MATHEMATICS OF COMPUTATION
Volume 74, Number 249, Pages 291–301
S 0025-5718(04)01646-1
Article electronically published on March 23, 2004

THE INEXACT, INEXACT PERTURBED, AND QUASI-NEWTON METHODS ARE EQUIVALENT MODELS

EMIL CĂTINAŞ

Abstract. A classical model of Newton iterations which takes into account some error terms is given by the quasi-Newton method, which assumes perturbed Jacobians at each step. Its high convergence orders were characterized by Dennis and Moré [Math. Comp. **28** (1974), 549–560]. The inexact Newton method constitutes another such model, since it assumes that at each step the linear systems are only approximately solved; the high convergence orders of these iterations were characterized by Dembo, Eisenstat and Steihaug [SIAM J. Numer. Anal. **19** (1982), 400–408]. We have recently considered the inexact perturbed Newton method [J. Optim. Theory Appl. **108** (2001), 543–570] which assumes that at each step the linear systems are perturbed and then they are only approximately solved; we have characterized the high convergence orders of these iterates in terms of the perturbations and residuals.

In the present paper we show that these three models are in fact equivalent, in the sense that each one may be used to characterize the high convergence orders of the other two. We also study the relationship in the case of linear convergence and we deduce a new convergence result.

1. Introduction

Consider a nonlinear system $F(x) = 0$, where $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. The local convergence of the Newton iterates

$$\begin{aligned} &F'(x_k)\, s_k = -F(x_k), \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D, \end{aligned}$$

to a solution $x^* \in \operatorname{int} D$ is usually studied under the following conditions, which will be implicitly assumed throughout this paper:

- the mapping F is Fréchet differentiable on $\operatorname{int} D$, with F' continuous at x^*;
- the Jacobian $F'(x^*)$ is invertible.

Given an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^n$, these hypotheses assure the existence of a radius $r > 0$ such that the Newton iterates converge superlinearly to x^* for any initial approximation x_0 with $\|x_0 - x^*\| < r$ [27, Th.10.2.2] (see also [33, Th.4.4]).

Received by the editor July 23, 2001 and, in revised form, May 3, 2003.

2000 *Mathematics Subject Classification.* Primary 65H10.

Key words and phrases. Inexact, inexact perturbed and quasi-Newton methods, convergence orders.

This research has been supported by the Romanian Academy under grant GAR 97/1999, and by the National Agency for Science, Technology and Innovation under grant GANSTI 6100 GR/2000.

Recall that an arbitrary sequence $(x_k)_{k\geq 0} \subset \mathbb{R}^n$ is said to converge q-superlinearly (superlinearly, for short) to its limit $\bar{x} \in \mathbb{R}^n$ if

$$(1.1) \qquad Q_1\{x_k\} := \limsup_{k\to\infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} = 0, \quad \text{(assuming } x_k \neq \bar{x} \text{ for all } k \geq k_0\text{)},$$

also denoted by $\|x_{k+1} - \bar{x}\| = o\left(\|x_k - \bar{x}\|\right)$, as $k \to \infty$. For rigorous definitions and results concerning the high convergence orders, we refer the reader to [27, ch.9] and [31] (see also [33, ch.3] and [32]).

However, in many situations, different elements from the Newton iterations are only approximately determined. The first such case considers approximate Jacobians at each step, and leads to the quasi-Newton (QN) iterates

$$\begin{aligned} &B_k s_k = -F\left(x_k\right), \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D. \end{aligned}$$

There exist a number of studies dealing with the approximation of $F'\left(x_k\right)$ by various techniques (see for instance [27], [15], [23], [33] and the references therein). The superlinear convergence of these sequences was characterized by Dennis and Moré. We state here a slightly weaker form of this result.

Theorem 1 ([13]). *Consider a sequence $(B_k)_{k\geq 0} \subset \mathbb{R}^{n\times n}$ of invertible matrices and an initial approximation $x_0 \in D$. If the* QN *iterates converge to x^*, then they converge superlinearly if and only if*

$$(1.2) \qquad \frac{\left\|\left(B_k - F'(x^*)\right)\left(x_{k+1} - x_k\right)\right\|}{\|x_{k+1} - x_k\|} \to 0 \quad \textit{as } k \to \infty.$$

Another practical model of Newton iterates assumes that the linear systems from each step are not solved exactly:

$$\begin{aligned} &F'\left(x_k\right) s_k = -F\left(x_k\right) + r_k, \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D. \end{aligned}$$

The terms $r_k \in \mathbb{R}^n$ represent the residuals of the approximate solutions s_k. Dembo, Eisenstat and Steihaug characterized the superlinear convergence of the inexact Newton (IN) method above.

