_{opt}. The direct calculation gives us

$$\tau_{opt} = \begin{cases} 1, & \text{for} \quad 0 < a_k \leqslant \frac{1}{2}, \\ \frac{1}{2a_k}, & \text{for} \quad \frac{1}{2} < a_k < 1, \\ \frac{1}{a_k} - \varepsilon, & \text{for} \quad a_k \geqslant 1, \end{cases}$$

$$\psi(\tau_{opt}) = \begin{cases} a_k, & \text{for} \quad 0 < a_k \leqslant \frac{1}{2}, \\ 1 - \frac{1}{4a_k} < \frac{3}{4}, & \text{for} \quad \frac{1}{2} < a_k < 1, \\ 1 - \varepsilon + \varepsilon^2 a_k, & \text{for} \quad a_k \geqslant 1. \end{cases} \tag{9}$$

Here ε is a small number. Some particular choices of the parameter, in contrast to the optimal one, may be useful. We consider

$$\tau_k = \frac{-1 + \sqrt{1 + 8a_k}}{4a_k}, \tag{10}$$

which is a middle point of the admissible interval I_k. Since $\tau_k < 1$, then from (6) we obtain

$$\psi(\tau_k) = \frac{3}{2}(1 - \tau_k). \tag{11}$$

In (11) we can see that $\psi(\tau_k) < 1$ if $\tau_k > 1/3$. On the other hand, the τ_k is decreasing and $\psi(\tau_k)$ is increasing with respect to a_k. Therefore from (10) and (11) we get

$$\frac{1}{3} < \tau_k < 1, \quad 0 < \psi(\tau_k) < 1, \quad \tau_k a_k < 1$$

under condition $0 < a_k < 3$. Thus, the choice (10) allows us to weaken the condition imposed on a_k. From (10), (11) it follows that $\tau_k \to 1$ and $\psi(\tau_k) \to 0$ at $a_k \to 0$. Moreover, if $a_{k+1} < a_k$ for $k = 0, 1, \dots$, then we have

$$\tau_0 < \tau_1 < \dots < \tau_k < \dots < 1, \quad 0 < \psi(\tau_k) < \dots < \psi(\tau_1) < \psi(\tau_0).$$

It should be pointed out that, according to affine invariance property [3] we can also take

$$\overline{\tau}_k = \frac{-1 + \sqrt{1 + 2ba_k}}{ba_k}, \quad \text{for} \quad \forall b > 0, \quad k = 0, 1, \dots, \tag{12}$$

which has the same properties as τ_k given by (10), i.e. $\overline{\tau}_k : (0; \infty) \to (0, 1)$ and decreasing with respect to a_k.

Theorem 2. *Suppose the conditions* $(c_1) - (c_4)$ *are satisfied and* τ_k *is given by* (12) *with* $0 < a_k < \frac{b+2}{2}$. *Then the sequence* $\{x_k\}$ *obtained by* (1) *is well defined and remains in* L_α *and converges to some* x^* *with* $F(x^*) = 0$.

Proof. By virtue of (6) and (12) we have

$$\psi(\overline{\tau}_k) = \frac{b+2}{b}(1 - \overline{\tau}_k).$$

Since $\overline{\tau}_k(a_k)$ is a decreasing function with respect to a_k and $\psi(\overline{\tau}_k)$ is decreasing too with respect to $\overline{\tau}_k$, and $0 < a_k < \dfrac{b+2}{2}$, then we get $\dfrac{2}{b+2} < \overline{\tau}_k < 1$, and $0 < \psi(\overline{\tau}_k) < 1$, $\overline{\tau}_k a_k < 1$. The proof follows immediately from (5) and (8). □

The above derived admissible interval (2) and the theoretical optimal value as well as the particular choices (10), (12) cannot be implemented directly, since the arising quantities a_k are computationally unavailable due to the arising Lipschitz constant, $\|\Gamma_k\|$ and $\|\Gamma_k F(x_k)\|$. However, the obtained theoretical results can be useful for the construction of computational strategies. Namely, from the assumption (c_4) it is evident that $\|F(x_k)\| \to 0$ if $a_k \to 0$, and $a_k \to \infty$ if $\|F(x_k)\| \to \infty$.

