Necessary Optimality Conditions for Interval Optimization Problems with Functional and Abstract Constraints

Fabiola Roxana Villanueva[1] · **Valeriano Antunes de Oliveira**[2]

Abstract

This work addresses interval optimization problems in which the objective function is interval-valued while the constraints are given in functional and abstract forms. The functional constraints are described by means of both inequalities and equalities. The abstract constraint is expressed through a closed and convex set with a nonempty interior. Necessary optimality conditions are derived, given in a multiplier rule structure involving the gH-gradient of the interval objective function along with the (classical) gradients of the constraint functions and the normal cone to the set related to the abstract constraint. The main tool is a specification of the Dubovitskii–Milyutin formalism. We defined an appropriated notion of directions of decrease to an interval-valued function, using the lower–upper partial ordering of the interval space (LU order).

1 Introduction

The interval-valued optimization problems have been of ample interest since real-world optimization problems often involve inexactness or imprecise data due to some unpredictable circumstances or measurement miscalculations in the objective function and/or in the constraints. In recent years, interval optimization theory and its applica-

---

✉ Valeriano Antunes de Oliveira
 valeriano.oliveira@unesp.br

 Fabiola Roxana Villanueva
 olita_villanueva@hotmail.com

[1] Universidad Mayor de San Andrés, Carrera de Matemática, La Paz, Bolivia

[2] São Paulo State University, São José do Rio Preto, São Paulo, Brazil

tions have been studied by many authors. For instance, in Cao et al. [7] a vehicle routing problem with interval demands is investigated, where the demands of customers are uncertain. Costa et al. [9] used interval analysis in the calculation of the 3D structure of a protein molecule: the measurements of the distance between atoms obtained from nuclear magnetic resonance are not precise (such measures are important to determine the protein structure). In Das et al. [10], an imperfect production inventory model is developed in the interval environment. Kumar, Behera, and Bhurjee [16] studied a portfolio selection model with interval-type random parameters considering risk measures as value-at-risk. In Wen et al. [25], an interval optimization method is developed to determine the optimal size of an energy storage system in a hybrid ship power system to reduce the fuel cost, capital cost of the system, and emissions of greenhouse gases.

With respect to the theoretical developments, see, e.g., Ahmad et al. [1, 2], Antczak [3], Chalco-Cano et al. [8], Kummari and Ahmad [17], Singh et al. [19], Stefanini and Arana-Jiménez [21], Tung [23], Van Luu and Mai [24], Wu [27], Zhang et al. [28] and Zhao and Bin [29]. In all of the aforementioned studies, optimality conditions for interval-valued optimization problems were derived. To express such conditions, a suitable concept of derivative for interval-valued functions should be chosen. Most of the authors use Hukuhara differentiability (see Hukuhara [13]) or generalized Hukuhara differentiability (see Chalco-Cano [8]). The first one is known to be very restrictive. Despite the fact that the generalized Hukuhara differentiability is less restrictive than Hukuhara differentiability, in the case of many variables it does not generalize the one-dimensional generalized Hukuhara differentiability given in Stefanini and Bede [22]. Recently, Stefanini and Arana-Jiménez [21] give a new definition of a generalized Hukuhara differentiability for functions with many variables which does extend the one-dimensional case.

Necessary optimality conditions of Karush–Kuhn–Tucker (KKT) type for interval optimization problems were given, for instance, in [2, 8, 17, 21, 24, 28, 29] in the mono-objective (scalar) case. KKT necessary conditions for the multi-objective case were furnished in [1, 3, 23].

Sufficient optimality conditions under convexity can be found in, among others, [21] for scalar problems and in [3, 19, 23, 27] for multi-objective problems; under generalized convexity, they can be found in [2, 17, 24], for scalar problems.

Duality theory for interval problems was developed, for example, in [17, 24] in the scalar case, and in [3, 23] in the vector case.

Hukuhara differentiability was used in [27, 28]. In other studies by [1, 2, 19], generalized Hukuhara differentiability (given in [8]) was utilized. In all of the papers cited in the previous sentence, the constraint functions are assumed to be differentiable (in the classic sense). Problems in which the constraint functions are non-smooth were treated in [3, 17, 23, 24, 29], where neither Hukuhara nor generalized Hukuhara differentiability was used. Assumptions were made directly on the extremum functions that define the interval function. Clarke subdifferential was used in [3, 29], limiting subdifferential in [17], the subdifferential of convex analysis in [23], and convexificators (a generalized notion of subdifferential) in [24].

Most of the papers mentioned so far deal with interval problems posed in finite-dimensional Euclidean spaces and with a finite number of constraints. Stefanini and

Arana-Jiménez [21], actually, developed necessary and sufficient optimality conditions for fuzzy optimization problems with inequality constraints. However, such conditions are easily converted to interval problems. Tung [23] studied a semi-infinite interval problem, that is, the problem is posed in $\mathbb{R}^n$ and has possibly an infinite number of (inequality) constraints. Van Luu and Mai [24] considered a very general interval problem, posed in a real Banach space (possibly infinite dimensional), with equality, inequality and set constraints, and within a non-smooth context. Despite possessing, possibly, an infinite number of variables, the number of equality and inequality constraints is finite. Zhao and Bin [29] also considered an interval optimization problem posed in a Banach space in the non-smooth context but with (a finite number of) inequality constraints only. The novelty was the fact that the constraints were allowed to be parametrized into compact subsets of some given topological spaces. Such problems are called robust optimization problems.

The main goal of this work is to present necessary optimality conditions for interval-valued optimization problems in several variables with inequality and equality functional as well as abstract (or set) constraints. We use the new gH-differentiability given in Stefanini and Arana-Jiménez [21] and the LU order relation. We present necessary optimality conditions by means of a multiplier rule of KKT type. A specification of the Dubovitskii–Milyutin formalism is utilized to establish such optimality conditions.

The Dubovitskii–Milyutin formalism can be found, for example, in Girsanov [12]. This formalism consists of a unified functional-analytic approach to treat the so-called extremum problems. A great diversity of problems can be attacked through this approach, from classical mathematical programming to optimal control; very general problems fit into such a framework since the only requirement in the space in which the problem is posed is to be a locally convex topological linear one. However, the theory described in [12] can deal only with problems in which the objective function is a functional, which is not the case in interval optimization. So, we propose a modification of the Dubovitskii–Milyutin formalism to comprise optimization problems in which the objective function is interval-valued. To be specific, we define an appropriate notion of directions of decrease for interval-valued functions, by using the LU order relation, in such a way that it is possible to characterize geometrically the optimal solutions. The set of all directions of decrease is, then, described by means of the gH-gradient of the objective function.

The KKT-type necessary optimality conditions furnished here are similar to those given in Stefanini and Arana-Jiménez [21], expressed in terms of the gH-gradient of the objective function. In other studies cited above, the KKT conditions do not involve the gH-gradient, but the classic gradient (or a subdifferential, in the case of non-smooth functions) of the extremum functions which define the interval objective function. A distinction between our results and those of Stefanini and Arana-Jiménez resides in the constraint qualification used: while we require a positive linear independence condition, they assume a linear independence one (which is stronger). Moreover, the interval problem we worked on is more general, since equality, inequality, and abstract constraints are allowed, whereas in [21] only inequality constraints are present.

As mentioned, Van Luu and Mai [24] worked with a very general problem in the non-smooth context. Notwithstanding, they made strong assumptions on the data. Despite the fact that the objective extremum functions and equality constraints are merely

assumed to be locally Lipschitz and the inequality constraints to be continuous, it is supposed that the objective extremum functions and inequality constraints admit certain convexificators and that the absolute values of the equality constraints are Clarke regular. Further, they also need a regularity condition in the sense of Ioffe. The KKT necessary conditions are obtained under a Mangasarian–Fromovitz-type constraint qualification and formulated by means of subdifferentials of the functions that define the problem. Although their results are established in a more general setting, no kind of derivative for interval-valued functions is used. The KKT conditions we derive here are stated through the gH-derivative concept (given in [21]), which is a genuine derivative concept for interval-valued functions. Moreover, owing to the methodology we made use of, the Dubovitskii–Milyutin formalism, our results are easily extended to include problems posed in Banach spaces.

The work is organized as follows. In Sect. 2, we give basic definitions and important results from the literature, and present the interval optimization problem. Section 3 is dedicated to obtain the necessary optimality conditions. Some examples and possible applications are discussed in Sect. 4. To finish, in Sect. 5 we present some conclusions of this study.

2 Preliminaries

Here, we set the notation, give some basic definitions, state some important results from the literature and pose the interval-valued optimization problem the paper is concerned with.

2.1 Basic Definitions and Important Properties

The interval-valued objective function we work with is defined from $\mathbb{R}^n$ into the space $\mathcal{K}_C = \{[\underline{a}, \overline{a}] : \underline{a}, \overline{a} \in \mathbb{R}, \ \underline{a} \leq \overline{a}\}$. In such a space, the following interval arithmetic operations are defined. Given $A = [\underline{a}, \overline{a}]$, $B = [\underline{b}, \overline{b}]$, $C = [\underline{c}, \overline{c}] \in \mathcal{K}_C$ and $\lambda \in \mathbb{R}$,

$$A + B = [\underline{a} + \underline{b}, \overline{a} + \overline{b}] \quad \text{and} \quad \lambda \cdot A = \begin{cases} [\lambda \underline{a}, \lambda \overline{a}], & \text{if } \lambda \geq 0, \\ [\lambda \overline{a}, \lambda \underline{a}], & \text{if } \lambda < 0. \end{cases}$$

The gH-difference of two intervals $A, B \in \mathcal{K}_C$ (see Stefanini [20]) is defined by:

$$A \ominus_{gH} B = C \Leftrightarrow \begin{cases} A = B + C, \text{ or} \\ B = A + (-1) \cdot C. \end{cases}$$

Note that $A \ominus_{gH} A = [0, 0]$. Moreover, the gH-difference of two intervals always exists and, if $A = [\underline{a}, \overline{a}]$, $B = [\underline{b}, \overline{b}] \in \mathcal{K}_C$, it holds that $A \ominus_{gH} B = \left[(\underline{a} - \underline{b}) \vee (\overline{a} - \overline{b}) \right]$, where $[\underline{a} \vee \overline{a}] := \left[\min\{\underline{a}, \overline{a}\}, \max\{\underline{a}, \overline{a}\} \right]$.

The Cartesian product $\mathcal{K}_C \times \cdots \times \mathcal{K}_C$, of n-factors $\mathcal{K}_C$, is denoted by $\mathcal{K}_C^n$. If $A = (A_1, \ldots, A_n)$, $B = (B_1, \ldots, B_n) \in \mathcal{K}_C^n$ and $\lambda \in \mathbb{R}$, the sum and the scalar product in $\mathcal{K}_C^n$ are defined as follows:

$$A \oplus B = (A_1, \ldots, A_n) \oplus (B_1, \ldots, B_n) = (A_1 + B_1, \ldots, A_n + B_n),$$
$$\lambda \odot A = \lambda \odot (A_1, \ldots, A_n) = (\lambda \cdot A_1, \ldots, \lambda \cdot A_n).$$

Note that any real number a can be seen as a degenerated interval $[a, a]$. Then, for all $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ and $A = (A_1, \ldots, A_n) \in \mathcal{K}_C^n$, in which $A_i = \left[\underline{a}_i, \overline{a}_i\right] \in \mathcal{K}_C$, $i = 1, \ldots, n$, it follows that

$$A \oplus v = \left(\left[\underline{a}_1, \overline{a}_1\right], \ldots, \left[\underline{a}_n, \overline{a}_n\right]\right) + ([v_1, v_1], \ldots, [v_n, v_n])$$
$$= \left(\left[\underline{a}_1, \overline{a}_1\right] + [v_1, v_1], \ldots, \left[\underline{a}_n, \overline{a}_n\right] + [v_n, v_n]\right).$$

We denote $v \in A$ to mean that $v_i \in A_i$ for all $i = 1, \ldots, n$. Given $S \subset \mathbb{R}^n$, we define the sum $A \oplus S$ as

$$A \oplus S = \{a + s : a \in A, \ s \in S\}.$$

Next, we have the definition of the LU order relation which will be used here.