Theorem 2 ([12]). *Assume that the* IN *iterates converge to x^*. Then the convergence is superlinear if and only if*

$$(1.3) \qquad \|r_k\| = o\big(\|F(x_k)\|\big) \quad \textit{as } k \to \infty.$$

They also obtained the following local convergence result.

Theorem 3 ([12]). *Given $\eta_k \leq \bar{\eta} < t < 1,\ k = 0, 1, \ldots,$ there exists $\varepsilon > 0$ such that for any initial approximation x_0 with $\|x_0 - x^*\| \leq \varepsilon$, the sequence of the* IN *iterates $(x_k)_{k\geq 0}$ satisfying*

$$(1.4) \qquad \|r_k\| \leq \eta_k \left\|F\left(x_k\right)\right\|, \quad k = 0, 1, \ldots,$$

converges to x^*. Moreover, the convergence is linear in the sense that

$$(1.5) \qquad \|x_{k+1} - x^*\|_* \leq t\,\|x_k - x^*\|_*, \quad k = 0, 1, \ldots,$$

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

We have recently considered in [8] the inexact perturbed Newton (IPN) method

$$\begin{aligned} (F'(x_k) + \Delta_k) s_k &= (-F(x_k) + \delta_k) + \hat{r}_k, \\ x_{k+1} &= x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D, \end{aligned}$$

where $\Delta_k \in \mathbb{R}^{n\times n}$ represent perturbations to the Jacobians, $\delta_k \in \mathbb{R}^n$ perturbations to the function evaluations, while $\hat{r}_k \in \mathbb{R}^n$ are the residuals of the approximate solutions s_k of the perturbed linear systems $(F'(x_k) + \Delta_k) s = -F(x_k) + \delta_k$.

We have obtained the following results.[1]

Theorem 4 ([8]). *Assume that the* IPN *iterates are uniquely defined* (*i.e., the perturbations* $(\Delta_k)_{k\geq 0}$ *are such that the matrices* $F'(x_k) + \Delta_k$ *are invertible for* $k = 0, 1, \ldots$) *and converge to* x^*. *Then the convergence is superlinear if and only if*

$$\left\| \Delta_k (F'(x_k) + \Delta_k)^{-1} F(x_k) + \big(I - \Delta_k (F'(x_k) + \Delta_k)^{-1}\big)(\delta_k + \hat{r}_k) \right\| = o\big(\|F(x_k)\|\big),$$

as $k \to \infty$.

Theorem 5 ([8]). *Given* $\eta_k \leq \bar{\eta} < t < 1$, $k = 0, 1, \ldots$, *there exists* $\varepsilon > 0$ *such that if* $\|x_0 - x^*\| \leq \varepsilon$ *and the* IPN *iterates are uniquely defined, satisfying*

$$\left\| \Delta_k (F'(x_k) + \Delta_k)^{-1} F(x_k) + \big(I - \Delta_k (F'(x_k) + \Delta_k)^{-1}\big)(\delta_k + \hat{r}_k) \right\| \leq \eta_k \|F(x_k)\|,$$

where $k = 0, 1, \ldots$, *then these iterates converge to* x^* *at the linear rate*

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

The same conclusion holds if the above condition is replaced by

$$\begin{aligned} &\left\| \Delta_k (F'(x_k) + \Delta_k)^{-1} \right\| \leq q_1 \eta_k \quad \text{and} \\ &\qquad \|\delta_k\| + \|\hat{r}_k\| \leq \tfrac{q_2}{1+q_1} \eta_k \, \|F(x_k)\|, \quad \text{for } k = 0, 1, \ldots, \end{aligned}$$

where $0 < q_2 < 1 - q_1$ *and* $t \in (q_1 + q_2, 1)$.

Remark 1. It is not difficult to prove that, in fact, the above theorem also holds with $q_1 + q_2 = 1$ and $\bar{\eta} < t < 1$ (instead of $0 < q_1 + q_2 < t < 1$).

The aim of this paper is to perform an analysis of the three methods mentioned in order to reveal the natural connection between them. This will allow us to obtain sharp conditions ensuring the local convergence of the inexact perturbed, and quasi-Newton methods.

