Cerfacs **Technical Report No. TR/PA/00/56, revised August 2001**

Nicholas I.M. Gould, Dominique Orban, Annick Sartenaer and Philippe L. Toint

Componentwise fast convergence in the solution of full-rank systems of nonlinear equations*

September 14, 2001

Abstract. The asymptotic convergence of parameterized variants of Newton's method for the solution of nonlinear systems of equations is considered. The original system is perturbed by a term involving the variables and a scalar parameter which is driven to zero as the iteration proceeds. The exact local solutions to the perturbed systems then form a differentiable path leading to a solution of the original system, the scalar parameter determining the progress along the path. A path-following algorithm, which involves an inner iteration in which the perturbed systems are approximately solved, is outlined. It is shown that asymptotically, a single linear system is solved per update of the scalar parameter. It turns out that a *componentwise* Q-superlinear rate may be attained, both in the direct error and in the residuals, under standard assumptions, and that this rate may be made arbitrarily close to quadratic. Numerical experiments illustrate the results and we discuss the relationships that this method shares with interior methods in constrained optimization.

Key words. nonlinear systems of equations, path-following methods, componentwise Q-superlinear convergence

1. Introduction

We consider the solution of a nonlinear system of algebraic equations

$$F(x) = 0, \tag{1.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function of the vector of unknowns x. Let $f_i(x)$ be the i-th component function of $F(x)$. We assume that the Jacobian matrix $J(x)$ of $F(x)$, *i.e.*, the matrix whose i-th row is $\nabla_x f_i(x)^T$, is full rank at a root of interest, x^*, and Lipschitz continuous close to this root, that is

N. I. M. Gould: Rutherford Appleton Laboratory, Computational Science and Engineering Department, Chilton, Oxfordshire, England. e-mail: `n.gould@rl.ac.uk`

D. Orban: CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 1, France. e-mail: `Dominique.Orban@cerfacs.fr`

A. Sartenaer: Research Associate of the Belgian National Fund for Scientific Research. Facultés Universitaires Notre-Dame de la Paix, 61, rue de Bruxelles, B-5000 Namur, Belgium. e-mail: `Annick.Sartenaer@fundp.ac.be`

Ph. L. Toint: Facultés Universitaires Notre-Dame de la Paix, 61, rue de Bruxelles, B-5000 Namur, Belgium. e-mail: `Philippe.Toint@fundp.ac.be`

Mathematics Subject Classification (1991): 65K05, 90C26, 90C51

* This work was supported by the MNRT Grant for joint Ph.D. Support.

A1. $J(x^*)$ is non-singular, and
A2. $J(x)$ is Lipschitz continuous in some open neighbourhood of x^*.

It is well-known that in this case, if x^* is a limit point of the Newton iteration

$$x^{k+1} = x^k - J^{-1}(x^k)F(x^k) \tag{1.2}$$

started from some x^1 sufficiently close to x^*, the complete sequence $\{x^k\}$ converges *in norm* at a Q-quadratic rate, that is to say that there is a constant $\kappa \geq 0$ and integer k_q for which

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^2} \leq \kappa \tag{1.3}$$

for all $k \geq k_q$ (see, for example, [8,16]), where $\|\cdot\|$ denotes any norm on $\mathbb{R}^n$.

While this result is entirely satisfactory if all we are concerned about is a fast normwise convergence rate, it is less useful if our interest is in methods that attain a fast *componentwise* rate of convergence. Let us denote by z_i the i-th component of any vector $z \in \mathbb{R}^n$. It follows immediately from (1.3) that there exists $\kappa_\infty > 0$ such that

$$|x_i^{k+1} - x_i^*| \leq \|x^{k+1} - x^*\|_\infty \leq \kappa_\infty \; \|x^k - x^*\|_\infty^2$$

for every component x_i^k of x^k, and thus that $\{x_i^k\}$ converges R-quadratically. However this is not sufficient to guarantee that each (or any) component converges Q-quadratically, and indeed this may not be the case, as we show in the following example.

Example 1.1. For $1 \leq \ell \leq n$, let e_ℓ be the ℓ-th column of the identity matrix of order n. Consider the nonlinear system of equations

$$x_i^2 + x_{i+1} = 0, \quad \text{for} \quad 1 \leq i \leq n-1, \quad \text{and} \quad x_n^2 + x_1 = 0, \tag{1.4}$$

and suppose that the iterate x^k is equal to $x_\ell \, e_\ell$, ($|x_\ell| < 1$) for some index $1 \leq \ell \leq n$. It is possible to show (see the Appendix) that a Newton step from x^k results in the improved estimate

$$x^{k+1} = \begin{cases} x_\ell^2 \, e_{\ell+1} & \text{for } 1 \leq \ell \leq n-1 \\ x_n^2 \, e_1 & \text{for } \ell = n, \end{cases} \tag{1.5}$$

of the root $x^* = 0$. Hence, if the Newton iteration is started at $x^1 = x_\ell \, e_\ell$ for some $1 \leq \ell \leq n$, the resulting sequence of iterates has the property that each component is non-zero precisely once in every cycle of n iterations. Thus, although the norm of the iterates converges Q-quadratically, this is not true of individual components—they do, however, converge at an n-step Q-2^n rate. □

In this paper, we are interested in Newton-like methods that produce iterates whose components converge at the same Q-rate since then we can be sure that, asymptotically, each iteration results in an improvement in any particular component(s). This is particularly important in large-scale computations, where every iteration is expensive, and when only a subset of the variables are of interest. Our aim then is not to spend time on unnecessary computation, which might happen if the components of interest do not improve over the course of a number of iterations. For example, if $x = (x_{\mathrm{P}}, y)$ where x_{P} is the vector of primal variables in a basic SQP Newton method for constrained optimization, and where y is the vector of associated Lagrange multipliers, it is known that the pair (x_{P}, y) asymptotically converges at a Q-quadratic rate under reasonable assumptions, while often it is only the primal variables x_{P} that are of interest, and these may not inherit the overall rate of the combined pair. Similarly, in some control problems, the control u may be of greater interest than the state x, or the convergence properties of a particular combination of some components of u with some components of x is important, and this combination might not possess the same convergence properties as the complete solution pair (x, u).

Another example where convergence does not take place componentwise is in the solution to the Navier-Stokes equations when studying turbulent fluid flow around a complex geometry, say, an aircraft. In the far field, the flow settles after a few iterations and fast convergence occurs, whereas in the vicinity of the aircraft, convergence may be much slower and variables may be of widely different magnitude, because of the effects of the complexity of the flowfield and of the local mesh refinement. It may thus be desirable to encourage faster convergence in certain, predefined, regions of the mesh—such as boundary layers, wakes, recirculation regions and shock waves—and thus to have control over the convergence of certain, predefined, variables.

Recently, it has been shown that the iterates generated by certain interior-point methods for constrained optimization converge at a componentwise fast (almost quadratic) Q-rate [12]. The methods we consider in the present paper are a variation, and generalization, on this theme.

Homotopy methods [24,25,35], designed to be globally convergent—*i.e.*, convergent from a remote starting point—with probability one, perturb (1.1) in such a way that solutions to the perturbed problem converge to a solution of (1.1) as the perturbation gradually vanishes. This perturbation usually depends on a scalar parameter $\mu > 0$ which iteratively decreases to eventually equal zero, at which point the perturbed problem coincides with (1.1). Some probability-one homotopy convergence theorems for both unconstrained and constrained optimization are derived in [28]. We cast problem (1.1) into a path-following framework by allowing complete control of the rate at which the perturbation vanishes. This is at variance with homotopy methods in which global convergence partially relies on a “search” in the perturbation parameter space. This distinction, and the path-following framework, are further described in the next sections.

We consider the case of regular solutions, *i.e.*, Assumption **A1** is satisfied. This assumption, in conjunction with the others and the implicit-function theorem, guarantees that a locally unique parameterized path emanates from x^* and can therefore be traced numerically. Were **A1** not satisfied, local uniqueness would no longer take place and we would be confronted with a *bifurcation* problem. Bifurcation problems in the framework of nonlinear parametric programming problems are described in [14,17,19,21]. An occurrence of such a situation in a practical application is presented in [18] in the context of analysis and design of elastic/plastic structures, and where the failure of **A1** is due to the Mangasarian and Fromovitz constraint qualification being satisfied in place of the linear independence constraint qualification.

Rheinboldt [19] examines and compares a class of *continuation* methods proposed by a number of authors, based on augmented systems of the form

$$\begin{bmatrix} F(x) \\ \phi(x,\lambda) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where we often have $\phi : \mathbb{R}^{n+p} \to \mathbb{R}$, and $\lambda \in \mathbb{R}^p$ is a suitable vector of parameters. These methods are designed to numerically trace continuously differentiable manifolds of regular solutions, when this system is first turned into an initial-value problem. His analysis is more general, fits in the framework of differential geometry and encompasses a large class of methods, both in finite and infinite-dimensional cases, including that presented here.

In this paper, we are only interested in the asymptotic behaviour of a sequence of iterates converging to a solution of (1.1). More precisely, we implicitly assume that some globally-convergent algorithm—based, for instance, on line-searches, trust regions or homotopies—has been successful in providing a point sufficiently close to a solution of (1.1), from which a local procedure is started. This local procedure is the main concern of the analysis presented in the next sections and is a phase sometimes referred to as the "end game" in the context of homotopy methods [20], as opposed to the strategy used to properly initiate the globalization—the "opening game". Hereafter, this end game is a Newton-like method resulting in componentwise Q-superlinear convergence as outlined above. As it turns out, the same componentwise Q-superlinear rate also applies to the vector of residuals.

In what follows, we use the order notation for conciseness and clarity. If $\{\alpha^k\}$ and $\{\beta^k\}$ are two sequences of positive numbers converging to zero, we say that $\alpha^k = o(\beta^k)$ if $\lim_{k\to\infty} \alpha^k/\beta^k = 0$ and we say that $\alpha^k = O(\beta^k)$ if there exists a constant $\kappa > 0$ such that $\alpha^k \le \kappa\beta^k$ for all sufficiently large k. We say that $\alpha^k = \Omega(\beta^k)$ if $\beta^k = O(\alpha^k)$, and we write $\alpha^k = \Theta(\beta^k)$ if $\alpha^k = O(\beta^k)$ and $\alpha^k = \Omega(\beta^k)$. We also say that $\alpha^k = \Theta(1)$ if there exist constants $\kappa^{\rm L}, \kappa^{\rm U} > 0$ such that $\kappa^{\rm L} \le \alpha^k \le \kappa^{\rm U}$ for all sufficiently large k.

If $\{z^k\}$ is a sequence in $\mathbb{R}^n$, and if $\mathcal{K}$ is an index set, the subsequence of $\{z^k\}$ indexed by $\mathcal{K}$ will be denoted by $\{z^k\}_{k\in\mathcal{K}}$.

We use the following terminology. A method which is convergent from a remote starting point is said to be *globally* convergent. Otherwise, it is said to be *locally* convergent.

The paper is organized as follows. In §2, we recall the methodology of the Newton homotopy, describe the alternative approach we adopt to solve (1.1) and outline a generic algorithm. We then examine the fast asymptotic properties of this algorithm in §3 and give an explicit componentwise Q-rate of convergence. We discuss the method and its relations with interior-point methods for constrained optimization in §4 and comment on our assumptions and other related issues in §5. Finally, we illustrate the results numerically in §6 and conclude in §7.