Definition 2.1 (Kulish and Miranker [15]) Let $A = \left[\underline{a}, \overline{a}\right]$ and $B = \left[\underline{b}, \overline{b}\right] \in \mathcal{K}_C$. The lower and upper order relation, LU in short, is defined by

(i) $A \leqq_{LU} B$ if, and only if, $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$;
(ii) $A \leq_{LU} B$ if, and only if, either $\underline{a} < \underline{b}$ and $\overline{a} \leq \overline{b}$ or $\underline{a} \leq \underline{b}$ and $\overline{a} < \overline{b}$;
(iii) $A <_{LU} B$ if, and only if, $\underline{a} < \underline{b}$ and $\overline{a} < \overline{b}$.

An application $F : S \subseteq \mathbb{R}^n \to \mathcal{K}_C$ such that $F(x) = \left[\underline{f}(x), \overline{f}(x)\right]$ for all $x \in S$, where $\underline{f}, \overline{f} : S \to \mathbb{R}$, with $\underline{f}(x) \leq \overline{f}(x)$ for all $x \in S$, is called an interval-valued function of several variables. Functions $\underline{f}$ and $\overline{f}$ are its extremum functions.

The Pompeiu–Hausdorff metric between $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}] \in \mathcal{K}_C$ is defined as $d_H(A, B) := \max\{|\underline{a} - \underline{b}|, |\overline{a} - \overline{b}|\}$. In Diamond and Kloeden [11] it is shown that $(\mathcal{K}_C, d_H)$ is a complete and separable metric space. It is natural, then, to consider the limit and continuity concepts of interval-valued functions defined by means of the Pompeiu–Hausdorff metric d_H. Before enunciating definitions and results involving the concepts of limit and continuity, let us state the technical lemma below which will be useful later in the paper.

Lemma 2.1 *Given $A = [\underline{a}, \overline{a}]$, $B = [\underline{b}, \overline{b}] \in \mathcal{K}_C$, then*

$$A \ominus_{gH} B \leqq_{LU} [d_H(A, B), d_H(A, B)].$$

Proof We have that

$$A \ominus_{gH} B = \left[\underline{a} - \underline{b} \vee \overline{a} - \overline{b}\right] \leqq_{LU} \left[|\underline{a} - \underline{b}| \vee |\overline{a} - \overline{b}|\right]$$
$$\leqq_{LU} \left[\max\{|\underline{a} - \underline{b}|, |\overline{a} - \overline{b}|\}, \max\{|\underline{a} - \underline{b}|, |\overline{a} - \overline{b}|\}\right]$$
$$= [d_H(A, B), d_H(A, B)].$$

$\square$

Theorem 2.1 (Aubin and Cellina [4]) *Let $F : S \subseteq \mathbb{R} \to \mathcal{K}_C$ be an interval-valued function such that $F(x) = \left[\underline{f}(x), \overline{f}(x)\right]$ for all $x \in S$. Given $x_0 \in S$,*

$$\lim_{x \to x_0} F(x) = \left[\lim_{x \to x_0} \underline{f}(x), \lim_{x \to x_0} \overline{f}(x)\right].$$

The usual inner product in $\mathbb{R}^n$ is denoted by $\langle \cdot, \cdot \rangle$. The Euclidean norm is denoted by $\| \cdot \|$. Given $\epsilon > 0$, $N_\epsilon(x_0)$ denotes the ϵ-neighborhood of $x_0 \in \mathbb{R}^n$, that is, $N_\epsilon(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \epsilon\}$.

Definition 2.2 (Luhandjula and Rangoaga [18]) Let $F : S \subseteq \mathbb{R}^n \to \mathcal{K}_C$, $F(x) = \left[\underline{f}(x), \overline{f}(x)\right]$. F is said to be d_H-continuous at $x_0 \in S$ if for every $\epsilon > 0$, there exists $\delta > 0$ such that $\|x - x_0\| < \delta$ implies that $d_H(F(x), F(x_0)) < \epsilon$.

Definition 2.3 Let $S \subseteq \mathbb{R}^n$ be open and nonempty and $F : S \to \mathcal{K}_C$.

(i) F is said to be Lipschitz continuous of rank $L > 0$ if

$$d_H(F(x_1), F(x_2)) \leq L\|x_1 - x_2\| \ \forall x_1, x_2 \in S.$$

(ii) F is said to be Lipschitz continuous locally near a given point $x \in S$ of rank $L > 0$ if, for some $\epsilon > 0$,

$$d_H(F(x_1), F(x_2)) \leq L\|x_1 - x_2\| \ \forall x_1, x_2 \in N_\epsilon(x).$$

All intervals $A = [\underline{a}, \overline{a}] \in \mathcal{K}_C$ can be expressed through their center-radius representation $A = (a^C; a^R)$ (see Ishibuchi and Tanaka [14] and Stefanini and Arana-Jiménez [21]), where $a^C = \frac{\underline{a}+\overline{a}}{2}, a^R = \frac{\overline{a}-\underline{a}}{2}, \underline{a} = a^C - a^R$ and $\overline{a} = a^C + a^R$. Then, the concepts of limit and continuity of interval-valued functions defined by means of the Pompeiu–Hausdorff metric d_H can be interpreted through their respective versions for real-valued functions (see Stefanini and Bede [22]). Based on the center-radius representation of intervals, the definitions of gH-derivative, gH-gradient, gH-directional derivative and gH-partial derivatives are defined in Stefanini and Arana-Jiménez [21]. In what follows, we state the definitions of such concepts and some of their properties.

Definition 2.4 (Stefanini and Arana Jiménez [21]) Let $F : S \subseteq \mathbb{R}^n \to \mathcal{K}_C$, $F(x) = (f^C(x); f^R(x))$, and let $x_0 \in S$ be such that $x_0 + h \in S$ for all $h \in \mathbb{R}^n$ with $\|h\| < \delta$ for a given $\delta > 0$. F is said to be gH-differentiable at x_0 if there exist vectors $w^C, w^R \in \mathbb{R}^n$ and functions $\epsilon^C, \epsilon^R : \mathbb{R}^n \to \mathbb{R}$ with $\epsilon^C(h) \to 0$ and $\epsilon^R(h) \to 0$, when $h \to 0$, such that, for all $h \neq 0$,

$$f^C(x_0 + h) - f^C(x_0) = \langle h, w^C \rangle + \|h\|\epsilon^C(h)$$

and

$$\left|f^R(x_0 + h) - f^R(x_0)\right| = \left|\langle h, w^R \rangle + \|h\|\epsilon^R(h)\right|.$$

The interval-valued function $D_{gH}F(x_0) : \mathbb{R}^n \to \mathcal{K}_C$ defined, for $h \in \mathbb{R}^n$, by

$$D_{gH}F(x_0)(h) = \left(\langle h, w^C \rangle; \left| \langle h, w^R \rangle \right| \right)$$

is called the total gH-derivative of F at x_0, and $D_{gH}F(x_0)(h)$ is the interval-valued differential of F at x_0 with respect to h.

Definition 2.5 (Stefanini and Arana Jiménez [21]) Let S be a nonempty open subset of $\mathbb{R}^n$ and $F : S \to \mathcal{K}_C$, and let $x_0 \in S$ be such that $x_0 + h \in S$ for all $h \in \mathbb{R}^n$ with $\|h\| < \delta$ for a given $\delta > 0$. The directional generalized Hukuhara derivative (directional gH-derivative, for short) of F at x_0 in the direction $d \in \mathbb{R}^n$ is defined by

$$F'_{gH}(x_0; d) = \lim_{h \to 0^+} \frac{1}{h} \cdot (F(x_0 + hd) \ominus_{gH} F(x_0)),$$

if it exists.

The partial interval-valued gH-derivative of F at $x_0 \in S$ with respect to x_j, $j \in \{1, \dots, n\}$, is defined by

$$\frac{\partial_{gH}F}{\partial x_j}(x_0) = \lim_{h \to 0} \frac{1}{h} \cdot (F(x_0 + he_j) \ominus_{gH} F(x_0)),$$

if it exists. (Above, e_j denotes the j-th canonical vector in $\mathbb{R}^n$.)

Proposition 2.1 (Stefanini and Arana Jiménez [21]) *Let $F : S \to \mathcal{K}_C$ and S be an open subset of $\mathbb{R}^n$. Assume that $F = [\underline{f}, \overline{f}]$ is gH-differentiable at $x_0 \in S$ and suppose that the lateral partial derivatives of $\underline{f}$ and $\overline{f}$ exist at x_0 for a given index $j \in \{1, \dots, n\}$. Denote $\underline{u}_j^- = \frac{\partial \underline{f}}{\partial x_j^-}(x_0)$, $\underline{u}_j^+ = \frac{\partial \underline{f}}{\partial x_j^+}(x_0)$, $\overline{u}_j^- = \frac{\partial \overline{f}}{\partial x_j^-}(x_0)$ and $\overline{u}_j^+ = \frac{\partial \overline{f}}{\partial x_j^+}(x_0)$.*

(i) *If $\underline{u}_j^- = \underline{u}_j^+ =: m_j$ and $\overline{u}_j^- = \overline{u}_j^+ =: n_j$, then*

$$\frac{\partial_{gH}F}{\partial x_j}(x_0) = [m_j \vee n_j] = \left(\frac{m_j + n_j}{2}; \left| \frac{m_j - n_j}{2} \right| \right).$$

(ii) *If $\underline{u}_j^- = \overline{u}_j^+ =: p_j$ and $\overline{u}_j^- = \underline{u}_j^+ =: q_j$, then*

$$\frac{\partial_{gH}F}{\partial x_j}(x_0) = [p_j \vee q_j] = \left(\frac{p_j + q_j}{2}; \left| \frac{p_j - q_j}{2} \right| \right).$$

Theorem 2.2 (Stefanini and Arana Jiménez [21]) *Let $F = (f^C; f^R)$ be an interval-valued function in an open set $S \subseteq \mathbb{R}^n$. If F is gH-differentiable at x_0, then all the*

interval-valued partial gH-derivatives of F exist and, following the notation used in Definition 2.4,

$$\frac{\partial_{gH} F}{\partial x_j}(x_0) = F'_{gH}(x_0; e_j) = \left(w_j^C; \left| w_j^R \right| \right)$$

for $j \in \{1, \ldots, n\}$. Furthermore, all directional gH-derivatives of F exist and

$$F'_{gH}(x_0; d) = \left(\langle w^C, d \rangle; \left| \langle w^R, d \rangle \right| \right).$$

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ and $x_0 \in S$, where S is open, be such that every lateral partial derivative of f exists at x_0. We will use the following notation:

$$(\nabla f)_- (x_0) = \left(\frac{\partial f}{\partial x_1^-}(x_0), \ldots, \frac{\partial f}{\partial x_n^-}(x_0) \right)$$

$$(\nabla f)_+ (x_0) = \left(\frac{\partial f}{\partial x_1^+}(x_0), \ldots, \frac{\partial f}{\partial x_n^+}(x_0) \right).$$

