Math. Program., Ser. B (2018) 168:475–507
https://doi.org/10.1007/s10107-016-1047-x

FULL LENGTH PAPER

Critical solutions of nonlinear equations: stability issues

A. F. Izmailov[1,2] · **A. S. Kurennoy**[3] · **M. V. Solodov**[4]

Received: 25 November 2015 / Accepted: 23 June 2016 / Published online: 1 July 2016

Abstract It is known that when the set of Lagrange multipliers associated with a stationary point of a constrained optimization problem is not a singleton, this set may contain so-called critical multipliers. This special subset of Lagrange multipliers defines, to a great extent, stability pattern of the solution in question subject to parametric perturbations. Criticality of a Lagrange multiplier can be equivalently characterized by the absence of the local Lipschitzian error bound in terms of the natural residual of the optimality system. In this work, taking the view of criticality as that associated to the error bound, we extend the concept to general nonlinear equations (not necessar-

Dedicated to Professor R. Tyrrell Rockafellar.

This research is supported in part by the Russian Foundation for Basic Research Grant 14-01-00113, by the Russian Science Foundation Grant 15-11-10021, by the Grant of the Russian Federation President for the state support of leading scientific schools NSh-8215.2016.1, by Volkswagen Foundation, by CNPq Grants PVE 401119/2014-9 and 303724/2015-3, and by FAPERJ.

✉ M. V. Solodov
solodov@impa.br

A. F. Izmailov
izmaf@ccas.ru

A. S. Kurennoy
akurennoy@cs.msu.ru

1 OR Department, VMK Faculty, Lomonosov Moscow State University (MSU), Uchebniy Korpus 2, Leninskiye Gory, Moscow, Russia 119991

2 RUDN University, Miklukho-Maklaya Str. 6, Moscow, Russia 117198

3 Department of Mathematics, Physics and Computer Sciences, Derzhavin Tambov State University, TSU, Internationalnaya 33, Tambov, Russia 392000

4 IMPA – Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil

ily with primal–dual optimality structure). Among other things, we show that while singular noncritical solutions of nonlinear equations can be expected to be stable only subject to some poor "asymptotically thin" classes of perturbations, critical solutions can be stable under rich classes of perturbations. This fact is quite remarkable, considering that in the case of nonisolated solutions, critical solutions usually form a thin subset within all the solutions. We also note that the results for general equations lead to some new insights into the properties of critical Lagrange multipliers (i.e., solutions of equations with primal–dual structure).

Keywords Nonlinear equations · Error bound · Critical Lagrange multipliers · Critical solutions · Stability · Sensitivity · 2-Regularity

Mathematics Subject Classification 90C33 · 65K10 · 49J53

1 Introduction

Consider a generic nonlinear equation without any special structure:

$$\Phi(u) = 0, \tag{1}$$

where $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ is some given mapping.

As is well known, if Φ is differentiable at a solution $\bar{u} \in \mathbb{R}^p$ of Eq. (1), then

$$T_{\Phi^{-1}(0)}(\bar{u}) \subset \ker \Phi'(\bar{u}), \tag{2}$$

where $T_U(u)$ stands for the contingent cone to the set U at a point $u \in U$, i.e. the tangent cone as defined in [33, Definition 6.1]. The following notion of critical/noncritical solutions of general nonlinear equations, formulated here for the first time, is the key to this work; it employs Clarke-regularity of a set, for which we refer to [33, Definition 6.4] (see also the original definition in [8, Definition 2.4.6]).

Definition 1 Assuming that Φ is differentiable at a solution $\bar{u}$ of Eq. (1), this solution is referred to as *noncritical* if the set $\Phi^{-1}(0)$ is Clarke-regular at $\bar{u}$, and

$$T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u}). \tag{3}$$

Otherwise, solution $\bar{u}$ is referred to as *critical*.

We shall show that noncriticality of a solution $\bar{u}$ is closely related to the local Lipschitzian error bound:

$$\operatorname{dist}(u,\ \Phi^{-1}(0)) = O(\|\Phi(u)\|) \tag{4}$$

holds as $u \in \mathbb{R}^p$ tends to $\bar{u}$. We shall also establish that singular noncritical solutions can be expected to be stable only subject to some poor "asymptotically thin" classes

of perturbations. By contrast, critical solutions can be stable under rich classes of perturbations.

To explain the origins of the notion of critical/noncritical solutions for the general Eq. (1), consider the equality-constrained optimization problem

$$\begin{array}{l} \text{minimize } f(x) \\ \text{subject to } h(x) = 0, \end{array} \tag{5}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ are smooth. The Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$ of this problem is given by

$$L(x,\, \lambda) = f(x) + \langle \lambda,\, h(x) \rangle.$$

Then stationary points and associated Lagrange multipliers of the problem (5) are characterized by the Lagrange optimality system

$$\frac{\partial L}{\partial x}(x,\, \lambda) = 0, \quad h(x) = 0, \tag{6}$$

with respect to $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^l$. Let $\mathscr{M}(\bar{x})$ stand for the set of Lagrange multipliers associated to a stationary point $\bar{x}$ of the problem (5), i.e.,

$$\mathscr{M}(\bar{x}) = \left\{ \lambda \in \mathbb{R}^l \;\middle|\; \frac{\partial L}{\partial x}(\bar{x},\, \lambda) = 0 \right\}.$$

When the multiplier set $\mathscr{M}(\bar{x})$ is nonempty but is not a singleton, it is an affine manifold of a positive dimension. It has been observed that in the latter cases, there is often a special subset of Lagrange multipliers, called critical; see Definition 2 below (this notion was first introduced in [19]). It turned out that this kind of multipliers are important for a good number of reasons, including convergence properties of Newton-type methods, error bounds, and stability of problems under perturbations. We refer to [12,20,23–25,27,29] and discussions therein; see also the book [26].

Definition 2 A Lagrange multiplier $\bar{\lambda} \in \mathbb{R}^l$ associated to a stationary point $\bar{x}$ of the optimization problem (5) is called *critical* if there

$$\text{exists } \xi \in \ker h'(\bar{x}) \backslash \{0\} \text{ such that } \frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda})\xi \in \operatorname{im}(h'(\bar{x}))^{\mathrm{T}}, \tag{7}$$

and *noncritical* otherwise.

In other words, $\bar{\lambda}$ is critical if the corresponding reduced Hessian of the Lagrangian (i.e., the symmetric matrix $H(\bar{\lambda}) = H(\bar{x},\, \bar{\lambda})$ of the quadratic form $\xi \to \langle \frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda})\xi,\, \xi \rangle : \ker h'(\bar{x}) \to \mathbb{R}$) is singular. As we shall show (see Proposition 2 below), if $\bar{\lambda}$ is a noncritical Lagrange multiplier, then $\bar{u} = (\bar{x},\, \bar{\lambda})$ is a noncritical solution of the equation representing the Lagrange optimality system (6). Moreover,

if $\bar{x}$ is an isolated stationary point of the optimization problem (5), then $\bar{u} = (\bar{x}, \bar{\lambda})$ is a critical solution of the Lagrange system if and only if $\bar{\lambda}$ is a critical Lagrange multiplier.

For the purposes of this work, it is useful to point out the following characterization of critical and noncritical Lagrange multipliers [26, Proposition 1.43]. A related result can be found in [16, Lemma 2].

Theorem 1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be twice differentiable at $\bar{x} \in \mathbb{R}^n$. Let $\bar{x}$ be a stationary point of problem (5), and let $\bar{\lambda} \in \mathbb{R}^l$ be an associated Lagrange multiplier.*

Then the following three properties are equivalent:

(a) The multiplier $\bar{\lambda}$ is noncritical.

(b) The error bound

$$\|x - \bar{x}\| + \operatorname{dist}(\lambda, \mathscr{M}(\bar{x})) = O\left(\left\|\left(\frac{\partial L}{\partial x}(x, \lambda), h(x)\right)\right\|\right)$$

holds as $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ tends to $(\bar{x}, \bar{\lambda})$.

(c) For every $w = (a, b) \in \mathbb{R}^n \times \mathbb{R}^l$, any solution $(x(w), \lambda(w))$ of the canonically perturbed Lagrange system

$$\frac{\partial L}{\partial x}(x, \lambda) = a, \quad h(x) = b,$$

which is close enough to $(\bar{x}, \bar{\lambda})$, satisfies the estimate

$$\|x(w) - \bar{x}\| + \operatorname{dist}(\lambda(w), \mathscr{M}(\bar{x})) = O(\|w\|)$$

as $w \to 0$.

In particular, criticality of a Lagrange multiplier can be equivalently characterized by the *lack* of the Lipschitzian error bound (the bound on the distance to the primal–dual solution set in terms of the residual of the Lagrange optimality system). This issue had been emphasized in the discussion associated to [27] (see [28]), and it was conjectured that the notion of critical solutions might be relevant beyond optimality systems with primal–dual structure as in (6). The present work is devoted precisely to this subject. It is taking the view of existence or not of a Lipschitzian error bound for the general Eq. (1), that we arrived to the notion of criticality stated in Definition 1; for the precise relations, see Sect. 2 and Theorem 2 in particular. We also show that this notion is central to stability patterns of solutions of nonlinear equations subject to perturbations; see Sect. 3. Going back to optimization and critical Lagrange multipliers, some new insights are given in Sect. 4.

We finish this section with some words about our notation. Throughout, $\|\cdot\|$ stands for the Euclidian norm; $B(x, \delta)$ is an open ball centered at x, of radius δ; $\operatorname{dist}(u, U) = \inf\{\|u - \hat{u}\| \mid \hat{u} \in U\}$. Along with the contingent cone $T_U(u)$, we shall make use of the regular tangent cone $\widehat{T}_U(u)$ to U at u, as defined in [33, Definition 6.25]. The polar

(negative dual) cone to a cone K is denoted by K°. Then $\widehat{N}_U(u)$ stands for the regular normal cone to U at u, as defined in [33, Definition 6.3], i.e., it is $(T_U(u))^\circ$ (see [33, Theorem 6.28 (a)]). For a smooth manifold S, its dimension is $\dim S = \dim T_S(u)$ for all $u \in S$ (in this case, $T_S(u)$ is a linear subspace). For a matrix A, $\ker A$ is its null space and $\mathrm{im}\, A$ is its range space. By I we denote the identity matrix of any dimension (always clear from the context). The orthogonal projector onto a linear subspace M is denoted by $P_M^\perp$.

Recall finally that a set U is called star-like with respect to $u \in U$ if $t\hat{u}+(1-t)u \in U$ for all $\hat{u} \in U$ and all $t \in [0,\ 1]$. For such a set, v is referred to as an excluded direction if $u + tv \notin U$ for all $t > 0$.

2 Noncritical solutions and the error bound

Given a solution $\bar{u}$ of the Eq. (1), we shall be saying that Φ is *strictly differentiable* at $\bar{u}$ *with respect to the null set* $\Phi^{-1}(0)$ if it is differentiable at $\bar{u}$, and

$$\|\Phi(u) - \Phi'(\bar{u})(u - \hat{u})\| = o(\|u - \hat{u}\|) \tag{8}$$

as $u \in \mathbb{R}^p$ and $\hat{u} \in \Phi^{-1}(0)$ tend to $\bar{u}$. Note that this property is weaker than the usual strict differentiability (for example, if $\bar{u}$ is an isolated solution of (1), then strict differentiability of Φ at $\bar{u}$ with respect to the null set is equivalent to differentiability of Φ at $\bar{u}$).

The key features of noncritical solutions are exposed by the following Theorem 2. After the proof, we shall illustrate this theorem by some examples, and discuss some subtleties of its assertions and assumptions. In particular, we shall show that Clarke-regularity and the equality (3) in the definition of noncriticality are independent (neither property implies the other); that the equivalent properties in Theorem 2 can hold even when $\Phi^{-1}(0)$ is not a smooth manifold near $\bar{u}$; and that in general strict differentiability of Φ at $\bar{u}$ with respect to the null set $\Phi^{-1}(0)$ cannot be replaced by differentiability at $\bar{u}$. See also Remark 3 below for another justification of the smoothness assumptions in Theorem 2, coming from the context of the optimization problem (5) and Theorem 1.

Theorem 2 *Let $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ be continuous near a solution $\bar{u} \in \mathbb{R}^p$ of Eq. (1), and strictly differentiable at $\bar{u}$ with respect to the null set $\Phi^{-1}(0)$.*

Then the following three properties are equivalent:

(a) Solution $\bar{u}$ is noncritical.
(b) The error bound (4) holds as $u \in \mathbb{R}^p$ tends to $\bar{u}$.
(c) Any solution $u(w)$ of the perturbed equation

$$\Phi(u) = w, \tag{9}$$

close enough to $\bar{u}$, satisfies the upper Lipschitzian property

$$\mathrm{dist}(u(w),\ \Phi^{-1}(0)) = O(\|w\|)$$

as $w \in \mathbb{R}^q$ tends to 0.

We emphasize that item (c) above does not claim the existence of solutions of perturbed problems (the same concerns the corresponding part of Theorem 1). The upper-Lipschitzian property only means that if a solution close enough to the basic one exists, it satisfies the stated estimate.

For the proof we shall need the following.

Lemma 1 *For any $U \subset \mathbb{R}^p$, any $u \in U$ and $v \in \mathbb{R}^p$ satisfying* $\operatorname{dist}(v,\ T_U(u)) > 0$, *and any $\delta \in (0,\ \operatorname{dist}(v,\ T_U(u)))$, it holds that* $\operatorname{dist}(u+tv,\ U) \geq \delta t$ *for all $t > 0$ small enough.*

Proof We argue by contradiction: suppose that there exists a sequence of reals $\{t_k\}$ such that $\{t_k\} \to 0+$ and

$$\operatorname{dist}(u + t_k v,\ T_U(u)) < \delta t_k$$

for all k. Then for every k there exists $\hat{u}^k \in U$ such that $\|u + t_k v - \hat{u}^k\| < \delta t_k$, and hence,

$$\|v - (\hat{u}^k - u)/t_k\| < \delta. \tag{10}$$

This implies, in particular, that the sequence $\{(\hat{u}^k - u)/t_k\}$ is bounded, and thus, has an accumulation point $\hat{v}$, which belongs to $T_U(u)$ by the definition of the latter. Then (10) yields

$$\operatorname{dist}(v,\ T_U(u)) \leq \delta,$$

contradicting the choice of δ. □

Proof (of Theorem 2) The equivalence between properties (b) and (c) is obvious, and it is valid without any differentiability assumptions: for each $u \in \mathbb{R}^p$, just set $w = \Phi(u)$ by definition. We next prove the equivalence between items (a) and (b).