2. A parameterized variant of Newton's method

Suppose $h : \mathbb{R}_+ \to \mathbb{R}^n$ is a continuously differentiable vector-valued function of the positive real parameter μ such that $h(0) = 0$. A common way of solving problem (1.1) is instead to consider solving a sequence of systems

$$F(x) = h(\mu), \tag{2.1}$$

as μ is progressively driven to zero (see for instance [13,25,35]). At iteration k, a typical method sets the value of μ to μ^k, generates a starting point $x_{\rm S}^k$, possibly from the solution x^k corresponding to μ^{k-1}, and uses a globally convergent method—which we refer to as the *inner iteration*—to find x^{k+1} satisfying

$$\|F(x^{k+1}) - h(\mu^k)\| \le \epsilon^k \tag{2.2}$$

for some small $\epsilon^k > 0$. The sequences $\{\mu^k\}$ and $\{\epsilon^k\}$ are chosen to converge to zero. The procedure (2.1)–(2.2) is usually referred to as the *Newton homotopy* [35]. Our major concern in what follows will be the choice of the starting point $x_{\rm S}^k$ rather than the details of the globally convergent procedure—we simply mention that a trust-region variant of Newton's method is suitable, provided that it ensures convergence to a solution of (1.1) [5]. The most obvious starting point is simply the result of the previous inner iteration, x^k. However, we intend to show that a superior point may be obtained as an *approximation* to the Newton point for the next parameterized subproblem

$$F(x) = h(\mu^k)$$

from x^k. Applying an approximate Newton correction, we set

$$x_{\rm S}^k = x^k + s^k, \tag{2.3}$$

where s^k satisfies

$$J(x^k)s^k = h(\mu^k) - F(x^k) + r^k \quad \text{and} \quad \|r^k\| \le \eta^k, \tag{2.4}$$

for some residual vector r^k and a suitable sequence $\{\eta^k\}$ of positive values whose limit is zero. We might then examine the possible fast componentwise convergence of the sequence $\{x^k\}$ to a solution of (1.1), under appropriate assumptions on h, when the starting point (2.3) is used at each iteration.

In the remainder of this paper, our intent is to analyze methods similar to (2.1), but to allow more generality on the perturbation function h, and to permit the use of approximate Jacobian matrices $G(x^k) \approx J(x^k)$. More precisely, consider the function $h : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$ of (x, μ) and suppose it satisfies the following properties

A3. The derivatives of $h(x, \mu)$ with respect to x and μ exist and are Lipschitz continuous over a neighbourhood of $(x^*, 0)$, and

A4. $h(x^*, 0) = 0$, $h(x, \mu)$ is nonzero for all $x \ne x^*$ sufficiently close to x^* and all sufficiently small $\mu > 0$, and satisfies $\nabla_x h(x^*, 0) = 0$.

We shall henceforth consider the parameterized system

$$F(x) = h(x, \mu) \tag{2.5}$$

as a function of μ as μ decreases to zero. Since (2.1) is a special case of (2.5), we only consider and refer to (2.5) in the remainder of this paper.

Assumption **A4** will be used several times in the sequel, in relation with the implicit-function theorem, noting that when $(x, \mu) = (x^*, 0)$, the system (2.5) reduces to $F(x^*) = 0$ because of the first part of **A4**, and that its Jacobian is $J(x^*)$ because of the last part of **A4**. In the remainder of this paper, whenever **A4** is used in conjunction with the implicit-function theorem, we simply refer to **A4** rather than to its first and last parts. We further comment Assumption **A4** in §5.

Let $x(\mu)$ be the root of (2.5) for which $\lim_{\mu\to 0} x(\mu) = x^*$. From the implicit-function theorem, such a root exists for sufficiently small μ by virtue of **A1**–**A4** and may be expressed as a continuously differentiable function of μ. This claim will be more formally stated in Lemma 3.1.

We shall refer to the method outlined above as a *path-following* method and the simple idea behind it is to track the path of exact solutions $x(\mu)$ of (2.5) to x^*, using a convergent iterative method to solve (2.5) with fixed μ. Note that this procedure differs from globally-convergent homotopy methods, which attempt to solve (2.5) for the joint variables (x, μ). An overview of homotopy methods may be found in [13,16,35]. Deeper motivation appears in [24] together with numerous references to practical situations, basic work and sparse and dense linear algebra concerns in the course of the solution to (2.5).

In our more general context, we still use (2.3) where s^k is now required to satisfy

$$G(x^k)s^k = h(x^k, \mu^k) - F(x^k) + r^k \quad \text{and} \quad \|r^k\| \le \eta^k, \tag{2.6}$$

for some residual vector r^k and a suitable sequence $\{\eta^k\}$ of positive values whose limit is zero. In this equation, the matrix $G(x^k)$ is either $J(x^k)$ or is a suitable approximation thereof, and is required to satisfy the following condition

A5. There exists a constant $\kappa_{\rm G} > 0$ such that for all x^k sufficiently close to x^*,

$$\left\| \left[G(x^k) - J(x^k) \right] s^k \right\| \le \kappa_{\rm G}\, \mu^{k-1}\, \|s^k\|, \tag{2.7}$$

where s^k is defined by (2.6).

Condition (2.7) is similar to the Dennis-Moré condition [8, Theorem 8.2.4], often used in quasi-Newton methods, but is stronger in that it makes the accuracy of the Jacobian approximation along the normalized step explicitly dependent on μ^{k-1}, that is, as we show in Lemma 3.3, on the distance to the solution. It is indeed easy to see, using **A1**–**A2**, that condition (2.7) implies the Dennis-Moré condition if we assume, as will be the case later in this paper, that a subsequence $\{x^k\}_{k\in\mathcal{K}}$ of $\{x^k\}$ converges to x^*. This and other aspects of Assumption **A5** are further commented in §5.

Since no derivative of h with respect to x appears in (2.6), a possible interpretation of (2.3) is as an approximation to the Newton point for the next parameterized system $F(x) = h(x, \mu^k)$ starting from x^k, where the derivatives of h with respect to x, or any approximation thereof, have been discarded from $G(x)$. This is justified if, for instance, $\nabla_x h(x, \mu)$ is expensive to compute or is perhaps unavailable. An alternative interpretation is as an approximation to the Newton point for $F(x) = h(x^k, \mu^k)$, starting from x^k. However, the absence of $\nabla_x h(x, \mu)$ in (2.6) is a key point to the results to come and is strongly related to the asymptotic behaviour of interior methods in optimization, as we will further comment in §4.

Let $\{\mu^k\}$, $\{\epsilon^k\}$ and $\{\eta^k\}$ be strictly decreasing sequences of parameters whose limit are zero, and consider the iteration given in Algorithm 2.1 for solving (1.1) by way of (2.5).

Algorithm 2.1: Parameterized root-finding procedure

Given $x^1 \in \mathbb{R}^n$, $\mu^0 > 0$ and $\epsilon^0 > \|F(x^1) - h(x^1, \mu^0)\|$, set $k = 1$ and perform the following steps:

Step 1 [*Update*]. Compute the new parameters μ^k, ϵ^k and η^k in such a way that they form strictly decreasing sequences converging to zero.

Step 2 [*Starting point*]. Generate the starting point x_{S}^k from x^k using (2.3) and (2.6).

Step 3 [*Inner iteration*]. Starting from x_{S}^k, find x^{k+1} satisfying

$$\|F(x^{k+1}) - h(x^{k+1}, \mu^k)\| \le \epsilon^k, \tag{2.8}$$

stopping at the first inner iterate for which this condition is satisfied.

Step 4 [*Loop*]. Set $k \leftarrow k + 1$. Return to Step 1.

The rationale behind updating the parameter μ at Step 1 of Algorithm 2.1 will emerge in Theorem 3.1 and Theorem 3.2 from which we see that the stopping condition of the overall outer iteration may be conveniently based on the magnitude of the parameter μ.

For future reference, we write the stopping rule (2.8) at iteration $k-1$, that is

$$\|F(x^k) - h(x^k, \mu^{k-1})\| \le \epsilon^{k-1}. \tag{2.9}$$

In order to find a point satisfying (2.8), we may for instance apply Newton's method to (2.5). However, in what follows, we will not be concerned with the exact mechanism of the inner iteration, merely than it starts from x_{S}^k and returns as soon as a point x^{k+1} satisfying (2.8) is discovered.

Notice that the step s^k computed in Step 2 of Algorithm 2.1 might have been rejected by the mechanism of the inner iteration, perhaps by a linesearch or trust-region acceptance rule, had we used (2.6) as inner iteration procedure, and started from x^k. By computing such a step outside the inner iteration, we may choose to ignore such (often overly cautious) restrictions. Notice also that the flexibility implied by (2.6) may be exploited by allowing early termination of iterative methods (such as Krylov subspace methods) applied to the system

$$G(x^k)s = h(x^k, \mu^k) - F(x^k)$$

in the unknown s.

Our aim is thus to propose suitable values for the parameters $\{\mu^k\}$ and the accuracy tolerances $\{\epsilon^k\}$ and $\{\eta^k\}$, and to examine the effect such choices have on the rate of convergence of the overall iteration.

3. Fast local convergence

In this section, we propose a set of updating rules for Step 1 of Algorithm 2.1. While it is possible to derive more abstract conditions on general parameters $\{\mu^k\}$ and accuracy tolerances $\{\epsilon^k\}$ and $\{\eta^k\}$ that permit fast convergence, we restrict ourselves here to the update

$$\mu^{k+1} = \tau_\mu^k \; (\mu^k)^{\theta_\mu^k} \tag{3.1}$$

and accuracy tolerances

$$\epsilon^k = \tau_\epsilon^k \; (\mu^k)^{\theta_\epsilon^k} \quad \text{and} \tag{3.2}$$

$$\eta^k = \tau_\eta^k \; (\mu^k)^{\theta_\eta^k}, \tag{3.3}$$

for appropriate scale factors

$$0 < \tau_{\min} \le \tau_\mu^k, \tau_\epsilon^k, \tau_\eta^k \le \tau_{\max}, \tag{3.4}$$

and exponents

$$\theta_\mu^k, \theta_\epsilon^k, \theta_\eta^k \ge 1, \tag{3.5}$$

where $\tau_{\min}$ and $\tau_{\max}$ are given constants. Our aim is to show that there are suitable values of these exponents (3.5) that allow fast componentwise convergence. As asymptotic properties are sought, boundedness of the scale factors (3.4) is sufficient, and we will not be concerned by their exact values.

Our first result considers the behaviour of $x(\mu)$ near x^*.