Theorem 2.3 *Let $S \subseteq \mathbb{R}^n$ be an open and nonempty set and $F : S \to \mathcal{K}_C$, with $F(x) = \left[\underline{f}(x), \overline{f}(x) \right]$ for all $x \in S$. If F is gH-differentiable at $x_0 \in S$, then for all $d \in \mathbb{R}^n$, one of the following cases holds:*

(i) $\nabla \underline{f}(x_0)$ *and* $\nabla \overline{f}(x_0)$ *exist, and*

$$F'_{gH}(x_0; d) = \left[\left\langle \nabla \underline{f}(x_0), d \right\rangle \vee \left\langle \nabla \overline{f}(x_0), d \right\rangle \right].$$

Particularly,

$$\frac{\partial_{gH} F}{\partial x_i}(x_0) = \left[\frac{\partial \underline{f}}{\partial x_i}(x_0) \vee \frac{\partial \overline{f}}{\partial x_i}(x_0) \right], \quad i \in \{1, \ldots, n\}.$$

(ii) $\left(\nabla \underline{f} \right)_- (x_0), (\nabla \overline{f})_- (x_0), \left(\nabla \underline{f} \right)_+ (x_0), (\nabla \overline{f})_+ (x_0)$ *exist, satisfy*

$$\left\langle \left(\nabla \underline{f} \right)_- (x_0), d \right\rangle = \left\langle (\nabla \overline{f})_+ (x_0), d \right\rangle,$$

$$\left\langle (\nabla \overline{f})_- (x_0), d \right\rangle = \left\langle \left(\nabla \underline{f} \right)_+ (x_0), d \right\rangle,$$

and

$$F'_{gH}(x_0; d) = \left[\left\langle \left(\nabla \underline{f} \right)_- (x_0), d \right\rangle \vee \left\langle (\nabla \overline{f})_- (x_0), d \right\rangle \right]$$

$$= \left[\left\langle \left(\nabla \underline{f} \right)_+ (x_0), d \right\rangle \vee \left\langle (\nabla \overline{f})_+ (x_0), d \right\rangle \right].$$

Particularly,

$$\frac{\partial_{gH} F}{\partial x_i}(x_0) = \left[\frac{\partial \underline{f}}{\partial x_i^-}(x_0) \vee \frac{\partial \overline{f}}{\partial x_i^-}(x_0)\right] = \left[\frac{\partial \underline{f}}{\partial x_i^+}(x_0) \vee \frac{\partial \overline{f}}{\partial x_i^+}(x_0)\right], \ i \in \{1, \ldots, n\}.$$

Proof It follows directly from Proposition 2.1 and Theorem 2.2. □

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ and $x_0 \in S$, where S is open, be such that every lateral partial derivative of f exist at x_0. The following notation will be used later in the paper:

$$\nabla_\sharp f(x^*) = \begin{cases} \nabla f(x^*), & \text{if it exists,} \\ \nabla f_-(x^*), & \text{otherwise,} \end{cases}$$

$$\frac{\partial f}{\partial x_i^\sharp}(x^*) = \begin{cases} \frac{\partial f}{\partial x_i}(x^*), & \text{if it exists,} \\ \frac{\partial f}{\partial x_i^-}(x^*), & \text{otherwise.} \end{cases}$$

Given $F : S \subseteq \mathbb{R}^n \to \mathcal{K}_C$, with $F(x) = \left[\underline{f}(x), \overline{f}(x)\right] \forall x \in S$, and $x_0 \in S$, when F is gH-differentiable at x_0, we will denote:

$$\frac{\partial_{gH} F}{\partial x_i}(x^*) = \begin{cases} \left[\frac{\partial \underline{f}}{\partial x_i}(x^*) \vee \frac{\partial \overline{f}}{\partial x_i}(x^*)\right], & \text{if they exist,} \\ \left[\frac{\partial \underline{f}}{\partial x_i^-}(x^*) \vee \frac{\partial \overline{f}}{\partial x_i^-}(x^*)\right] = \left[\frac{\partial \underline{f}}{\partial x_i^+}(x_0) \vee \frac{\partial \overline{f}}{\partial x_i^+}(x_0)\right], & \text{otherwise} \end{cases}$$

$$= \left[\frac{\partial \underline{f}}{\partial x_i^\sharp}(x^*) \vee \frac{\partial \overline{f}}{\partial x_i^\sharp}(x^*)\right]$$

and

$$\nabla_{gH} F(x^*) = \left(\frac{\partial_{gH} F}{\partial x_1}(x^*), \ldots, \frac{\partial_{gH} F}{\partial x_n}(x^*)\right).$$

Next, we reproduce some key definitions and results from Girsanov [12] related to the Dubovitskii–Milyutin formalism.

Definition 2.6 Let E be a topological linear space. A nonzero continuous linear functional f is said to be a supporting functional for a set $A \subset E$ at $x_0 \in A$ if $f(x) \geq f(x_0)$ for all $x \in A$.

Definition 2.7 Let E be a topological linear space and E' its topological dual space.

(i) A set $\mathcal{K}$ in E is called a cone with apex at 0 if $x \in \mathcal{K}$ implies that $\lambda x \in \mathcal{K}$ for all $\lambda > 0$.
(ii) If $\mathcal{K}$ is a cone in E, its dual cone, denoted by $\mathcal{K}^*$, is defined as the set of all continuous linear functionals which are nonnegative on $\mathcal{K}$, that is,

$$\mathcal{K}^* = \{f \in E' : f(x) \geq 0 \ \forall x \in \mathcal{K}\}.$$

Lemma 2.2 (Dubovitskii–Milyutin, see Girsanov [12]) *Let $\mathcal{K}_1, \ldots, \mathcal{K}_l, \mathcal{K}_{l+1}$ be convex cones with apex at the origin, where $\mathcal{K}_1, \ldots, \mathcal{K}_l$ are open. Then, $\bigcap_{i=1}^{l+1} \mathcal{K}_i = \emptyset$ if, and only if, there exist linear functionals $f_i \in \mathcal{K}_i^*$, $i \in \{1, \ldots, l+1\}$, not all zero, such that $f_1 + \ldots + f_l + f_{l+1} = 0$.*

Definition 2.8 Let E be a topological linear space and $Q \subseteq E$.

(i) A vector d is a feasible direction for Q at a point x_0 if there exists a neighborhood U of d and a positive scalar ϵ_0 such that $x_0 + \epsilon \tilde{d} \in Q$ for all $\epsilon \in (0, \epsilon_0)$ and all $\tilde{d} \in U$.
The set of all feasible directions will denoted by $\mathcal{K}_b$.
(ii) A vector d is a tangent direction to Q at a point x_0 if there exist a positive scalar ϵ_0 and a map $r : (0, \epsilon_0) \rightarrow E$ such that $x(\epsilon) := x_0 + \epsilon d + r(\epsilon) \in Q$ for all $\epsilon \in (0, \epsilon_0)$, where for any neighborhood U of zero, $\frac{1}{\epsilon} r(\epsilon) \in U$ for all small $\epsilon > 0$ (in a Banach space this is replaced by the simpler condition $\|r(\epsilon)\| = o(\epsilon)$).
The set of all tangent directions will be denoted by $\mathcal{K}_k$.

We keep the classical notation $\mathcal{K}_b$ and $\mathcal{K}_k$ (Russian the initials of "feasible cone" and "tangent cone").

Proposition 2.2 (Girsanov [12]) *The set of all feasible directions $\mathcal{K}_b$ is an open cone with apex at zero.*

Proposition 2.3 (Girsanov [12]) *The set of all tangent directions $\mathcal{K}_k$ is a cone with apex at zero.*

Theorem 2.4 (Girsanov [12]) *Let E be a Banach space and $f : E \rightarrow \mathbb{R}$ be Fréchet differentiable at $x_0 \in E$. Denote $Q = \{x : f(x) \leq f(x_0)\}$. If $f'(x_0) \neq 0$, then*

$$\mathcal{K}_b = \{d : f'(x_0; d) < 0\}.$$

Theorem 2.5 (Girsanov [12]) *Let Q be a closed convex set. If $\mathrm{int}(Q) \neq \emptyset$, then*

$$\begin{aligned} \mathcal{K}_b &= \{\lambda[\mathrm{int}(Q) - x_0] : \lambda > 0\} \\ &= \{d : d = \lambda(x - x_0), \ x \in \mathrm{int}(Q), \ \lambda > 0\}. \end{aligned}$$

Theorem 2.6 (Lyusternik, see Girsanov [12]) *Let E_1 and E_2 be Banach spaces, $P : E_1 \rightarrow E_2$ be a Fréchet differentiable operator in a neighborhood of a point x_0 with $P(x_0) = 0$. Let $P'(x)$ be continuous in a neighborhood of x_0 and suppose that $P'(x_0)$ maps E_1 onto E_2. Then,*

$$\mathcal{K}_k = \{d : P'(x_0)h = 0\}.$$

Theorem 2.7 (Girsanov [12]) *Let $f \in E'$ and denote*

$$\mathcal{K}_I = \{x : f(x) = 0\}, \quad \mathcal{K}_{II} = \{x : f(x) \geq 0\} \quad and \quad \mathcal{K}_{III} = \{x : f(x) > 0\}.$$

Then,

$$\mathcal{K}_I^* = \{\lambda f : \lambda \in \mathbb{R}\}, \quad \mathcal{K}_{II}^* = \{\lambda f : 0 \leq \lambda < \infty\}, \quad \mathcal{K}_{III}^* = \begin{cases} E', & \text{if } f = 0, \\ \mathcal{K}_{II}^*, & \text{if } f \neq 0. \end{cases}$$

Theorem 2.8 (Girsanov [12]) *Let A be a $k \times l$ matrix and $K = \{x \in \mathbb{R}^l : Ax = 0\}$. Then, $K^* = \{A^T y : y \in \mathbb{R}^k\}$.*

Theorem 2.9 (Girsanov [12]) *Let Q be a closed convex set. If $\mathrm{int}(Q) \neq \emptyset$, then $\mathcal{K}_b^* = Q^* = \{f \in E' : f(x) \geq f(x_0) \; \forall x \in Q\}$.*

2.2 The Interval Optimization Problem

We consider the following interval-valued optimization problem with functional and abstract constraints:

$$\begin{aligned}
&\text{minimize } F(x) = [\underline{f}(x), \overline{f}(x)] \\
&\text{subject to } g_i(x) \leq 0, \; i \in I = \{1, \ldots, p\}, \\
&\qquad\qquad h_j(x) = 0, \; j \in J = \{1, \ldots, q\}, \\
&\qquad\qquad x \in S \subseteq \mathbb{R}^n,
\end{aligned} \qquad (\text{IOP})$$

where $F = [\underline{f}, \overline{f}] : \mathbb{R}^n \to \mathcal{K}_C$, $\underline{f}, \overline{f} : \mathbb{R}^n \to \mathbb{R}$, $\underline{f}(x) \leq \overline{f}(x)$ for all $x \in \mathbb{R}^n$, $g_i : \mathbb{R}^n \to \mathbb{R}, i \in I, h_j : \mathbb{R}^n \to \mathbb{R}, j \in J$, and S is a nonempty set in $\mathbb{R}^n$.

The set of all feasible solutions to (IOP) will be denoted by $\mathcal{F}$, that is,

$$\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, \; i \in I, \; h_j(x) = 0, \; j \in J, \; x \in S\}.$$

The set of indices related to the active constraints at a feasible point $x^* \in \mathcal{F}$ will be denoted by $I(x^*)$:

$$I(x^*) = \{i \in I : g_i(x^*) = 0\}.$$

Next, we give the concepts of solutions that will be used for the problem (IOP). They are based in the order relation LU.

