APPROXIMATE SOLUTIONS FOR ABSTRACT INEQUALITY SYSTEMS

CHONG LI∗ AND K. F. NG†

Abstract.

Key words.

AMS subject classifications. 47J15 65H10 Secondary, 41A29

1. Introduction. In this paper, we study approximate solutions for conic inequality systems of the following type:

\[ F(x) \geq_K 0, \]  \hspace{1cm} (1.1)

where \( F \) is a twice Fréchet differentiable function from a Hilbert space \( X \) to another one \( Y \) with the partial order (or pre-order) \( \geq_K \) defined by a nonempty convex cone \( K \subseteq Y \). We assume that system (1.1) has an “approximate solution” \( x \) in the sense that \( F(x) + \epsilon \geq_K 0 \) for some \( \epsilon \in Y \). In more concrete situation we will consider as applications the following system with \( Y = \mathbb{R}^l \):

\[
\begin{align*}
&f_i(x) > 0 & i = 1, \ldots, l_S, \\
&f_i(x) \geq 0 & i = l_S + 1, \ldots, l,
\end{align*}
\]  \hspace{1cm} (1.2)

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

\[
\begin{align*}
&f_i(x) > 0 & i = 1, \ldots, l_S, \\
&f_i(x) \geq 0 & i = l_S + 1, \ldots, l, \\
g_j(x) = 0 & j = 1, \ldots, m.
\end{align*}
\]  \hspace{1cm} (1.3)

Here each \( f_i \) and each \( g_j \) are analytic functions on \( X \), and \( l, l_S, m \) are nonnegative integers such that \( 0 \leq l_S \leq l \). Associated with (1.2) and (1.3), let \( f : X \to \mathbb{R}^l, \ g : X \to \mathbb{R}^m \) and \((f, g) : X \to \mathbb{R}^l \times \mathbb{R}^m \) be defined respectively by

\[
\begin{align*}
f(x) := (f_1(x), \ldots, f_l(x)), & \quad g(x) := (g_1(x), \ldots, g_m(x)) \quad \text{and} \quad (f, g)(x) := (f(x), g(x))
\end{align*}
\]  \hspace{1cm} (1.4)

for each \( x \in X \). For each natural number \( k \), let \( F^{(k)}(x) \) denote the \( k \)-th derivative of a analytic function \( F \) at \( x \in X \) with the norm \( \|F^{(k)}(x)\| \) defined by

\[
\|F^{(k)}(x)\| := \sup\{|F^{(k)}(x)(u_1, \ldots, u_k)| : \text{ each } u_i \in X \text{ with } \|u_i\| \leq 1\}.
\]

∗Department of Mathematics, Zhejiang University, Hangzhou 310027, P. R. China, (cli@zju.edu.cn). This author 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 was supported by a General Research Fund (GRF) grant from the Research Grant Council of Hong Kong.
Further, let

\[ \Gamma(f; x) := \sup_{k \geq 2} \left\| \frac{f^{(k)}(x)}{k!} \right\|_{\infty}, \quad \Gamma(f, g; x) := \sup_{k \geq 2} \left\| \frac{(f, g)^{(k)}(x)}{k!} \right\|_{\infty}, \]

(1.5)

\[ \delta(f; x) := \|(f'(x)f'(x)^* + \text{Diag}(4f(x)^+))^{\dagger}\|^{\frac{1}{2}} \quad \text{and} \quad \delta(f, g; x) := \|G(x)\|^{\frac{1}{2}} \]

(1.6)

where, for \( d = (d_i) \in \mathbb{R}^l \), we use \( d^+, d^- \) respectively to denote the vector whose \( i \)-th coordinate is \( d_i^+ := \max\{0, d_i\} \) and \( d^- := \max\{0, -d_i\} \), \( \text{Diag}(d) \) is the diagonal matrix with diagonal entries \( d_1, \ldots, d_l \), \( A^\dagger \) and \( A^\ast \) denote respectively the Moore-Penrose generalized inverse and the conjugate of matrices (bounded linear operators) \( A \) (see Section 3 for details); while \( G(x) \) is the matrix defined by

\[ G(x) := \begin{pmatrix} f'(x)f'(x)^* + \text{Diag}(4f(x)^+) & f'(x)g'(x)^* \\ g'(x)f'(x)^* & g'(x)g'(x)^* \end{pmatrix} \]

(1.7)

and the norm is the usual operator norm for matrices. If the norms \( \|f(x)^{-}\| \) of \( f(x)^{-} \) is small then \( x \) “approximately satisfies” (1.2) in the sense that

\[ f_i(x) + \epsilon_i \geq 0 \quad \text{for each} \quad i = 1, \ldots, l \]

(1.8)

with some small real numbers \( \epsilon_i \). One of the main results of Dedieu in [7, Theorem?] for system (1.2) with \( X = \mathbb{R}^n \) can be stated as follows: if \( x_0 \in \mathbb{R}^n \) and \( \sigma := \sum_{k=0}^{\infty} \left( \frac{1}{2} \right)^{x^k-1} = 1.63284 \cdots \) are such that

\[ f_i(x_0) > (\sigma \delta(f; x_0) \|f(x_0)^{-}\|)^2 \quad \text{for each} \quad i = 1, \ldots, l_S, \]

(1.9)

the matrix \( (f'(x_0)f'(x_0)^* + \text{Diag}(4f(x_0)^+)) \) is nonsingular

(1.10)

and

\[ \|f(x_0)^{-}\| < \frac{1}{8(1 + \Gamma(f; x_0)) \delta(f; x_0) \max\{1, \delta(f; x_0)\}}. \]

(1.11)

then the solution set (denoted by \( f^\circ \)) of (1.2) is nonempty and its distance from \( x_0 \) satisfies that

\[ d(x_0, f^\circ) \leq \sigma \delta(f; x_0) \|f(x_0)^{-}\|. \]

(1.12)

Similar treatments have also been done for system (1.3) with \( X = \mathbb{R}^n \) by Dedieu, it was proved in [7, Theorem?] that if \( x_0 \in \mathbb{R}^n \) is such that (1.9) holds,

the matrix \( G(x_0) \) is nonsingular,

(1.13)

and

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

(1.14)

then the solution set (denoted by \( (f, g)^\circ \)) of (1.3) is nonempty and its distance from \( x_0 \) satisfies that

\[ d(x_0, (f, g)^\circ) \leq \sigma \delta(f, g; x_0) (\|f(x_0)^{-}\|^2 + \|g(x_0)\|^2)^{\frac{1}{2}}; \]

(1.15)
see [7], [] and references therein for more details and background informations. From some points of view, assumptions (1.9) and (1.13) are too restrictive, for example, when all \( f_i \) are zero functions with \( l_S < i \leq l \), that is when (1.2) is defined by only strict inequalities, (1.9) itself trivially entails that \( x_0 \) is a solution of (1.2).