We shall show first that each one of the inexact, inexact perturbed, and quasi-Newton methods may be used to characterize the high convergence orders of the other two. In this sense, we remark (see [8]) that the proofs of Theorems 4 and 5 were obtained by rewriting the IPN iterations as IN iterations having the residuals

$$r_k = \Delta_k (F'(x_k) + \Delta_k)^{-1} F(x_k) + \left(I - \Delta_k (F'(x_k) + \Delta_k)^{-1}\right)(\delta_k + \hat{r}_k),$$

and then applying Theorems 2 and 3, respectively. We also note that the IN model is a particular instance of the IPN model. These facts show the equivalence of these

[1] We have recently noticed that Wilkinson [36] has previously considered the iterates

$$x_{k+1} = x_k - \big(F'(x_k) + E_k\big)^{-1}\big(F(x_k) + e_k\big) + g_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D,$$

very similar to the IPN ones, while Potra [29], [30] has considered some inexact perturbed secant iterates:

$$x_{k+1} = x_k - \big([x_{k-1}, x_k; F] + E_k\big)^{-1}\big(F(x_k) + e_k\big) + g_k, \qquad k = 1, 2, \ldots, \quad x_0, x_1 \in D,$$

where $[x, y; F]$ represents the first order divided difference of F on x and y. However, these authors have not dealt with the local convergence of these iterates.

two models regarding their linear and superlinear convergence; the same connection appears in fact for the convergence orders $1+p$, $p \in (0,1]$, under supplementary Hölder continuity conditions on F' at x^*.

It remains therefore to analyze the connection between the inexact and the quasi-Newton iterations. This will be done in the following section, while in §3 we shall give a new local linear convergence result and relate some existing ones.

2. Superlinear convergence of inexact and quasi-Newton methods

We begin this section by presenting some auxiliary results.

Walker has shown that the convergence of an arbitrary sequence from $\mathbb{R}^n$ is tightly connected to the convergence of its corrections.

Lemma 1 ([35]). *Consider an arbitrary sequence $(x_k)_{k\geq 0} \subset \mathbb{R}^n$ converging to some element $\bar{x} \in \mathbb{R}^n$. Then the convergence is superlinear if and only if the corrections $(x_{k+1}-x_k)_{k\geq 0}$ converge superlinearly to zero. In case of superlinear convergence it follows that*

$$\lim_{k\to\infty} \frac{\|x_k-\bar{x}\|}{\|x_{k+1}-x_k\|} = 1.$$

The last affirmation of this lemma was known for a longer time (see [13]).

The following result was given by Dembo, Eisenstat and Steihaug.[2]

Lemma 2 ([12]). *Let $\beta = \|F'(x^*)^{-1}\|$ and $\alpha = \max\left\{ \|F'(x^*)\| + \frac{1}{2\beta},\ 2\beta\right\}$. Then there exists $\varepsilon > 0$ such that*

$$\tfrac{1}{\alpha}\|x-x^*\| \leq \|F(x)\| \leq \alpha\|x-x^*\|, \quad \text{when } \|x-x^*\| < \varepsilon.$$

Before stating our results, denote $\Delta_k = B_k - F'(x_k)$; the quasi-Newton iterates are transcribed as

$$\begin{aligned} &(F'(x_k)+\Delta_k)\, s_k = -F(x_k), \\ &x_{k+1} = x_k + s_k, \qquad k=0,1,\ldots, \quad x_0 \in D, \end{aligned}$$

and condition (1.2) characterizing their superlinear convergence becomes

$$\left\|\left(F'(x_k)+\Delta_k - F'(x^*)\right)s_k\right\| = o\left(\|s_k\|\right) \quad \text{as } k\to\infty. \tag{2.1}$$

Now we are able to present the results relating the superlinear convergence of the IN and QN methods. First, we shall regard the QN iterates as IN iterates:

$$F'(x_k)\, s_k = -F(x_k) - \Delta_k s_k, \qquad k=0,1,\ldots, \quad x_0\in D;$$

condition (1.3) characterizing their superlinear convergence becomes

$$\|\Delta_k s_k\| = o\left(\|F(x_k)\|\right) \quad \text{as } k\to\infty. \tag{2.2}$$

The first step is accomplished by the following result.

Theorem 6. *Conditions* (2.1) *and* (2.2) *are equivalent.*

Proof. Some obvious reasons show that (2.1) holds iff

$$\|\Delta_k s_k\| = o\left(\|s_k\|\right) \quad \text{as } k\to\infty.$$

Lemmas 1 and 2 show that the sequences $(x_k-x^*)_{k\geq 0}$, $(s_k)_{k\geq 0}$, and $(F(x_k))_{k\geq 0}$ converge superlinearly to zero only at the same time, which ends the proof. $\square$

[2] A more general form of this lemma was previously obtained by Dennis and Schnabel [15, Lm.4.1.16]; other variants may be found in [23, Lm.5.2.1], [18] and [33, Th.4.2].

