REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
**Rev. Anal. Numér. Théor. Approx., vol. 40 (2011) no. 2, pp. 182–189**
ictp.acad.ro/jnaat

# RELATIONSHIP BETWEEN THE INEXACT NEWTON METHOD AND THE CONTINUOUS ANALOGY OF NEWTON'S METHOD

T. ZHANLAV*, O. CHULUUNBAATAR*† and G. ANKHBAYAR*

**Abstract.** In this paper we propose two new strategies to determine the forcing terms that allow one to improve the efficiency and robustness of the inexact Newton method. The choices are based on the relationship between the inexact Newton method and the continuous analogy of Newton's method. With the new forcing terms, the inexact Newton method is locally Q-superlinearly and quadratically convergent. Numerical results are presented to support the effectiveness of the new forcing terms.

**MSC 2000.** 97N40.

**Keywords.** inexact Newton method, continuous analogy of Newton's method.

## 1. INTRODUCTION

Consider a nonlinear system

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

where $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable nonlinear mapping. Among all kinds of methods for solving the nonlinear equations (1), the Newton method is perhaps the most elementary, popular and important [1, 2, 3, 4, 5, 6]. One of the advantages of the method is its local quadratic convergence. However, its computational cost is expensive, particularly when the size of the problem is very large, because the Newton equations

$$F(x_k) + F'(x_k)s_k = 0 \tag{2}$$

should be solved at each iteration step. To reduce the computational cost of the Newton method, Dembo, Eisenstat and Steihaug [7] proposed an inexact Newton (IN) method,

$$\begin{aligned} F'(x_k)\bar{s}_k &= -F(x_k) + r_k, \\ x_{k+1} &= x_k + \bar{s}_k, \quad k = 0, 1, \ldots, \quad x_0 \in D. \end{aligned} \tag{3}$$

*School of Mathematics and Computer Science, National University of Mongolia, Ulan-Bator, Mongolia, e-mail: tzhanlav@yahoo.com, chuka@jinr.ru, anxaa_w@yahoo.com.

†Joint Institute for Nuclear Research, Dubna, 141980 Moscow region, Russia.

The terms $r_k \in \mathbb{R}^n$ represent the residuals of the approximate solutions $\bar{s}_k$ of the Newton equation (2), i.e., inexactly solve the Newton equation (2) and obtain a step $\bar{s}_k$ such that

$$\|r_k\| = \|F(x_k) + F'(x_k)\bar{s}_k\| \leq \eta_k \|F(x_k)\|, \tag{4}$$

where $\eta_k \in [0, 1)$ is the forcing term. In each iteration step of the IN method, a real number $\eta_k$ should be chosen first, and then an IN step $\bar{s}_k$ is obtained by solving the Newton equations approximately with an efficient iteration solver for systems of linear equations. The forcing terms play an important role both in reducing the residuals of Newton equations and in increasing accuracy of the method. In particular, if $\eta_k = 0$ for all $k$, then the IN method reduces to the Newton method. The IN method, like the Newton method, is locally convergent.

Theorem 1.1. [7]. *Assume that the IN iterates converge to $x^*$. Then the convergence is superlinear if and only if*

$$\|r_k\| = o(\|F(x_k)\|) \text{ as } k \to \infty. \tag{5}$$

Theorem 1.2. [7]. *Given $\eta_k \leq \overline{\eta} < t < 1,\ k = 0, 1, \ldots$ there exists $\varepsilon > 0$ such that for any initial approximation $x_0$ with $\|x_0 - x^*\| \leq \varepsilon$, the sequence of the IN iteration $x_k$ satisfying* (4) *converges to $x^*$. Moreover, the convergence is linear in the sense that*