Therefore, replacing a_k in (12) by $y_k = \|F(x_k)\|$, we obtain

$$\tau_k^* = \frac{2}{1+\sqrt{1+2by_k}} \in (0,1), \quad b > 0. \tag{13}$$

The function $\tau_k^* = \tau_k^*(y_k)$ decreases with respect to y_k. From (13) it is also clear that $\tau_k^* \to 1$ if $y_k \to 0$, and $\tau_k^* \to 0$ if $y_k \to \infty$. We have the following theorem.

Theorem 3. *Suppose the conditions $(c_1) - (c_4)$ are satisfied and τ_k is given by* (13). *Suppose also that $0 < a_k < \dfrac{\sqrt{1+2by_k}}{2}$. Then the sequence $\{x_k\}$ obtained by* (1) *is well defined and remains in L_α and converges to some x^* with $F(x^*) = 0$.*

Proof. By virtue of (6) and (13) we have

$$\psi(\tau_k^*) = \frac{by_k + 2a_k}{1 + by_k + \sqrt{1+2by_k}}. \tag{14}$$

From (13), (14) it is easy to show that the next two inequalities hold simultaneously

$$\tau_k^* a_k < 1 \quad \text{and} \quad \psi(\tau_k^*) < 1 \tag{15}$$

under the condition

$$0 < a_k < \frac{1+\sqrt{1+2by_k}}{2} \equiv M_k \tag{14a}$$

or equivalently

$$y_k > \frac{2a_k(a_k-1)}{b}. \tag{14b}$$

Since $0 < \tau_k^* < 1$ and $\tau_k^* a_k < 1$ then $\tau_k^* \in I_k$. As a consequence, the assertion of the theorem follows directly from Theorem 1. □

Corollary. Theorem 3 is valid for any choice $\tau_k \in (0, \tau_k^*]$.

Indeed, $\tau_k a_k \leqslant \tau_k^* a_k < 1$ and for such τ_k we have $\psi(\tau_k) = 1 - \tau_k + a_k \tau_k^2 < 1$, that assures a reduction of the residual norm, i.e. $\|F(x_{k+1})\| < \|F(x_k)\|$. However the choice τ_k^* is preferable in the sense that $\tau_k^* \to 1$ as $y_k \to 0$ and the method (1) asymptotically leads to quadratically convergent Newton method. Nevertheless, there arises a question of the choice of the parameter b in (13). In our opinion, it must be done in such a way that the damped Newton's method converges rapidly.

3. A Semilocal Convergence of CANM

We define the sequence

$$a_{n+1} = f(a_n, \tau_n)^2 g(a_n, \tau_n) a_n, \quad a_0 = L\beta\eta, \quad n = 0, 1, \ldots, \tag{16}$$

where

$$f(x,\tau)=\frac{1}{1-\tau x},\quad g(x,\tau)=|1-\tau|+\tau^2 x. \tag{17}$$

We need the following auxiliary lemmas, whose proofs are trivial from [2].

Lemma 1. *Let f and g be two real functions given by (17). Then*

(i) For a fixed $\tau\in(0,1]$, $f(x,\tau)$ and $g(x,\tau)$ increase as a function of $x\in(0,1/\tau)$ and $f(x,\tau)>1$,

(ii) $f(\gamma x;\tau)<f(x;\tau)$ and $g(\gamma x,\tau)\leqslant\gamma^p g(x,\tau)$ for $\gamma\in(0,1]$,

where

$$p=\begin{cases}0, & \text{if}\quad \tau\neq 1,\\ 1, & \text{if}\quad \tau=1.\end{cases}$$

