Adv Syst Sci Appl 2024; **01**:29–39
Published online at https://ijassa.ipu.ru.

The Piecewise Levenberg–Marquardt Method

Alexey F. Izmailov[1*], Evgeniy I. Uskov[2], Yan Zhibai[1]

[1]*Lomonosov Moscow State University, Moscow, Russia*
[2]*Derzhavin Tambov State University, Tambov, Russia*

Abstract: We develop sharp local superlinear convergence results for the Levenberg–Marquardt method applied to constrained piecewise smooth equations, under the local Lipschitzian error bound condition for active selections, allowing for nonisolated solutions. We also characterize the level of inexactness allowed when solving subproblems, such that it does not interfere with superlinear convergence rate. Applications to constrained reformulations of complementarity systems, using the "min" complementarity function are also discussed.

Keywords: nonlinear equation, constrained equation, piecewise smooth equation, nonisolated solution, local error bound, Levenberg–Marquardt method

1. INTRODUCTION

We consider the constrained equation

$$\Phi(u) = 0, \quad u \in P, \tag{1.1}$$

where $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ is a given mapping, and $P \subset \mathbb{R}^p$ is a given nonempty closed convex set. Let U stand for the solution set of (1.1).

The case of $P = \mathbb{R}^p$ and $p = q$ is the classical setting of an unconstrained nonlinear equation, and assuming that Φ is smooth, the fundamental approach for solving such problems is the Newton method; see, e.g., [12]. Local superlinear convergence of the basic form of the Newton method requires a solution in question to be nonsingular, and in particular isolated. The Levenberg–Marquardt (LM) method for unconstrained nonlinear equations [17, 19] (see also [20, Section 10.3]) is a classical regularization technique for handling cases when a solution can be singular, and possibly nonisolated, and possibly with $p \neq q$.

A natural way to approach the LM method is to consider first the (constrained) Gauss–Newton (GN) method. For the current iterate $u \in P$, the GN method defines the next iterate as $u + v$, where v minimizes the (squared) residual of the linearized equation from (1.1) over $P - u$, i.e., v is a solution of the optimization problem

$$\text{minimize} \quad \frac{1}{2}\|\Phi(u) + \Phi'(u)v\|^2 \quad \text{subject to} \quad u + v \in P. \tag{1.2}$$

Due to the Frank–Wolfe Theorem [10], this subproblem is always solvable when P is polyhedral, but a solution need not be unique.

The potential lack of uniqueness of solutions in the subproblem (1.2) is one of the reasons regularize it. This results in the constrained LM method [2, 15] for (1.1), defining the next

*Corresponding author: izmaf@cs.msu.ru

iterate as $u+v$, where v solves the optimization problem

$$\text{minimize} \quad \frac{1}{2}\|\Phi(u)+G(u)v\|^2+\frac{1}{2}\sigma(u)\|v\|^2 \quad \text{subject to} \quad u+v\in P, \tag{1.3}$$

with a function $\sigma : P \to \mathbb{R}_+$ defining the values of the regularization parameter, and with $G(u)$ being $\Phi'(u)$, or its suitable substitute when the derivative does not exist or is not available. Observe that if $\sigma(u) > 0$, and if Euclidean norms are used, the objective function of (1.3) is strongly convex quadratic, and hence, this subproblem has the unique solution. Moreover, if P is a polyhedral set, then (1.3) is a strongly convex quadratic programming problem.

When $P = \mathbb{R}^p$, the constrained LM method reduces to the classical (sometimes called unconstrained) LM method with its origins in [17, 19].

For the convergence properties for $P = \mathbb{R}^p$ in the case of smooth Φ and isolated solutions, as well as related references, see [3, Theorem 10.2.6]. The cases of nonsmooth (semismooth) Φ with isolated solutions was considered in [5]; see also [13, 14].

In this work we deal with local convergence analysis based on weaker assumptions of the local Lipschitzian error bound type allowing, in particular, for nonisolated solutions, and originating in the context of LM methods from [21]. Moreover, we concentrate on the peculiarities of the case when Φ in (1.1) is piecewise smooth, but need not be smooth. As argued in [4], the combination of nonisolated solutions with nonsmoothness is an especially challenging situation.

For a recent survey of modern local convergence theories for the LM method, including some applications, the issues of convergence globalization, etc., see [8].

Some comments on our notation are in order. The Euclidean inner product for $u, v \in \mathbb{R}^p$ is denoted by $\langle u, v\rangle$, and to avoid any confusion, let $\|\cdot\|$ stand for the Euclidean norm throughout. For a set $U \subset \mathbb{R}^p$ and a point $u \in \mathbb{R}^p$, $\mathrm{dist}(x, U) = \inf_{v\in U}\|v-u\|$, and $B(u, \delta) = \{v \in \mathbb{R}^p \mid \|v-u\| \leq \delta\}$ is the closed ball of radius $\delta \geq 0$ centered at u. For a given index set $H \subset \{1, \ldots, p\}$, by u_J we denote the subvector of $u \in \mathbb{R}^p$, with components u_j, $j \in J$.