More generally, we study the corresponding issues for the general system (1.1). Our results cover the case when \( K \) is not necessarily closed and/or the case when the nonsingularity is not assumed. As applications, we extend and improve the results of Dedieu mentioned above for system (1.2) and system (1.3). In particular, we show that for (1.15) to hold, both (1.13) and (1.14) can be replaced by weaker assumptions. Another byproduct of our main results can be cast in the stability point of view (Robinson initiated the study of stability in [?] for (1.1)). Under suitable assumptions in terms of the \( \gamma \)-conditions, we show that if \( \bar{x} \) solves (1.1), then perturbed systems

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

(1.16)

(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 sequel are listed in Section 3. Section 4 contains the proofs of the main theorems; while applications are provided in Section 5.

2. Main results. Let \( F'(x) \) and \( F''(x) \) denote the first and second Fréchet derivatives of \( F \) at \( x \in X \), and we consider the conic inequality system (1.1), where \( F : X \to Y \) with the given data: \( X \) and \( Y \) are Hilbert spaces and \( K \) is a convex cone in \( Y \). When uniquely exists, we use \( F'(x)^\dagger \) to denote the Moore-Penrose generalized inverse of \( F'(x) \), and note that the composition \( F'(\bar{x})^\dagger F''(x) \) (with \( \bar{x}, x \in X \)) of \( F'(\bar{x})^\dagger \) with \( F''(x) \) is then a bounded bi-linear operator from \( X \times X \) to \( Y \) (its norm will be denoted by \( \| F'(\bar{x})^\dagger F''(x) \| \)). We say that the map \( F \) satisfies the \( \gamma \)-condition at \( \bar{x} \) on an open ball \( B(\bar{x}, r) \) (with center \( \bar{x} \in X \) and radius \( r > 0 \)) if \( 0 \leq r_\gamma \leq 1 \) and

\[
\| F'(\bar{x})^\dagger F''(x) \| \leq \frac{2\gamma}{(1 - \gamma \|x - \bar{x}\|)^3}
\]

(2.1)

(see the next section for more details of definitions and notations involved). For the whole paper, we make the follow convention that \( \frac{a}{0} = +\infty \) for any \( a \geq 0 \). The following theorem is our main result regarding on the solution set \( S := \{ x \in X : F(x) \in K \} \) of the system (1.1).

**Theorem 2.1.** Consider the conic inequality system (1.1) and the solution set \( S \) (where \( F, X, Y, K \) are as already set at the beginning of this section). Let \( x_0 \in X \) be such that the image \( \text{im}F'(x_0) \) is closed. Let \( \gamma, \tau, \xi \in \mathbb{R} \) be such that

\[
\gamma \in [0, +\infty), \quad \tau \in \left[ 1, \frac{2 + \sqrt{2}}{2} \right]
\]

(2.2)

and

\[
\xi := \| F'(x_0)^\dagger (K - F(x_0)) \| \leq \frac{\tau - 1}{\gamma \tau (2\tau - 1)}.
\]

(2.3)

where the “norm” of \( F'(x_0)^\dagger (K - F(x_0)) \) is, by definition, the distance to the origin from the set \( \{ F'(x_0)^\dagger (y - F(x_0)) : y \in K \} \). Let \( r^* \) be defined by

\[
r^* := \frac{1 + \gamma \xi - \sqrt{(1 + \gamma \xi)^2 - 8\gamma \xi}}{4\gamma},
\]

(2.4)
and let $\bar{r} \geq r^*$ be such that $F$ satisfies the $\gamma$-condition on $B(x_0, \bar{r})$ and
\[ K \cup \{ F(x_0) \} \cup \text{im}F''(x) \subseteq \text{im}F'(x_0) \quad \text{for each } x \in B(x_0, \bar{r}). \quad (2.5) \]

Then there exists some element $x^* \in X$ satisfying $F(x^*) \in \overline{K}$ such that
\[ \| x_0 - x^* \| \leq \tau \| F'(x_0) \| d(F(x_0), K). \quad (2.6) \]

Furthermore, $S$ is nonempty and
\[ d(x_0, S) \leq \tau \| F'(x_0) \| d(F(x_0), K) \quad (2.7) \]

provided that
\begin{itemize}
  \item[(a)] $K$ is closed, or
  \item[(b)] $\bar{r} > r^*$, $\xi < \frac{3-2\sqrt{2}}{7}$ and the relative interior $\text{ri} K$ of $K$ is nonempty.
\end{itemize}

Remark 2.1. If $\xi = 0$ in (2.3), then assumption (2.5) entails that $F(x_0) \in \overline{K}$. In fact, assume that $\xi = 0$ in (2.3) and (2.5) holds. Then, $\| F'(x_0)^\dagger (K - F(x_0)) \| = 0$ and, by definition, we can choose $\{ u_n \} \subseteq K$ such that $F'(x_0)^\dagger (u_n - F(x_0)) \to 0$. Thus $F'(x_0)F'(x_0)^\dagger (u_n - F(x_0)) \to 0$. Since $F'(x_0)F'(x_0)^\dagger$ is simply the projection $\Pi_{\text{im}F'(x_0)}$ (see (3.4) of the next section) and since $u_n - F(x_0)$ lines in im$F'(x_0)$ by (2.5), 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 $\bar{x}$ on $B(\bar{x}, \bar{r})$, together with assumption (2.5) entails that $F''(x) = 0$ for each $x \in B(\bar{x}, \bar{r})$, that is $F$ is affine on $B(\bar{x}, \bar{r})$.

Remark 2.2. In the case when $F$ is analytic at $x_0$, we define
\[ \gamma_F(x_0) := \sup_{k \geq 2} \left\| \frac{F'(x_0)^\dagger F^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}. \quad (2.8) \]

Then, $\gamma_F < +\infty$. Furthermore, by [1], if $F$ is analytic on $B(x_0, r_0)$, then $F$ satisfies the $\gamma$-condition at $\bar{x}$ on $B(x_0, r)$ with $\gamma := \gamma_F$ and $r := \min \left\{ r_0, \frac{1}{2} \right\}$.

The following Theorem 2.2 improves the corresponding result of Dedieu mentioned in Section 1 in two aspects: we drop his assumption (1.9) and replace the constant $1/8$ in the right-hand side of (1.11) by a bigger constant.

Theorem 2.2. Consider the inequality system (1.2). Let $\tau \in (1, \frac{2+\sqrt{2}}{2}]$ and $x_0 \in X$ be such that assumption (1.10) is satisfied and
\[ \| f(x_0)^\gamma \| < \frac{\tau - 1}{\tau(2\tau - 1)(1 + \Gamma(f; x_0))d(f; x_0)\max\{1, \delta(f; x_0)\}}. \quad (2.9) \]

Suppose that $f$ is analytic on $B \left( x_0, \frac{2-\sqrt{2}}{2(1+\Gamma(f; x_0))} \right)$. Then the solution set $f^>$ of (1.2) is nonempty and the following estimate holds:
\[ d(x_0, f^>) \leq \tau \delta(f; x_0) \| f(x_0)^\gamma \|. \quad (2.10) \]
Similarly we also extend the result of [7, Theorem?] for system (1.3) as follows:

**Theorem 2.3.** Consider the inequality system (1.3), and let $F : X \to \mathbb{R}^l \times \mathbb{R}^m$ be defined by $F := (f, g)$. Let $\tau \in \{1, \frac{2+\sqrt{2}}{2}\}$ and $x_0 \in X$ be such that

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

(2.11)

