Chapter 7
Semi-infinite Programs with Some Convexity

This and the next chapters of the book contain mainly some recent applications of the constructions and results of variational analysis and generalized differentiation presented above, as well as new developments required for such applications, to a remarkable class of optimization problems unified under the name of *semi-infinite programming* (SIP). We also use the abbreviations "SIP" for a particular semi-infinite program and "SIPs" as plural. The SIP terminology comes from the fact that originally this class of optimization problems concerned minimizing real-valued functions on *finite*-dimensional spaces subject to *infinitely* many inequality constraints usually indexed by a compact set. Over the years, the theory and applications of SIP have been evolved to include optimization problems with noncompact index sets and on infinite-dimensional spaces. Sometimes SIPs with infinite-dimensional decision spaces are labeled as problems of "infinite programming," while here we prefer to use the conventional SIP terminology regardless of the decision space dimension. As seen, the underlying style in the previous chapters was to present major results in finite-dimensional spaces and then to discuss infinite-dimensional extensions only in exercise and commentary sections. In contrast, the standing framework of this and the next chapters is, unless otherwise stated, the general *Banach space* setting. The main reasons for it are as follows:

(1) Due to their essence, SIPs always contain an infinite-dimensional part and require the usage of infinite-dimensional analysis for their investigation.

(2) The major results obtained below are formulated exactly in the same way in both cases of finite-dimensional and Banach decision spaces.

(3) Many practically meaningful models can be described as SIPs with infinite-dimensional decision spaces. In particular, this is the case of the water resource optimization problem, which is formulated and solved in Section 7.2 by using the necessary optimality conditions obtained therein.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_7

7.1 Stability of Infinite Linear Inequality Systems

In this section, we study the sets of feasible solutions to SIPs described by the parameterized *infinite systems of linear inequalities*

$$\mathcal{F}(p) := \big\{x \in X \big| \ \langle a_t^*, x\rangle \le b_t + p_t,\ t \in T\big\}, \quad p = (p_t)_{t\in T}, \tag{7.1}$$

with an *arbitrary index set* T, where $x \in X$ is a *decision* variable belonging to a *Banach* space X and where $p = (p_t)_{t\in T} \in P$ is a functional *parameter* taking values in the prescribed Banach space P of perturbations specified below. The data of (7.1) are given as follows:

• $a_t^* \in X^*$ are fixed for all $t \in T$. We use the same notation for the given norm $\|\cdot\|$ on X and the corresponding dual norm on X^* defined by

$$\|x^*\| := \sup\big\{\langle x^*, x\rangle \big|\ \|x\| \le 1\big\}, \quad x^* \in X^*.$$

• $b_t \in \mathbb{R}$ are fixed for all $t \in T$. We identify the collection $\{b_t |\ t \in T\}$ with the real-valued function $b\colon T \to \mathbb{R}$.

• $p_t = p(t) \in \mathbb{R}$ for all $t \in T$. These functional parameters $p\colon T \to \mathbb{R}$ are our varying perturbations, which are taken from the Banach parameter space $P := l^\infty(T)$ of all bounded functions on T with the supremum norm $\|p\|_\infty := \sup\big\{|p(t)| \ \big|\ t \in T\big\}$. When T is compact and $p(\cdot)$ are restricted to be continuous on T, the parameter space P reduces to $\mathcal{C}(T)$.

It is obvious that the space $l^\infty(T)$ is never finite-dimensional when the index set T is infinite. Moreover, in the infinite-dimensional case, the space $l^\infty(T)$ is *never Asplund*; see [638, Example 1.21].

The primary goal of this section is to calculate the *coderivative* of the set-valued mapping $\mathcal{F}$ defined in (7.1) as well as the *coderivative norm* of $\mathcal{F}$ at the reference point entirely in terms of the initial data of (7.1). Based on this, we derive here a complete *coderivative characterization* of the *Lipschitz-like property* of $\mathcal{F}$ in the form identical to the finite-dimensional setting of Chapter 3. Furthermore, the obtained coderivative calculation is the key of deriving necessary optimality conditions for SIPs with linear inequality constraints of type (7.1) in Section 7.2 and then in turn becomes crucial to investigate SIPs described by convex inequalities and the like in the subsequent Section 7.3.

Recall that the coderivative of any mapping $F\colon X \rightrightarrows Y$ between Banach spaces studied in this and the next chapters is considered in the usual "normal" sense as in finite dimensions. This means that, given any $(\bar{x}, \bar{y}) \in \text{gph}\, F$, the coderivative of F at $(\bar{x}, \bar{y})$ is the mapping $F\colon Y^* \rightrightarrows X^*$ defined by

$$D^*F(\bar{x}, \bar{y})(y^*) := \big\{x^* \in X^* \big|\ (x^*, -y^*) \in N\big((\bar{x}, \bar{y}); \text{gph}\, F\big)\big\} \tag{7.2}$$

for $y^* \in Y^*$ via the corresponding normal cone to the graph of F at $(\bar{x}, \bar{y})$.

7.1.1 Lipschitz-Like Property and Strong Slater Condition

Since we are in the general Banach space setting, the symbol w^*-lim signifies here the weak* *topological* limit in the dual space in question. This corresponds to the convergence of *nets* denoted usually by $\{x_\nu^*\}_{\nu\in\mathcal{N}}$. In the case of sequences, we replace the symbol $\mathcal{N}$ by the standard natural series notion $\mathbb{N} = \{1, 2, \ldots\}$. For an arbitrary index set T, denote by $\mathbb{R}^T$ the *product space* of $\lambda = (\lambda_t |\ t \in T)$ with $\lambda_t \in \mathbb{R}$ for all $t \in T$. Finally, let $\mathbb{R}^{(T)}$ be the collection of multipliers $\lambda \in \mathbb{R}^T$ such that $\lambda_t \neq 0$ for *finitely many* $t \in T$, and let $\mathbb{R}_+^{(T)}$ be the *positive cone* in $\mathbb{R}^{(T)}$ defined by

$$\mathbb{R}_+^{(T)} := \big\{\lambda \in \mathbb{R}^{(T)} \big|\ \lambda_t \geq 0 \ \text{ for all } \ t \in T\big\}. \tag{7.3}$$

Note also that throughout this chapter, the symbol "cone Ω" stands for the *convex* conic hull of the set in question.

Let us now recall a well-recognized *qualification condition* for SIPs with infinite linear inequality constraints and then show that it provides, along with other conditions, an equivalent description of the Lipschitz-like property of the constraint mapping $\mathcal{F}$ from (7.1).

Definition 7.1 (Strong Slater Condition). *We say that the infinite linear inequality system* (7.1) *satisfies the* STRONG SLATER CONDITION (SSC) *at* $p = (p_t)_{t\in T}$ *if there exists* $\widehat{x} \in X$ *such that*

$$\sup_{t\in T}\big[\langle a_t^*, \widehat{x}\rangle - b_t - p_t\big] < 0. \tag{7.4}$$

Furthermore, every point $\widehat{x} \in X$ *satisfying condition* (7.4) *is a* STRONG SLATER POINT *for system* (7.1) *at* $p = (p_t)_{t\in T}$.

Define further the parametric *characteristic sets*

$$C(p) := \text{co}\big\{(a_t^*, b_t + p_t)\big|\ t \in T\big\}, \quad p \in l^\infty(T), \tag{7.5}$$

and suppose without loss of generality that $\bar{p} = 0 \in l^\infty(T)$ is the designated *nominal parameter*. First, we verify the following equivalences.

Theorem 7.2 (Equivalent Descriptions of the Lipschitz-Like Property for Infinite Linear Systems). *Given* $p \in \text{dom}\,\mathcal{F}$ *for* (7.1) *in the Banach decision space* X, *the following properties are equivalent:*

(i) $\mathcal{F}$ *is Lipschitz-like around* (p, x) *for all* $x \in \mathcal{F}(p)$.
(ii) $p \in \text{int}(\text{dom}\,\mathcal{F})$.
(iii) $\mathcal{F}$ *satisfies the strong Slater condition at* p.
(iv) $(0, 0) \notin \text{cl}^*C(p)$ *via the characteristic set in* (7.5).

Finally, the boundedness of $\big\{a_t^* \mid t \in T\big\}$ *ensures the equivalence of* (i)–(iv) *to:*

(v) *there exists* $\widehat{x} \in X$ *such that* $(p, \widehat{x}) \in \text{int}(\text{gph}\,\mathcal{F})$.

Proof. The equivalence between (i) and (ii) is a consequence of the Robinson-Ursescu theorem and the equivalence between the Lipschitz-like property of the convex-graph mapping $\mathcal{F}$ and the metric regularity/covering properties of its inverse; see Theorem 3.2 and Corollary 3.6 together with the corresponding exercises and commentaries in Sections 3.4 and 3.5.

To verify implication (iii)⇒(ii), suppose that $\widehat{x}$ is a strong Slater point for system (7.1) at p and find $\vartheta > 0$ such that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t \le -\vartheta \quad \text{for all} \quad t \in T.$$

Then it is obvious that for any $q \in l^\infty(T)$ with $\|q\| < \vartheta$, we have $\widehat{x} \in \mathcal{F}(p + q)$. Therefore $p + q \in \operatorname{dom}\mathcal{F}$, and thus (ii) holds. To justify further the converse implication (ii)⇒(iii), take $p \in \operatorname{int}(\operatorname{dom}\mathcal{F})$, and then get $p + q \in \operatorname{dom}\mathcal{F}$ provided that $q_t = -\vartheta$ as $t \in T$ and that $\vartheta > 0$ is sufficiently small. Thus every $\widehat{x} \in \mathcal{F}(p+q)$ is a strong Slater point for the infinite system (7.1) at p.

Next we show that (iii)⇒(iv). Arguing by contradiction, suppose that $(0, 0) \in \operatorname{cl}^* C(p)$. Then there exists a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \in \mathbb{R}_+^{(T)}$ satisfying the equality $\sum_{t\in T} \lambda_{t\nu} = 1$ for all $\nu \in \mathcal{N}$ and the limiting condition

$$(0, 0) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\big(a_t^*, b_t + p_t\big). \tag{7.6}$$

If $\widehat{x}$ is a strong Slater point for system (7.1) at p, we find $\vartheta > 0$ such that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t \le -\vartheta \quad \text{for all} \quad t \in T.$$

Then condition (7.6) leads us to the contradiction

$$0 = \langle 0, \widehat{x}\rangle + 0 \cdot (-1) = \lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\big(\langle a_t^*, \widehat{x}\rangle + (b_t + p_t) \cdot (-1)\big) \le -\vartheta,$$

which thus justifies (iii)⇒(iv). To verify the converse implication (iv)⇒(iii), we employ the dual description of the *consistency* in (7.1) given by

$$p \in \operatorname{dom}\mathcal{F} \Longleftrightarrow (0, -1) \notin \operatorname{cl}^*\operatorname{cone}\big\{(a_t^*, b_t + p_t)\big|\ t \in T\big\}, \tag{7.7}$$

which is discussed in Exercise 7.71 and the commentaries in Section 7.7. Then the classical strong separation theorem gives us $(0, 0) \ne (v, \alpha) \in X \times \mathbb{R}$ with

$$\big\langle a_t^*, v\big\rangle + \alpha(b_t + p_t) \le 0, \ t \in T, \quad \text{and} \quad \langle 0, v\rangle + (-1)\alpha = -\alpha > 0. \tag{7.8}$$

Using (iv), we get $(0, 0) \ne (z, \beta) \in X \times \mathbb{R}$ and $\gamma \in \mathbb{R}$ for which

$$\langle a_t^*, z\rangle + \beta(b_t + p_t) \le \gamma < 0 \quad \text{whenever} \quad t \in T. \tag{7.9}$$

Consider now the combination $(u, \eta) := (z, \beta) + \lambda(v, \alpha)$, and select $\lambda > 0$ such that $\eta < 0$. Defining $\widehat{x} := -\eta^{-1}u$, we deduce from (7.8) and (7.9) that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t = -\eta^{-1}\big(\langle a_t^*, u\rangle + \eta(b_t + p_t)\big) \le -\eta^{-1}\gamma < 0.$$

Hence $\widehat{x}$ is a strong Slater point for system (7.1) at p, i.e., (iii) holds.

It remains to consider condition (v). It is easy to see that (v) always implies (iv) and so the other conditions of the theorem. Suppose now that the set $\{a_t^* \mid t \in T\}$ is bounded, and show that (iii) implies (v). Select $M \ge 0$ such that $\|a_t^*\| \le M$ for every $t \in T$, and take $\widehat{x} \in X$ satisfying (7.4). Denote

$$\gamma := -\sup_{t\in T}\big[\langle a_t^*, \widehat{x}\rangle - b_t - p_t\big] > 0$$

and consider any pair $(p', u) \in l^\infty(T) \times X$ such that

$$\|u\| \le \eta := \gamma/(M+1) > 0 \ \text{ and } \ \|p'\| \le \eta.$$

It is easy to see that for such (p', u) and every $t \in T$, we have

$$\langle a_t^*, \widehat{x} + u\rangle - b_t - p_t - p_t' \le -\gamma + M\|u\| + \|p'\| \le \eta(M+1) - \gamma = 0,$$

and so $(p + p', \widehat{x} + u) \in \operatorname{gph}\mathcal{F}$. Thus $(p, \widehat{x}) \in \operatorname{int}(\operatorname{gph}\mathcal{F})$, which verifies implication (iii)$\Rightarrow$(v) and completes the proof of the theorem. △

7.1.2 Coderivatives for Parametric Infinite Linear Systems

In this subsection, we calculate the coderivative $D^*\mathcal{F}(0, \bar{x})$ as in (7.2) of the parametric infinite system (7.1) at the reference point $(0, \bar{x})$ and also its norm $\|D^*\mathcal{F}(0, \bar{x})\|$ entirely via the initial data of (7.1). Recall that the dual space $l^\infty(T)^*$ to the parameter space in (7.1) is isometric to the space $ba(T)$ of all the bounded and additive measures $\mu(\cdot)$ on subsets of T with the norm

$$\|\mu\| := \sup_{A\subset T}\mu(A) - \inf_{B\subset T}\mu(B).$$

In what follows a dual element $p^* \in l^\infty(T)^*$ is identified with the corresponding measure $\mu \in ba(T)$ satisfying the canonical duality relationship

$$\langle \mu, p\rangle = \int_T p_t\,\mu(dt), \quad p = (p_t)_{t\in T}.$$

To proceed further, we need the following extension of the classical *Farkas lemma* to the case of infinite linear inequality systems; see Exercise 7.73 and the corresponding commentaries in Section 7.7.

Proposition 7.3 (Extended Farkas Lemma for Infinite Linear Inequalities). *Let $p \in \operatorname{dom}\mathcal{F}$ for the infinite system (7.1), and let $(x^*, \alpha) \in X^* \times \mathbb{R}$. The following assertions are equivalent:*

(i) *We have $\langle x^*, x\rangle \le \alpha$ whenever $x \in \mathcal{F}(p)$, i.e.,*

$$\big[\langle a_t^*, x\rangle \le b_t + p_t \ \textit{for all} \ t \in T\big] \Longrightarrow \big[\langle x^*, x\rangle \le \alpha\big].$$

(ii) *The pair* (x^*, α) *satisfies the inclusion*

$$(x^*, \alpha) \in \text{cl}^*\text{cone}\Big[\Big\{(a_t^*, b_t + p_t)\Big|\ t \in T\Big\} \cup \{(0, 1)\}\Big] \ \textit{with} \ 0 \in X^*.$$

Using Proposition 7.3, we first describe the normal cone to the graph

$$\text{gph}\,\mathcal{F} = \Big\{(p, x) \in l^\infty(T) \times X\Big|\ \langle a_t^*, x\rangle \le b_t + p_t \ \text{ for all } \ t \in T\Big\}$$

at the reference point $(0, \bar{x}) \in \text{gph}\,\mathcal{F}$. Recall that δ_t stands for the classical *Dirac function/measure* at $t \in T$ satisfying

$$\langle \delta_t, p\rangle = p_t \ \text{ as } \ t \in T \ \text{ for } \ p = (p_t)_{t \in T} \in l^\infty(T). \tag{7.10}$$

Proposition 7.4 (Graphical Normals for Infinite Linear Systems). *Let* $(0, \bar{x}) \in \text{gph}\,\mathcal{F}$ *for the mapping* $\mathcal{F}$ *from* (7.1)*, and let* $(p^*, x^*) \in l^\infty(T)^* \times X^*$*. Then we have* $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ *if and only if*

$$\big(p^*, x^*, \langle x^*, \bar{x}\rangle\big) \in \text{cl}^*\text{cone}\Big[\Big\{(-\delta_t, a_t^*, b_t)\Big|\ t \in T\Big\} \cup \{(0, 0, 1)\}\Big], \tag{7.11}$$

where $0 \in l^\infty(T)^*$ *and* $0 \in X^*$ *stand for the first and second entry of the last triple, respectively. Furthermore, the inclusion* $(p^*, x^*) \in N\left((0, \bar{x})\,; \text{gph}\,\mathcal{F}\right)$ *implies that* $p^* \le 0$ *in the space* ba (T)*, i.e.,* $p^*(A) \le 0$ *for all* $A \subset T$.