For a closed convex set $U \subset \mathbb{R}^p$, by $N_U(u)$ we denote the normal cone to U at u, i.e., $N_U(u) = \{v \in \mathbb{R}^p \mid \langle v, \widetilde{u}-u\rangle \leq 0,\ \forall\, \widetilde{u} \in U\}$ if $u \in U$, and $N_U(u) = \emptyset$ otherwise.

For a sequence $\{u^k\} \subset \mathbb{R}^p$ convergent to some $\bar{u} \in \mathbb{R}^p$, we say that convergence is of Q-order $\theta > 1$ if there exists $c > 0$ such that $\|u^{k+1}-\bar{u}\| \leq c\|u^k-\bar{u}\|^\theta$ for all k large enough. Such rate of convergence is superlinear, and it is at least quadratic if $\theta \geq 2$. We say that $\{u^k\}$ converges to $\bar{u}$ with R-order θ if there exist $c > 0$ and a sequence $\{t_k\} \subset \mathbb{R}_+$ converging to 0 with Q-order θ, such that $\|u^{k+1}-\bar{u}\| \leq ct_k$ for all k large enough.

The rest of this paper is organized as follows. In Section 2, we recall the local convergence framework from [9], that will serve as a main tool in our analysis. Section 3 contains our new local convergence result for the piecewise smooth case, and Section 4 is devoted to the effect of approximate solution of subproblems. Finally, in Section 5, we discuss applications of the results obtained to constrained reformulations of complementarity systems, using the “min” complementarity function.

2. LOCAL CONVERGENCE FRAMEWORK

We start with a discussion of the abstract local convergence framework recently proposed in [9]. For that purpose, it is convenient to consider a problem setting with a single scalar equation:

$$\varphi(u) = 0, \quad u \in P, \tag{2.4}$$

with $\varphi : \mathbb{R}^p \to \mathbb{R}_+$ being a given scalar-valued function, and $P \subset \mathbb{R}^p$ being a given nonempty closed set. Obviously, the constrained equation (1.1) can be stated in the form (2.4) with, say,

$\varphi(u) = \|\Phi(u)\|$. To that end, there will be no confusion if in this section we will use U for the solution set of (2.4).

Let $\Psi : P \to P$ be a given mapping, and consider an abstract iterative process updating the current iterate $u \in P$ to the new one of the form $\Psi(u)$. The following is a version of [9, Theorem 2.1], somehow simplified for our purposes.

Theorem 2.1:
Let $\varphi : \mathbb{R}^p \to \mathbb{R}_+$ be a continuous function, $P \subset \mathbb{R}^p$ be a nonempty closed set, $\bar{u} \in U$, and assume that

$$\varphi(u) = O(\mathrm{dist}(u, U)) \tag{2.5}$$

as $u \in P$ tends to $\bar{u}$.

Moreover, let $\Psi : P \to P$ be a mapping such that, with some $\tau > 1$,

$$\Psi(u) - u = O(\varphi(u)) \tag{2.6}$$

and

$$\varphi(\Psi(u)) = O((\varphi(u))^\tau) \tag{2.7}$$

as $u \in P$ tends to $\bar{u}$.

Then, for every $\delta > 0$, and every $u^0 \in P$ close enough to $\bar{u}$, the sequence $\{u^k\}$ defined by $u^{k+1} = \Psi(u^k)$ for all k is contained in $B(\bar{u}, \delta)$ and converges to some $u^* \in U$, and the rate of convergence is superlinear with the Q-order τ.

3. THE PIECEWISE SMOOTH CASE: LOCAL CONVERGENCE

Our main focus in this work is on the case when Φ in (1.1) is piecewise smooth near a solution $\bar{u}$ in question: for a finite collection of selection mappings $\Phi^1, \ldots, \Phi^s : \mathbb{R}^p \to \mathbb{R}^q$ which are continuously differentiable near $\bar{u}$, it holds that

$$\Phi(u) \in \{\Phi^1(u), \ldots, \Phi^s(u)\} \quad \forall\, u \in \mathbb{R}^p,$$

and Φ is continuous near $\bar{u}$. Taking $s = 1$ recovers the smooth case, i.e. the case when Φ is continuously differentiable near $\bar{u}$. According to [11, Theorem 2.1], any mapping that is piecewise smooth near some point is necessarily Lipschitz-continuous near that point.

For every $u \in \mathbb{R}^p$, define the set

$$\mathcal{A}(u) = \{j \in \{1, \ldots, s\} \mid \Phi(u) = \Phi^j(u)\} \tag{3.8}$$

of indices of all selection mappings active at u. Let G be any mapping such that

$$G(u) \in \{(\Phi^j)'(u) \mid j \in \mathcal{A}(u)\} \tag{3.9}$$

for $u \in \mathbb{R}^p$ close enough to $\bar{u}$. The algorithm with the subproblem (1.3) employing G defined by (3.8) and (3.9), is naturally referred to as the constrained piecewise LM method. Observe that with this definition of G, for any $u \in \mathbb{R}^p$ close enough to $\bar{u}$, the subproblem (1.3) can be written in the form

$$\text{minimize} \quad \frac{1}{2}\|\Phi^j(u) + (\Phi^j)'(u)v\|^2 + \frac{1}{2}\sigma(u)\|v\|^2 \quad \text{subject to} \quad u + v \in P, \tag{3.10}$$

with *some* $j \in \mathcal{A}(u)$, and this can be seen as the subproblem of the LM method applied to a *smooth* constrained equation

$$\Phi^j(u) = 0, \quad u \in P. \tag{3.11}$$

Needless to say, the index j can vary from one iteration to another.