Remark 2. As noticed in [14], condition $\Delta_k \to 0$, as $k \to \infty$, is sufficient but not also necessary for (2.2) to hold.

Formulas (2.1) and (2.2) do not explicitly require the invertibility of the perturbed Jacobians at each step. Consequently, one may restate Theorem 1 by demanding the corresponding iterates only to be well defined; i.e., the linear systems $(F'(x_k) + \Delta_k)\, s = -F(x_k)$ to be compatible. In this sense, Theorem 1 (as well as, in fact, Theorem 2) can be retrieved from the following extension of Theorem 4.

Theorem 7. *Assume that the* IPN *iterates are well defined and converge to* x^*. *Then the convergence is superlinear if and only if*

$$\|-\Delta_k s_k + \delta_k + \hat{r}_k\| = o\big(\|F(x_k)\|\big) \quad \text{as } k \to \infty. \tag{2.3}$$

Proof. One may use Theorem 2 in a straightforward manner by writing the IPN iterates as IN iterates with residuals $r_k = -\Delta_k s_k + \delta_k + \hat{r}_k$. ☐

The utility of this result comes out for example when analyzing the local convergence of the two Newton-Krylov methods described below.

Example 1. a) Given a linear system $Ax = b$, $A \in \mathbb{R}^{n\times n}$ nonsingular, $b \in \mathbb{R}^n$, an arbitrary initial approximation x_0 to the solution of this linear system, and denoting $r_0 = b - Ax_0$, the GMBACK solver [21] determines an approximation $x_m^{GB} \in x_0 + \mathcal{K}_m = x_0 + \operatorname{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$ by the minimization problem[3]

$$\left\|\Delta_m^{GB}\right\|_F = \min_{x_m \in x_0 + \mathcal{K}_m} \|\Delta_m\|_F \quad \text{w.r.t. } (A - \Delta_m)x_m = b.$$

As with all the Krylov solvers, the method is advantageous when good approximations are obtained for small subspace dimensions $m \in \{1, \ldots, n\}$, n being supposed large. Depending on the parameters, the problem may have no solution at all, a unique solution x_m^{GB}, or several (at most m) solutions. In the first case the algorithm may be continued either by increasing m or by restarting with a different x_0. In the second case the matrix $A - \Delta_m^{GB}$ is invertible, while in the third case this matrix is not invertible, but the linear system $(A - \Delta_m^{GB})x = b$ is still compatible.

The superlinear convergence of the Newton-GMBACK iterates written as

$$\begin{aligned} &(F'(x_k) - \Delta_{k,m_k}^{GB})s_{k,m_k}^{GB} = -F(x_k), \\ &x_{k+1} = x_k + s_{k,m_k}^{GB}, \quad k = 0, 1, \ldots, \end{aligned}$$

may therefore be characterized by Theorem 7, taking $\Delta_k = -\Delta_{k,m_k}^{GB}$ and $\delta_k = \hat{r}_k = 0$, since if the iterates converge we do not mind if they are not uniquely defined.

Apart from theoretical interest, the use of the QN model for these iterations is worth considering also from the computational standpoint, when the residuals are expensive to evaluate. Indeed, according to Remark 2, condition $\|\Delta_{k,m_k}^{GB}\| \to 0$ as $k \to \infty$ is sufficient for the converging Newton-GMBACK iterates to attain superlinear rate (see also [8]). This provides an alternative in controlling the convergence rate of the above method, since the magnitude of Δ_m^{GB} may be estimated by computing the smallest eigenvalue of a generalized eigenproblem of (low) dimension $m + 1$, during the same process of determining x_m^{GB}.

[3]We shall use $\|\cdot\|_F$ to denote the Frobenius norm of a matrix, $\|Z\|_F = \operatorname{tr}(ZZ^t)^{1/2}$, and $\|\cdot\|_2$ to denote the Euclidean norm from $\mathbb{R}^n$ and its induced operator norm.

b) The other Krylov solver we mention is the MINPERT method [22], which minimizes the joint backward error $[\Delta_m \ \delta_m] \in \mathbb{R}^{n\times(n+1)}$:

$$\left\|[\Delta_m^{MP} \ \delta_m^{MP}]\right\|_F = \min_{x_m\in x_0+\mathcal{K}_m} \left\|[\Delta_m \ \delta_m]\right\|_F \quad \text{w.r.t. } (A-\Delta_m)x_m = b+\delta_m.$$

Theorem 7 is again the choice for characterizing the superlinear rate of the Newton-MINPERT iterations when framed in the perturbed Newton method

$$\begin{aligned}&\big(F'(x_k) - \Delta_{k,m_k}^{MP}\big)s_{k,m_k}^{MP} = -F(x_k) + \delta_{k,m_k}^{MP},\\ &x_{k+1} = x_k + s_{k,m_k}^{MP}, \quad k = 0,1,\ldots,\end{aligned}$$

with the remark that the convergence of these iterates may be characterized by eigenvalues computed in the inner steps (see [8]).