Proof. It is easy to see that

$$\text{gph}\,\mathcal{F} = \Big\{(p, x) \in l^\infty(T) \times X\Big|\ \langle a_t^*, x\rangle - \langle \delta_t, p\rangle \le b_t \ \text{ for all } \ t \in T\Big\},$$

and therefore we have $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ if and only if

$$\langle p^*, p\rangle + \langle x^*, x\rangle \le \langle x^*, \bar{x}\rangle \ \text{ for every } \ (p, x) \in \text{gph}\,\mathcal{F}. \tag{7.12}$$

Employing now the equivalence between (i) and (ii) in Proposition 7.3, we conclude that $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ if and only if inclusion (7.11) holds.

To justify the last statement of the proposition, for every set $A \subset T$, consider its *characteristic function* $\chi_A\colon T \to \{0, 1\}$ defined by

$$\chi_A(t) := \begin{cases} 1 & \text{if } \ t \in A, \\ 0 & \text{if } \ t \notin A. \end{cases}$$

It is obvious that the inclusion $(p, x) \in \text{gph}\,\mathcal{F}$ implies that $(p + \lambda\chi_A, x) \in \text{gph}\,\mathcal{F}$ for each $\lambda > 0$. Replacing now in (7.12) the pair (p, x) by $(p + \lambda\chi_A, x)$, dividing both sides of the inequality by λ, and then letting $\lambda \to \infty$ give us

$$\big\langle p^*, \chi_A\big\rangle = \int_T \chi_A(t)\, p^*(dt) = p^*(A) \le 0,$$

which completes the proof of the proposition. △

The representation of graphical normals obtained in Proposition 7.4 is crucial to calculate the coderivative of $D^*\mathcal{F}(0,\bar{x})$ defined via the normal cone to the gph $\mathcal{F}$ at $(0,\bar{x})$ according to (7.2).

Theorem 7.5 (Coderivative Calculation). *Given $\bar{x} \in \mathcal{F}(0)$ for the infinite system (7.1), we have that $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ if and only if*

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{cone}\big\{(-\delta_t, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.13}$$

Proof. It follows from the coderivative definition and Proposition 7.4 that $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ if and only if

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{cone}\Big[\big\{(-\delta_t, a_t^*, b_t)\big|\ t \in T\big\} \cup \{(0,0,1)\}\Big]. \tag{7.14}$$

To justify the coderivative representation claimed in the theorem, we need to show that inclusion (7.14) yields in fact the "smaller" one in (7.13). Assuming indeed that (7.14) holds, we find by (7.14) some nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ and $\{\gamma_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+$ satisfying the limiting relationship

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big(\sum_{t\in T}\lambda_{t\nu}(-\delta_t, a_t^*, b_t) + \gamma_\nu(0,0,1)\Big), \tag{7.15}$$

where $\lambda_{t\nu}$ stands for the t-entry of $\lambda_\nu = (\lambda_{t\nu})_{t\in T}$ as $\nu \in \mathcal{N}$. The component structure of (7.15) tells us that

$$0 = \langle p^*, 0\rangle + \langle -x^*, \bar{x}\rangle + (-\langle x^*, \bar{x}\rangle)(-1) = \lim_{\nu\in\mathcal{N}}\Big(\sum_{t\in T}\lambda_{t\nu}\big(\langle a_t^*, \bar{x}\rangle - b_t\big) - \gamma_\nu\Big).$$

Taking into account the definition (7.3) of the positive cone $\mathbb{R}_+^{(T)}$ and that $(0,\bar{x})$ satisfies the infinite inequality system in (7.1), we get $\lim_{\nu\in\mathcal{N}}\gamma_\nu = 0$. This justifies (7.13) and thus completes the proof of the theorem. △

The next consequence of Theorem 7.5 is useful in what follows.

Corollary 7.6 (Limiting Coderivative Description). *If $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ in the setting of Theorem 7.5, then there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ with*

$$\sum_{t\in T}\lambda_{t\nu} \to \|p^*\| = -\big\langle p^*, e\big\rangle,\ \sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} -x^*,\ \ \textit{and}\ \ \sum_{t\in T}\lambda_{t\nu}b_t \to -\big\langle x^*, \bar{x}\big\rangle.$$

Proof. Theorem 7.5 gives us a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ such that

$$\sum_{t\in T}\lambda_{t\nu}\delta_t \xrightarrow{w^*} -p^*,\quad \sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} -x^*,\quad \text{and}\ \sum_{t\in T}\lambda_{t\nu}b_t \to -\big\langle x^*, \bar{x}\big\rangle.$$

This readily implies the relationships

$$\left\langle \sum_{t\in T} \lambda_{t\nu}\delta_t, e\right\rangle = \sum_{t\in T} \lambda_{t\nu} \to \left\langle p^*, -e\right\rangle =: \lambda \in [0,\infty)\,.$$

Since the dual norm on X^* is w^*-lower semicontinuous, we have

$$\|p^*\| \le \liminf_{\nu\in\mathcal{N}} \left\| \sum_{t\in T} \lambda_{t\nu}\delta_t \right\| \le \liminf_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu} = \lambda.$$

Furthermore, it follows from the norm definition that

$$\|p^*\| = \sup_{\|p\|\le 1} \left\langle p^*, p\right\rangle \ge \left\langle p^*, -e\right\rangle = \lambda,$$

which yields $\|p^*\| = -\langle p^*, e\rangle$ and thus completes the proof. △

Now we proceed with the exact calculation of the *coderivative norm*

$$\|D^*\mathcal{F}(0,\bar{x})\| := \sup\left\{\|p^*\| \;\middle|\; p^* \in D^*\mathcal{F}(0,\bar{x})(x^*),\ \|x^*\| \le 1\right\} \tag{7.16}$$

entirely via the initial data of the infinite linear inequality system (7.1). A part of our analysis is the following proposition on properties of the characteristic set (7.5) at $p = 0$ in connection with strong Slater points of (7.1).

Proposition 7.7 (Strong Slater Points Relative to the Characteristic Set). *Given $\bar{x} \in \mathcal{F}(0)$, consider the set*

$$S := \left\{x^* \in X^* \middle|\, \left(x^*, \langle x^*, \bar{x}\rangle\right) \in \mathrm{cl}^*C(0)\right\} \tag{7.17}$$

built upon $C(0)$ from (7.5). The following assertions hold:

(i) *Let $\bar{x}$ be not a strong Slater point of the infinite system (7.1) at $p = 0$, and let the coefficient collection $\{a_t^*|\ t \in T\}$ be bounded in X^*. Then the set S in (7.17) is nonempty and w^*-compact in X^*.*

(ii) *Let $\bar{x}$ be a strong Slater point of (7.1) at $p = 0$. Then $S = \emptyset$ in (7.17).*

Proof. To justify (i), assume that $\bar{x}$ is not a strong Slater point for the infinite system (7.1) at $p = 0$. Then there is a sequence $\{t_k\}_{k\in\mathbb{N}} \subset T$ such that $\lim_k(\langle a_{t_k}^*, \bar{x}\rangle - b_{t_k}) = 0$. The boundedness of $\{a_t^*|\ t \in T\}$ implies by the classical Alaoglu-Bourbaki theorem that this set is relatively w^*-compact in X^*, i.e., there is a subnet $\{a_{t_\nu}^*\}_{\nu\in\mathcal{N}}$ of the latter sequence that w^*-converges to some element $u^* \in \mathrm{cl}^*\{a_t^*|\ t \in T\}$. This yields $\lim_{\nu\in\mathcal{N}} b_{t_\nu} = \langle u^*, \bar{x}\rangle$ and

$$\left(u^*, \langle u^*, \bar{x}\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \left(a_{t_\nu}^*, b_{t_\nu}\right) \in \mathrm{cl}^*C(0),$$

which justifies the nonemptiness of the set S in (7.17).

To verify the w^*-compactness of S, observe that the boundedness of the set $A := \{a_t^* |\ t \in T\}$ implies this property of $\mathrm{cl}^*\mathrm{co}A$; the latter set is actually w^*-compact due to its automatic w^*-closedness. Note further that the set S in (7.17) is a preimage of $\mathrm{cl}^*C(0)$ under the w^*-continuous mapping $x^* \mapsto (x^*, \langle x^*, \bar{x}\rangle)$, and thus it is w^*-closed in X^*. Since S is a subset of $\mathrm{cl}^*\mathrm{co}A$, it is also bounded and hence w^*-compact in X^*. We are done with (i).

To proceed with (ii), let $\bar{x}$ be a strong Slater point of system (7.1) at $p = 0$, and let $\gamma := -\sup_{t\in T}\left\{\left\langle a_t^*, \bar{x}\right\rangle - b_t\right\}$. Then we have the inequality

$$\left\langle x^*, \bar{x}\right\rangle \le \beta - \gamma \quad \text{whenever} \quad \left(x^*, \beta\right) \in \mathrm{cl}^*C\left(0\right),$$

which justifies (ii) and thus completes the proof of the proposition. △

Now we are ready to calculate the coderivative norm $\|D^*\mathcal{F}(0, \bar{x})\|$ entirely in terms of the given data of the infinite system (7.1) in Banach spaces.

Theorem 7.8 (Calculating the Coderivative Norm). *Let $\bar{x} \in \mathrm{dom}\,\mathcal{F}$ for the infinite system (7.1), which satisfies the strong Slater condition at $p = 0$. Then the following assertions hold under the boundedness of $\{a_t^* |\ t \in T\}$:*

(i) *If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then $\|D^*\mathcal{F}(0, \bar{x})\| = 0$.*

(ii) *If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then the coderivative norm (7.16) is positive and is calculated by*

$$\|D^*\mathcal{F}(0, \bar{x})\| = \max\left\{\|x^*\|^{-1} \,\Big|\, \left(x^*, \langle x^*, \bar{x}\rangle\right) \in \mathrm{cl}^*C(0)\right\}. \tag{7.18}$$

Proof. To verify assertion (i), suppose that $\bar{x}$ is a strong Slater point for the system $\mathcal{F}$ at $p = 0$. It follows from the proof of implication (iii)⇒(v) in Theorem 7.2 that $(0, \bar{x}) \in \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ and hence $N((0, \bar{x}); \mathrm{gph}\,\mathcal{F}) = \{(0, 0)\}$. Thus (i) follows from definitions of the coderivative and its norm.

To prove assertion (ii), take $x^* \in X^*$ such that $(x^*, \langle x^*, \bar{x}\rangle) \in \mathrm{cl}^*C(0)$; the latter set is nonempty according to Proposition 7.7. Then there exists a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ with $\sum_{t\in T}\lambda_{t\nu} = 1$ for all $\nu \in \mathcal{N}$ such that

$$\sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} x^* \quad \text{and} \quad \sum_{t\in T}\lambda_{t\nu}b_t \to \left\langle x^*, \bar{x}\right\rangle.$$

Form further the net elements $p_\nu^* \in l^\infty(T)^*$ by

$$p_\nu^* := -\sum_{t\in T}\lambda_{t\nu}\delta_t, \quad \text{with} \quad \left\|p_\nu^*\right\| = \left\langle p_\nu^*, -e\right\rangle = 1, \;\; \nu \in \mathcal{N},$$

and find by the Alaoglu-Bourbaki theorem a convergent subnet $p_\nu^* \xrightarrow{w^*} p^*$ for some $p^* \in l^\infty(T)^*$ with $\|p^*\| \le 1$. Employing the same arguments as in the proof of Corollary 7.6, we conclude that

$$1 = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu} = \left\|p^*\right\| = \left\langle p^*, -e\right\rangle. \tag{7.19}$$

Furthermore, it follows by passing to the limit that

$$\left(p^*, x^*, \langle x^*, \bar{x}\rangle\right) \in \mathrm{cl}^*\mathrm{co}\left\{\left(-\delta_t, a_t^*, b_t\right) \,\middle|\, t \in T\right\},$$

which implies by the coderivative calculation of Theorem 7.5 that

$$p^* \in D^*\mathcal{F}(0, \bar{x})(-x^*). \tag{7.20}$$

Suppose now that $x^* = 0$ in (7.20). Since $p^* \neq 0$ by (7.19), we get from (7.20) that $D^*\mathcal{F}(0, \bar{x})(0) \neq \{0\}$. It tells us by Exercise 3.35(i) and the graph convexity of $\mathcal{F}$ that $\mathcal{F}$ is *not Lipschitz-like* around $(0, \bar{x})$ and therefore it cannot satisfy the strong Slater condition by implication (iii)⇒(i) in Theorem 7.2. This contradicts the assumption imposed in the theorem.

Thus $x^* \neq 0$ in (7.20), and we derive from the latter relationship that

$$\|x^*\|^{-1} p^* \in D^*\mathcal{F}(0, \bar{x})\left(-\|x^*\|^{-1} x^*\right),$$

which gives us in turn the estimate

$$\|D^*\mathcal{F}(0, \bar{x})\| \geq \left\| \|x^*\|^{-1} p^* \right\| = \|x^*\|^{-1}$$

and hence justifies the inequality "≥" in (7.18).

It remains to prove the opposite inequality in (7.18). For the nonempty and w^*-compact set S in (7.17), we have $0 \notin S$ by Theorem 7.2, and the function $x^* \mapsto \|x^*\|^{-1}$ is w^*-upper semicontinuous of on S. Thus the supremum in the right-hand side of (7.18) is attained and belongs to $(0, \infty)$. Then condition (v) in Theorem 7.2 implies that $(0, \widehat{x}) \in \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ for some $\widehat{x} \in X$ and so $0 \in \mathrm{int}(\mathrm{dom}\,\mathcal{F})$. Moreover, we have that $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$ if and only if $(p^*, x^*) \in N\left((0, \bar{x})\,;\mathrm{gph}\,\mathcal{F}\right)$, which is equivalent to

$$\langle p^*, p\rangle + \langle x^*, x\rangle \leq \langle x^*, \bar{x}\rangle \quad \text{for all} \quad (p, x) \in \mathrm{gph}\,\mathcal{F}. \tag{7.21}$$

This allows us, by taking into account that $0 \in \mathrm{int}(\mathrm{dom}\,\mathcal{F})$, to arrive at

$$p^* \in D^*\mathcal{F}(0, \bar{x})\,(0) \Longleftrightarrow \langle p^*, p\rangle \leq 0 \;\text{ for all }\; p \in \mathrm{dom}\mathcal{F} \Longleftrightarrow p^* = 0. \tag{7.22}$$

Observe furthermore that, since $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, we have $(0, \bar{x}) \notin \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ and thus conclude by the classical separation theorem that there is a pair $(p^*, x^*) \neq (0, 0)$ for which condition (7.21) holds. Employing (7.22), we have $x^* \neq 0$ and $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$.