We will need to assume the P-*property* at $\bar{u}$, introduced in [7, p. 434]. It consists of saying that near $\bar{u}$ the constraint set P excludes all zeroes of smooth selections active at $\bar{u}$, which are not zeroes of Φ. Formally it means that $U_j \subset U$ near $\bar{u}$ for all $j \in \mathcal{A}(\bar{u})$, where U_j stands for the solution set of (3.11). The P-property at $\bar{u}$ evidently implies that

$$\mathrm{dist}(u,\, U) \leq \mathrm{dist}(u,\, U_j) \tag{3.12}$$

for all $u \in P$ close enough to $\bar{u}$. We note that the P-property can be easily guaranteed for reformulations of complementarity systems by means of the "min" (natural residual) complementarity function, see [7] and Section 5 below.

Another assumption employed in Theorem 3.1 below is the constrained error bound for each active selection, that is

$$\mathrm{dist}(u,\, U_j) = O(\|\Phi^j(u)\|) \quad \forall\, j \in \mathcal{A}(\bar{u}) \tag{3.13}$$

as $u \in P$ tends to $\bar{u}$. Under this assumption, the P-property at $\bar{u}$ is equivalent to the condition

$$\Phi(u) = O(\|\Phi^j(u)\|) \quad \forall\, j \in \mathcal{A}(\bar{u})$$

as $u \in P$ tends to $\bar{u}$. Observe also that, employing the inclusion $\mathcal{A}(u) \subset \mathcal{A}(\bar{u})$ for $u \in \mathbb{R}^p$ close enough to $\bar{u}$, one can easily see that the P-property at $\bar{u}$, and the condition (3.13) as $u \in P$ tends to $\bar{u}$, imply the constrained error bound

$$\mathrm{dist}(u,\, U) = O(\|\Phi(u)\|) \tag{3.14}$$

as $u \in P$ tends to $\bar{u}$.

We proceed with the following characterization of local superlinear convergence of the piecewise LM method, which is the main result of this section.

Theorem 3.1:
Let $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ be a given mapping, $P \subset \mathbb{R}^p$ a nonempty closed convex set, and $\bar{u} \in U$. Assume that Φ is piecewise smooth near $\bar{u}$, and the derivatives of its smooth selection mappings $\Phi^1, \ldots, \Phi^s : \mathbb{R}^p \to \mathbb{R}^q$ are Lipschitz-continuous near $\bar{u}$. Let the P-property at $\bar{u}$ and condition (3.13) *be satisfied as $u \in P$ tends to $\bar{u}$. Moreover, let $G : \mathbb{R}^p \to \mathbb{R}^{q\times p}$ be a fixed mapping satisfying* (3.9)*, and assume that the function $\sigma : P \to \mathbb{R}_+$ defining the regularization parameter satisfies*

$$\|\Phi(u)\|^\theta = O(\sigma(u)), \quad \sigma(u) = O(\|\Phi(u)\|^\theta) \tag{3.15}$$

as $u \in P$ tends to $\bar{u}$, with some fixed $\theta \in (0,\, 2]$.

Then, for every $u^0 \in P$, there exists the unique sequence $\{u^k\}$ such that for every k, the displacement $u^{k+1} - u^k$ is a solution of (1.3) *with $u = u^k$, and with the additional convention that $u^{k+1} = u^k$ if $u^k \in U$. For any $\delta > 0$, if $u^0 \in P$ is close enough to $\bar{u}$, then this sequence is contained in $B(\bar{u},\, \delta)$ and converges to some $u^* \in U$, and the rate of convergence superlinear with the Q-order* $\min\{\theta + 1,\, 2\}$.

Proof
The first estimate in (3.15) implies that $\sigma(u) > 0$ for all $u \in P \setminus U$ close enough to $\bar{u}$, and hence, for such u, the LM subproblem (1.3) has the unique solution $v(u)$. Furthermore, for $u \in \mathbb{R}^p$ close to $\bar{u}$, and for $j \in \mathcal{A}(u) \subset \mathcal{A}(\bar{u})$ such that $G(u) = (\Phi^j)'(u)$ (see (3.9)), we have that $\Phi^j(u) = \Phi(u)$ (see (3.8)). For any such j, (3.15) implies the estimates

$$\|\Phi^j(u)\|^\theta = O(\sigma(u)), \quad \sigma(u) = O(\|\Phi^j(u)\|^\theta) \tag{3.16}$$

as $u \in P$ tends to $\bar{u}$, and an iteration of the constrained piecewise LM method can be interpreted as an iteration of the usual constrained LM method for the smooth constrained equation (3.11), with the subproblem (3.10), and with the regularization parameter satisfying (3.16).