Returning to the analysis of the IN and QN methods, it remains to write the IN as QN iterates. We get

$$\big(F'(x_k) - \tfrac{1}{\|s_k\|_2^2} r_k s_k^t\big)s_k = -F(x_k), \quad k = 0,1,\ldots,$$

condition (2.1) characterizing their superlinear convergence being transcribed as

$$\left\|\big(F'(x_k) - \tfrac{1}{\|s_k\|_2^2} r_k s_k^t - F'(x^*)\big)s_k\right\| = o\big(\|s_k\|\big) \quad \text{as } k\to\infty. \tag{2.4}$$

The equivalence of the QN and IN models is completed by the following result, which again has a straightforward proof.

Theorem 8. *Conditions* (1.3) *and* (2.4) *are equivalent.*

Remark 3. a) In case of superlinear convergence of the IN iterates, the invertibility of the matrices $F'(x_k) - (1/\|s_k\|_2^2)\cdot r_k s_k^t$ is automatically satisfied from a certain step. Indeed, since

$$\left\|r_k s_k^t\right\|_2 = \|r_k\|_2 \|s_k\|_2$$

(see [20, P.2.3.9]), some standard arguments show that the assumptions on the mapping F assure the stated property.

b) Condition (1.3) appeared in a natural way in characterizing the convergence orders of the IN iterates; it is especially suitable for example in the case of the standard Newton-GMRES method, when the norms of the residuals may be cheaply computed at each inner step $m = 1,2,\ldots,\bar{m}$, $\bar{m}\in\{1,\ldots,n\}$, without the cost of forming the actual corrections [34]. However, in some situations this condition may require unnecessarily small residuals (oversolving), as reported in several papers (see, e.g., [18]).

According to Lemmas 1 and 2, the sequences $(x_k - x^*)_{k\ge 0}$, $(x_{k+1}-x_k)_{k\ge 0}$, and $(F(x_k))_{k\ge 0}$ converge superlinearly to zero only at the same time, and therefore one may devise some combinations to use instead of (1.3). We mention the following condition, which characterizes the quadratic convergence of the IN iterations:

$$\frac{\|r_k\|}{\|F(x_k)\| + \|s_k\|} = \mathcal{O}\big(\|F(x_k)\|\big) \quad \text{as } k\to\infty.$$

It emerged naturally by backward error analysis [7], and it clearly shows that the oversolving does not appear when the corrections are sufficiently large. We intend to analyze the controlling of the convergence orders in a future work.

3. Local linear convergence of the IPN method

Morini [26] and Gasparo and Morini [19] have obtained some local linear convergence results for the iterates

$$(F'(x_k) + \Delta_k)s_k = -F(x_k) + \hat{r}_k, \qquad (3.1)$$
$$x_{k+1} = x_k + s_k, \quad k = 0, 1, \ldots.$$

We shall relate them with Theorem 5, but in its special instances for the QN and IN sequences.

We notice first that, similarly to Theorem 7, one may easily prove the following result.

Theorem 9. *Given $\eta_k \le \bar{\eta} < t < 1$, $k = 0, 1, \ldots$, there exists $\varepsilon > 0$ such that for any initial approximation x_0 with $\|x_0 - x^*\| \le \varepsilon$, if the* IPN *iterates $(x_k)_{k\ge 0}$ are well defined and satisfy*

$$\|-\Delta_k s_k + \delta_k + \hat{r}_k\| \le \eta_k \|F(x_k)\|, \quad k = 0, 1, \ldots, \qquad (3.2)$$

then they converge to x^* and obey

$$\|x_{k+1} - x^*\|_* \le t\,\|x_k - x^*\|_*, \quad k = 0, 1, \ldots.$$

Since condition (1.4) is known to be sharp for ensuring the local convergence of the IN iterates, the same property follows for (3.2) concerning the IPN method.

We may also obtain the sharp condition regarding the QN model by taking $\delta_k = \hat{r}_k = 0$ in (3.2). When the QN iterates are uniquely defined, Theorem 5 yields another sufficient condition for convergence (by Remark 1, we took $q_1 = 1$),

$$\left\|\Delta_k(F'(x_k) + \Delta_k)^{-1}\right\| \le \eta_k, \quad k = 0, 1, \ldots, \qquad (3.3)$$

which is not always sharp since it is deduced using an estimate of the form $\|Av\| \le \|A\|\,\|v\|$.

