APPROXIMATE SOLUTIONS FOR ABSTRACT INEQUALITY SYSTEMS

CHONG LI† AND K. F. NG‡

Abstract. We consider conic inequality systems of the type $F(x) \geq_K 0$, with approximate solution $x_0$ associated to a parameter $\tau$, where $F$ is a twice Fréchet differentiable function between Hilbert spaces $X$ and $Y$, and $\geq_K$ is the partial order in $Y$ defined by a nonempty convex (not necessarily closed) cone $K \subseteq Y$. We prove that, under the suitable conditions, the system $F(x) \geq_K 0$ is solvable, and the ratio of the distance from $x_0$ to the solution set $S$ over the distance from $F(x_0)$ to the cone $K$ has an upper bound given explicitly in terms of $\tau$ and $x_0$. We show that the upper bound is sharp. Applications to analytic function inequality/equality systems on Euclidean spaces are given, and the corresponding results of Dedieu [SIAM J. Optim., 11 (2000), pp. 411–425] are extended and significantly improved.

Key words. conic inequality systems, approximate solution, stability

AMS subject classifications. 47J99, 90C31

DOI. 10.1137/120885176

1. Introduction. The constraint sets in most constrained optimization problems can be described as conic inequality systems of the following type:

\begin{equation}
F(x) \geq_K 0,
\end{equation}

where $F$ is a function between Banach spaces $X$, $Y$ and $Y$ is endowed with the partial order (or preorder) $\geq_K$ defined by a nonempty convex cone $K \subseteq Y$. One of the most important issues for conic inequality systems is the issue of error bounds. We say that a constant $c$ is a (local) error bound for (1.1) at $x_0 \in X$ if there exists a neighborhood $U$ of $x_0$ such that

\begin{equation}
d(x, S) \leq c \cdot e(x) \quad \text{for any } x \in U,
\end{equation}

where $S$ denotes the solution set of (1.1) (consisting of all $x$ satisfying (1.1)), $d(\cdot, S)$ is the distance function associated with $S$ defined by

\[d(x, S) := \inf \{\|x - z\| : z \in S\}\]

for each $x \in X$, and $e(\cdot)$ is some suitable residual function measuring a degree of violation with respect to (1.1).

In the case when $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, and $K = \mathbb{R}^m_{\geq 0}$ (consisting of all $m$-vector $(\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$ such that $\alpha_i \leq 0$ for all $i$), system (1.1) is reduced to the finitely many inequalities. In this case, the most important and celebrated result regarding

*Received by the editors July 19, 2012; accepted for publication (in revised form) March 19, 2013; published electronically June 20, 2013.
†Department of Mathematics, Zhejiang University, Hangzhou 310027, P. R. China (cli@zju.edu.cn). This author’s work was supported in part by the National Natural Science Foundation of China (grant 11171300) and by Zhejiang Provincial Natural Science Foundation of China (grant Y6110006).
‡Department of Mathematics (and IMS), Chinese University of Hong Kong, Hong Kong, P. R. China (kfng@math.cuhk.edu.hk). This author’s work was supported by an earmarked grant (GRF) from the Research Grant Council of Hong Kong (Projects CUHK 403110 and 402612).
the error bound is due to Hoffman, who proved that if $F$ is affine and the solution set $S$ is nonempty, then there exists $c > 0$ such that (1.2) holds with $e(\cdot) = d(F(\cdot), R^m_\infty)$ and $U = X$ (see [6, 7]). Extensions of Hoffman’s result to convex inequalities have been well established in the literature; see, for example, [8, 9, 10, 11, 12, 19, 20, 28] and references therein. However, study of error bound for nonconvex inequalities is not so rich. In the special case when (1.1) is a single inequality system $F(x) \leq 0$ (i.e., $m = 1$), Luo and Luo [13] and Luo and Pang [14] established global error bound inequality (1.2), respectively for polynomial and analytic inequality but with an unknown fractional exponents to $d(F(x), R_\infty)$. Other related works in this direction can be found in [15, 17, 18, 27].

For the general case when $K$ is a closed convex cone, Ng and Yang [16] established some error bound results for abstract linear inequality systems (i.e., $F$ is affine) in general Banach spaces, which in particular extend the corresponding results in finite dimensional spaces due to Robinson [21]. For the general differential system, Robinson [22] proved that if the so-called Robinson constraint qualification (RCQ) holds at $x_0 \in S$, then there exist $c > 0$ and a neighborhood $U$ of $x_0$ such that (1.2) holds with $e(x) = d(F(x), K)$. By weakening the RCQ assumption, this result has been extended in [2] by He and Sun.

We observe that all results mentioned above are done under two additional assumptions: (a) $K$ is closed, and (b) $S$ is preassumed nonempty with $x_0 \in S$. In this paper we assume that $X$ and $Y$ are Hilbert spaces and that system (1.1) has an “approximate solution” $x_0$ in the sense that $F(x_0) + c \geq K 0$ for some $c \in Y$ with small norm or small residual value, but we drop the two additional assumptions mentioned above. Given an approximate solution $x_0$, the main concern of this paper is to show, under suitable assumptions, that the solution set $S$ of (1.1) is nonempty and its distance $d(x_0, S)$ from $x_0$ has an upper estimate in terms of $x_0$ together with a parameter. In a more concrete situation we will consider, as applications, the following system with $Y = \mathbb{R}^l$,

$$
\begin{align*}
&f_i(x) > 0, \quad i = 1, \ldots, l_S, \\
&f_i(x) \geq 0, \quad i = l_S + 1, \ldots, l,
\end{align*}
$$

and the system with $Y = \mathbb{R}^{m+l}$ involving inequality and equality constraints,

$$
\begin{align*}
&f_i(x) > 0, \quad i = 1, \ldots, l_S, \\
&f_i(x) \geq 0, \quad i = l_S + 1, \ldots, l, \\
g_j(x) = 0, \quad j = 1, \ldots, m.
\end{align*}
$$

Here $f_i$ and $g_j$ are real-valued functions on $X$, and $l, l_S, m$ are nonnegative integers such that $0 \leq l_S \leq l$. Associated with (1.3) and (1.4), let $f : X \rightarrow \mathbb{R}^l$, $g : X \rightarrow \mathbb{R}^m$, and $(f, g) : X \rightarrow \mathbb{R}^l \times \mathbb{R}^m$ be defined, respectively, by

$$
\begin{align*}
f(x) := (f_1(x), \ldots, f_l(x)), \\
g(x) := (g_1(x), \ldots, g_m(x)), \quad \text{and} \quad (f, g)(x) := (f(x), g(x))
\end{align*}
$$

for each $x \in X$. We assume throughout that the involved functions $f_i$, $g_i$ (and so $f$, $g$ and $(f, g)$) are analytic (at least locally around the point $x$ under consideration); thus we have the following real constants:

$$
\begin{align*}
\Gamma(f; x) := \sup_{k \geq 2} \left\| \frac{f^{(k)}(x)}{k!} \right\| \text{ and } \Gamma(f, g; x) := \sup_{k \geq 2} \left\| \frac{(f, g)^{(k)}(x)}{k!} \right\|,
\end{align*}
$$
where, e.g., \( f^{(k)}(x) \) denotes the \( k \)th derivative of \( f \) at \( x \) and
\[
\| f^{(k)}(x) \| := \sup\{ \| f^{(k)}(x)(u_1, \ldots, u_k) \| : \text{each } u_i \in X \text{ with } \| u_i \| \leq 1 \}.
\]

Let \( f'(x)^* \) denote the conjugate (adjoint) of \( f'(x) \); thus \( f'(x)f''(x)^* \) is a linear operator from \( \mathbb{R}^l \) into itself and is, as usual, represented as an \( l \times l \) matrix. For a matrix \( A \), we use \( A^\dagger \) to denote the Moore–Penrose generalized inverse of \( A \), and \( \| A \| \) the usual operator norm of \( A \). In connection with (1.4), one defines an \((l + m) \times (l + m)\) matrix \( G(x) \) for \( x \in X \) by
\[
G(x) := \begin{pmatrix}
( f'(x)f'(x)^* + \text{Diag}(4f(x)^+) ) & ( f'(x)g'(x)^* ) \\
( g'(x)f'(x)^* ) & ( g'(x)g'(x)^* )
\end{pmatrix},
\]
where \( \text{Diag}(4f(x)^+) \) is the diagonal matrix with diagonal entries \( 4f_i(x)^+ \), \( i = 1, \ldots, l \).

Here and throughout, we use the notations \( a^+ := \max\{a, 0\} \) and \( a^- := \max\{-a, 0\} \) for any real number \( a \). Let
\[
\delta(f; x) := \| ( f'(x)f'(x)^* + \text{Diag}(4f(x)^+) ) \|_F^\frac{1}{2} \quad \text{and} \quad \delta(f, g; x) := \| G(x) \|_F^\frac{1}{2}.
\]

One of the main results of Dedieu in [3] for system (1.4) with \( X = \mathbb{R}^n \) can be stated as follows: if \( x_0 \in \mathbb{R}^n \) and \( \sigma := \sum_{k=0}^{\infty} \left( \frac{1}{2} \right)^{2k-1} = 1.63284 \ldots \) are such that
\[
f_i(x_0) > (\sigma \delta(f; x_0) \| f(x_0)^{-} \|)^2 \quad \text{for each } i = 1, \ldots, l_S,
\]
(1.11) the matrix \( G(x_0) \) is nonsingular,

and
\[
\| f(x_0)^{-} \|^2 + \| g(x_0) \|^2 \leq \frac{1}{8(1 + \Gamma(f, g; x_0))\delta(f, g; x_0) \max\{1, \delta(f, g; x_0)\}},
\]
(1.12) \( \| f(x_0)^{-} \|^2 + \| g(x_0) \|^2 \leq \sigma \delta(f, g; x_0) (\| f(x_0)^{-} \|^2 + \| g(x_0) \|^2)^\frac{1}{2} \).