Employing the restriction $\theta \in (0,\, 2]$, from (3.13) and from the first estimate in (3.16) it follows that

$$\frac{(\mathrm{dist}(u,\, U_j))^4}{\sigma(u)} = (\mathrm{dist}(u,\, U_j))^2 O\left(\frac{\|\Phi^j(u)\|^2}{\sigma(u)}\right) = O((\mathrm{dist}(u,\, U_j))^2) \tag{3.17}$$

as $u \in P \setminus U$ tends to $\bar{u}$.

Let $\widehat{u}^j$ stand for a metric projection of u onto U_j:

$$\|u - \widehat{u}^j\| = \mathrm{dist}(u,\, U_j). \tag{3.18}$$

Since $v(u)$ is a global solution of (3.10), it holds that

$$\begin{aligned}\|\Phi^j(u) + (\Phi^j)'(u)v(u)\|^2 + \sigma(u)\|v(u)\|^2 \;\; \leq \;\; & \|\Phi^j(u) + (\Phi^j)'(u)(\widehat{u}^j - u)\|^2 \\ & +\sigma(u)\|\widehat{u}^j - u\|^2.\end{aligned} \tag{3.19}$$

Since $j \in \mathcal{A}(\bar{u})$, it holds that $\bar{u} \in U_j$. Therefore, evidently, $\widehat{u}^j \to \bar{u}$ as $u \to \bar{u}$, and from (3.18)–(3.19), applying again the Mean-Value Theorem [12, Theorem A.10], we derive that

$$\begin{aligned}\|v(u)\|^2 \;\; & \leq \;\; \frac{1}{\sigma(u)}\left(\|\Phi^j(u) + (\Phi^j)'(u)(\widehat{u}^j - u)\|^2 + \sigma(u)\|\widehat{u}^j - u\|^2\right) \\ & = \;\; \frac{1}{\sigma(u)}\|\Phi^j(u) - \Phi^j(\widehat{u}^j) - (\Phi^j)'(u)(u - \widehat{u}^j)\|^2 + \|u - \widehat{u}^j\|^2 \\ & \leq \;\; \frac{1}{\sigma(u)} \sup_{\tau \in [0,\, 1]} \|(\Phi^j)'(\tau u + (1-\tau)\widehat{u}^j) - (\Phi^j)'(u)\|^2\|u - \widehat{u}^j\|^2 + \|u - \widehat{u}^j\|^2 \\ & = \;\; \|u - \widehat{u}^j\|^2 + O\left(\frac{\|u - \widehat{u}^j\|^4}{\sigma(u)}\right) \\ & = \;\; O((\mathrm{dist}(u,\, U_j))^2),\end{aligned} \tag{3.20}$$

where the last estimate is by (3.17). Hence, employing again (3.13),

$$v(u) = O(\mathrm{dist}(u,\, U_j)) = O(\|\Phi^j(u)\|) = O(\|\Phi(u)\|) \tag{3.21}$$

as $u \in P \setminus U$ tends to $\bar{u}$.

Furthermore, any $u \in U_j$ is a (global) solution of the optimization problem

$$\text{minimize} \quad \frac{1}{2}\|\Phi^j(u)\|^2 \quad \text{subject to} \quad u \in P,$$

and the objective function of this problem is differentiable at u, with the gradient being $((\Phi^j)'(u))^\top \Phi^j(u)$. Therefore, any such u must satisfy the first-order necessary optimality condition

$$((\Phi^j)'(u))^\top \Phi^j(u) + N_P(u) \ni 0. \tag{3.22}$$

Employing now [2, Lemma 1] demonstrating that the constrained error bound implies the upper Lipschitzian property of the solutions set of the generalized equation (3.22) subject to the right-hand side perturbations, we obtain that under (3.13), for any solution u of the perturbed generalized equation

$$((\Phi^j)'(u))^\top \Phi^j(u) + N_P(u) \ni \omega, \tag{3.23}$$

close enough to $\bar{u}$, it holds that

$$\mathrm{dist}(u,\, U_j) = O(\|\omega\|)$$

as $\omega \to 0$.

The first-order necessary optimality condition for the subproblem (3.10) has the form

$$((\Phi^j)'(u))^\top (\Phi^j(u) + (\Phi^j)'(u)v) + \sigma(u)v + N_P(u+v) \ni 0. \tag{3.24}$$

For any given $u \in \mathbb{R}^p$, (3.22) is equivalent to saying that

$$((\Phi^j)'(u+v))^\top \Phi^j(u+v) + N_P(u+v) \ni 0 \tag{3.25}$$

holds with $v = 0$. Then (3.24) can be regarded as a perturbation of the generalized equation (3.25). Specifically, if we set

$$\begin{aligned} \omega(u,\, v) &= ((\Phi^j)'(u+v))^\top \Phi^j(u+v) - ((\Phi^j)'(u))^\top (\Phi^j(u) + (\Phi^j)'(u)v) - \sigma(u)v \\ &= \left(((\Phi^j)'(u+v))^\top - ((\Phi^j)'(u))^\top\right) \Phi^j(u+v) \\ &\quad + ((\Phi^j)'(u))^\top (\Phi^j(u+v) - \Phi^j(u) - (\Phi^j)'(u)v) \\ &\quad - \sigma(u)v, \end{aligned} \tag{3.26}$$

then (3.24) can be written in the form

$$((\Phi^j)'(u+v))^\top \Phi^j(u+v) + N_P(u+v) \ni \omega(u,\, v).$$

From Lipschitz-continuity of $(\Phi^j)'$ near $\bar{u}$, and from the second estimate in (3.15) and (3.21), employing again the Mean-Value Theorem [12, Theorem A.10], one can readily derive the estimates

$$\begin{aligned} \omega(u,\, v(u)) &= O(\sigma(u)\|v(u)\|) + O(\|v(u)\|^2) \\ &= O(\|\Phi(u)\|^{\theta+1}) + O(\|\Phi^j(u)\|^2) \\ &= O(\|\Phi(u)\|^{\min\{\theta+1,\, 2\}}) \end{aligned} \tag{3.27}$$

as $u \in P \setminus U$ tends to $\bar{u}$.