Definition 2.9 (*See* Wu [26]) Let $x^* \in S$. Then,

(i) x^* is said to be a *local LU-solution* of (IOP) if there exists $\epsilon > 0$ such that there does not exist $x \in N_\epsilon(x^*) \cap \mathcal{F}$ with $F(x) \leq_{LU} F(x^*)$.
(ii) x^* is said to be a *local weak LU-solution* of (IOP) if there exists $\epsilon > 0$ such that there does not exist $x \in N_\epsilon(x^*) \cap \mathcal{F}$ with $F(x) <_{LU} F(x^*)$.

If $N_\epsilon(x^*)$ above is replaced by the whole space $\mathbb{R}^n$, we have the corresponding definitions of global solutions for (IOP).

3 Necessary Optimality Conditions

In this section, necessary optimality conditions for (IOP) will be established through an adaptation of the Dubovitskii–Milyutin formalism. The cones of feasible and tangent directions and the respective duals can be found in Girsanov [12]. We define an appropriate cone of directions of decrease for interval-valued functions and give it a characterization by means of the gH-derivative of the function. A geometric necessary optimality condition involving the mentioned cones is derived. Algebraic necessary optimality conditions are obtained after an application of the Dubovitskii–Milyutin Lemma (see Lemma 5.11 in [12]).

The following assumptions will be required. Let $x^* \in \mathcal{F}$.

(A1) $F(x)$ is Lipschitz continuous locally near x^*.
(A2) F is continuously gH-differentiable in a neighborhood of x^*.
(A3) $g_i, i \in I, h_j, j \in J$, are continuously differentiable in a neighborhood of x^*.
(A4) S is closed and convex, and $\mathrm{int}(S) \neq \emptyset$.

Now, we define the directions of decrease for the interval-valued function F. Then, some of its properties will be developed.

Definition 3.1 A direction $d \in \mathbb{R}^n$ is said to be a *LU-direction of decrease* of F at $x^* \in \mathcal{F}$ if there exist a neighborhood N of d, a positive scalar ϵ^* and an interval $[\underline{\alpha}, \overline{\alpha}] <_{LU} [0, 0]$ such that

$$F(x^* + \epsilon \tilde{d}) <_{LU} F(x^*) + \epsilon[\underline{\alpha}, \overline{\alpha}] \ \forall \epsilon \in (0, \epsilon^*), \ \forall \tilde{d} \in N.$$

The set of all LU-directions of decrease of F at x^* will be denoted by $\mathcal{K}_y^{LU}$.

It is important to emphasize that in Definition 3.1, both $\underline{\alpha}$ and $\overline{\alpha}$ depend on F, x^*, and $\bar{x}$.

Proposition 3.1 *The set of all LU-directions of decrease $\mathcal{K}_y^{LU}$ is an open convex cone.*

The proof of Proposition 3.1 will be postponed to the end of the next subsection. Next, we have a geometric characterization of weak LU-solutions.

Theorem 3.1 *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP). Then,*

$$\mathcal{K}_y^{LU} \cap \mathcal{K}_b \cap \mathcal{K}_k = \emptyset.$$

Proof Suppose that there exists $d \in \mathcal{K}_y^{LU} \cap \mathcal{K}_b \cap \mathcal{K}_y$. Then, by Definition 3.1, there exists a neighborhood N_0 of d, a positive scalar ϵ_0 and an interval $[\underline{\alpha}, \overline{\alpha}] <_{LU} [0, 0]$ such that

$$F(x^* + \epsilon \tilde{d}) <_{LU} F(x^*) + \epsilon [\underline{\alpha}, \overline{\alpha}] \ \forall \epsilon \in (0, \epsilon_0), \ \forall \tilde{d} \in N_0. \tag{1}$$

By Definition 2.8-(i), it follows that there exists a neighborhood N_1 of d and a scalar $\epsilon_1 > 0$ such that
$$x^* + \epsilon \tilde{d} \in \mathcal{F}_1 \ \forall \epsilon \in (0, \epsilon_1), \ \forall \tilde{d} \in N_1, \tag{2}$$

where $\mathcal{F}_1 = \{x \in \mathbb{R}^n : g_i(x) \leq 0, \ i \in I, \ x \in S\}$. We also have, by Definition 2.8-(ii), that there exist a scalar $\epsilon_2 > 0$ and a map $r : (0, \epsilon_2) \to \mathbb{R}^n$ such that

$$x^* + \epsilon d + r(\epsilon) \in \mathcal{F}_2 \ \forall \epsilon \in (0, \epsilon_2), \tag{3}$$

where $\|r(\epsilon)\| = o(\epsilon)$ and $\mathcal{F}_2 = \{x \in \mathbb{R}^n : h_j(x) = 0, \ j \in J\}$. Let $d_\epsilon := d + \frac{1}{\epsilon} r(\epsilon)$, $\epsilon \in (0, \epsilon_3)$, where $\epsilon_3 > 0$ is small enough so that $d_\epsilon \in N := N_0 \cap N_1, \epsilon \in (0, \epsilon_3)$. Provided x^* is a local weak LU-solution of (IOP), the exists $\epsilon_4 > 0$ such that there does not exist $x \in N_{\epsilon_4}(x^*) \cap \mathcal{F}$ with $F(x) \leq_{LU} F(x^*)$. If $\epsilon_5 := \min\{\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4\}$, it follows, by (2)–(3), that

$$x_\epsilon := x^* + \epsilon d_\epsilon \in N_{\epsilon_5}(x^*) \cap \mathcal{F}_1 \cap \mathcal{F}_2 = N_{\epsilon_5}(x^*) \cap \mathcal{F} \ \forall \epsilon \in (0, \epsilon_5).$$

Moreover, by (1),

$$F(x_\epsilon) = F(x^* + \epsilon d_\epsilon) <_{LU} F(x^*) + \epsilon \left[\underline{\alpha}, \overline{\alpha}\right] <_{LU} F(x^*) \ \forall \epsilon \in (0, \epsilon_5).$$

Therefore, $x_\epsilon \in N_{\epsilon_5}(x^*) \cap \mathcal{F}$ satisfies $F(x_\epsilon) <_{LU} F(x^*)$, which is a contradiction to the assumption that $x^* \in \mathcal{F}$ is a local weak LU-solution. $\square$

Characterization of the Cone of LU-Directions of Decrease

In order to transcribe the necessary optimality condition given in geometric form in Theorem 3.1 into an algebraic one, it is necessary to characterize the cones of LU-directions of decrease and of feasible and tangent directions. The characterizations of the cones of feasible and tangent directions are already known in the literature and we have already stated them in Theorems 2.4, 2.5, and 2.6. So, we proceed to characterize $\mathcal{K}_y^{LU}$.

Let $x^* \in S$. In order to characterize the cone of LU-directions of decrease in terms of the gH-derivative of F, we define

$$\mathcal{K}_0 = \left\{d \in \mathbb{R}^n : F'_{gH}(x^*; d) <_{LU} [0, 0]\right\}.$$

Theorem 3.2 *Assume that (A2) holds. Then, $\mathcal{K}_y^{LU} \subset \mathcal{K}_0$.*

Proof Let $d \in \mathcal{K}_y^{LU}$. Then, there exists a neighborhood N of d, a positive scalar ϵ^* and an interval $[\underline{\alpha}, \overline{\alpha}] <_{LU} [0, 0]$ such that

$$F(x^* + \epsilon \tilde{d}) <_{LU} F(x^*) + \epsilon[\underline{\alpha}, \overline{\alpha}] \ \forall \epsilon \in (0, \epsilon^*), \ \forall \tilde{d} \in N. \tag{4}$$

We have that

$$\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \cdot (F(x^* + \epsilon d) \ominus_{gH} F(x^*)) <_{LU} [0, 0].$$

Indeed, let us set, for $\epsilon \in (0, \epsilon^*)$,

$$A_\epsilon := [\underline{a}_\epsilon, \overline{a}_\epsilon] = F(x^* + \epsilon d),$$
$$B_\epsilon := [\underline{b}_\epsilon, \overline{b}_\epsilon] = F(x^*),$$
$$C_\epsilon := [\underline{c}_\epsilon, \overline{c}_\epsilon] = A_\epsilon \ominus_{gH} B_\epsilon = F(x^* + \epsilon d) \ominus_{gH} F(x^*).$$

Since $A_\epsilon \ominus_{gH} B_\epsilon = C_\epsilon$, we have that

$$\begin{cases} A_\epsilon = B_\epsilon + C_\epsilon \\ \text{or} \\ B_\epsilon = A_\epsilon - C_\epsilon \end{cases} \Leftrightarrow \begin{cases} \underline{c}_\epsilon = \underline{a}_\epsilon - \underline{b}_\epsilon \ \text{ and } \ \overline{c}_\epsilon = \overline{a}_\epsilon - \overline{b}_\epsilon \\ \text{or} \\ \overline{c}_\epsilon = \underline{a}_\epsilon - \underline{b}_\epsilon \ \text{ and } \ \underline{c}_\epsilon = \overline{a}_\epsilon - \overline{b}_\epsilon. \end{cases} \tag{5}$$

By (4), we have that

$$A_\epsilon <_{LU} B_\epsilon + \epsilon[\underline{\alpha}, \overline{\alpha}] \Leftrightarrow [\underline{a}_\epsilon, \overline{a}_\epsilon] <_{LU} [\underline{b}_\epsilon + \epsilon\underline{\alpha}, \overline{b}_\epsilon + \epsilon\overline{\alpha}],$$

so that

$$\frac{1}{\epsilon}(\underline{a}_\epsilon - \underline{b}_\epsilon) < \underline{\alpha} \ \text{ and } \ \frac{1}{\epsilon}(\overline{a}_\epsilon - \overline{b}_\epsilon) < \overline{\alpha}.$$

Let us assume that the first case in (5) holds. If the second case occurs, we proceed similarly. Then,

$$\lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\underline{c}_\epsilon < \underline{\alpha} \ \text{ and } \ \lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\overline{c}_\epsilon < \overline{\alpha},$$

that is,

$$\left[\lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\underline{c}_\epsilon, \ \lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\overline{c}_\epsilon\right] <_{LU} [\underline{\alpha}, \overline{\alpha}] <_{LU} [0, 0].$$

It follows from Theorem 2.1 that

$$\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \cdot C_\epsilon <_{LU} [0, 0].$$

By Definition 2.5, $F'_{gH}(x^*; d) <_{LU} [0, 0]$, that is, $d \in \mathcal{K}_0$.   □

Now, we establish the converse of Theorem 3.2:

Theorem 3.3 *Suppose that (A1)–(A2) holds. Then, $\mathcal{K}_0 \subset \mathcal{K}_y^{LU}$.*