Suppose that (a) holds, but (b) does not, i.e., there exists a sequence $\{u^k\} \subset \mathbb{R}^p \backslash \Phi^{-1}(0)$ such that $\{u^k\} \to \bar{u}$, and

$$\frac{\operatorname{dist}(u^k,\ \Phi^{-1}(0))}{\|\Phi(u^k)\|} \to \infty \tag{11}$$

as $k \to \infty$. By the continuity of Φ near $\bar{u}$, the set $\Phi^{-1}(0)$ is closed near $\bar{u}$. Hence, for each k sufficiently large there exists a projection of u^k onto $\Phi^{-1}(0)$. Let $\hat{u}^k$ be any projection of u^k onto $\Phi^{-1}(0)$, and define $v^k = (u^k - \hat{u}^k)/\|u^k - \hat{u}^k\|$ (recall that $u^k \notin \Phi^{-1}(0)$). Then $\{\hat{u}^k\}$ converges to $\bar{u}$, and without loss of generality we can assume that $\{v^k\}$ converges to some $v \in \mathbb{R}^p$, $\|v\| = 1$. From strict differentiability of Φ at $\bar{u}$ with respect to the null set $\Phi^{-1}(0)$, we then obtain that

$$\|\Phi(u^k) - \Phi'(\bar{u})(u^k - \hat{u}^k)\| = o(\|u^k - \hat{u}^k\|)$$

as $k \to \infty$. Therefore,

$$\frac{\|\Phi(u^k)\|}{\operatorname{dist}(u^k,\ \Phi^{-1}(0))} = \|\Phi'(\bar{u})v^k\| + \frac{o(\|u^k - \hat{u}^k\|)}{\|u^k - \hat{u}^k\|},$$

as $k \to \infty$. According to (11), the left-hand side in the latter relation tends to 0, while the right-hand side tends to $\Phi'(\bar{u})v$. We conclude that $v \in \ker \Phi'(\bar{u})$.

On the other hand, by [33, Example 6.16], for all k it holds that

$$u^k - \hat{u}^k \in \widehat{N}_{\Phi^{-1}(0)}(\hat{u}^k).$$

Therefore, $v^k \in \widehat{N}_{\Phi^{-1}(0)}(\hat{u}^k)$. Then, by Clarke-regularity of $\Phi^{-1}(0)$ at $\bar{u}$ (which is part of item (a); recall Definition 1), we obtain that

$$v \in \widehat{N}_{\Phi^{-1}(0)}(\bar{u}) = (T_{\Phi^{-1}(0)}(\bar{u}))^\circ = (\ker \Phi'(\bar{u}))^\perp,$$

where the last equality is by (3). Combining this with the inclusion $v \in \ker \Phi'(\bar{u})$, we get a contradiction, because $v \neq 0$.

Suppose now that (b) holds. Using again the fact that $\Phi^{-1}(0)$ is closed near $\bar{u}$, by [33, Corollary 6.29 (b)] we conclude that the needed Clarke-regularity of $\Phi^{-1}(0)$ at $\bar{u}$ is equivalent to the equality $T_{\Phi^{-1}(0)}(\bar{u}) = \widehat{T}_{\Phi^{-1}(0)}(\bar{u})$. The inclusion $\widehat{T}_{\Phi^{-1}(0)}(\bar{u}) \subset T_{\Phi^{-1}(0)}(\bar{u})$ is always valid [33, Theorem 6.26]. Thus we need to prove the converse inclusion.

Let there exists $v \in T_{\Phi^{-1}(0)}(\bar{u}) \backslash \widehat{T}_{\Phi^{-1}(0)}(\bar{u})$. Employing again [33, Theorem 6.26], this implies the existence of a sequence $\{u^k\} \subset \Phi^{-1}(0)$ such that $\{u^k\} \to \bar{u}$, and for any choices of $v^k \in T_{\Phi^{-1}(0)}(u^k)$ the sequence $\{v^k\}$ does not converge to v. Then passing onto a subsequence if necessary, we can assume that there exists $\gamma > 0$ such that for all k

$$\operatorname{dist}(v,\ T_{\Phi^{-1}(0)}(u^k)) \geq \gamma.$$

Then by Lemma 1 we conclude that for all k

$$\operatorname{dist}(u^k + tv,\ \Phi^{-1}(0)) \geq \frac{\gamma}{2} t$$

for all $t > 0$ small enough. This implies that we can choose a sequence of reals $\{t_k\}$ such that $\{t_k\} \to 0+$, and for all k

$$\operatorname{dist}(u^k + t_k v,\ \Phi^{-1}(0)) \geq \frac{\gamma}{2} t_k. \tag{12}$$

On the other hand, by strict differentiability of Φ at $\bar{u}$ with respect to the null set $\Phi^{-1}(0)$, we have that

$$\|\Phi(u^k + t_k v) - t_k \Phi'(\bar{u})v\| = o(t_k)$$

as $k \to \infty$, where $v \in \ker \Phi'(\bar{u})$ due to (2). Therefore,

$$\|\Phi(u^k + t_k v)\| = o(t_k)$$

as $k \to \infty$. Combining this estimate with (12), we get a contradiction with (4).

It remains to establish (3). This relation follows from (4) in a standard way (and the only assumption needed is differentiability of Φ at $\bar{u}$). Specifically, recalling again that (2) is automatic, for every $v \in \ker \Phi'(\bar{u})$ we have by (4) that

$$\operatorname{dist}(\bar{u} + tv,\ \Phi^{-1}(0)) = O(\|\Phi(\bar{u} + tv)\|) = O(t\|\Phi'(\bar{u})v\|) + o(t) = o(t)$$

as $t \to 0$, implying that $v \in T_{\Phi^{-1}(0)}(\bar{u})$. □

We next illustrate Theorem 2 by some examples, and in particular discuss some subtleties of its assertions and assumptions.

To begin with, it was demonstrated in [6] that, assuming continuous differentiability of Φ near $\bar{u}$, the error bound (4) implies that $\Phi^{-1}(0)$ is a smooth manifold near $\bar{u}$. Hence, in this case, it is automatically Clarke-regular at $\bar{u}$ [33, Example 6.8]. We next exhibit that under the smoothness assumptions of Theorem 2, the equivalent properties (a–c) may hold even when $\Phi^{-1}(0)$ is *not* a smooth manifold near $\bar{u}$.

Example 1 Consider the function $\varphi : [-1,\ 1] \to \mathbb{R}$ whose graph is shown in [10, left graph of Figure 1.7]. This function is continuous, it holds that $\varphi(0) = 0$, $\varphi(\pm 1/k) = 1/(k^2)$, and it is affine on the intervals $(-1/k,\ -1/(k+1))$ and $(1/(k+1),\ 1/k)$, $k = 1,\ 2,\ \ldots$. This function is strictly differentiable at 0, with $\varphi'(0) = 0$, but every neighborhood of 0 contains points where φ is not differentiable. Define $\Phi : \mathbb{R}^2 \to \mathbb{R}$ in such a way that $\Phi(u) = u_2 - \varphi(u_1)$ when $u_1 \in [0,\ 1]$. Then Φ is continuous near $\bar{u} = 0$ and strictly differentiable at $\bar{u}$, with $\Phi'(\bar{u}) = (0,\ 1)$. Furthermore, the null set $\Phi^{-1}(0)$ near $\bar{u}$ coincides with the graph of φ, and its intersection with any neighborhood of $\bar{u}$ is not a smooth manifold. At the same time, this set is evidently Clarke-regular at $\bar{u}$, and (3) holds because $T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u}) = \{v \in \mathbb{R}^2 \mid v_2 = 0\}$. In other words, $\bar{u}$ is a noncritical solution of (1), and hence, by Theorem 2, error bound (4) and the upper Lipschitzian property both hold for this solution. □

Evidently, regardless of any smoothness assumptions, Clarke-regularity does not imply (3). Indeed, if $\Phi^{-1}(0)$ is a singleton $\{\bar{u}\}$, it is certainly Clarke-regular at $\bar{u}$. But if $\ker \Phi'(\bar{u}) \neq \{0\}$, then (3) is violated. (Take, e.g., $\Phi : \mathbb{R} \to \mathbb{R}$, $\Phi(u) = u^2$). The converse implication [of Clarke-regularity by (3)] is also not valid, as demonstrated by the next example. Therefore, Clarke regularity and (3) are indeed independent ingredients of the definition of noncriticality.

Example 2 Take any closed set $U \subset \mathbb{R}^p$ such that it is not Clarke-regular at some $\bar{u} \in U$, and it holds that $T_U(\bar{u}) = \mathbb{R}^p$ (e.g., two closed balls in $\mathbb{R}^p$ with the only common point $\bar{u}$). According to the remarkable theorem due to Whitney (see, e.g., [3, Theorem 2.3.1]), there exists an infinitely differentiable function $\Phi : \mathbb{R}^p \to \mathbb{R}$ such that $U = \Phi^{-1}(0)$. From (2) it then follows that $\Phi'(\bar{u}) = 0$. Hence, (3) holds, which demonstrates that the latter does not imply Clarke-regularity under any smoothness assumptions. Therefore, according to Theorem 2, error bound (4) (and the upper Lipschitzian property) cannot hold for any choice of an appropriate mapping Φ. □

The next two examples demonstrate that strict differentiability with respect to the null set in Theorem 2 cannot be replaced by only differentiability at the solution in question.

Example 3 Define the function $\Phi : \mathbb{R}^2 \to \mathbb{R}$,

$$\Phi(u) = \begin{cases} \varphi(u) & \text{if } u_2 \geq u_1^2, \\ u_2\varphi(u) & \text{if } 0 < u_2 < u_1^2, \\ u_2 & \text{if } u_2 \leq 0, \end{cases}$$

where $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $\varphi(u) = u_2 - u_1^2$. This Φ is everywhere continuous, and

$$\Phi^{-1}(0) = \{u \in \mathbb{R}^2 \mid \varphi(u) = 0 \text{ or } u_2 = 0\} \tag{13}$$

consists of the parabola and the straight line which are tangent to each other at $\bar{u} = 0$. The set $\Phi^{-1}(0)$ is evidently Clarke-regular at every point.

We first show that Φ is differentiable at $\bar{u}$, with $\Phi'(\bar{u}) = (0,\ 1)$. If this were not the case, there would exist $\gamma > 0$ and a sequence $\{u^k\} \subset \mathbb{R}^2 \backslash \{0\}$ such that $\{u^k\} \to \bar{u}$, and for all k it holds that

$$\gamma \leq \frac{\Phi(u^k) - \langle (0,\ 1),\ u^k \rangle}{\|u^k\|} = \frac{\Phi(u^k) - u_2^k}{\|u^k\|}. \tag{14}$$

Since infinitely many elements of the sequence $\{u^k\}$ satisfy at least one of the inequalities $u_2^k \geq (u_1^k)^2$, $0 < u_2^k < (u_1^k)^2$, or $u_2^k \leq 0$, passing onto a subsequence if necessary, we can assume without loss of generality that one of these inequalities holds for all k. If the first inequality holds, then

$$\Phi(u^k) - u_2^k = u_2^k - (u_1^k)^2 - u_2^k = -(u_1^k)^2,$$

which contradicts (14). If the second inequality holds, then

$$|\Phi(u^k) - u_2^k| = |u_2^k\varphi(u^k) - u_2^k| = |u_2^k| + o(|u_2^k|) \leq (u_1^k)^2 + o((u_1^k)^2),$$

which again contradicts (14). Finally, if the third inequality holds, then

$$\Phi(u^k) - u_2^k = u_2^k - u_2^k = 0,$$

which again contradicts (14).

We conclude that Φ is differentiable at $\bar{u}$ and $\Phi'(\bar{u}) = (0, 1)$. In particular, $T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u}) = \{v \in \mathbb{R}^2 \mid v_2 = 0\}$ (the latter is evident, but also follows from [17, Theorem F]). Thus, (3) holds.

We next show that in spite of all the nice properties shown above, the error bound (4) does not hold as $u \to \bar{u}$. Observe first that the function φ is everywhere continuously differentiable, and hence, Lipschitz-continuous near $\bar{u}$ with some constant $\ell > 0$. Denoting by $\hat{u}$ any projection of u onto $\varphi^{-1}(0)$, and observing that $\hat{u} \to \bar{u}$ as $u \to \bar{u}$, we obtain that

$$|\varphi(u)| = |\varphi(u) - \varphi(\hat{u})| \leq \ell\|u - \hat{u}\| = \ell \mathrm{dist}(u,\ \varphi^{-1}(0)) \tag{15}$$

for all $u \in \mathbb{R}^2$ close enough to $\bar{u}$.

For each k take $u^k = (1/k,\ 1/(2k^2))$. Since $\hat{u}^k = (1/k,\ 1/k^2) \in \varphi^{-1}(0)$, we have that

$$\operatorname{dist}(u^k,\ \varphi^{-1}(0)) \le \|u^k - \hat{u}^k\| = |u_2^k - \hat{u}_2^k| = \frac{1}{2k^2}.$$

Therefore, by (13) and (15), it holds that

$$\operatorname{dist}(u^k,\ \Phi^{-1}(0)) = \min\{\operatorname{dist}(u^k,\ \varphi^{-1}(0)),\ |u_2^k|\} = \min\left\{\operatorname{dist}(u^k,\ \varphi^{-1}(0)),\ \frac{1}{2k^2}\right\}$$
$$= \operatorname{dist}(u^k,\ \varphi^{-1}(0)) \ge \frac{1}{\ell}|\varphi(u^k)| = \frac{1}{2\ell k^2}.$$

On the other hand, since $0 < u_2^k < (u_1^k)^2$, we have that

$$\Phi(u^k) = u_2^k \varphi(u^k) = o(u_2^k) = o\left(\frac{1}{k^2}\right),$$

and hence, (4) cannot hold.