(we adopt the convention that the matrix $G$ is a linear operator on $\mathbb{R}^l \times \mathbb{R}^m$) and

$$
\left(\|f(x_0)\|^2 + \|g(x_0)\|^2\right)^{\frac{1}{2}} < \frac{\tau - 1}{\tau(2\tau - 1)(1 + \Gamma(f, g; x_0))}\delta(f, g; x_0) \max\{1, \delta(f, g; x_0)\}.
$$

(2.12)

Suppose that $f$ and $g$ are analytic on $B \left(x_0, \frac{2-\sqrt{2}}{2(1+1/(f, g; x_0))}\right)$. Then the solution set $S(f, g)$ of (1.3) is nonempty and the following estimate holds:

$$
d(x_0, S(f, g)) \leq \delta(f, g; x_0) \left(\|f(x_0)\|^2 + \|g(x_0)\|^2\right)^{\frac{1}{2}}.
$$

(2.13)

**Remark 2.3.** (a) If we take $\tau = \sigma = \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^{2^k - 1} = 1.63284 \ldots$ in Theorems 2.2 and 2.3, then

$$
\frac{\tau - 1}{\tau(2\tau - 1)} = \frac{\sigma - 1}{\sigma(2\sigma - 1)} = 0.1710\ldots > \frac{1}{8}.
$$

Finally, as a direct application of the above theorem, we get the following stability result for the perturbed system (1.16).

**Theorem 2.4.** Let $F, X, Y, K$ and $S$ be as in Theorem 2.1 and let $\gamma \in (0, +\infty)$. Let $\bar{x} \in S$ and suppose that $F$ satisfies the $\gamma$-condition on $B \left(\bar{x}, \frac{2-\sqrt{2}}{2\gamma}\right)$ and

$$
K \cup \text{im} F''(x) \subseteq \text{im} F'((\bar{x}) \quad \text{for each } x \in B \left(\bar{x}, \frac{2-\sqrt{2}}{2\gamma}\right) \cdot
$$

(2.14)

If either $K$ is closed, or the relative interior $\text{ri} K$ of $K$ is nonempty, then, for any $y \in B \left(0, \frac{3-2\sqrt{2}}{\gamma\|F'(\bar{x})\|}\right) \cap \text{im} F'(\bar{x})$, the perturbed system (1.16) is solvable and the solution map $S(\cdot)$, restricted on $\text{im} F'(\bar{x})$, is calm at $(0, \bar{x})$, that is

$$
d(\bar{x}, S(y)) \leq \frac{2 + \sqrt{2}}{2} \|F'(\bar{x})\| \|y\| \quad \text{for any } y \in B \left(0, \frac{3-2\sqrt{2}}{\gamma\|F'(\bar{x})\|}\right) \cap \text{im} F'(\bar{x}).
$$

(2.15)

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

$$
F_y(\cdot) := F(\cdot) - y \quad \text{and} \quad \xi := \|F'_y((\bar{x})^\dagger(K - F_y(\bar{x}))\|).
$$

(2.16)

Note that $F'_y = F'$, $F''_y = F''$ and $F(\bar{x}) \in K$ (as $\bar{x} \in S$); hence $y = F(\bar{x}) - F_y(\bar{x}) \in K - F_y(\bar{x})$ and

$$
\xi \leq \|F'(\bar{x})\| < \frac{3 - 2\sqrt{2}}{\gamma}.
$$

(2.17)
Therefore $r^*$ (defined as in (2.4)) satisfies that
\[ r^* \leq \frac{1 + \gamma \xi}{4\gamma} < \frac{1 + 3 - 2\sqrt{2}}{4\gamma} = \frac{2 - \sqrt{2}}{2\gamma}. \]

This together with the given assumptions implies that $F_y$ satisfies $\gamma$-condition on $B(\bar{x}, r^*)$, and
\[
\{ F_y(\bar{x}) \} \cup K \cup \text{im}F_y''(x) \subseteq \text{im}F_y'(x) \quad \text{for each } x \in B(\bar{x}, r^*)
\]
(note that $\text{im}F_y''(x) = \text{im}F'(x)$ contains $y$ and $K$, and $F_y(\bar{x}) \in K - y$). Therefore, one can apply Theorem 2.1 to conclude that, if $K$ is closed or the relative interior $\text{ri} \ K$ of $K$ is nonempty, then the perturbed problem
\[
F(x) - y = F_y(x) \geq_K 0
\]
is solvable and its solution set $S(y)$ satisfies that
\[
d(\bar{x}, S(y)) \leq \tau\|F_y'({\bar{x}})\|d(F_y(\bar{x}), K) \leq \tau\|F_y'(\bar{x})\|\|F_y(\bar{x}) - F(\bar{x})\| = \frac{2 + \sqrt{2}}{2}\|F'(\bar{x})\|\|y\|.
\]

This proves (2.15) and the proof is complete. \hfill \Box

3. Notations, notions and auxiliary lemmas. Assume that $X$ and $Y$ are Hilbert spaces for the whole paper. As usual, for a bounded linear operator $A$, we use $\|A\|$, $\text{ker}A$ and $\text{im}A$ to denote respectively the operator norm of $A$, the kernel and the image of $A$. More general, for a $k$-multilinear bounded operator $A : X^k \to Y$, with any natural number $k$, we define
\[
\text{im}A := \{ A(x_1, ..., x_k) : (x_1, ..., x_k) \in X^k \}
\]
and
\[
\|A\| := \sup\{ \|A(x_1, ..., x_k)\| : (x_1, ..., x_k) \in X^k, \|x_i\| \leq 1 \text{ for each } i \}.
\]
Furthermore, for a nonempty set $D$ in $X$ (or in other normed spaces), it would be convenient to use the notation $\|D\|$ to denote its distance to the origin, that is,
\[
\|D\| = \inf\{\|a\| : a \in D \}, \quad (3.1)
\]
with the convention that $\|\emptyset\| = +\infty$. We also make the convention that $D + \emptyset = \emptyset$ for each set $D$.

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
\[
AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (A^\dagger A)^* = (A^\dagger A) \quad \text{and} \quad (AA^\dagger)^* = (AA^\dagger), \quad (3.2)
\]
where $A^*$ denotes the adjoint operator of $A$. For convenience, we list some facts regarding the Moore-Penrose inverse:
\[
(A^*)^\dagger = (A^\dagger)^*, \quad A^\dagger = (A^* A)^\dagger A^* = A^* (A A^*)^\dagger, \quad (AA^\dagger)^\dagger = (A^*)^\dagger A^\dagger, \quad (3.3)
\]
and
\[
A^\dagger A = \Pi_{(\text{ker}A)^\perp}, \quad AA^\dagger = \Pi_{\text{im}A}, \quad A^\dagger \Pi_{\text{im}A} = A^\dagger AA^\dagger = A^\dagger, \quad (3.4)
\]
where $\Pi_D$ denotes the projection on subset $D$; see for example [11] for more details.

Notice by (3.3) that $(A^\dagger)^* A^\dagger = (A^*)^\dagger A^\dagger = (A A^*)^\dagger$ and so
\[
\|(A A^*)^\dagger\| = \|(A^*)^\dagger A^\dagger\| = \|A^\dagger\|^2.
\] (3.5)

**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 $\{u \in D : \|A^\dagger u\| = \|A^\dagger D\|\} \neq \emptyset$.