Similar treatments have also been done for system (1.3) with \( X = \mathbb{R}^n \) by Dedieu; see [3] and references therein for more details and background information. From some points of view, assumptions (1.10) and (1.11) are too restrictive; for example, when all \( f_i \) with \( 1 \leq i \leq l \) and all \( g_j \) are zero functions, because the matrix \( G(x_0) \) is clearly singular even (1.10) itself trivially entails that \( x_0 \) is a solution of (1.4), the results of Dedieu mentioned above cannot be applied.

More generally, we study the corresponding issues for the general system (1.1) for which we assume from now on that \( F \) is twice Fréchet differentiable (at least locally around \( x \in X \) with the first and the second derivatives \( F'(\cdot) \) and \( F''(\cdot) \)). We say that \( F \) satisfies the \( \gamma \)-condition at \( x_0 \in X \) on the open ball \( B(x_0, r) \) (with center \( x_0 \in X \) and radius \( r > 0 \)) if \( 0 \leq r \gamma \leq 1 \) and
\[
\| F'(x_0)^{\dagger} F''(x) \| \leq \frac{2\gamma}{(1 - \gamma \| x - x_0 \|)^3} \quad \text{for each } x \in B(x_0, r),
\]
(1.14)
where $F'(x_0)^d$ is the Moore–Penrose generalized inverse of $F'(x_0)$ and the norm of $\|F'(x_0)^d F''(x_0)\|$ is defined similarly as done in (1.7); see section 3 for more notation matters. The $\gamma$-condition (in the special case when $F'(x_0)$ is nonsingular) was introduced by Wang (see [25, 26]) for improving the famous point estimate theory presented by Smale [24]. An important example is of course the case when $F$ is analytic: if $F: X \to Y$ is analytic on $B(x_0, r)$, then $F$ satisfies the $\gamma$-condition at $x_0$ on $B(x_0, r_0)$ with any $\gamma \in [\gamma_F(x_0), +\infty)$ and $r_0 := \min\{r, \frac{1}{2}\}$, where the quantity $\gamma_F(x_0)$ was introduced by Shub and Smale [23] (see also [1]) and is defined by

\[
\gamma_F(x_0) := \sup_{k \geq 2} \frac{\|F'(x_0)^d F^{(k)}(x_0)\|}{k!}.
\]

Our main results are described in terms of the $\gamma$-condition and cover cases when $K$ is not necessarily closed and/or the case when the nonsingularity is not assumed. As we will see later, an advantage of our approach is that the corresponding constant $c$ in the error bound result can be pointwise determined explicitly. In particular, our estimate is optimal in some sense; see Example 2.1. As applications, we extend and improve the results of Dedieu mentioned above for system (1.3) and system (1.4). In particular, we show that for (1.13) to hold, both (1.11) and (1.12) can be replaced by weaker assumptions, while (1.10) is not needed, as explained before Theorem 2.2 in section 2. Another by-product of our main results can be cast in the stability point of view (Robinson initiated the study of stability in [21, 22] for (1.1)). Under suitable assumptions, we show that if (1.1) is solvable, then the perturbed systems

\[
F(x) - y \geq_K 0
\]

(with small enough $\|y\|$) are also solvable, and the solution map $y \mapsto S(y)$ enjoys some calmness property.

The paper is organized as follows. In section 2, we summarize the main theorems and their corollaries. Some basic concepts and known facts needed in the rest of the paper are listed in section 3. Section 4 contains the proofs of the main theorems, while applications are provided in section 5.

2. Main results. Throughout the paper let $X$ and $Y$ be (real) Hilbert spaces, and $K \subseteq Y$ be a cone which defines a binary relation $\geq_K$ by

\[
y_1 \geq y_2 \Leftrightarrow y_1 - y_2 \in K.
\]

Given a bounded linear operator $A$ (or a matrix), let $\ker A$, $\im A$, and $A^*$ respectively denote the kernel, the image, and the conjugate operator of $A$. Let $A^\dagger$ denote the Moore–Penrose generalized inverse of $A$ when it exists uniquely (e.g., if $\im A$ is closed).

The following theorem is our main result regarding the solution set $S := \{x \in X : F(x) \in K\}$ of system (1.1), where we assume that $F$ is a function from $X$ into $Y$. The usual convention that $\frac{a}{a} = +\infty$ is always adopted for any $a > 0$. Further, for a nonempty set $D$ in $X$ (or in other normed spaces), we use the notation $\|D\|$ to denote the distance from the origin to $D$, namely $\|D\| := \inf\{\|u\| : u \in D\}$, with the convention that $\|\emptyset\| = +\infty$. The closure and the relative interior of $D$ are denoted by $\overline{D}$ and $\ri D$, respectively. For an element $x \in X$, the notation $D - x$ stands for the set consisting of all elements $d - x$ with $d \in D$. 
THEOREM 2.1. Let $x_0 \in X$ and $\bar{r}, \gamma \in [0, +\infty)$. Let $F : X \to Y$ be continuous on $B(x_0, \bar{r})$ and twice Fréchet differentiable on $B(x_0, \bar{r})$ with the following properties:

1. $\operatorname{im} F'(x_0)$ is closed;
2. $F$ satisfies the $\gamma$-condition at $x_0$ on $B(x_0, \bar{r})$;
3. $(2.1) \quad K \cup \{F(x_0)\} \cup \operatorname{im} F''(x) \subseteq \operatorname{im} F'(x_0)$ for each $x \in B(x_0, \bar{r})$.

Let $\xi, r^*$ be defined by

$$(2.2) \quad \xi := \|F'(x_0)\|^\top(K - F(x_0))$$

and

$$(2.3) \quad r^* := \left\{ \begin{array}{l} \frac{1 + \gamma \xi - \sqrt{(1 + \gamma \xi)^2 - 8 \gamma \xi}}{4 \gamma}, \\
\xi, \quad 0 < \gamma \leq \frac{3 - 2 \sqrt{2}}{\xi}, \\
\gamma = 0. \end{array} \right.$$ 

Suppose that $r^* \leq \bar{r}$ and that there exists $\tau \in \mathbb{R}$ such that

$$(2.4) \quad 1 < \tau \leq \frac{2 \sqrt{2} + 2}{2} \quad \text{and} \quad \xi \leq \frac{\tau - 1}{\gamma \tau (2\tau - 1)}.$$ 

Then there exists $x^* \in X$ such that $F(x^*) \in \overline{K}$ and

$$(2.5) \quad \|x_0 - x^*\| \leq \tau \|F'(x_0)\|^\top d(F(x_0), K).$$

Moreover, if

(a) $K$ is closed
or
(b) $r^* < \bar{r}, \xi < \frac{3 - 2 \sqrt{2}}{\xi}$, and the relative interior $\text{ri} K$ of $K$ is nonempty,

then $S$ is nonempty and

$$(2.6) \quad d(x_0, S) \leq \tau \|F'(x_0)\|^\top d(F(x_0), K).$$

Remark 2.1. If $\xi = 0$ in (2.2), then assumption (2.1) entails that $F(x_0) \in \overline{K}$. In fact, assume that $\xi = 0$ in (2.2) and (2.1) holds. Then, $\|F'(x_0)^\top(K - F(x_0))\| = 0$, and, by definition, we can choose $\{u_n\} \subseteq K$ such that $F'(x_0)^\top(u_n - F(x_0)) \to 0$. Thus $F'(x_0)F'(x_0)^\top(u_n - F(x_0)) \to 0$. Since $F'(x_0)F''(x_0)$ is simply the projection $\Pi_{\text{im} F'(x_0)}$ (see (3.3) of the next section) and since $u_n - F(x_0)$ lies in $\operatorname{im} F'(x_0)$ by (2.1), it follows that $u_n - F(x_0) \to 0$. Therefore $F(x_0) \in \overline{K}$, as claimed.

Similarly, if $\gamma = 0$, then the $\gamma$-condition at $x_0$ on $B(x_0, \bar{r})$ together with assumption (2.1) entails that $F''(x) = 0$ for each $x \in B(x_0, \bar{r})$; that is, $F$ is affine on $B(x_0, \bar{r})$.

The following example shows that the bound $\tau \|F'(x_0)\|^\top$ before $d(F(x_0), K)$ in (2.6) cannot be improved.