Proof Let $d \in \mathcal{K}_0$. By Definition 2.5,

$$F'_{gH}(x^*; d) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \cdot (F(x^* + \epsilon d) \ominus_{gH} F(x^*)) =: [\underline{\delta}, \overline{\delta}] \in \mathcal{K}_C,$$

where $\underline{\delta} \leq \overline{\delta}$. Then, provided $d \in \mathcal{K}_0$,

$$[0, 0] >_{LU} F'_{gH}(x^*; d) = \left[\underline{\delta}, \overline{\delta}\right],$$

which implies that $\left[\underline{\delta}, \overline{\delta}\right] <_{LU} [0, 0]$, that is, $\underline{\delta} < 0$ and $\overline{\delta} < 0$. By definition of limit, given $\epsilon' \in (0, -\overline{\delta}/2)$, there exists $\delta' > 0$ such that

$$d_H \left(\frac{1}{\epsilon} \cdot (F(x^* + \epsilon d) \ominus_{gH} F(x^*)), [\underline{\delta}, \overline{\delta}] \right) < \epsilon'$$

for all $\epsilon \in (0, \delta')$. By setting

$$\frac{1}{\epsilon} \cdot (F(x^* + \epsilon d) \ominus_{gH} F(x^*)) =: K_\epsilon = \left[\underline{k}_\epsilon, \overline{k}_\epsilon\right],$$

where $\epsilon \in (0, \delta')$, we have that

$$\max\{|\underline{k}_\epsilon - \underline{\delta}|, |\overline{k}_\epsilon - \overline{\delta}|\} = d_H(K_\epsilon, [\underline{\delta}, \overline{\delta}]) < \epsilon'$$

$$\Rightarrow \underline{k}_\epsilon - \underline{\delta} \leq |\underline{k}_\epsilon - \underline{\delta}| < \epsilon' \quad \text{and} \quad \overline{k}_\epsilon - \overline{\delta} \leq |\overline{k}_\epsilon - \overline{\delta}| < \epsilon'$$

$$\Rightarrow \underline{k}_\epsilon < \epsilon' + \underline{\delta} < -\frac{\overline{\delta}}{2} + \underline{\delta} < -\frac{\delta}{2} + \underline{\delta} = \frac{\delta}{2} \quad \text{and} \quad \overline{k}_\epsilon < \epsilon' + \overline{\delta} < -\frac{\overline{\delta}}{2} + \overline{\delta} = \frac{\overline{\delta}}{2}$$

$$\Rightarrow K_\epsilon = \left[\underline{k}_\epsilon, \overline{k}_\epsilon\right] <_{LU} \left[\frac{\underline{\delta}}{2}, \frac{\overline{\delta}}{2}\right]$$

$$\Rightarrow F(x^* + \epsilon d) \ominus_{gH} F(x^*) <_{LU} \epsilon \left[\frac{\underline{\delta}}{2}, \frac{\overline{\delta}}{2}\right] \quad \forall \epsilon \in (0, \delta'). \tag{6}$$

Since $F(x)$ satisfies the Lipschitz continuous condition in a neighborhood of x^*, there exist $L > 0$ and $\epsilon^* > 0$ such that

$$d_H(F(x_1), F(x_2)) \leq L\|x_1 - x_2\| \tag{7}$$

for all x_1 and x_2 in which $\|x_1 - x^*\| < \epsilon^*$ and $\|x_2 - x^*\| < \epsilon^*$.

Let $\tilde{d} \in \mathbb{R}^n$ be arbitrary but such that

$$\|\tilde{d} - d\| < -\frac{\overline{\delta}}{4L}. \tag{8}$$

Whenever $\epsilon \in (0, \hat{\epsilon})$, where $\hat{\epsilon} < \min \left\{ \frac{4\epsilon^* L}{-\overline{\delta} + 4\|d\|L}, \frac{\epsilon^*}{\|d\|} \right\}$, we obtain that

$$\|(x^* + \epsilon\tilde{d}) - x^*\| \leq \epsilon(\|\tilde{d} - d\| + \|d\|) < \frac{4\epsilon^* L}{-\overline{\delta} + 4|d\|L} \left(-\frac{\overline{\delta}}{4L} + \|d\| \right) = \epsilon^*$$

and

$$\|(x^* + \epsilon d) - x^*\| = \epsilon \|d\| < \frac{\epsilon^*}{\|d\|} \|d\| = \epsilon^*.$$

Hence, it follows from (7) that

$$d_H(F(x^* + \epsilon \tilde{d}), F(x^* + \epsilon d)) \le L\|(x^* + \epsilon \tilde{d}) - (x^* + \epsilon d)\|$$

for all $\epsilon \in (0, \hat{\epsilon})$. Then, by Lemma 2.1 and (8),

$$
\begin{aligned}
F(x^* + \epsilon \tilde{d}) \ominus_{gH} F(x^* + \epsilon d) \\
\le_{LU} [d_H(F(x^* + \epsilon \tilde{d}), F(x^* + \epsilon d)), d_H(F(x^* + \epsilon \tilde{d}), F(x^* + \epsilon d))] \\
\le_{LU} [L\|(x^* + \epsilon \tilde{d}) - (x^* + \epsilon d)\|, L\|(x^* + \epsilon \tilde{d}) - (x^* + \epsilon d)\|] \\
<_{LU} \left[-L\epsilon \frac{\overline{\delta}}{4L}, -L\epsilon \frac{\overline{\delta}}{4L} \right] \\
= \epsilon \left[-\frac{\overline{\delta}}{4}, -\frac{\overline{\delta}}{4} \right] \quad \forall \epsilon \in (0, \hat{\epsilon}),
\end{aligned}
$$

i.e.,

$$F(x^* + \epsilon \tilde{d}) \ominus_{gH} F(x^* + \epsilon d) <_{LU} \epsilon \left[-\frac{\overline{\delta}}{4}, -\frac{\overline{\delta}}{4} \right] \quad \forall \epsilon \in (0, \hat{\epsilon}). \qquad (9)$$

Let $\tilde{\epsilon} = \min \{\delta', \hat{\epsilon}\}$ and $\underline{\alpha}' = \overline{\alpha}' = \frac{\overline{\delta}}{4}$. Given $0 < \epsilon < \tilde{\epsilon}$, we have

$$F(x^* + \epsilon \tilde{d}) <_{LU} F(x^*) + \epsilon [\underline{\alpha}', \overline{\alpha}'].$$

Indeed, set, for $\epsilon \in (0, \tilde{\epsilon})$,

$$
\begin{aligned}
A_\epsilon &:= [\underline{a}_\epsilon, \overline{a}_\epsilon] = F(x^* + \epsilon d), \\
B_\epsilon &:= [\underline{b}_\epsilon, \overline{b}_\epsilon] = F(x^*), \\
C_\epsilon &:= [\underline{c}_\epsilon, \overline{c}_\epsilon] = A_\epsilon \ominus_{gH} B_\epsilon = F(x^* + \epsilon d) \ominus_{gH} F(x^*), \\
D_\epsilon &:= [\underline{d}_\epsilon, \overline{d}_\epsilon] = F(x^* + \epsilon \tilde{d}), \\
E_\epsilon &:= [\underline{e}_\epsilon, \overline{e}_\epsilon] = D_\epsilon \ominus_{gH} A_\epsilon = F(x^* + \epsilon \tilde{d}) \ominus_{gH} F(x^* + \epsilon d).
\end{aligned}
$$

Since $A_\epsilon \ominus_{gH} B_\epsilon = C_\epsilon$, by (6), we have that

$$
\begin{cases}
A_\epsilon = B_\epsilon + C_\epsilon \\
\text{or} \\
B_\epsilon = A_\epsilon - C_\epsilon, \\
\text{with} \\
C_\epsilon <_{LU} \left[\frac{\epsilon\underline{\delta}}{2}, \frac{\epsilon\overline{\delta}}{2} \right]
\end{cases}
\Leftrightarrow
\begin{cases}
[\underline{a}_\epsilon, \overline{a}_\epsilon] = [\underline{b}_\epsilon, \overline{b}_\epsilon] + [\underline{c}_\epsilon, \overline{c}_\epsilon] \\
\text{or} \\
[\underline{b}_\epsilon, \overline{b}_\epsilon] = [\underline{a}_\epsilon, \overline{a}_\epsilon] + [-\overline{c}_\epsilon, -\underline{c}_\epsilon], \\
\text{with} \\
[\underline{c}_\epsilon, \overline{c}_\epsilon] <_{LU} \left[\frac{\epsilon\underline{\delta}}{2}, \frac{\epsilon\overline{\delta}}{2} \right]
\end{cases}
$$

and since $D_\epsilon \ominus_{gH} A_\epsilon = E_\epsilon$, by (9), we have that

$$
\begin{cases}
D_\epsilon = A_\epsilon + E_\epsilon \\
\text{or} \\
A_\epsilon = D_\epsilon - E_\epsilon, \\
\text{with} \\
E_\epsilon <_{LU} \left[-\frac{\epsilon\bar{\delta}}{4}, -\frac{\epsilon\bar{\delta}}{4}\right]
\end{cases}
\Leftrightarrow
\begin{cases}
[\underline{d}_\epsilon, \bar{d}_\epsilon] = [\underline{a}_\epsilon, \bar{a}_\epsilon] + [\underline{e}_\epsilon, \bar{e}_\epsilon] \\
\text{or} \\
[\underline{a}_\epsilon, \bar{a}_\epsilon] = [\underline{d}_\epsilon, \bar{d}_\epsilon] + [-\bar{e}_\epsilon, -\underline{e}_\epsilon], \\
\text{with} \\
[\underline{e}_\epsilon, \bar{e}_\epsilon] <_{LU} \left[-\frac{\epsilon\bar{\delta}}{4}, -\frac{\epsilon\bar{\delta}}{4}\right].
\end{cases}
$$

We have one of the following cases:

(i) $\left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] = \left[\underline{b}_\epsilon, \bar{b}_\epsilon\right] + \left[\underline{c}_\epsilon, \bar{c}_\epsilon\right]$ and $\left[\underline{d}_\epsilon, \bar{d}_\epsilon\right] = \left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] + \left[\underline{e}_\epsilon, \bar{e}_\epsilon\right]$;
(ii) $\left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] = \left[\underline{b}_\epsilon, \bar{b}_\epsilon\right] + \left[\underline{c}_\epsilon, \bar{c}_\epsilon\right]$ and $\left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] = \left[\underline{d}_\epsilon, \bar{d}_\epsilon\right] + \left[-\bar{e}_\epsilon, -\underline{e}_\epsilon\right]$;
(iii) $\left[\underline{b}_\epsilon, \bar{b}_\epsilon\right] = \left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] + \left[-\bar{c}_\epsilon, -\underline{c}_\epsilon\right]$ and $\left[\underline{d}_\epsilon, \bar{d}_\epsilon\right] = \left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] + \left[\underline{e}_\epsilon, \bar{e}_\epsilon\right]$;
(iv) $\left[\underline{b}_\epsilon, \bar{b}_\epsilon\right] = \left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] + \left[-\bar{c}_\epsilon, -\underline{c}_\epsilon\right]$ and $\left[\underline{a}_\epsilon, \bar{a}_\epsilon\right] = \left[\underline{d}_\epsilon, \bar{d}_\epsilon\right] + \left[-\bar{e}_\epsilon, -\underline{e}_\epsilon\right]$;

always with

$$
\left[\underline{c}_\epsilon, \bar{c}_\epsilon\right] <_{LU} \left[\frac{\epsilon\delta}{2}, \frac{\epsilon\bar{\delta}}{2}\right] \quad \text{and} \quad \left[\underline{e}_\epsilon, \bar{e}_\epsilon\right] <_{LU} \left[-\frac{\epsilon\bar{\delta}}{4}, -\frac{\epsilon\bar{\delta}}{4}\right].
$$

In each of the cases (i)–(iv) above we have $\left[\underline{d}_\epsilon, \bar{d}_\epsilon\right] <_{LU} \left[\underline{b}_\epsilon, \bar{b}_\epsilon\right] + \epsilon\left[\underline{\alpha}', \bar{\alpha}'\right]$, with $\underline{\alpha}' = \bar{\alpha}' = \frac{\bar{\delta}}{4}$, that is, $D_\epsilon <_{LU} B_\epsilon + \epsilon\left[\underline{\alpha}', \bar{\alpha}'\right]$, which is equivalent to

$$
F(x^* + \epsilon\tilde{d}) <_{LU} F(x^*) + \epsilon\left[\underline{\alpha}', \bar{\alpha}'\right],
$$

with $\left[\underline{\alpha}', \bar{\alpha}'\right] = \left[\frac{\bar{\delta}}{4}, \frac{\bar{\delta}}{4}\right] <_{LU} [0, 0]$, since $\bar{\delta} < 0$. It follows from Definition 3.1 that $d \in \mathcal{K}_y^{LU}$. $\qquad\square$

By Theorems 3.2 and 3.3, we have:

Theorem 3.4 *Suppose that (A1)–(A2) holds. Then, the cone of LU-directions of decrease $\mathcal{K}_y^{LU}$ coincides with $\mathcal{K}_0$.*

If (A2) holds, by Theorem 2.3, we have

$$
F'_{gH}(x^*; d) = \left[\left\langle \nabla_\sharp \underline{f}(x^*), d\right\rangle \vee \left\langle \nabla_\sharp \bar{f}(x^*), d\right\rangle\right], \, d \in \mathbb{R}^n,
$$

so that we can write

$$
\mathcal{K}_0 = \left\{d \in \mathbb{R}^n : \left[\left\langle \nabla_\sharp \underline{f}(x^*), d\right\rangle \vee \left\langle \nabla_\sharp \bar{f}(x^*), d\right\rangle\right] <_{LU} [0, 0]\right\}.
$$

Now, let us set

$$
\underline{\mathcal{K}}_0 := \left\{d \in \mathbb{R}^n : \left\langle \nabla_\sharp \underline{f}(x^*), d\right\rangle < 0\right\}, \tag{10}
$$

$$\overline{\mathcal{K}}_0 := \left\{ d \in \mathbb{R}^n : \langle \nabla_\sharp \overline{f}(x^*), d \rangle < 0 \right\}. \tag{11}$$

It is clear that $\mathcal{K}_0 = \underline{\mathcal{K}}_0 \cap \overline{\mathcal{K}}_0$.