Lemma 3.1. *Suppose that* **A1**–**A4** *hold. Then*

$$x(\mu) = x^* + \mu x' + O(\mu^2), \tag{3.6}$$

for all sufficiently small μ, *where*

$$x' = J(x^*)^{-1} \nabla_\mu h(x^*, 0). \tag{3.7}$$

Proof. The definition of $x(\mu)$ and (2.5) give that

$$F(x(\mu)) = h(x(\mu), \mu).$$

By the implicit-function theorem and **A1**–**A4**, there exists a vector x' such that $x(\mu)$ may be expressed in the form

$$x(\mu) = x^* + \mu x' + \varepsilon(\mu) \quad \text{where} \quad \|\varepsilon(\mu)\| = o(\mu). \tag{3.8}$$

Note that (3.8) immediately implies that

$$x(\mu) - x^* = O(\mu). \tag{3.9}$$

A Taylor expansion of $F(x(\mu))$ involving (3.8) and (3.9) may be combined with **A2** and the identity $F(x^*) = 0$ to give

$$\begin{aligned} F(x(\mu)) &= F(x^*) + J(x^*)(\mu x' + \varepsilon(\mu)) + O(\mu^2) \\ &= \mu J(x^*) x' + J(x^*)\varepsilon(\mu) + O(\mu^2). \end{aligned} \tag{3.10}$$

On the other hand, a second-order expansion of $h(x(\mu), \mu)$ about $(x^*, 0)$ gives

$$\begin{aligned} h(x(\mu), \mu) &= h(x^*, 0) + \nabla_x h(x^*, 0)(x(\mu) - x^*) + \nabla_\mu h(x^*, 0)\mu + O(\mu^2) \\ &= \nabla_\mu h(x^*, 0)\mu + O(\mu^2), \end{aligned} \tag{3.11}$$

where we used **A3** and **A4** and (3.9).

A matched asymptotic expansion between (3.10) and (3.11) yields (3.7) and, using **A1**, the estimate $\|\varepsilon(\mu)\| = O(\mu^2)$. Such an estimate combines with (3.8) to give (3.6). ∎

In the remainder of this section, we consider a subsequence $\{x^k\}_{k\in\mathcal{K}}$ converging to x^*. Unless otherwise specified, the following results involve only $k \in \mathcal{K}$. Our main theorems will then be concerned with the whole sequence $\{x^k\}$. Before proceeding, we need the following technical result, inspired by [12, Lemma 3.1] and [33, Lemma 3.1], that will provide a bound, given in Lemma 3.3, on the distance between x^k satisfying (2.9) and x^* for all sufficiently large $k \in \mathcal{K}$.

Lemma 3.2. *Suppose that* **A1**–**A4** *hold and that the vector* $x(\mu, u) \in \mathbb{R}^n$ *is defined implicitly as the solution to the nonlinear system*

$$\Psi(x, \mu, u) \stackrel{\text{def}}{=} F(x) - h(x, \mu) - u = 0, \tag{3.12}$$

for $\mu \in \mathbb{R}_+$ *and* $u \in \mathbb{R}^n$. *Then there exists a constant* $\delta > 0$ *such that the following statements hold.*

(i) The derivatives of $x(\mu, u)$ *with respect to* μ *and* u *exist and are Lipschitz-continuous functions over the neighbourhood*

$$\mathcal{N}_\delta \stackrel{\text{def}}{=} \{(\mu, u) \in \mathbb{R}_+ \times \mathbb{R}^n \mid \mu + \|u\| \le \delta\}.$$

(ii) If $(\mu_1, u_1), (\mu_2, u_2) \in \mathcal{N}_\delta$, *we have*

$$\|x(\mu_1, u_1) - x(\mu_2, u_2)\| = \Theta\left(|\mu_1 - \mu_2| + \|u_1 - u_2\|\right). \tag{3.13}$$

Proof. Since **A1** and **A4** imply that the Jacobian matrix $\nabla_x \Psi(x^*, 0, 0) = J(x^*)$ is non-singular, since $\Psi(x^*, 0, 0) = 0$ and since **A2** and **A3** imply that $\Psi(x, \mu, u)$ has Lipschitz-continuous partial derivatives in a neighbourhood of $(x^*, 0, 0)$, (i) follows from the implicit-function theorem. Moreover, still from the implicit-function theorem, the Jacobian of the function $x(\mu, u)$ found in (i) is given by

$$\nabla_{\mu,u} x(\mu, u) = \left[J(x(\mu, u)) - \nabla_x h(x(\mu, u), \mu)\right]^{-1} \left[\nabla_\mu h(x(\mu, u), \mu) \;\, I_n\right],$$

where I_n is the identity matrix of size n. If $(\mu_1, u_1), (\mu_2, u_2) \in \mathcal{N}_\delta$, a Taylor expansion thus yields

$$x(\mu_1, u_1) - x(\mu_2, u_2) = \nabla_{\mu,u} x(\mu_2, u_2) \begin{pmatrix} \mu_1 - \mu_2 \\ u_1 - u_2 \end{pmatrix} + O\left(\left\| \begin{pmatrix} \mu_1 - \mu_2 \\ u_1 - u_2 \end{pmatrix} \right\|^2 \right). \tag{3.14}$$

Notice that if (μ, u) is close to $(0, 0)$, the Jacobian $\nabla_{\mu,u} x(\mu, u)$ remains uniformly bounded and has full-rank in a neighbourhood of $(0, 0)$ because of **A1**–**A4**. Its smallest singular value thus remains bounded away from zero, and the first term on the right-hand side of (3.14) is therefore dominant. These two last facts combine with (3.14) to give (3.13). ■

Note that because of Lemma 3.1, the exact solution $x(\mu)$ to (2.5) is unique for all sufficiently small μ. Therefore, we have that $x(\mu, 0) = x(\mu)$, with $x(\mu, 0)$ as defined in Lemma 3.2.

Lemma 3.2 now allows us to derive the following bound on the distance between x^k and x^*.

Lemma 3.3. *Suppose that* **A1**–**A4** *hold and that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then, for sufficiently large* $k \in \mathcal{K}$,

$$\|x^k - x^*\| = O(\mu^{k-1}). \tag{3.15}$$

Proof. Inequality (2.9) may be written as

$$F(x^k) - h(x^k, \mu^{k-1}) = u^k \quad \text{with} \quad \|u^k\| \leq \epsilon^{k-1}.$$

For sufficiently large $k \in \mathcal{K}$, we have from the fact that $\mu^k \to 0$ and (3.2) that (μ^{k-1}, u^k) lies in the neighbourhood $\mathcal{N}_\delta$ defined in Lemma 3.2. We may therefore apply (3.13) with the parameters (μ^{k-1}, u^k) and $(0,0)$ to obtain

$$\|x^k - x^*\| = \Theta(\mu^{k-1} + \|u^k\|) = O(\mu^{k-1} + \epsilon^{k-1}) = O(\mu^{k-1}),$$

where we used the identities $x(\mu^{k-1}, u^k) = x^k$, $x(0,0) = x^*$ and (3.2). ■

The following technical result relates the norm of the step s^k computed from (2.6) to the norm of $G(x^k)s^k$.

Lemma 3.4. *Suppose that* **A1**–**A5** *hold, that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5), and that the step* s^k *is computed from (2.6). Then, for sufficiently large* $k \in \mathcal{K}$,

$$\|G(x^k)s^k\| = \Theta(\|s^k\|). \tag{3.16}$$

Proof. From **A2**, there exists a $\kappa_{\rm J} > 0$ such that $\|J(x^k)\| \leq \kappa_{\rm J}$ for all sufficiently large $k \in \mathcal{K}$, and using **A5**, we have

$$\begin{aligned}
\|G(x^k)s^k\| &= \left\|J(x^k)s^k + \left[G(x^k) - J(x^k)\right] s^k\right\| \\
&\leq \|J(x^k)\|\,\|s^k\| + \kappa_{\rm G}\mu^{k-1}\|s^k\| \\
&\leq \left(\kappa_{\rm J} + \kappa_{\rm G}\mu^{k-1}\right)\|s^k\| \\
&= O(\|s^k\|).
\end{aligned} \tag{3.17}$$

On the other hand, **A1** implies that $\|J(x^*)v\| \geq \sigma_1\|v\|$ for all $v \in \mathbb{R}^n$, where $\sigma_1 > 0$ is the smallest singular value of $J(x^*)$. Hence, a continuity argument yields that, in particular, $\|J(x^k)s^k\| \geq \frac{1}{2}\sigma_1\|s^k\|$ for all sufficiently large $k \in \mathcal{K}$. Therefore, we have, using **A5**,

$$\begin{aligned}
\|G(x^k)s^k\| &= \left\|J(x^k)s^k - \left[J(x^k) - G(x^k)\right] s^k\right\| \\
&\geq \|J(x^k)s^k\| - \left\|\left[G(x^k) - J(x^k)\right] s^k\right\| \\
&\geq \tfrac{1}{2}\sigma_1\|s^k\| - \kappa_{\rm G}\mu^{k-1}\|s^k\| \\
&= \left(\tfrac{1}{2}\sigma_1 - \kappa_{\rm G}\mu^{k-1}\right)\|s^k\| \\
&= \Omega(\|s^k\|),
\end{aligned} \tag{3.18}$$

since μ^{k-1} converges to zero. Combining (3.17) with (3.18), we obtain (3.16). ■

We now give an upper bound on the size of the step s^k computed from (2.6).

Lemma 3.5. *Suppose that* **A1**–**A5** *hold, that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5), and that the step* s^k *is computed from (2.6). Then, for sufficiently large* $k \in \mathcal{K}$,

$$\|s^k\| = O(\mu^{k-1}). \tag{3.19}$$

Proof. The rule (2.6) shows that

$$\begin{aligned} G(x^k)s^k &= h(x^k,\mu^k) - F(x^k) + r^k \\ &= h(x^k,\mu^k) - h(x^k,\mu^{k-1}) + h(x^k,\mu^{k-1}) - F(x^k) + r^k, \end{aligned} \tag{3.20}$$

where $\|r^k\| \le \eta^k$. In addition, **A3**, **A4**, (3.15) and the fact that $\mu^k < \mu^{k-1}$ imply that for sufficiently large $k \in \mathcal{K}$,

$$h(x^k,\mu^{k-1}) = \nabla_\mu h(x^*,0)\mu^{k-1} + O\left((\mu^{k-1})^2\right), \tag{3.21}$$

and

$$h(x^k,\mu^k) = \nabla_\mu h(x^*,0)\mu^k + O\left((\mu^{k-1})^2\right), \tag{3.22}$$

so that

$$h(x^k,\mu^k) - h(x^k,\mu^{k-1}) = O(|\mu^k - \mu^{k-1}|) + O\left((\mu^{k-1})^2\right) = O(\mu^{k-1}). \tag{3.23}$$

Combining (3.20) with (2.9), (3.2), (3.3) and (3.23), we thus find that

$$\begin{aligned} \|G(x^k)s^k\| &\le \|h(x^k,\mu^k) - h(x^k,\mu^{k-1})\| + \epsilon^{k-1} + \eta^k \\ &\le O(\mu^{k-1}) + \epsilon^{k-1} + \eta^k \\ &= O(\mu^{k-1}). \end{aligned} \tag{3.24}$$

Combining (3.24) with (3.16), we obtain $\|s^k\| = O(\mu^{k-1})$, which proves (3.19). ■

Our next result shows the benefit of using the advanced starting point (2.3) with s^k computed from (2.6).