Example 2.1. Let $\gamma \in (0, +\infty)$ and $\tau \in (1, \frac{2 \sqrt{2}}{2})$ (so $\frac{\tau - 1}{\gamma \tau (2\tau - 1)} \leq 3 - 2 \sqrt{2}$; see (4.2) and (4.3)). Define $F : (-\infty, \frac{1}{\gamma}) \subseteq \mathbb{R} \to \mathbb{R}$ by

$$(2.7) \quad F(x) := \frac{\tau - 1}{\gamma \tau (2\tau - 1)} - x + \frac{\gamma x^2}{1 - \gamma x} \quad \text{for each} \quad x \in \left(-\infty, \frac{1}{\gamma}\right).$$
Consider the inequality system (1.1) with $K := \mathbb{R}_-$ and $F$ defined above, namely

\begin{equation}
F(x) \leq 0.
\end{equation}

Let $\bar{r} := \frac{1}{2\gamma}$ and $x_0 = 0$. Then $F'(x_0) = -1$, and it is routine to check that $F$ possesses properties (1)-(3) of Theorem 2.1. Moreover, we have that

\begin{equation}
\xi = \|F'(x_0)^\dagger(K - F(x_0))\| = \frac{\tau - 1}{\gamma(2\tau - 1)} \leq \frac{3 - 2\sqrt{2}}{\gamma} < \frac{1}{\gamma},
\end{equation}

so (2.4) is satisfied and $r^* < \bar{r}$, where $r^*$ is defined by (2.3). Note that the solution set $S = [r^*, r^{**}]$, where $r^{**} := \frac{1 + \gamma_0 + \sqrt{(1 + \gamma_0)^2 - 8\gamma_0}}{4\gamma} \leq \bar{r}$. Therefore, $d(x_0, S) = \|0 - r^*\| = r^*$. Noting that $\xi \neq 0$ by (2.9), one can check by element calculation that

\begin{equation}
\frac{1 + \gamma_0 - \sqrt{(1 + \gamma_0)^2 - 8\gamma_0}}{4\gamma} = \tau_\xi \iff \xi = \frac{\tau - 1}{\gamma(2\tau - 1)}.
\end{equation}

It follows from (2.9) that $r^* = \frac{1 + \gamma_0 - \sqrt{(1 + \gamma_0)^2 - 8\gamma_0}}{4\gamma} = \tau_\xi$. Thus we have

\[ d(x_0, S) = r^* = \tau \|F'(x_0)^{-1}\| d(F(x_0), K) \]

because $d(F(x_0), K) = \xi$ and $\|F'(x_0)^{-1}\| = 1$. This shows that the bound $\tau \|F'(x_0)^{-1}\|$ is optimal.

The following theorem, Theorem 2.2, improves the corresponding result of Dedieu mentioned in section 1 in three aspects: we drop his assumption (1.10), weaken his nonsingularity assumption, and replace the constant $\frac{1}{s}$ in the right-hand side of (1.12) by a bigger constant. Indeed, if we take $\tau = \sigma = \sum_{k=0}^{\infty} \left(\frac{1}{4}\right)^{2^k - 1} = 1.63284\ldots$ in Theorems 2.2 and 2.3, then

\begin{equation}
\frac{\tau - 1}{\tau(2\tau - 1)} = \frac{\sigma - 1}{\sigma(2\sigma - 1)} = 0.1710\ldots \geq \frac{1}{8}.
\end{equation}

Recall that $\Gamma$ and $\delta$ are defined as in (1.6) and (1.9).

**Theorem 2.2.** Consider the inequality/equality system (1.4), where each $f_i$ and each $g_j$ are real-valued functions on $X$, and $(f, g)$ are defined as in (1.5). Let $x_0 \in X$ and suppose that $(f, g)$ is analytic on $B(x_0, 2(1+\Gamma(f,g,x_0)))$ such that

\begin{equation}
(\mathbb{R}^l \times \{0_m\}) \cup \{(f(x_0), g(x_0))\} \cup \text{im}(f^{(k)}(x_0), g^{(k)}(x_0)) \subseteq \text{im}G(x_0) \quad \text{for each } k = 2,\ldots
\end{equation}

e.g., $G(x_0)$ is nonsingular, where $G(x_0)$ is defined as in (1.8). Let $\tau \in (1, \frac{2\sqrt{2}}{2})$ be such that

\begin{equation}
\|f(x_0)^{-1}\|^2 + \|g(x_0)^2\|^2 < \frac{\tau - 1}{\tau(2\tau - 1)(1 + \Gamma(f,g,x_0))\delta(f,g,x_0)\max\{1, \delta(f,g,x_0)\}}.
\end{equation}

Then the solution set $S(f,g)$ of (1.4) is nonempty, and the following estimate holds:

\begin{equation}
d(x_0, S(f,g)) \leq \tau \delta(f,g,x_0) (\|f(x_0)^{-1}\|^2 + \|g(x_0)^2\|^2)^{\frac{1}{2}}.
\end{equation}

Similarly we also extend the result of [3, Theorem 1] for system (1.3) as follows.
Theorem 2.3. Consider system (1.3), where each $f_i$ is a real-valued function on $X$ and $f$ is defined as in (1.5). Let $x_0 \in X$, and suppose that $f$ is analytic on $B(x_0, \frac{2 - \sqrt{2}}{2(1 + \Gamma(f(x_0)))})$ such that

\begin{equation}
(2.13)
\text{the matrix } (f'(x_0)f'(x_0)^* + \text{Diag}(f(x_0)^*)) \text{ is nonsingular.}
\end{equation}

Let $\tau \in (1, \frac{2 + \sqrt{2}}{2}]$ be such that

\begin{equation}
(2.14)
\|f(x_0)^-\| < \frac{\tau - 1}{\sqrt{2}(2\tau - 1)(1 + \Gamma(f;x_0))}\delta(f;x_0)\max\{1, \delta(f;x_0)\}.
\end{equation}

Then the solution set $f^>$ of (1.3) is nonempty, and the following estimate holds:

\begin{equation}
(2.15)\quad d(x_0, f^>) \leq \tau \delta(f;x_0)\|f(x_0)^-\|.
\end{equation}

We defer the proofs of the above three theorems to sections 4 and 5. As a direct application of Theorem 2.1, we present below a proof for the following stability result for the perturbed system (1.16).

Theorem 2.4. Let $\gamma \in [0, +\infty)$, and let $x_0 \in X$ be a solution of system (1.1), where the function $F : X \to Y$ is twice Fréchet differentiable on $B(x_0, \frac{2 - \sqrt{2}}{2\gamma})$ with properties (1)–(3) in Theorem 2.1 but replacing $\tilde{r}$ by $\frac{2 - \sqrt{2}}{2\gamma}$. If $K$ is closed or the relative interior $\text{ri} \ K$ of $K$ is nonempty, then, for any $y \in B(0, \frac{3 - 2\sqrt{2}}{\gamma\|F'(x_0)\|^*}) \cap \text{im}F'(x_0)$, the perturbed system (1.16) is solvable, and its solution set $S(y)$ satisfies that

\begin{equation}
(2.16)
d(x_0, S(y)) \leq \frac{2 + \sqrt{2}}{2}\|F'(x_0)\|\|y\| \quad \text{for any } y \in B(0, \frac{3 - 2\sqrt{2}}{\gamma\|F'(x_0)\|^*}) \cap \text{im}F'(x_0).
\end{equation}

(In particular, the solution map $S(\cdot)$, restricted to $\text{im}F'(x_0)$, is calm at $(0, x_0)$.)

Proof. Consider $\tau := \frac{2 + \sqrt{2}}{2}$, and note that $\frac{\tau - 1}{\sqrt{2}(2\tau - 1)} = 3 - 2\sqrt{2}$. Let $y \in B(0, \frac{3 - 2\sqrt{2}}{\gamma\|F'(x_0)\|^*}) \cap \text{im}F'(x_0)$, define the function

\begin{equation}
(2.17)
F_y(\cdot) := F(\cdot) - y, \quad \text{and let } \xi_y := \|F_y'(x_0)^*\|(K - F_y(x_0))\|.
\end{equation}

Then $F_y' = F'$, $F''_y \equiv F''$. Since $x_0$ is a solution of system (1.1), one has $F(x_0) \in K$, and so $y = F(x_0) - F_y(x_0) \in K - F_y(x_0)$; it follows that

\begin{equation}
(2.18)
\xi_y \leq \|F_y'(x_0)^*\|\|y\| = \|F'(x_0)^*\|\|y\| \leq \frac{3 - 2\sqrt{2}}{\gamma}\left(\frac{\tau - 1}{\sqrt{2}(2\tau - 1)}\right).
\end{equation}

Therefore the corresponding $r_y^*$ (defined as in (2.3) but replacing $\xi$ by $\xi_y$) satisfies that

$$
r_y^* \leq \frac{1 + \gamma \xi_y}{4\gamma} < \frac{1 + 3 - 2\sqrt{2}}{4\gamma} = \frac{2 - \sqrt{2}}{2\gamma}.
$$