**Proof.** Choose a sequence $\{u_n\} \subseteq D$ such that $\|A^\dagger u_n\| \to \|A^\dagger D\|$. Then $\{A A^\dagger u_n\}$ is bounded. Since $\{u_n\} \subseteq D \subseteq \text{im} A$ by assumption, and since $A A^\dagger$ is the projection $\Pi_{\text{im} A}$ on $\text{im} A$ (see (3.4)), one has that $u_n = A A^\dagger u_n$ 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, there exists a sequence $\{v_n\} \subseteq D$, each element of which can be represented a convex combination of some elements of $\{u_n\}$, such that $v_n$ converges to $u_0$ strongly. Thus, by the continuity and linearity of $A^\dagger$, it is routine to check that $\|A^\dagger u_0\| = \|A^\dagger D\|$ and so we complete the proof. □

**Lemma 3.2.** Let $0 < r \leq \frac{1}{2}$ and let $\bar{x} \in X$. Suppose that
\[
\text{im} F''(x) \subseteq \text{im} F'(\bar{x}) \quad \text{for each } x \in B(\bar{x}, r)
\] (3.6)
and that $F$ satisfies the $\gamma$-condition on $B(\bar{x}, r)$. Let $x \in X$ satisfy that
\[
\|x - \bar{x}\| < \min \left\{ r, \frac{2 - \sqrt{2}}{2\gamma} \right\}.
\] (3.7)

Then
\[
\text{im} F'(x) = \text{im} F'(\bar{x})
\] (3.8)
and
\[
\|F'(x)^\dagger F'(\bar{x})\| \leq \left(2 - \frac{1}{1 - \gamma \|x - \bar{x}\|^2}\right)^{-1} = \frac{(1 - \gamma \|x - \bar{x}\|^2)}{1 - 4\gamma \|x - \bar{x}\|^2 + 2\gamma^2 \|x - \bar{x}\|^2}.
\] (3.9)

Moreover, if assume additionally that $F(\bar{x}) \in \text{im} F''(\bar{x})$ then
\[
F(x) \in \text{im} F'(\bar{x}) = \text{im} F'(x).
\] (3.10)

**Proof.** (i). Since
\[
F'(x) = F'(\bar{x}) + \int_0^1 F''(\bar{x} + \tau(x - \bar{x}))(x - \bar{x})d\tau,
\] (3.11)
assumption (3.6) implies that the image of $F'(x)$ is contained in $\text{im} F'(\bar{x})$. Since, restricted on $\text{im} F'(\bar{x})$, $F'(\bar{x})F'(\bar{x})^\dagger$ is the identity map (see (3.4)), 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 (3.2) that
\[
F'(x) = F'(\bar{x})F'(\bar{x})^\dagger (F'(x) - F'(\bar{x})) + F'(\bar{x}) = F'(\bar{x})(I_X - W).
\] (3.12)
Assuming that $I_X - W$ is invertible, (3.8) is seen to hold (by (3.12)), and $F'(\bar{x}) = F'(x)(I_X - W)^{-1}$ and so
\[
\|F'(x)^\dagger F'(\bar{x})\| = \|F'(x)(I_X - W)^{-1}\| \leq \|(I_X - W)^{-1}\|.
\] (3.13)
because $F'(x)^\dagger F'(x)$ is a projection operator (so of norm bounded by 1) by (3.4). Thus (3.9) will also be shown provided that
\[
\|(I_X - W)^{-1}\| \leq \frac{(1 - \gamma\|x - \bar{x}\|)^2}{1 - 4\gamma\|x - \bar{x}\| + 2\gamma^2\|x - \bar{x}\|^2}.
\] (3.14)
To accomplish this, we use (3.11) and the assumed $\gamma$-condition to note that
\[
\|W\| \leq \int_0^1 \|F'(\bar{x})^\dagger F''(\bar{x} + \tau(x - \bar{x}))\|\|x - \bar{x}\|d\tau
\leq \int_0^1 \frac{2\gamma}{(1 - \gamma\|x - \bar{x}\|)^2}\|x - \bar{x}\|d\tau
= \frac{1}{(1 - \gamma\|x - \bar{x}\|)^2} - 1.
\] (2.5)
Further, since $\gamma\|x - \bar{x}\| \leq \frac{2 - 2\sqrt{2}}{2}$ by (3.7), we also see that $\|W\| < 1$ as
\[
\frac{1}{(1 - \gamma\|x - \bar{x}\|)^2} < \frac{1}{(1 - 2\sqrt{2})^2} = 2.
\]
By the Banach Lemma, $I_X - W$ is therefore invertible and (3.14) holds.

Finally, suppose that $F(\bar{x}) \in \text{im}F'(\bar{x})$. To establish (3.10), we need only show that $F(x) \in \text{im}F'(\bar{x})$. But this is clearly true from the identity
\[
F(x) = F(\bar{x}) + \int_0^1 F'(\bar{x} + \tau(x - \bar{x}))(x - \bar{x})d\tau
\]
because $F'(\bar{x} + \tau(x - \bar{x}))(x - \bar{x}) \in \text{im}F'(\bar{x})$ by (3.8) for any $\tau \in [0, 1]$. $\spadesuit$

Let $\xi, \gamma \in [0, +\infty)$ and define the function $\phi_{\xi, \gamma}$ by
\[
\phi_{\xi, \gamma}(t) := \xi - t + \frac{\gamma t^2}{1 - \gamma t} \quad \text{for each } t \in [0, 1/\gamma).
\] (3.15)
If $0 \leq \xi \leq \frac{3 - 2\sqrt{2}}{\gamma}$, then the smaller root $r^*$ of $\phi_{\xi, \gamma}$ on $[0, 1/\gamma)$ is given by (2.4), that is
\[
r^* := 1 + \gamma \xi - \sqrt{(1 + \gamma \xi)^2 - 8\gamma \xi}.
\] (3.16)
which satisfies that
\[
0 \leq r^* \leq \frac{1 + \gamma \xi}{4\gamma} \leq \frac{2 - \sqrt{2}}{2\gamma}.
\] (3.17)
For convenience, we define the function $w : [0, 3 - 2\sqrt{2}] \to [1, \frac{2 + \sqrt{2}}{2}]$ by
\[
w(t) := \frac{2}{1 + t + \sqrt{(1 + t)^2 - 8t}} \quad \text{for each } t \in [0, 3 - 2\sqrt{2}].
\] (3.18)
Then $w$ is strictly increasing in $[0, 3 - 2\sqrt{2}]$ and its inverse function is given by
\[
w^{-1}(t) := \frac{t - 1}{t(2t - 1)} \quad \text{for each } t \in [1, \frac{2 + \sqrt{2}}{2}].
\] (3.19)
4. Proofs of main results. Through the whole section, we assume that $F : X \to Y$ is twice Fréchet differential and that $K$ is a nonempty convex cone. As before, let $S$ denote the solution set of (1.1).