Summarizing the considerations above, $u + v(u)$ is a solution of the generalized equation (3.23) with $\omega = \omega(u,\, v(u))$, and it holds that $u + v(u) \to \bar{u}$ and $\omega(u,\, v(u)) \to 0$ as $u \in P \setminus U$ tends to $\bar{u}$. Therefore, [2, Lemma 1] allows to conclude that

$$\mathrm{dist}(u + v(u),\, U_j) = O(\omega(u,\, v(u))) = O(\|\Phi(u)\|^{\min\{\theta+1,\, 2\}}),$$

where the second estimate is by (3.27). Therefore, since Φ is Lipschitz-continuous near $\bar{u}$, employing the inequality (3.12) (following from the P-property) we derive

$$\Phi(u + v(u)) = O(\mathrm{dist}(u + v(u),\, U)) = O(\mathrm{dist}(u + v(u),\, U_j)) = O(\|\Phi(u)\|^{\min\{\theta+1,\, 2\}}) \tag{3.28}$$

as $u \in P \setminus U$ tends to $\bar{u}$.

Estimates (3.21) and (3.28) yield (2.6) and (2.7) in Theorem 2.1 with $\Psi(u) = u + v(u)$, $\varphi(u) = \|\Phi(u)\|$, and $\tau = \min\{\theta + 1,\, 2\}$, while (2.5) is satisfied again because of Lipschitz continuity of Φ. The needed conclusion now follows by application of Theorem 2.1. □

For the case when $\theta = 2$, the result of Theorem 3.1 was essentially obtained in [7, Theorems 1, 2]; see also [8, Theorem 4.1], where it is emphasized that the convexity assumption on P is actually not needed in that case. On the other hand, in the smooth case, Theorem 3.1 is a particular case of [2, Theorem 1].

As demonstrated in [8, Example 3.2], the convergence rate estimate in Theorem 3.1 is sharp even in the smooth unconstrained case, and even when the solution $\bar{u}$ in question is isolated. This example also shows that the constrained error bound condition (3.14), corresponding to (3.13) in the smooth case, is essential for this analysis.

Regarding the restriction $\theta \in (0,\, 2]$ in Theorem 3.1 on the exponent in the regularization parameter, [8, Example 3.2] demonstrates that at least the requirement $\theta < 4$ cannot be avoided, even in the smooth case. The possibility to obtain a counterpart of Theorem 3.1 for the values $\theta \in (2,\, 4)$ seems to remain an open question. The lack of quadratic convergence is claimed in [1, Example 4.2] for $\theta > 3$, but this claim relies on some numerical observations only. In the smooth case, superlinear convergence to 0 of the sequence $\{\mathrm{dist}(u^k,\, U)\}$ for $\theta \in (2,\, 3)$ was established in [6], with some similar results and extensions in the recent work [22].

4. THE PIECEWISE SMOOTH CASE: EFFECT OF INEXACTNESS

Since solving the constrained LM subproblems exactly can be too expensive, or even impossible, a natural issue consists of characterization of the "level" of controllable inexactness when solving subproblems that does not interfere with local convergence and rate of convergence properties of the LM method established in Section 3.

Yet again, for $u \in \mathbb{R}^p$ close to $\bar{u}$, let $j \in \mathcal{A}(u) \subset \mathcal{A}(\bar{u})$ be such that $G(u) = (\Phi^j)'(u)$ (see (3.9)); it also necessarily holds that $\Phi^j(u) = \Phi(u)$ (see (3.8)). Recall that an iteration of the constrained piecewise LM method can be interpreted as an iteration of the usual constrained LM method for the smooth constrained equation (3.11), with the subproblem (3.10), and with the regularization parameter satisfying (3.16).

The first-order necessary optimality condition for (3.10) has the form (3.24), and it is natural to consider the version of the inexact constrained LM method, with inexactness measured by the violation of (3.24). Specifically, the process of solving the subproblem (3.10) is terminated once

$$((\Phi^j)'(u))^\top(\Phi^j(u) + (\Phi^j)'(u)v) + \sigma(u)v + N_P(u+v) \ni w \tag{4.29}$$

is satisfied with some $w \in \mathbb{R}^p$ smaller than the given tolerance that conforms with the exponent θ in (3.15) (and in (3.16)) as follows:

$$w = O(\|\Phi^j(u)\|^{\theta+1}) \tag{4.30}$$

as $u \in P$ tends to $\bar{u}$.