There exist few local linear convergence results for the QN method in the literature, despite the frequent use of this model. A first result was obtained by Ortega and Rheinboldt [27, p.311], who considered a mapping $B : D \to \mathbb{R}^{n\times n}$ and

$$x_{k+1} = x_k - B(x_k)^{-1}F(x_k), \quad k = 0, 1, \ldots.$$

The local linear convergence result for the above sequence is followed by the Ostrowski attraction fixed point theorem, under the strong basic assumption that B is continuous at x^* and, moreover,

$$\rho(I - B(x^*)^{-1}F'(x^*)) < 1,$$

where $\rho(A) = \max\{|\lambda| : \lambda \in \mathbb{C}, \lambda \text{ eigenvalue of } A\}$ denotes the spectral radius of A. In our notation, the above condition becomes

$$\rho\left((F'(x^*) + \Delta(x^*))^{-1}\Delta(x^*)\right) < 1,$$

and is implied, for example, when $\left\|(F'(x^*) + \Delta(x^*))^{-1}\Delta(x^*)\right\| < 1$.

Other results we are aware of can be retrieved from those in [26] and [19], as we shall see in the following. In these papers conditions were assumed of the form $\|P_k\hat{r}_k\| \le \theta_k \|P_k F(x_k)\|$, $k = 0, 1, \ldots$. The invertible matrices P_k arise in the context of preconditioning strategies for solving linear systems. We shall consider here the case $P_k = I$, $k = 0, 1, \ldots$, in order to be able to relate with the previous results. We recall that the condition number of $A \in \mathbb{R}^{n\times n}$ in norm $\|\cdot\|$ is given by $\operatorname{cond}(A) = \|A\| \cdot \|A^{-1}\|$. The mapping F was assumed to belong to the class $\mathcal{F}(\omega, \Lambda^*)$, i.e., obeying the following additional assumptions:

- the set D (on which F is defined) is open;
- the derivative F' is continuous on D;
- the solution x^* is unique in the ball $\bar{B}_\omega(x^*) = \{x \in \mathbb{R}^n : \|x - x^*\| \leq \omega\}$ and $\bar{B}_\omega(x^*) \subseteq D$;
- for all $x, y \in \bar{B}_\omega(x^*)$ one has

$$\|F'(x^*)^{-1}(F'(y) - F'(x))\| \leq \Lambda^* \|y - x\|.$$

These hypotheses implied the existence of $\sigma < \min\{\omega, 1/\Lambda^*\}$ such that $F'(x)$ is invertible for all $x \in \bar{B} = \bar{B}_\sigma(x^*)$ and

$$\|F'(x)^{-1}(F'(y) - F'(x))\| \leq \Lambda \|y - x\|,$$

where $\Lambda = \Lambda^*/(1 - \Lambda^*\sigma)$.

The following result was obtained.

Theorem 10 ([26]). *Let the approximations $F'(x) + \Delta(x)$ to $F'(x)$ be invertible and satisfy for all $x \in \bar{B}$ the properties*

$$\|(F'(x) + \Delta(x))^{-1}\Delta(x)\| \leq \tau_1, \tag{3.4}$$
$$\|(F'(x) + \Delta(x))^{-1}F'(x)\| \leq \tau_2.$$

Let $F \in \mathcal{F}(\omega, \Lambda^*)$, $\|x_0 - x^*\| \leq \delta$, denote $\nu_k = \theta_k \operatorname{cond}(F'(x_k) + \Delta_k)$, with $\nu_k \leq \bar{\nu} < \nu$. If

$$\alpha = \rho(\rho + \tau_1 + \nu\tau_2) + \tau_1 + \nu\tau_2 < 1,$$

where $\rho = \frac{1}{2}\Lambda\delta(1+\nu)\tau_2$, then the sequence $(x_k)_{k\geq 0}$ given by (3.1) and obeying $\|\hat{r}_k\| \leq \theta_k \|F(x_k)\|$, $k = 0, 1, \ldots$, is uniquely defined and converges to x^*, with

$$\|x_{k+1} - x^*\| \leq (\tau_1 + \nu\tau_2)\|x_k - x^*\|, \quad \textit{for all } k \textit{ sufficiently large.}$$

For the case of the quasi-Newton iterates we take $\nu = 0$ in the above theorem, being lead to relation

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq \tau_1, \quad \text{for all } k \text{ sufficiently large,} \tag{3.5}$$

while conditions (3.3) imply the convergence rate (1.5), which can be estimated in norm $\|\cdot\|$ by

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq t \operatorname{cond}(F'(x^*)), \quad k = 0, 1, \ldots. \tag{3.6}$$

Though assumptions (3.4) and (3.3) seem to be somehow similar, the above estimated upper bound appears to be larger than the one in (3.5).