Let $\xi, \gamma \in \mathbb{R}$ be such that
\[
\gamma \geq 0 \quad \text{and} \quad 0 \leq \xi \leq \frac{3 - 2\sqrt{2}}{\gamma}.
\]
(4.1)

Let $\phi := \phi_{\xi, \gamma}$ be defined as in (3.15), with the corresponding smaller root $r^*$ (see (3.16)). Let $f : D \subseteq X \to Y$ 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 $F$ respectively with initial points $t_0 = 0$ and $x_0$, that is,
\[
t_{n+1} := t_n - \phi'(t_n)^{-1}\phi(t_n) \quad \text{for each } n = 0, 1, \ldots
\]
and
\[
x_{n+1} := x_n - f'(x_n)^{\dagger}f(x_n) \quad \text{for each } n = 0, 1, \ldots
\]
(4.2) (4.3)

The convergence property of the Newton method (4.3) is described in the following proposition, which is a direct consequence of [12, Theorem 3.3] (and its proof). Note that [12, Theorem 3.3] was proved for the case when $X$ and $Y$ are finite-dimensional spaces, but the proof also works for the general Hilbert space setting.

**Proposition 4.1.** Suppose that $f : \Omega \subseteq X \to Y$ is a nonlinear operator with continuous second Fréchet derivative on the interior $\text{int } \Omega$ of $\Omega$. Let $x_0 \in X$ be such that $B(x_0, r^*) \subseteq \Omega$, $f'(x_0)$ is surjective,
\[
\xi := \|f'(x_0)^{\dagger}f(x_0)\| \leq \frac{3 - 2\sqrt{2}}{\gamma},
\]
and that $f$ satisfies the $\gamma$-condition at $x_0$ on $B(x_0, r^*)$. Then the Newton method (4.3) with initial point $x_0$ is well-defined and the generated sequence $\{x_n\}$ converges to a zero $x^*$ of $f$ in $\overline{B}(x_0, r^*)$ satisfying
\[
\|x_n - x^*\| \leq r^* - t_n \quad \text{for each } n = 0, 1, \ldots
\]
(4.4)

We shall need to make use of the function $w$ defined as in (3.18), and recall in particular that
\[
w : [0, 3 - 2\sqrt{2}] \to [1, \frac{2 + \sqrt{2}}{2}] \text{ is strictly increasing and bijective.}
\]
(4.5)

Note that the inequality in (2.3) and the definition $r^*$ given by (2.4) can be rewritten as
\[
\gamma \xi \leq w^{-1}(\tau)
\]
(4.6)

and
\[
r^* = w(\gamma \xi) \xi
\]
(4.7)

respectively. 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
\[
\overline{K} \cup \{F(x_0)\} \cup \text{im} F''(x) \subseteq \text{im} F'(x_0) = Y_0 \quad \text{for each } x \in B(x_0, \bar{r})
\]
(4.8)
Set $\tilde{r}_0 := \min \left\{ \tilde{r}, \frac{2-\sqrt{2}}{2n} \right\}$ (so $r^* \leq \tilde{r}_0$ by (3.17)). Thanks to the given $\gamma$-condition on $\mathbf{B}(x_0, \tilde{r})$, we can apply Lemma 3.2 (to $x_0$, $\tilde{r}$ in place of $\tilde{x}$, $r$) to conclude that

$$ F(x) \in Y_0 = \text{im} F'(x) \quad \text{for each } x \in \mathbf{B}(x_0, \tilde{r}_0). $$

Moreover, by Lemma 3.1, the distance to the origin from the set $F'(x_0)^\dagger (\mathbf{K} - F(x_0))$ is attained: there exists some $c_0 \in \overline{\mathbf{K}}$ such that

$$ \|F'(x_0)^\dagger (c_0 - F(x_0))\| = \|F'(x_0)^\dagger (\mathbf{K} - F(x_0))\| = \|F'(x_0)^\dagger (K - F(x_0))\| = \xi, \quad (4.9) $$

where the last equality is because of the definition of $\xi$ given in (2.3). Now we define the map $f : \Omega := \mathbf{B}(x_0, \tilde{r}_0) \to Y_0$ by

$$ f(x) := F(x) - c_0 \quad \text{for each } x \in \Omega. $$

Then

$$ f'(x) = F'(x) \quad \text{and} \quad f''(x) = F''(x) \quad \text{for each } x \in \text{int} \ \Omega = \mathbf{B}(x_0, \tilde{r}_0). \quad (4.10) $$

In particular, $f'(x_0) = F'(x_0)$ and so $f'(x_0)^\dagger$ is surjective. Note further that $f'(x_0)^\dagger$ is exactly the restriction to $Y_0$ of the operator $F'(x_0)^\dagger$. This together with (4.8) implies that

$$ f'(x_0)^\dagger f(x_0) = F'(x_0)^\dagger (F(x_0) - c_0) \quad \text{and} \quad f'(x_0)^\dagger f''(x) = F'(x_0)^\dagger F''(x) \quad \text{for each } x \in \mathbf{B}(x_0, \tilde{r}_0). \quad (4.11) $$

It follows from the assumed $\gamma$-condition for $F$ that $f$ satisfies the $\gamma$-condition on $\mathbf{B}(x_0, \tilde{r}_0)$. Further, by (4.9), (4.11) and (4.6) together with (4.5), we have that

$$ \|f'(x_0)^\dagger f(x_0)\| = \xi \leq \frac{w^{-1}(\tau)}{\gamma} \leq \frac{3 - 2\sqrt{2}}{\gamma}. \quad (4.12) $$

Finally, since $r^* \leq \tilde{r}_0$, one has that $\mathbf{B}(x_0, r^*) \subseteq \Omega$. Thus, all assumptions of Proposition 4.1 are checked and so Proposition 4.1 is applicable to concluding that the sequence $x_n$ generated by the Newton method (4.3) for $f$ with initial point $x_0$ converges to a zero $x^*$ of $f$ and (4.4) holds. Hence, $F(x^*) = c_0 \in \overline{\mathbf{K}}$ and, (4.4) in particular implies that

$$ d(x_0, S) \leq \|x_0 - x^*\| \leq r^* = w(\gamma) \xi \leq \tau \xi \quad (4.13) $$

(see (4.6) and (4.7)). Since $K - F(x_0) \subseteq \text{im} F'(x_0)$ by (4.8), one has

$$ \|F'(x_0)^\dagger (K - F(x_0))\| \leq \|F'(x_0)^\dagger\| d(F(x_0), K). $$

This together with (4.9) implies that $\xi \leq \|F'(x_0)^\dagger\| d(F(x_0), K)$, which shows (2.6) thanks to (4.13). 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

$$ \tilde{r} > r^*, \quad \xi < \frac{3 - 2\sqrt{2}}{\gamma} \quad \text{and} \quad ri K \neq \emptyset. \quad (4.14) $$

Then there exists a sequence $\{c_n\} \subseteq ri K$ be such that $c_n \to 0$. Without loss of generality, we may assume further that, for each $n = 1, 2, \ldots$

$$ \tilde{\xi}_n := \xi + \|F'(x_0)^\dagger\| ||c_n|| < \frac{3 - 2\sqrt{2}}{\gamma}, \quad (4.15) $$
\[ r_n := w(\gamma \xi_n) \in \left[ 1, \frac{2 + \sqrt{2}}{2} \right) \]  

\[ r^*_n := w(\gamma \xi_n) \tilde{c}_n < \bar{r} \]  