The analysis in [2] relies on the observation that (4.29) is a necessary and sufficient optimality condition for the following convex optimization problem, which is a perturbation of (3.10):

$$\text{minimize} \quad \frac{1}{2}\|\Phi^j(u) + (\Phi^j)'(u)v\|^2 + \frac{1}{2}\sigma(u)\|v\|^2 - \langle w,\, v\rangle \quad \text{subject to} \quad u + v \in P. \tag{4.31}$$

Then we follow the argument used above to prove (3.21), but taking into account that now the objective function of (4.31) has an extra term involving w.

Employing again a metric projection $\widehat{u}^j$ of $u \in P \setminus U$ onto U_j, for the unique global solution $v(u)$ of (4.31), similarly to (3.19) we obtain that

$$\begin{aligned}\|\Phi^j(u) + (\Phi^j)'(u)v\|^2 + \sigma(u)\|v(u)\|^2 - 2\langle w,\, v(u)\rangle \;\; &\le \;\; \|\Phi^j(u) + (\Phi^j)'(u)(\widehat{u}^j - u)\|^2 \\ &\quad + \sigma(u)\|\widehat{u}^j - u\|^2 - 2\langle w,\, \widehat{u}^j - u\rangle.\end{aligned}$$

Then, similarly to (3.20) (and in particular, making use of (3.18)), we derive the estimate

$$\begin{aligned}\|v(u)\|^2 &\leq \frac{1}{\sigma(u)}\big(\|\Phi^j(u)+(\Phi^j)'(u)(\widehat{u}^j-u)\|^2+\sigma(u)\|\widehat{u}^j-u\|^2\\ &\quad +2\langle w,\, v(u)\rangle-2\langle w,\, \widehat{u}^j-u\rangle\big)\\ &\leq \frac{2\|w\|}{\sigma(u)}(\|v(u)\|+\operatorname{dist}(u,\, U_j))+O((\operatorname{dist}(u,\, U_j))^2)\\ &\leq O(\|\Phi^j(u)\|(\|v(u)\|+\operatorname{dist}(u,\, U_j)))+O((\operatorname{dist}(u,\, U_j))^2)\\ &= O(\operatorname{dist}(u,\, U_j)\|v(u)\|)+O((\operatorname{dist}(u,\, U_j))^2) \end{aligned} \tag{4.32}$$

as $u \in P \setminus U$ tends to $\bar{u}$, where the next-to-the-last estimate is by (4.30) and the first relation in (3.16), and the last one is the Lipschitz continuity of Φ^j near $\bar{u}$. Evidently, (4.32) implies (3.21) as $u \in P \setminus U$ tends to $\bar{u}$.

The remaining part of the argument consists of the reasoning used above to prove Theorem 3.1, but with $\omega(u,\, v)$ defined in (3.26) replaced by $\omega(u,\, v) + w$: taking again into account (4.30), one can see that this modification does not affect (3.27) and the subsequent analysis. This yields

Theorem 4.1:
Under the assumptions of Theorem 3.1, let the function $\psi : P \to \mathbb{R}_+$ satisfy $\psi(u) = O(\|\Phi(u)\|^{\theta+1})$ as $u \in P$ tends to $\bar{u}$.

Then, for every $u^0 \in P$, there exists a sequence $\{u^k\}$ such that for every k, the displacement $u^{k+1} - u^k$ is the solution of (4.29) *with $u = u^k$, with some $w \in \mathbb{R}^p$ satisfying $\|w\| \leq \psi(u^k)$, and with the additional convention that $u^{k+1} = u^k$ if $u^k \in U$. For any $\delta > 0$, if $u^0 \in P$ is close enough to $\bar{u}$, any such sequence is contained in $B(\bar{u},\, \delta)$ and converges to some $u^* \in U$, and the rate of convergence is superlinear with the Q-order* $\min\{\theta + 1,\, 2\}$.

For the case when $\theta = 2$, a result related to Theorem 4.1 was obtained in [4, Theorems 3], but for a somewhat different kind of inexactness. See also [8, Section 5] for a discussion of these issues in the smooth case, including establishing the relations between these two kinds of inexactness.

The requirement (4.30) on allowed inexactness in Theorem 4.1 cannot be relaxed even in the smooth case, as demonstrated by an example in [2, Section 5].

5. APPLICATION TO REFORMULATIONS OF COMPLEMENTARITY SYSTEMS

In this section we deal with a nonlinear complementarity system of the form

$$a(x) = 0, \quad b(x) \geq 0, \quad c(x) \geq 0, \quad d(x) \geq 0, \quad \langle c(x),\, d(x)\rangle = 0, \tag{5.33}$$

with given smooth mappings $a : \mathbb{R}^n \to \mathbb{R}^l$, $b : \mathbb{R}^n \to \mathbb{R}^m$, $c : \mathbb{R}^n \to \mathbb{R}^r$, $d : \mathbb{R}^n \to \mathbb{R}^r$, i.e., possessing Lipschitz-continuous derivatives near a solution $\bar{x}$ of (5.33). Various important problem classes can be modeled as (5.33), including the so-called (mixed) complementarity problems, and in particular, Karush–Kuhn–Tucker optimality systems, to mention some.