According to Theorem 2, in the current example the only possible reason for the lack of the error bound can be that Φ is not strictly differentiable at $\bar{u}$, and even not strictly differentiable with respect to $\Phi^{-1}(0)$. Indeed, for the sequences defined above,

$$|\Phi(u^k) - \Phi'(\bar{u})(u^k - \hat{u}^k)| = |u_2^k - \hat{u}_2^k| + o\left(\frac{1}{k^2}\right) = \frac{1}{2k^2} + o\left(\frac{1}{k^2}\right),$$

while $\|u^k - \hat{u}^k\| = 1/(2k^2)$, contradicting (8). □

As mentioned at the very end of the proof of Theorem 2, the error bound (4) implies (3) assuming only that Φ is differentiable at $\bar{u}$. However, without strict differentiability with respect to the null set, the error bound (4) does not necessarily imply Clarke-regularity. We show this next.

Example 4 As in Example 1, define $\Phi : \mathbb{R}^2 \to \mathbb{R}$ as $\Phi(u) = u_2 - \varphi(u_1)$, where now $\varphi : \mathbb{R} \to \mathbb{R}$ is given by

$$\varphi(t) = \begin{cases} t^2 \sin(1/t) & \text{if } t \ne 0, \\ 0 & \text{if } t = 0. \end{cases}$$

This Φ is everywhere continuous, and $\Phi^{-1}(0)$ is the graph of φ, which is not Clarke-regular at $\bar{u} = 0$.

It can be easily seen that Φ is differentiable at $\bar{u}$, with $\Phi'(\bar{u}) = (0,\ 1)$, and as in Examples 1 and 3, it holds that $T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u}) = \{v \in \mathbb{R}^2 \mid v_2 = 0\}$. In particular, (3) holds.

Furthermore, no matter what is taken as φ, for every $u \in \mathbb{R}^2$ it holds that $(u_1,\ \varphi(u_1)) \in \Phi^{-1}(0)$. Hence,

$$\operatorname{dist}(u,\ \Phi^{-1}(0)) \le |u_2 - \varphi(u_1)| = |\Phi(u)|,$$

giving the error bound (4).

According to Theorem 2, the only possible reason for the lack of Clarke-regularity is again the lack of strict differentiability of Φ with respect to the null set. Indeed, for each k take $\hat{u}^k = (1/(\pi k),\ 0) \in \Phi^{-1}(0)$ and $u^k = (2/(\pi(1+2k)),\ 0)$. Then

$$|\Phi(u^k) - \Phi'(\bar{u})(u^k - \hat{u}^k)| = |\hat{u}_2^k - \varphi(u_1^k)| = |\varphi(u_1^k)| = (u_1^k)^2 = \frac{4}{(\pi(1+2k))^2},$$

while

$$\|u^k - \hat{u}^k\| = |u_1^k - \hat{u}_1^k| = \frac{1}{\pi k(1+2k)} = O\left(\frac{1}{k^2}\right),$$

contradicting (8). □

In both Examples 3 and 4, the regularity condition

$$\operatorname{rank}\Phi'(\bar{u}) = q \tag{16}$$

holds. Therefore, these examples demonstrate that in the absence of strict differentiability with respect to the null set, the regularity condition (16) does not guarantee neither the error bound, nor Clarke-regularity. However, from [17, Theorem F] it immediately follows that (16) guarantees (3). At the same time, under strict differentiability with respect to the null set, (16) implies the error bound, which (by Theorem 2) implies Clarke-regularity, and thus noncriticality of the solution in question.

Theorem 3 *Under the assumptions of Theorem 2, if the condition* (16) *is satisfied, then the error bound* (4) *holds as* $u \in \mathbb{R}^p$ *tends to* $\bar{u}$*, and in particular,* $\bar{u}$ *is a noncritical solution of Eq.* (1).

Proof Fix any matrix $A \in \mathbb{R}^{(p-q)\times p}$ such that

$$\ker \Phi'(\bar{u}) \cap \ker A = \{0\}$$

(such matrix exists due to (16)). Define the mapping $F : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^p$,

$$F(u,\ r) = (\Phi(u+r),\ Ar). \tag{17}$$

Then $F(\bar{u},\ 0) = 0$, and

$$\frac{\partial F}{\partial r}(\bar{u},\ 0) = \begin{pmatrix} \Phi'(\bar{u}) \\ A \end{pmatrix} \tag{18}$$

is a nonsingular square matrix. Applying [17, Theorem C] (which is the implicit function theorem not assuming strict differentiability), we obtain the existence of a neighborhood O of $\bar{u}$ and of a mapping $r(\cdot) : O \to \mathbb{R}^p$ such that $r(\bar{u}) = 0$, r is continuous at $\bar{u}$, and

$$F(u,\ r(u)) = 0 \quad \forall\, u \in O. \tag{19}$$

According to (17), the last relation implies that

$$\Phi(u + r(u)) = 0 \quad \forall\, u \in O. \tag{20}$$

Furthermore, since the matrix in (18) is nonsingular, there exists $\gamma > 0$ such that

$$\left\| \frac{\partial F}{\partial r}(\bar{u},\, 0)v \right\| \geq \gamma \|v\| \quad \forall\, v \in \mathbb{R}^p.$$

Then from (19) we obtain that

$$\|\Phi(u)\| = \|F(u,\, 0)\| \geq \left\| \frac{\partial F}{\partial r}(\bar{u},\, 0)r(u) \right\| - \left\| F(u,\, 0) - F(u,\, r(u)) + \frac{\partial F}{\partial r}(\bar{u},\, 0)r(u) \right\|$$
$$\geq \gamma \|r(u)\| - \|\Phi(u) - \Phi(u + r(u)) + \Phi'(\bar{u})r(u)\| = \gamma \|r(u)\| + o(\|r(u)\|)$$

as $u \to \bar{u}$, where the second inequality is by (17), and the last equality is by (8) and (20). This yields (4). □

Of course, under any smoothness assumptions, solution $\bar{u}$ can be noncritical when (16) does not hold. The simplest example is by taking $\Phi \equiv 0$. We also note that a mapping can be strictly differentiable with respect to the solution set but not strictly differentiable in the classical sense, even when the regularity condition (16) holds. To see this, augment the mapping from Example 4 by u_1 as the second component (i.e, the system now has two equations: one defined in Example 4, and the second is $u_1 = 0$). Then $\Phi'(\bar{u})$ is square and nonsingular (thus (16) holds), Φ is not strictly differentiable, but it is strictly differentiable with respect to the null set because $\bar{u}$ is an isolated solution.

Relations between various properties involved in the discussion above are summarized in Fig. 1. Full lines with arrows correspond to the established implications, while dotted ones indicate implications which do not hold. The labels "D" (for differentiability at the solution) and "SDNS" (for strict differentiability at the solution with respect to the null set) indicate the smoothness requirements under which the implication holds or does not hold. If this information is missing, the corresponding implication does not hold under any smoothness assumptions.

We complete this section giving some more examples of noncritical and critical solutions, which will be useful also further below to illustrate some stability results. In all these examples the solutions are singular, by which we mean that the regularity condition (16) is violated. In fact, this situation is the main case of interest in the rest of this paper. Note that degeneracy is automatic if $p = q$ and $\bar{u}$ is a nonisolated solution of Eq. (1).

Example 5 Consider $\Phi : \mathbb{R}^p \to \mathbb{R}^p, \Phi(u) = (u_1^2,\, \ldots,\, u_p^2)$. Then the unique solution of (1) is $\bar{u} = 0$. As $\ker \Phi'(\bar{u}) = \mathbb{R}^p$, it is clear that $\bar{u}$ is a critical solution. □

We proceed with examples where $p = q$ (as in Example 5), but solution sets contain manifolds of positive dimension.

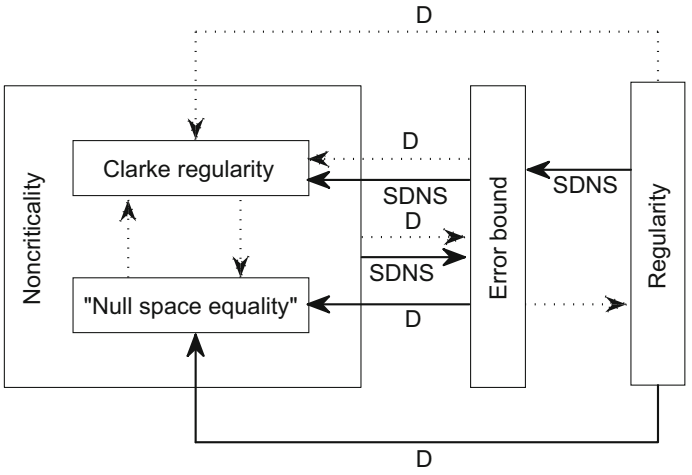

Fig. 1 Relations between the properties involved

Example 6 Consider $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$, $\Phi(u) = (\varphi(u)\varphi_1(u),\ \varphi(u)\varphi_2(u))$, where the functions $\varphi,\ \varphi_1,\ \varphi_2 : \mathbb{R}^2 \to \mathbb{R}$ are continuously differentiable functions. Then $\Phi^{-1}(0) \supset \varphi^{-1}(0)$, and if for some $\bar{u} \in \varphi^{-1}(0)$ it holds that $\varphi'(\bar{u}) \neq 0$, then near $\bar{u}$, the set $\varphi^{-1}(0)$ is a smooth manifold S of dimension 1. Furthermore, if for all $u \in \mathbb{R}^2 \setminus \{\bar{u}\}$ close enough to $\bar{u}$ it holds that $\varphi_1(u) \neq 0$ or $\varphi_2(u) \neq 0$ (e.g., when the gradients $\varphi_1'(\bar{u})$ and $\varphi_2'(\bar{u})$ are linearly independent), then $\Phi^{-1}(0) = S$ near $\bar{u}$. Since

$$\Phi'(\bar{u}) = \begin{pmatrix} \varphi_1(\bar{u})\varphi'(\bar{u}) \\ \varphi_2(\bar{u})\varphi'(\bar{u}) \end{pmatrix},$$

it holds that if $\varphi_1(\bar{u}) \neq 0$ or $\varphi_2(\bar{u}) \neq 0$, then $\dim \ker \Phi'(\bar{u}) = 1 = \dim S$, implying that $\bar{u}$ is a noncritical solution. On the other hand, if $\varphi_1(\bar{u}) = \varphi_2(\bar{u}) = 0$, then $\dim \ker \Phi'(\bar{u}) = 2 > 1 = \dim S$, and hence, $\bar{u}$ is a critical solution. □

Example 7 Consider $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$, $\Phi(u) = (u_1u_2,\ u_1u_3,\ u_2u_3)$. Then $\Phi^{-1}(0)$ is the union of three linear subspaces, all of dimension 1: $\{u \in \mathbb{R}^3 \mid u_1 = 0,\ u_2 = 0\}$, $\{u \in \mathbb{R}^3 \mid u_1 = 0,\ u_3 = 0\}$, and $\{u \in \mathbb{R}^3 \mid u_2 = 0,\ u_3 = 0\}$. Any nonzero solution $\bar{u}$ in any of these subspaces is noncritical, since

$$\Phi'(\bar{u}) = \begin{pmatrix} \bar{u}_2 & \bar{u}_1 & 0 \\ \bar{u}_3 & 0 & \bar{u}_1 \\ 0 & \bar{u}_3 & \bar{u}_2 \end{pmatrix},$$

implying that $\dim \ker \Phi'(\bar{u}) = 1$. However, $\bar{u} = 0$ belongs to all the specified subspaces and is critical, since $\dim \ker \Phi'(0) = 3 > 1$. □

Example 8 Consider $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$, $\Phi(u) = (u_1,\ u_1u_3,\ u_2u_3)$. Then $\Phi^{-1}(0)$ is the union of two linear subspaces, both of dimension 1: $\{u \in \mathbb{R}^3 \mid u_1 = 0,\ u_2 = 0\}$ and

$\{u \in \mathbb{R}^3 \mid u_1 = 0,\ u_3 = 0\}$. Any nonzero solution $\bar{u}$ in any of these subspaces is noncritical, since

$$\Phi'(\bar{u}) = \begin{pmatrix} 1 & 0 & 0 \\ \bar{u}_3 & 0 & \bar{u}_1 \\ 0 & \bar{u}_3 & \bar{u}_2 \end{pmatrix},$$

and hence, $\dim \ker \Phi'(\bar{u}) = 1$. However, $\bar{u} = 0$ belongs to both specified subspaces and is critical, since $\dim \ker \Phi'(0) = 2 > 1$. □

3 Further stability issues

The next result is a generalization of [19, Proposition 7], which analyzed stability properties of noncritical Lagrange multipliers. Here, we demonstrate that noncritical singular solutions of general nonlinear equations can be stable subject to very special perturbations only. In particular, see Remark 1 below.

Proposition 1 *Let $\Phi : \mathbb{R}^s \times \mathbb{R}^p \to \mathbb{R}^q$ be continuous near $(\bar{\sigma},\ \bar{u}) \in \mathbb{R}^s \times \mathbb{R}^p$, where $\bar{u}$ is a noncritical solution of the equation*

$$\Phi(\bar{\sigma},\ u) = 0. \tag{21}$$

Let Φ be strictly differentiable at $(\bar{\sigma},\ \bar{u})$ with respect to its null set (in the space $\mathbb{R}^s \times \mathbb{R}^p$). Let $\{\sigma^k\} \subset \mathbb{R}^s \backslash \{\bar{\sigma}\}$ and $\{u^k\} \subset \mathbb{R}^p$ be any sequences such that $\{\sigma^k\} \to \bar{\sigma}$, $\{u^k\} \to \bar{u}$, and for each k it holds that

$$\Phi(\sigma^k,\ u^k) = 0. \tag{22}$$

For each k, let $\hat{u}^k$ be any projection of u^k onto the solution set of the Eq. (21).