(noting (4.5) and that \( \xi_n \to \xi \) and \( r^*_n \to r^* \) by (4.7)). 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:

\[ F_n(x) \geq_K 0. \]  

Then, \( F'_n = F' \), and so \( F'(x_0)^\dagger = F'_n(x_0)^\dagger \) and \( F_n(x_0) \in \text{im} F'_n(x_0) \) thanks to assumption (4.8). Moreover, we have that

\[ \tau_n := w(\gamma \xi_n) \leq w(\gamma \xi) = \tau_n \in \left[ 1, \frac{2 + \sqrt{2}}{2} \right) \]  

and

\[ r^*_n := w(\gamma \xi_n) \xi_n \leq w(\gamma \xi_n) \tilde{c}_n = \tilde{r}_n < \bar{r} \]  

thanks to (4.5). This and the given assumptions imply that \( F_n \) satisfies the \( \gamma \)-condition on \( B(x_0, r^*_n) \), and both (2.3) and (2.5) hold with \( F_n, \tau_n \) in place of \( F, \tau \) (by (4.20) and the definition of \( w^{-1} \) in (3.19), \( \gamma \xi_n = w^{-1}(\tau_n) = \frac{\tau_n^r - 1}{\tau_n(2\tau_n - 1)} \)). Moreover, by (4.7), the quantity \( r^*_n \) defined above is exactly the corresponding quantity \( r^* \) defined by (2.4) with \( \xi_n \) in place of \( \xi \). Then we apply the first conclusion of Theorem 2.1 just proved to conclude that there exists \( x_n^* \in X \) satisfying \( F_n(x_n^*) \in \overline{K} \) such that

\[ d(x_0, S_n) \leq \|x_0 - x_n^*\| \leq \tau_n \|F'(x_0)^\dagger\|d(F_n(x_0), K). \]

By the choice of \( c_n \), we have that \( F(x_n^*) \in \overline{K} + c_n \subseteq rK \subseteq K \). Hence \( x_n^* \in S \) (and so \( S \) is nonempty) and it follows that

\[ d(x_0, S) \leq \|x_0 - x_n^*\| \leq \tau_n \|F'(x_0)^\dagger\|d(F(x_0), K). \]

Letting \( n \to \infty \), we establish (2.7) as it is clear that \( \tau_n \to \tau \) and \( d(F_n(x_0), K) \to d(F(x_0), K) \). Thus the proof is complete. \( \Box \)

5. Applications. Before our discussion on the applications of Theorem 2.1 to the study for the systems (1.2) and (1.3), 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 \( \tau \in (0, +\infty) \) and \( x_0 \in X \). Let \( F : X \to Y \) be analytic on \( B(x_0, r) \) and let \( \gamma := \gamma_F(x_0) \) be defined by (2.8). 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
\[
\text{im} F''(x) \subseteq \text{im} F'(x) \quad \text{for each } x \in \mathbf{B}(x_0, r_0).
\]

Proof. Assertion (i) can be proved with a almost the same argument as that for [1, 9]. Below we verify the second assertion. For this purpose, assume that \( \text{im} F^{(k)}(x_0) \subseteq \text{im} F'(x_0) \) for each \( k = 2, \ldots \), and let \( x \in \mathbf{B}(x_0, r_0) \). Since \( F \) is analytic on \( \mathbf{B}(x_0, r) \), it follows from the Taylor expansion that
\[
F''(x) = \sum_{k=0}^{\infty} \frac{F^{(k+2)}(x_0)}{k!}(x - x_0)^k.
\]

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

As applications of Theorem 2.1 to systems (1.2) and (1.3) respectively, we derive the following corollaries, where we assume that the involved functions \( f \) and \( g \) are defined by (1.4) and are analytic at the involved point \( x_0 \).

**Corollary 5.2.** Consider the inequality system (1.2). Let \( \tau \in (1, \frac{2 + \sqrt{2}}{2} \), \( \gamma \in [\gamma_f(x_0), +\infty) \) and \( x_0 \in X \) be such that \( f'(x_0) \) is surjective and
\[
\| f(x_0) \| < \frac{\tau - 1}{\tau(2\tau - 1)\gamma} \| f(x_0) \|, \tag{5.1}
\]
where \( \gamma_f(x_0) \) is defined by (2.8) (with \( f \) in place of \( F \)). Suppose that \( f \) is analytic on \( \mathbf{B}\left(x_0, \frac{2 + \sqrt{2}}{2\gamma}\right) \). Then the solution set \( f^> \) of (1.2) (consisting of all \( x \) satisfying (1.2)) is nonempty and the following estimate holds:
\[
d(x_0, f^>) \leq \tau \| f(x_0) \|^\dagger \| f(x_0) \|^{-\dagger}. \tag{5.2}
\]

Proof. Let \( K := \{(u_i) \in \mathbb{B}_1^+: u_i > 0, \forall 1 \leq i \leq l, s\} \) and consider the function \( F := f \) with \( f \) defined by (1.4). Then the inequality system (1.2) is the same as the inequality system (1.1), and \( d(F(x_0), K) = \| f(x_0) \|^{-\dagger} \). Thus,
\[
\| F'(x_0) \|^\dagger \| F(x_0) \| \leq \| F'(x_0) \|^\dagger \| d(F(x_0), K) = \| f'(x_0) \|^\dagger \| f(x_0) \|^{-\dagger}. \tag{5.3}
\]
Then, (5.3) and (5.1) imply that
\[
\xi = \| F'(x_0) \|^\dagger (K - F(x_0)) \| < \frac{\tau - 1}{\tau(2\tau - 1)\gamma} \leq \frac{3 - 2\sqrt{2}}{\gamma},
\]
that is, (2.3) holds and \( \xi < \frac{3 - 2\sqrt{2}}{\gamma} \). Since
\[
\tilde{r} := \frac{2 - \sqrt{2}}{2\gamma} < \frac{1}{\gamma} \leq \frac{1}{\gamma f(x_0)},
\]
it follows from Lemma 5.1 that \( F \) satisfies the \( \gamma \)-condition at \( x_0 \) on \( \mathbf{B}(x_0, \tilde{r}) \); while assumption (2.5) holds trivially because \( F'(x_0) \) is surjective by assumptions. Moreover, it is clear that \( \text{ri} K \neq \emptyset \). Recall that \( r^* \) is defined by (2.4) and so it satisfies that
\[
r^* \leq \frac{1 + \gamma \xi}{4\gamma} < \frac{1 + 3 - 2\sqrt{2}}{4\gamma} = \frac{2 - \sqrt{2}}{2\gamma} = \tilde{r},
\]
(because $\gamma \xi < 3 - 2\sqrt{2}$). Thus, Theorems 2.1 is applicable to concluding that the solution set $f^\gamma$ of (1.1) (and (1.2)) is nonempty, and
\[
\text{d}(x_0, f^\gamma) \leq \gamma \|F'(x_0)^\dagger\| \text{d}(F(x_0), K) = \tau \|f(x_0)^\dagger\| \|f(x_0)^-\|,
\]
where the last equality is because of (5.3). This shows (5.2) and the proof is complete. 