One of the most successful approaches to (5.33) consists of reformulating it as a system of equations by means of the so-called complementarity functions, perhaps the simples (but at the same time prominent) one being the “min” (natural residual) complementarity function: using auxiliary slack variables, system (5.33) can be equivalently written as (1.1) with $p = n + 2r + m$, $q = l + 3r + m$,

$$\Phi(u) = (a(x),\, \min\{y,\, z\},\, c(x) - y,\, d(x) - z,\, b(x) - \mu), \quad u = (x,\, y,\, z,\, \mu), \tag{5.34}$$

$$P = \mathbb{R}^n \times \mathbb{R}^r_+ \times \mathbb{R}^r_+ \times \mathbb{R}^m_+. \tag{5.35}$$

The mapping Φ defined in (5.34) is evidently piecewise smooth. More precisely, for a current iterate $u^k = (x^k,\, y^k,\, z^k,\, \mu^k) \in \mathbb{R}^p$, the smooth selections active at u^k have the form

$$\Phi^J(u) = (a(x),\, y_{I_c(u^k)\cap J},\, z_{I_d(u^k)\cap(I_0(u^k)\setminus J)},\, c(x) - y,\, d(x) - z,\, b(x) - \mu), \tag{5.36}$$

where

$$\begin{array}{rcl} I_c(u^k) & = & \{i \in \{1,\, \ldots,\, r\} \mid y_i^k < z_i^k\}, \\ I_d(u^k) & = & \{i \in \{1,\, \ldots,\, r\} \mid y_i^k > z_i^k\}, \\ I_0(u^k) & = & \{i \in \{1,\, \ldots,\, r\} \mid y_i^k = z_i^k\}, \end{array}$$

and $J \subset I_0(u^k)$. Therefore, the iteration of the piecewise LM method consists of choosing such J, setting $G(u^k) = (\Phi^J)'(u^k)$, and solving (1.3) with $u = u^k$ for the displacement $v^k = u^{k+1} - u^k$.

According to (5.36), for any solution $\bar{x}$ of system (5.33), and for the corresponding solution $\bar{u} = (\bar{x},\, \bar{y},\, \bar{z},\, \bar{\mu})$ of problem (1.1) with Φ and P defined in (5.34) and (5.35), respectively, with $\bar{y} = c(\bar{x})$, $\bar{z} = d(\bar{x})$, and $\bar{\mu} = b(\bar{x})$, the smooth selections of Φ active at $\bar{u}$ have the form

$$\Phi^J(u) = (a(x),\, y_{I_c\cap J},\, z_{I_d\cap(I_0\setminus J)},\, c(x) - y,\, d(x) - z,\, b(x) - \mu), \tag{5.37}$$

where

$$\begin{array}{rcccl} I_c & = & I_c(\bar{u}) & = & \{i \in \{1,\, \ldots,\, r\} \mid c_i(\bar{x}) = 0 < d_i(\bar{x})\}, \\ I_d & = & I_b(\bar{u}) & = & \{i \in \{1,\, \ldots,\, r\} \mid c_i(\bar{x}) > 0 = d_i(\bar{x})\}, \\ I_0 & = & I_0(\bar{u}) & = & \{i \in \{1,\, \ldots,\, r\} \mid c_i(\bar{x}) = 0 = d_i(\bar{x})\}, \end{array}$$

and $J \subset I_0$.

Observe that the nonnegativity conditions on y and z in the definition of P in (5.35) may seem extraneus for this reformulation as they are guaranteed by the equality $\min\{y,\, z\} = 0$. However, these constraints appear essential for guaranteing the P property to hold at $\bar{u}$. Indeed, if for a given $J \subset I_0$, it holds that $\Phi^J(u) = 0$ for some $u \in P$, then according to (5.35), $y_{I_d\cap(I_0\setminus J)} \geq 0$ and $z_{I_c\cap J} \geq 0$, and hence, according to (5.34) and (5.37), it holds that $\Phi(u) = 0$. At the same time, if $u \in \mathbb{R}^p \setminus P$, then some components of, say, $z_{I_c\cap J}$ can be negative, and then $\min\{y_{I_c\cap J},\, z_{I_c\cap J}\} = \min\{0,\ z_{I_c\cap J}\}$ need not be equal to 0.

The remaining ingredient for applicability of Theorems 3.1 and 4.1 is the constrained local Lipschitzian error bound (3.13) for selections active at $\bar{u}$. For every $J \subset I_0$, let

$$X_J = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} a(x) = 0,\, c_{I_c\cap J}(x) = 0,\, d_{I_d\cap(I_0\setminus J)}(x) = 0, \\ b(x) \geq 0,\, c(x) \geq 0,\, d(x) \geq 0 \end{array} \right\}.$$