Then it holds that

$$\|u^k - \hat{u}^k\| = O(\|\sigma^k - \bar{\sigma}\|) \tag{23}$$

as $k \to \infty$, the sequence $\{(\sigma^k - \bar{\sigma},\ u^k - \hat{u}^k)/\|\sigma^k - \bar{\sigma}\|\}$ has accumulation points, and any such accumulation point $(d,\ v)$ satisfies the equality

$$\frac{\partial \Phi}{\partial \sigma}(\bar{\sigma},\ \bar{u})d + \frac{\partial \Phi}{\partial u}(\bar{\sigma},\ \bar{u})v = 0. \tag{24}$$

Proof Estimate (23) follows from (4), which holds under the stated assumptions, according to Theorem 2. Indeed,

$$\begin{aligned} \|u^k - \hat{u}^k\| &= O(\|\Phi(\bar{\sigma},\ u^k)\|) = O(\|\Phi(\bar{\sigma},\ u^k) - \Phi(\sigma^k,\ u^k)\|) \\ &= O\left(\left\|\frac{\partial \Phi}{\partial \sigma}(\bar{\sigma},\ \bar{u})(\sigma^k - \bar{\sigma})\right\|\right) + o(\|\sigma^k - \bar{\sigma}\|) = O(\|\sigma^k - \bar{\sigma}\|) \end{aligned}$$

as $k \to \infty$, where the first equality is by the noncriticality of $\bar{u}$ as a solution of (21) (in particular, by (4)), the second equality is by (22), and the third is by strict differentiability of Φ at $(\bar{\sigma}, \bar{u})$ with respect to the null set.

Note that, by its definition, $\{\hat{u}^k\}$ converges to $\bar{u}$. We then derive that

$$0 = \|\Phi(\sigma^k, u^k)\| = \|\Phi(\sigma^k, u^k) - \Phi(\bar{\sigma}, \hat{u}^k)\|$$
$$= \left\| \frac{\partial \Phi}{\partial \sigma}(\bar{\sigma}, \bar{u})(\sigma^k - \bar{\sigma}) + \frac{\partial \Phi}{\partial u}(\bar{\sigma}, \bar{u})(u^k - \hat{u}^k) \right\| + o(\|(\sigma^k - \bar{\sigma}, u^k - \hat{u}^k)\|)$$

as $k \to \infty$, where the last equation is again by strict differentiability of Φ at $(\bar{\sigma}, \bar{u})$ with respect to the null set. Taking into account (23), this implies (24). □

We next discuss why the results of Proposition 1 mean that singular noncritical solutions can be expected to be stable only under some poor/special classes of perturbations.

Remark 1 Note that (24) implies the inclusion

$$\frac{\partial \Phi}{\partial \sigma}(\bar{\sigma}, \bar{u})d \in \operatorname{im} \frac{\partial \Phi}{\partial u}(\bar{\sigma}, \bar{u}). \tag{25}$$

If the solution $\bar{u}$ of (21) is singular, i.e.,

$$\operatorname{rank} \frac{\partial \Phi}{\partial u}(\bar{\sigma}, \bar{u}) < q,$$

then the right-hand side of (25) is a proper linear subspace in $\mathbb{R}^q$. Hence, in this case, (25) can hold only for very special sequences $\{\sigma^k\}$, unless

$$\operatorname{im} \frac{\partial \Phi}{\partial \sigma}(\bar{\sigma}, \bar{u}) \subset \operatorname{im} \frac{\partial \Phi}{\partial u}(\bar{\sigma}, \bar{u}).$$

But the latter property is clearly atypical, and can only hold for very special (in a sense, poor) parameterizations. For instance, it does not hold for parameterizations allowing arbitrary right-hand side perturbations: singular noncritical solutions usually do not "survive" such perturbations. In particular, stability of a noncritical solution subject to arbitrary right-hand side perturbation implies the nondegeneracy condition (16). □

We proceed to give some illustrations of the discussion above.

Example 6 (continued) Consider the mapping Φ from Example 6 with $\varphi(u) = u_1$, $\varphi_1(\cdot) \equiv 1$, $\varphi_2(u) = u_2$. Then $\Phi^{-1}(0) = \varphi^{-1}(0) = \{u \in \mathbb{R}^2 \mid u_1 = 0\}$ is a linear subspace of dimension 1. Since φ_1 never equals zero, every solution $\bar{u}$ is noncritical. For any $w \in \mathbb{R}^2 \backslash \{0\}$, the perturbed Eq. (9) is solvable only when $w_1 \neq 0$, in which case the unique solution has the form

$$u(w) = (w_1, w_2/w_1). \tag{26}$$

Suppose that $\{w^k\} \subset \mathbb{R}^2$ converges to 0, $w_1^k \neq 0$ for all k, and $\{u(w^k)\}$ converges to some $\bar{u} \in \Phi^{-1}(0)$. Then by (26), it necessarily holds that $w_2^k/w_1^k \to \bar{u}_2$, implying that for any accumulation point d of the sequence $\{w^k/\|w^k\|\}$ it holds that $d_2 = \bar{u}_2 d_1$. This fully agrees with (25), since $\mathrm{im}\Phi'(\bar{u}) = \{w \in \mathbb{R}^2 \mid w_2 = \bar{u}_2 w_1\}$. Therefore, each solution (recall that they are all noncritical) can be stable only subject to perturbations tangential to very special directions, forming a linear subspace $\mathrm{im}\Phi'(\bar{u})$ of dimension 1 in the space of right-hand side perturbations, of dimension 2.

Now let $\varphi(u) = \varphi_1(u) = u_1$, $\varphi_2(u) = u_2$. Then the solution set is the same, but the solution $\bar{u} = 0$ is now critical, with all the other solutions being noncritical. For any $w \in \mathbb{R}^2 \backslash \{0\}$, the perturbed Eq. (9) is solvable only when $w_1 > 0$, in which case the solutions have the form

$$u(w) = \left(\pm\sqrt{w_1},\ \pm\frac{w_2}{\sqrt{w_1}}\right). \tag{27}$$

Suppose that $\{w^k\} \subset \mathbb{R}^2$ converges to 0, $w_1^k > 0$ for all k, and $\{u(w^k)\}$ converges to some $\bar{u} \in \Phi^{-1}(0)$. Then by (27), it necessarily holds that $|w_2^k|/\sqrt{w_1^k} \to \bar{u}_2$. Therefore, if $\bar{u}_2 \neq 0$, then for any accumulation point d of the sequence $\{w^k/\|w^k\|\}$ it holds that either $d = (0,\ 1)$ or $d = (0,\ -1)$. This again fully agrees with (25), since $\mathrm{im}\Phi'(\bar{u}) = \{w \in \mathbb{R}^2 \mid w_1 = 0\}$. At the same time, it can be easily seen that the unique critical solution $\bar{u} = 0$ is stable subject to a wide class of right-hand side perturbations, and this fact is explained by Theorem 4 below. □

We next discuss some further examples, showing that our considerations are relevant for perturbations of optimization problems with inequality constraints (at least if strict complementarity holds), and even for generalized Nash equilibrium problems [11].

Example 9 (DEGEN 20103 [9]) Consider the canonically perturbed inequality-constrained optimization problem

$$\begin{array}{l} \text{minimize } -x^2 - \chi x \\ \text{subject to } x^2 \leq y, \end{array} \tag{28}$$

where $w = (\chi,\ y) \in \mathbb{R} \times \mathbb{R}$ is a parameter. For $w = (0,\ 0)$, the unique solution of this problem is $\bar{x} = 0$.

The Karush–Kuhn–Tucker (KKT) optimality system with respect to $(x,\ \mu) \in \mathbb{R} \times \mathbb{R}$, characterizing stationary points and associated Lagrange multipliers of problem (28), has the form

$$-2x - \chi + 2\mu x = 0, \quad \mu \geq 0,\ x^2 \leq y,\ \mu(x^2 - y) = 0. \tag{29}$$

This system has no solutions if $y < 0$, and if $\chi \neq 0$, $y = 0$. For $w = (0,\ 0)$, the solution set is $\{\bar{x}\} \times \mathbb{R}_+$. If $y > 0$, this system has the solution $(x(w),\ \mu(w)) = (\sqrt{y},\ 1 + \chi/(2\sqrt{y}))$ when $-2\sqrt{y} \leq \chi$; the solution $(x(w),\ \mu(w)) = (-\sqrt{y},\ 1 - \chi/(2\sqrt{y}))$ when $\chi \leq 2\sqrt{y}$; and also the solution $(x(w),\ \mu(w)) = (-\chi/2,\ 0)$ when $-2\sqrt{y} < \chi < 2\sqrt{y}$. Solutions of the first two families tend to $(\bar{x},\ 1)$ if $\chi = o(\sqrt{y})$,

while solutions of the last family always tend to $(\bar{x},\ 0)$ as $w \to (0,\ 0)$. Therefore, the two solutions $(\bar{x},\ 1)$ and $(\bar{x},\ 0)$ of the unperturbed KKT system are stable subject to wide classes of specified perturbations. Other solutions can "survive" very special perturbations only, i.e., those satisfying $y = O(\chi^2)$. Observe that, for every sequence $\{w^k\} \subset \mathbb{R} \times \mathbb{R}$ such that $w^k = (\chi_k,\ y_k)$, $\chi_k \to 0$, and $y_k = O(\chi_k^2)$, any accumulation point d of the sequence $\{w^k/\|w^k\|\}$ is either $d = (1,\ 0)$ or $d = (-1,\ 0)$.

We next relate these observations about stability patterns in this problem to the results obtained above. Note that for any $\bar{\mu} > 0$, near the solution $\bar{u} = (\bar{x},\ \bar{\mu})$ of the unperturbed KKT system, and for w close enough to $(0,\ 0)$, system (29) reduces to the system of Eqs. (9) with $p = 2$, $u = (x,\ \mu)$, $w = (\chi,\ y)$,

$$\Phi(u) = (-2x(1-\mu),\ x^2). \tag{30}$$

It is easy to see that $\bar{u} = (\bar{x},\ 1)$ is a critical solution of Eq. (1), since $\Phi'(\bar{u}) = 0$, while $T_{\Phi^{-1}(0)}(\bar{u}) = \{0\} \times \mathbb{R}$. Other solutions $\bar{u} = (\bar{x},\ \bar{\mu})$ with $\bar{\mu} > 0$ are noncritical, with $\mathrm{im}\Phi'(\bar{u}) = \{w = (\chi,\ y) \in \mathbb{R} \times \mathbb{R} \mid y = 0\}$.

Furthermore, using the smooth complementarity function, we can equivalently reformulate (29) as the parametric system of equations

$$\Phi(\sigma,\ u) = 0, \tag{31}$$

where $\sigma = (\chi,\ y)$ and

$$\Phi(\sigma,\ u) = (-2x(1-\mu) - \chi,\ -2\mu(x^2 - y) - (\min\{0,\ \mu - x^2 + y\})^2).$$

It is easy to see that both $\bar{u} = (\bar{x},\ 1)$ and $\bar{u} = (\bar{x},\ 0)$ are critical solutions of (31) for $\sigma = 0$ (the latter solution corresponds to the unique multiplier violating strict complementarity, and $T_{\Phi^{-1}(0)}(\bar{u}) = \{0\} \times \mathbb{R}_+$). All the other solutions are noncritical. □

Example 10 ([11, Example 1.1]) Consider the canonically perturbed generalized Nash equilibrium problem

$$\begin{array}{ll} \text{minimize}_{x_1}\ (x_1 - 1)^2 - \chi_1 x_1 & \text{minimize}_{x_2}\ (x_2 - 1/2)^2 - \chi_2 x_2 \\ \text{subject to}\quad x_1 + x_2 \le y_1, & \text{subject to}\quad x_1 + x_2 \le y_2, \end{array} \tag{32}$$

where $w = (\chi,\ y) \in \mathbb{R}^2 \times \mathbb{R}^2$ is a parameter.

The KKT-type system of problem (32) has the form

$$2(x_1 - 1) - \chi_1 + \mu_1 = 0, \quad 2(x_2 - 1/2) - \chi_2 + \mu_2 = 0, \tag{33}$$

$$\mu_1 \ge 0,\ x_1 + x_2 \le y_1,\ \mu_1(x_1 + x_2 \le y_1) = 0,$$

$$\mu_2 \ge 0,\ x_1 + x_2 \le y_2,\ \mu_2(x_1 + x_2 \le y_2) = 0. \tag{34}$$

For $w = (\chi,\ y) \in \mathbb{R}^2 \times \mathbb{R}^2$ close enough to $(0,\ 0)$, this system has the solution

$$(x(w),\ \mu(w)) = \left(\left(\frac{1}{2} - \frac{1}{2}\chi_2 + y_1,\ \frac{1}{2} - \frac{1}{2}\chi_2\right),\ (1 + \chi_1 + \chi_2 - 2y_1,\ 0)\right)$$

if $y_1 < y_2$; the set of solutions

$$\left\{ (x,\ \mu) \;\middle|\; \begin{array}{c} x = (t,\ 1 + y_1 - t),\ \mu = \left(2(1-t) + \chi_1,\ 2\left(t - \frac{1}{2} - y_1\right) + \chi_2\right), \\ t \in \left[\frac{1}{2} - \frac{1}{2}\chi_2 + y,\ 1 + \frac{1}{2}\chi_1\right] \end{array} \right\}$$

if $y_1 = y_2$; and the solution

$$(x(w),\ \mu(w)) = \left(\left(\frac{1}{2} + \frac{1}{2}\chi_1,\ -\frac{1}{2}\chi_1 + y_2,\ \frac{1}{2} + \frac{1}{2}\chi_2\right),\ (0,\ 1 + \chi_1 + \chi_2 - 2y_2)\right)$$

if $y_1 > y_2$. In particular, for $w = (0,\ 0)$, the solution set of system (33)–(34) has the form

$$\left\{ (x,\ \mu) \;\middle|\; x = (t,\ 1-t),\ \mu = \left(2(1-t),\ 2\left(t - \frac{1}{2}\right)\right),\ t \in \left[\frac{1}{2},\ 1\right] \right\}.$$

Solutions of the first family tend to $((1/2,\ 1/2),\ (1,\ 0))$, while solutions of the third family tend to $((1,\ 0),\ (0,\ 1))$ as $w \to (0,\ 0)$. Hence, the two specified solutions of the unperturbed KKT-type system are stable subject to wide classes of specified perturbations. Solutions of the remaining second family may tend to any solution of the unperturbed KKT-type system, depending on the control of t, but this family exists for very special perturbations only, i.e., those with $y_1 = y_2$.