Take now $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$ with $\|x^*\| \leq 1$, and suppose that $x^* \neq 0$; the arguments above ensure the existence of such an element. By Corollary 7.6, there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ for which

$$\gamma_\nu := \sum_{t\in T} \lambda_{t\nu} \to \|p^*\| = -\langle p^*, e\rangle, \;\; x_\nu^* := \sum_{t\in T} \lambda_{t\nu} a_t^* \xrightarrow{w^*} x^*, \;\; \sum_{t\in T} \lambda_{t\nu} b_t \to \langle x^*, \bar{x}\rangle.$$

Taking $M \geq \|a_t^*\|$ for every $t \in T$, we get the estimate

$$\|x_\nu^*\| \leq M\gamma_\nu \quad \text{whenever} \quad \nu \in \mathcal{N}$$

and also the limiting relationships

$$0 < \|x^*\| \leq \liminf_{\nu \in \mathcal{N}} \|x_\nu^*\| \leq M \liminf_{\nu \in \mathcal{N}} \gamma_\nu = M\|p^*\|,$$

which ensure that $p^* \neq 0$. It follows furthermore that

$$\|p^*\|^{-1}\big(x^*, \langle x^*, \bar{x}\rangle\big) \in \text{cl}^* C(0).$$

Remembering finally that $0 < \|x^*\| \leq 1$, we arrive at the estimates

$$\|p^*\| \leq \Big\|\|p^*\|^{-1} x^*\Big\|^{-1} \leq \max\Big\{\|u^*\|^{-1} \Big|\ \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \text{cl}^* C(0)\Big\},$$

which justify the inequality "$\leq$" in (7.18) and thus complete the proof. $\triangle$

7.1.3 Coderivative Characterization of Lipschitzian Stability

In this subsection, we employ the above coderivative analysis married to appropriate techniques in linear SIPs to establish the *coderivative criterion* of Lipschitzian stability (in the sense of the validity of the Lipschitz-like property) for infinite linear systems (7.1) with precise calculation of the *exact Lipschitzian bound* $\text{lip}\,\mathcal{F}(0, \bar{x})$. Surprisingly, the obtained results look exactly like in the finite-dimensional setting of Theorem 3.3 for general closed-graph multifunctions, while in the case here we can express both the coderivative criterion and exact Lipschitzian bound entirely in terms of the given data of (7.1).

First, we present necessary and sufficient condition for the Lipschitz-like property of $\mathcal{F}$ around the reference point $(0, \bar{x})$ in the form of (3.9).

Theorem 7.9 (Coderivative Criterion for the Lipschitz-Like Property of Linear Infinite Systems). *Let $\bar{x} \in \mathcal{F}(0)$ for the infinite inequality system* (7.1). *Then $\mathcal{F}$ is Lipschitz-like around* $(0, \bar{x})$ *if and only if*

$$D^*\mathcal{F}(0, \bar{x})(0) = \{0\}. \tag{7.23}$$

Proof. The "only if" part follows from the proof in Step 1 of Theorem 3.3 valid in arbitrarily Banach spaces. To justify now the "if" part of the theorem, suppose on the contrary that $D^*\mathcal{F}(0, \bar{x})(0) = \{0\}$, while the mapping $\mathcal{F}$ is *not* Lipschitz-like around $(0, \bar{x})$. Then, by the equivalence between properties (i) and (iv) in Theorem 7.2, we get the inclusion

$$(0, 0) \in \text{cl}^*\text{co}\big\{(a_t^*, b_t) \in X^* \times \mathbb{R}\big|\ t \in T\big\}$$

meaning that there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \in \mathbb{R}^{(T)}_+$ with $\sum_{t\in T}\lambda_{t\nu} = 1$, $\nu \in \mathcal{N}$, and

$$w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(a_t^*, b_t) = (0,0). \tag{7.24}$$

Since the net $\{\sum_{t\in T}\lambda_{t\nu}(-\delta_t)\}_{\nu\in\mathcal{N}}$ is obviously bounded in $l^\infty(T)^*$, the Alaoglu-Bourbaki theorem ensures the existence of its subnet (no relabeling) that w^*-converges to some element $p^* \in l^\infty(T)^*$, i.e.,

$$p^* = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(-\delta_t). \tag{7.25}$$

It follows from (7.25) by the Dirac function definition that

$$\langle p^*, -e\rangle = \lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu} = 1, \quad \text{where } \ e = (e_t)_{t\in T} \ \text{ with } \ e_t = 1 \ \text{ for all } \ t \in T,$$

which yields $p^* \neq 0$. Furthermore, combining (7.24) and (7.25) tells us that

$$(p^*, 0, 0) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(-\delta_t, a_t^*, b_t) \ \text{ with } \ p^* \neq 0,$$

and therefore, by the explicit coderivative description of Theorem 7.5, we get the inclusion $p^* \in D^*\mathcal{F}(0,\bar{x})(0)\setminus\{0\}$, which contradicts the assumed condition (7.23). This verifies the sufficiency part of the coderivative criterion (7.23) for the Lipschitz-like property and thus completes the proof of the theorem. △

Our next goal is to calculate the *exact Lipschitzian bound* $\operatorname{lip}\mathcal{F}(0,\bar{x})$. To proceed, observe the following limiting representation of $\operatorname{lip} F(\bar{x},\bar{y})$ via the distance function to a set that holds for any mapping $F\colon X \rightrightarrows Y$:

$$\operatorname{lip} F(\bar{z},\bar{y}) = \limsup_{(z,y)\to(\bar{z},\bar{y})} \frac{\operatorname{dist}\big(y; F(z)\big)}{\operatorname{dist}\big(z; F^{-1}(y)\big)} \quad \text{where } \ 0/0 := 0. \tag{7.26}$$

To begin with, form the closed affine half-space

$$H(x^*,\alpha) := \big\{x \in X \big|\ \langle x^*, x\rangle \le \alpha\big\} \ \text{ for } \ (x^*,\alpha) \in X^* \times \mathbb{R}$$

and derive the distance function representation known as the *Ascoli formula*.

Proposition 7.10 (Ascoli Formula). *We have*

$$\operatorname{dist}\big(x; H(x^*,\alpha)\big) = \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}, \tag{7.27}$$

where $[\gamma]_+ := \max\{\gamma, 0\}$ *for* $\gamma \in \mathbb{R}$ *and* $0/0 := 0$.

Proof. In the case of $x \in H(x^*,\alpha)$, representation (7.27) is obvious. Consider now that case of $x \notin H(x^*,\alpha)$, and define the associated *optimization problem*

$$\text{minimize } \|u - x\| \text{ subject to } u \in H(x^*, \alpha), \tag{7.28}$$

where an optimal solution exists; see Exercise 7.75. Let $\bar{u} \in H(x^*, \alpha)$ be any solution to (7.28). Applying the generalized Fermat rule and then the subdifferential sum rule b valid due to the continuity of $u \mapsto \|u - x\|$ yields

$$0 \in \partial\| \cdot -x\|(\bar{u}) + N\big(\bar{u}; H(x^*, \alpha)\big) \tag{7.29}$$

with $\bar{u} \neq x$. Since we have in this case that

$$\partial\| \cdot -x\|(\bar{u}) = \big\{u^* \in X^* \big|\ \|u^*\| = 1,\ \langle u^*, \bar{u} - x\rangle = \|\bar{u} - x\|\big\}$$

and that $N(\bar{u}; H(x^*, \alpha)) = \text{cone}\{x^*\}$ if $\langle x^*, \bar{u}\rangle = \alpha$ with $N(\bar{u}; H(x^*, \alpha)) = \{0\}$ otherwise, it tells us by (7.29) that

$$\langle x^*, \bar{u}\rangle = \alpha \ \text{ and } \ \|x^*\| \cdot \|\bar{u} - x\| = \langle x^*, x - \bar{u}\rangle.$$

This implies in turn the equalities

$$\|\bar{u} - x\| = \frac{\langle x^*, x\rangle - \langle x^*, \bar{u}\rangle}{\|x^*\|} = \frac{\langle x^*, x\rangle - \alpha}{\|x^*\|} = \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}$$

and thus justifies the Ascoli formula (7.27). △

The next two propositions, which are certainly of their own interest, establish extensions of the Ascoli formula first to the case of *convex* inequalities and then to *infinite* systems of linear inequalities instead of the single one as in (7.27). These results play a significant role in what follows for computing the exact Lipschitzian bound $\text{lip}\,\mathcal{F}(0, \bar{x})$. In their proofs, we use elements of the classical *duality theory* of convex analysis in Banach spaces; see, e.g., [757].

Given a proper (may not be convex) function $\varphi\colon X \to \overline{\mathbb{R}}$, recall that its (always convex) *Fenchel conjugate* $\varphi^*\colon X^* \to \overline{\mathbb{R}}$ is defined by

$$\varphi^*\left(x^*\right) := \sup\big\{\left\langle x^*, x\right\rangle - \varphi\left(x\right) \big|\ x \in X\big\}. \tag{7.30}$$

First, we provide an extension of the Ascoli formula from (single) linear to convex inequalities by using the Fenchel conjugate (7.30).

Proposition 7.11 (Extended Ascoli Formula for Single Convex Inequalities). *Let $g\colon X \to \overline{\mathbb{R}}$ be a (proper) convex function, and let*

$$Q := \big\{y \in X \big|\ g(y) \leq 0\big\}. \tag{7.31}$$

Assume the fulfillment of the classical Slater condition: there is $\widehat{x} \in X$ with $g(\widehat{x}) <$ 0. Then the distance function to the set Q in (7.31) is calculated by

$$\text{dist}(x; Q) = \max_{(x^*, \alpha) \in \text{epi}\, g^*} \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}. \tag{7.32}$$

Proof. Observe that the nonemptiness of Q in (7.31) yields $\alpha \geq 0$ whenever $(0, \alpha) \in \operatorname{epi} g^*$ and that the possibility of $x^* = 0$ is not an obstacle in (7.32) under the convention $0/0 := 0$. The distance function $\operatorname{dist}(x; Q)$ is nothing else but the optimal *value function* in the parametric *convex optimization* problem.

$$\text{minimize} \ \ \|y - x\| \ \ \text{subject to} \ \ g(y) \leq 0. \tag{7.33}$$

Since the Slater condition holds for (7.33) by our assumption, we have the strong Lagrange duality in (7.33) by, e.g., [757, Theorem 2.9.3], which yields

$$\begin{aligned}
\operatorname{dist}(x; Q) &= \max_{\lambda \geq 0} \inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\} \\
&= \max\left\{\max_{\lambda > 0} \inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\},\ \inf_{y \in X} \|y - x\|\right\} \\
&= \max\left\{\max_{\lambda > 0} \inf_{y \in X} \big\{\, \|y - x\| + \lambda g(y)\big\},\ 0\right\}.
\end{aligned}$$

Applying now the classical Fenchel duality theorem to the inner infimum problem above for a fixed $\lambda > 0$, we get

$$\inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\} = \max_{y^* \in X^*} \big\{ - \| \cdot - x\|^*(-y^*) - (\lambda g)^*(y^*)\big\}. \tag{7.34}$$

Furthermore, it is well known in convex analysis that

$$\|\cdot - x\|^* \,(-y^*) = \begin{cases} \langle -y^*, x\rangle & \text{if} \ \ \|y^*\| \leq 1, \\ \infty & \text{otherwise.} \end{cases}$$

Substituting it into formula (7.34) leads us to

$$\begin{aligned}
\inf_{y \in X} \big\{\, \|y - x\| + \lambda g\,(y)\,\big\} &= \max_{\|y^*\| \leq 1} \big\{\, \big\langle y^*, x\big\rangle - (\lambda g)^* \big(y^*\big)\,\big\} \\
&= \max_{\|y^*\| \leq 1,\ (\lambda g)^*(y^*) \leq \eta} \big\{\, \big\langle y^*, x\big\rangle - \eta\big\} \\
&= \max_{\|y^*\| \leq 1,\ \lambda g^*(y^*/\lambda) \leq \eta} \big\{\, \big\langle y^*, x\big\rangle - \eta\big\} \\
&= \max_{\|y^*\| \leq 1,\ (1/\lambda)(y^*, \eta) \in \operatorname{epi} g^*} \big\{\, \big\langle y^*, x\big\rangle - \eta\big\}.
\end{aligned}$$

This ensures, by denoting $x^* := (1/\lambda) y^*$ and $\alpha := (1/\lambda)\eta$, that

$$\inf_{y \in X} \big\{\, \|y - x\| + \lambda g(y)\big\} = \max_{(x^*, \alpha) \in \operatorname{epi} g^*,\ \|x^*\| \leq 1/\lambda} \lambda \big\{\, \big\langle x^*, x\big\rangle - \alpha\big\}.$$

Combining the latter with the formulas above, we arrive at

$$\begin{aligned}\operatorname{dist}(x;Q) &= \max\Big\{\max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\lambda\big\{\langle x^*,x\rangle-\alpha\big\},\ 0\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}.\end{aligned}\tag{7.35}$$

It is easy to observe the following relationships hold for any $\lambda>0$:

$$\begin{aligned}\max_{(0,\alpha)\in\operatorname{epi}g^*}\lambda\big\{\langle 0,x\rangle-\alpha\big\} &= \max_{g^*(0)\le\alpha}\lambda\big(\langle 0,x\rangle-\alpha\big)=\lambda\big(-g^*(0)\big)\\ &\le\lambda\inf_{x\in X}g(x)\le\lambda g(\widehat{x})<0.\end{aligned}$$

Taking this into account, we deduce from (7.35) the equalities

$$\begin{aligned}\operatorname{dist}(x;Q) &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*}\ \max_{\|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|},\end{aligned}$$

which justify (7.32) and thus complete the proof of the proposition. △

The next proposition provides the required extension of the Ascoli formula (7.27) to the case of the *infinite* inequality systems (7.1) in Banach spaces.

Proposition 7.12 (Extended Ascoli Formula for Infinite Linear Systems). *Assume that the infinite linear system* (7.1) *satisfies the strong Slater condition at* $p=(p_t)_{t\in T}$. *Then for any* $x\in X$ *and* $p\in l^\infty(T)$, *we have*

$$\operatorname{dist}\big(x;\mathcal{F}(p)\big)=\max_{(x^*,\alpha)\in\operatorname{cl}^*C(p)}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|}.\tag{7.36}$$

If in addition X *is reflexive, then* (7.36) *can be simplified by*

$$\operatorname{dist}\big(x;\mathcal{F}(p)\big)=\max_{(x^*,\alpha)\in C(p)}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|}.\tag{7.37}$$

Proof. Observe that the infinite linear system (7.1) can be represented as

$$\mathcal{F}(p=\big\{x\in X\big|\ g(x)\le 0\big\},\tag{7.38}$$

where the convex function $g\colon X\to\overline{\mathbb{R}}$ is given in the supremum form

$$g(x):=\sup_{t\in T}\big(f_t(x)-p_t\big)\ \text{ with }\ f_t(x):=\langle a_t^*,x\rangle-b_t.\tag{7.39}$$

The assumed strong Slater condition for $\mathcal{F}(p)$ ensures the validity of the classical Slater condition for g from Proposition 7.11. To employ the result therein in the framework of (7.38), we need to calculate the Fenchel conjugate to the supremum function in (7.39). It can be done by (see Exercise 7.77)

$$\begin{cases} \text{epi}\, g^* = \text{epi}\left\{\sup_{t\in T}(f_t - p_t)\right\}^* = \text{cl}^*\text{co}\left(\bigcup_{t\in T}\text{epi}\left(f_t - p_t\right)^*\right) \\ \qquad\quad = \text{cl}^* C(p) + \mathbb{R}_+(0,1) \ \text{ with } \ 0 \in X^*, \end{cases} \tag{7.40}$$

where the weak* closedness of the set $\text{cl}^* C(p) + \mathbb{R}_+(0,1)$ is a consequence of the classical Dieudonné theorem; see, e.g., [757, Theorem 1.1.8]. Thus we get the distance formula (7.36) from Proposition 7.11 in general Banach spaces.