The proof of the following corollary is similar and so we omit it here.

**Corollary 5.3.** Consider the inequality system (1.3), and let $F : X \to \mathbb{R}^l \times \mathbb{R}^m$ be defined by $F := (f, g)$. Let $\tau \in (1, \frac{2 + \sqrt{2}}{2}, \gamma \in [\gamma F(x_0), +\infty)$ and $x_0 \in X$ be such that
\[
(R^l \times \{0_m\}) \cup \{(f(x_0), g(x_0))\} \cup \text{im}F(k)x_0) \subseteq \text{im}F'(x_0) \quad \text{for each } k = 2, \ldots
\]

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

Suppose that $f$ and $g$ are analytic on $B\left(x_0, \frac{2 + \sqrt{2}}{2}\right)$. Then the solution set $S(f, g)$ of (1.3) (consisting of all $x$ satisfying (1.3)) is nonempty and
\[
\text{d}(x_0, S(f, g)) \leq \tau \|F'(x_0)^\dagger\| \|f(x_0)^-\|^2 + \|g(x_0)^2\|^\frac{1}{2}.
\]

**Remark 5.1.** Consider the inequality system (1.2) for the special case when $x_0$ satisfies that
\[
f_i(x_0) < 0 \quad \text{for each } 1 \leq i \leq l.
\]
Then Corollary 5.2 is a stronger result than Theorem 2.2. To see this, we suppose that (5.7) holds and choose $\gamma = \gamma_f(x_0)$. Then $\text{Diag}(f(x_0)^+) = 0$. Hence
\[
\delta(f; x_0) = \|f'(x_0)f'(x_0)^+\|^\frac{1}{2} = \|f'(x_0)^\dagger\|,
\]
(see (3.5)) and assumption (1.10)$\iff (f'(x_0)$ is surjective), but (2.9)$\implies (5.1)$ because (5.8) holds and
\[
\gamma = \gamma_f(x_0) = \sup_{k \geq 2} \left\| \frac{f'(x_0)^\dagger f'(x_0)^k(x_0)}{k!} \right\|^{\frac{1}{k \tau}} \leq \sup_{k \geq 2} \left\| f'(x_0)^\dagger \right\|^\frac{1}{k \tau} \Gamma(f, x_0) = \max\{1, \delta(f; x_0)\} \Gamma(f, x_0)
\]
(so the quantity given in the right-hand side in (5.1) is smaller than that in (2.9)). Similarly, Corollary 5.3 is a stronger result than Theorem 2.3 for the system (1.3) in the special case when $x_0$ satisfies (5.7).

To prove the theorems of Proofs 2.2 and 2.3, we need first to do some preparations. 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
\[
(F'(x); \text{Diag}(v))(z; u) := F'(x)z + \text{Diag}(v)u; 0 \quad \text{for each } (z; u) \in X \times \mathbb{R}^l.
\]
Using the “matrix” notation, the operator $(F'(x); \text{Diag}(v))$ can be reexpressed as
\[
(F'(x); \text{Diag}(v))(z; u) = \begin{pmatrix}
F'(x) & \text{Diag}(v) \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
\begin{pmatrix}
z \\
u
\end{pmatrix}
\end{pmatrix}
\quad \text{for each } (z; u) \in X \times \mathbb{R}^l.
\]
Let $G(x; v)$ be the $(l + m) \times (l + m)$ matrix defined by
\[ G(x; v) := \begin{pmatrix} f'(x)f'(x)^* + \text{Diag}(v^2) & f'(x)g'(x)^* \\ g'(x)f'(x)^* & g'(x)g'(x)^* \end{pmatrix}, \tag{5.10} \]
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 is useful.

**Lemma 5.4.** Let $(x; v) \in X \times \mathbb{R}^l$. Then we have the following assertions:
\[ (F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^* = G(x; v) \tag{5.11} \]
and
\[ \text{im} (F'(x); \text{Diag}(v)) = \text{im} G(x; v). \tag{5.12} \]

**Proof.** We assert that
\[ (F'(x); \text{Diag}(v))^*(w_1, w_2) = \begin{pmatrix} f'(x)^* & g'(x)^* \\ \text{Diag}(v) & 0_{l \times m} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \text{ for each } (w_1; w_2) \in \mathbb{R}^l \times \mathbb{R}^m. \tag{5.13} \]
In fact, by definition, for any $(z, u) \in X \times \mathbb{R}^l$, one checks that
\[ \langle (z, u), (F'(x); \text{Diag}(v))^*(w_1, w_2) \rangle = \langle (F'(x); \text{Diag}(v))(z, u), (w_1, w_2) \rangle = \langle f'(x)z, w_1 \rangle + \langle \text{Diag}(v)u, w_1 \rangle + \langle g'(x)z, w_2 \rangle = \langle z, f'(x)^*w_1 \rangle + \langle u, \text{Diag}(v)w_1 \rangle + \langle z, g'(x)^*w_2 \rangle = \langle (z, u), \begin{pmatrix} f'(x)^* & g'(x)^* \\ \text{Diag}(v) & 0_{l \times m} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \rangle, \]
and (5.13) is established. Therefore, by (5.9) and (5.13), one has that
\[ (F'(x); \text{Diag}(v)) \circ (F'(x); \text{Diag}(v))^* = \begin{pmatrix} f'(x) & \text{Diag}(v) \\ g'(x) & 0_{m \times l} \end{pmatrix} \begin{pmatrix} f'(x)^* & g'(x)^* \\ \text{Diag}(v) & 0_{l \times m} \end{pmatrix} \]
and (5.11) is seen to hold.

To show (5.12), we set $Y_0 := \text{im}(F'(x); \text{Diag}(v))$. By (5.11), 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 $(F'(x); \text{Diag}(v))$ on $Y_0$, 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. \( \square \)

**Proofs of Theorems 2.2 and 2.3.** As before, we here only present the proof for Theorem 2.3 as the proof for Theorem 2.2 is similar.

Let $H : X \times \mathbb{R}^l \to \mathbb{R}^l \times \mathbb{R}^m$ be the operator defined by $H := (h; g)$ with each $h = (h_i)$ and each $g = (g_j)$ given by
\[ \begin{cases} h_i(x; v) := f_i(x) - v_i^2, & 1 \leq i \leq l \\
g_j(x; v) := g_j(x), & 1 \leq j \leq m \end{cases} \tag{5.14} \]
for each \((x; v) \in X \times \mathbb{R}^l\) with \(v = (v_1, ..., v_l) \in \mathbb{R}^l\). Consider the following inequality system on \(X \times \mathbb{R}^l\):

\[
\begin{align*}
    &h_i(x; v) > 0 & i = 1, ..., l_S, \\
    &h_i(x; v) \geq 0 & i = l_S + 1, ..., l, \\
    &g_j(x; v) = 0 & j = 1, ..., m.
\end{align*}
\] (5.15)

As before, we use \(S(h; g)\) to denote the solution set of 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,

\[
S_0 := \{ x \in X : \exists v \in \mathbb{R}^l \text{ s.t. } (x; v) \in S(h; g) \} \subseteq S(f; g).
\] (5.16)