All these observations fully agree with the results obtained above, the same way as in Example 9, by considering separately those solutions satisfying strict complementarity (corresponding to $t \in (1/2,\ 1)$), and by treating the remaining two solutions via the smooth equation reformulation of the KKT-type system. □

We proceed to prove some formal results showing that, unlike noncritical solutions, critical ones can indeed be expected to be stable under some rich classes of perturbations. To this end, the notion of 2-regularity of a mapping will be useful.

Consider a mapping $\Phi : \mathbb{R}^p \to \mathbb{R}^q$, which is twice differentiable at $\bar{u} \in \mathbb{R}^p$. Let Π be the projector in $\mathbb{R}^q$ onto an arbitrary fixed complementary subspace of $\mathrm{im}\Phi'(\bar{u})$ along this subspace. For each $v \in \mathbb{R}^p$, define the $q \times p$-matrix

$$\Psi(\bar{u};\ v) = \Phi'(\bar{u}) + \Pi\Phi''(\bar{u})[v]. \tag{35}$$

The mapping Φ is referred to as *2-regular* at the point $\bar{u}$ in the direction $v \in \mathbb{R}^p$ if

$$\mathrm{rank}\Psi(\bar{u};\ v) = q.$$

It can be easily seen that the 2-regularity property is invariant with respect to the choice of Π, and to the norm of v, and it is stable subject to small perturbations of v. Moreover, 2-regularity in a direction v implies 2-regularity in the direction $-v$ as well.

The notion of 2-regularity proved to be a useful tool in nonlinear analysis and optimization theory; see, e.g., the book [2] and references therein. If Φ is regular at $\bar{u}$ in the sense of (16), then it is 2-regular at this point in every direction. However, in the singular case when (16) does not hold, the linear approximation of Φ is not adequate, and second-order information needs to be employed. This is where the notion of 2-regularity comes into play, and helps to extend various results to the singular case (for some applications, see, e.g., [4,14,15,21,22]).

Here, we use 2-regularity in the context of implicit function theorems. One important theorem of this kind was derived in [5], but it is not applicable in the irregular case with $p = q$, which is the setting of principal interest in the present work. A more general implicit function theorem was established in [18]. It is free from the above disadvantage, and contains the result of [5] as a particular case. The following assertions are obtained applying the implicit function theorem of [18] to the case of the right-hand side perturbations.

Theorem 4 *Let $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ be twice differentiable near $\bar{u} \in \mathbb{R}^p$, and let its second derivative be continuous at $\bar{u}$. Let $\bar{u}$ be a solution of Eq. (1). Let $K \subset \mathbb{R}^p$ be a closed cone such that the mapping Φ is 2-regular at $\bar{u}$ in every direction $v \in K \setminus \{0\}$. Let Π be the projector in $\mathbb{R}^q$ onto some complementary subspace of* $\mathrm{im}\Phi'(\bar{u})$ *along this subspace. Define the set*

$$W = W(K, \Pi) = \tilde{\Phi}(K), \tag{36}$$

where $\tilde{\Phi} : \mathbb{R}^p \to \mathbb{R}^q$,

$$\tilde{\Phi}(u) = \Phi'(\bar{u})u + \frac{1}{2}\Pi\Phi''(\bar{u})[u, u].$$

Then there exist $\varepsilon = \varepsilon(K, \Pi) > 0$ and $C = C(K, \Pi) > 0$ such that for every $w \in W \cap B(0, \varepsilon)$ the equation $\Phi(u) = w$ has a solution $u(w)$ such that

$$\|u(w) - \bar{u}\| \le C(\|(I - \Pi)w\| + \sqrt{\|\Pi w\|}). \tag{37}$$

We next provide some examples showing that, unlike for noncritical solutions, Theorem 4 can guarantee stability of critical solutions subject to wide classes of perturbations, allowing for star-like domains of "good" parameter values, with nonempty interior (and in particular, not "asymptotically thin").

Example 6 (continued) Consider again the mapping Φ from Example 6 with $\varphi(u) = u_1$, $\varphi_1(\cdot) \equiv 1$, $\varphi_2(u) = u_2$. Consider any noncritical solution, say $\bar{u} = (0, 1)$. Let Π be the orthogonal projector onto $(\mathrm{im}\Phi'(\bar{u}))^\perp = \{w \in \mathbb{R}^2 \mid w_1 + w_2 = 0\}$. We have that

$$\Pi = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}, \quad \Phi''(\bar{u})[v] = \begin{pmatrix} 0 & 0 \\ v_2 & v_1 \end{pmatrix}, \quad \Psi(\bar{u}; v) = \begin{pmatrix} 1 - v_2/2 & -v_1/2 \\ 1 + v_2/2 & v_1/2 \end{pmatrix}.$$

Therefore, $\det \Psi(\bar{u}; v) = v_1$, and hence, Φ is 2-regular at $\bar{u}$ in any direction v such that $v_1 \ne 0$. In particular, for every $\gamma > 0$, the mapping Φ is 2-regular at $\bar{u}$ in any

direction v from the closed cone

$$K_\gamma = \{v \in \mathbb{R}^2 \mid |v_1| \geq \gamma |v_2|\}. \tag{38}$$

Furthermore,

$$\tilde{\Phi}(u) = (u_1 - u_1 u_2/2,\ u_1 + u_1 u_2/2)\,,$$

and hence, the equation

$$\tilde{\Phi}(u) = w \tag{39}$$

is solvable for $w \neq 0$ if and only if $w_1 + w_2 \neq 0$, with the unique solution being

$$u(w) = \left(\frac{1}{2}(w_1 + w_2),\ \frac{w_2 - w_1}{w_1 + w_2}\right).$$

This readily implies that, for the cone K_γ defined in (38), the set defined according to (36) has the form

$$W(K_\gamma,\ \Pi) = \left\{ w \in \mathbb{R}^2 \,\middle|\, (w_1 + w_2)^2/2 \geq \gamma |w_2 - w_1| \right\}.$$

This set is shown in Fig. 2 as the area between the two symmetric parabolas; it is "asymptotically thin" near 0, which means that the ratio of the "size" (e.g., the Lebesgue measure) of the intersection of this area with $B(0,\ \delta)$ and the "size" of $B(0,\ \delta)$ tends to zero as $\delta \to 0+$. Theorem 4 can be applied with $K = K_\gamma$, and it claims that for every $\gamma > 0$ there exist $\varepsilon(\gamma) > 0$ and $C(\gamma) > 0$ such that for every $w \in W(K_\gamma,\ \Pi)$ satisfying $\|w\| < \varepsilon(\gamma)$, the perturbed Eq. (9) has a solution $u(w)$ satisfying

$$\|u(w) - \bar{u}\| \leq C(\gamma)(|w_1 + w_2| + \sqrt{|w_1 - w_2|}). \tag{40}$$

Smaller values of $\gamma > 0$ give larger sets $W(K_\gamma,\ \Pi)$ (see Fig. 2), and in the limit as $\gamma \to 0$, they give the entire plane with excluded nonzero points on the line $w_1 + w_2 = 0$. However, the domain of "appropriate" values of w remains "asymptotically thin", even if we give up with the estimate (40): according to Proposition 1, for every $d \in \mathbb{R}^2$ with $d_1 \neq d_2$ it holds that $w(t) = td$ does not belong to this domain for all $t > 0$ small enough.

We next turn to the case when $\varphi(u) = \varphi_1(u) = u_1$, $\varphi_2(u) = u_2$. Consider any noncritical solution, say $\bar{u} = (0,\ 1)$. Let Π be the orthogonal projector onto $(\mathrm{im}\Phi'(\bar{u}))^\perp = \{w \in \mathbb{R}^2 \mid w_1 = 0\}$. We have that

$$\Pi = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi''(\bar{u})[v] = \begin{pmatrix} 2v_1 & 0 \\ v_2 & v_1 \end{pmatrix}, \quad \Psi(\bar{u};\ v) = \begin{pmatrix} 2v_1 & 0 \\ 1 & 0 \end{pmatrix}.$$

This matrix is singular whatever is taken as v, and hence, Φ is not 2-regular at $\bar{u}$ in any direction. Therefore, Theorem 4 is not applicable at such solutions.

Consider now the unique critical solution $\bar{u} = 0$. We have: $\Phi'(\bar{u}) = 0$, $\Pi = I$, $\Psi(\bar{u};\ v) = \Phi''(\bar{u})[v]$. Therefore, $\det \Psi(\bar{u};\ v) = v_1$, and hence, Φ is 2-regular at $\bar{u}$ in

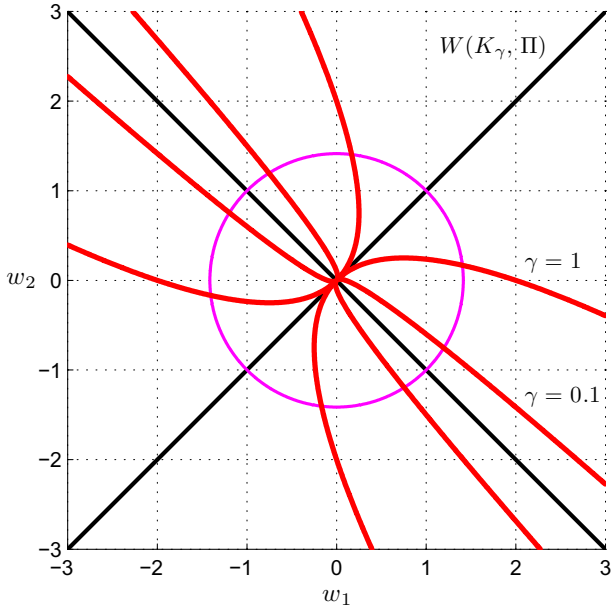

Fig. 2 Set $W(K_\gamma, \Pi)$

any direction v such that $v_1 \neq 0$. Furthermore, $\tilde{\Phi} = \Phi$ and for the cone K_γ defined in (38), we have that

$$W(K_\gamma, \Pi) = \{w \in \mathbb{R}^2 \mid w_1 \geq \gamma |w_2|\}.$$

Observe that, as a consequence of full degeneracy, in this case $W(K_\gamma, \Pi)$ is always a cone; see Fig. 3. Theorem 4 applied with $K = K_\gamma$ claims that for every $\gamma > 0$ there exist $\varepsilon(\gamma) > 0$ and $C(\gamma) > 0$ such that for every $w \in W(K_\gamma, \Pi)$ satisfying $\|w\| < \varepsilon(\gamma)$, the perturbed Eq. (9) has a solution $u(w)$ satisfying

$$\|u(w) - \bar{u}\| \leq C(\gamma)\sqrt{\|w\|}. \tag{41}$$

In the limit as $\gamma \to 0$, the sets $W(K_\gamma, \Pi)$ cover the entire open right half-plane with the added zero point. More precisely, for every $d \in \mathbb{R}^2$ with $\|d\| = 1$ and $d_1 > 0$ there exists $\gamma = \gamma(d) > 0$ such that $d \in W(K_\gamma, \Pi)$. Fix any $\beta > 0$, set $\tilde{\varepsilon}(d) = \min\{\varepsilon(\gamma),\ 1/(C(\gamma))^{2(1+\beta)}\}$, and define the set

$$W = \left\{w \in \mathbb{R}^2 \mid w_1 > 0,\ \|w\| < \tilde{\varepsilon}\,(w/\|w\|)\right\}.$$

Observe that this set is star-like with respect to 0, with the excluded directions being only those $d \in \mathbb{R}^2$ satisfying $d_1 \leq 0$; see Fig. 3. Then for every $w \in W$ the perturbed Eq. (9) has a solution $u(w)$ satisfying (41) with $\gamma = \gamma(w/\|w\|)$. This implies that $u(w) \to \bar{u}$ as $w \to 0$. Indeed, consider any sequence $\{w^k\} \subset W$ converging to zero. If

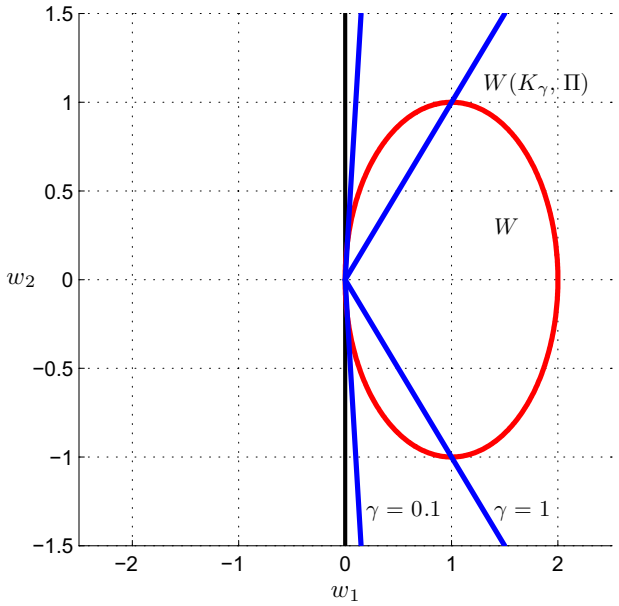

Fig. 3 Sets $W(K_\gamma, \Pi)$ and W

the sequence $\{C(\gamma(w^k/\|w^k\|))\}$ is bounded, then $\{u(w^k)\}$ converges to $\bar{u}$ according to (41). On the other hand, if $\{C(\gamma(w^k/\|w^k\|))\} \to \infty$, then from (41) and the definition of $\tilde{\varepsilon}(w^k/\|w^k\|)$ we have

$$\|u(w^k) - \bar{u}\| \leq (C(\gamma(w^k/\|w^k\|)))^{-\beta} \to 0$$

as $k \to \infty$.

Observe, however, that the estimate (41) with $C(\gamma)$ replaced by some $C > 0$ independent of γ does not hold for all $w \in W$. Specifically, for any choice of $C > 0$, such estimate does not hold along any sequence $\{w^k\} \subset W$ convergent to zero and such that $w_1^k = o(\|w_2^k\|)$. Indeed, from (27) we then have

$$|u_2(w^k)| = \frac{|w_2^k|}{\sqrt{w_1^k}} = \sqrt{\frac{|w_2^k|}{w_1^k}}\sqrt{|w_2^k|} > C\sqrt{\|w^k\|}$$

for all k large enough. □

Motivated by the example above, in the rest of this section we shall provide conditions ensuring that a given solution is stable subject to the right-hand side perturbations in a star-like domain with nonempty interior, in particular, not "asymptotically thin".