Lemma 3.6. *Suppose that* **A1**–**A5** *hold and that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then, if*

$$\liminf_{\substack{k\to\infty \\ k\in\mathcal{K}}} \frac{\min(2,\theta_\eta^k)}{\theta_\epsilon^k\,\theta_\mu^{k-1}} > 1, \tag{3.25}$$

it follows that

$$\|F(x_s^k) - h(x_s^k,\mu^k)\| \le \epsilon^k, \tag{3.26}$$

for all sufficiently large $k \in \mathcal{K}$, *where the starting point* x_s^k *is computed as in Step 2 of Algorithm 2.1.*

Proof. It follows from **A2**, (2.3), (2.6) and **A5** that for all sufficiently large $k \in \mathcal{K}$,

$$\begin{aligned} F(x_S^k) &= F(x^k + s^k) && (3.27)\\ &= F(x^k) + J(x^k)s^k + O\left(\|s^k\|^2\right) && (3.28)\\ &= F(x^k) + G(x^k)s^k + [J(x^k) - G(x^k)]s^k + O\left(\|s^k\|^2\right) && (3.29)\\ &= h(x^k,\mu^k) + r^k + O(\mu^{k-1}||s^k||) + O\left(\|s^k\|^2\right) && (3.30)\\ &= h(x^k,\mu^k) + O(\eta^k) + O(\mu^{k-1}||s^k||) + O\left(\|s^k\|^2\right). && (3.31) \end{aligned}$$

Hence (3.3), Lemma 3.5, (3.31) and the fact that $\mu^k < \mu^{k-1}$ reveal that

$$\|F(x_{\mathrm{S}}^k) - h(x^k, \mu^k)\| = O(\eta^k) + O\left((\mu^{k-1})^2\right) = O\left((\mu^{k-1})^{\min(2,\theta_\eta^k)}\right). \tag{3.32}$$

On the other hand, using **A3**, **A4** and the expansions

$$h(x^k, \mu^k) = \nabla_\mu h(x^*, 0)\mu^k + O\left((\mu^{k-1})^2\right),$$

and

$$h(x_{\mathrm{S}}^k, \mu^k) = \nabla_\mu h(x^*, 0)\mu^k + O\left((\mu^{k-1})^2\right),$$

where we have used the fact that $\mu^k < \mu^{k-1}$, (3.15) and (3.19), we have

$$h(x^k, \mu^k) - h(x_{\mathrm{S}}^k, \mu^k) = O\left((\mu^{k-1})^2\right). \tag{3.33}$$

Thus, combining (3.32) and (3.33), we have

$$\begin{aligned}
\|F(x_{\mathrm{S}}^k) - h(x_{\mathrm{S}}^k, \mu^k)\| &\le \|F(x_{\mathrm{S}}^k) - h(x^k, \mu^k)\| + \|h(x^k, \mu^k) - h(x_{\mathrm{S}}^k, \mu^k)\| \\
&= O\left((\mu^{k-1})^{\min(2,\theta_\eta^k)}\right) + O\left((\mu^{k-1})^2\right) \\
&= O\left((\mu^{k-1})^{\min(2,\theta_\eta^k)}\right).
\end{aligned} \tag{3.34}$$

Equalities (3.1) and (3.2) give

$$\epsilon^k = \tau_\epsilon^k \ (\mu^k)^{\theta_\epsilon^k} = \tau_\epsilon^k \ (\tau_\mu^{k-1})^{\theta_\epsilon^k} \ (\mu^{k-1})^{\theta_\mu^{k-1}\theta_\epsilon^k}, \tag{3.35}$$

and observe that because of (3.5), (3.25) is equivalent to

$$\min(2, \theta_\eta^k) \ge \theta_\epsilon^k \theta_\mu^{k-1} + \nu, \tag{3.36}$$

for all sufficiently large $k \in \mathcal{K}$ and for some constant $\nu > 0$. The required bound (3.26) now follows, from (3.34), (3.35) and (3.36). ∎

It follows immediately from Lemma 3.6 that for all sufficiently large $k \in \mathcal{K}$, the starting point x_{S}^k satisfies the inner-iteration stopping rule (2.8) provided that the exponents θ_μ^{k-1}, θ_ϵ^k, and θ_η^k satisfy (3.25), and hence that we will choose, for all sufficiently large $k \in \mathcal{K}$,

$$x^{k+1} = x_{\mathrm{S}}^k, \tag{3.37}$$

since Step 3 of Algorithm 2.1 requires that the first point satisfying (2.8) is taken.

We now show that, up to at least first order, the asymptotic form (3.6) of $x(\mu)$ is inherited by any x^k close to x^*, satisfying (2.9).

Lemma 3.7. *Suppose that* **A1**–**A4** *hold, that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5) and that*

$$\liminf_{\substack{k\to\infty \\ k\in\mathcal{K}}} \theta_\epsilon^k > 1. \tag{3.38}$$

Then for all sufficiently large $k \in \mathcal{K}$,

$$x^k = x^* + \mu^{k-1}x' + o(\mu^{k-1}), \tag{3.39}$$

where x' *is given by (3.7).*

Proof. Inequality (2.9) may be written as

$$F(x^k) = h(x^k, \mu^{k-1}) + u^k, \quad \text{where} \quad \|u^k\| \le \epsilon^{k-1}.$$

It follows from **A2**, the fact that x^* is a root of (1.1) and (3.15), that

$$F(x^k) = J(x^*)(x^k - x^*) + w^k, \quad \text{where} \quad \|w^k\| = O(\|x^k - x^*\|^2) = O((\mu^{k-1})^2),$$

for all sufficiently large $k \in \mathcal{K}$. Hence, **A1** implies that

$$x^k - x^* = J^{-1}(x^*)(h(x^k, \mu^{k-1}) + u^k - w^k). \tag{3.40}$$

On the other hand, we have from **A3**, **A4** and (3.15) that

$$h(x^k, \mu^{k-1}) = \nabla_\mu h(x^*, 0)\mu^{k-1} + O\left((\mu^{k-1})^2\right), \tag{3.41}$$

which combines with (3.40) to give

$$x^k - x^* = J(x^*)^{-1}\left(\nabla_\mu h(x^*, 0)\mu^{k-1} + u^k - w^k + O\left((\mu^{k-1})^2\right)\right). \tag{3.42}$$

Observe now that (3.38) is equivalent to $\theta_\epsilon^k \ge \theta_\epsilon^*$ for some $\theta_\epsilon^* > 1$ and for all sufficiently large $k \in \mathcal{K}$, which, together with (3.2) implies that $\epsilon^{k-1} = o(\mu^{k-1})$. Combining this observation with the identities (3.7) and (3.42), we obtain (3.39). ■

Finally, we show that an estimate of the form (3.39) continues to hold at x^{k+1}.

Lemma 3.8. *Suppose that* **A1**–**A5** *hold, that* $\{x^k\}_{k\in\mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5), that (3.25) and (3.38) hold, that we have*

$$\liminf_{\substack{k\to\infty \\ k\in\mathcal{K}}} \theta_\eta^k > 1, \tag{3.43}$$

and

$$\limsup_{\substack{k\to\infty \\ k\in\mathcal{K}}} \theta_\mu^{k-1} < 2. \tag{3.44}$$

Then for all sufficiently large $k \in \mathcal{K}$,

$$x^{k+1} = x^* + \mu^k x' + o(\mu^k), \tag{3.45}$$

where x' *is given by (3.7).*

Proof. Applying the arguments used to derive (3.22) and (3.31), it follows from (2.3), (2.6), **A2**–**A5**, (3.37), Lemmas 3.3 and 3.5, and the inequality $\mu^k < \mu^{k-1}$ that

$$\begin{aligned} F(x^{k+1}) &= F(x^k + s^k) \\ &= h(x^k, \mu^k) + O(\eta^k) + O\left(\mu^{k-1}||s^k||\right) + O\left(\|s^k\|^2\right) \\ &= \nabla_\mu h(x^*, 0)\mu^k + O(\eta^k) + O\left((\mu^{k-1})^2\right), \end{aligned} \tag{3.46}$$

for sufficiently large $k \in \mathcal{K}$. On the other hand, using **A2**, the identity $F(x^*) = 0$, (3.15), (3.19), and the inequality $\|x^{k+1} - x^*\| \leq \|x^k - x^*\| + \|s^k\| = O(\mu^{k-1})$, a second-order Taylor expansion gives

$$\begin{aligned} F(x^{k+1}) &= F(x^*) + J(x^*)(x^{k+1} - x^*) + O\left(\|x^{k+1} - x^*\|^2\right) \\ &= J(x^*)(x^{k+1} - x^*) + O\left((\mu^{k-1})^2\right). \end{aligned} \tag{3.47}$$

Observe that because of (3.5), (3.44) is equivalent to $1 \leq \theta_\mu^{k-1} \leq \theta_\mu^*$ for some $\theta_\mu^* < 2$ and for all sufficiently large $k \in \mathcal{K}$, which combines with (3.1) to give $(\mu^{k-1})^2 = o(\mu^k)$. Similarly, (3.43) is equivalent to $\theta_\eta^k \geq \theta_\eta^*$ for some $\theta_\eta^* > 1$ and for all sufficiently large $k \in \mathcal{K}$, which, together with (3.3) ensures that $\eta^k = \tau_\eta^k \ (\mu^k)^{\theta_\eta^k} = o(\mu^k)$. Using these observations and combining (3.46) and (3.47), **A1** implies that

$$\begin{aligned} x^{k+1} - x^* &= \mu^k J(x^*)^{-1} \nabla_\mu h(x^*, 0) + O(\eta^k) + O\left((\mu^{k-1})^2\right) \\ &= \mu^k J(x^*)^{-1} \nabla_\mu h(x^*, 0) + o(\mu^k), \end{aligned}$$

which completes the proof. ∎

Before moving on to the main theorems of this section, we now collect all the previous assumptions on the exponents (3.5) in a single assumption. For convenience, the main theorems will refer to this single assumption rather than to the previous individual requirements. These requirements are put together in Assumption **A6**.

A6. The exponents (3.5) used in the updating rules

$$\mu^{k+1} = \tau_\mu^k \ (\mu^k)^{\theta_\mu^k}, \quad \epsilon^k = \tau_\epsilon^k \ (\mu^k)^{\theta_\epsilon^k}, \quad \text{and} \quad \eta^k = \tau_\eta^k \ (\mu^k)^{\theta_\eta^k}$$

satisfy

$$\limsup_{\substack{k \to \infty \\ k \in \mathcal{K}}} \theta_\mu^{k-1} < 2, \quad \liminf_{\substack{k \to \infty \\ k \in \mathcal{K}}} \theta_\epsilon^k > 1, \quad \liminf_{\substack{k \to \infty \\ k \in \mathcal{K}}} \theta_\eta^k > 1,$$

and

$$\liminf_{\substack{k \to \infty \\ k \in \mathcal{K}}} \frac{\min(2, \theta_\eta^k)}{\theta_\epsilon^k \, \theta_\mu^{k-1}} > 1,$$

along the subsequence $\{x^k\}_{k \in \mathcal{K}}$ of interest.

Let us now define the set

$$\mathcal{I} = \{i = 1, \ldots, n \mid x'_i \neq 0\}, \tag{3.48}$$

where x' is defined by (3.7). Strictly speaking, the index set $\mathcal{I}$ depends on the function h, but we keep this dependence implicit as no particular instances of h are considered in this section.

Note that the fundamental difference between Lemma 3.7 and Lemma 3.8 is that (3.39) is concerned with an iterate x^k belonging to a converging subsequence indexed by $\mathcal{K}$, while in (3.45), there is no assumption that $k + 1 \in \mathcal{K}$. This fact is the key to the proof of the next result.