Lemma 2. *Let $0<a_0<\dfrac{3-\sqrt{5}}{2}$ and $\tau_k\in(0,1)$ for all $k=0,1,\ldots$ Then the sequence $\{a_n\}$ is decreasing, i.e.*

$$0<\cdots<a_n<a_{n-1}<\cdots<a_1<a_0<\frac{3-\sqrt{5}}{2}.$$

Lemma 3. *Let us suppose the conditions of Lemma 2 are satisfied and define $\gamma=a_1/a_0$. Then*

(i) $\gamma=f(a_0,\tau_0)^2 g(a_0,\tau_0)\in(0,1)$,

(ii) $a_n\leqslant\gamma^{(1+p)^{n-1}}a_{n-1}\leqslant\cdots\leqslant\gamma^{\frac{(1+p)^n-1}{p}}a_0$,

(iii) $f(a_n,\tau_n)g(a_n,\tau_n)\leqslant\gamma^{(1+p)^n}\Delta$ with $\Delta=\dfrac{1}{f(a_0,\tau_0)}<1$.

Notice that

$$L\,\|\Gamma_0\|\,\|\Gamma_0F(x_0)\|\leqslant a_0,$$

$$\|x_1-x_0\|\leqslant\|\Gamma_0F(x_0)\|\leqslant\eta<R\eta\quad\text{with}\quad R=\frac{1}{1-\Delta\gamma}>1,$$

i.e. $x_1\in B(x_0,R\eta)=\{x\in X;\ \|x-x_0\|<R\eta\}$. In these conditions we prove, for $n\geqslant1$, the following statements:

(I_n) $\|\Gamma_n\|\leqslant f(a_{n-1},\tau_{n-1})\,\|\Gamma_{n-1}\|$,

(II_n) $\|\Gamma_nF(x_n)\|\leqslant f(a_{n-1},\tau_{n-1})g(a_{n-1},\tau_{n-1})\,\|\Gamma_{n-1}F(x_{n-1})\|$,

(III_n) $L\,\|\Gamma_n\|\,\|\Gamma_nF(x_n)\|\leqslant a_n$,

(IV_n) $x_{n+1}\in B(x_0,R\eta)$.

Assuming $\tau_0a_0<1$ and $x_1\in\Omega$, we have

$$\|I-\Gamma_0F'(x_1)\|\leqslant\|\Gamma_0\|\,\|F'(x_0)-F'(x_1)\|\leqslant L\,\|\Gamma_0\|\,\|x_1-x_0\|=$$
$$=\tau_0L\,\|\Gamma_0\|\,\|\Gamma_0F(x_0)\|\leqslant\tau_0a_0<1.$$

Then by the Banach lemma, Γ_1 exists and

$$\|\Gamma_1\|\leqslant\frac{\|\Gamma_0\|}{1-\|\Gamma_0\|\,\|F'(x_0)-F'(x_1)\|}\leqslant f(a_0,\tau_0)\,\|\Gamma_0\|,$$

and (I_1) is true. On the other hand, according to (1), we have

$$F(x_{n+1})=(1-\tau_n)F(x_n)+(F'(\xi_n)-F'(x_n))(x_{n+1}-x_n),$$

where $\xi_n=\theta x_n+(1-\theta)x_{n+1}$, $\theta\in(0,1)$, and

$$F'(x_{n+1})=F'(x_n)(I-P_n),\quad P_n=\Gamma_n(F'(x_n)-F'(x_{n+1})).$$

If $x_n, x_{n+1} \in \Omega$ following

$$\|P_n\| \leqslant L\,\|\Gamma_n\|\,\|x_{n+1} - x_n\| \leqslant \tau_n L\,\|\Gamma_n\|\,\|\Gamma_n F(x_n)\| \leqslant \tau_n a_n < 1.$$