Pick $\bar{r} \in \left(r_y^*, \frac{2 - \sqrt{2}}{2\gamma}\right)$. This together with the given assumptions implies that $F_y$ satisfies the $\gamma$-condition on $B(x_0, \bar{r})$, and that

$$
K \cup \{F_y(x_0)\} \cup \text{im}F_y'(x) \subseteq \text{im}F_y'(x_0) \quad \text{for each } x \in B(x_0, \bar{r}).
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
(noting that \( \text{im} F'_y(x_0) = \text{im} F'_y(x_0) \) contains \( y \) and \( K \), and that \( F_y(x_0) \in K - y \). Therefore, if \( K \) is closed or the ri \( K \) is nonempty, one can apply Theorem 2.1, and we conclude that the perturbed problem

\[
F(x) - y = F_y(x) \geq_K 0
\]

is solvable, and its solution set \( S(y) \) satisfies that

\[
d(x_0, S(y)) \leq \tau \|F'_y(x_0)\| \|d(F_y(x_0), K) \leq \tau \|F'_y(x_0)\| \|F_y(x_0) - F(x_0)\|
\]

\[
= \frac{2 + \sqrt{2}}{2} \|F'_y(x_0)\| \|y\|.
\]

This proves (2.16), and the proof is complete. \( \square \)

3. Notations, notions, and auxiliary lemmas. Let \( \mathbb{N} \) denote the set of all natural numbers, and let \( k \in \mathbb{N} \). We consider a \( k \)-multilinear bounded operator \( \Lambda : X^k \to Y \). As in (1.7), we define the norm \( \|\Lambda\| \) by

\[
\|\Lambda\| := \sup\{\|\Lambda(x_1, \ldots, x_k)\| : (x_1, \ldots, x_k) \in X^k, \|x_i\| \leq 1 \text{ for each } i\};
\]

also, let \( \text{im} \Lambda \) denote the image of \( \Lambda \):

\[
\text{im} \Lambda := \{\Lambda(x_1, \ldots, x_k) : (x_1, \ldots, x_k) \in X^k\}.
\]

Recall that if \( A \) is a bounded linear operator from \( X \) to \( Y \) such that \( \text{im} A \) is closed, then it has a unique Moore–Penrose generalized inverse (to be denoted by \( A^\dagger \)), which is, by definition, the bounded linear operator from \( Y \) to \( X \) such that

\[
(3.1) \quad AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (A^\dagger A)^* = (A^\dagger A), \quad \text{and } (AA^\dagger)^* = (AA^\dagger),
\]

where \( A^\ast \) denotes the adjoint operator of \( A \). For convenience, we list some facts regarding the Moore–Penrose inverse:

\[
(3.2) \quad (A^\ast)^\dagger = (A^\dagger)^*, \quad A^\dagger = (A^\ast A)^\dagger A^\ast = (AA^\ast)^\dagger, \quad (AA^\ast)^\dagger = (A^\ast)^\dagger A^\dagger,
\]

and

\[
(3.3) \quad A^\dagger A = \Pi_{(\ker A)^\perp}, \quad A^\dagger A = \Pi_{\text{im} A}, \quad A^\dagger \Pi_{\text{im} A} = A^\dagger AA^\dagger = A^\dagger,
\]

where \( \Pi_D \) denotes the projection on subset \( D \); see, for example, [4] for more details.

Notice by (3.2) that \( (AA^\ast)^\dagger = (A^\ast)^\dagger A^\dagger = (A^\ast)^{\dagger} A^\dagger \). It follows that

\[
(3.4) \quad \|AA^\ast\| = \|A^\dagger\|^2.
\]

**Lemma 3.1.** Let \( A \) be a bounded linear operator from \( X \) to \( Y \) with closed image. Let \( D \) be a closed convex subset of \( \text{im} A \). Then there exists \( u_0 \in D \) such that \( \|A^\dagger u_0\| = \|A^\dagger D\| \).

**Proof.** Choose a sequence \( \{u_n\} \subseteq D \) such that \( \|A^\dagger u_n\| \to \|A^\dagger D\| \). Then \( \{AA^\dagger u_n\} \) is bounded. Since \( \{u_n\} \subseteq D \subseteq \text{im} A \) by assumption, and since \( AA^\dagger \) is the projection \( \Pi_{\text{im} A} \) on \( \text{im} A \) (see (3.3)), one has that \( u_n = AA^\dagger u_n \), and so \( \{u_n\} \) is bounded. Thus \( \{u_n\} \) has a subsequence, denoted by itself, which converges weakly to some point \( u_0 \in D \); consequently, by a standard result in functional analysis, there exists a sequence \( \{v_n\} \subseteq D \) with each of its elements representable as a convex combination...
of some elements of \{u_n\} such that \(u_n\) converges to \(u_0\) strongly. By the continuity and linearity of \(A^t\), it is then routine to check that \(\|A^t u_0\| = \|A^t D\|\), and so we complete the proof. □

Lemma 3.2. Let \(\gamma > 0\), and suppose that \(F\) satisfies the \(\gamma\)-condition at some \(\bar{x} \in X\) on \(\mathcal{B}(\bar{x}, r)\) with some \(r \in (0, \frac{1}{\gamma}]\). Suppose further that

\[
\text{im}F''(x) \subseteq \text{im}F'({\bar{x}}) \quad \text{for each } x \in \mathcal{B}(\bar{x}, r).
\]

Let \(x \in X\) be such that

\[
\|x - \bar{x}\| < \min \left\{ r, \frac{2 - \sqrt{2}}{2\gamma} \right\}.
\]

Then

\[
\text{im}F'(x) = \text{im}F'({\bar{x}})
\]

and

\[
\|F'(x)^\dagger F'({\bar{x}})\| \leq \left( 2 - \frac{1}{1 - \gamma \|x - \bar{x}\|^2} \right)^{-1} = \frac{1 - \gamma \|x - \bar{x}\|^2}{1 - 4\gamma \|x - \bar{x}\| + 2\gamma^2 \|x - \bar{x}\|^2};
\]

moreover, if \(F(\bar{x}) \in \text{im}F'({\bar{x}})\), then

\[
F(x) \in \text{im}F'(\bar{x}) = \text{im}F'(x).
\]

Proof. Since

\[
F'(x) = F'(\bar{x}) + \int_0^1 F''(\bar{x} + t(x - \bar{x}))(x - \bar{x})dt,
\]

assumption (3.5) implies that the image of \(F'(x)\) is contained in \(\text{im}F'(\bar{x})\). Since, when restricting to \(\text{im}F'(\bar{x})\), \(F'(\bar{x})F'(\bar{x})^\dagger\) is the identity map (see (3.3)), we then have that \(F'(\bar{x})F'(\bar{x})^\dagger F'(x) = F'(x)\). Letting \(W := F'(\bar{x})^\dagger (F'(\bar{x}) - F'(x))\) and \(I_X\) denote the identity map on \(X\), it follows from (3.1) that

\[
F'(x) = F'(\bar{x})F'(\bar{x})^\dagger (F'(\bar{x}) - F'(x)) + F'(\bar{x}) = F'(\bar{x})(-W + I_X).
\]

Assuming that \(I_X - W\) is invertible, (3.7) is seen to hold (by (3.11)), and \(F'(\bar{x}) = F'(x)(I_X - W)^{-1}\), and so

\[
\|F'(x)^\dagger F'(\bar{x})\| = \|F'(x)^\dagger F'(x)(I_X - W)^{-1}\| \leq \|(I_X - W)^{-1}\|
\]

because \(F'(x)^\dagger F'(x)\) is a projection operator (and thus of norm bounded by 1) by (3.3). Thus (3.8) will also be shown, provided that

\[
\|(I_X - W)^{-1}\| \leq \frac{(1 - \gamma \|x - \bar{x}\|^2)}{1 - 4\gamma \|x - \bar{x}\| + 2\gamma^2 \|x - \bar{x}\|^2}.
\]

To accomplish this, we use (3.10) and the assumed \(\gamma\)-condition to note that

\[
\|W\| \leq \int_0^1 \|F'(\bar{x})^\dagger F''(\bar{x} + t(x - \bar{x}))(x - \bar{x})\|dt
\]

\[
\leq \int_0^1 \frac{2\gamma}{1 - \gamma t \|x - \bar{x}\|^2} \|x - \bar{x}\|dt
\]

\[
= \frac{1}{(1 - \gamma \|x - \bar{x}\|^2)} - 1.
\]
Further, since \( \gamma \| x - \bar{x} \| \leq \frac{2\sqrt{2}}{x} \) by (3.6), we also see that \( \| W \| < 1 \) as
\[
\frac{1}{(1 - \gamma \| x - \bar{x} \|)^2} < \frac{1}{(1 - \frac{2\sqrt{2}}{x})^2} = 2.
\]
By the Banach lemma, \( I_X - W \) is therefore invertible, and (3.13) holds.

Finally, suppose that \( F(\bar{x}) \in \text{im} F'(\bar{x}) \). To establish (3.9), we need only show that \( F(x) \in \text{im} F'(\bar{x}) \). But this is clearly true from the identity
\[
F(x) = F(\bar{x}) + \int_0^1 F'(\bar{x} + t(x - \bar{x}))(x - \bar{x})dt
\]
because \( F'(\bar{x} + t(x - \bar{x}))(x - \bar{x}) \in \text{im} F'(\bar{x}) \) by (3.7) for any \( t \in [0, 1] \).