To justify the simplified distance formula (7.37) in the case of reflexive spaces, suppose on the contrary that it doesn't hold. Then there is a scalar $\beta \in \mathbb{R}$ such that we have the strict inequalities

$$\max_{(x^*,\alpha)\in \text{cl}^* C(p)} \frac{\left[\langle x^*, x\rangle - \alpha\right]_+}{\|x^*\|} > \beta > \sup_{(x^*,\alpha)\in C(p)} \frac{\left[\langle x^*, x\rangle - \alpha\right]_+}{\|x^*\|}. \tag{7.41}$$

This yields the existence of $(\bar{x}^*, \bar{\alpha}) \in \text{cl}^* C(p)$ with $\bar{x}^* \neq 0$ satisfying

$$\frac{\left[\langle \bar{x}^*, x\rangle - \bar{\alpha}\right]_+}{\|\bar{x}^*\|} > \beta.$$

Taking into account that X is reflexive and that $C(p)$ is convex and then employing the Mazur weak closure theorem, we can replace the weak* closure of $C(p)$ by its norm closure in X^*. This allows us to find a sequence $(x_k^*, \alpha_k) \in C(p)$ converging by norm to $(\bar{x}^*, \bar{\alpha})$ as $k \to \infty$. Hence we get

$$\lim_{k\to\infty} \frac{\left[\left\langle x_k^*, x\right\rangle - \alpha_k\right]_+}{\left\|x_k^*\right\|} = \frac{\left[\langle \bar{x}^*, x\rangle - \bar{\alpha}\right]_+}{\|\bar{x}^*\|} > \beta,$$

and therefore there exists $k_0 \in \mathbb{N}$ for which

$$\frac{\left[\left\langle x_{k_0}^*, x\right\rangle - \alpha_{k_0}\right]_+}{\left\|x_{k_0}^*\right\|} > \beta.$$

The latter contradicts (7.41) and thus completes the proof. △

The following example shows that the reflexivity of the decision space X is an essential requirement for the validity of the simplified distance formula (7.37), even in the framework of (nonreflexive) Asplund spaces.

Example 7.13 (Failure of the Simplified Distance Formula in Nonreflexive Asplund Spaces). Consider the classical space c_0 of real number sequences converging to zero and endowed with the supremum norm. This space is known to be Asplund while not reflexive. Let us show that the simplified distance formula (7.37) *fails* in $X = c_0$ (the classical space of sequences converging to zero, with the supremum norm) for a rather plain linear system of *countable* inequalities. Of course, we need to demonstrate that the inequality "$\le$" is generally violated in (7.37), since the opposite inequality holds in any Banach space. Form the infinite (countable) linear inequality system

$$\mathcal{F}(0) := \left\{x \in c_0 \middle| \langle e_1^* + e_t^*, x\rangle \le -1,\ t \in \mathbb{N}\right\}, \tag{7.42}$$

where $e_t^* \in l_1$ has 1 as its t^{th}-component, while all the remaining components are zeros. System (7.42) can be rewritten as

$$x \in \mathcal{F}(0) \Longleftrightarrow x(1) + x(t) \le -1 \ \text{ for all }\ t \in \mathbb{N}.$$

Observe that for $z = 0$, we have $\text{dist}(0; \mathcal{F}(0)) = 1$, and the distance is realized at, e.g., $u = (-1, 0, 0, \ldots)$. Indeed, passing to the limit in $x(1)+x(t) \le -1$ as $t \to \infty$ and taking into account that $x(t) \to 0$ by the structure of the space of c_0, we get $x(1) \le -1$. Furthermore, it can be checked that

$$\begin{aligned}&(e_1^*, -1) \in \text{cl}^*C(0),\ \ \langle e_1^*, x-u\rangle \le 0 \ \text{ for all }\ x \in \mathcal{F}(0),\\ &\text{dist}\big(z; \mathcal{F}(0)\big) = \|z-u\| = \langle e_1^*, z-u\rangle = \frac{\langle e_1^*, z\rangle - (-1)}{\|e_1^*\|}.\end{aligned}$$

On the other hand, for the pair $(x^*, \alpha) \in X^* \times \mathbb{R}$ given by

$$(x^*, \alpha) := \Big(e_1^* + \sum_{t\in\mathbb{N}} \lambda_t e_t^*, -1\Big) \in C(0) \ \text{ with }\ \lambda \in \mathbb{R}_+^{(\mathbb{N})} \ \text{ and }\ \sum_{t\in\mathbb{N}} \lambda_t = 1,$$

we can directly verify that $\|x^*\| = 2$ and hence

$$\frac{\big[\langle x^*, z\rangle - \alpha\big]_+}{\|x^*\|} = \frac{1}{2},$$

which shows that the equality in (7.37) is violated for the countable system (7.42) in the nonreflexive Asplund space $X = c_0$.

Prior to deriving the main result of this subsection on the precise calculation of the exact Lipschitzian bound for the infinite system (7.1) at the reference point, we need the following technical assertion.

Lemma 7.14 (Closed-Graph Property of Characteristic Sets). *The set-valued mapping $l^\infty(T) \ni p \mapsto \text{cl}^*C(p) \subset X^* \times \mathbb{R}$ generated by the characteristic sets (7.5) is closed-graph in the norm×weak* topology of $\ell^\infty(T) \times (X^* \times \mathbb{R})$, i.e., for any nets $\{p_\nu\}_{\nu\in\mathcal{N}} \subset l^\infty(T)$, $\{x_\nu^*\}_{\nu\in\mathcal{N}} \subset X^*$, and $\{\beta_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}$, satisfying the conditions $p_\nu \to p$, $x_\nu^* \xrightarrow{w^*} x^*$, $\beta_\nu \to \beta$, and $\left(x_\nu^*, \beta_\nu\right) \in \text{cl}^*C(p_\nu)$ for every $\nu \in \mathcal{N}$, we have the inclusion $(x^*, \beta) \in \text{cl}^*C(p)$.*

Proof. Arguing by contradiction, suppose that $(x^*, \beta) \notin \text{cl}^* C(p)$. Then the classical strict separation indexconvex separation theorem allows us to find a nonzero pair $(x, \alpha) \in X \times \mathbb{R}$ and real numbers γ and γ' satisfying

$$\langle x^*, x\rangle + \beta\alpha < \gamma' < \gamma \leq \langle a_t^*, x\rangle + (b_t + p_t)\,\alpha \quad \text{for all} \quad t \in T.$$

Hence there exists a net index $\nu_0 \in \mathcal{N}$ such that

$$\langle x_\nu^*, x\rangle + \beta_\nu\alpha < \gamma' \quad \text{and} \quad \|\alpha\,(p - p_\nu)\| \leq \gamma - \gamma' \quad \text{whenever} \quad \nu \succeq \nu_0.$$

This ensures therefore the validity of the estimates

$$\begin{aligned}\langle a_t^*, x\rangle + \alpha\,(b_t + p_{t\nu}) &= \langle a_t^*, x\rangle + \alpha\,(b_t + p_t) + \alpha\,(p_{t\nu} - p_t)\\ &\geq \gamma - \|\alpha\,(p_\nu - p)\| \geq \gamma' \quad \text{for all} \quad t \in T.\end{aligned}$$

The latter implies that $\gamma' \leq \langle z^*, x\rangle + \eta\alpha$ for all $(z^*, \eta) \in \text{cl}^* C(p_\nu)$ whenever $\nu \succeq \nu_0$. Thus we arrive at the contradiction

$$\langle x_\nu^*, x\rangle + \beta_\nu\alpha < \gamma' \leq \langle x_\nu^*, x\rangle + \beta_\nu\alpha, \qquad \nu \succeq \nu_0,$$

which completes the proof of the lemma. △

Now we are ready to provide a precise calculation of the exact Lipschitzian bound of $\mathcal{F}$ around $(0, \bar{x})$ in the general Banach space setting.

Theorem 7.15 (Calculating the Exact Lipschitzian Bound of Infinite Linear Systems). *Let $\bar{x} \in \mathcal{F}(0)$ for the linear infinite inequality system (7.1). Suppose that $\mathcal{F}$ satisfies the strong Slater condition at $p = 0$ and that the coefficient set $\{a_t^* |\ t \in T\}$ is bounded in X^*. The following assertions hold:*

(i) *If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then* $\text{lip}\,\mathcal{F}(0, \bar{x}) = 0$.

(ii) *If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then the exact of $\mathcal{F}$ around $(0, \bar{x})$ is calculated by*

$$\text{lip}\,\mathcal{F}(0, \bar{x}) = \max\left\{\|x^*\|^{-1} \,\middle|\, \left(x^*, \langle x^*, \bar{x}\rangle\right) \in \text{cl}^* C(0)\right\} > 0 \tag{7.43}$$

via the w^*-closure of the characteristic set (7.5) at $p = 0$.

Proof. To verify (i), recall from the proof of Theorem 7.8(i) that the assumptions made imply that $(0, \bar{x}) \in \text{int}(\text{gph}\,\mathcal{F})$, which in turn yields $\text{lip}\,\mathcal{F}(0, \bar{x}) = 0$ by the definition of the exact Lipschitzian bound.

Next we justify the more difficult assertion (ii) of the theorem while assuming that $\bar{x}$ is *not* a strong Slater point for $\mathcal{F}$ at $p = 0$. Observe that by Proposition 7.7, the set under the maximum operation on the right-hand side in (7.43) is nonempty and w^*-compact in X^*. Thus the maximum over this set is realized and is finite. The inequality "$\geq$" in (7.43) follows from the estimate

$$\text{lip}\,\mathcal{F}(0, \bar{x}) \geq \|D^*\mathcal{F}(0, \bar{x})\|$$

taken from Exercise 3.35(i) and then combined with formula (7.18) for calculating the coderivative norm of the infinite inequality system (7.1) derived in Theorem 7.8. Thus it remains to verify the opposite inequality "$\leq$" in (7.43).

To proceed, let $M := \sup_{t\in T}\left\|a_t^*\right\| < \infty$, and observe that the inequality "$\leq$" in (7.43) is obvious when $L := \operatorname{lip}\mathcal{F}(0,\bar{x}) = 0$. Suppose now that $L > 0$, and consider any pair (p, x) sufficiently close to $(0, \bar{x})$ in representation (7.26) of the exact Lipschitzian bound $\operatorname{lip}\mathcal{F}(0,\bar{x})$. By $L > 0$, we can confine ourselves to the case of $(p, x) \notin \operatorname{gph}\mathcal{F}$. It follows from the structure of $\mathcal{F}$ that

$$0 < \operatorname{dist}\bigl(p; \mathcal{F}^{-1}(x)\bigr) = \sup_{t\in T}\bigl[\langle a_t^*, x\rangle - b_t - p_t\bigr]_+. \tag{7.44}$$

Moreover, we have the relationships

$$\begin{aligned}\bigl\langle a_t^*, x\bigr\rangle - b_t - p_t &= \bigl\langle a_t^*, x-\bar{x}\bigr\rangle + \bigl\langle a_t^*, \bar{x}\bigr\rangle - b_t - p_t\\ &\leq M\,\|x-\bar{x}\| + \|p\| \quad \text{for all} \;\; t\in T,\end{aligned}$$

which allow us to conclude that

$$\begin{aligned}0 &< \sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\bigl[\bigl\langle x^*, x\bigr\rangle - \beta\bigr]_+ = \sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\bigl\{\bigl\langle x^*, x\bigr\rangle - \beta\bigr\}\\ &\leq M\,\|x-\bar{x}\| + \|p\| \quad \text{for all} \;\; x\in X \;\; \text{and} \;\; p\in P.\end{aligned} \tag{7.45}$$

Consider further the set

$$C_+(p,x) := \bigl\{\bigl(x^*,\beta\bigr)\in \mathrm{cl}^*C(p) \bigm| \bigl\langle x^*, x\bigr\rangle - \beta > 0\bigr\},$$

which is obviously nonempty, and denote

$$M_{(p,x)} := \sup\bigl\{\,\bigl\|x^*\bigr\|^{-1} \bigm| \bigl(x^*,\beta\bigr)\in C_+(p,x)\,\bigr\}.$$

In our setting, we get $0 \in \operatorname{int}(\operatorname{dom}\mathcal{F})$ (see Exercise 7.72(i)) and therefore $p \in \operatorname{dom}\mathcal{F}$ for all $p \in l^\infty(T)$ sufficiently close to the origin. In this case, the set $C_+(p,x)$ cannot contain any element of the form $(0,\beta)$, since the contrary would yield $\beta < 0$ by the definition of $C_+(p,x)$, while Proposition 7.3 tells us that $\beta \geq 0$. Thus we conclude that $0 < \|x^*\| \leq M$ whenever $(x^*,\beta) \in C_+(p,x)$ and, in particular, $M_{(p,x)} \in (0,\infty]$. It follows furthermore that

$$\frac{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)} \dfrac{\bigl[\langle x^*, x\rangle - \beta\bigr]_+}{\|x^*\|}}{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\bigl[\langle x^*, x\rangle - \beta\bigr]_+} = \frac{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)} \dfrac{\langle x^*, x\rangle - \beta}{\|x^*\|}}{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\bigl\{\langle x^*, x\rangle - \beta\bigr\}} \leq M_{(p,x)},$$

where the latter inequality ensures the estimate

$$L \leq \limsup_{(p,x)\to(0,\bar{x}),\; x\notin\mathcal{F}(p)\neq\emptyset} M_{(p,x)} := K.$$

Considering next a sequence $(p_k, x_k) \to (0, \bar{x})$ with $x_k \notin \mathcal{F}(p_k) \neq \emptyset$ and

$$L \leq \lim_{k\to\infty} M_{(p_k,x_k)} = K,$$

we select a sequence $\{\alpha_k\}_{k=1}^{\infty} \subset \mathbb{R}$ such that

$$\lim_{k\to\infty} \alpha_k = K \text{ and } 0 < \alpha_k < M_{(p_k,x_k)} \text{ as } k \in \mathbb{N}.$$

Take now $\left(x_k^*, \beta_k\right) \in C_+(p_k, x_k)$ with $\alpha_k < \left\|x_k^*\right\|^{-1}$ for all $k \in \mathbb{N}$. Since the sequence $\left\{x_k^*\right\}_{k\in\mathbb{N}} \subset X^*$ is bounded, it contains a subnet $\left\{x_\nu^*\right\}_{\nu\in\mathcal{N}}$ that w^*-converges to some $x^* \in X^*$. Denoting by $\{p_\nu\}$, $\{x_\nu\}$, $\{\beta_\nu\}$, and $\{\alpha_\nu\}$ the corresponding subnets of $\{p_k\}$, $\{x_k\}$, $\{\beta_k\}$, and $\{\alpha_k\}$, we get from (7.45) that

$$0 < \left\langle x_\nu^*, x_\nu\right\rangle - \beta_\nu \leq M\left\|x_\nu - \bar{x}\right\| + \left\|p_\nu\right\|.$$

Hence $\left\langle x_\nu^*, x_\nu\right\rangle - \beta_\nu \to 0$, which implies by the constructions above that $\beta_\nu \to \langle x^*, \bar{x}\rangle$. We deduce from Lemma 7.14 that

$$\left(x^*, \left\langle x^*, \bar{x}\right\rangle\right) \in \mathrm{cl}^* C\,(0)\,,$$

and then Theorem 7.2 ensures that $x^* \neq 0$.

To finalize verifying the inequality "$\leq$" in (7.43), observe that

$$\left\|x^*\right\| \leq \liminf_{\nu\in\mathcal{N}} \left\|x_\nu^*\right\| \leq \lim_{\nu\in\mathcal{N}} \frac{1}{\alpha_\nu} = \frac{1}{K}$$

due to $\left\|x_\nu^*\right\| \leq \alpha_\nu^{-1}$ and $\lim_{\nu\in\mathcal{N}} \alpha_\nu = K$, which gives us

$$L \leq K \leq \frac{1}{\|x^*\|} \leq \max\left\{ \left\|z^*\right\|^{-1} \,\middle|\, \left(z^*, \left\langle z^*, \bar{x}\right\rangle\right) \in \mathrm{cl}^* C\,(0) \right\}.$$

Remembering the notation above, we complete the proof of the theorem. △

Summarizing the obtained results on the calculations of the coderivative norm in Theorem 7.8 and the exact Lipschitzian bound in Theorem 7.15 allows us to arrive at the *unconditional* relationship between these quantities for the infinite linear inequality system $\mathcal{F}$ with an arbitrary Banach decision space X that is expressed by the same formula as the one (3.10) derived in Theorem 3.3 for set-valued mappings between finite-dimensional spaces.