Then $\Gamma_{n+1} = (I - P_n)^{-1}\Gamma_n$, and for $n = 0$, if $x_0, x_1 \in \Omega$, we have

$$\Gamma_1 F(x_1) = (I - P_0)^{-1}\Gamma_0\{(1-\tau_0)F(x_0) + (F'(\xi_0) - F'(x_0))(x_1 - x_0)\} = \\ = (I - P_0)^{-1}\{(1-\tau_0) - \tau_0\Gamma_0(F'(\xi_0) - F'(x_0))\}\Gamma_0 F(x_0).$$

Thus, we get

$$\|\Gamma_1 F(x_1)\| \leqslant \frac{1}{1-\tau_0 a_0}\{|1-\tau_0| + \tau_0^2 a_0\}\,\|\Gamma_0 F(x_0)\| = f(a_0,\tau_0)g(a_0,\tau_0)\,\|\Gamma_0 F(x_0)\|,$$

i.e. (II_1) is true. To prove (III_1) and (IV_1), we use

$$L\,\|\Gamma_1\|\,\|\Gamma_1 F(x_1)\| \leqslant f(a_0,\tau_0)^2 g(a_0,\tau_0) L\,\|\Gamma_0\|\,\|\Gamma_0 F(x_0)\| \leqslant f(a_0,\tau_0)^2 g(a_0,\tau_0)a_0 = a_1$$

and

$$\|x_2 - x_0\| \leqslant \|x_2 - x_1\| + \|x_1 - x_0\| \leqslant \tau_1\|\Gamma_1 F(x_1)\| + \tau_0\|\Gamma_0 F(x_0)\| \leqslant \\ \leqslant [\tau_1 f(a_0,\tau_0)g(a_0,\tau_0) + \tau_0]\,\|\Gamma_0 F(x_0)\| \leqslant \eta\left(1 + \frac{\gamma}{f(a_0,\tau_0)}\right) = \\ = \eta(1 + \Delta\gamma) < \frac{\eta}{1-\Delta\gamma} = R\eta,$$

i.e. $x_2 \in B(\alpha_0, R\eta)$ and this proof holds by induction for all $n \in \mathbb{N}$. Now following an inductive procedure and assuming $x_{n+1} \in \Omega$ the items (I_n)–(IV_n) are proved.

To establish the convergence of $\{x_n\}$, we only have to prove that it is a Cauchy sequence and that $x_{n+1} \in \Omega$. We note that

$$\|\Gamma_n F(x_n)\| \leqslant \|\Gamma_0 F(x_0)\| G_n, \quad G_n = \prod_{k=0}^{n-1} f(a_k,\tau_k)g(a_k,\tau_k).$$

As a consequence of Lemma 3 it follows that

$$G_n \leqslant \Delta^n \prod_{k=0}^{n-1} \gamma^{(1+p)^k} = \Delta^n \gamma^{\frac{(1+p)^n - 1}{p}}.$$

So, from $\Delta < 1$, we deduce that $\lim_{n\to\infty} G_n = 0$. We can now formulate the following result on convergence for the iteration (1).

Theorem 4. *In the conditions indicated for the operator F, let us assume that $\Gamma_0 = (F'(x_0))^{-1} \in L(Y,X)$ exists at some $x_0 \in \Omega$ and $(c_1) - (c_4)$ are satisfied. Suppose that τ_k is given by*

$$\tau_k = \begin{cases} \tau_k^* = \dfrac{-1 + \sqrt{1 + 2by_k}}{by_k}, & for \quad 1 - \tau_k^* > \varepsilon, \\ 1, \quad for \quad 1 - \tau_k^* \leqslant \varepsilon. \end{cases} \tag{18}$$