4. Proofs of main results. For convenience, we shall use \( w \) to denote the elementary function defined by
\[
w(t) := \frac{2}{1 + t + \sqrt{(1 + t)^2 - 8t}} \quad \text{for each } t \in [0, 3 - 2\sqrt{2}].
\]
Then
\[
w : [0, 3 - 2\sqrt{2}] \rightarrow \left[ 1, \frac{2 + \sqrt{2}}{2} \right]
\]
is strictly increasing and bijective with the inverse map \( w^{-1} : \left[ 1, \frac{2 + \sqrt{2}}{2} \right] \rightarrow [0, 3 - 2\sqrt{2}] \) given by
\[
w^{-1}(t) := \frac{t - 1}{t(2t - 1)} \quad \text{for each } t \in \left[ 1, \frac{2 + \sqrt{2}}{2} \right].
\]
Let \( \xi, \gamma \in [0, +\infty) \), and define the function \( \phi_{\xi, \gamma} \) by
\[
\phi_{\xi, \gamma}(t) := \xi - t + \frac{\gamma t^2}{1 - \gamma t} \quad \text{for each } t \in \left( -\infty, \frac{1}{\gamma} \right).
\]
Suppose that
\[
0 \leq \xi \leq \frac{3 - 2\sqrt{2}}{\gamma}.
\]
Then the smaller root \( r^* \) of \( \phi_{\xi, \gamma} \) on \([0, 1/\gamma)\) is given by (2.3), that is,
\[
r^* := \begin{cases} 
\frac{1 + \gamma \xi - \sqrt{(1 + \gamma \xi)^2 - 8\xi}}{4\gamma}, & 0 < \gamma \leq \frac{3 - 2\sqrt{2}}{\xi}, \\
\xi, & \gamma = 0;
\end{cases}
\]
we note that
\[
0 \leq r^* \leq \frac{1 + \gamma \xi}{4\gamma} \leq \frac{2 - \sqrt{2}}{2\gamma}.
\]
Notice further that the inequality in (2.4) and the definition for \( r^* \) given by (4.6) can be rewritten as
\[
\gamma \xi \leq w^{-1}(\tau)
\]
and
\begin{equation}
(4.9)
\quad r^* = w(\gamma \xi) \xi,
\end{equation}
respectively. If there is no ambiguity, we shall write \( \phi \) for \( \phi_{\xi, \gamma} \) (which is defined as in (4.4)). Let \( h : D \subseteq X \to Y \) be a nonlinear map with continuous Fréchet derivative. Let \( \{t_n\} \) and \( \{x_n\} \) be sequences generated by the (Gauss-)Newton method for \( \phi \) and for \( h \), respectively, with initial points \( t_0 = 0 \) and \( x_0 \); that is,
\begin{equation}
(4.10)
\quad t_{n+1} := t_n - \phi'(t_n)^{-1} \phi(t_n) \quad \text{for each } n = 0, 1, \ldots
\end{equation}
and
\begin{equation}
(4.11)
\quad x_{n+1} := x_n - h'(x_n)^{\dagger} h(x_n) \quad \text{for each } n = 0, 1, \ldots.
\end{equation}

The main part of the following result is taken from [5].

**Proposition 4.1.** Let \( \gamma \in [0, +\infty) \), \( \Omega \subseteq X \), and let \( x_0 \) be an interior point of \( \Omega \). Let \( h : \Omega \to Y \) be continuous such that it is twice Fréchet differentiable on the interior of \( \Omega \) with the surjective derivative \( h'(x_0) \). Let \( \xi \) be defined by
\begin{equation}
(4.12)
\quad \xi := \| h'(x_0) h(x_0) \|,
\end{equation}
let \( r^* \) be defined by (4.6), and let \( \phi := \phi_{\xi, \gamma} \). Suppose that
\begin{equation}
(4.13)
\quad \overline{B(x_0, r^*)} \subseteq \Omega,
\end{equation}
and \( h \) satisfies the \( \gamma \)-condition at \( x_0 \) on \( B(x_0, r^*) \). Then the Newton method (4.11) with initial point \( x_0 \) is well defined, and the generated sequence \( \{x_n\} \) converges to a zero \( x^* \) of \( h \) with \( x^* \in \overline{B(x_0, r^*)} \) satisfying
\begin{equation}
(4.14)
\quad \| x_n - x^* \| \leq r^* - t_n \quad \text{for each } n = 0, 1, \ldots.
\end{equation}

**Proof.** If \( \xi = 0 \), then, by definition, one has \( x_n = x_0 = x^* \) and \( t_n = r^* = 0 \) for each \( n \), and (4.14) holds trivially. Suppose that \( \xi > 0 \) but \( \gamma = 0 \). Then, by the given assumptions of surjectivity, \( \gamma \)-condition, and continuity, one has that \( h \) is affine on \( \overline{B(x_0, r^*)} \):
\begin{equation}
(4.15)
\quad h(x) = h(x_0) + h'(x_0)(x - x_0) \quad \text{for each } x \in \overline{B(x_0, r^*)}.
\end{equation}
Thus we have that
\begin{equation}
\quad x_n = x^* = x_0 + h'(x_0)^{\dagger} h(x_0) \quad \text{and} \quad t_n = r^* = \xi \quad \text{for each } n > 0.
\end{equation}
Therefore, (4.14) is seen to hold because
\begin{equation}
\| x_n - x^* \| = \left\{ \begin{array}{ll}
\xi = r^* - t_0, & n = 0, \\
0 = r^* - t_n, & n > 0.
\end{array} \right.
\end{equation}
Finally, for the remaining case (namely, \( \xi > 0, \gamma > 0 \)), the proof given in [5, Theorem 3.3] can be adopted here although the setting of [5] is in the finite dimensional spaces \( X \) and \( Y \). \( \square \)
Now we are ready to give the proof of Theorem 2.1.

Proof of Theorem 2.1. By the assumptions, $Y_0 := \text{im}F'(x_0)$ is a closed subspace such that
\begin{equation}
(4.15) \quad \overline{K} \cup \{F(x_0)\} \cup \text{im}F''(x) \subseteq \text{im}F'(x_0) = Y_0 \quad \text{for each } x \in B(x_0, \bar{r}).
\end{equation}

Set $\bar{r}_0 := \min\{\bar{r}, \frac{z}{2\gamma}\}$ (so $r^* \leq \bar{r}_0$ by (4.7) and thanks to the assumption $r^* \leq \bar{r}$). By the given $\gamma$-condition on $B(x_0, \bar{r})$, we can apply Lemma 3.2 (to $x_0$, $\bar{r}$ in place of $\bar{x}$, $r$) to conclude that
\begin{equation}
F(x) \in Y_0 = \text{im}F'(x_0) \quad \text{for each } x \in B(x_0, \bar{r}_0) \quad (\text{and so for each } x \in \overline{B(x_0, \bar{r}_0)}).
\end{equation}

Moreover, by Lemma 3.1, the distance from the origin to the set $F'(x_0)\overline{(K - F(x_0))}$ is attained: there exists some $c_0 \in \overline{K}$ (so $F(x) - c_0 \in Y_0$ for each $x \in \overline{B(x_0, \bar{r}_0)}$) such that
\begin{equation}
(4.16) \quad \|F'(x_0)\overline{(c_0 - F(x_0))}\| = \|F'(x_0)\overline{(K - F(x_0))}\| = \|F'(x_0)\overline{(K - F(x_0))}\| = \xi,
\end{equation}
where the last equality holds by the definition of $\xi$ given in (2.2). Let $\Omega := \overline{B(x_0, \bar{r}_0)}$ and define $h : \Omega \to Y_0$ by
\begin{equation}
(4.17) \quad h(x) := F(x) - c_0 \quad \text{for each } x \in \Omega.
\end{equation}

Then $h$ is continuous on $\Omega$, and
\begin{equation}
(4.18) \quad h'(x) = F'(x) \quad \text{and} \quad h''(x) = F''(x) \quad \text{for each } x \in \text{int}\Omega = B(x_0, \bar{r}_0).
\end{equation}

In particular, $h'(x_0) = F'(x_0)$, and so $h'(x_0) : X \to Y_0$ is surjective. Note further that $h'(x_0)\overline{}$ is exactly the restriction to $Y_0$ of the operator $F'(x_0)\overline{ }$. This together with (4.15) implies that
\begin{equation}
(4.19) \quad \|h'(x_0)\overline{h(x_0)}\| = \xi \leq \frac{w^{-1}(\tau)}{\gamma} \leq \frac{3 - 2\sqrt{2}}{\gamma}.
\end{equation}

Finally, since $r^* \leq \bar{r}_0$ as noted earlier, one has that $\overline{B(x_0, r^*)} \subseteq \Omega$. Thus, Proposition 4.1 is applicable to concluding that the sequence $x_n$ generated by the Newton method (4.11) for $h$ with initial point $x_0$ converges to a zero $x^*$ of $h$ and (4.14) holds. Hence, $F(x^*) = c_0 \in \overline{K}$, and (4.14) in particular implies that
\begin{equation}
(4.20) \quad \|x_0 - x^*\| \leq r^* = w(\gamma)\xi \leq \bar{r}\xi
\end{equation}
(see (4.8) and (4.9)). Clearly, one has
\begin{equation}
\|F'(x_0)\overline{(K - F(x_0))}\| \leq \|F'(x_0)\overline{d(F(x_0), K)}.
\end{equation}