Theorem 3.1. *Suppose that* **A1**–**A6** *hold, that* $\{x^k\}_{k \in \mathcal{K}} \to x^*$ *where* $\{x^k\}$ *is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then the* complete *sequence* $\{x^k\}$ *converges to* x^* *and for all sufficiently large* k*, we have* *(i).* $\quad x^{k+1} = x^* + \mu^k x' + o(\mu^k) \quad$ *where* $x' = J(x^*)^{-1}\nabla_\mu h(x^*, 0)$,

(ii). $\quad \dfrac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^{\theta_\mu^{k-1}}} = \Theta(1) \quad$ *provided that* $x' \neq 0$*, and*

(iii). $\quad \dfrac{|x_i^{k+1} - x_i^*|}{|x_i^k - x_i^*|^{\theta_\mu^{k-1}}} = \Theta(1) \quad$ *for every component* $i \in \mathcal{I}$*,*

where the set $\mathcal{I}$ *is defined in (3.48). As a consequence, the entire sequence* $\{x^k\}$ *converges to* x^* *at a (componentwise) Q-superlinear rate determined by* $\limsup_{k\to\infty} \theta_\mu^k$.

Proof. It follows from Lemma 3.8 and the fact that $\mu^k \to 0$ that the sequence $\{x^{k+1}\}_{k\in\mathcal{K}}$ also converges to x^*. Applying again Lemma 3.7 and Lemma 3.8 to the subsequence of $\{x^k\}$ defined by the index set $\mathcal{K}^+ = \mathcal{K} \cup \{k+1 \mid k \in \mathcal{K}\}$, we conclude that the sequence $\{x^{k+1}\}_{k\in\mathcal{K}^+}$ also converges to x^*. An inductive argument proves (i) and thus shows that the whole sequence $\{x^k\}$ converges to x^*. To prove (ii), let us assume that $x' \neq 0$. We then have from (i) that $\|x^k - x^*\| = \Theta(\mu^{k-1})$ and $\|x^{k+1} - x^*\| = \Theta(\mu^k)$, which we combine with (3.1) to obtain the desired result. We prove (iii) similarly, by considering equations (3.39) and (3.45) componentwise. ∎

Interestingly, the results of Theorem 3.1 carry over to the sequence $\{F(x^k)\}$. Indeed, upon defining the index set

$$\mathcal{J} = \{j = 1, \ldots, n \mid [\nabla_\mu h(x^*, 0)]_j \neq 0\}, \tag{3.49}$$

we have the following result.

Theorem 3.2. *Suppose that* **A1**–**A6** *hold, that* $\{x^k\}$ *is the convergent sequence described in Theorem 3.1, generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then the sequence* $\{F(x^k)\}$ *converges to zero, and for all sufficiently large* k*, we have*

(i). $\quad F(x^{k+1}) = \mu^k F' + o(\mu^k), \quad$ *where* $F' = \nabla_\mu h(x^*, 0)$,

(ii). $\quad \dfrac{\|F(x^{k+1})\|}{\|F(x^k)\|^{\theta_\mu^{k-1}}} = \Theta(1) \quad$ *provided that* $F' \neq 0$*, and*

(iii). $\quad \dfrac{|f_j(x^{k+1})|}{|f_j(x^k)|^{\theta_\mu^{k-1}}} = \Theta(1) \quad$ *for every component* $j \in \mathcal{J}$*,*

where the set $\mathcal{J}$ *is defined in (3.49). As a consequence, the entire sequence* $\{F(x^k)\}$ *converges to zero at a (componentwise) Q-superlinear rate determined by* $\limsup_{k\to\infty} \theta_\mu^k$.

Proof. From Theorem 3.1, the expansion (3.45) is valid for all index k. Therefore, denoting $\overline{x}^{k+1} = x^{k+1} - x^*$, and combining (3.45) with a Taylor expansion of F about x^*, we have

$$\begin{aligned} F(x^{k+1}) &= F(x^* + \overline{x}^{k+1}) \\ &= F(x^*) + J(x^*)\overline{x}^{k+1} + o(\|\overline{x}^{k+1}\|) \\ &= J(x^*)\left(\mu^k x' + o(\mu^k)\right) + o(\mu^k) \\ &= \mu^k \nabla_\mu h(x^*, 0) + o(\mu^k), \end{aligned} \tag{3.50}$$

where we have used (3.7) and the identity $F(x^*) = 0$. This yields (i). Results (ii) and (iii) are proven along the same lines as (i) and (ii) of Theorem 3.1. ∎

The combination of Theorem 3.1 and Theorem 3.2 results in the following conclusion. The sequences $\{x^k\}$ and $\{F(x^k)\}$ both converge at the *exact same Q-rate* to x^* and zero respectively, and so do all the sequences of components $\{x_i^k\}$ for $i \in \mathcal{I}$ and $\{f_j(x^k)\}$ for $j \in \mathcal{J}$. Moreover, if x' and F' are nonzero, the identities

$$\|F(x^{k+1})\| = \Theta(\mu^k) = \Theta(\|x^{k+1} - x^*\|),$$

and

$$f_j(x^{k+1}) = \Theta(\mu^k) = \Theta(|x_i^{k+1} - x_i^*|) \quad i \in \mathcal{I},\ j \in \mathcal{J},$$

hold for all k and may be seen as a theoretical justification to basing a stopping criterion for Algorithm 2.1 on the magnitude of μ^k. Note that whenever $i \notin \mathcal{I}$ and $j \notin \mathcal{J}$, we have $|x_i^{k+1} - x_i^*| = o(\mu^k)$ and $f_j(x^{k+1}) = o(\mu^k)$. It is also interesting to remark that the two tangent vectors x' and F' are obtained from each other using the simple relationships

$$x' = J(x^*)^{-1}F' \qquad \text{and} \qquad F' = J(x^*)x',$$

and hence that x' and F' are simultaneously zero or nonzero.

The asymptotic componentwise Q-rate of convergence of $\{x^k\}$ to x^* and of $\{F(x^k)\}$ to zero is determined by the upper limit $\limsup_{k\to\infty} \theta_\mu^k$. Under the assumption that $x_i' \neq 0$ for all $i = 1, \ldots, n$, and according to the restrictions imposed by **A6**, a componentwise Q-rate of convergence arbitrarily close to 2 is achieved by choosing $\{\theta_\mu^k\}$ and $\{\theta_\epsilon^k\}$ so that $\limsup_{k\to\infty} \theta_\mu^k = 2 - \alpha$, $\liminf_{k\to\infty} \theta_\epsilon^k = 1 + \beta$ and $\theta_\eta^k > 2$ for all sufficiently large k, where $\alpha, \beta \in (0, 1)$ can be chosen as small as desired but must satisfy $\beta < \alpha/(2 - \alpha)$. A Q-rate of 1.5 may be obtained by setting $\theta_\mu^k = 1.5$, $\theta_\epsilon^k = 1.01$ for all k and θ_η^k such that $\liminf_{k\to\infty} \theta_\eta^k > 1.515$, while a Q-rate of 1.9 may be obtained by setting $\theta_\mu^k = 1.9$, $\theta_\epsilon^k = 1.01$ for all k and θ_η^k such that $\liminf_{k\to\infty} \theta_\eta^k > 1.919$. The ultimate rate of convergence, entirely determined by $\limsup_{k\to\infty} \theta_\mu^k$, is therefore prescribed by the user. In this respect the convergence properties of Algorithm 2.1 are similar to those of [7].

4. Applications

We first examine in this section the relationships between the class of algorithms described in §2–§3 and interior-point methods for nonlinear programming.

Consider the nonlinear program

$$\text{NLP} \equiv \begin{cases} \min & f(x) \\ \text{s.t.} & g(x) = 0 \\ & c(x) \geq 0, \end{cases} \tag{4.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $c : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be thrice continuously differentiable functions. We denote by g_j $(j = 1, \dots, m)$ and c_i $(i = 1, \dots, p)$ the components of g and c respectively. If the *strictly feasible set* for NLP

$$\mathcal{F}^\circ = \{x \in \mathbb{R}^n \mid g(x) = 0,\ c(x) > 0\}$$

is nonempty, a common procedure is to replace NLP by a sequence of *barrier problems* of the form

$$\mathrm{BS}(\mu) \equiv \begin{cases} \min & \phi(x,\mu) \\ \text{s.t.} & g(x) = 0, \end{cases} \tag{4.2}$$

for a sequence $\{\mu_k\}$ of positive barrier parameters converging to zero and where $\phi(x,\mu) = f(x) - \mu \sum_{i=1}^p \log c_i(x)$ is the *logarithmic barrier*. The procedure starts with a strictly feasible point $x^0 \in \mathcal{F}^\circ$ and at iteration k, fixes μ to a positive value μ^k and seeks an approximate minimizer x^{k+1} of $\mathrm{BS}(\mu^k)$. The parameter μ is then decreased so as to converge to zero, and attention turns to the next barrier problem [10,15,30,32,34]. This procedure is a path-following method as, under reasonable regularity assumptions, exact solutions to $\mathrm{BS}(\mu)$ form a locally-unique trajectory which can be numerically tracked and which leads to a local solution of NLP. It turns out that these exact solutions are precisely the exact solutions to the square *primal-dual system*

$$\Psi(x,y,z;\mu) \equiv \begin{bmatrix} \nabla_x \mathcal{L}(x,y,z) \\ g(x) \\ C(x)z \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \mu e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.3}$$

where $\mathcal{L}(x,y,z) = f(x) + g(x)^T y - c(x)^T z$ is the *Lagrangian function* for problem NLP, $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^p_+$ are the *Lagrange multipliers* corresponding to the equality and inequality constraints respectively, $C(x) = \mathrm{diag}(c(x))$ and e is the vector of all ones of appropriate dimension. The path of exact solutions is referred to as the *primal-dual central path*.

When $\mu = 0$, the system (4.3) coincides with the optimality conditions for problem NLP, so that the solution of NLP may be recast in the framework of this paper, identifying the composite vector (x,y,z) with the vector x of (1.1), $F(x)$ with the first term of the left-hand side of (4.3), and $h(x,\mu)$ with the second term of the left-hand side of (4.3). Under regularity assumptions, and assuming that some globally-convergent algorithm produced a point (x^0, y^0, z^0) close enough to a solution (x^*, y^*, z^*) of NLP, Algorithm 2.1 may be applied to system (4.3) and the results described in §3 will follow. It should be noted however that the corresponding tangent vector x' given by (3.7) is not guaranteed to have the desired nonzero components. Nevertheless, we conclude that for all such components, a Q-superlinear rate of convergence is possible, if μ is decreased accordingly. Theorem 3.2 guarantees that all the products $c_i(x)z_i$ converge at the same rate to zero, and that this rate may be as close to quadratic as desired. Such conclusions hold for the methods studied in [11,12].

We may of course choose a different $h(x,\mu)$ in (4.3), perhaps at the risk of loosing the interpretation in terms of barrier functions. Note that this may also result in an infeasible algorithm. For instance, choosing $h(x,\mu)$ such that the tangent vector x' only has nonzero components would imply that all the components of x, y and z converge at the same rate to their limit, as would any subvector of (x,y,z). Choosing $h(x,\mu)$ such that the tangent vector $F' = \nabla_\mu h(x^*, 0)$

only has nonzero components would imply that all the partial derivatives of the Lagrangian function, all the components of the infeasibility $g(x)$ and all the components of the complementary slackness $C(x)z$ converge to zero at the same rate.