Then if $\overline{B(x_0, R\eta)} = \{x \in X;\ \|x - x_0\| \leqslant R\eta\} \in \Omega$, the sequence $\{x_n\}$ defined by (1) and starting at x_0 has, at least, R–order $(p+1)$ and converges to a solution x^* of the equation $F(x) = 0$. In this case, the solution x^* and the iterations x_n

belong to $\overline{B(x_0, R\eta)}$ *and* x^* *is the only solution to* $F(x) = 0$ *in* $B\left(x_0, \frac{2}{L\beta} - R\eta\right) \cap \Omega$. *Furthermore, we can give the following error estimates:*

$$\|x^* - x_{n+m}\| \leqslant \frac{\Delta^{n+m}}{1-\Delta\gamma}\gamma^{2^n-1+m}\|\Gamma_0 F(x_0)\|, \quad \gamma = \frac{|1-\tau_0| + a_0\tau_0^2}{(1-\tau_0 a_0)^2}. \tag{19}$$

Proof. We have proved above that the assumption (c_4) is satisfied. Then by virtue of Theorem 3 the damped Newton's iteration converges. Namely the residual norm $\|F(x_k)\|$ decrease as k increase. Then τ_k^* tends to units as k increases. Hence the condition $1 - \tau_k^* \leqslant \varepsilon$ will holds starting at some number $k = m$. Then by virtue of (18) we have $\tau_k \equiv 1$ for all $k \geqslant m$. On the other hand, $x_n \in B(x_0, R\eta)$ for all $n \in \mathbb{N}$, then $x_n \in \Omega$, $n \in \mathbb{N}$. Now we prove that $\{x_n\}$ is a Cauchy sequence. To do this, we consider $n, m \geqslant 1$:

$$\|x_{m+n+e} - x_{m+n}\| \leqslant \sum_{l=0}^{e-1}\|x_{m+n+l+1} - x_{m+n+l}\| \leqslant \sum_{l=0}^{e-1}\|\Gamma_{m+n+l}F(x_{m+n+l})\| \leqslant$$
$$\leqslant \|\Gamma_{m+n}F(x_{m+n})\|\sum_{l=0}^{e-1}\Delta^l\gamma^{2^{m+n}(2^l-1)} \leqslant \|\Gamma_{m+n}F(x_{m+n})\|\sum_{l=0}^{e-1}(\Delta\gamma)^{e-1} =$$
$$= \frac{1-(\Delta\gamma)^e}{1-\Delta\gamma}\|\Gamma_{m+n}F(x_{m+n})\|, \tag{20}$$

in which we have used the well-known inequality $(1+x)^k > 1 + kx$.

According to Lemma 3, we have

$$\|\Gamma_{m+n}F(x_{m+n})\| \leqslant \Delta^n\gamma^{2^m(2^n-1)}\|\Gamma_m F(x_m)\| < \Delta^n\gamma^{2^n-1}\|\Gamma_m F(x_m)\| \tag{21}$$

and

$$\|\Gamma_m F(x_m)\| \leqslant (\Delta\gamma)^m\|\Gamma_0 F(x_0)\|. \tag{22}$$

Substituting (21), (22) into (20), we obtain

$$\|x_{m+n+e} - x_{m+n}\| \leqslant \frac{1-(\Delta\gamma)^e}{1-\delta\gamma}\Delta^{m+n}\gamma^{2^n-1+m}\|\Gamma_0 F(x_0)\|, \tag{23}$$

then $\{x_n\}$ is a Cauchy sequence. Now by letting $e \to \infty$ in (23), we obtain (19).