By (4.16), this means that $\xi \leq \|F'(x_0)\overline{d(F(x_0), K)}$, and so (2.5) holds thanks to (4.20). In particular, if $K$ is closed, then $x^* \in S$, and the conclusion is proved for case (a).
It remains to consider case (b): we assume that
\[(4.21)\quad r^* < \bar{r}, \quad \xi < \frac{3 - 2\sqrt{2}}{\gamma}, \quad \text{and} \quad \text{ri} K \neq \emptyset.\]
Then there exists a sequence \(\{c_n\} \subseteq \text{ri} K\) such that \(c_n \to 0\). Without loss of generality, we may assume further that, for each \(n = 1, 2, \ldots\)
\[(4.22)\quad \bar{\xi}_n := \xi + \|F'(x_0)^\dagger\|\|c_n\| < \frac{3 - 2\sqrt{2}}{\gamma},\]
\[(4.23)\quad \bar{\tau}_n := w(\gamma \bar{\xi}_n) \in \left(1, \frac{2 + \sqrt{2}}{2}\right),\]
and
\[(4.24)\quad \bar{r}^* := w(\gamma \bar{\xi}_n) \bar{\xi}_n < \bar{r}\]
(noticing (4.2) and that \(\bar{\xi}_n \to \xi\) and \(\bar{r}^* \to r^*\) by (4.9)). Let \(F_n : X \to Y\) be the map defined by \(F_n(\cdot) := F(\cdot) - c_n\), and consider the inequality system (1.1) with \(F_n\) in place of \(F\), that is, the inequality system
\[(4.25)\quad F_n(x) \geq_K 0.\]
Then, \(F_n = F'\), and so \(F_n'(x_0)^\dagger = F'(x_0)^\dagger\) and \(F_n(x_0) \in \text{im} F_n'(x_0)\), thanks to assumption (4.15). Moreover, we have \(F_n'(x_0)^\dagger(K - F_n(x_0)) = F'(x_0)^\dagger(K - F(x_0) - c_n)\), and so
\[(4.26)\quad \xi_n := \|F_n'(x_0)^\dagger(K - F_n(x_0))\| \leq \xi + \|F'(x_0)^\dagger\|\|c_n\| = \bar{\xi}_n,\]
\[(4.27)\quad \tau_n := w(\gamma \xi_n) \leq w(\gamma \bar{\xi}_n) = \bar{\tau}_n \in \left(1, \frac{2 + \sqrt{2}}{2}\right),\]
and
\[r_n^* := w(\gamma \xi_n) \xi_n \leq w(\gamma \bar{\xi}_n) \bar{\xi}_n = \bar{r}_n ^* < \bar{r}\]
thanks to (4.2). This and the given assumptions imply that \(F_n\) satisfies the \(\gamma\)-condition on \(B(x_0, r_n^*)\), and both (2.4) and (2.1) hold with \(F_n, \tau_n\) in place of \(F, \tau\). (By (4.27) and the definition of \(w^{-1}\) in (4.3), \(\gamma \xi_n = w^{-1}(\tau_n) = \frac{\bar{\xi}_n}{\tau_n/\sqrt{2} - 1}\)) Moreover, by (4.9), the quantity \(r_n^*\) defined above is exactly the corresponding quantity \(r^*\) defined by (2.3) with \(\xi_n\) in place of \(\xi\). Then we apply the (already established) first conclusion of Theorem 2.1 and conclude that there exists \(x_n^* \in X\) satisfying \(F_n(x_n^*) \in K\) such that
\[\|x_0 - x_n^*\| \leq \tau_n\|F'(x_0)^\dagger\|d(F_n(x_0), K).\]
By the choice of \(c_n\), we have that \(F(x_n^*) \in K + c_n \subseteq \text{ri} K \subseteq K\). Hence \(x_n^* \in S\) (and so \(S\) is nonempty), and it follows that
\[d(x_0, S) \leq \|x_0 - x_n^*\| \leq \tau_n\|F'(x_0)^\dagger\|d(F_n(x_0), K).\]
Letting \(n \to \infty\), we establish (2.6), as it is clear that \(d(F_n(x_0), K) \to d(F(x_0), K)\) and \(\tau_n = w(\gamma \xi_n) \to w(\gamma \xi) \leq \tau\). Thus the proof is complete. \(\square\)
5. Proofs of Theorems 2.2 and 2.3. Before our discussion on the applications of Theorem 2.1 to the study of systems (1.3) and (1.4), let us prove a simple lemma regarding the \( \gamma \)-condition for analytic functions. As assumed in the previous sections, let \( X \) and \( Y \) be Hilbert spaces.

Lemma 5.1. Let \( F : X \to Y \) be analytic on \( B(x_0, r) \) for some \( x_0 \in X \) and \( r \in (0, +\infty) \). Let \( \gamma \in \gamma_F(x_0), +\infty \), where \( \gamma_F(x_0) \) is defined by (1.15), and let \( r_0 := \min \{r, \frac{1}{\gamma} \} \). Then the following assertions hold:

(i) \( F \) satisfies the \( \gamma \)-condition at \( x_0 \) on \( B(x_0, r_0) \).

(ii) If \( \text{im} F'(x_0) \) is closed and \( \text{im} F^{(k)}(x_0) \subseteq \text{im} F'(x_0) \) for each \( k = 2, \ldots \), then

\[
\text{im} F''(x) \subseteq \text{im} F'(x_0) \quad \text{for each} \; x \in B(x_0, r_0).
\]

Proof. The proof of assertion (i) is standard and is a direct application of the Taylor expansion; see, for example, [25, Theorem 1]. Below we verify assertion (ii). For this purpose, assume that \( \text{im} F^{(k)}(x_0) \subseteq \text{im} F'(x_0) \) for each \( k = 2, \ldots \), and let \( x \in B(x_0, r_0) \). Since \( F \) is analytic on \( B(x_0, r) \), it follows from the Taylor expansion that

\[
F''(x) = \sum_{k \geq 0} \frac{F^{(k+2)}(x_0)}{k!}(x - x_0)^k.
\]

Hence the conclusion follows, as \( \text{im} F'(x_0) \) is closed. \( \square \)

As applications of Theorem 2.1 to systems (1.3) and (1.4), respectively, we derive the following corollaries, where we assume that the involved functions \( f \) and \( g \) are defined by (1.5) and are analytic around the involved point \( x_0 \).

Corollary 5.2. Consider the inequality system (1.4), and let \( F : X \to \mathbb{R}^l \times \mathbb{R}^m \) be defined by \( F := (f, g) \). Let \( \tau \in (1, \frac{2+\sqrt{\gamma}}{2}], \gamma \in [\gamma_F(x_0), +\infty) \), and \( x_0 \in X \) be such that

\[
(\mathbb{R}^l \times \{0_m\}) \cup \{F(x_0)\} \cup \text{im} F^{(k)}(x_0) \subseteq \text{im} F'(x_0) \quad \text{for each} \; k = 2, \ldots.
\]

and

\[
(\|f(x_0)^{-}\|^2 + \|g(x_0)^{-}\|^2)^{\frac{1}{2}} < \frac{\tau - 1}{\gamma \gamma(2\tau - 1)} \|F'(x_0)^{-}\|.
\]

Suppose that \( f \) and \( g \) are analytic on \( B(x_0, \frac{2-\sqrt{\gamma}}{2\gamma}) \). Then the solution set \( S(f, g) \) of (1.4) is nonempty and

\[
d(x_0, S(f, g)) \leq \tau \|F'(x_0)^{-}\| (\|f(x_0)^{-}\|^2 + \|g(x_0)^{-}\|^2)^{\frac{1}{2}}.
\]

Proof. Let \( Y := \mathbb{R}^l \times \mathbb{R}^m \), and define

\[
K := \{(u, 0_m) \in \mathbb{R}^l \times \mathbb{R}^m : u_i > 0 \; \forall 1 \leq i \leq l_S\}.
\]

Then, with \( F := (f; g) \), systems (1.1) and (1.4) are identical, \( d(F(x_0), K) = (\|f(x_0)^{-}\|^2 + \|g(x_0)^{-}\|^2)^{\frac{1}{2}} \), and \( F'(x_0) = (f'(x_0); g'(x_0)) \) satisfies by the assumptions that

\[
K \cup \{F(x_0)\} \cup \text{im} F^{(k)}(x_0) \subseteq \text{im} F'(x_0) \quad \text{for each} \; k = 2, \ldots.
\]

Set \( \xi := \|F'(x_0)^{-}\| (K - F(x_0)) \|. \) Since by (5.3)

\[
\|F'(x_0)^{-}\|d(F(x_0), K) = \|F'(x_0)^{-}\| (\|f(x_0)^{-}\|^2 + \|g(x_0)^{-}\|^2)^{\frac{1}{2}} < \frac{\tau - 1}{\gamma \gamma(2\tau - 1)},
\]
it follows that

\[(5.7) \quad \xi \leq \|F'(x_0)\| d(F(x_0), K) \leq \frac{\tau - 1}{\gamma (2\tau - 1)} \leq \frac{3 - 2\sqrt{2}}{\gamma}\]