Consider any $w \in W(K, \Pi)$ for some cone $K \subset \mathbb{R}^p$ satisfying $K = -K$, i.e., there exists $u \in K$ satisfying (39). For convenience, let Π be the orthogonal projector

onto $(\mathrm{im}\Phi'(\bar{u}))^{\perp}$. Then for every $t \in \mathbb{R}$,

$$\tilde{\Phi}(tu) = t\Phi'(\bar{u})u + \frac{1}{2}t^2\Pi\Phi''(\bar{u})[u,\ u] = t(I-\Pi)w + t^2\Pi w.$$

Therefore, for the function $\omega_w : \mathbb{R} \to \mathbb{R}^q$, $\omega_w(t) = t(I-\Pi)w + t^2\Pi w$, we conclude that the parabolic curve defined by this function, passing through w (for $t = 1$), is contained in $W(K,\ \Pi)$, i.e., $\omega_w(t) \in W(K,\ \Pi)$ for all $t \in \mathbb{R}$.

Another observation is the following. For a given $\bar{v} \in \mathbb{R}^p$ such that Φ is 2-regular at $\bar{u}$ in this direction, set

$$\bar{w} = \tilde{\Phi}(\bar{v}). \tag{42}$$

Then $\tilde{\Phi}'(\bar{v}) = \Psi(\bar{u};\ \bar{v})$ has rank q, and applying the standard covering theorem to $\tilde{\Phi}$ at $\bar{v}$, we obtain the existence of $\delta > 0$ such that for every $w \in \mathbb{R}^q$ satisfying $\|w - \bar{w}\| < \delta$, Eq. (39) has a solution $u(w)$ tending to $\bar{v}$ as w tends to $\bar{w}$. By stability of 2-regularity with respect to small perturbations of a direction, there exists a closed cone $K \subset \mathbb{R}^p$ such that Φ is 2-regular at $\bar{u}$ in every direction $v \in K\backslash\{0\}$, and $\bar{v} \in \mathrm{int}K$. Therefore, if $\delta > 0$ is taken small enough, then

$$B(\bar{w},\ \delta) \subset W(K,\ \Pi). \tag{43}$$

Assume now that Φ is 2-regular at $\bar{u}$ in a direction $\bar{v} \in \ker\Phi'(\bar{u})$. We next show that if $p = q$, this assumption can be expected to hold only if $\bar{u}$ is a critical solution of Eq. (1). Indeed, if $\bar{u}$ is a noncritical solution, then for every $v \in \ker\Phi'(\bar{u})$ it holds that $v \in T_{\Phi^{-1}(0)}(\bar{u})$, by (3). Thus, there exist a sequence $\{t_k\}$ of positive reals and a sequence $\{r^k\} \subset \mathbb{R}^p$ such that $\{t_k\} \to 0$, $\|r^k\| = o(t_k)$, and for all k it holds that

$$0 = \|\Phi(\bar{u} + t_k v + r^k)\| = \left\|\Phi'(\bar{u})r^k + \frac{1}{2}t_k^2\Phi''(\bar{u})[v,\ v]\right\| + o(t_k^2).$$

Hence,

$$\frac{1}{2}t_k^2\|\Pi\Phi''(\bar{u})[v,\ v]\| = \|\Pi\Phi'(\bar{u})r^k\| + o(t_k^2) = o(t_k^2),$$

so that

$$\Pi\Phi''(\bar{u})[v,\ v] = 0.$$

Then, from (35) we obtain that $v \in \ker\Psi(\bar{u};\ v)$. If $v \neq 0$, the latter implies that $\Psi(\bar{u};\ v)$ is singular, and hence, Φ cannot be 2-regular at $\bar{u}$ in the direction v. In particular, if $\Phi'(\bar{u})$ is singular, then Φ cannot be 2-regular at $\bar{u}$ in any direction $v \in \ker\Phi'(\bar{u})$. Therefore, for a singular (e.g., nonisolated) but noncritical solution $\bar{u}$, there exists no $\bar{v}$ with the needed properties.

On the other hand, if $\bar{u}$ is a critical solution, the needed $\bar{v}$ can exist even when $p = q$. In the last example considered above, for the unique critical solution $\bar{u} = 0$ any $\bar{v} \in \mathbb{R}^2$ with $\bar{v}_1 \neq 0$ is appropriate. For the mapping Φ from Example 7, for

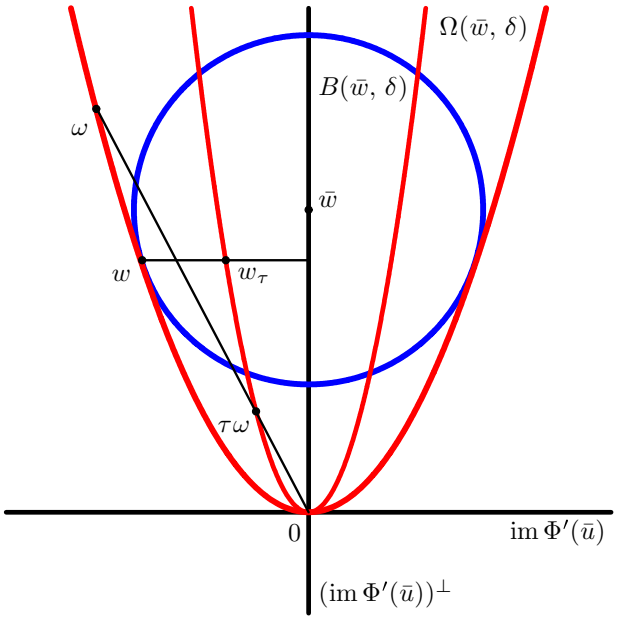

Fig. 4 Illustration of Lemma 2

$\bar{u} = 0$ the appropriate $\bar{v} \in \mathbb{R}^3$ are those satisfying $\bar{v}_1\bar{v}_2\bar{v}_3 \neq 0$. At the same time, for the mapping Φ from Example 8, for the unique critical solution $\bar{u} = 0$ there are no appropriate $\bar{v}$. For the mapping Φ from Example 5, for the unique solution $\bar{u} = 0$ the appropriate $\bar{v} \in \mathbb{R}^p$ are those satisfying $\bar{v}_1 \ldots \bar{v}_p \neq 0$.

Let $\bar{w}$ be defined according to (42) (and hence, $\bar{w} = \Pi\Phi''(\bar{u})[\bar{v},\,\bar{v}]/2$). From inclusion (43), which holds in this case with some $\delta > 0$, it further follows that $W(K,\,\Pi)$ contains the entire collection of parabolic curves specified above, passing through every point of the ball $B(\bar{w},\,\delta)$:

$$\Omega(\bar{w},\,\delta) \subset W(K,\,\Pi), \tag{44}$$

where

$$\Omega(\bar{w},\,\delta) = \{\omega_w(t) \mid w \in B(\bar{w},\,\delta), t \in \mathbb{R}\}. \tag{45}$$

The following Lemma 2, and its proof, are illustrated in Fig. 4.

Lemma 2 *Let $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ be differentiable at $\bar{u} \in \mathbb{R}^p$, and let $\bar{w} \in (\operatorname{im}\Phi'(\bar{u}))^\perp$. Then for every $\delta > 0$ the set $\Omega(\bar{w},\,\delta)$, defined in (45), is star-like with respect to 0.*

Proof We need to show that for every $\omega \in \Omega(\bar{w},\,\delta)$ and every $\tau \in [0,\,1]$, it holds that $\tau\omega \in \Omega(\bar{w},\,\delta)$.

Take any $w \in B(\bar{w},\,\delta)$ and $t \in \mathbb{R}$ such that $\omega = \omega_w(t)$ [they exist according to (45)], and define $w_\tau = \sqrt{\tau}(I - \Pi)w + \Pi w$. Since $\bar{w} \in (\operatorname{im}\Phi'(\bar{u}))^\perp$, we have that $(I - \Pi)\bar{w} = 0$, $\Pi\bar{w} = \bar{w}$. Using also that Π and $I - \Pi$ are the orthogonal projectors

onto two subspaces which are orthogonal complements to each other, we obtain that

$$\|w_\tau - \bar{w}\|^2 = \tau\|(I-\Pi)(w-\bar{w})\|^2 + \|\Pi(w-\bar{w})\|^2 \\ \leq \|(I-\Pi)(w-\bar{w})\|^2 + \|\Pi(w-\bar{w})\|^2 = \|w-\bar{w}\|^2 < \delta^2.$$

Therefore, $w_\tau \in B(\bar{w},\ \delta)$, and hence, by (45), we conclude that

$$\Omega(\bar{w},\ \delta) \ni \omega_{w_\tau}(t\sqrt{\tau}) = t\tau(I-\Pi)w + t^2\tau\Pi w = \tau\omega_w(t) = \tau\omega.$$

□

Remark 2 If $\bar{w} = 0$, then $\Omega(\bar{w},\ \delta) = \mathbb{R}^q$. On the other hand, if $\bar{w} \neq 0$ and $\mathrm{rank}\,\Phi'(\bar{u}) = q-1$, then for every $d \in \mathbb{R}^q$ satisfying $\langle \bar{w},\ d\rangle > 0$, it holds that $\tau d \in \Omega(\bar{w},\ \delta)$ for all $\tau > 0$ small enough, and therefore, $\Omega(\bar{w},\ \delta)$ is asymptotically dense within the half-space $\{w \in \mathbb{R}^q \mid \langle \bar{w},\ w\rangle \geq 0\}$. □

Combining Theorem 4 with (44) and Lemma 2, we finally obtain the following.

Theorem 5 *Let $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ be twice differentiable near $\bar{u} \in \mathbb{R}^p$, and let its second derivative be continuous at $\bar{u}$. Let $\bar{u}$ be a solution of Eq. (1). Let Φ be 2-regular at $\bar{u}$ in a direction $\bar{v} \in \ker \Phi'(\bar{u})$. Let Π be the orthogonal projector onto $(\mathrm{im}\,\Phi'(\bar{u}))^\perp$.*

Then there exist a set $W = W(\bar{v}) \subset \mathbb{R}^q$ and $C = C(\bar{v}) > 0$ such that W is star-like with respect to 0, estimate (37) holds for every $w \in W$, and there exist $\varepsilon = \varepsilon(\bar{v}) > 0$ and $\delta = \delta(\bar{v}) > 0$ such that $B(\varepsilon\Pi\Phi''(\bar{u})[\bar{v},\ \bar{v}],\ \delta) \subset W$.

4 Back to Lagrange multipliers

We now get back to the Lagrange optimality system (6) for the equality-constrained optimization problem (5). We shall relate our new results for general equations to the notions of critical/noncritical Lagrange multipliers [20,23–25,27] (see also the book [26]), and derive some new insights into properties of the latter.

The Lagrange optimality system (6) is a special case of Eq. (1), setting $p = q = n + l$, $u = (x,\ \lambda)$,

$$\Phi(u) = \left(\frac{\partial L}{\partial x}(x,\ \lambda),\ h(x)\right). \tag{46}$$

If $\bar{x} \in \mathbb{R}^n$ is a stationary point of problem (5), then $\Phi^{-1}(0)$ contains the affine manifold $S = \{\bar{x}\} \times \mathscr{M}(\bar{x})$. Therefore,

$$T_S(\bar{u}) \subset T_{\Phi^{-1}(0)}(\bar{u}), \tag{47}$$

where $\bar{u} = (\bar{x},\ \bar{\lambda})$, for every $\bar{\lambda} \in \mathscr{M}(\bar{x})$. Furthermore,

$$\dim S = \dim \ker (h'(\bar{x}))^{\mathrm{T}} = l - \mathrm{rank}\, h'(\bar{x}).$$

In particular, $\dim S > 0$ if and only if the regularity condition

$$\operatorname{rank} h'(\bar{x}) = l \tag{48}$$

is violated.

Since

$$\Phi'(u) = \begin{pmatrix} \dfrac{\partial^2 L}{\partial x^2}(x,\ \lambda) & (h'(x))^{\mathrm{T}} \\ h'(x) & 0 \end{pmatrix}, \tag{49}$$

we obtain that

$$\ker \Phi'(\bar{u}) = \left\{ (\xi,\ \eta) \in Q(\bar{x},\ \bar{\lambda}) \times \mathbb{R}^l \;\middle|\; (h'(\bar{x}))^{\mathrm{T}}\eta = -\frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})\xi \right\}, \tag{50}$$

where the linear subspace $Q(\bar{x},\ \bar{\lambda})$ is given by

$$Q(\bar{x},\ \bar{\lambda}) = \left\{ \xi \in \ker h'(\bar{x}) \;\middle|\; \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})\xi \in \operatorname{im}\left(h'(\bar{x})\right)^{\mathrm{T}} \right\}. \tag{51}$$

From (50) and (51), it can be readily seen that

$$\dim \ker \Phi'(\bar{u}) = \dim Q(\bar{x},\ \bar{\lambda}) + \dim \ker (h'(\bar{x}))^{\mathrm{T}}. \tag{52}$$

Hence, $\dim \ker \Phi'(\bar{u}) > \dim S$ if and only if $Q(\bar{x},\ \bar{\lambda}) \neq \{0\}$, which is equivalent to saying that $\bar{\lambda}$ is a critical Lagrange multiplier (see (7)). In particular, by (2) and (47), if $\bar{\lambda}$ is a noncritical multiplier, then $\bar{u}$ is necessarily noncritical as a solution of (1) with Φ given by (46). Moreover, if $\bar{x}$ is an isolated stationary point, then $\Phi^{-1}(0) = S$ near $\bar{u} = (\bar{x},\ \bar{\lambda})$ for every $\bar{\lambda} \in \mathscr{M}(\bar{x})$. Hence, in this case, $\bar{u}$ is a critical solution of (1) if and only if $\bar{\lambda}$ is a critical Lagrange multiplier.

We summarize the above relations in the following.