Proposition 3.2 *Assume that (A2) holds. Then,*

(i) $\underline{\mathcal{K}}_0$ *is an open convex cone;*
(ii) $\overline{\mathcal{K}}_0$ *is an open convex cone.*

Proof The proof is straightforward and has been omitted. □

We end this subsection by providing the proof of Proposition 3.1.

Proof of Proposition 3.1 It follows directly from Theorem 3.4 and Proposition 3.2. □

Characterization of the Cone of Feasible Directions

As already mentioned, the characterization of the cone of feasible directions can be obtained by applying Theorems 2.4 and 2.5. Here, we assume that (A3)–(A4) holds.

For each $i \in I$, we set $\mathcal{F}_{1i} = \{x \in \mathbb{R}^n : g_i(x) \leq 0\}$ and denote by $\mathcal{K}_{bi}$ the cone of feasible directions for $\mathcal{F}_{1i}$ at x^*. For $i \in I(x^*)$, by Theorem 2.4, if $\nabla g_i(x^*) \neq 0$,

$$\mathcal{K}_{bi} = \{d \in \mathbb{R}^n : \langle \nabla g_i(x^*), d \rangle < 0\}. \tag{12}$$

If $g_i(x^*) < 0$, then clearly $\mathcal{K}_{bi} = \mathbb{R}^n$. If $\mathcal{F}_{1s} = \{x \in \mathbb{R}^n : x \in S\}$ and $\mathcal{K}_{bs}$ denotes the cone of feasible directions for $\mathcal{F}_{1s}$ at x^*, by Theorem 2.5,

$$\mathcal{K}_{bs} = \left\{ d \in \mathbb{R}^n : d = \lambda(x - x^*), \ x \in \text{int}(S), \ \lambda > 0 \right\}. \tag{13}$$

It is easy to see that

$$\mathcal{F} = \left(\bigcap_{i \in I} \mathcal{F}_{1i} \right) \cap \mathcal{F}_{1s} \quad \text{and} \quad \mathcal{K}_b = \left(\bigcap_{i \in I} \mathcal{K}_{bi} \right) \cap \mathcal{K}_{bs}. \tag{14}$$

Characterization of the Cone of Tangent Directions

The characterization of the cone of tangent directions is done by making use of Theorem 2.6. Assume that (A3) is valid. We denote $\mathcal{F}_2 = \{x \in \mathbb{R}^n : h_j(x) = 0, \ j \in J\}$. If $\{\nabla h_j(x^*)\}_{j \in J}$ is a linearly independent set, by Theorem 2.6,

$$\mathcal{K}_k = \{d \in \mathbb{R}^n : \langle \nabla h_j(x^*), d \rangle = 0, \ j \in J\}. \tag{15}$$

The Multiplier Rule

The theorem below is a first step in transcribing the geometric necessary optimality condition into an algebraic one.

Theorem 3.5 *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP). Assume that (A1)–(A2) hold. Then,*

$$\underline{\mathcal{K}}_0 \cap \overline{\mathcal{K}}_0 \cap \left(\bigcap_{i \in I} \mathcal{K}_{bi} \right) \cap \mathcal{K}_{bs} \cap \mathcal{K}_k = \emptyset.$$

Proof The result is a direct consequence of Theorems 3.1 and 3.4 and (14). □

In order to establish the main result of the paper, we will need a constraint qualification. Before stating the constraint qualification, we recall the definition of a normal cone.

Definition 3.2 The normal cone to a convex set $S \subset \mathbb{R}^n$ at the point $x^* \in S$, denoted by $N_S(x^*)$, is defined as

$$N_S(x^*) = \{v \in \mathbb{R}^n : \langle v, x - x^* \rangle \leq 0 \ \forall x \in S\}.$$

Definition 3.3 The constraints of (IOP) are said to satisfy the positive linear independence constraint qualification (PLICQ) at a feasible point $x^* \in \mathcal{F}$ if given $\mu_i \geq 0$, $i \in I(x^*)$, and $\lambda_j \in \mathbb{R}$, $j \in J$, such that

$$\sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) + \sum_{j \in J} \lambda_j \nabla h_j(x^*) \in -N_S(x^*)$$

one has

$$\mu_i = 0, \ i \in I(x^*), \quad \text{and} \quad \lambda_j = 0, \ j \in J.$$

Theorem 3.6 (Necessary optimality conditions for (IOP)) *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP). Assume that (A1)–(A4) and (PLICQ) hold at x^*. Then, there exist $\mu_i \geq 0$, $i \in I(x^*)$, $\lambda_j \in \mathbb{R}$, $j \in J$, and $y \in \mathbb{R}^n$ such that*

$$y \in \nabla_{gH} F(x^*) \oplus \sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) \oplus \sum_{j \in J} \lambda_j \nabla h_j(x^*)$$

and

$$\langle y, x - x^* \rangle \geq 0 \ \forall x \in S.$$

Proof Provided (PLICQ) is valid at x^*, we have that $\nabla g_i(x^*) \neq 0$, $i \in I(x^*)$, and $\nabla h_j(x^*) \neq 0$, $j \in J$. Moreover, the gradients of the equality constraints are linearly independent at x^*.

Suppose that $\underline{\mathcal{K}}_0 \neq \emptyset$ and $\overline{\mathcal{K}}_0 \neq \emptyset$. The exceptional cases (when this supposition is not valid) will be discussed at the end of the proof.

We know by Propositions 2.2, 2.3 and 3.2 and (12), (13) and (15) that $\underline{\mathcal{K}}_0$, $\overline{\mathcal{K}}_0$, $\mathcal{K}_{bi}$, $i \in I$, and $\mathcal{K}_{bs}$ are open convex cones and $\mathcal{K}_k$ is a convex cone. It follows from Theorem 3.5 and the Dubovitskii–Milyutin Lemma 2.2 that there exist linear functionals $\underline{f}_0 \in \underline{\mathcal{K}}_0^*$, $\overline{f}_0 \in \overline{\mathcal{K}}_0^*$, $f_{1i} \in \mathcal{K}_{bi}^*$, $i \in I$, $f_{1s} \in \mathcal{K}_{bs}^*$ and $f_2 \in \mathcal{K}_k^*$, not all zero, such that, for all $d \in \mathbb{R}^n$,

$$\underline{f}_0(d) + \overline{f}_0(d) + \sum_{i \in I} f_{1i}(d) + f_{1s}(d) + f_2(d) = 0. \tag{16}$$

By (10), (11), (13) and Theorem 2.7, we see that

$$\underline{\mathcal{K}}_0^* = \{-\underline{\lambda}\nabla_\sharp \underline{f}(x^*) : \underline{\lambda} \geq 0\}, \quad \overline{\mathcal{K}}_0^* = \{-\overline{\lambda}\nabla_\sharp \overline{f}(x^*) : \overline{\lambda} \geq 0\},$$
$$\mathcal{K}_{bi}^* = \{-\mu_i \nabla g_i(x^*) : \mu_i \geq 0\}, \ i \in I(x^*), \quad \mathcal{K}_{bi}^* = \{0\}, \ i \in I \backslash I(x^*).$$

From (13) and Theorem 2.9, we obtain

$$\mathcal{K}_{bs}^* = S^* = \{f \in (\mathbb{R}^n)' : f(x) \geq f(x^*) \ \forall x \in S\}.$$

By (15) and Theorem 2.8,

$$\mathcal{K}_k^* = \left\{ \sum_{j \in J} \lambda_j \nabla h_j(x^*) : \lambda_j \in \mathbb{R}, \ j \in J \right\}.$$

Therefore, from (16), there exist $\underline{\lambda} \geq 0$, $\overline{\lambda} \geq 0$, $\mu_i \geq 0$, $i \in I(x^*)$, $\mu_i := 0$, $i \in I \backslash I(x^*)$ and $\lambda_j \in \mathbb{R}$, such that, for all $d \in \mathbb{R}^n$,

$$f_{1s}(d) = \underline{\lambda}\langle \nabla_\sharp \underline{f}(x^*), d\rangle + \overline{\lambda}\langle \nabla_\sharp \overline{f}(x^*), d\rangle$$
$$+ \sum_{i \in I} \mu_i \langle \nabla g_i(x^*), d\rangle + \sum_{j \in J} \lambda_j \langle \nabla h_j(x^*), d\rangle. \tag{17}$$

Since $S^* \subset (\mathbb{R}^n)' \approx \mathbb{R}^n$, by the Riesz Representation Theorem (see Brezis [6]), there exists a unique $y \in \mathbb{R}^n$ such that

$$f_{1s}(x) = \langle y, x\rangle \ \forall x \in \mathbb{R}^n. \tag{18}$$

Provided f_{1s} is a supporting functional of S at $x^* \in S$, we have that

$$\langle y, x\rangle \geq \langle y, x^*\rangle \Leftrightarrow \langle y, x - x^*\rangle \geq 0 \ \forall x \in S. \tag{19}$$

It follows from (17) and (18) that

$$\left\langle y - \underline{\lambda}\nabla_\sharp \underline{f}(x^*) - \overline{\lambda}\nabla_\sharp \overline{f}(x^*) - \sum_{i \in I} \mu_i \nabla g_i(x^*) - \sum_{j \in J} \lambda_j \nabla h_j(x^*), d \right\rangle = 0$$

for all $d \in \mathbb{R}^n$. Then,

$$y = \underline{\lambda}\nabla_\sharp \underline{f}(x^*) + \overline{\lambda}\nabla_\sharp \overline{f}(x^*) + \sum_{i \in I} \mu_i \nabla g_i(x^*) + \sum_{j \in J} \lambda_j \nabla h_j(x^*). \tag{20}$$

Note that $\underline{\lambda} = 0$ and $\overline{\lambda} = 0$ cannot occur simultaneously. Indeed, if $\underline{\lambda} = \overline{\lambda} = 0$, by (19) and (20) and the fact that (PLICQ) is valid at x^*, it follows that $\mu_i = 0$, $i \in I$, and $\lambda_j = 0$, $j \in J$. This implies that $\underline{f}_0(d) = \overline{f}_0(d) = f_{1i}(d) = f_{1s}(d) = f_2(d) = 0$, $i \in I$, $j \in J$, for all $d \in \mathbb{R}^n$, which is contradiction. Therefore, we can assume that $\underline{\lambda} + \overline{\lambda} = 1$.