$$\|x_{k+1} - x^*\|_* \leq t\|x_k - x^*\|_*,\ k = 0, 1, \ldots, \tag{6}$$

*where $\|y\|_* = \|F'(x^*)y\|$.*

Theorem 1.3. [7]. *Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable, $x^* \in \mathbb{R}^n$ such that $F(x^*) = 0$ and $F'(x^*)$ is nonsingular. If the sequence $x_k$ generated by IN iterates converges to $x^*$, then*
*(1) $x_k$ converges to $x^*$ superlinearly when $\eta_k \to 0$;*
*(2) $x_k$ converges to $x^*$ quadratically if $\eta_k = O(\|F(x_k)\|)$ and $F'(x)$ is Lipschitz continuous at $x^*$. In* [8] *the condition*

$$\frac{\|r_k\|}{\|F(x_k)\| + \|\bar{s}_k\|} = O(\|F(x_k)\|) \text{ as } k \to \infty \tag{7}$$

*is proposed which characterizes the quadratic convergence of the IN iterations.*

In [8] it is shown that the IN, inexact perturbed and quasi-Newton methods are equivalent models. At present there are some strategies for choosing the forcing terms [9, 10, 11].

## 2. THE CONTINUOUS ANALOGY OF NEWTON'S METHOD (CANM)

One of the modifications of Newton method is the well known CANM or damped Newton method [1, 4, 5, 6]

$$F'(x_k)v_k = -F(x_k),\ x_{k+1} = x_k + \tau_k v_k,\ k = 0, 1, \ldots, \tag{8}$$

where $\tau_k > 0$ is an iteration parameter. The suitable choice of the parameter allows us to speed-up the convergence and to enlarge the convergence domain. If $\tau_k \equiv 1$ for all $k$, then CANM reduces to Newton method. There exists a closed relationship between the CANM and the IN iterates. Indeed, the CANM (8) can be considered as the IN iteration (3) with the residual

$$r_k = F'(x_k)\bar{s}_k + F(x_k) = (1 - \tau_k)F(x_k). \tag{9}$$

From here we get

$$\|r_k\| = \eta_k \|F(x_k)\| \tag{10}$$

with

$$\eta_k = |1 - \tau_k|. \tag{11}$$

We have a local and semi-local convergence results:

THEOREM 2.1. *There exits $\varepsilon > 0$ such that for any initial approximation $x_0$ with $\|x_0 - x^*\| \le \varepsilon$, the sequence generated by* (8) *with parameter $\tau_k \in (0, 2)$ converges to $x^*$.*

*Proof.* As mentioned above, the CANM is equivalent to IN iterates (9), for which the residual satisfies inequality (4) with forcing term $\eta_k \in [0, 1)$ under condition $\tau_k \in (0, 2)$. Then by theorem 1.2 the sequence $x_k$ generated by (3) or by (8) converges to $x^*$. □

Let $D_0$ be a convex set with $\overline{D}_0 \subseteq D$, and $(F'(x))^{-1}$ exists for all $x \in D_0$.

THEOREM 2.2. *We assume that*

(i) $\|F''(x)\| \le M,\ x \in D_0,$
(ii) $\|F'(x_0)^{-1}\| \le \beta,$
(iii) $\|F'(x_0)^{-1}F(x_0)\| \le \eta,\ a_0 = M\beta\eta,$
(iv) $M\|F'(x_k)^{-1}\|\|F'(x_k)^{-1}F(x_k)\| \le a_k < 2,\ k = 0, 1, \ldots$

*and*

$$0 < \tau_k < \tfrac{-1+\sqrt{1+4a_k}}{a_k}. \tag{12}$$

*Then the sequence $\{x_n\}$ defined by* (8) *and starting at $x_0 \in D_0$ converges to a solution $x^*$ of* (1)*.*

*Proof.* The Taylor expansion gives

$$F(x_{k+1}) = (1 - \tau_k)F(x_k) + \tfrac{F''(\xi_k)}{2}\tau_k^2(F'(x_k)^{-1}F(x_k))^2, \tag{13}$$

where $\xi_k = \theta x_k + (1 - \theta)x_{k+1}$, $\theta \in (0, 1)$. If we use the assumption (iv), then from (13) it follows that

$$\|F(x_{k+1})\| < \left(|1 - \tau_k| + \tfrac{a_k}{2}\tau_k^2\right)\|F(x_k)\| < \|F(x_k)\|$$

under condition (12), i.e., $\{\|F(x_k)\|\}$ is a decreasing sequence provided the iteration parameter $\tau_k$ is chosen in the interval $\left(0, \frac{-1+\sqrt{1+4a_k}}{a_k}\right)$. □