Corollary 7.16 (Relationship Between the Exact Lipschitzian Bound and Coderivative Norm). *Let $\bar{x} \in \mathcal{F}(0)$ for the infinite system (7.1) satisfying the strong Slater condition at $p = 0$, and let the coefficient set $\{a_t^* |\ t \in T\}$ be bounded in X^*. Then we have the equality*

$$\mathrm{lip}\,\mathcal{F}(0, \bar{x}) = \|D^*\mathcal{F}(0, \bar{x})\|. \tag{7.46}$$

Proof. If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then we get equality (7.46) by comparing assertions (i) in Theorem 7.8 and Theorem 7.15, which yield

$$\operatorname{lip}\mathcal{F}(0,\bar{x}) = \|D^*\mathcal{F}(0,\bar{x})\| = 0.$$

If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then (7.46) follows from comparing assertions (ii) in Theorem 7.8 and Theorem 7.15, which give us the same formula for calculating both $\|D^*\mathcal{F}(0,\bar{x})\|$ and $\operatorname{lip}\mathcal{F}(0,\bar{x})$. △

7.2 Optimization Under Infinite Linear Constraints

In this section, we derive necessary optimality conditions for SIPs with general nonsmooth cost functions over feasible solution sets governed by infinite linear constraint systems of type (7.1). The calculation of the coderivative of the feasible solution map given in Section 7.1 plays a crucial role in deriving necessary optimality conditions of both upper and lower subdifferential types presented below. The results obtained are then applied to solving an optimization problem of a practical interest arising in water resource modeling.

7.2.1 Two-Variable SIPs with Infinite Inequality Constraints

We deal here with the following SIP problem:

$$\text{minimize } \varphi(p,x) \text{ subject to } x \in \mathcal{F}(p), \tag{7.47}$$

where $\varphi\colon P \times X \to \overline{\mathbb{R}} := (-\infty,\infty]$ is an extended-real-valued *cost* function (generally nonsmooth and nonconvex) defined on the product of *Banach* spaces and where $\mathcal{F}\colon P \rightrightarrows X$ is a set-valued mapping of *feasible solutions*

$$\mathcal{F}(p) := \left\{x \in X \,\middle|\, \langle a_t^*, x\rangle \le b_t + \langle c_t^*, p\rangle,\quad t \in T\right\} \tag{7.48}$$

with an *arbitrary* (possibly infinite) *index set* T and with some fixed elements $a_t^* \in X^*$, $c_t^* \in P^*$, and $b_t \in \mathbb{R}$ for all $t \in T$. Note that our considerations in Section 7.1, conducted mainly from the viewpoints of Lipschitzian stability of parametric mappings $\mathcal{F}(p)$, concern the case of (7.48) with $P = l^\infty(T)$ and $c_t^* = \delta_t$ (Dirac measure), but the coderivative calculation given therein can be easily adapted to the case of (7.48).

Observe that the optimization in (7.47) is taken with respect to both variables (p,x), which are interconnected through the infinite inequality system (7.48). This means in fact that we have *two groups of decision variables* represented by x and p. One player specifies p, and the other solves (7.47) in x subject to (7.48) with the specified p as a parameter. The first one, having the same objective, varies his/her parameter p to get the best outcome via the so-called *optimistic approach*. We could treat this as a *two-level design*: optimizing the basic parameter p at the upper level, while at the lower level, the cost function is optimized with respect to x for the given p. The reader is referred to, e.g., [442], and the bibliography therein for various tuning and tolerancing problems of such types arising in engineering design. Another area where two-variable SIPs governed by (7.47) and (7.48) with Banach decision spaces X and P naturally appear concerns optimization of water resources. A practical problem of this type is introduced and studied in Subsection 7.2.4.

We can notice some similarity between the two-variable optimization problem in (7.47) and (7.48), treated above as a two-level optimistic design, and the optimistic model of bilevel programming that was considered in Chapter 6 for finitely many constraints and will be studied in Section 7.5.4 for infinitely many ones. The main difference between these classes is that (7.48) is a *constraint system* described by finitely many or infinitely many inequalities, while the corresponding parameter-dependent set $S(\cdot)$ at the upper level of bilevel programming is given by a *variational system* of optimal solutions to a lower-level problem of parametric optimization.

Keeping the same notation as in Section 7.1, we proceed now with deriving two types of necessary optimality conditions for the SIP given in (7.47) and (7.48).

7.2.2 Upper Subdifferential Optimality Conditions for SIPs

Let us begin with upper subdifferential optimality conditions for problem (7.47) and (7.48) that utilize the upper regular subdifferential (6.2) of the cost function (7.47) along with the precise coderivative calculation for the infinite inequality constraint system in (7.48).

Recall the well-known *Farkas-Minkowski property* for (7.48) that amounts to saying that the conic convex hull

$$\operatorname{cone}\big\{(-c_t^*, a_t^*, b_t) \in P^* \times X^* \times \mathbb{R}\,\big|\; t \in T\}\tag{7.49}$$

is weak* closed in the dual space $P^* \times X^* \times \mathbb{R}$.

Now we are ready to formulate and prove upper subdifferential necessary optimality conditions for the SIP in (7.47) and (7.48) in general Banach spaces.

Theorem 7.17 (Upper Subdifferential Conditions for SIPs with Linear Inequality Constraints). *Let* $(\bar{p}, \bar{x}) \in \operatorname{gph}\mathcal{F} \cap \operatorname{dom}\varphi$ *be a local minimizer for the two-variable SIP given by* (7.47) *and* (7.48). *Then every upper regular subgradient* $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$ *satisfies the asymptotic optimality condition*

$$-\big(p^*, x^*, \langle p^*, \bar{p}\rangle + \langle x^*, \bar{x}\rangle\big) \in \operatorname{cl}^*\operatorname{cone}\Big\{(-c_t^*, a_t^*, b_t)\,\Big|\; t \in T\Big\}.\tag{7.50}$$

If furthermore the Farkas-Minkowski property (7.49) *holds for* (7.48), *then* (7.50) *can be equivalently written in the upper subdifferential KKT form: for every* $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, *there are multipliers* $\lambda = (\lambda_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ *satisfying*

$$(p^*, x^*) + \sum_{t \in T(\bar{p},\bar{x})} \lambda_t(-c_t^*, a_t^*) = 0, \tag{7.51}$$

where $\mathbb{R}_+^{(T)}$ *is defined in* (7.3) *and where*

$$T(\bar{p}, \bar{x}) := \big\{ t \in T \big|\ \langle a_t^*, \bar{x}\rangle - \langle c_t^*, \bar{p}\rangle = b_t \big\}. \tag{7.52}$$

Proof. Pick any $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, and, employing the first part of Theorem 1.27 which holds in arbitrary Banach spaces (see Exercise 1.64), construct a function $s\colon P \times X \to \mathbb{R}$ such that

$$s(\bar{p}, \bar{x}) = \varphi(\bar{p}, \bar{x}), \quad \varphi(p, x) \le s(p, x) \ \text{ for all } \ (p, x) \in P \times X, \tag{7.53}$$

and $s(\cdot)$ is Fréchet differentiable at $(\bar{p}, \bar{x})$ with $\nabla s(\bar{p}, \bar{x}) = (p^*, x^*)$. Taking into account that $(\bar{p}, \bar{x})$ is a local minimizer in (7.47), (7.48) and that

$$s(\bar{p}, \bar{x}) = \varphi(\bar{p}, \bar{x}) \le \varphi(p, x) \le s(p, x) \ \text{ for all } \ (p, x) \in \operatorname{gph} \mathcal{F} \ \text{ near } \ (\bar{p}, \bar{x})$$

by (7.53), we deduce that $(\bar{p}, \bar{x})$ is a local minimizer for the auxiliary problem

$$\text{minimize} \ \ s(p, x) \ \text{ subject to } \ (p, x) \in \operatorname{gph} \mathcal{F} \tag{7.54}$$

with the objective $s(\cdot)$ that is Fréchet differentiable at $(\bar{p}, \bar{x})$. Rewriting (7.54) in the infinite-penalty unconstrained form

$$\text{minimize} \ \ s(p, x) + \delta\big((p, x); \operatorname{gph} \mathcal{F}\big)$$

via the indicator function of $\operatorname{gph} \mathcal{F}$, observe directly from definition (1.33) of the regular subdifferential at a local minimizer that

$$(0, 0) \in \widehat{\partial}\big[s + \delta(\cdot; \operatorname{gph} \mathcal{F})\big](\bar{p}, \bar{x}). \tag{7.55}$$

Since $s(\cdot)$ is Fréchet differentiable at $(\bar{p}, \bar{x})$, we easily get from (7.55) that

$$(0, 0) \in \nabla s(\bar{p}, \bar{x}) + N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big),$$

which implies by $\nabla s(\bar{p}, \bar{x}) = (p^*, x^*)$ and the coderivative definition (1.15) that $-p^* \in D^*\mathcal{F}(\bar{p}, \bar{x})(x^*)$. It follows from the proof of Theorem 7.5 that the latter coderivative condition can be constructively described in terms of the initial problem data as follows:

$$\big(-p^*, -x^*, -(\langle p^*, \bar{p}\rangle + \langle x^*, \bar{x}\rangle)\big) \in \text{cl}^*\text{cone}\big\{(-c_t^*, a_t^*, b_t) \big|\ t \in T\big\}. \tag{7.56}$$

Thus (7.56) justifies the asymptotic condition (7.50) for the given upper subgradient $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$. If the Farkas-Minkowski property (7.49) is satisfied, then the operation cl^* in (7.50) can be omitted, and we arrive at the KKT condition (7.51) while completing the proof of the theorem. △

The essence of upper subdifferential conditions in the general framework of minimization has been discussed above in Remark 6.2, which equally applies to the SIP setting of Theorem 7.17. The following consequence of the obtained results is used in Subsection 7.2.4 when both spaces X and P are Banach.

Corollary 7.18 (Necessary Conditions for SIPs with Fréchet Differentiable Costs). *In the setting of Theorem 7.17, suppose that the cost function φ is Fréchet differentiable at the local optimal solution $(\bar{p}, \bar{x})$ with the derivative $(p^*, x^*) = \nabla\varphi(\bar{p}, \bar{x})$. Then (7.50) holds and further reduces to (7.51) if in addition system (7.48) enjoys the Farkas-Minkowski property.*

Proof. It follows directly from Theorem 7.17 since in this case we have $\widehat{\partial}^+\varphi(\bar{p}, \bar{x}) = \{\nabla\varphi(\bar{p}, \bar{x})\}$ for the regular upper subdifferential of φ. △

Observe that in the general settings of Theorem 7.17 and Corollary 7.18, the necessary optimality condition (7.50) is obtained in the *normal form* meaning that we have a nonzero ($\lambda_0 = 1$) multiplier associated with the cost function without any constraint qualification. However, this condition is expressed in the *asymptotic form* involving the weak* closure of the set on the right-hand side of (7.50). This feature partly relates to considering arbitrary index sets in the SIP constraint (7.48) but may also be exhibited in problems with compact index sets as shown in Subsection 7.2.4.

The latter phenomenon doesn't appear under the validity of Farkas-Minkowski property (7.49), which ensures the more conventional KKT form (7.51). Let us present another consequence of Theorem 7.17, where the Farkas-Minkowski property holds and gives us KKT (7.51).

To proceed, we need the following adaptation of the *strong Slater condition* (SSC) from Definition 7.1 to the case of the constraint system (7.48): SSC holds for (7.48) if there is a pair $(\widehat{p}, \widehat{x}) \in P \times X$ such that

$$\sup_{t\in T}\big[\langle a_t^*, \widehat{x}\rangle - \langle c_t^*, \widehat{p}\rangle - b_t\big] < 0. \tag{7.57}$$

The reader can easily check the validity of the equivalent descriptions of SSC for (7.48) similar to those given in Theorem 7.2.

Corollary 7.19 (Upper Subdifferential Conditions in KKT Form). *Suppose that T is a compact Hausdorff space, that both X and P are finite-dimensional, that the mapping $t \mapsto (a_t^*, c_t^*, b_t)$ is continuous on T, and that SSC (7.57) holds. Then for any $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, there are multipliers $\lambda = (\lambda_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ such that the KKT condition (7.51) is satisfied.*

Proof. To check the fulfillment of the Farkas-Minkowski property under the assumptions imposed in the corollary, we observe first that the boundedness and closedness of the set $\{(c_t^*, a_t^*, b_t)|\ t \in T\}$ (and hence of its convex hull by the classical Carathéodory theorem) follow from the continuity of $t \mapsto (c_t^*, a_t^*, b_t)$ and compactness of T. Using this boundedness and the equivalence (ii)⇔(iii) in the counterpart of Theorem 7.2 for (7.48) gives us the condition

$$(0,0,0) \notin \mathrm{co}\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.58}$$

As well known in convex analysis (see, e.g., [667, Corollary 9.6.1]), the validity of (7.58) in this setting yields the closedness of the convex conic hull of $\{-c_t^*, a_t^*, b_t^*)|\ t \in T\}$, and thus the Farkas-Minkowski property holds. △

7.2.3 Lower Subdifferential Optimality Conditions for SIPs

Now we turn to lower subdifferential optimality conditions for the SIP under consideration, which use the basic subgradients (1.24) of the cost function φ in (7.47). Our standing assumption in this subsection is that both spaces X and P are *Asplund*. Recall also that the lower semicontinuity of φ, which is the standing assumption in this book, is essential here, while it is not needed for the upper subdifferential results of Subsection 7.2.2.

The lower subdifferential conditions for the SIP in (7.47) and (7.48) derived below differ from their upper subdifferential counterparts in assumptions as well as in conclusions even for the case of finite-dimensional decision spaces. Observe that the following theorem utilizes both basic (1.24) and singular (1.25) subgradients of the cost function.

Theorem 7.20 (Lower Subdifferential Conditions for SIPs with Linear Inequality Constraints). *Let* $(\bar p, \bar x) \in \mathrm{gph}\,\mathcal{F} \cap \mathrm{dom}\,\varphi$ *be a local minimizer for the SIP under consideration. Suppose also that:*

(a) *either* φ *is locally Lipschitzian around* $(\bar p, \bar x)$*;*

(b) *or* $\mathrm{int}(\mathrm{gph}\,\mathcal{F}) \neq \emptyset$ (*which is true, in particular, when SSC* (7.57) *holds and the set* $\{(a_t^*, c_t^*)|\ t \in T\}$ *is bounded in* $X^* \times P^*$) *and the system*

$$\begin{gathered}(p^*, x^*) \in \partial^\infty \varphi(\bar p, \bar x),\\ -\big(p^*, x^*, \langle (p^*, x^*), (\bar p, \bar x)\rangle\big) \in \mathrm{cl}^*\mathrm{cone}\,\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\}\end{gathered} \tag{7.59}$$

admits only the trivial solution $(p^*, x^*) = (0,0)$.

Then there is a basic subgradient pair $(p^*, x^*) \in \partial\varphi(\bar p, \bar x)$ *satisfying the asymptotic optimality condition* (7.50)*. If in addition the Farkas-Minkowski property* (7.49) *holds for* (7.48)*, then there are subgradients* $(p^*, x^*) \in \partial\varphi(\bar p, \bar x)$ *and multipliers* $\lambda = (\lambda_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ *satisfying the KKT condition*

$$(p^*, x^*) + \sum_{t \in T(\bar p, \bar x)} \lambda_t(-c_t^*, a_t^*) = 0 \tag{7.60}$$

with the active index set $T(\bar p, \bar x)$ *defined in* (7.52)*.*

Proof. The SIP in (7.47) and (7.48) can be equivalently written as

$$\text{minimize}\ \ \varphi(p,x) + \delta\big((p,x); \mathrm{gph}\,\mathcal{F}\big). \tag{7.61}$$

Applying the generalized Fermat rule to $(\bar{p}, \bar{x})$ in (7.61) gives us

$$(0, 0) \in \partial\big[\varphi + \delta(\cdot; \operatorname{gph} \mathcal{F})\big](\bar{p}, \bar{x}) \tag{7.62}$$

in terms of the basic subdifferential of the summation function in (7.62). By using an Asplund space version of the subdifferential sum rule from Exercise 2.54(i) with taking into account that the product of two Asplund spaces is Asplund and that the SNC property holds for solid convex sets by Exercise 2.29(ii), we deduce from (7.62) the validity of the inclusion

$$(0, 0) \in \partial\varphi(\bar{p}, \bar{x}) + N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big) \tag{7.63}$$

provided that either φ is locally Lipschitzian around $(\bar{p}, \bar{x})$ as assumed in (a) or the interior of gph $\mathcal{F}$ is nonempty and the qualification condition

$$\partial^\infty\varphi(\bar{p}, \bar{x}) \cap \big[-N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big)\big] = \{(0, 0)\} \tag{7.64}$$

is satisfied as assumed in (b). It follows from the proof of Theorem 7.2 that the strong Slater condition (7.57) and the boundedness of $\{(a_t^*, c_t^*)|\ t \in T\}$ surely imply that the interior of gph $\mathcal{F}$ is nonempty. Using now the coderivative description obtained in Theorem 7.5 while modifying it for the case of $\mathcal{F}$ from (7.48) shows that the qualification condition (7.64) can be equivalently written as the triviality of solutions to system (7.59) imposed above. In the same way, we reduce (7.63) to the validity of (7.50) for some $(p^*, x^*) \in \partial\varphi(\bar{p}, \bar{x})$. If furthermore the Farkas-Minkowski property (7.49) is satisfied for (7.48), then the operation cl* in (7.50) can be omitted. Thus we arrive at the KKT condition (7.60) and complete the proof of the theorem. △

Similarly to Subsection 7.2.2, we can derive from Theorem 7.20 the lower subdifferential counterpart of Corollary 7.19. Observe that the corresponding consequence of Theorem 7.20 involving an appropriate differentiability of the cost function in (7.47) holds under more restrictive assumptions in comparison with Corollary 7.18: besides the Asplund property of X and P, we have to assume the strict differentiability of φ at $(\bar{p}, \bar{x})$.