(see (4.2) and (4.3)). Set \(r_0 := \frac{2 - \sqrt{2}}{2\gamma} (\leq \frac{1}{2})\). Thus, by Lemma 5.1, \(F\) satisfies the \(\gamma\)-condition at \(x_0\) on \(B(x_0, r_0)\), and (5.1) holds thanks to (5.5). Therefore, (2.1) is satisfied. Moreover, \(r^* < r_0\), where \(r^*\) is defined as in (2.3) (see also (4.6)). To see this, we may suppose that \(\gamma \neq 0\). Then by (2.3) and (5.7), one has

\[r^* = \frac{1 + \gamma \xi}{4\gamma} < \frac{1 + 3 - 2\sqrt{2}}{4\gamma} = r_0,\]

as claimed. Let \(\bar{r} \in (r^*, r_0)\). Then \(\overline{B(x_0, \bar{r})} \subseteq B(x_0, \frac{2 - \sqrt{2}}{2\gamma})\). Therefore, \(F\) is continuous on \(\overline{B(x_0, \bar{r})}\), as \(f\) and \(g\) are analytic on \(B(x_0, \frac{2 - \sqrt{2}}{2\gamma})\) by assumption. Noting \(\text{ri} \ K \neq \emptyset\), the assumption stated in (b) of Theorem 2.1 is satisfied. Thus Theorem 2.1 is applicable here, and we conclude that the solution set \(S(f, g)\) of (1.1) (namely that of (1.4)) is nonempty, and

\[d(x_0, S(f, g)) \leq \tau \|F'(x_0)\| d(F(x_0), K);\]

i.e., \(d(x_0, S(f, g)) \leq \tau \|f'(x_0)\| (\|f(x_0)^-\|^2 + \|g(x_0)^-\|^2)^{\frac{1}{2}}\). This completes the proof of the corollary.

**Corollary 5.3. Consider the inequality system (1.3). Let \(\tau \in (1, \frac{2 + \sqrt{2}}{2\gamma}], \gamma \in [\gamma_f(x_0), +\infty), x_0 \in X, and suppose that \(f\) is analytic on \(B(x_0, \frac{2 - \sqrt{2}}{2\gamma})\) such that \(f'(x_0) : X \to \mathbb{R}^l\) is surjective and**

\[(5.8) \quad \|f(x_0)^-\| < \frac{\tau - 1}{\gamma (2\tau - 1)} \|f'(x_0)^-\|.\]

**Then the solution set \(f^>\) of (1.3) is nonempty, and the following estimate holds:**

\[(5.9) \quad d(x_0, f^>) \leq \tau \|F'(x_0)^-\| \|f(x_0)^-\|;\]

**Proof.** Consider the inequality system (1.4) with \(f\) given as in the corollary and \(g = 0\). Then \(f\) and \(g\) are analytic on \(B(x_0, \frac{2 - \sqrt{2}}{2\gamma})\), and the solution set \(S(f; g)\) of system (1.4) coincides with the solution set \(f^>\) of (1.3). Let \(F := (f; g)\). Then

\[(5.10) \quad F'(x_0)^- = (f'(x_0)^-; 0)^T \quad \text{and} \quad F^{(k)}(x_0) = (f^{(k)}(x_0); 0_m) \quad \text{for each} \quad k.\]

This implies that

\[\|F'(x_0)^-\| = \|f'(x_0)^-\|, \quad \gamma_F(x_0) = \gamma_f(x_0),\]

and

\[\text{im}F^{(k)}(x_0) = \text{im}f^{(k)}(x_0) \times \{0_m\} \subseteq \mathbb{R}^l \times \{0_m\} = \text{im}F'(x_0) \quad \text{for any} \quad k \geq 2,\]
thanks to the surjectivity assumption of \( f'(x_0) \). Thus assumptions (5.2) and (5.3) in Corollary 5.2 are satisfied. Hence Corollary 5.2 is applicable to concluding that \( f^\to = S(f; g) \) is nonempty and

\[
d(x_0, f^\to) = d(x_0, S(f, g)) \\
\leq \tau \| F'(x_0)^\dag \| (\| f(x_0) \| + \| g(x_0) \|^2)^{\frac{1}{2}} \\
= \tau \| f(x_0) \| \| f(x_0) \|. 
\]

The proof is complete. \( \square \)

**Remark 5.1.** Consider the inequality system (1.3) for the special case when \( x_0 \) satisfies that

\[
(5.11) \quad f_i(x_0) < 0 \quad \text{for each } 1 \leq i \leq l.
\]

Then, for this \( x_0 \), Corollary 5.3 is a stronger result than Theorem 2.3. To see this, we suppose that (5.11) holds and choose \( \gamma = \gamma_f(x_0) \). Note then that

\[
(5.12) \quad \text{Diag}(f(x_0)^\dag) = 0,
\]

and hence assumption (2.13) simply means that \( f'(x_0) \) is surjective. Moreover, by (5.12), (1.9), and (3.4),

\[
(5.13) \quad \delta(f; x_0) = \| (f'(x_0)f'(x_0)^\star)^\dag \| = \| f'(x_0)^\dag \|.
\]

Finally, recalling (1.6) and an elementary fact,

\[
(5.14) \quad \sup_{k \geq 2} t^{\frac{k-1}{k}} = \begin{cases} 
  t & \text{if } t > 1, \\
  1 & \text{if } t \leq 1,
\end{cases}
\]

we have

\[
\gamma = \gamma_f(x_0) = \sup_{k \geq 2} \| f'(x_0)^\dag f^{(k)}(x_0) \|^{\frac{1}{k-1}} \\
\leq \sup_{k \geq 2} \| f'(x_0)^\dag \|^\frac{1}{k-1} \Gamma(f, x_0) \\
= \max \{1, \delta(f; x_0)\} \Gamma(f, x_0),
\]

and so the quantity given on the right-hand side of (5.8) is smaller than that of (2.14); thus (2.14) \( \Rightarrow \) (5.8).

Similarly, Corollary 5.2 is a stronger result than Theorem 2.2 for system (1.4) in the special case when \( x_0 \) satisfies (5.11).

For the proofs of Theorems 2.2 and 2.3, it would be convenient to do some preparations first. Let \( (x; v) \in X \times \mathbb{R}^l \). We use \( (F'(x); \text{Diag}(v)) \) to denote the operator from \( X \times \mathbb{R}^l \) to \( \mathbb{R}^l \times \mathbb{R}^m \) defined by

\[
(F'(x); \text{Diag}(v))(z; u) := F'(x)z + (\text{Diag}(v)u; 0_m) \quad \text{for each } (z; u) \in X \times \mathbb{R}^l.
\]

Using the matrix notation, the operator \( (F'(x); \text{Diag}(v)) \) can be represented as

\[
(5.15) \quad (F'(x); \text{Diag}(v))(z; u) = \begin{pmatrix} f'(x) & \text{Diag}(v) \end{pmatrix} \begin{pmatrix} z \\ 0_{m \times l} \end{pmatrix} \quad \text{for each } (z; u) \in X \times \mathbb{R}^l.
\]
Let \( G(x;v) \) be the \((l+m) \times (l+m)\) matrix defined by

\[
G(x;v) := \begin{pmatrix}
f'(x)f'(x)^* + \text{Diag}(v^2) & f'(x)g'(x)^* \\
g'(x)f'(x)^* & g'(x)g'(x)^*
\end{pmatrix},
\]

where, for any \( v = (v_i) \in \mathbb{R}^l \), \( v^2 \) stands for the vector \((v_i^2)\) in \( \mathbb{R}^l \). The following lemma will be useful for us.

**Lemma 5.4.** Let \( (x;v) \in X \times \mathbb{R}^l \). Then we have the following assertions:

\[
(F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^* = G(x;v)
\]

and

\[
im (F'(x); \text{Diag}(v)) = \text{im} G(x;v).
\]

**Proof.** We claim that

\[
(F'(x); \text{Diag}(v))^*(w_1, w_2) = \begin{pmatrix}
f'(x)^* \\
\text{Diag}(v)
\end{pmatrix} \begin{pmatrix}
w_1 \\
w_2
\end{pmatrix}
\]

for each \((w_1; w_2) \in \mathbb{R}^l \times \mathbb{R}^m\).

In fact, by definition, for any \((z, u) \in X \times \mathbb{R}^l\), one checks that

\[
((z, u), (F'(x); \text{Diag}(v))^*(w_1, w_2)) = (F'(x); \text{Diag}(v))(z, u), (w_1, w_2)
\]

\[
= \langle f'(x)z, w_1 \rangle + \langle \text{Diag}(v)u, w_1 \rangle + \langle g'(x)z, w_2 \rangle
\]

\[
= \langle z, f'(x)^*w_1 \rangle + \langle u, \text{Diag}(v)w_1 \rangle + \langle z, g'(x)^*w_2 \rangle
\]

\[
= \langle (z, u), \begin{pmatrix}
f'(x)^* \\
\text{Diag}(v)
\end{pmatrix} \begin{pmatrix}
w_1 \\
w_2
\end{pmatrix} \rangle,
\]

and so (5.19) is established. This together with (5.15) implies that

\[
(F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^* = \begin{pmatrix}
f'(x) \\
\text{Diag}(v)
\end{pmatrix} \begin{pmatrix}
f'(x)^* \\
\text{Diag}(v)^*
\end{pmatrix},
\]

and (5.17) is seen to hold.