Remark 2.1. It is easy to show that

$$1 < \frac{-1+\sqrt{1+4a_k}}{a_k} < 2, \text{ when } 0 < a_n < 2.$$

This means that $\tau_k \in (0, 2)$. □

Remark 2.2. Assume that the sequence $x_k$ generated by CANM converges to $x^*$. Then

(1) $\{x_k\}$ converges to $x^*$ superlinearly when $\tau \to 1$,

(2) $\{x_k\}$ converges to $x^*$ quadratically if

$$|1-\tau_k| = O(\|F(x_k)\|) \text{ or } |1-\tau_k| = O(\|F(x_{k-1})\|), \tag{14}$$

because of theorem 1.3. □

## 3. SOME CHOICES OF ITERATION PARAMETERS

From (13) it follows that

$$\frac{\|F(x_k)\| - |1-\tau_{k-1}|\|F(x_{k-1})\|}{\|F(x_{k-1})\|} = O(\|F(x_{k-1})\|)$$

and

$$\frac{\|F(x_k)\| - |1-\tau_{k-1}|\|F(x_{k-1})\|}{\|F(x_k)\|} = O(\|F(x_{k-1})\|)$$

Then by Remark 2.2, one can choose $\tau_k$, such that

$$|1-\tau_k| = \frac{|\|F(x_k)\| - |1-\tau_{k-1}|\|\|F(x_{k-1})\|}{\|F(x_{k-1})\|} \tag{16$a$}$$

or

$$|1-\tau_k| = \frac{|\|F(x_k)\| - |1-\tau_{k-1}|\|F(x_{k-1})\||}{\|F(x_k)\|}, \tag{16$b$}$$

which allows us the quadratic convergence of CANM. The formulas (16) can be rewritten in term of the forcing term $\eta_k$ as

$$\eta_k = \frac{|\alpha_k \eta_{k-1} - 1|}{\alpha_k} \tag{17$a$}$$

and respectively

$$\eta_k = |\alpha_k \eta_{k-1} - 1|, \tag{17$b$}$$

where

$$\alpha_k = \frac{\|F(x_{k-1})\|}{\|F(x_k)\|}. \tag{18}$$

Assume that

$$\|F(x_k)\| \le \eta_{k-1}\|F(x_{k-1})\|, \ 0 \le \eta_{k-1} < 1. \tag{19}$$

Then $\alpha_k > 1$ and $\alpha_k \ \eta_{k-1} \ge 1$. As a consequence, the minimum of the possible choices (17) is

$$\eta_n = \frac{\alpha_k \eta_{k-1} - 1}{\alpha_k}.$$

If the inequality (19) is not true, then (17) give us $\eta_n = 1 - \eta_{k-1}\alpha_k$. Thus we have

$$\eta_k = \begin{cases} 1 - \eta_{k-1}\alpha_k, & \text{when } \eta_{k-1}\alpha_k < 1, \\ -\frac{1-\eta_{k-1}\alpha_k}{\alpha_k}, & \text{when } \eta_{k-1}\alpha_k \ge 1. \end{cases} \tag{20}$$

The second choice in (20) allows us to decrease $\eta_k$, i.e., $0 < \eta_k < \eta_{k-1}$, while the first choice in (20) implies that $0 < \eta_k < 1$. In both cases we have

$0 < \eta_k < 1$. According to (4), it is possible that (19) is true and thereby the second choice in (20) allows us to decrease $\eta_k$, i.e., $\eta_k \to 0$ as $k \to \infty$. In terms of $\tau_k$ we have the following choice

$$|1 - \tau_k| = \eta_k. \tag{21}$$

From this it follows that, if $0 < \eta_k < 1$ and $\eta_k \to 0$ as $k \to \infty$, then we have $0 < \tau_k < 2$ and $\tau_k \to 1$ as $k \to \infty$. Thus, the choice of the iteration parameter given by (20), (21) can be used in CANM.