Proposition 2 *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be twice differentiable at a stationary point $\bar{x} \in \mathbb{R}^n$ of optimization problem (5), and let $\bar{\lambda} \in \mathbb{R}^l$ be an associated Lagrange multiplier.*

If $\bar{\lambda}$ is a noncritical Lagrange multiplier, then $\bar{u} = (\bar{x},\ \bar{\lambda})$ is a noncritical solution of Eq. (1) with Φ defined in (46).

Moreover, if $\bar{x}$ is an isolated stationary point, then $\bar{u} = (\bar{x},\ \bar{\lambda})$ is a critical solution of (1) if and only if $\bar{\lambda}$ is a critical Lagrange multiplier.

However, if $\bar{x}$ is a nonisolated stationary point, $\bar{\lambda}$ can be critical when $\bar{u} = (\bar{x},\ \bar{\lambda})$ is noncritical. This is illustrated by the following.

Example 11 Consider $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1^2$, $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x) = x_1^2 x_2$. Then $\bar{x} = 0$ is a (nonisolated) stationary point of problem (5), $\mathscr{M}(0) = \mathbb{R}$, and every multiplier in this set is critical.

We have that $\Phi(u) = (2x_1(1+\lambda x_2),\ \lambda x_1^2,\ x_1^2 x_2)$, and $\Phi^{-1}(0)$ is the linear subspace of dimension 2, defined by the equation $x_1 = 0$. As for $\bar{u} = (\bar{x},\ \bar{\lambda})$ we have

$$\Phi'(\bar{u}) = \begin{pmatrix} 2\,0\,0 \\ 0\,0\,0 \\ 0\,0\,0 \end{pmatrix},$$

it holds that $\dim \ker \Phi'(\bar{u}) = 2$, whatever is taken as $\bar{\lambda}$. Therefore, $\bar{u}$ is noncritical. □

Another useful observation is the following.

Remark 3 Note that twice differentiability of f and h at an isolated stationary point $\bar{x}$ of problem (5) implies strict differentiability of Φ defined in (46), at $\bar{u} = (\bar{x},\ \bar{\lambda})$ for every $\bar{\lambda} \in \mathscr{M}(\bar{x})$, with respect to its null set which locally coincides with $S = \{\bar{x}\} \times \mathscr{M}(\bar{x})$. Indeed,

$$\begin{aligned}
&\left\| \frac{\partial L}{\partial x}(x,\ \lambda) - \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})(x-\bar{x}) - (h'(\bar{x}))^{\mathrm{T}}(\lambda - \hat{\lambda}) \right\| \\
&= \left\| \frac{\partial L}{\partial x}(x,\ \hat{\lambda}) + (h'(x))^{\mathrm{T}}(\lambda - \hat{\lambda}) - \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})(x-\bar{x}) - (h'(\bar{x}))^{\mathrm{T}}(\lambda - \hat{\lambda}) \right\| \\
&= \left\| \frac{\partial L}{\partial x}(x,\ \hat{\lambda}) - \frac{\partial L}{\partial x}(\bar{x},\ \hat{\lambda}) - \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})(x-\bar{x}) \right\| + o(\|x-\bar{x}\|) \\
&= \left\| \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \hat{\lambda})(x-\bar{x}) - \frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})(x-\bar{x}) \right\| + o(\|x-\bar{x}\|) \\
&= o(\|x-\bar{x}\|), \\
&\|h(x) - h(\bar{x}) - h'(\bar{x})(x-\bar{x})\| = o(\|x-\bar{x}\|)
\end{aligned}$$

as $x \in \mathbb{R}^n$ tends to $\bar{x}$, and $\lambda \in \mathbb{R}^l$ and $\hat{\lambda} \in \mathscr{M}(\bar{x})$ tend to $\bar{\lambda}$, yielding the needed property. In particular, it follows that Theorem 2 implies Theorem 1, while Proposition 1 implies the corresponding result in [19].

Observe that any stronger smoothness properties of Φ, like strict differentiability at $\bar{u}$, are not implied by twice differentiability of f and h. □

The next task is to understand what the 2-regularity conditions, used above in the case of general equations, mean when the Lagrange optimality system is considered.

Observe that, when $p = q$ (as in the case in question), according to (35), Φ is not 2-regular at $\bar{u}$ in a direction $v \in \mathbb{R}^p$ if and only if there exists $u \in \mathbb{R}^p \backslash \{0\}$ such that

$$u \in \ker \Phi'(\bar{u}), \quad \Phi''(\bar{u})[v,\ u] \in \mathrm{im}\,\Phi'(\bar{u}). \tag{53}$$

Let Φ be defined in (46). We first derive the characterization of 2-regularity of Φ at $\bar{u} = (\bar{x},\ \bar{\lambda})$ in a direction $v = (\xi,\ \eta) \in \mathbb{R}^n \times \mathbb{R}^l$, where $\bar{\lambda} \in \mathscr{M}(\bar{x})$.

Define the linear operator $\Lambda(\bar{x},\,\bar{\lambda}) : Q(\bar{x},\,\bar{\lambda}) \to \mathrm{im}h'(\bar{x})$ putting in correspondence to every $\xi \in Q(\bar{x},\,\bar{\lambda})$ the unique solution of the linear system

$$(h'(\bar{x}))^{\mathrm{T}}\eta = -\frac{\partial^2 L}{\partial x^2}(\bar{x},\,\bar{\lambda})\xi \tag{54}$$

in $\mathrm{im}h'(\bar{x}) = (\ker(h'(\bar{x}))^{\mathrm{T}})^{\perp}$. [This operator is correctly defined, due to (51).] It has been shown in [19, Proposition 3] that

$$\mathrm{im}\Phi'(\bar{u}) = \left\{(x,\,y) \in \mathbb{R}^n \times \mathrm{im}h'(\bar{x}) \,\middle|\, x + (\Lambda(\bar{x},\,\bar{\lambda}))^* y \in (Q(\bar{x},\,\bar{\lambda}))^{\perp}\right\}, \tag{55}$$

where Λ^* stands for the adjoint of a linear operator Λ.

Assuming that f and h are three times differentiable, from (49) we obtain that for $v = (\xi,\,\eta) \in \mathbb{R}^n \times \mathbb{R}^l$ and $u = (x,\,\lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ it holds that

$$\Phi''(\bar{u})[v,\,u] = \begin{pmatrix} \dfrac{\partial^3 L}{\partial x^3}(\bar{x},\,\bar{\lambda})[\xi,\,x] + (h''(\bar{x})[x])^{\mathrm{T}}\eta + (h''(\bar{x})[\xi])^{\mathrm{T}}\lambda \\ h''(\bar{x})[\xi,\,x] \end{pmatrix}. \tag{56}$$

Therefore, according to Φ is not 2-regular in a direction $v = (\xi,\,\eta)$ if and only if there exists $(x,\,\lambda) \in (\mathbb{R}^n \times \mathbb{R}^l)\backslash\{(0,\,0)\}$ such that

$$x \in \ker h'(\bar{x}), \quad \frac{\partial^2 L}{\partial x^2}(\bar{x},\,\bar{\lambda})x + (h'(\bar{x}))^{\mathrm{T}}\lambda = 0, \tag{57}$$

$$\frac{\partial^3 L}{\partial x^3}(\bar{x},\,\bar{\lambda})[\xi,\,x] + (h''(\bar{x})[x])^{\mathrm{T}}\eta + (h''(\bar{x})[\xi])^{\mathrm{T}}\lambda + (\Lambda(\bar{x},\,\bar{\lambda}))^* h''(\bar{x})[\xi,\,x] \in (Q(\bar{x},\,\bar{\lambda}))^{\perp}, \tag{58}$$

$$h''(\bar{x})[\xi,\,x] \in \mathrm{im}h'(\bar{x}). \tag{59}$$

The next lemma gives a sufficient condition for 2-regularity.

Lemma 3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be three times differentiable at $\bar{x} \in \mathbb{R}^n$. For a given pair $(\xi,\,\eta) \in \mathbb{R}^n \times \mathbb{R}^l$, and for some $\bar{\lambda} \in \mathbb{R}^l$, assume that*

$$\frac{\partial^3 L}{\partial x^3}(\bar{x},\,\bar{\lambda})[\xi,\,x,\,x] + \langle \eta,\, h''(\bar{x})[x,\,x]\rangle + 2\langle \Lambda(\bar{x},\,\bar{\lambda})x,\, h''(\bar{x})[\xi,\,x]\rangle \neq 0 \tag{60}$$

for all $x \in Q(\bar{x},\,\bar{\lambda})\backslash\{0\}$ satisfying (59), and

$$\mathrm{im}h'(\bar{x}) + h''(\bar{x})[\xi,\,Q(\bar{x},\,\bar{\lambda})] = \mathbb{R}^l. \tag{61}$$

Then the mapping Φ defined in (46) is 2-regular at $\bar{u} = (\bar{x},\,\bar{\lambda})$ in the direction $v = (\xi,\,\eta)$.

Proof Suppose that, on the contrary, there exists $(x,\ \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l)\backslash\{(0,\ 0)\}$ satisfying (57)–(59). Multiplying the left-hand side of (58) by x (which belongs to $Q(\bar{x},\ \bar{\lambda})$ according to (51) and (57)), we then obtain

$$\frac{\partial^3 L}{\partial x^3}(\bar{x},\ \bar{\lambda})[\xi,\ x,\ x] + \langle \eta,\ h''(\bar{x})[x,\ x]\rangle + \langle \lambda + \Lambda(\bar{x},\ \bar{\lambda})x,\ h''(\bar{x})[\xi,\ x]\rangle = 0.$$

By the second relation in (57), λ is a solution of Eq. (54). Hence, by (59) and the definition of $\Lambda(\bar{x},\ \bar{\lambda})$, it holds that

$$\langle \lambda,\ h''(\bar{x})[\xi,\ x]\rangle = \langle \Lambda(\bar{x},\ \bar{\lambda})x,\ h''(\bar{x})[\xi,\ x]\rangle.$$

Hence, the left-hand side of (60) equals zero, which is only possible if $x = 0$. Then from (57)–(58) we obtain that

$$(h'(\bar{x}))^{\mathrm{T}}\lambda = 0, \quad (h''(\bar{x})[\xi])^{\mathrm{T}}\lambda \in (Q(\bar{x},\ \bar{\lambda}))^{\perp}.$$

By (61), this implies that $\lambda = 0$, giving a contradiction. □

Using the characterization of 2-regularity provided above, one can apply Theorem 4 to the Lagrange optimality system, specifying appropriate cones K. Here, we shall restrict ourselves to deciphering Theorem 4 in this context.

In [19, Proposition 4], the following projector onto an appropriate complementary subspace to $\mathrm{im}\Phi'(\bar{u})$ is constructed: for $(\chi,\ y) \in \mathbb{R}^n \times \mathbb{R}^l$

$$\Pi(\chi,\ y) = \left(P_Q^{\perp}\chi + (\Lambda(\bar{x},\ \bar{\lambda}))^* P^{\perp}_{\mathrm{im}h'(\bar{x})}y,\ P^{\perp}_{(\mathrm{im}h'(\bar{x}))^{\perp}}y\right).$$

With this choice of Π, the set $W(K,\ \Pi)$ defined in (36) consists of $(\chi,\ y) \in \mathbb{R}^n \times \mathbb{R}^l$ such that there exists $(x,\ \lambda) \in K$ satisfying

$$\frac{\partial^2 L}{\partial x^2}(\bar{x},\ \bar{\lambda})x + (h'(\bar{x}))^{\mathrm{T}}\lambda + \frac{1}{2}P_Q^{\perp}\left(\frac{\partial^3 L}{\partial x^3}(\bar{x},\ \bar{\lambda})[x,\ x] + 2(h''(\bar{x})[x])^{\mathrm{T}}\lambda\right)$$
$$+ (\Lambda(\bar{x},\ \bar{\lambda}))^* P^{\perp}_{\mathrm{im}h'(\bar{x})}h''(\bar{x})[x,\ x] = \chi, \tag{62}$$

$$h'(\bar{x})x + \frac{1}{2}P^{\perp}_{(\mathrm{im}h'(\bar{x}))^{\perp}}h''(\bar{x})[x,\ x] = y. \tag{63}$$

Given the constructions above, Theorem 4 results in the following.

Proposition 3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be three times differentiable near $\bar{x} \in \mathbb{R}^n$, and let their third derivatives be continuous at $\bar{x}$. Let $\bar{x}$ be a stationary point of problem (5), and let $\bar{\lambda} \in \mathscr{M}(\bar{x})$. Let $K \subset \mathbb{R}^n \times \mathbb{R}^l$ be a closed cone such that for every $(\xi,\ \eta) \in K\backslash\{(0,\ 0)\}$ there exists no $(x,\ \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l)\backslash\{(0,\ 0)\}$ satisfying (57)–(59).*

Then there exist $\varepsilon = \varepsilon(K) > 0$ and $C = C(K) > 0$ such that for every $w = (\chi,\ y) \in B(0,\ \varepsilon)$ satisfying (62)–(63) with some $(x,\ \lambda) \in K$, there exists $(x(w),\ \lambda(w)) \in \mathbb{R}^n \times \mathbb{R}^l$ satisfying

$$\frac{\partial L}{\partial x}(x,\,\lambda)=\chi,\quad h(x)=y,$$

and

$$\|(x(w)-\bar{x},\,\lambda(w)-\bar{\lambda})\|\le C\left(\left\|\left(P^{\perp}_{Q^{\perp}}\chi-(\Lambda(\bar{x},\,\bar{\lambda}))^{*}P^{\perp}_{\mathrm{im}\,h'(\bar{x})}y,\;P^{\perp}_{\mathrm{im}\,h'(\bar{x})}y\right)\right\|\right.$$
$$\left.+\sqrt{\left\|\left(P^{\perp}_{Q}\chi+(\Lambda(\bar{x},\,\bar{\lambda}))^{*}P^{\perp}_{\mathrm{im}\,h'(\bar{x})}y,\;P^{\perp}_{(\mathrm{im}\,h'(\bar{x}))^{\perp}}y\right)\right\|}\right).$$

Proposition 3 establishes Hölder stability of primal–dual solutions of optimization problem (5) subject to wide classes of canonical perturbations. For other results on Hölder stability of solutions and solution sets, see, e.g., [1,13,30–32,34] and [7, Chapter 4]. One feature distinguishing Proposition 3 from the cited works is that it deals with stability of a specific dual solution. A result related to Proposition 3 was established in [19], but for directional (one-dimensional) perturbations only.