7.2.4 Applications to Water Resource Optimization

This subsection provides applications of the obtained general results for SIPs to a water resource optimization problem of a practical interest. We formulate the water recourse model and reduce it to a two-variable SIP over a compact index set with Banach decision spaces. The usage of the necessary optimality conditions for such problems established above allows us to determine optimal decision strategies and suggest efficient ways of their realizations.

The *water resource problem* under consideration is inspired by a continuous-time network flow model formulated in [15]. Consider a system of n reservoirs $R_1, \ldots, R_n$ from which a time-varying water demand is required during a fixed continuous-time period $T = [\underline{t}, \overline{t}]$. Let c_i be the capacity of the reservoir R_i, and let water flow into R_i at rate $r_i(t)$ for each $i = 1, \ldots, n$ and $t \in T$. Denote by $D(t)$ the rate of water demand at t, and suppose that all these nonnegative functions $r_1, \ldots, r_n$ and D are piecewise continuous on the compact interval T and are known in advance. If there is enough water to fill all the reservoir capacity, then the rest can be sold to a neighboring dry area provided that the demand is satisfied. Conversely, if the inflows are short and the reservoirs have free capability for holding additional water, then some water can be bought from outside to meet the inner demand in the region; see Fig. 7.1.

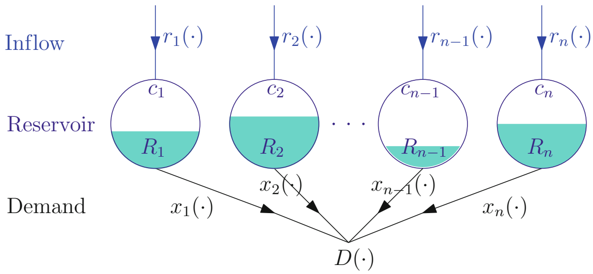

Fig. 7.1 Reservoirs.

Denote by $x_i(t)$ the rate at which water is fed from the reservoir R_i at time $t \in T$. It is natural to assume in our basic model that $x_i \in \mathcal{C}(T)$ for all $i = 1, \ldots, n$. The *feeder constraints* can be expressed by

$$0 \leq x_i(t) \leq \eta_i, \quad i = 1, \ldots, n, \tag{7.65}$$

with fixed bounds $\eta_i \geq 0$. The *selling rate* of water from the reservoir R_i at time t is given by $dp_i(t)$, which means that $p_i(t)$ is the quantity of water sold until instant t and depending on t continuously on the time interval T. Without loss of generality, suppose that $p_i\left(\underline{t}\right) = 0$ for all $i = 1, \ldots, n$. Observe that we are actually buying water at time $t \in T$ if the selling rate $dp_i(t)$ is negative. Denoting further by $s_i \geq 0$ the amount of water initially stored in R_i, we describe the *storage constraints* by

$$\begin{aligned} 0 &\leq \int_{\underline{t}}^{t} \big[r_i(\tau) - x_i(\tau)\big]\, d\tau - \int_{\underline{t}}^{t} dp_i(\tau) + s_i \\ &= \int_{\underline{t}}^{t} \big[r_i(\tau) - x_i(\tau)\big]\, d\tau - p_i(t) + s_i \\ &\leq c_i \quad \text{for all} \quad t \in T \quad \text{and} \quad i = 1, \ldots, n \end{aligned} \tag{7.66}$$

and arrive at the following problem of *water resource optimization*:

$$\begin{cases} \text{minimize } \ \varphi(p,x) \ \text{ subject to } \ (7.65), (7.66), \\ \text{and } \displaystyle\sum_{i=1}^{n} x_i(t) \geq D(t) \ \text{ for all } \ t \in T, \end{cases} \tag{7.67}$$

where the cost function $\varphi(p, x)$ is determined by the cost of water, environmental requirements in the region, and the technology of reservoir processes in the water resource problem. It is clear that we should impose the relationship

$$D(t) \leq \sum_{i=1}^{n} \eta_i, \quad t \in T,$$

in order to ensure the consistency of the constraints in (7.67).

Let us show that problem (7.67) can be reduced to the SIP form in (7.47), (7.48) with two groups of variables $(p, x) \in \mathcal{C}(T)^n \times \mathcal{C}(T)^n$. To proceed, define the following t-parametric families of functions on T:

$$\delta_t(\tau) := \begin{cases} 0 & \text{if } \ \underline{t} \leq \tau < t, \\ 1 & \text{otherwise;} \end{cases} \qquad \alpha_t(\tau) := \begin{cases} \tau & \text{if } \ \underline{t} \leq \tau < t, \\ t & \text{otherwise.} \end{cases}$$

Both families $\{\delta_t |\ t \in T\}$ and $\{\alpha_t |\ t \in T\}$ are subsets of the dual space $\mathcal{C}(T)^*$. In fact, the Riesz representation theorem ensures that each function $\gamma\colon T \to \mathbb{R}$ of bounded variation on T determines a linear functional on $\mathcal{C}(T)$ by

$$z \mapsto \langle \gamma, z \rangle := \int_{\underline{t}}^{\overline{t}} z(\tau)\, d\gamma(\tau), \quad z \in \mathcal{C}(T),$$

via the Stieltjes integral. It is easy to check that

$$\int_{\underline{t}}^{t} x_i(\tau)\, d\tau = \langle \alpha_t, x_i \rangle, \quad d\alpha_t(\tau) = \chi_{[\underline{t},t]}(\tau)\, d\tau \ \text{ for } \ t \in T,$$

where $\chi_{[\underline{t},t]}$ is the standard characteristic function of the interval $[\underline{t}, t]$. Moreover, for each element $z \in \mathcal{C}(T)$, we have

$$\langle \delta_t, z \rangle = z\,(t)\,, \quad t \in T,$$

and thus δ_t can be identified in this context with the Dirac measure at t, which justifies the δ-notation above. Consider further the functions

$$\beta_i(t) := \int_{\underline{t}}^{t} r_i(\tau)\, d\tau \ \text{ for } \ i = 1, \ldots, n, \quad t \in T,$$

and notice that the constraints in (7.66) can be rewritten as

$$\begin{cases} \langle \delta_t, p_i \rangle + \langle \alpha_t, x_i \rangle \le \beta_i(t) + s_i, \\ -\langle \delta_t, p_i \rangle - \langle \alpha_t, x_i \rangle \le c_i - s_i - \beta_i(t), \end{cases} \tag{7.68}$$

while the one in (7.67) admits the form

$$\sum_{i=1}^{n} \langle \delta_t, x_i \rangle \ge D(t), \quad t \in T. \tag{7.69}$$

Observing finally that the constraints in (7.65) can be equivalently given by

$$0 \le \langle \delta_t, x_i \rangle \le \eta_i, \quad i = 1, \ldots, n, \quad t \in T, \tag{7.70}$$

we arrive at the following reduction result.

Proposition 7.21 (Water Resource Problem as SIP in Banach Spaces). *The problem of water resource optimization* (7.67) *is equivalent to the two-variable SIP of type* (7.47) and (7.48) *in the space* $\mathcal{C}(T) \times \mathcal{C}(T)$*:*

$$\textit{minimize} \;\; \varphi(p, x) \;\; \textit{subject to} \;\; (7.68), (7.69), \;\; \textit{and} \;\; (7.70) \tag{7.71}$$

with the data δ_t, α_t, β_t, c_i, s_i, η_i, *and* D *defined above.*

Now we examine the possibility to apply the obtained necessary optimality conditions for SIPs to the case of the water resource model (7.71). Since the space $\mathcal{C}(T)$ for both variables x and p in our model is not Asplund, we proceed with applying the upper subdifferential optimality conditions of Theorem 7.17 and consider for definiteness the case where the cost function φ is Fréchet differentiable at the reference point, i.e., apply the optimality conditions of Corollary 7.18. For simplicity of notation, suppose in what follows that $n = 1$ in (7.71), and write $(p, x, \beta, c, s, \eta)$ instead of $(p_1, x_1, \beta_1, c_1, s_1, \eta_1)$.

Using the initial data of problem (7.71), define the following convex conic hull in the dual space $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ by

$$K(T) := \text{cone}\left\{ \begin{matrix} \Big[\big(\delta_t, \alpha_t, \beta(t) + s\big), \big(-\delta_t, -\alpha_t, c - s - \beta(t)\big), \\ \big(0, -\delta_t, -D(t)\big), \big(0, \delta_t, \eta\big) \;\text{ over all }\; t \in T\Big] \end{matrix} \right\}, \tag{7.72}$$

which is a specification of (7.49) for the water recourse problem (7.71). Given a solution pair $(\bar{p}, \bar{x})$, consider the sets of *active indices* corresponding to all the inequality constraints in (7.71) formed as

$$\begin{cases} T_1(\bar{p}, \bar{x}) := \big\{t \in T \;\big|\; \langle \delta_t, \bar{p} \rangle + \langle \alpha_t, \bar{x} \rangle = \beta(t) + s\big\}, \\ T_2(\bar{p}, \bar{x}) := \big\{t \in T \;\big|\; -\langle \delta_t, \bar{p} \rangle - \langle \alpha_t, x \rangle = c - s - \beta(t)\big\}, \\ T_3(\bar{p}, \bar{x}) := \big\{t \in T \;\big|\; -\langle \delta_t, x \rangle = -D(t)\big\}, \\ T_4(\bar{p}, \bar{x}) := \big\{t \in T \;\big|\; \langle \delta_t, \bar{x} \rangle = \eta\big\}. \end{cases} \tag{7.73}$$

The next result provides necessary conditions for local minimizers in the water recourse optimization problem (7.71).

Proposition 7.22 (Necessary Optimality Conditions for Water Resource Optimization). *Let $(\bar{p}, \bar{x})$ be a local minimizer in problem (7.71). Assume that the cost function $\varphi\colon \mathcal{C}(T) \times \mathcal{C}(T) \to \overline{\mathbb{R}}$ is Fréchet differentiable at $(\bar{p}, \bar{x})$, and consider the cone $K(T)$ defined in (7.72). Then we have the inclusion*

$$-\big(\nabla_p\varphi(\bar{p}, \bar{x}), \nabla_x\varphi(\bar{p}, \bar{x}), \langle\nabla_p\varphi(\bar{p}, \bar{x}), \bar{p}\rangle + \langle\nabla_x\varphi(\bar{p}, \bar{x}), \bar{x}\rangle\big) \in \mathrm{cl}^* K(T).$$

If furthermore the cone $K(T)$ is weak closed, then there exist generalized multipliers $\lambda = (\lambda_t)_{t\in T}$, $\mu = (\mu_t)_{t\in T}$, $\gamma = (\gamma_t)_{t\in T}$, and $\rho = (\rho_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ satisfying the following KKT relationship:*

$$\begin{cases} -\big(\nabla_p\varphi(\bar{p}, \bar{x}), \nabla_x\varphi(\bar{p}, \bar{x})\big) = \displaystyle\sum_{t\in T_1(\bar{p},\bar{x})} \lambda_t\,(\delta_t, \alpha_t) \\ + \displaystyle\sum_{t\in T_2(\bar{p},\bar{x})} \mu_t\,(-\delta_t, -\alpha_t) + \sum_{t\in T_3(\bar{p},\bar{x})} \gamma_t\,(0, -\delta_t) + \sum_{t\in T_4(\bar{p},\bar{x})} \rho_t\,(0, \delta_t)\,, \end{cases} \tag{7.74}$$

where the sets of active indices $T_i(\bar{p}, \bar{x})$, $i = 1, \ldots, 4$, are defined in (7.73).

Proof. This follows from the necessary optimality conditions in Corollary 7.18 applied to problem (7.71) taking into account the specification of the characteristic cone (7.49) for problem (7.71) obtained in (7.72) and then expressed via the active index sets from (7.73) corresponding to the infinite inequality constraints in (7.68)–(7.70). △

Observe that the optimality conditions obtained in Proposition 7.22 provide a valuable insight to our understanding of *optimal strategies* for the water resource problem. Indeed, it follows from the structures of constraints in (7.71) and their active index sets that the time inclusion $t \in T_1(\bar{p}, \bar{x})$ means that at this moment t the reservoir is empty, while the one of $t \in T_2(\bar{p}, \bar{x})$ means that at this time the quantity of water inside the reservoir given by $\langle\delta_t, p\rangle + \langle\alpha_t, x\rangle - s - \beta(t)$ attains its maximum level c, i.e., the reservoir is full. Similarly the inclusions $t \in T_i(\bar{p}, \bar{x})$ for $i = 3, 4$ signify, respectively, that the water is flowing at its minimum rate or at its maximum rate to satisfy the demand. The KKT relationship (7.74), valid under the Farkas-Minkowski condition, reflects therefore that the "dual action" (p^*, x^*) is a linear combination of these "bang-bang" strategies with the corresponding weights $(\lambda, \mu, \gamma, \rho)$. The general *asymptotic* optimality condition of the proposition indicates from this viewpoint that, in the absence of the Farkas-Minkowski property, the optimal impulse can be approximated by such combinations.

Finally in this section, we fully characterize the setting of Proposition 7.22 in which the Farkas-Minkowski property is satisfied for problem (7.71).

Proposition 7.23 (Farkas-Minkowski Property in Water Resource Optimization). *Let $\widetilde{T}$ be a nonempty subset of the time interval $T = [\underline{t}, \bar{t}]$ in (7.71). Then the cone $K(\widetilde{T})$ from (7.72) is weak* closed in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ if and only if the set $\widetilde{T}$ consists of a finite number of indices.*

Proof. The "if" part easily follows from the definitions. Let us justify the "only if" part arguing by contradiction and taking into account that the space $\mathcal{C}(T)$ is separable. Suppose that the set $\widetilde{T}$ is infinite and pick for simplicity a strictly monotone (increasing or decreasing) sequence $\{t_k\}_{k\in\mathbb{N}}$ in $\widetilde{T}$, which therefore converges to some point of T. It is not hard to check that the sequence in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ given by

$$\left\{\sum_{j=1}^{k} \frac{1}{j^2}\big(\delta_{t_j}, \alpha_{t_j}, \beta(t_j) + s\big)\right\}_{k\in\mathbb{N}} \tag{7.75}$$

weak* converges to the triple (δ, α, b) defined by

$$\langle(\delta, \alpha, b), (p, x, q)\rangle := \langle\delta, p\rangle + \langle\alpha, x\rangle + bq \tag{7.76}$$

via the componentwise relationships

$$\langle\delta, p\rangle := \sum_{j=1}^{\infty} \frac{1}{j^2} p\left(t_j\right), \;\; \langle\alpha, x\rangle := \sum_{j=1}^{\infty} \frac{1}{j^2} \int_{\underline{t}}^{t_j} x(t)\, dt, \;\; b := \sum_{j=1}^{\infty} \frac{1}{j^2}\big(\beta(t_j) + s\big).$$

Indeed, the weak* convergence of the above sequence follows directly from the boundedness of the set $\left\{\big(\delta_{t_j}, \alpha_{t_j}, \beta(t_j) + s\big)\right\}_{k\in\mathbb{N}}$ in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ and the convergence of the series $\sum_{j=1}^{\infty} \frac{1}{j^2}$.