In [6] another choice for CANM was proposed:

$$\tau_n = \frac{2}{1+\sqrt{1+2b\|F(x_n)\|}}. \tag{22}$$

According to (21), the formula (22) in term of $\eta_n$ reads as

$$\eta_n = \frac{\sqrt{1+2b\|F(x_n)\|}-1}{\sqrt{1+2b\|F(x_n)\|}+1}. \tag{23}$$

which can be used in the IN as a forcing term. From (20) we see that the inexact Newton method with forcing term given by (20) is locally Q-superlinearly convergent, while IN with forcing term given by (23) converges quadratically because of $\eta_n = O(\|F(x_n)\|)$ (see theorem 1.3).

## 4. NUMERICAL EXPERIMENTS

In this section we present numerical examples to demonstrate the efficiency of the new strategies to choose forcing terms. We compare the strategies with some known strategies on their numerical behavior.

We have used three test problems that are typical systems of nonlinear equations in literature, with each of its own name and standart initial guess, say $x_s$. The problems are listed as follows [11]:

Problem 4.1 (Generalized function of Rosenbrock)

$$\begin{cases} f_1(x) = -4c(x_2 - x_1^2)x_1 - 2(1 - x_1), \\ f_i(x) = 2c(x_i - x_{i-1}^2) - 4c(x_{i+1} - x_i^2)x_i - 2(1 - x_i), \ i = 2, 3, \ldots, n-1, \\ f_n(x) = 2c(x_n - x_{n-1}^2), \end{cases}$$

with $c = 2$ and $x_s = (1.2, 1.2, ..., 1.2)^T$.

Problem 4.2 (Tridiagonal system)

$$\begin{cases} f_1(x) = 4(x_1 - x_2^2), \\ f_i(x) = 8x_i(x_i^2 - x_{i-1}) - 2(1 - x_i) + 4(x_i - x_{i+1}^2), \ i = 2, 3, \ldots, n-1, \\ f_n(x) = 8x_n(x_n^2 - x_{n-1}) - 2(1 - x_n), \end{cases}$$

with $x_s = (12, 12, ..., 12)^T$.

Table 1. Results for problems 4.1–4.3.

| Choice | | GF of Rosenbrock | | | Tridiagonal system | | | Five-diagonal system | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $x^s$ | $3x^s$ | $-3x^s$ | $x^s$ | $2x^s$ | $-2x^s$ | $x^s$ | $2x^s$ | $-2x^s$ |
| EW1 | NI | 6 | 17 | 18 | 24 | 23 | 27 | 41 | 19 | 16 |
| | GI | 60 | 165 | 122 | 193 | 203 | 251 | 386 | 121 | 125 |
| | CT | 1.91 | 5.66 | 2.19 | 6.21 | 7.15 | 7.81 | 12.88 | 5.11 | 3.17 |
| EW2 | NI | 6 | 16 | 35 | 31 | 40 | 43 | 19 | 47 | 17 |
| | GI | 85 | 110 | 129 | 147 | 206 | 167 | 94 | 156 | 85 |
| | CT | 2.47 | 3.91 | 3.81 | 5.85 | 7.21 | 6.25 | 3.66 | 5.21 | 2.21 |
| APR | NI | 9 | 11 | 21 | 12 | 20 | 46 | 15 | 17 | 16 |
| | GI | 81 | 93 | 116 | 88 | 128 | 216 | 97 | 98 | 95 |
| | CT | 2.66 | 3.13 | 4.94 | 3.25 | 4.35 | 8.28 | 4.03 | 2.58 | 2.02 |
| Str (20) | NI | 10 | 19 | 35 | 33 | 31 | 42 | 33 | 32 | 19 |
| | GI | 81 | 105 | 140 | 141 | 128 | 145 | 128 | 116 | 88 |
| | CT | 2.50 | 3.88 | 5.22 | 5.37 | 5.17 | 6.16 | 6.38 | 3.30 | 3.00 |
| Str (23) | NI | 6 | 14 | 31 | 19 | 28 | 30 | 25 | 17 | 12 |
| | GI | 64 | 87 | 115 | 102 | 135 | 127 | 126 | 91 | 70 |
| | CT | 1.94 | 3.19 | 4.65 | 3.69 | 4.96 | 3.34 | 5.78 | 2.11 | 1.91 |