Since $\underline{\lambda} + \overline{\lambda} = 1$ with $\underline{\lambda}, \overline{\lambda} \geq 0$, we have that

$$\underline{\lambda} \frac{\partial \underline{f}}{\partial x_i^{\sharp}}(x^*) + \overline{\lambda} \frac{\partial \overline{f}}{\partial x_i^{\sharp}}(x^*) \in \left[\frac{\partial \underline{f}}{\partial x_i^{\sharp}}(x^*) \vee \frac{\partial \overline{f}}{\partial x_i^{\sharp}}(x^*) \right] = \frac{\partial_{gH} F}{\partial x_i}(x^*), \ i = 1, \ldots, n.$$

Then,

$$\underline{\lambda} \nabla_{\sharp} \underline{f}(x^*) + \overline{\lambda} \nabla_{\sharp} \overline{f}(x^*) \in \nabla_{gH} F(x^*). \tag{21}$$

From (20), (21) and the fact that $\mu_i = 0$, $i \in I \backslash I(x^*)$, we get

$$y \in \nabla_{gH} F(x^*) \oplus \sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) \oplus \sum_{j \in J} \lambda_j \nabla h_j(x^*). \tag{22}$$

The result follows from (19) and (22).

Finally, we discuss the exceptional cases. Suppose that $\underline{\mathcal{K}}_0 = \emptyset$, that is,

$$\langle \nabla_{\sharp} \underline{f}(x^*), d \rangle = 0 \ \forall d \in \mathbb{R}^n. \tag{23}$$

Taking $\underline{\lambda} = 1$, $\overline{\lambda} = 0$, $\mu_i = 0$, $i \in I$, $\lambda_j = 0$, $j \in J$, there exists $y = \nabla_{\sharp} \underline{f}(x^*) \in \mathbb{R}^n$ such that

$$y = \underline{\lambda} \nabla_{\sharp} \underline{f}(x^*) + \overline{\lambda} \nabla_{\sharp} \overline{f}(x^*) + \sum_{i \in I} \mu_i \nabla g_i(x^*) + \sum_{j \in J} \lambda_j \nabla h_j(x^*)$$

$$\in \nabla_{gH} F(x^*) \oplus \sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) \oplus \sum_{j \in J} \lambda_j \nabla h_j(x^*).$$

Moreover, (23) becomes $\langle y, d \rangle = 0$ for all $d \in \mathbb{R}^n$, which implies that $y = 0$, so that

$$\langle y, x - x^* \rangle \geq 0 \ \forall x \in S.$$

In the case $\overline{\mathcal{K}}_0 = \emptyset$, we proceed analogously. The proof is complete. □

Note that the condition $\langle y, x - x^* \rangle \geq 0$ for all $x \in S$ in Theorem 3.6 can also be expressed as $-y \in N_S(x^*)$. Thus, the necessary optimality conditions for (IOP) can be stated in the following alternative form.

Theorem 3.7 *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP). Assume that (A1)–(A4) and (PLICQ) hold at x^*. Then, there exist $\mu_i \geq 0$, $i \in I(x^*)$, $\lambda_j \in \mathbb{R}$, $j \in J$, such that*

$$0 \in \nabla_{gH} F(x^*) \oplus \sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) \oplus \sum_{j \in J} \lambda_j \nabla h_j(x^*) \oplus N_S(x^*).$$

In the following, we reproduce, as particular cases, the last result for some special forms of (IOP). First, we consider the problem without the abstract constraint:

$$\text{minimize } F(x) = [\underline{f}(x), \overline{f}(x)]$$
$$\text{subject to } g_i(x) \leq 0, \ i \in I = \{1, \ldots, p\}, \qquad \text{(IOP1)}$$
$$h_j(x) = 0, \ j \in J = \{1, \ldots, q\}.$$

In this case, the constraint qualification reads

Definition 3.4 The constraints of (IOP1) are said to satisfy the positive linear independence constraint qualification (PLICQ) at a feasible point $x^* \in \mathcal{F}$ if given $\mu_i \geq 0$, $i \in I(x^*)$, and $\lambda_j \in \mathbb{R}$, $j \in J$, such that

$$\sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) + \sum_{j \in J} \lambda_j \nabla h_j(x^*) = 0$$

one has

$$\mu_i = 0, \ i \in I(x^*), \quad \text{and} \quad \lambda_j = 0, \ j \in J.$$

Theorem 3.8 (Necessary optimality conditions for (IOP1)) *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP1). Assume that (A1)–(A3) and (PLICQ) hold at x^*. Then, there exist $\mu_i \geq 0$, $i \in I(x^*)$, $\lambda_j \in \mathbb{R}$, $j \in J$, such that*

$$0 \in \nabla_{gH} F(x^*) \oplus \sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*) \oplus \sum_{j \in J} \lambda_j \nabla h_j(x^*).$$

Now, we consider the problem with the abstract constraint only:

$$\text{minimize } F(x) = [\underline{f}(x), \overline{f}(x)] \qquad \text{(IOP2)}$$
$$\text{subject to } x \in S.$$

The constraint qualification is unnecessary.

Theorem 3.9 (Necessary optimality conditions for (IOP2)) *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP2). Assume that (A1), (A2) and (A4) hold at x^*. Then, there exists $y \in \mathbb{R}^n$ such that*

$$y \in \nabla_{gH} F(x^*)$$

and

$$\langle y, x - x^* \rangle \geq 0 \ \forall x \in S.$$

Alternatively,

Theorem 3.10 *Let $x^* \in \mathcal{F}$ be a local weak LU-solution of (IOP2). Assume that (A1), (A2) and (A4) hold at x^*. Then,*

$$0 \in \nabla_{gH} F(x^*) \oplus N_S(x^*).$$

We finalize this section by observing that, since every LU-solution is also a weak LU-solution, all the results above on necessary optimality conditions can be applied to characterize LU-solutions as well.

4 Examples and Possible Applications

As already mentioned in the introduction, interval optimization problems arise naturally in practice. In this section, we will briefly discuss some possible applications whose modeling falls into a framework often referred to as discrete-time optimal control. We first describe the discrete-time optimal control problems with an interval-valued objective function. Afterward, we present an application on reservoir regulation.

A discrete-time optimal control problem with an interval-valued objective function may be posed as follows:

$$\min F(x, u) = \sum_{k=0}^{N-1} F_k(x_k, u_k) + F_N(x_N)$$
$$\text{s.t. } x_{k+1} = f_k(x_k, u_k), \ k \in \{0, \dots, N-1\},$$
$$(x_k, u_k) \in S_k, \ k \in \{0, \dots, N-1\},$$

where $F_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathcal{K}_C$, $k \in \{0, \dots, N-1\}$, $F_N : \mathbb{R}^n \to \mathcal{K}_C$, $f_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $k \in \{0, \dots, N-1\}$, $S_k \subset \mathbb{R}^n \times \mathbb{R}^m$, $k \in \{0, \dots, N-1\}$. Vectors $u_k \in \mathbb{R}^m$ are called control vectors while $x_k \in \mathbb{R}^n$ are termed state vectors. We refer to $u = (u_0, \dots, u_{N-1})$ as a control policy. The corresponding $x = (x_0, \dots, x_N)$ is said to be the state trajectory. The integer N is the number of periods. The vector difference equation $x_{k+1} = f_k(x_k, u_k)$ is called system equation and describes the "dynamics" of the problem.

Regardless of its peculiar structure, the discrete optimal control problems above fit thoroughly as an (IOP) without inequality constraints. Then, the results established in the last section can be applied.

Discrete-time optimal control problem appear often in applications. One instance of such problems is the reservoir regulation. The deterministic case was described in Bertsekas [5], for example. What follows is an adaptation to the interval case. We denote by x_k the volume of water held in a reservoir at the k-th of N time periods and by u_k the water used for some productive purpose in period k. Then, the volume evolves according to

$$x_{k+1} = x_k - u_k, \ k \in \{0, \dots, N-1\}.$$

The outflow u_k is seen as the controls, while the volume x_k as the states. We may consider a cost, say $F_N(x_N)$, associated to the terminal state, and costs, say $F_k(u_k)$, related to the outflow u_k at period k. Due to operational and market uncertainties, it is interesting to consider such costs as interval-valued functions. When u_k is used for electric power generation, for example, $F_k(u_k)$ may be an interval function of the

value of power produced from u_k. The outflows $u_0, u_1, \ldots, u_{N-1}$ are to be chosen so that $\sum_{k=0}^{N-1} F_k(u_k) + F_N(x_N)$ is LU-minimized subject to some constraints on the volume and on the outflow (for example, x_k should be kept between some upper and lower bounds and u_k should have nonnegative entries). If we use S_k to denote such constraints, from the system equation $x_{k+1} = x_k - u_k$, we can write the problem as

$$\min \; F(x, u) = \sum_{k=0}^{N-1} F_k(x_k - x_{k+1}) + F_N(x_N)$$
$$\text{s.t.} \;\; (x_k, u_k) \in S_k, \; k \in \{0, \ldots, N-1\},$$

and Theorem 3.10 can be applied when $F(x, u)$ is gH-differentiable.

We close the paper with a numerical example to better illustrate the application of our results. With the purpose of illustrating the application of the main result, Theorem 3.6 or its equivalent formulation Theorem 3.7, we will not use the condensed form above. Let us consider the particular case below:

$$\min \; F(x, u) = \left[-u_0^2, 0\right] + \left[-u_1^2, 0\right] + \left[-\frac{(x_2 - 2)^3}{8}, -\frac{(x_2 - 2)^3}{8} - \frac{x_2^2 - 4x_2}{8}\right]$$
$$\text{s.t.} \;\; x_1 = x_0 - u_0,$$
$$x_2 = x_1 - u_1,$$
$$(x_0, u_0) \in \{5\} \times [0, 5],$$
$$(x_1, u_1) \in [3, 4] \times [0, 3].$$

From the system equation $x_1 = x_0 - u_0$ and the constraints $x_0 \in \{5\}$ and $x_1 \in [3, 4]$, we get $u_0 \in [1, 2]$. From the system equation $x_2 = x_1 - u_1$ and the constraints $x_1 \in [3, 4]$ and $u_1 \in [0, 3]$ we see that $x_2 \in [0, 4]$. Therefore, the feasible set $\mathcal{F}$ is contained in $\{5\} \times [3, 4] \times [0, 4] \times [1, 2] \times [0, 3]$. With respect to the notation of (IOP), we clearly identify

$$\underline{f}(x_0, x_1, x_2, u_0, u_1) = -u_0^2 - u_1^2 - \frac{(x_2 - 2)^3}{8},$$
$$\overline{f}(x_0, x_1, x_2, u_0, u_1) = -\frac{(x_2 - 2)^3}{8} - \frac{x_2^2 - 4x_2}{8},$$
$$h_1(x_0, x_1, x_2, u_0, u_1) = x_1 - x_0 + u_0,$$
$$h_2(x_0, x_1, x_2, u_0, u_1) = x_2 - x_1 + u_1,$$
$$S = \{5\} \times [3, 4] \times \mathbb{R} \times [0, 5] \times [0, 3].$$