It is important to note that one could solve (4.3) by applying a globally-convergent optimization algorithm to minimize

$$\tfrac{1}{2}\,\|\Psi(x,y,z;\mu)\|_2^2, \tag{4.4}$$

in the variables (x, y, z). However, the generated sequence might quite possibly converge to a *local* minimizer of (4.4) which is not a solution of (4.3), while the solutions of (4.3) correspond to *global* minimizers of (4.4). Such situations may arise for functions as simple as one-dimensional polynomials, such as $\psi(x) = x^3 - 2x^2 + 5$, whose only real root is -1.2419, and corresponds to a global minimizer of (4.4), while in this case, it is easily verified that (4.4) possesses another, non-global, local minimizer at 4/3. Note, however, that in general, since the ℓ_2-norm is used, any local minimizer of (4.4) that satisfies assumption **A1** is necessarily a global minimizer. The distinction between *local* and *global* solutions is fundamental in the context of nonlinear programming and is in fact at the roots of *continuation* and *homotopy* methods [1,23,25–27,29].

We now turn to a possible justification for computing the step s^k using (2.3) and (2.6). Our motivation is that using (2.3) and (2.6) may be interpreted as a method which extrapolates along the path of exact solutions $x(\mu)$ to (2.5), in a manner similar to that used in *primal* interior-point methods for inequality-constrained optimization. In that framework, if x^k is the current estimate of $x(\mu^{k-1})$ and if the parameter μ^{k-1} is updated to μ^k, the unit Newton step $d^{\text{N}}(x^k, \mu^k)$ for (2.5)—thus computed using the Jacobian matrix $\nabla_x h(x^k, \mu^k)$ or a nonzero approximation thereof—is likely to be infeasible with respect to the inequality constraints, no matter how close x^k is to x^* [31]. One thus has to cut back quite severely, which may considerably thwart fast convergence. To circumvent this problem, Dussault [9] computes an extrapolation step, $d^{\text{EX}}(x^k, \mu^{k-1}, \mu^k)$, to the trajectory $x(\mu)$ at the current iterate x^k, following the tangential to the trajectory of an amount equal to $\mu^{k-1} - \mu^k$, followed by the Newton step $d^{\text{N}}(x^k, \mu^{k-1})$ that would have been taken from x^k using the *old* barrier parameter—the sum of these two steps $d^{\text{EX}}(x^k, \mu^{k-1}, \mu^k) + d^{\text{N}}(x^k, \mu^{k-1})$ being similar to the step s^k given by (2.6) despite the absence of μ^{k-1} in (2.6) [12]. This process turns out to be superlinearly convergent with limiting exponent 4/3. Similar conclusions hold for equality and inequality-constrained nonlinear programs solved using an exponential penalty function [3]. In a primal-dual interior-point framework, the aforementioned infeasibility problem of the Newton step does not arise, and the extrapolation method detailed by Dussault coincides with Newton's method, due to the special structure of the primal-dual system. In other words, we have $d^{\text{EX}}(x^k, \mu^{k-1}, \mu^k) + d^{\text{N}}(x^k, \mu^{k-1}) = d^{\text{N}}_{\text{PD}}(x^k, \mu^k)$, where $d^{\text{N}}_{\text{PD}}(x^k, \mu^k)$ is the primal-dual Newton step, *i.e.*, the Newton step for (4.3), taken from x^k using the new parameter μ^k. The process may therefore result in a componentwise nearly quadratic Q-rate of convergence [12]. The structure of the primal-dual function is such that the Jacobian matrix $\nabla_x h(x^k, \mu^k)$ is absent from the Newton system, which is our main motivation for computing the step s^k using (2.6). Not surprisingly, similar conclusions hold for methods using a quadratic penalty function [11] or an exponential penalty function [4].

Interestingly, our approach can also easily be applied to the framework of nonlinear least-squares problems, where one solves

$$\min_x \tfrac{1}{2}\|R(x)\|_2^2 \tag{4.5}$$

for some smooth function R from $\mathbb{R}^n$ into $\mathbb{R}^m$ with $m \geq n$. If we assume that the Jacobian matrix of R, $\nabla_x R(x)$ is of full rank at the solution x^* of (4.5), the normal equations then provide a link to our approach. Indeed, we may replace (2.8) by

$$\|\nabla_x R(x^{k+1})^T R(x^{k+1}) - h(x^{k+1}, \mu^k)\| \leq \epsilon^k$$

and compute the step to x_{S}^k using

$$G(x^k)s^k = h(x^k, \mu^k) - \nabla_x R(x^k)^T R(x^k) + r^k \quad \text{and} \quad \|r^k\| \leq \eta^k,$$

where now $G(x^k) \approx \nabla_x R(x^k)^T \nabla_x R(x^k)$. Note however that for this $G(x)$ to satisfy **A1** and **A5**, we must be seeking a global solution to (4.5), *i.e.*, $R(x^*) = 0$.

5. Discussion

So far, we have not commented on appropriate choices for $h(x, \mu)$. Perhaps the most obvious and simplest choice is the linear function

$$h(x, \mu) = \mu e, \tag{5.1}$$

where e is the vector of all ones, for which $\nabla_\mu h(x^*, 0) = e$. As we have seen in (4.3), a similar choice is made in primal-dual interior-point methods where the complementarity condition is perturbed and where we choose

$$h(x, \mu) = \mu \tilde{e}, \tag{5.2}$$

where the fixed vector $\tilde{e}$ is made of zeros and ones. We refer the interested reader to [12] for more details on a primal-dual interior-point method in relation with the componentwise convergence of the iterates. In view of Theorem 3.1, ideally we would like to use an $h(x, \mu)$ for which every component of x' is nonzero. We have no such guarantee if we use (5.1) or (5.2). For instance,

$$\nabla_\mu h(x^*, 0) = J(x^*)e, \tag{5.3}$$

such as would arise if

$$h(x, \mu) = \mu J(x^*)e, \tag{5.4}$$

would be ideal for then $x' = e$, but of course depends on the (unknown) $J(x^*)$. If F is twice continuously differentiable, the choice

$$h(x, \mu) = \mu J(x)e, \tag{5.5}$$

which also aims for $x' = e$, is more similar to a primal interior-point method, and differs from (5.4) in a number of respects. In the framework of interior-point methods, the most obvious of these respects is probably that if $\nabla_x[F(x) - h(x, \mu)]$ is independent of μ, as is the case in (5.2) and (5.4), then so is the region of convergence of the Newton-like method based on (2.6) [2, Lemma 2.1], [16, §3.2.12]. Indeed, in primal-dual methods, similar to using (5.2), the size of this region of convergence is $\Theta(1)$ while in primal methods, similar to using (5.5), its size is $\Theta(\mu)$ and thus shrinks at the rate at which μ decreases to zero [22]. This behaviour is observed in the numerical tests of Section 6. A particular form of $h(x, \mu)$ may emerge naturally from the problem itself, or may need to have properties dictated by the problem, for instance, a higher level of nonlinearity in μ. See for instance [24] for references to such work. If we are only interested in the convergence properties of certain components then, in view of Theorem 3.1 (ii), the perturbation function $h(x, \mu)$ may be chosen such that the asymptotic behaviour of these components is controlled, *i.e.*, such that the corresponding components of the vector x' are nonzero. In any case, we certainly want to choose $h(x, \mu)$ and μ^0 such that a solution to $F(x) = h(x, \mu^0)$ is either known or easy to find.

We now briefly comment on the assumptions we made in §2. While the first part of Assumption **A4** is clearly implied by our intent to use a path-following method, the last part limits the level of dependence of h on x, at least close to x^*, which is in line with our initial motivation where h is independent of x. It is easily verified that the examples of this section and of §6 satisfy **A4**. This assumption guarantees that the componentwise behaviour described in Theorem 3.1 (ii) occurs. It could be replaced by any more general assumption on h guaranteeing that for all sufficiently small μ, the bound $\|x(\mu) - x^*\| = \Theta(\mu)$ holds. It is the case under **A4** if in addition $\nabla_\mu h(x^*, 0)$ is such that the tangent vector x' is nonzero (see (3.6)–(3.7)). An example of a function failing to satisfy **A4** is $h(x, \mu) = \mu e + (x - x^*)$.

Regarding **A5**, one could argue that having the right-hand side of (2.7) depend on μ is not intuitive since its left-hand side does not. In view of Lemma 3.5, it is possible to replace the right-hand side of (2.7) by $\kappa_{\rm G} \, ||s^k||^2$—which would be stronger since the bound $\mu^{k-1} = O(\|s^k\|)$ does not necessarily hold—or even by $\kappa_{\rm G} \, ||s^k||^{\theta_S^k}$ for some θ_S^k satisfying $\liminf_{k\to\infty} \theta_S^k > 1$. This last choice would however slightly modify condition (3.25) which we would need to replace by

$$\liminf_{\substack{k\to\infty \\ k\in\mathcal{K}}} \frac{\min(2, \theta_\eta^k, \theta_S^k)}{\theta_\epsilon^k \, \theta_\mu^{k-1}} > 1. \tag{5.6}$$

In all cases, these conditions imply that

$$\left\| \left[G(x^k) - J(x^k) \right] s^k \right\| = o(\|s^k\|), \tag{5.7}$$

and it is easily seen that **A1**–**A2** and (5.7) together imply that the Dennis-Moré condition holds. Indeed, since the whole sequence $\{x^k\}$ converges to x^*, we have

$$\begin{aligned} \left\| \left[G(x^k) - J(x^*) \right] s^k \right\| &\le \left\| \left[G(x^k) - J(x^k) \right] s^k \right\| + \left\| \left[J(x^k) - J(x^*) \right] s^k \right\| \\ &\le o(\|s^k\|) + L_{\rm J} \, \|x^k - x^*\| \; \|s^k\| \\ &= o(\|s^k\|), \end{aligned}$$

which is exactly [8, Condition (8.2.16)], where the constant $L_{\rm J} > 0$ comes from **A2**.

A weaker assumption would be to replace the right-hand side of (2.7) by $\kappa_{\rm G} \, (\mu^{k-1})^2$, or, weaker still, $\kappa_{\rm G} \, (\mu^{k-1})^{\theta_{\rm G}^k}$ for some $\theta_{\rm G}^k$ such that $\liminf_{k\to\infty} \theta_{\rm G}^k > 1$, in which case (3.25) becomes

$$\liminf_{\substack{k\to\infty \\ k\in\mathcal{K}}} \frac{\min(2, \theta_\eta^k, \theta_{\rm G}^k)}{\theta_\epsilon^k \, \theta_\mu^{k-1}} > 1.$$

In this last case again, the left and right-hand sides of (2.7) would not involve quantities that are directly comparable.

Finally, a comment on the accuracy requirement in (2.6) is in order. Although not particularly restrictive, this requirement matches a classical pattern in inexact quasi-Newton methods [6], in which it is shown that if $\|r^k\| = o(\|F(x^{k+1}) - h(x^{k+1}, \mu^k)\|)$, the sequence $\{x^k\}$ generated by the iteration (2.6) is superlinearly convergent. However, the stopping criterion (2.8) and the updating rule (3.3) give to the results of this paper a different flavour. Indeed, it follows from Theorem 3.2, (3.25) and the updating rule (3.3) that

$$\|r^k\| \le \eta^k = \Theta\left((\mu^k)^{\theta_\eta^k}\right) = O\left((\mu^k)^{\theta_\epsilon^k \theta_\mu^{k-1}}\right) = O\left(\|F(x^{k+1})\|^{\theta_\epsilon^k \theta_\mu^{k-1}}\right),$$

which obviously implies that $\|r^k\| = o(\|F(x^{k+1})\|)$ but not necessarily that $\|r^k\| = o(\|F(x^{k+1}) - h(x^{k+1}, \mu^k)\|)$.