We next study the cases when Φ can (or cannot) be 2-regular at $\bar{u}=(\bar{x},\,\bar{\lambda})$ in some direction $v=(\xi,\,\eta)\in\ker\Phi'(\bar{u})$. Note that if a direction $v\in\ker\Phi'(\bar{u})$ for which 2-regularity holds exists, then Theorem 5 guarantees stability of the solution $\bar{u}$ (with this specific $\bar{\lambda}\in\mathscr{M}(\bar{x})$!) with respect to a wide class of right-hand side perturbations of the Lagrange optimality system.

According to Proposition 2, if $\bar{\lambda}$ is a noncritical Lagrange multiplier, then $\bar{u}=(\bar{x},\,\bar{\lambda})$ is a noncritical solution of Eq. (1). Furthermore, as discussed above, if $\bar{u}$ is a noncritical solution and $\Phi'(\bar{u})$ is singular, then Φ cannot be 2-regular at $\bar{u}$ in any direction $v\in\ker\Phi'(\bar{u})$. Therefore, according to (52), in the case of violation of the constraints regularity condition (48) we can expect 2-regularity in the needed directions only when $\bar{\lambda}$ is a critical multiplier, i.e., when $Q(\bar{x},\,\bar{\lambda})\ne 0$.

Recall also that according to (50) and (51), v belongs to $\ker\Phi'(\bar{u})$ if and only if

$$\xi\in\ker h'(\bar{x}),\quad \frac{\partial^2 L}{\partial x^2}(\bar{x},\,\bar{\lambda})\xi+(h'(\bar{x}))^{\mathrm{T}}\eta=0. \tag{64}$$

We next consider some special cases, with conclusions summarized in Proposition 4 below. Observe first that, if $\xi=0$, then relations (57)–(59) are satisfied by $x=0$ and by every $\lambda\in\ker(h'(\bar{x}))^{\mathrm{T}}$, where the subspace $\ker(h'(\bar{x}))^{\mathrm{T}}$ is nontrivial when the constraints regularity condition (48) does not hold. Hence, 2-regularity is not possible in such directions.

Furthermore, let $\xi\ne 0$, and consider the case of $\dim Q(\bar{x},\,\bar{\lambda})=1$, i.e., $Q(\bar{x},\,\bar{\lambda})$ is spanned by some $\bar{\xi}\in\mathbb{R}^n\backslash\{0\}$ (in this case, $\bar{\lambda}$ is referred to as a multiplier critical of order 1 [27]). Then (51) and (64) imply that ξ is a nonzero multiple of $\bar{\xi}$, and taking $x=0$ in (57)–(59) reduces these relations to

$$(h'(\bar{x}))^{\mathrm{T}}\lambda=0,\quad \langle h''(\bar{x})[\bar{\xi},\,\bar{\xi}],\,\lambda\rangle=0.$$

If $h''(\bar{x})[\bar{\xi},\,\bar{\xi}]\in\mathrm{im}\,h'(\bar{x})$, then this system always has a nontrivial solution when the constraints regularity condition (48) is violated. Otherwise, this system reduces

to a system consisting of $\operatorname{rank} h'(\bar{x}) + 1$ linearly independent linear equations in l variables. In particular, if $\operatorname{rank} h'(\bar{x}) \leq l-2$, then 2-regularity in the needed directions is not possible. This case is especially difficult, as it allows for nonisolated critical multipliers.

Suppose now that $l = 1$. Then violation of constraints regularity condition (48) means full degeneracy: $h'(\bar{x}) = 0$. Then it holds that

$$Q(\bar{x}, \bar{\lambda}) = \ker \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}), \quad \Lambda(\bar{x}, \bar{\lambda}) = 0.$$

Therefore, system (57)–(59) takes the form

$$\begin{aligned} &\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})x = 0, \quad \frac{\partial^3 L}{\partial x^3}(\bar{x}, \bar{\lambda})[\xi, x] + \eta(h''(\bar{x})[x])^{\mathrm{T}} \\ &\qquad + \lambda(h''(\bar{x})[\xi])^{\mathrm{T}} \in \left(\ker \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\right)^{\perp}, \\ &h''(\bar{x})[\xi, x] = 0. \end{aligned}$$

If $\xi \neq 0$ and $\dim Q(\bar{x}, \bar{\lambda}) = 1$, then these relations reduce to the system

$$t\frac{\partial^3 L}{\partial x^3}(\bar{x}, \bar{\lambda})[\bar{\xi}, \bar{\xi}, \bar{\xi}] + (t\eta + \lambda)h''(\bar{x})[\bar{\xi}, \bar{\xi}] = 0, \quad th''(\bar{x})[\bar{\xi}, \bar{\xi}] = 0$$

with respect to $(t, \lambda) \in \mathbb{R} \times \mathbb{R}$, where we set $x = t\bar{\xi}$. This system has only the trivial solution if and only if

$$h''(\bar{x})[\bar{\xi}, \bar{\xi}] \neq 0. \tag{65}$$

Therefore, in the case of $l = 1$, and when constraints regularity condition (48) does not hold and $Q(\bar{x}, \bar{\lambda})$ is spanned by $\bar{\xi}$, we conclude that Φ is 2-regular at $\bar{u}$ in the directions $v = (\bar{\xi}, \eta) \in \ker \Phi'(\bar{u})$ for all $\eta \in \mathbb{R}$ if and only if (65) holds.

We summarize the above considerations in the following.

Proposition 4 *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ be three times differentiable at a stationary point $\bar{x} \in \mathbb{R}^n$ of optimization problem (5), and let $\bar{\lambda}$ be a Lagrange multiplier associated to $\bar{x}$. Let $Q(\bar{x}, \bar{\lambda})$ be spanned by some $\bar{\xi} \in \mathbb{R}^n \backslash \{0\}$, i.e., $\bar{\lambda}$ is a critical multiplier of order 1.*

If $\operatorname{rank} h'(\bar{x}) = l - 1$, then $\ker \Phi'(\bar{u})$ contains elements of the form $v = (\bar{\xi}, \eta)$ with some $\eta \in \mathbb{R}^l$, and Φ is 2-regular at $\bar{u}$ in every such direction if and only if $h''(\bar{x})[\bar{\xi}, \bar{\xi}] \notin \operatorname{im} h'(\bar{x})$.

If $\operatorname{rank} h'(\bar{x}) \leq l-2$, then Φ cannot be 2-regular at $\bar{u}$ in any direction $v \in \ker \Phi'(\bar{u})$.

If $h'(\bar{x}) = 0$, and $l \geq 2$ or (65) does not hold, then Φ cannot be 2-regular at $\bar{u}$ in any direction $v \in \ker \Phi'(\bar{u})$.

Example 12 (DEGEN 20101 [9]) Consider $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x) = x^2$. Then $\bar{x} = 0$ is the unique solution of problem (5), $h'(\bar{x}) = 0$, and $\mathscr{M}(\bar{x}) = \mathbb{R}$. Furthermore,

$$\frac{\partial^2 L}{\partial x^2}(x,\ \lambda) = 2(1+\lambda),$$

and hence, the only critical multiplier is $\bar{\lambda} = -1$.

For the mapping Φ defined in (46), equation (9) with right-hand side perturbation $w = (\chi,\ y) \in \mathbb{R} \times \mathbb{R}$ (corresponding to canonical perturbation of problem (5)) has the solutions $(x(w),\ \lambda(w)) = (\pm\sqrt{y},\ -1 \pm \chi/(2\sqrt{y}))$ when $y > 0$, and no solutions for other $w \neq 0$. If $\chi = o(\sqrt{y})$, both these solutions tend to $\bar{u} = (\bar{x},\ \bar{\lambda})$ as $w \to 0$. Other points in $\{\bar{x}\} \times \mathscr{M}(\bar{x})$ can be stable only subject to special perturbations w satisfying $y = O(\chi^2)$, thus with $w/\|w\|$ tending to $d = (1,\ 0)$.

Observe that here $\dim Q(\bar{x},\ \bar{\lambda}) = 1$, $l = 1$, and (65) holds. Hence, according to Proposition 4, Φ is 2-regular at $\bar{u}$ in the directions $v = (\bar{\xi},\ \eta) \in \ker \Phi'(\bar{u})$ for every $\eta \in \mathbb{R}$. □

We conclude by mentioning that the case when $\dim Q(\bar{x},\ \bar{\lambda}) \geq 2$ (i.e., when $\bar{\lambda}$ is critical of order higher than 1) opens wide possibilities for 2-regularity in the needed directions, and hence, for stability subject to wide classes of perturbations.

Finally, it is worth making the following simple but useful observation: all the results and discussions above readily extend to KKT systems involving inequality constraints (arising from optimization or variational problems), to KKT-type systems for equilibrium problems (including GNEPs), and to more general complementarity systems, assuming that solution in question satisfies strict complementarity. Near such solutions, complementarity systems naturally (without using any complementarity functions) reduce to a smooth system of equations. Such cases have already been illustrated by Examples 9 and 10. For instance, a critical solution $\bar{u} = (\bar{x},\ 1)$ in Example 9 can be treated the same way as the unique critical solution in Example 12, with the same conclusions for the corresponding mapping Φ defined in (30).

Acknowledgements The authors thank the two anonymous referees for their comments, which improved the original presentation.

References

1. Alt, W.: Lipschitzian perturbations of infinite optinization problems. In: Fiacco, A.V. (ed.) Mathematical Programming with Data Perturbations, pp. 7–21. M. Dekker, New York (1983)
2. Arutyunov, A.V.: Optimality Conditions: Abnormal and Degenerate Problems. Kluwer Academic Publishers, Dordrecht (2000)
3. Aubin, J.P., Ekeland, I.: Applied Nonlinear Analysis. Wiley, Hoboken (1984)
4. Avakov, E.R.: Extremum conditions for smooth problems with equality-type constraints. USSR Comput. Math. Math. Phys. **25**, 24–32 (1985)
5. Avakov, E.R.: Theorems on estimates in the neighborhood of a singular point of a mapping. Math. Notes **47**, 425432 (1990)
6. Behling, R., Iusem, A.: The effect of calmness on the solution set of systems of nonlinear equations. Math. Program. **137**, 155–165 (2013)
7. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
8. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
9. DEGEN. http://w3.impa.br/~optim/DEGEN_collection.zip
10. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings, 2nd edn. Springer, New York (2014)

11. Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. Ann. Oper. Res. **175**, 177–211 (2010)
12. Fernández, D., Izmailov, A.F., Solodov, M.V.: Sharp primal superlinear convergence results for some Newtonian methods for constrained optimization. SIAM J. Optim. **20**, 3312–3334 (2010)
13. Gfrerer, H., Klatte, D.: Lipschitz and Hölder stability of optimization problems and generalized equations. Math. Program. (2015). doi:10.1007/s10107-015-0914-1
14. Gfrerer, H., Mordukhovich, B.S.: Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions. SIAM J. Optim. **25**, 2081–2119 (2015)
15. Gfrerer, H., Outrata, J.V.: On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications. Optimization (2016). doi:10.1080/02331934.2015.1066372
16. Hager, W.W., Gowda, M.S.: Stability in the presence of degeneracy and error estimation. Math. Program. **85**, 181–192 (1999)
17. Halkin, H.: Implicit functions and optimization problems without continuous differentiability of the data. SIAM J. Control **12**, 229–236 (1974)
18. Izmailov, A.F.: On some generalizations of Morse lemma. Proc. Steklov Inst. Math. **220**, 138–153 (1998)
19. Izmailov, A.F.: On the analytical and numerical stability of critical Lagrange multipliers. Comput. Math. Math. Phys. **45**, 930–946 (2005)
20. Izmailov, A.F., Kurennoy, A.S., Solodov, M.V.: A note on upper Lipschitz stability, error bounds, and critical multipliers for Lipschitz-continuous KKT systems. Math. Program. **142**, 591–604 (2013)
21. Izmailov, A.F., Solodov, M.V.: Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications. Math. Program. **89**, 413–435 (2001)
22. Izmailov, A.F., Solodov, M.V.: The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. Math. Oper. Res. **27**, 614–635 (2002)
23. Izmailov, A.F., Solodov, M.V.: On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions. Math. Program. **117**, 271–304 (2009)
24. Izmailov, A.F., Solodov, M.V.: Examples of dual behaviour of Newton-type methods on optimization problems with degenerate constraints. Comput. Optim. Appl. **42**, 231–264 (2009)
25. Izmailov, A.F., Solodov, M.V.: On attraction of linearly constrained Lagrangian methods and of stabilized and quasi-Newton SQP methods to critical multipliers. Math. Program. **126**, 231–257 (2011)
26. Izmailov, A.F., Solodov, M.V.: Newton-Type Methods for Optimization and Variational Problems. Springer series in operations research and financial engineering. Springer, Switzerland (2014)
27. Izmailov, A.F., Solodov, M.V.: Critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it. TOP **23**, 1–26 (2015)
28. Izmailov, A.F., Solodov, M.V.: Rejoinder on: critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it. TOP **23**, 48–52 (2015)
29. Izmailov, A.F., Uskov, E.I.: Attraction of Newton method to critical Lagrange multipliers: fully quadratic case. Math. Program. **152**, 33–73 (2015)
30. Klatte, D.: A Note on Quantitative Stability Results in Nonlinear Optimization. Seminarbericht Nr. 90. Sektion Mathematik, pp. 77–86. Humboldt-Universität zu Berlin, Berlin (1987)
31. Klatte, D.: On quantitative stability for non-isolated minima. Control Cybern. **23**, 183–200 (1994)
32. Kummer, B.: Inclusions in general spaces: Hoelder stability, solution schemes and Ekelands principle. J. Math. Anal. Appl. **358**, 327–344 (2009)
33. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
34. Shapiro, A.: Perturbation analysis of optimization problems in Banach spaces. Numer. Func. Anal. Optim. **13**, 97–116 (1992)