To prove that $F(x^*) = 0$, notice that $\|\Gamma_n F(x_n)\| \to 0$ by letting $n \to \infty$. As $\|F(x_n)\| \leqslant \|F'(x_n)\|\|\Gamma_n F(x_n)\|$ and $F'(x_n)$ is a bounded sequence, we deduce $\|F(x_n)\| \to 0$ and then $F(x^*) = 0$ by the continuity of F. Now, to show the uniqueness, suppose that $y^* \in B\left(x_0, \frac{2}{L\beta} - R\eta\right) \cap \Omega$ is another solution to $F(x) = 0$. Then

$$0 = F(y^*) - F(x^*) = \int_0^1 F'(x^* + t(y^* - x^*))\mathrm{d}t(y^* - x^*). \tag{24}$$

Also we use the estimation

$$\left\|I - \Gamma_0\int_0^1 F'(x^* + t(y^* - x^*))\mathrm{d}t\right\| \leqslant \|\Gamma_0\|\int_0^1\|F'(x^* + t(y^* - x^*)) - F'(x_0)\|\mathrm{d}t \leqslant$$

$$\leqslant L\beta \int_0^1 \|x^* - t(y^* - x^*) - x_0\| \mathrm{d}t \leqslant L\beta \int_0^1 ((1-t)\|x^* - x_0\| + t\|y^* - x_0\|)\mathrm{d}t <$$

$$< \frac{L\beta}{2}(R\eta + \frac{2}{L\beta} - R\eta) = 1. \quad (25)$$

We see that the operator $\int_0^1 F'(x^* + t(y^* - x^*))\mathrm{d}t$ has an inverse and consequently from (24), we get $y^* = x^*$. □

Remark. In [4] the following choice was proposed

$$\tau_n = \frac{\|F(x_{n-1})\|}{\|F(x_n)\|}\tau_{n-1}, \quad n = 1, 2, \ldots, \quad \tau_0 \approx 0.1, \quad (26)$$

and were the error bounds of kind of (19) derived under the conditions $\|F'(x)^{-1}\| \leqslant \beta$, $\|F''(x)\| \leqslant M$. Here we derive the error bounds (19) without such conditions. There is a closed relationship between (13) and (26). Namely, (13) gives

$$\frac{\tau_n^*}{\tau_{n-1}^*} = \frac{\|F(x_{n-1})\|}{\|F(x_n)\|}A_n, \quad A_n = \frac{-1 + \sqrt{1 + 2b\|F(x_n)\|}}{-1 + \sqrt{1 + 2b\|F(x_{n-1})\|}}.$$

It is easy to show that $A_n \to 1$ at $\max(\|F(x_{n-1})\|, \|F(x_n)\|) \to 0$, i.e. in the limit these two choices coincide. Since $\tau_k \equiv 1$ for all $k \geqslant m$ our iteration (1) can be considered as a Newton iteration starting at $\tilde{a_0} = a_m$, where

$$a_m = L\,\|\Gamma_m\|\,\|\Gamma_m F(x_m)\| \leqslant a_0\gamma^m < \frac{3 - \sqrt{5}}{2}, \quad (27)$$

which obviously, holds for large m, and for any a_0. On the other hand, the Kantorovich semilocal convergence theorem was proved under the condition $a_0 < 1/2$, i.e. the local Newton methods, require sufficiently good initial guesses. Unlike, our iteration (1), as Newton's method is able to compensate for bad initial guesses by virtue of damping strategies. The inequality (27) has shown that the damped Newton's method at first m-stage allows one to extend the convergence domain of the initial guesses.

4. Numerical Results and Discussion

Now let us give some numerical examples that confirm the theoretical results. We use the following test examples $f_i(x) = 0$, $i = 1, \ldots, 5$, which are the same as in [5]:

$f_1(x) = \ln x = 0$, $x^* = 1.0$,
$f_2(x) = e^{x^2+7x-30} - 1 = 0$, $x^* = 3.0$,
$f_3(x) = 1/x - 1 = 0$, $x^* = 1$,
$f_4(x) = x^3 + 4x^2 - 10 = 0$, $x^* = 1.3652300134140968457$,
$f_5(x) = \arctan x = 0$, $x^* = 0.0$.