6. Numerical experiments

In this section we illustrate the above results on Example 1.1 when $n = 5$ and when $G(x^k) = J(x^k)$ for all k. The starting point is chosen as $x^1 = 0.8\, e_3$. The following tables indicate the values taken by the components of the vector x and by $\|x\|_2$ along the iterations when solving the problem given in Example 1.1 for the indicated functions $h(x, \mu)$. The first row of each table indicates the starting values and the iteration counter increases when reading the table from top to bottom. In Table 6.1, a pure Newton scheme is applied, which corresponds to $h(x, \mu) = 0$, and we observe the behaviour described in Example 1.1. In this and subsequent experiments, all computation is performed in extended-precision arithmetic under Mathematica, so as to see the predicted rates of convergence, and the numbers have been subsequently truncated to four significant digits unless the extended-precision number already had less significant digits. Table 6.2 corresponds to the case where we set $h(x, \mu) = \mu J(x^*) e$, where $\mu^0 = 0.9$ and is updated according to $\mu^{k+1} = (\mu^k)^{1.9}$, which corresponds to using $\theta_\mu^k = 1.9$ for all k. We also use the values $\epsilon^0 = \|F(x^1) - h(x^1, \mu^0)\|$, $\theta_\epsilon^k = 1.05$, $r^k = 0$—*i.e.*, the system (2.6) is solved exactly—and $\tau_{\min} = \tau_\mu^k = \tau_\epsilon^k = \tau_\eta^k = \tau_{\max} = 1$ for all k. In this case, a single step (2.6) suffices for each value of μ^k, because x^k happens to lie inside the region of convergence of Newton's method for all k. In Table 6.3, we set $h(x, \mu) = \mu J(x) e$ and use the same rules for μ. In the inner iteration, we choose to take steps given by (2.6) as well. This time, two steps (2.6) are required for each of the first 8 iterations—a possible reason for this is that the region of convergence of Newton's method shrinks faster as x^* is approached than when using $h(x, \mu) = \mu J(x^*) e$. For the remaining iterations, a single step is sufficient. Notice that a single step suffices precisely where the (nearly) quadratic convergence shows up. It is also worth noticing that $\|x\|_2$ converges almost quadratically to zero in all cases, but that this behaviour appears later in the course of the iterations whenever $h(x, \mu) \ne 0$.

In Table 6.4, Table 6.5 and Table 6.6, the corresponding values of $F(x^k)$ and $\|F(x^k)\|_2$ along the course of the iterations is shown, for the same perturbation functions, illustrating Theorem 3.2. The present example has the particularity that the solution to $F(x) = 0$ is precisely $x^* = 0$, that $J(x^*)$ is the identity matrix and therefore that $x' = F' = e$. As a consequence, all the components of

k	x_1	x_2	x_3	x_4	x_5	$\|x\|_2$
1	0	0	0.8	0	0	0.8
2	0	0	0	0.64	0	0.64
3	0	0	0	0	0.4096	0.4096
4	0.1678	0	0	0	0	0.1678
5	0	0.0281	0	0	0	0.0281
6	0	0	0.0008	0	0	0.0008
7	0	0	0	6.2771$e-$7	0	6.2771$e-$7
8	0	0	0	0	3.9402$e-$13	3.9402$e-$13
9	1.5525$e-$25	0	0	0	0	1.5525$e-$25
10	0	2.4103$e-$50	0	0	0	2.4103$e-$50
11	0	0	5.8096$e-$100	0	0	5.8096$e-$100
12	0	0	0	3.3751$e-$199	0	3.3751$e-$199

Table 6.1. Components of x^k for $h(x,\mu) = 0$

k	x_1	x_2	x_3	x_4	x_5	$\|x\|_2$
1	0	0	0.8	0	0	0.8
2	0.8186	0.8186	0.8186	0.1488	0.8186	1.6439
3	0.4926	0.5471	0.4579	0.6041	0.5259	1.1804
4	0.3392	0.3939	0.3538	0.3711	0.4020	0.8335
5	0.2255	0.2154	0.2388	0.2095	0.2355	0.5037
6	0.0916	0.0832	0.0842	0.0904	0.0796	0.1921
7	0.0113	0.0133	0.0117	0.0121	0.0130	0.0276
8	0.0002	0.0002	0.0002	0.0002	0.0002	0.0005
9	6.6918$e-$8	7.6882$e-$8	5.8296$e-$8	8.1508$e-$8	6.2239$e-$8	1.5590$e-$7
10	5.5901$e-$15	6.1944$e-$15	7.6272$e-$15	5.1148$e-$15	8.3599$e-$15	1.4961$e-$14
11	1.5777$e-$28	1.1912$e-$28	1.2779$e-$28	1.4594$e-$28	1.1360$e-$28	2.9936$e-$28
12	0.	0.	0.	0.	0.	0.

Table 6.2. Components of x^k for $h(x,\mu) = \mu J(x^*)e$

k	x_1	x_2	x_3	x_4	x_5	$\|x\|_2$
1	0	0	0.8	0	0	0.8
2	1.0435	1.2237	0.8994	1.4830	1.1436	2.6273
3	0.9314	1.3006	0.8849	1.1926	1.2267	2.5038
4	0.6189	0.9793	0.8038	0.7443	1.0460	1.9069
5	0.4577	0.3339	0.7409	0.1549	0.6364	1.1397
6	0.1479	0.4608	0.0354	0.3604	0.0145	0.6046
7	0.0539	0.0054	0.0167	0.0720	0.0028	0.0917
8	0.0002	0.0001	0.0002	0.0001	0.0001	0.0004
9	8.1276$e-$5	8.1310$e-$5	8.1275$e-$5	8.1290$e-$5	8.1275$e-$5	0.0002
10	2.3538$e-$8	2.3538$e-$8	2.3544$e-$8	2.3538$e-$8	2.3541$e-$8	5.2636$e-$8
11	2.2706$e-$15	2.2704$e-$15	2.2705$e-$15	2.2707$e-$15	2.2704$e-$15	5.0771$e-$15
12	9.348$e-$29	9.348$e-$29	9.348$e-$29	9.348$e-$29	9.348$e-$29	2.0903$e-$29
13	0.	0.	0.	0.	0.	0.

Table 6.3. Components of x^k for $h(x,\mu) = \mu J(x)e$

x and $F(x)$ converge at the exact same Q-rate—equal to 1.9 in this illustration—to zero, along the same tangent vector as the solution is approached.

7. Conclusion

In this paper, we have studied a path-following method for solving square systems of non-singular, nonlinear equations, which may be viewed as the "end game" of a globally-convergent homotopy algorithm. Not surprisingly, since the method presented may be seen as a quasi-Newton method obeying the Dennis-Moré condition [8, Theorem 8.2.4], normwise Q-superlinear convergence of the sequence of iterates is achieved and observed numerically. More interestingly, the method exhibits an asymptotic convergence rate that is essentially as fast

k	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$\|F(x)\|_2$
1	0	0.8	0.64	0	0	1.0245
2	0	0	0.64	0.4096	0	0.75985
3	0	0	0	0.4096	0.16777	0.44263
4	0.028147	0	0	0	0.16777	0.17012
5	0.028147	0.00079228	0	0	0	0.02816
6	0	0.00079228	$6.2771e{-7}$	0	0	0.0008
7	0	0	$6.2771e{-7}$	$3.9402e{-13}$	0	$6.2771e{-7}$
8	0	0	0	$3.9402e{-13}$	$1.5525e{-25}$	$3.9402e{-13}$
9	$2.4103e{-50}$	0	0	0	$1.5525e{-25}$	$1.5525e{-25}$
10	$2.4103e{-50}$	$5.8096e{-100}$	0	0	0	$2.4103e{-50}$
11	0	$5.8096e{-100}$	$3.3752e{-199}$	0	0	$5.8096e{-100}$
12	0	0	$3.3752e{-199}$	$1.13917e{-397}$	0	$3.3752e{-199}$

Table 6.4. Components of $F(x^k)$ for $h(x, \mu) = 0$

k	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$\|F(x)\|_2$
1	0	0.8	0.64	0	0	1.0245
2	1.4886	1.4886	0.8189	0.8407	1.4886	2.83297
3	0.7899	0.7573	0.8137	0.8908	0.7693	1.80134
4	0.5090	0.5089	0.4963	0.5397	0.5008	1.14303
5	0.2663	0.2852	0.2665	0.2794	0.2810	0.61672
6	0.0916	0.0911	0.0975	0.0878	0.0980	0.208554
7	0.0135	0.0119	0.0123	0.0137	0.0115	0.027919
8	0.0002	0.0002	0.0002	0.0002	0.0002	0.000511321
9	$7.6882e{-8}$	$5.8296e{-8}$	$8.1508e{-8}$	$6.2239e{-8}$	$6.6918e{-8}$	$1.55899e{-7}$
10	$6.1944e{-15}$	$7.6272e{-15}$	$5.1148e{-15}$	$8.3599e{-15}$	$5.5901e{-15}$	$1.49614e{-14}$
11	$1.1912e{-28}$	$1.2779e{-28}$	$1.4594e{-28}$	$1.1357e{-28}$	$1.5777e{-28}$	$2.99743e{-28}$
12	0.	0.	0.	0.	0.	0.

Table 6.5. Components of $F(x^k)$ for $h(x, \mu) = \mu J(x^*)e$

k	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$\|F(x)\|_2$
1	0	0.8	0.64	0	0	1.0245
2	2.3126	2.3968	2.2920	3.3430	2.3514	5.7490
3	2.1682	2.5764	1.9756	2.6490	2.4363	5.3098
4	1.3623	1.7628	1.3904	1.5999	1.7131	3.5201
5	0.5434	0.8524	0.7038	0.6605	0.8628	1.6424
6	0.4827	0.2477	0.3617	0.1444	0.1482	0.6840
7	0.0083	0.0167	0.0723	0.0080	0.0539	0.0924
8	0.0001	0.0002	0.0001	0.0001	0.0002	0.0004
9	$8.1317e{-5}$	$8.1281e{-5}$	$8.1297e{-5}$	$8.1281e{-5}$	$8.1283e{-5}$	$1.8177e{-4}$
10	$2.3538e{-8}$	$2.3544e{-8}$	$2.3538e{-8}$	$2.3541e{-8}$	$2.3538e{-8}$	$5.2636e{-8}$
11	$2.2704e{-15}$	$2.2705e{-15}$	$2.2707e{-15}$	$2.2704e{-15}$	$2.2706e{-15}$	$5.0771e{-15}$
12	$9.348e{-29}$	$9.348e{-29}$	$9.348e{-29}$	$9.348e{-29}$	$9.348e{-29}$	$2.0903e{-28}$
13	0.	0.	0.	0.	0.	0.