Let us now show that $(\delta, \alpha, b) \notin K(\widetilde{T})$, and thus the cone $K(\widetilde{T})$ is not weak* closed. To verify it, observe that the inclusion $(\delta, \alpha, b) \in K(\widetilde{T})$ yields

$$\delta = \sum_{t\in\widetilde{T}} \lambda_t \delta_t \;\text{ for some }\; \lambda \in \mathbb{R}_+^{(\widetilde{T})},$$

which gives us a function $\delta \in \mathcal{C}(T)^*$ that is discontinuous only on a finite subset of T. It is easy to check at the same time that this component δ of the triple above is the weak* limit of the functions $\sum_{j=1}^{k} \frac{1}{j^2}\delta_{t_j}$ as $k \to \infty$, and hence it is discontinuous on the infinite set $\{t_k\}_{k\in\mathbb{N}}$. The obtained contradiction completes the proof of the proposition. △

One of the remarkable consequences of Proposition 7.23 is that the Farkas-Minkowski property *doesn't* hold for the water resource problem (7.71) on the compact continuous-time interval $T = [\underline{t}, \bar{t}]$. On the other hand, this result justifies yet another interpretation of the optimality conditions of Proposition 7.22 corresponding to the efficient realization of control strategies for reservoirs. Since in practice the measuring and control processes for the water resource model under consideration are implemented only at discrete instants of time, we can consider a *discretization* $\widetilde{T}$ of the time interval T and then apply the KKT conditions of Proposition 7.22 on $\widetilde{T}$.

7.3 Infinite Linear Systems Under Block Perturbations

In this section, we consider a class of infinite inequality constraint systems under *block perturbations*. Besides being of an undoubted interest in semilinear programming for its own sake, systems of this type eventually cover infinite *convex* inequality systems by using Fenchel duality. For brevity, we consider only the issues related to coderivative analysis of infinite linear block-perturbed and convex systems and its applications to characterizing Lipschitzian stability, i.e., we aim to develop convex counterparts of the results given in Section 7.1. It is not hard to observe that the coderivatives results obtained in this way can be equally applied to deriving both upper and lower subdifferential optimality conditions for SIPs with infinite constraints under consideration similarly to those obtained in Section 7.2 for the linear ones.

Our approach is as follows. We first consider infinite linear systems with block perturbations and extend to this case the results of Section 7.1. Then the results obtained are applied to infinite convex systems by using their *linearization* via Fenchel conjugates. As a by-product of our developments, we remove the boundedness assumption previously imposed on the coefficient of linear and convex systems in the case of reflexive decision spaces.

7.3.1 Description of Infinite Linear Block-Perturbed Systems

Given an arbitrary set $T \neq \emptyset$, consider its *partition*

$$\mathcal{J} := \left\{ T_j \mid j \in J \right\} \quad \text{with} \quad T_j \neq \emptyset \quad \text{for all} \quad j \in J$$

indexed by a fixed set $J \neq \emptyset$ so that we have

$$T = \bigcup_{j \in J} T_j \quad \text{with} \quad T_i \cap T_j = \emptyset \quad \text{if} \quad i \neq j,$$

where the sets T_j, $j \in J$, in the partition are referred to as *blocks*.

Given further a decision Banach space and coefficients $(a_t^*, b_t) \in X^* \times \mathbb{R}$, $t \in T$, consider the *block-perturbed* system

$$\sigma_{\mathcal{J}}(p) := \left\{ \langle a_t^*, x \rangle \leq b_t + p_j, \; t \in T_j, \; j \in J \right\} \tag{7.77}$$

with the perturbation parameter $p = \left(p_j\right)_{j \in J}$ ranging in the Banach space $l^\infty(J)$. The zero function $\bar{p} = 0$ is regarded as the *nominal parameter*, which corresponds to the *nominal system*

$$\sigma(0) := \left\{ \langle a_t^*, x \rangle \le b_t, \ t \in T \right\} \tag{7.78}$$

independently on the partition $\mathcal{J}$. The two extreme partitions

$$\mathcal{J}_{\min} := \{T\} \ \text{ and } \ \mathcal{J}_{\max} := \left\{ \{t\} \,\middle|\, t \in T \right\} \tag{7.79}$$

are called the *minimum partition* and the *maximum partition*, respectively.

Our major attention is focused in what follows on coderivative analysis of the *feasible solution map* $\mathcal{F}_{\mathcal{J}} : l^\infty(J) \rightrightarrows X$ generated by (7.77) as

$$\mathcal{F}_{\mathcal{J}}(p) := \left\{ x \in X \,\middle|\, x \ \text{ is a solution to } \ \sigma_{\mathcal{J}}(p) \right\} \tag{7.80}$$

and its applications to a complete characterization of Lipschitzian stability for (7.80) via the given data of the nominal system (7.78). Then we proceed with further applications to infinite convex inequality systems.

7.3.2 Stability of Block-Perturbed Systems via Coderivatives

First, we present the following coderivative calculation for $\mathcal{F}_{\mathcal{J}}$ at the reference point, where δ_j stands for the Dirac measure at $j \in J$ given by

$$\langle \delta_j, p \rangle := p_j \ \text{ for } \ p = (p_j)_{j \in J} \in l^\infty(J).$$

Proposition 7.24 (Coderivative Calculation for Block-Perturbed Linear Systems). *Let $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ for the mapping $\mathcal{F}_{\mathcal{J}} : l^\infty(J) \rightrightarrows X$ from (7.80). Then we have $p^* \in D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})(x^*)$ if and only if*

$$\left(p^*, -x^*, -\langle x^*, \bar{x} \rangle\right) \in \text{cl}^*\text{cone}\left\{ \left(-\delta_j, a_t^*, b_t\right) \,\middle|\, j \in J, \ t \in T_j \right\}.$$

Proof. It can be done by following the lines in the proof of Theorem 7.5 and the preceding propositions of Subsection 7.1.2. △

Similarly to (7.5), define the characteristic set for (7.77) by

$$C_{\mathcal{J}}(p) := \text{co}\left\{ (a_t^*, b_t + p_j) \,\middle|\, t \in T_j, \ j \in J \right\} \subset X^* \times \mathbb{R} \tag{7.81}$$

at $p \in l^\infty(J)$ and consider its specification at $p = 0$, which actually doesn't depend on $\mathcal{J}$ but just on the nominal system (7.78):

$$C(0) = \text{co}\left\{ \left(a_t^*, b_t\right) \,\middle|\, t \in T \right\}.$$

The strong Slater condition (SSC) for the nominal system $\sigma(0)$ and the corresponding strong Slater point $\widehat{x}$ are specifications of Definition 7.1 for $p = 0$.

We have the following equivalent relationships, which extend the equivalencies in Theorem 7.2 to the case of linear block-perturbed systems with taking into account some other results and proofs developed in Section 7.1.

Proposition 7.25 (Characterizations of the Lipschitz-Like Property for Linear Systems Under Block Perturbations). *Given* $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ *for the feasible solution map* (7.80)*, the following are equivalent:*

(i) $\mathcal{F}_{\mathcal{J}}$ *is Lipschitz-like around* $(0, \bar{x})$.
(ii) $D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})(0) = \{0\}$.
(iii) *SSC holds for* $\sigma(0)$.
(iv) $0 \in \text{int}(\text{dom}\, \mathcal{F}_{\mathcal{J}})$.
(v) $\mathcal{F}_{\mathcal{J}}$ *is Lipschitz-like around* $(0, x)$ *for all* $x \in \mathcal{F}_{\mathcal{J}}(0)$.
(vi) $(0, 0) \notin \text{cl}^*\, C(0)$.

Proof. Implication (i)⇒(ii) is verified, due to $D^*_M\mathcal{F}_{\mathcal{J}}(0, \bar{x}) = D^*_N\mathcal{F}_{\mathcal{J}}(0, \bar{x})$ by the graph convexity of $\mathcal{F}_{\mathcal{J}}(0, \bar{x})$, in Step 1 of Theorem 3.3 the proof of which holds without change in any change in arbitrary Banach space; see Exercise 3.35. The verification of the converse application (ii)⇒(i) follows the lines in the proof of Theorem 7.9 with the usage of Proposition 7.24. Since the conditions involved in (iii) and (vi) don't depend on partitions, the equivalence between them reduces to (iii)⇔(iv) for $p = 0$ in Theorem 7.2. Following the proof of (ii)⇔(iii) in Theorem 7.2 allows us to establish the equivalence between (iii) and (iv) for the maximum partition $\mathcal{J} = \mathcal{J}_{\max}$ in (7.79), which obviously implies that (iii)⇒(iv) for an arbitrary partition $\mathcal{J}$. The converse implication (iv)⇒(iii) holds by considering a constant perturbation $p \equiv \varepsilon$ with $\varepsilon > 0$ being sufficiently small to ensure that $p \in \text{int}(\text{dom}\, \mathcal{F}_{\mathcal{J}})$ by taking into account that constant perturbations (corresponding to the minimum partition $\mathcal{J} = \mathcal{J}_{\min}$ in (7.79)) are surely a particular case of block perturbations. The equivalent relationships in (i)⇔(iv) and (iv)⇔(v) follow from the classical Robinson-Ursescu theorem and the equivalence between the Lipschitz-like property of a mapping and the metric regularity/covering properties of the inverse; see Theorem 3.2, Corollary 3.6, and the corresponding commentaries in Section 3.5. This completes the proof of the proposition. △

Now we proceed with evaluating the exact Lipschitzian bound of the mapping (7.80) under block perturbations. Prior to establishing the main result in this direction, we present several propositions of their independent interest.

Proposition 7.26 (Relationships Between Exact Lipschitzian Bounds of Block-Perturbed Systems). *Let* $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ *for the feasible solution map from* (7.80)*. Then we have in the notation of* (7.79) *that*

$$\text{lip}\, \mathcal{F}_{\min}(0, \bar{x}) \leq \text{lip}\, \mathcal{F}_{\mathcal{J}}(0, \bar{x}) \leq \text{lip}\, \mathcal{F}_{\max}(0, \bar{x})\,.$$

Proof. We rely on the Lipschitzian bound representation given in (7.26). Consider the nontrivial case where SSC is satisfied at the nominal system $\sigma(0)$; otherwise all the exact Lipschitzian bounds above are equal to ∞ according to the equivalence (i)⇔(iii) in Proposition 7.25. Note that the mappings $\mathcal{F}_{\min}$, $\mathcal{F}_{\mathcal{J}}$, and $\mathcal{F}_{\max}$ act in the spaces $\mathbb{R}$, $l^\infty(J)$, and $l^\infty(T)$, respectively. For each $\rho \in \overline{\mathbb{R}}$, let p_ρ be the constant function $p_\rho \equiv \rho$ on J, and for each $p \in l^\infty(J)$, denote by p_T the piecewise constant function on T defined as p_j on the block T_j, $j \in J$. Let us further verify the two inequalities:

$$\text{dist}\left(\rho;\mathcal{F}_{\min}^{-1}(x)\right)\geq\text{dist}\left(p_\rho;\mathcal{F}_{\mathcal{J}}^{-1}(x)\right),\ \text{dist}\left(p;\mathcal{F}_{\mathcal{J}}^{-1}(x)\right)\geq\text{dist}\left(p_T;\mathcal{F}_{\max}^{-1}(x)\right)$$

valid for any $x \in X$. Indeed, we obviously have that $\mathcal{F}_{\mathcal{J}}^{-1}(x) = \emptyset$ yields $\mathcal{F}_{\min}^{-1}(x) = \emptyset$ and similarly for the second inequality above.

Consider now the nontrivial case where both of these sets are nonempty. Thus we get for some sequence $\{\rho_r\}_{r\in\mathbb{N}} \subset \mathcal{F}_{\min}^{-1}(x)$ that

$$\text{dist}\left(\rho;\mathcal{F}_{\min}^{-1}(x)\right) = \lim_{r\in\mathbb{N}}|\rho-\rho_r| = \lim_{r\in\mathbb{N}}\left\|p_\rho - p_{\rho_r}\right\| \geq \text{dist}\left(p_\rho;\mathcal{F}_{\mathcal{J}}^{-1}(x)\right)$$

by taking into account that $\rho_r \in \mathcal{F}_{\min}^{-1}(x)$ if and only if $p_{\rho_r} \in \mathcal{F}_{\mathcal{J}}^{-1}(x)$.

Finally, we appeal to representation (7.26) of the exact Lipschitzian bound combined with the directly verifiable equalities

$$\mathcal{F}_{\min}(\rho) = \mathcal{F}_{\mathcal{J}}\left(p_\rho\right) \ \text{ and } \ \mathcal{F}_{\mathcal{J}}(p) = \mathcal{F}_{\max}(p_T),$$

which thus allow us to complete the proof of the proposition. △

The next proposition establishes relationships between the coderivative norms of (7.80) corresponding to different partitions.

Proposition 7.27 (Coderivative Norms for Block-Perturbed Systems). *Consider the feasible solution mappings (7.80) corresponding to an arbitrary partition $\mathcal{J}$ and to the minimum one (7.79). Then for any $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$, we have*

$$\left\|D^*\mathcal{F}_{\min}(0,\bar{x})\right\| \leq \left\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\right\|. \tag{7.82}$$

Proof. Observe that $\mathcal{F}_{\mathcal{J}}(0) = \mathcal{F}_{\min}(0)$ since both sets therein reduce to the nominal one; hence $\bar{x} \in \mathcal{F}_{\min}(0)$. According to the coderivative norm definition, pick arbitrarily $x^* \in X^*$ with $\|x^*\| \leq 1$, and consider the nontrivial case where there exists $\mu \in \mathbb{R}\backslash\{0\}$ with $\mu \in D^*\mathcal{F}_{\min}(0,\bar{x})(x^*)$. The coderivative calculation in Proposition 7.24 entails the existence of a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t\in T} \in \mathbb{R}_+^{(T)}$ as $\nu \in \mathcal{N}$ satisfying the condition

$$\left(\mu, -x^*, -\left\langle x^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\left(-1, a_t^*, b_t\right). \tag{7.83}$$

Looking at the first coordinates in (7.83) and setting $\gamma_\nu := \sum_{t\in T}\lambda_{t\nu}$, we get $-\mu = \lim_{\nu\in\mathcal{N}}\gamma_\nu > 0$, and hence $\gamma_\nu > 0$ for ν sufficiently advanced in the directed set $\mathcal{N}$, say for all ν without loss of generality. This gives us

$$\left(\mu^{-1}x^*, \left\langle\mu^{-1}x^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\gamma_\nu^{-1}\lambda_{t\nu}\left(a_t^*, b_t\right) \in \text{cl}^*C(0). \tag{7.84}$$

For each $\nu \in \mathcal{N}$, consider next $\eta_\nu = \left(\eta_{j\nu}\right)_{j\in J} \in \mathbb{R}_+^{(J)}$ with $\eta_{j\nu} := \sum_{t\in T_j}\gamma_\nu^{-1}\lambda_{t\nu}$, which yields $\sum_{j\in J}\eta_{j\nu} = 1$. Since the net $\{\sum_{j\in J}\eta_{j\nu}\left(-\delta_j\right)\}_{\nu\in\mathcal{N}}$ is contained in the ball $\mathbb{B}_{l^\infty(J)^*}$, the Alaoglu-Bourbaki theorem tells us that a certain subnet (indexed without relabeling by $\nu \in \mathcal{N}$) weak* converges to some $p^* \in l^\infty(J)^*$ with $\|p^*\| \leq 1$. Denoting by $e \in l^\infty(J)$ the function whose coordinates are identically 1, we get the equality

$$\langle p^*, -e\rangle = \lim_{\nu\in\mathcal{N}} \sum_{j\in J} \eta_{j\nu} = 1, \quad \text{and so} \quad \|p^*\| = 1.$$

Appealing now to (7.84) shows for the subnet under consideration that

$$\left(p^*, \mu^{-1}x^*, \left\langle \mu^{-1}x^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{j\in J}\sum_{t\in T_j} \gamma_\nu^{-1}\lambda_{t\nu}\left(-\delta_j, a_t^*, b_t\right).$$

Employing then the coderivative description from Proposition 7.24 yields

$$p^* \in D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\left(-\mu^{-1}x^*\right).$$

Since $-\mu > 0$, the positive homogeneity of the coderivative implies that

$$-\mu p^* \in D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\left(x^*\right),$$

which ensures in turn by the coderivative norm definition that

$$\left\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\right\| \geq \left\|-\mu p^*\right\| = -\mu = |\mu|.$$

Since the number $\mu \in D^*\mathcal{F}_{\min}(0,\bar{x})(x^*)$ was chosen arbitrarily, we arrive at (7.82) and thus complete the proof of the proposition. △

To proceed further, we make for notational convenience the convention that $\sup\emptyset := 0$, which allows us to get the equality

$$\sup\left\{\|u^*\|^{-1} \mid \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \text{cl}^*C(0)\right\} = 0$$

for a strong Slater point $\bar{x}$ of $\sigma(0)$. Indeed, it is easy to check that for such a point $\bar{x}$, there is no element $u^* \in X^*$ satisfying $\left(u^*, \langle u^*, \bar{x}\rangle\right) \in \text{cl}^*C(0)$.