All these computations are carried out with a double arithmetic precision and the number of iterations n such that $|f(x_n)| < 1.0e - 16$ is tabulated in Table 1. Note that for the last two calculations for (10) and (9) used $a_k = |f''(x_k)(f'(x_k))^{-2}f(x_k)|$. Table 1 also gives the number of iterations of the simple Newton's method, and damped Newton's method with [6]

$$\tau_k = \frac{\delta(0)}{\delta(0) + \delta(1)}, \quad \delta(\theta) = f^2\left(x_k - \theta\frac{f(x_k)}{f'(x_k)}\right). \quad (28)$$

Table 1

The number of iterations at $\epsilon = 1e - 16$

			Functions														
			$f_1(x)$			$f_2(x)$			$f_3(x)$			$f_4(x)$			$f_5(x)$			
x_0			6.4	4.0	2.0	3.5	4.2	5.55	0.9	2.01	2.4	-0.5	0.1	1.0	2.0	1.7	1.4	1.0
Newton			-	-	6	12	22	45	4	-	-	108	10	5	-	-	-	5
τ_k by (28)			11	10	5	13	25	50	5	-	-	-	141	5	5	4	3	4
τ_k by (13)	b	3.0	5	6	6	81	-	-	5	8	12	19	23	10	7	7	6	5
		2.0	8	6	5	68	-	-	5	8	24	480	24	9	7	6	6	5
		1.0	9	8	5	51	-	-	5	9	-	73	25	7	9	7	6	5
		0.1	-	-	6	23	2260	-	5	13	-	40	22	5	-	-	7	5
τ_k by (10)			7	7	6	20	41	86	6	7	7	-	7	6	5	5	5	5
τ_k by (9)			19	20	1	15	26	48	5	4	6	-	8	5	5	5	5	3

References

1. *Zhanlav T., Chuluunbaatar O.* Convergence of the Continuous Analogy of Newton's Method for Solving Nonlinear Equations // Numerical Methods and Programming, Moscow State University. — 2009. — Vol. 10. — Pp. 402–407.
2. *Hernandez M. A., Salanova M. A.* Modification of the Kantorovich Assumptions for Semilocal Convergence of the Chebyshev Method // J. Comput. Appl. Math. — 2000. — Vol. 126. — Pp. 131–143.
3. *Deuflhard P.* Newton Methods for Nonlinear Problems. Affine Invariance and Adoptive Algorithms. — Springer International, 2002.
4. *Zhanlav T., Puzynin I. V.* The Convergence of Iteration Based on a Continuous Analogue of Newton's Method // Comput.Math and Math. Phys. — 1992. — Vol. 32. — Pp. 729–737.
5. *Zhanlav T., Chuluunbaatar O.* High-Order Convergent Iterative Methods for Solving Nonlinear Equations // Bulletin of Peoples' Friendship University of Russia. Series "Mathematics. Information Sciences. Physics". — 2009. — No 3. — Pp. 70–78.
6. *Ermakov V. V., Kalitkin N. N.* The Optimal Step and Regularization for Newton's Method // USSR Comp. Phys. and Math. Phys. — 1981. — No 21(2). — Pp. 235–242.

УДК 519.6

Локальная и полулокальная сходимость непрерывного аналога метода Ньютона

Т. Жанлав*, О. Чулуунбаатар†

* *Кафедра прикладной математики*
Монгольский государственный университет, Монголия
† *Лаборатория информационных технологий*
Объединённый институт ядерных исследований
ул. Жолио-Кюри д.6, Дубна, Московская область, 141980, Россия

В данной работе определена область сходимости непрерывного аналога метода Ньютона и предложен оптимальный выбор параметра τ. Для затухающего метода Ньютона доказана глобальная сходимость и получены оценки погрешности. Стратегии затухания позволяют расширить область начальных параметров, при которых метод сходится. Дано сравнение различных стратегий затухания. Приведённые численные примеры подтверждают теоретические результаты.

Ключевые слова: нелинейные уравнения в банаховых пространствах; затухающий метод ньютона; рекуррентные соотношения; оценка погрешности.