Table 6.6. Components of $F(x^k)$ for $h(x, \mu) = \mu J(x)e$

as that of Newton's method but has the property that all the components of the iterates and of the residuals converge to their limit at comparable rates. This property may be desirable in a number of practical applications. The results are illustrated on an example which exhibits a normwise Q-quadratic convergence which does not occur componentwise, neither in the iterates nor in the residuals, when a pure Newton scheme is applied. The method relies on the computation of a suitable starting point for the modified Newton iteration, which turns out to readily satisfy the stopping conditions. Asymptotically, a single step is then sufficient and leads to a componentwise Q-superlinear convergence rate, which may be as close to quadratic as desired. It is noticeable that the results do not depend on how the exponents θ_ϵ^k, θ_μ^k and θ_η^k are updated, merely than they should satisfy **A6**. The method leaves freedom to the user in two major respects.

The first, and most important, is the choice of the perturbation function $h(x, \mu)$ and its possible dependence on x, subject to **A4**. The second is to permit inexact Jacobians. The method presented in this paper was inspired by similar behaviour in primal-dual interior-point methods [12] and exterior penalty methods [11] for constrained optimization. It is still an issue to determine whether a particular $h(x, \mu)$ is advantageous, or should be avoided; in particular, there might or might not exist a perturbation function $h(x, \mu)$ that depends on x and which is better than the usual perturbation in primal interior-point methods for linear and nonlinear programming. The application of the technique presented in this paper to nonlinear least-squares problems is also the object of continuing research.

Acknowledgments

The authors wish to thank John Dennis for his useful comments on earlier drafts of this work, Layne Watson for his insightful comments on homotopy methods, and two anonymous referees for useful remarks and improvements.

Appendix.

We justify the claims made in Example 1.1.

Theorem A.0.1. *A single Newton iteration for the nonlinear system (1.4) from the point* $x^k = x_\ell e_\ell$, *for some index* $1 \le \ell \le n$, *results in the point* x^{k+1} *satisfying (1.5).*

Proof. On writing (1.4) as

$$F(x) = \begin{pmatrix} x_1^2 + x_2 \\ x_2^2 + x_3 \\ \cdots \\ x_{n-1}^2 + x_n \\ x_n^2 + x_1 \end{pmatrix} = 0,$$

it follows that

$$J(x) = \begin{bmatrix} 2x_1 & 1 & & & \\ & 2x_2 & 1 & & \\ & & \cdot & \cdot & \\ & & & 2x_{n-1} & 1 \\ 1 & & & & 2x_n \end{bmatrix} = P + 2\sum_{i=1}^{n} x_i e_i e_i^T,$$

where P is the permutation matrix

$$P = \begin{bmatrix} & 1 & & & \\ & & 1 & & \\ & & & \cdot & \\ & & & & 1 \\ 1 & & & & \end{bmatrix}.$$

At the point of interest,

$$J(x_\ell e_\ell) = P + 2x_\ell e_\ell e_\ell^T$$

and hence the Sherman-Morrison-Woodbury formula along with the fact that P is a permutation, and thus that $P^{-1} = P^T$, reveals that

$$J^{-1}(x_\ell e_\ell) = P^T - \frac{2x_\ell}{1 + 2x_\ell e_\ell^T P^T e_\ell} P^T e_\ell (Pe_\ell)^T.$$

Note that

$$Pe_\ell = \begin{cases} e_n & \text{for } \ell = 1 \\ e_{l-1} & \text{for } 2 \le \ell \le n \end{cases} \quad \text{and} \quad P^T e_\ell = \begin{cases} e_{l+1} & \text{for } 1 \le \ell \le n-1 \\ e_1 & \text{for } \ell = n. \end{cases}$$

There are now three cases to consider.

(i). If $\ell = 1$,

$$J^{-1}(x_1 e_1) = P^T - \frac{2x_1}{1 + 2x_1 e_1^T e_2} e_2 e_n^T = P^T - 2x_1 e_2 e_n^T.$$

But, since $F(x_1 e_1) = x_1^2 e_1 + x_1 e_n$, the Newton correction is

$$-J^{-1}(x_1 e_1) F(x_1 e_1) = -(P^T - 2x_1 e_2 e_n^T)(x_1^2 e_1 + x_1 e_n) = x_1^2 e_2 - x_1 e_1,$$

and hence $x^{k+1} = x_1^2 e_2$.

(ii). Secondly, if $1 < \ell < n$,

$$J^{-1}(x_\ell e_\ell) = P^T - \frac{2x_\ell}{1 + 2x_\ell e_\ell^T e_{\ell+1}} e_{\ell+1} e_{\ell-1}^T = P^T - 2x_\ell e_{\ell+1} e_{\ell-1}^T.$$

Since $F(x_\ell e_\ell) = x_\ell^2 e_\ell + x_\ell e_{\ell-1}$, the Newton correction in this case is

$$-J^{-1}(x_\ell e_\ell) F(x_\ell e_\ell) = -(P^T - 2x_\ell e_{\ell+1} e_{\ell-1}^T)(x_\ell^2 e_\ell + x_\ell e_{\ell-1}) = x_\ell^2 e_{\ell+1} - x_\ell e_\ell,$$

and hence $x^{k+1} = x_\ell^2 e_{\ell+1}$.

(iii). Finally, if $\ell = n$,

$$J^{-1}(x_n e_n) = P^T - \frac{2x_n}{1 + 2x_n e_n^T e_1} e_1 e_{n-1}^T = P^T - 2x_n e_1 e_{n-1}^T.$$

Since $F(x_n e_n) = x_n^2 e_n + x_n e_{n-1}$, the Newton correction is now

$$-J^{-1}(x_n e_n) F(x_n e_n) = -(P^T - 2x_n e_1 e_{n-1}^T)(x_n^2 e_n + x_n e_{n-1}) = x_n^2 e_1 - x_n e_n,$$

and hence $x^{k+1} = x_n^2 e_1$. ∎

References

1. E.L. Allgower and K. Georg. Continuation and path following. *Acta Numerica*, pages 1–64, 1992.

2. R.H. Byrd, G. Liu, and J. Nocedal. On the local behavior of an interior point method for nonlinear programming. In D.F. Griffiths and D.J. Higham, editors, *Numerical Analysis 1997*, pages 37–56. Addison Wesley Longman, 1997. Proceedings of the Dundee 1997 Conference on Numerical Analysis.
3. R. Cominetti and J.-P. Dussault. Stable exponential-penalty algorithm with superlinear convergence. *Journal of Optimization Theory and Applications*, 83(2):285–309, 1994.
4. R. Cominetti and J.M. Pérez-Cerda. Quadratic rate of convergence for a primal-dual exponential penalty algorithm. *Optimization*, 39(1):13–32, 1997.
5. A.R. Conn, N.I.M. Gould, and Ph.L. Toint. *Trust-Region Methods*. SIAM, Philadelphia, USA, 2000.
6. R.S. Dembo, S.C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19(2):400–408, 1982.
7. R.S. Dembo and T. Steihaug. Truncated-Newton algorithms for large-scale unconstrained optimization. *Mathematical Programming*, (26):190–212, 1983.
8. J. E. Dennis and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Prentice-Hall, Englewood Cliffs, USA, 1983. Reprinted as *Classics in Applied Mathematics 16*, SIAM, Philadelphia, USA.
9. J.-P. Dussault. Numerical stability and efficiency of penalty algorithms. *SIAM Journal on Numerical Analysis*, 32(1):296–317, 1995.
10. A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. J. Wiley and Sons, Chichester, England, 1968. Reprinted as *Classics in Applied Mathematics*, SIAM, Philadelphia, USA, 1990.
11. N.I.M. Gould. On the convergence of a sequential penalty function method for constrained minimization. *SIAM Journal on Numerical Analysis*, 26(1):107–126, 1989.
12. N.I.M. Gould, D. Orban, A. Sartenaer, and Ph.L. Toint. Superlinear Convergence of Primal-Dual Interior Point Algorithms for Nonlinear Programming. *SIAM Journal on Optimization*, 11(4):974–1002, 2001.
13. J.Nocedal and S.J. Wright. *Numerical Optimization*. Springer series in Operations Research. Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA, 1999.
14. B.N. Lundberg and A.B. Poore. Numerical continuation and singularity detection methods for parametric nonlinear programming. *SIAM Journal on Optimization*, 3(1):134–154, 1993.
15. Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, USA, 1994.
16. J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. Academic Press, New York, 1970.
17. A.B. Poore and C.A. Thiart. Bifurcation problems in nonlinear parametric programming. *Mathematical Programming*, (39):189–205, 1987.
18. J.R.J. Rao and R. Bradlaw. A unified characterization of nonunique response in elastic/perfectly-plastic and elastic/locking structures using constraint qualifications. Department of Mechanical Engineering, University of Houston, Houston, TX.
19. W.C. Rheinboldt. *Numerical Analysis of Parametrized Nonlinear Equations*, volume 7 of *The University of Arkansas Lecture Notes in the Mathematical Sciences*. J. Wiley and Sons, 1986.

20. M. Sosonkina, L.T. Watson, and D.E. Stewart. Note on the end game in homotopy zero curve tracking. *Transactions of the ACM on Mathematical Software*, 22(3):281–287, 1996.

21. C.A. Thiart and A.B. Poore. A bifurcation analysis of the nonlinear programming problem. *Mathematical Programming*, (47):117–141, 1990.

22. M.C. Villalobos, R.A. Tapia, and Y. Zhang. The Sphere of Convergence of Newton's Method on Two Equivalent Systems from Nonlinear Programming. Technical Report TR9913, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA, 1999.

23. L.T. Watson. Solving the nonlinear complementarity problem by a homotopy method. *SIAM Journal on Control and Optimization*, 17:36–46, 1979.

24. L.T. Watson. Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review*, 28(4):529–545, 1986.

25. L.T. Watson. Globally convergent homotopy methods: A tutorial. *Appl. Math. Comput.*, 31BK:369–396, 1989.

26. L.T. Watson. Globally convergent homotopy algorithms for nonlinear systems of equations. *Nonlinear Dynamics*, 1:143–191, 1990.

27. L.T. Watson. A survey of probability-one homotopy methods for engineering optimization. *The Arabian Journal for Science and Engineering*, 16(2B):297–323, 1991.

28. L.T. Watson. Theory of globally-convergent probability-one homotopies for nonlinear programming. *SIAM Journal on Optimization*, 11(3):761–780, 2000.

29. L.T. Watson, J.P. Bixler, and A.B. Poore. Continuous homotopies for the linear complementarity problem. *SIAM Journal on Matrix Analysis and Applications*, 10(2):259–277, 1989.

30. M.H. Wright. Interior methods for constrained optimization. *Acta Numerica*, 1:341–407, 1992.

31. M.H. Wright. Why a pure primal Newton barrier step may be infeasible. *SIAM Journal on Optimization*, 5(1):1–12, 1995.

32. M.H. Wright. The Interior-Point Revolution in Constrained Optimization. In R. DeLeone, A. Murli, P.M. Pardalos, and G. Toraldo, editors, *High-Performance Algorithms and Software in Nonlinear Optimization*, pages 359–381, Dordrecht, 1998. Kluwer Academic Publishers.

33. S. J. Wright and F. Jarre. The role of linear objective functions in barrier methods. *Mathematical Programming*, 84(2):357–373, 1999.

34. S.J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA, 1997.

35. W. I. Zangwill and C. B. Garcia. *Pathways to solutions, fixed points and equilibria*. Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1981.