Then according to the discussion in [7, Section 5] (see, in particular, [7, Figure 3]), condition (3.13) is implied by the following: for every $J \subset I_0$, it holds that

$$\begin{array}{rcl} \operatorname{dist}(x,\, X_J) & = & O(\|a(x)\| + \|c_{I_c\cap J}(x)\| + \|d_{I_d\cap(I_0\setminus J)}(x)\| + \|\min\{0,\, b(x)\}\| \\ & & +\|\min\{0,\, c_{I_d\cap(I_0\setminus J)}(x)\}\| + \|\min\{0,\, d_{I_c\cap J}(x)\}\|) \end{array} \tag{5.38}$$

as $x \in \mathbb{R}^n$ tends to $\bar{x}$. Condition (5.38) is nothing else but the local Lipschitzian error bound for the constraint system

$$a(x) = 0,\;\; c_{I_c\cap J}(x) = 0,\;\; d_{I_d\cap(I_0\setminus J)}(x) = 0,\;\; b(x) \geq 0,\;\; c_{I_d\cap(I_0\setminus J)}(x) \geq 0,\;\; d_{I_c\cap J}(x) \geq 0, \tag{5.39}$$

characterizing a branch of the solution set of the complementarity system (5.33) near a given solution $\bar{x}$. Therefore, (3.13) is impled by the so-called piecewise error bound condition for (5.33), that is, the local Lipschitzian error bound at $\bar{x}$ for every branch of the solution set, containing $\bar{x}$, i.e., for the constraint system (5.39) with every $J \subset I_0$. The latter property is implied by the so-called piecewise constrained qualifications at $\bar{x}$, such as the classical Mangasarian–Fromovits constraint qualification [18], or the relaxed constant rank constraint qualification [16], for (5.39) with every $J \subset I_0$.

ACKNOWLEDGEMENTS

This work was funded by the Russian Science Foundation Grant 24-21-00015 (https://rscf.ru/en/project/24-21-00015/).

REFERENCES

1. Behling, R. (2011) *The method and the trajectory of Levenberg–Marquardt*. PhD thesis. IMPA - Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro. https://impa.br/wp-content/uploads/2017/08/tese_dout_roger_behling.pdf
2. Behling, R. & Fischer, A. (2012) A unified local convergence analysis of inexact constrained Levenberg–Marquardt methods, *Optim. Lett.*, **6**, 927–940.
3. Dennis, J.E. & Schnabel, R.B. (1983) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
4. Facchinei, F., Fischer, A. & Herrich, M. (2013) A family of Newton methods for nonsmooth constrained systems with nonisolated solutions, *Math. Methods Oper. Res.*, **77**, 433–443.
5. Facchinei F. & Kanzow, C. (1997) A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Math. Program.*, **76**, 493–512.
6. Fan, J.-Y. & Pan, J.-Y. (2011) On the convergence rate of the inexact Levenberg–Marquardt method, *J. Ind. Manag. Optim.*, **7**, 199–210.
7. Fischer, A., Herrich, M., Izmailov, A.F. & Solodov, M.V. (2016) Convergence conditions for Newton-type methods applied to complementarity systems with nonisolated solutions, *Comput. Optim. Appl.*, **63**, 425–459.
8. Fischer, A., Izmailov, A.F. & Solodov, M.V. (2023) The Levenberg–Marquardt method: an overview of modern convergence theories and more. (Submitted).
9. Fischer, A. & Strasdat, N. (2023) An extended convergence framework applied to complementarity systems with degenerate and nonisolated solutions, *Pure Appl. Func. Analys.*, **8**, 1039–1054.
10. Frank, M. & Wolfe, P. (1956) An algorithm for quadratic programming, *Naval Research Logistics Quarterly*, **3**, 95–110.
11. Hager, W.W. (1979) Lipschitz continuity for constrained processes, *SIAM J. Control Optim.*, **17**, 321–338.
12. Izmailov, A.F. & Solodov, M.V. (2014) *Newton-Type Methods for Optimization and Variational Problems*. Cham, Germany: Springer.
13. Kanzow, C. & Petra, S. (2004) On a semismooth least squares formulation of complementarity problems with gap reduction, *Optim. Meth. Software*, **19**, 507–525.
14. Kanzow, C. & Petra, S. (2007) Projected filter trust region methods for a semismooth least squares formulation of mixed complementarity problems, *Optim. Meth. Software*, **22**, 713–735.
15. Kanzow, C., Yamashita, N. & Fukushima, M. (2004) Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex

constraints, *J. Comput. Appl. Math.*, **172**, 375–397.
16. Kyparisis, J. (1990) Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers, *Math. Oper. Res.*, **15**, 286–298.
17. Levenberg, K. (1944) A method for the solution of certain non-linear problems in least squares, *Quarterly of Appl. Math.*, **2**, 164–168.
18. Mangasarian, O.L. & Fromovitz, S. (1967) The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.*, **17**, 37–47.
19. Marquardt, D.W. (1963) An algorithm for least squares estimation of non-linear parameters, *SIAM J.*, **11**, 431–441.
20. Nocedal, J. & Wright, S.J. (2006) *Numerical Optimization. 2nd edition.* New York, NY: Springer New York.
21. Yamashita, N. & Fukushima, M. (2001) On the rate of convergence of the Levenberg–Marquardt method. In: Alefeld, G., Chen, X. (Eds.), *Topics in Numerical Analysis. Computing Supplementa, vol. 15* (pp. 239–249). Vienna, Austria: Springer.
22. Yin, J., Jian, J. & Ma, G. (2023) A modified inexact Levenberg–Marquardt method with the descent property for solving nonlinear equations, *Comput. Optim. Appl.* https://doi.org/10.1007/s10589-023-00513-z.