Problem 4.3 (Five-diagonal system)

$$\begin{cases} f_1(x) = 4(x_1 - x_2^2) + x_2 - x_3^2, \\ f_2(x) = 8x_2(x_2^2 - x_1) - 2(1 - x_2) + 4(x_2 - x_3^2) + x_3 - x_4^2, \\ f_i(x) = 8x_i(x_i^2 - x_{i-1}) - 2(1 - x_i) + 4(x_i - x_{i+1}^2) \\ \qquad + x_{i-1}^2 - x_{i-2} + x_{i+1} - x_{i+2}^2, \qquad i = 3, \dots, n-2, \\ f_{n-1}(x) = 8x_{n-1}(x_{n-1}^2 - x_{n-2}) - 2(1 - x_{n-1}) + 4(x_{n-1} - x_n^2) \\ \qquad + x_{n-2}^2 - x_{n-3}, \\ f_n(x) = 8x_n(x_n^2 - x_{n-1}) - 2(1 - x_n) + x_{n-1}^2 - x_{n-2}, \end{cases}$$

with $x^s = (-2, -2, ..., -2)^T$.

We test the case $n = 100$. Besides the standard initial guess $x^s$, also test other initial guesses such as $x^0 = \pm jx^s$, $j = 2, 3$. It is easy to see that $e = (1, 1, ..., 1)^T$ is a solution to each of the above three problems. For the convenience of reporting the results of different forcing terms, the following notations are used in Table 1:

EW1 : the first strategy given by Eisenstat and Walker [10],

EW2 : the second strategy given by Eisenstat and Walker [10],

APR : the actual reduction and the predicted reduction strategy given in [11] as well as strategies defined by formulas (20) and (23),

Str (20) : strategy given by (20),

Str (23) : strategy given by (23),

NI : denotes the total number of nonlinear iterations,

GI : denotes the total number of GMRES iterations,

CT : denotes $10^2\times$(CPU time).

The norm in our test is the Euclidean norm $\|\cdot\|_2$ and the stoping criteria is

$$\|F(x^k)\| \le \varepsilon = 10^{-12}.$$

Table 2. Iteration comparison between APR and new forcing terms on Rosenbrock system with initial guess $x^s$, tolerance $\varepsilon = 10^{-14}$ and parameter $b = 0.1$ in strategy (23).

| $k$ | Str (23) | | | APR | | |
|---|---|---|---|---|---|---|
| | GI | $\|F(x^k)\|$ | $\eta_k$ | GI | $\|F(x^k)\|$ | $\eta_k$ |
| 0 | 1 | 17.5015 | $5.0000 \cdot 10^{-1}$ | 1 | 17.5015 | $5.0000 \cdot 10^{-1}$ |
| 1 | 3 | 4.4680 | $1.5828 \cdot 10^{-1}$ | 2 | 4.4680 | $2.5000 \cdot 10^{-1}$ |
| 3 | 9 | $4.9646 \cdot 10^{-1}$ | $2.3662 \cdot 10^{-2}$ | 4 | 1.1195 | $1.2500 \cdot 10^{-1}$ |
| 4 | 11 | $1.0066 \cdot 10^{-1}$ | $4.9831 \cdot 10^{-3}$ | 8 | $1.0501 \cdot 10^{-1}$ | $6.250 \cdot 10^{-2}$ |
| 5 | 18 | $5.4711 \cdot 10^{-4}$ | $2.7354 \cdot 10^{-5}$ | 8 | $7.589 \cdot 10^{-2}$ | $6.250 \cdot 10^{-2}$ |
| 6 | 27 | $1.5473 \cdot 10^{-7}$ | $7.7363 \cdot 10^{-9}$ | 9 | $3.6494 \cdot 10^{-2}$ | $3.1250 \cdot 10^{-2}$ |
| 7 | 24 | $1.4223 \cdot 10^{-15}$ | $5.2340 \cdot 10^{-13}$ | 11 | $1.0827 \cdot 10^{-4}$ | $1.5625 \cdot 10^{-2}$ |
| 8 | - | - | - | 12 | $8.2737 \cdot 10^{-7}$ | $7.8125 \cdot 10^{-3}$ |
| 9 | - | - | - | 13 | $2.7747 \cdot 10^{-9}$ | $3.9063 \cdot 10^{-3}$ |
| 10 | - | - | - | 13 | $4.1806 \cdot 10^{-12}$ | $1.9531 \cdot 10^{-3}$ |
| 11 | - | - | - | 13 | $1.1678 \cdot 10^{-14}$ | $9.7656 \cdot 10^{-4}$ |