The point $p^* = (x_0^*, x_1^*, x_2^*, u_0^*, u_1^*) = (5, 4, 4, 1, 0)$ is a local LU-solution. Indeed, assume that, for an arbitrary ϵ-neighborhood $N_\epsilon(p^*)$ of p^*, there exists a feasible solution $p \in N_\epsilon(p^*)$ with $F(p) \leq_{LU} F(p^*)$. Then, if $p = (x_0, x_1, x_2, u_0, u_1)$, we would have either

$$\begin{cases} -u_0^2 - u_1^2 - \frac{(x_2-2)^3}{8} \leq -2 \\ -\frac{(x_2-2)^3}{8} - \frac{x_2^2-4x_2}{8} < -1 \end{cases}$$

or

$$\begin{cases} -u_0^2 - u_1^2 - \frac{(x_2-2)^3}{8} < -2 \\ -\frac{(x_2-2)^3}{8} - \frac{x_2^2 - 4x_2}{8} \leq -1. \end{cases}$$

In the first case, the second inequality implies in $x_2 > 4$, which is contradiction to the fact that p is a feasible solution. In the second case, the second inequality gives $x_2 \geq 4$, so that $x_2 = 4$ since p is feasible. From the system equations, we have

$$4 = x_2 = x_1 - u_1 = x_0 - u_0 - u_1 = 5 - u_0 - u_1 \Leftrightarrow u_0 + u_1 = 1.$$

Using this information together with $(u_0, u_1) \in [1, 2] \times [0, 3]$ (remember that $p \in \mathcal{F}$), we obtain $u_0 = 1$ and $u_1 = 0$. Substituting in the first inequality from the second case we get the contradiction $-2 < -2$. Let λ_1 and λ_2 be scalars such that $\lambda_1 \nabla h_1(p^*) + \lambda_2 \nabla h_2(p^*) \in -N_S(p^*)$. Then,

$$\lambda_1(-1, 1, 0, 1, 0) + \lambda_2(0, -1, 1, 0, 1) \in \mathbb{R} \times \mathbb{R}_- \times \{0\} \times \{0\} \times \mathbb{R}_+,$$

so that $\lambda_1 = \lambda_2 = 0$, that is, (PLICQ) holds at p^*. By Theorem 3.7, we know that there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ such that

$$0 \in \nabla_{gH} F(p^*) \oplus \lambda_1 \nabla h_1(p^*) \oplus \lambda_2 \nabla h_2(p^*) \oplus N_S(p^*).$$

The right-hand side above is equal to

$$\left([0, 0], [0, 0], \left[-\frac{3}{8}(x_2^* - 2)^2 \vee -\frac{3}{8}(x_2^* - 2)^2 - \frac{2}{8}x_2^* + \frac{4}{8} \right], [-2u_0^*, 0], [-2u_1, 0] \right)$$

$$\oplus \lambda_1(-1, 1, 0, 1, 0) \oplus \lambda_2(0, -1, 1, 0, 1) \oplus \mathbb{R} \times \mathbb{R}_+ \times \{0\} \times \{0\} \times \mathbb{R}_-.$$

Thence, $0 \in \nabla_{gH} F(p^*) \oplus \lambda_1 \nabla h_1(p^*) \oplus \lambda_2 \nabla h_2(p^*) \oplus N_S(p^*)$ if

$$\begin{cases} 0 \in [-\lambda_1, -\lambda_1] + \mathbb{R}, \\ 0 \in [\lambda_1 - \lambda_2, \lambda_1 - \lambda_2] + \mathbb{R}_+, \\ 0 \in [-2, -3/2] + [\lambda_2, \lambda_2], \\ 0 \in [-2, 0] + [\lambda_1, \lambda_1], \\ 0 \in [\lambda_2, \lambda_2] + \mathbb{R}_-. \end{cases}$$

There exist infinitely many multipliers that fulfill the system of inclusions above, $\lambda_1 = 1$ and $\lambda_2 = 7/4$, for example. It is easy to see that the set of all multipliers is a compact one.

5 Conclusion

We considered an interval optimization problem with equality and inequality constraints along with an abstract constraint. The necessary optimality conditions were

obtained using a specification of the Dubovitskii–Milyutin formalism. We defined the LU-directions of decrease for interval-valued functions. The characterization of such directions was done through the gH-derivative of the interval-valued objective function. After an application of the Dubovitskii–Milyutin Lemma, we came to the Euler–Lagrange equation. A multiplier rule of Karush–Kuhn–Tucker type was then derived.

Among others, dealing with abstract constraints is important in optimal control problems: in general, control constraints are given in the abstract form. In the case of discrete optimal control interval problems, our results can be directly applied, as mentioned in the last section. But even in the continuous case, although the optimality conditions developed here are not directly applicable, our results can be useful. It is well known that the necessary optimality conditions for optimal control problems given in the maximum principle of Pontryagin involve the maximization of the Hamiltonian function with respect to the control variable at each instant of time t. Despite the optimal control problem being posed in spaces of infinite dimensions (in the continuous case), the maximization of the Hamiltonian at a given t falls exactly in an (IOP2)-type problem.

Another important feature of our work has to do with the fact that the necessary optimality conditions were achieved after the application of a specification of the Dubovitskii–Milyutin formalism. This allows us to easily generalize the results for problems posed in Banach spaces, continuous-time interval programming problems, and optimal control interval problems.

For future research, it would be of interest to work with problems in which not only the objective function is interval-valued, but the constraints are as well. More challenging would be working with problems whose decision variables are intervals.

Data Availability Data sharing is not applicable to this article.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

1. Ahmad, I., Singh, D., Dar, B.A.: Optimality conditions in multiobjective programming problems with interval valued objective functions. Control Cybern. **44**(1), 19–45 (2015). https://doi.org/10.2298/FIL1608121A

2. Ahmad, I., Singh, D., Dar, B.A.: Optimality conditions for invex interval valued nonlinear programming problems involving generalized H-derivative. Filomat **30**(8), 2121–2138 (2016). https://doi.org/10.2298/FIL1608121A

3. Antczak, T.: Optimality conditions and duality results for nonsmooth vector optimization problems with the multiple interval-valued objective function. Acta Math. Sci. Ser. (Engl. Ed.) **B37**(4), 1133–1150 (2017). https://doi.org/10.1016/S0252-9602(17)30062-0

4. Aubin, J.P., Cellina, A.: Differential inclusions. In: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 264. Springer, Berlin. https://doi.org/10.1007/978-3-642-69512-4. Set-valued Maps And Viability Theory (1984)

5. Bertsekas, D.P.: Nonlinear Programming, Athena Scientific Optimization and Computation Series, 2nd edn. Athena Scientific, Belmont (1999)

6. Brezis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext. Springer, New York (2011)

7. Cao, E., Gao, R., Lai, M.: Research on the vehicle routing problem with interval demands. Appl. Math. Model. **54**, 332–346 (2018). https://doi.org/10.1016/j.apm.2017.09.050

8. Chalco-Cano, Y., Lodwick, W.A., Rufián-Lizana, A.: Optimality conditions of type KKT for optimization problem with interval-valued objective function via generalized derivative. Fuzzy Optim. Decis. Mak. **12**(3), 305–322 (2013). https://doi.org/10.1007/s10700-013-9156-y

9. Costa, T.M., Bouwmeester, H., Lodwick, W.A., Lavor, C.: Calculating the possible conformations arising from uncertainty in the molecular distance geometry problem using constraint interval analysis. Inf. Sci. **415**(416), 41–5 (2017). https://doi.org/10.1016/j.ins.2017.06.015

10. Das, S., Mondal, R., Shaikh, A.A., Bhunia, A.K.: An application of control theory for imperfect production problem with carbon emission investment policy in interval environment. J. Frankl. Inst. **359**(5), 1925–1970 (2022). https://doi.org/10.1016/j.jfranklin.2022.01.035

11. Diamond, P., Kloeden, P.E.: Metric Spaces of Fuzzy Sets: Theory and Applications. World Scientific, Singapore (1994)

12. Girsanov, I.V.: Lecture Notes in Economics and Mathematical Systems, Operations Research, Computer Science, Social Science, vol. 67. Springer, Berlin (1972)

13. Hukuhara, M.: Intégration des applications mesurables dont la valeur est un compact convex. Funkcial Ekvac **10**, 205–229 (1967)

14. Ishibuchi, H., Tanaka, H.: Multiobjective programming in optimization of the interval objective function. Eur. J. Oper. Res. **48**(2), 219–225 (1990)

15. Kulisch, U.W., Miranker, W.L.: Computer Arithmetic in Theory and Practice. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London. Computer Science and Applied Mathematics (1981)

16. Kumar, P., Behera, J., Bhurjee, A.K.: Solving mean-VaR portfolio selection model with interval-typed random parameter using interval analysis. Opsearch **59**(1), 41–77 (2022). https://doi.org/10.1007/s12597-021-00531-7

17. Kummari, K., Ahmad, I.: Sufficient optimality conditions and duality for nonsmooth interval-valued optimization problems via L-invex-infine functions. Politehn. Univ. Bucharest Sci. Bull. Ser. A Appl. Math. Phys. **82**(1), 45–54 (2020)

18. Luhandjula, M.K., Rangoaga, M.J.: An approach for solving a fuzzy multiobjective programming problem. Eur. J. Oper. Res. **232**, 249–255 (2014)

19. Singh, D., Dar, B., Kim, D.S.: KKT optimality conditions in interval valued multiobjective programming with generalized differentiable functions. Eur. J. Oper. Res. **254**(1), 29–39 (2016)

20. Stefanini, L.: A generalization of Hukuhara difference and division for interval and fuzzy arithmetic. Fuzzy Sets Syst. **161**(11), 1564–1584 (2010). https://doi.org/10.1016/j.fss.2009.06.009

21. Stefanini, L., Arana-Jiménez, M.: Karush–Kuhn–Tucker conditions for interval and fuzzy optimization in several variables under total and directional generalized differentiability. Fuzzy Sets Syst. **362**, 1–34 (2019). https://doi.org/10.1016/j.fss.2018.04.009

22. Stefanini, L., Bede, B.: Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Anal. **71**(3), 1311–1328 (2009). https://doi.org/10.1016/j.fss.2012.12.004

23. Tung, L.T.: Karush–Kuhn–Tucker optimality conditions and duality for convex semi-infinite programming with multiple interval-valued objective functions. J. Appl. Math. Comput. **62**, 67–91 (2020). https://doi.org/10.1007/s12190-019-01274-x

24. Van Luu, D., Mai, T.T.: Optimality and duality in constrained interval-valued optimization. 4OR **16**(3), 311–337 (2018). https://doi.org/10.1007/s10288-017-0369-8
25. Wen, S., Lan, H., Hong, Y.-Y., Yu, D.C., Zhang, L., Cheng, P.: Allocation of ESS by interval optimization method considering impact of ship swinging on hybrid PV/diesel ship power system. Appl. Energy **175**, 158–167 (2016). https://doi.org/10.1016/j.apenergy.2016.05.003
26. Wu, H.C.: The Karush–Kuhn–Tucker optimality conditions for the optimization problem with fuzzy-valued objective function. Math. Methods Oper. Res. **66**(2), 203–224 (2007). https://doi.org/10.1007/s00186-007-0156-y
27. Wu, H.C.: The Karush–Kuhn–Tucker optimality conditions in multiobjective programming problems with interval-valued objective functions. Eur. J. Oper. Res. **196**(1), 49–60 (2009). https://doi.org/10.1016/j.ejor.2008.03.012
28. Zhang, J., Liu, S., Li, L., Feng, Q.: The KKT optimality conditions in a class of generalized convex optimization problems with an interval-valued objective function. Optim. Lett. **8**(2), 607–631 (2012). https://doi.org/10.1007/s11590-012-0601-6
29. Zhao, J., Bin, M.: Karush–Kuhn–Tucker optimality conditions for a class of robust optimization problems with an interval-valued objective function. Open Math. **18**(1), 781–793 (2020). https://doi.org/10.1515/math-2020-0042