For the IN iterates we take $\tau_1 = 0$ and $\tau_2 = 1$ in Theorem 10, and denoting $\bar{\theta} = \limsup_{k\to\infty} \theta_k$ one gets the following upper bound for the q-factor defined in (1.1):

$$Q_1\{x_k\} = \limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq \bar{\theta} \operatorname{cond}(F'(x^*)).$$

Theorem 3 attracts (3.6), and since $t \in (\bar{\eta}, 1)$ is arbitrary, we arrive at a similar bound: $Q_1\{x_k\} \leq \bar{\eta} \operatorname{cond}(F'(x^*))$. However, the assumptions $\theta_k \operatorname{cond}(F'(x_k)) \leq \bar{\nu} < \nu < \alpha < 1$ in Theorem 10 are obviously stronger than $\eta_k \leq \bar{\eta} < t < 1$.

The results in [19] show somehow similar bounds for $Q_1\{x_k\}$, and again explicit inverse proportionality between the condition numbers of $F'(x_k) + \Delta_k$ and the forcing terms θ_k, but under weaker smoothness assumptions on F' (more exactly, requiring only continuity, and not also the Lipschitz-type condition involving the constant Λ^*).

The following aspects are known to occur in practical applications, when the condition numbers are large. First, linear convergence in norm $\|\cdot\|_*$ does not necessarily attract linear convergence (or linear convergence with sufficiently good rate) in norm $\|\cdot\|$, required in certain problems. Second, the excessively small residuals required to ensure good convergence properties affect the overall efficiency of the method (by additional inner iterates in solving the linear systems).

The results in [26] and [19] show that the use of preconditioners reducing the condition numbers allow larger forcing terms. Another important feature is that the condition number involved is not of the Jacobian at the solution (which is not known) but of the Jacobian (or preconditioned perturbed Jacobian) at the current approximation. The estimators of the condition numbers bring the practical utility of these results.

Conclusions

We have proved that the inexact, the inexact perturbed and the quasi-Newton methods are related in a natural way: the conditions for characterizing their high convergence orders remain invariant under reconsidering the source(s) of the error terms. This approach allowed us to obtain a new convergence result, but it also shows that any property specific to one model of perturbed Newton method may now be transcribed to the other models. For example, the affine invariant conditions for the Newton and the inexact Newton methods (see [16], [17], [37] and [26]) may be considered for the inexact perturbed and the quasi-Newton methods.

Another example of transposing a class of iterations in a different frame can be found in [9], where the successive approximations for smooth iteration mappings were regarded as IN sequences, and the Ostrowski attraction fixed point theorem was refined by characterizing the fast convergent trajectories. This approach is full of potential for further developments, and we should mention, for instance, the obtaining of estimates for the radius of the attraction balls.

Acknowledgment. The author would like to thank an anonymous referee for careful reading of the manuscript and for useful suggestions.

References

[1] I. Argyros, *Advances in the Efficiency of Computational Methods and Applications*, World Scientific Publishing Co., Inc., River Edge, NJ, 2000. MR **2002k:**65003

[2] I. Argyros, *Concerning the convergence of inexact Newton methods*, J. Comp. Appl. Math. **79** (1997), 235–247. MR **98c:**47077

[3] R. E. Bank and D. J. Rose, *Global approximate Newton methods*, Numer. Math. **37** (1981), 279–295. MR **83c:**65127

[4] ———, *Parameter selection for Newton-like methods applicable to nonlinear partial differential equations*, SIAM J. Numer. Anal. **17** (1980), 806–822. MR **81m:**65071

[5] P. N. Brown, *A local convergence theory for combined inexact-Newton/finite-difference projection methods*, SIAM J. Numer. Anal. **24** (1987), 407–434. MR **88d:**65088

[6] E. Cătinaş, *Newton and Newton-Krylov Methods for Solving Nonlinear Systems in $\mathbb{R}^n$*, Ph.D. thesis, "Babeş-Bolyai" University Cluj-Napoca, Romania, 1999.

[7] ———, *A note on the quadratic convergence of the inexact Newton methods*, Rev. Anal. Numér. Théor. Approx. **29** (2000), 129–134. MR **2003i:**65049

[8] ———, *Inexact perturbed Newton methods and applications to a class of Krylov solvers*, J. Optim. Theory Appl. **108** (2001), 543–570. MR **2002c:**65074

[9] ———, *On the superlinear convergence of the successive approximations method*, J. Optim. Theory Appl. **113** (2002), 473–485. MR **2003e:**65086

[10] ———, *Inexact perturbed Newton methods for nonlinear systems with singular Jacobians*, submitted.