The initial forcing term $\eta_0 = 0.5$ and additional parameters of $\eta_k$ are chosen following [11], for all the above iterations EW1, EW2, APR, Strategies (20) and (23).

The comparison of GMRES with different forcing terms was made by CPU time. From Table 1 we see that GMRES with the forcing term given by (23) is better than those with other forcing terms.

From Table 2 we see that the forcing term given by (23) decreases more rapidly than the forcing term APR that halved at each iteration. As results, the residual norm $\|F(x^k)\|$ in the first case decreases as rapidly as the forcing term.

Acknowledgement. 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”.

This work was partially sponsored by Foundation for Science and Technology of Ministry of Education, Culture and Science (Mongolia).

## REFERENCES

[1] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, Third edition, Springer, 2002.

[2] E. Zeidler, *Nonlinear functional analysis and its applications, part 1: Fixed-point theorems*, Springer-Verlag, 1986.

[3] L.V. Kantorovich, *Functional analysis and applied mathematics*, Uspekhi Mat. Nauk, **3:6(28)**, pp. 89–185, 1948, (in Russian).

[4] M.K. Gavurin, *Nonlinear functional equations and continuous analogues of iterative methods*, Izv. Vyssh. Uchebn. Zaved. Mat., **5(6)**, pp. 18–31, 1958.

[5] T. Zhanlav and I.V. Puzynin, *The convergence of iterations based on the continuous analogy of Newton's method*, Comput. Maths. Math. Phys., **32**, pp. 846–856, 1992. (in Russian)

[6] T. Zhanlav and O. Chuluunbaatar, *A local and semilocal convergence of the continuous analogy of Newton's method for solving nonlinear equations*, Bulletin of PFUR Series Mathematics. Information Sciences. Physics, **1**, pp. 34–43, 2012.

[7] R.S. Dembo, S.C. Eisenstat and T. Steihang, *Inexact Newton methods*, SIAM J. Numer. Anal., **19**, pp. 400–408, 1982.

[8] E. Cătinaş, *The inexact, inexact perturbed, and quasi-Newton methods are equivalent models*, Math. Comp., **74**, pp. 291–301, 2004.

[9] X.-C. Cai, W.D. Gropp, D.E. Keyes and M.D. Tidriti, *Newton-Krylov-Schwarz methods in CFD*, Proceeding of the international workshop on numerical methods for the Navier-stokes equations, Vieweg, Braunschwieg, pp. 17–30, 1995.

[10] S.C. Eisenstat and M.F. Walker, *Choosing the forcing term in an inexact Newton method*, SIAM J. Sci. Comput., **17**, pp. 16–32, 1996.

[11] H.-B. An, Z.-Y. Mo and X.-P. Liu, *A choice of forcing terms in inexact Newton method*, J. Comput. Appl. Math., **200**, pp. 47–60, 2007.

Received by the editors: January 11, 2012.