Let \(v_0 := (\sqrt{f^*(x_0)}') \in \mathbb{R}^l\). Then

\[
\|h(x_0; v_0)^-\| + \|g(x_0; v_0)^-\| = \|f(x_0)^-\|^2 + \|g(x_0)\|^2
\] (5.17)

and

\[
H(x_0; v_0) = F(x_0) - (f(x_0)^+; 0) \in F(x_0) + \mathbb{R}^l \times \{0_m\}.
\] (5.18)

Let \(\gamma := \max\{1, \delta(f; g; x)\}(1 + \Gamma(f; g; x))\) and we will apply Corollary 5.3 to the inequality system (5.15) (with \((x_0, v_0)\) in place of \(x_0\)). To do this, we note first that

\[
H'(x_0; v_0) = (F'(x_0); \text{Diag}(-2v_0)), \quad H^{(k)}(x_0; v_0) = F^{(k)}(x_0) \quad \text{for each } k > 2.
\] (5.19)

To express \(H''(x_0; v_0)\), we use \(D_{\text{li}} : \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}^l \times \mathbb{R}^m\) to denotes the operator defined by

\[
D_l(u; v) := \begin{pmatrix} \text{Diag}(u) & v \\ 0_m \end{pmatrix} \quad \text{for any } (u; v) \in \mathbb{R}^l \times \mathbb{R}^l.
\] (5.20)

Then we have

\[
H''(x_0; v_0) = (F''(x_0); -2D_l)
\] (5.21)

in the sense that

\[
(F''(x_0); -2D_l)(z_1; u_1)(z_2; u_2) := F''(x)z_1z_2 - 2D_l(u_1; u_2) \quad \text{for any } (z_1; u_1), (z_2; u_2) \in X \times \mathbb{R}^l,
\]

Therefore,

\[
\text{im}H^k(x_0; v_0) = \text{im}F^k(x_0) \quad \text{for each } k > 2
\] (5.22)

and

\[
\text{im}H''(x_0; v_0) = \text{im}F''(x_0) + \text{im}(-2D_l) \subseteq \text{im}F''(x_0) + \mathbb{R}^l \times \{0_m\}.
\] (5.23)

Moreover, by (5.10) and (5.19), one apply Lemma 5.4 to obtain that

\[
H'(x_0; v_0) \circ H'(x_0; v_0)^* = G(x_0; 2v_0) = G(x_0)
\] (5.24)

(noting that \((2v_0)^2 = 4f(x_0)^+\)) and

\[
\text{im}G(x_0) = \text{im}H'(x_0; v_0).
\] (5.25)
Combining this together with (5.18), (5.22) and (5.23), we conclude from assumption (2.11) that
\[ \mathbb{R}^l \times \{0_m\} \cup \{H(x_0; v_0)\} \cup \text{im} H^{(k)} \subseteq \text{im} H'(x_0; v_0) + \mathbb{R}^l \times \{0_m\} = \text{im} H'(x_0; v_0) \quad \text{for each } k \geq 2, \]
(5.26)
because $\text{im} H'(x_0; v_0)$ contains $\mathbb{R}^l \times \{0_m\}$ and is linear subspace. By (5.24) and the definition of $\beta_H$, we can apply (3.5) (to $H'(x_0; v_0)$ in place of $A$) to get that
\[ \beta_H(x_0; v_0) = \|H'(x_0; v_0)\| = \|G(x_0)\|^{\frac{1}{2}} = \delta(f, g; x_0). \]
(5.27)
Thus we conclude that
\[ \sup_{k \geq 2} \left\| \frac{H^{(k)}(x_0; v_0)}{k!} \right\|^{\frac{1}{k!}} = \max \left\{ \frac{\|H''(x_0; v_0)\|}{2}, \sup_{k \geq 2} \left\| \frac{H^{(k)}(x_0; v_0)}{k!} \right\|^{\frac{1}{k!}} \right\} \leq 1 + \Gamma(f, g; x_0) \]
(5.28)
as
\[ \|H''(x_0; v_0)\| \leq 2 + \|F''(x_0)\| \quad \text{and} \quad \|H^{(k)}(x_0; v_0)\| = \|F^{(k)}(x_0)\| \quad \text{for each } k > 2; \]
and, further by the definition of $\gamma_H$, that
\[ \gamma_H(x_0; v_0) \leq \sup_{k \geq 2} \|H'(x_0; v_0)\|^{\frac{1}{k!}} \cdot \sup_{k \geq 2} \left\| \frac{H^{(k)}(x_0; v_0)}{k!} \right\|^{\frac{1}{k!}} \]
\[ \leq \max \{1, \delta(f, g; x_0)\} \left(1 + \Gamma(f, g; x_0)\right) \]
(5.29)
where the last inequality is because of (5.28) and the following equality:
\[ \sup_{k \geq 2} t^{\frac{1}{k!}} = \begin{cases} t & \text{if } t > 1, \\ 1 & \text{if } t \leq 1. \end{cases} \]

It follows from (5.17), (5.27) and assumption (2.12) that
\[ \|h(x_0; v_0)\|^2 + \|g(x_0; v_0)\|^2 < \frac{\tau - 1}{\tau(2\tau - 1)(1 + \Gamma(f, g; x_0)) \delta(f, g; x_0) \max \{1, \delta(f, g; x_0)\}} \]
\[ = \frac{\tau - 1}{\tau(2\tau - 1)\gamma \|H'(x_0; v_0)\|}. \]
(5.30)
Since $1 + \Gamma(f, g; x_0) \leq \gamma$, it follows from the analyticity assumption that $H$ is analytic on $B\left(x_0, \frac{2\sqrt{\tau}}{2\tau}\right)$. This together with (5.26) and (5.30) implies that all assumptions in Corollary 5.3 for the inequality system (5.15) (with $(x_0, v_0)$ in place of $x_0$) are satisfied and then Corollary 5.3 is applicable to getting that $S(h; g) \neq \emptyset$ and
\[ d((x_0; v_0), S(h; g)) \leq \tau \|H'(x_0; v_0)\| (\|h(x_0; v_0)\|^2 + \|g(x_0; v_0)\|^2)^{\frac{1}{2}}. \]
Using (5.17) and (5.27), we obtain that
\[ d((x_0; v_0), S(h; g)) \leq \tau \delta(f, g; x_0) (\|f(x_0)\|^2 + \|g(x_0)\|^2)^{\frac{1}{2}}. \]
Let $\epsilon > 0$ and let $(x^*, v^*) \in S(h; g)$ be such that
\[ \|(x_0; v_0) - (x^*; v^*)\| \leq \tau \delta(f, g; x_0) (\|f(x_0)\|^2 + \|g(x_0)\|^2)^{\frac{1}{2}} + \epsilon. \]
(5.31)
Then $x^* \in S(f, g)$ by (5.16) and
\[ d(x_0, S(f, g)) \leq \|x_0 - x^*\| \leq \|(x_0; v_0) - (x^*; v^*)\|. \]
Thus (2.13) follows from (5.31) because $\epsilon > 0$ is arbitrary and the proof is complete. \(\Box\)
REFERENCES