[11] ———, *Accurate finite differences in the Newton-Krylov methods*, manuscript.

[12] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. **19** (1982), 400–408. MR **83b:**65056

[13] J. E. Dennis, Jr. and J. J. Moré, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp. **28** (1974), 549–560. MR **49:**8322

[14] ———, *Quasi-Newton methods, motivation and theory*, SIAM Rev. **19** (1977), 46–89. MR **56:**4146

[15] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall Series in Computational Mathematics, Englewood Cliffs, NJ, 1983. MR **85j:**65001

[16] P. Deuflhard and G. Heindl, *Affine invariant convergence theorems for Newton's method and extensions to related methods*, SIAM J. Numer. Anal. **16** (1979), 1–10. MR **80i:**65068

[17] P. Deuflhard and F. A. Potra, *Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal. **29** (1992), 1395–1412. MR **94a:**65031

[18] S. C. Eisenstat and H. F. Walker, *Choosing the forcing terms in an inexact Newton method*, SIAM J. Sci. Comput. **17** (1996), 16–32. MR **96k:**65037

[19] M. Gasparo and B. Morini, *Inexact methods: forcing terms and conditioning*, J. Optim. Theory Appl. **107** (2000), 573–589. MR **2001k:**65089

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third ed., The Johns Hopkins University Press, Baltimore, 1996. MR **97g:**65006

[21] E. M. Kasenally, *GMBACK: a generalised minimum backward error algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput. **16** (1995), 698–719. MR **95m:**65054

[22] E. M. Kasenally and V. Simoncini, *Analysis of a minimum perturbation algorithm for nonsymmetric linear systems*, SIAM J. Numer. Anal. **34** (1997), 48–66. MR **98a:**65040

[23] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995. MR **96d:**65002

[24] Şt. Măruşter, *Numerical Methods for Solving Nonlinear Equations*, Ed. Tehnică, Bucharest, Romania, 1981 (in Romanian).

[25] H. J. Martinez, Z. Parada and R. A. Tapia, *On the characterization of Q-superlinear convergence of quasi-Newton interior-point methods for nonlinear programming*, Bol. Soc. Mat. Mexicana **1** (1995), 137–148. MR **97d:**90090

[26] B. Morini, *Convergence behaviour of inexact Newton methods*, Math. Comp. **68** (1999), 1605–1613. MR **99m:**65114

[27] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970. MR **42:**8686

[28] I. Păvăloiu, *Introduction to the Theory of Approximating the Solutions of Equations*, Ed. Dacia, Cluj-Napoca, Romania, 1976 (in Romanian).

[29] F. A. Potra, *On the numerical stability of a class of iterative processes*, Itinerant Seminar of Functional Equations: Approximation and Convexity, Timişoara, Romania, 1983, 51–54.

[30] ———, *On a class of iterative procedures for solving nonlinear equations in Banach spaces*, Computational Mathematics, Banach Center Publications, PWN-Polish Scientific Publishers, Warsaw, Poland, vol. **13**, 607–621, 1984. MR **87c:**65068

[31] ———, *On Q-order and R-order of convergence*, J. Optim. Theory Appl. **63** (1989), 415–431. MR **91d:**65077

[32] ———, *Q-superlinear convergence of the iterates in primal-dual interior-point methods*, Math. Progr., to appear. MR **2002i:**90120

[33] W. C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, Second ed., SIAM, Philadelphia, PA, 1998. MR **2000c:**65001

[34] Y. Saad and M. H. Schultz, *GMRES: a minimum residual algorithm*, SIAM J. Sci. Statist. Comput. **7** (1986), 856–869. MR **87g:**65064

[35] H. F. Walker, *An approach to continuation using Krylov subspace methods*, in Computational Science in the 21st Century, M.-O. Bristeau, G. Etgen, W. Fitzgibbon, J. L. Lions, J. Periaux and M. F. Wheeler, eds., John Wiley and Sons, Ltd., 1997, 72–82.

[36] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR **32:**1894

[37] T. J. Ypma, *Local convergence of inexact Newton methods*, SIAM J. Numer. Anal. **21** (1984), 583–590. MR **85k:**65043

Romanian Academy, "T. Popoviciu" Institute of Numerical Analysis, P.O. Box 68–1, Cluj-Napoca 3400, Romania

E-mail address: ecatinas@ictp.acad.ro