To show (5.18), we set \( Y_0 := \text{im}(F'(x); \text{Diag}(v)) \). By (5.17), it suffices to check that \( \text{im}((F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^*) \supseteq Y_0 \), as the converse inclusion holds trivially. Since \( Y_0 \) is a finite dimensional subspace of \( \mathbb{R}^l \times \mathbb{R}^m \), we can choose finite dimensional subspace \( X_0 \) such that \( Y_0 = (F'(x); \text{Diag}(v))(X_0) \) and the restriction \( A := (F'(x); \text{Diag}(v))|_{X_0} \) on \( X_0 \) is a bijection. Hence \( A^* \) coincides with the restriction of \( (F'(x); \text{Diag}(v))^* \) to \( Y_0 \), and \( A^* \) is also a bijection. Consequently, the composite \( A \circ A^* \) is a bijection on \( Y_0 \). This implies that

\[
Y_0 = (A \circ A^*)(Y_0) \subseteq ((F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^*)(Y)
\]

\[
= \text{im}((F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^*),
\]

and the proof is complete. 

**Proof of Theorem 2.2.** Let \( H : X \times \mathbb{R}^l \to \mathbb{R}^l \times \mathbb{R}^m \) be the operator defined by \( H := (h; g) \) with \( h = (h_i) \) and \( g = (g_j) \), where each \( h_i \) and each \( g_j \) is given by

\[
h_i(x;v) := f_i(x) - v_i^2, \quad 1 \leq i \leq l,
\]

\[
g_j(x;v) := g_j(x), \quad 1 \leq j \leq m,
\]
for each \((x; v) \in X \times \mathbb{R}^l\) with \(v = (v_1, \ldots, v_l) \in \mathbb{R}^l\). Consider the following inequality system on \(X \times \mathbb{R}^l\):

\[
\begin{align*}
  h_i(x; v) &> 0, \quad i = 1, \ldots, l_S, \\
  h_i(x; v) &\geq 0, \quad i = l_S + 1, \ldots, l, \\
  g_j(x; v) &= 0, \quad j = 1, \ldots, m. 
\end{align*}
\]

(5.21)

As before, we use \(S(h; g)\) to denote the solution set of the above system. Then it is easy to check that \(x \in S(f; g)\) if \((x; v) \in S(h; g)\) for some \(v \in \mathbb{R}^l\), that is,

\[
S_0 := \{x \in X : \exists v \in \mathbb{R}^l \text{ s.t. } (x; v) \in S(h; g)\} \subseteq S(f; g).
\]

(5.22)

Let \(v_0 := (\sqrt{f(x_0)})^+ \in \mathbb{R}^l\), and notice that

\[
||h(x_0; v_0)||^2 + ||g(x_0; v_0)||^2 = ||f(x_0)||^2 + ||g(x_0)||^2,
\]

(5.23)

and that

\[
H(x_0; v_0) = F(x_0) - (f(x_0))^+; 0) \in F(x_0) + \mathbb{R}^l \times \{0_m\},
\]

(5.24)

and that

\[
G(x_0; \pm 2v_0) = G(x_0)
\]

(5.25)

by (5.16) and (1.8). Recalling (1.6) and (1.9), we define

\[
\gamma := \max\{1, \delta(f, g; x)\}(1 + \Gamma(f, g; x)).
\]

(5.26)

We will apply Corollary 5.2 to the inequality system (5.21) in place of (1.4) (so to \(H\) and \((x_0, v_0)\) in place of \(F\) and \(x_0\)). To do this, note first that, since \(\gamma \geq 1 + \Gamma(f, g; x_0)\), the analyticity assumption implies that \(H\) is analytic on \(B(x_0, 2\sqrt{2})\). Next note that

\[
H'(x_0; v_0) = (F'(x_0); \text{Diag}(-2v_0)), \quad H^{(k)}(x_0; v_0) = F^{(k)}(x_0) \quad \text{for each } k > 2,
\]

(5.27)

and

\[
H''(x_0; v_0) = (F''(x_0); -2\mathbb{D}_l),
\]

(5.28)

where \((F''(x_0); -2\mathbb{D}_l)\) is the bilinear map from \((X \times \mathbb{R}^l) \times (X \times \mathbb{R}^l)\) to \(\mathbb{R}^l \times \mathbb{R}^m\) defined by

\[
(F''(x_0); -2\mathbb{D}_l)(z_1; u_1)(z_2; u_2) := F''(x)z_1z_2 - 2\mathbb{D}_l(u_1; u_2)
\]

for any \((z_1; u_1), (z_2; u_2) \in X \times \mathbb{R}^l\), and \(\mathbb{D}_l : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l \times \mathbb{R}^m\) is the bilinear operator defined by

\[
\mathbb{D}_l(u; v) := \begin{pmatrix}
  \text{Diag}(u)v \\
  0_m
\end{pmatrix}
\]

for any \((u; v) \in \mathbb{R}^l \times \mathbb{R}^l\).

(5.29)

This implies in particular that

\[
\text{im}H''(x_0; v_0) = \text{im}F''(x_0) + \text{im}(-2\mathbb{D}_l) \subseteq \text{im}F''(x_0) + \mathbb{R}^l \times \{0_m\},
\]

(5.30)

and note also that

\[
\text{im}H^k(x_0; v_0) = \text{im}F^k(x_0) \quad \text{for each } k > 2.
\]

(5.31)
Moreover, by (5.25) and (5.27), we have from Lemma 5.4 that
\begin{equation}
G(x_0) = G(x_0; -2v_0) = H'(x_0; v_0) \circ H'(x_0; v_0)^\dagger
\end{equation}
and
\begin{equation}
\text{im} G(x_0) = \text{im} H'(x_0; v_0).
\end{equation}

By the given assumption (2.10), it follows that the vector space \( \text{im} H'(x_0; v_0) \) contains \( \mathbb{R}^d \times \{0_m\} \cup \{F(x_0)\} \cup \cup_{k\geq 2} \text{im} F^{(k)} \) (and so contains \( \mathbb{R}^d \times \{0_m\} \cup \{H(x_0; v_0)\} \cup \cup_{k\geq 2} \text{im} H^{(k)} \) by (5.24), (5.31), and (5.30)). Another consequence of (5.32) is \( \|G(x_0)^\dagger\| = \|H'(x_0; v_0)^\dagger\|^2 \) (see (3.4)). Thus, by (1.9),
\begin{equation}
\delta(f, g; x_0) = \|G(x_0)^\dagger\|^2 = \|H'(x_0; v_0)^\dagger\|.
\end{equation}

Combining this with (5.23) and (5.26), we see that assumption (2.11) means that
\begin{equation}
\|h(x_0; v_0)^\dagger\|^2 + \|g(x_0; v_0)^\dagger\|^2 < \frac{\tau - 1}{\gamma (2\tau - 1)} \|H'(x_0; v_0)^\dagger\|^2.
\end{equation}

Further, since
\[ \|H''(x_0; v_0)\| \leq 2 + \|F''(x_0)\| \quad \text{and} \quad \|H^{(k)}(x_0; v_0)\| = \|F^{(k)}(x_0)\| \quad \text{for each} \quad k > 2, \]
and making use of (5.14), one has the following estimate for \( \gamma_H(x_0, v_0) \) (see (1.15)):
\begin{equation}
\gamma_H(x_0; v_0) \leq \left( \sup_{k \geq 2} \|H'(x_0; v_0)^\dagger\|^2 \right) \left( \sup_{k \geq 2} \|H^{(k)}(x_0; v_0)\|^2 \right) \leq \max(1, \delta(f, g; x_0))(1 + \Gamma(f, g; x_0));
\end{equation}
thus one has \( \gamma \in (\gamma_H(x_0; v_0), +\infty) \) by (5.26). Therefore, one can apply Corollary 5.2 to the inequality system (5.21) (with \((x_0, v_0)\) in place of \(x_0\)) to get that \( S(h; g) \neq \emptyset \) and
\begin{equation}
d((x_0; v_0), S(h; g)) \leq \tau \|H'(x_0; v_0)^\dagger\| (\|h(x_0; v_0)^\dagger\|^2 + \|g(x_0; v_0)^\dagger\|^2)^{\frac{1}{2}}.
\end{equation}

Using (5.23) and (5.34), we obtain that
\begin{equation}
d((x_0; v_0), S(h; g)) \leq \tau \delta(f, g; x_0) (\|f(x_0)^\dagger\|^2 + \|g(x_0)^\dagger\|^2)^{\frac{1}{2}}.
\end{equation}

Let \( \epsilon > 0 \), and let \((x^*; v^*) \in S(h; g)\) be such that
\begin{equation}
\|x_0 - (x^*; v^*)\| \leq \tau \delta(f, g; x_0) (\|f(x_0)^\dagger\|^2 + \|g(x_0)^\dagger\|^2)^{\frac{1}{2}} + \epsilon.
\end{equation}

Then \( x^* \in S(f, g) \) by (5.22) and
\[ d(x_0, S(f, g)) \leq \|x_0 - x^*\| \leq \|(x_0; v_0) - (x^*; v^*)\|. \]
This together with (5.37) implies (2.12) because \( \epsilon > 0 \) is arbitrary. The proof is complete.

Proof of Theorem 2.3. This is similar to the proof for Corollary 5.3 by taking \( g = 0 \) in Theorem 2.2.
Acknowledgment. The authors are indebted to two anonymous referees for their helpful remarks and comments.

REFERENCES