Note that the converse statement doesn't hold in general. To illustrate it, consider the system $\sigma(0) := \{tx \leq 1/t \ \text{ as } \ t = 1, 2, \ldots\}$ in $\mathbb{R}$. On the one hand, observe that $\bar{x} = 0$ is not a strong Slater point of this system. On the other hand, we have $\{u^* \in \mathbb{R} |\ (u^*, \langle u^*, \bar{x}\rangle) \in \text{cl}^*C(0)\} = \emptyset$.

Recall also that the failure of SSC for $\sigma(0)$ tells us by Proposition 7.25 that $(0, 0) \in \text{cl}^*C(0)$, which ensures under the convention $1/0 := \infty$ that for any feasible point $\bar{x}$ of $\sigma(0)$, we have the relationship

$$\sup\left\{\|u^*\|^{-1} \mid \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \text{cl}^*C(0)\right\} = \infty.$$

These observations are useful in deriving the following lower estimate of the coderivative norm for the minimum partition, which is an important step to obtain the main result of this section.

Proposition 7.28 (Lower Estimate of the Coderivative Norm for the Minimum Partition). *Consider the mapping $\mathcal{F}_{\min} : \mathbb{R} \rightrightarrows X$ defined by the minimum partition $\mathcal{J}_{\min}$ in (7.79), and pick any $\bar{x} \in \mathcal{F}_{\min}(0)$. Then we have*

$$\sup\left\{\left\|u^*\right\|^{-1} \mid \left(u^*, \left\langle u^*, \bar{x}\right\rangle\right) \in \text{cl}^*C(0)\right\} \leq \left\|D^*\mathcal{F}_{\min}(0, \bar{x})\right\|. \tag{7.85}$$

Proof. Let us check first that $\|D^*\mathcal{F}_{\min}(0, \bar{x})\| = \infty$ provided that SSC for $\sigma(0)$. Indeed, in this case, Proposition 7.25 tells us that $(0, 0) \in \text{cl}^*C(0)$, which yields in turn the existence of a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t\in T} \in \mathbb{R}_+^{(T)}$ and $\sum_{t\in T}\lambda_{t\nu} = 1$ as $\nu \in \mathcal{N}$ satisfying the condition

$$\left(0, 0\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\left(a_t^*, b_t\right).$$

The latter obviously implies that $\left(-1, 0, 0\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\left(-1, a_t^*, b_t\right)$, i.e., by Proposition 7.24 we get the inclusion

$$-1 \in D^*\mathcal{F}_{\min}(0, \bar{x})(0).$$

Since $D^*\mathcal{F}_{\min}(0, \bar{x})$ is positively homogeneous, the coderivative norm definition ensures the validity of the claimed condition $\|D^*\mathcal{F}_{\min}(0, \bar{x})\| = \infty$.

Next we consider the nontrivial case where SSC holds for $\sigma(0)$ and the set of elements $u^* \in X^*$ with $(u^*, \langle u^*, \bar{x}\rangle) \in \text{cl}^*C(0)$ is nonempty. Take such an element u^*, and observe that the fulfillment of SSC for $\sigma(0)$ yields $u^* \neq 0$ according to Proposition 7.25. The choice of u^* allows us to find a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t\in T} \in \mathbb{R}_+^{(T)}$ and $\sum_{t\in T}\lambda_{t\nu} = 1$ as $\nu \in \mathcal{N}$ satisfying

$$\left(u^*, \left\langle u^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\left(a_t^*, b_t\right),$$

which can be equivalently rewritten in the form

$$\left(-1, u^*, \left\langle u^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\left(-1, a_t^*, b_t\right).$$

This implies that $-1 \in D^*\mathcal{F}_{\min}(0, \bar{x})(-u^*)$, and hence

$$-\left\|u^*\right\|^{-1} \in D^*\mathcal{F}_{\min}(0, \bar{x})\left(-\left\|u^*\right\|^{-1}u^*\right),$$

which ensures by the definition of the coderivative norm that

$$\left\|D^*\mathcal{F}_{\min}(0, \bar{x})\right\| \geq \left\|u^*\right\|^{-1}.$$

Since the element u^* was chosen arbitrarily from those satisfying the inclusion $(u^*, \langle u^*, \bar{x}\rangle) \in \text{cl}^*C(0)$, we arrive at the claimed lower estimate (7.85) of the coderivative norm and thus complete the proof of the proposition. △

Now we are ready to establish the main result of this subsection.

Theorem 7.29 (Evaluation of Coderivative Norms for Block-Perturbed Systems). *For any $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$, we have the relationships*

$$\sup\left\{\|u^*\|^{-1} \,\Big|\, \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \mathrm{cl}^*C(0)\right\} \le \|D^*\mathcal{F}_{\min}(0,\bar{x})\| \le \left\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\right\|$$
$$\le \mathrm{lip}\,\mathcal{F}_{\mathcal{J}}(0,\bar{x}) \le \mathrm{lip}\,\mathcal{F}_{\max}(0,\bar{x}).$$

Furthermore, if either the coefficient set $\{a_t^* \mid t \in T\}$ is bounded in X^* or the space X is reflexive, then all the above inequalities hold as equalities.

Proof. Recall as above that the lower estimate

$$\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\| \le \mathrm{lip}\,\mathcal{F}_{\mathcal{J}}(0,\bar{x}) \tag{7.86}$$

follows from the proof of Step 1 Theorem 3.3 in arbitrary Banach spaces. Applying now (in this order) Propositions 7.28 and 7.27, formula (7.86), and Proposition 7.26 verifies the chain of inequalities claimed in the theorem.

To verify the equalities therein under the additional assumptions made, consider first the case where the coefficient set $\{a_t^* \mid t \in T\}$ is bounded in X^*. Then using Theorem 7.15 adapted to the current notation gives us

$$\mathrm{lip}\,\mathcal{F}_{\max}(0,\bar{x}) \le \sup\left\{\left\|u^*\right\|^{-1} \,\Big|\, \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \mathrm{cl}^*C(0)\right\} \tag{7.87}$$

in the nontrivial case where SSC holds for the nominal system $\sigma(0)$.

It remains to consider the case where the space X is reflexive and to justify the upper estimate (7.87) provided the validity of SSC for $\sigma(0)$. Employing in this case the Mazur weak closure theorem allows us to replace the weak* closure $\mathrm{cl}^*C(0)$ of the convex set $C(0)$ by its norm closure $\mathrm{cl}\,C(0)$. Suppose that (7.87) fails, and choose $\beta > 0$ such that

$$\mathrm{lip}\,\mathcal{F}_{\max}(0,\bar{x}) > \beta > \sup\left\{\left\|u^*\right\|^{-1} \,\Big|\, \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \mathrm{cl}\,C(0)\right\}. \tag{7.88}$$

Using the distance representation (7.26) of the exact Lipschitzian bound and the first inequality in (7.88) gives us sequences $p_r = (p_{tr})_{t\in T} \to 0$ and $x_r \to \bar{x}$ along which we have the relationship

$$\mathrm{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) > \beta\,\mathrm{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) \text{ for all } r \in \mathbb{N}, \tag{7.89}$$

which readily implies that the quantity

$$\begin{aligned}\mathrm{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) &= \sup_{t\in T}\left[\langle a_t^*, x_r\rangle - b_t - p_{tr}\right]_+ \\ &= \sup_{(x^*,\alpha)\in C_{\max}(p_r)}\left[\langle x^*, x_r\rangle - \alpha\right]_+\end{aligned} \tag{7.90}$$

is finite. It follows from Proposition 7.25 due to the assumed SSC that $\mathcal{F}_{\max}(p_r) \neq \emptyset$ for $r \in \mathbb{N}$ sufficiently large, say, for all $r \in \mathbb{N}$ without loss of generality. Furthermore, under this condition, we have

$$\lim_{r\to\infty} \operatorname{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) = 0; \tag{7.91}$$

see Exercise 7.86 for more discussions. Assume without loss of generality the validity of SSC for the system $\sigma_{\max}(p_r)$ and then deduce from the extended Ascoli formula (7.37) for infinite linear systems in Proposition 7.12, which holds in reflexive spaces, the representation

$$\operatorname{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) = \sup_{(x^*,\alpha)\in C_{\max}(p_r)} \frac{\big[\langle x^*, x_r\rangle - \alpha\big]_+}{\|x^*\|}, \quad r \in \mathbb{N}.$$

This allows us to find $\big(x_r^*, \alpha_r\big) \in C_{\max}(p_r)$ as $r \in \mathbb{N}$ satisfying

$$0 < \operatorname{dist}\big(x_r, \mathcal{F}_{\max}(p_r)\big) - \frac{\big\langle x_r^*, x_r\big\rangle - \alpha_r}{\big\|x_r^*\big\|} < \frac{1}{r}. \tag{7.92}$$

Furthermore, by (7.89) and (7.90), we can choose $\big(x_r^*, \alpha_r\big)$ in (7.92) so that

$$\beta \operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) < \frac{\big\langle x_r^*, x_r\big\rangle - \alpha_r}{\big\|x_r^*\big\|} \le \frac{\operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big)}{\big\|x_r^*\big\|}. \tag{7.93}$$

Since $\operatorname{dist}(p_r; \mathcal{F}_{\max}^{-1}(x_r)) > 0$ (otherwise both sides of (7.89) would be equal to zero, which is not possible), it follows from (7.93) that $\|x_r^*\| < \beta^{-1}$ for all $r \in \mathbb{N}$. Thus, by the weak* sequential compactness of the unit balls in duals to reflexive spaces, we select a subsequence $\big\{x_{r_k}^*\big\}_{k\in\mathbb{N}}$ that weak* converges to some $x^* \in X^*$ with $\|x^*\| \le \beta^{-1}$. Then (7.91) and (7.92) yield

$$\lim_{k\in\mathbb{N}} \frac{\big\langle x_{r_k}^*, x_{r_k}\big\rangle - \alpha_{r_k}}{\big\|x_{r_k}^*\big\|} = 0, \quad \text{and so} \quad \lim_{k\in\mathbb{N}} \Big(\big\langle x_{r_k}^*, x_{r_k}\big\rangle - \alpha_{r_k}\Big) = 0.$$

The latter implies by the normal convergence of $\big\{x_{r_k}\big\}_{k\in\mathbb{N}}$ to $\bar{x}$ that

$$\lim_{k\in\mathbb{N}} \alpha_{r_k} = \lim_{k\in\mathbb{N}} \big\langle x_{r_k}^*, x_{r_k}\big\rangle = \big\langle x^*, \bar{x}\big\rangle.$$

Then we deduce from $\big(x_{r_k}^*, \alpha_{r_k}\big) \in C_{\max}\big(p_{r_k}\big)$ the existence of multiplies $\lambda_{r_k} = (\lambda_{tr_k})_{t\in T}$ such that $\lambda_{tr_k} \ge 0$, only finitely many of them are not zero, and

$$\sum_{t\in T} \lambda_{tr_k} = 1, \quad \text{and} \quad (x_{r_k}^*, \alpha_{r_k}) = \sum_{t\in T} \lambda_{tr_k} \big(a_t^*, b_t + p_{tr_k}\big), \quad k \in \mathbb{N}.$$

Combining the above equations gives us the relationships

$$\begin{aligned}\left(x^*, \langle x^*, \bar{x}\rangle\right) &= w^* \text{-}\lim_{k\in\mathbb{N}}(x^*_{r_k}, \alpha_{r_k}) = w^* \text{-}\lim_{k\in\mathbb{N}} \sum_{t\in T} \lambda_{tr_k} \left(a_t^*, b_t + p_{tr_k}\right) \\ &= w^* \text{-}\lim_{k\in\mathbb{N}} \sum_{t\in T} \lambda_{tr_k} \left(a_t^*, b_t\right) \in \operatorname{cl} C\,(0)\,,\end{aligned}$$

where the last equality comes from $\lim_{k\to\infty} \left\| p_{r_k} \right\| = 0$. Observe finally that $x^* \neq 0$ due to the validity of SSC for $\sigma\,(0)$ by Proposition 7.25. Hence

$$\sup\left\{ \left\|u^*\right\|^{-1} \mid \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \operatorname{cl} C\,(0)\right\} \geq \left\|x^*\right\|^{-1} \geq \beta,$$

which contradicts (7.88) and thus completes the proof of the theorem. △

7.3.3 Applications to Infinite Convex Inequality Systems

Here we consider *parameterized convex inequality systems* given by

$$\sigma(p) := \left\{\varphi_j\,(x) \leq p_j,\;\; j \in J\right\}, \tag{7.94}$$

where J is an arbitrary index set and where the functions $\varphi_j \colon X \to \overline{\mathbb{R}}$, $j \in J$, are l.s.c. (our standing assumption) and *convex* on the Banach space X. As above, the functional parameter p belongs to $l^\infty(J)$ and the zero function $\bar{p} = 0$ is the nominal parameter. Our goal is to characterize Lipschitzian stability of the convex system (7.94) around $\bar{p} = 0$ by applying the obtained results for block-perturbed linear systems. We can do it with the help of the *Fenchel conjugate* (7.30) defined for each function φ_j by

$$\varphi_j^*(u^*) := \sup\left\{ \langle u^*, x\rangle - \varphi_j(x) \,\middle|\, x \in X\right\} = \sup\left\{ \langle u^*, x\rangle - \varphi_j(x) \,\middle|\, x \in \operatorname{dom}\varphi_j\right\}.$$

Indeed, the classical Fenchel duality theorem tells us that relationship

$$\varphi_j^{**} = \varphi_j \;\text{ on }\; X \;\text{ with }\; \varphi_j^{**} := \left(\varphi_j^*\right)^*$$

holds under the assumptions made. Using this, we get for each $j \in J$ that the convex inequality $\varphi_j\,(x) \leq p_j$ turns out to be equivalent to the *linear system*

$$\left\{\langle u^*, x\rangle - \varphi_j^*\left(u^*\right) \leq p_j,\;\; u^* \in \operatorname{dom}\varphi_j^*\right\} \tag{7.95}$$

in the sense that they have the same solution sets. Denote

$$T := \left\{\left(j, u^*\right) \in J \times X^* \,\middle|\, u^* \in \operatorname{dom}\varphi_j^*\right\}$$

and observe that T can be partitioned as

$$T = \bigcup_{j \in J} T_j \quad \text{with} \quad T_j := \{j\} \times \operatorname{dom} \varphi_j^*. \tag{7.96}$$

In this way the right-hand side perturbations on the nominal convex system $\sigma(0)$ correspond to block perturbations of the linearized nominal system $\sigma_{\mathcal{J}}(0)$ with the partition $\mathcal{J} := \{T_j \mid j \in J\}$. It is important to realize to this end that the feasible solution map $\mathcal{F} : l^\infty(J) \rightrightarrows X$ to (7.94) given by

$$\mathcal{F}(p) := \{x \in X \big| \; x \text{ is a solution to } \sigma(p)\} \tag{7.97}$$

and the one for the block-perturbed linearized system $\mathcal{F}_{\mathcal{J}}$ with the partition $\mathcal{J} := \{T_j \mid j \in J\}$ are *exactly the same mapping*. This allows us to implement the results of Subsection 7.3.1 to characterizing Lipschitzian stability of infinite convex systems. It is not hard to check that the convex counterpart of the characteristic set $C_{\mathcal{J}}(p)$ from (7.81) is

$$\begin{aligned} C(p) &:= \operatorname{co}\Big\{\Big(u^*, \varphi_j^*(u^*) + p_j\Big) \Big| \; j \in J, \; u^* \in \operatorname{dom} \varphi_j^*\Big\} \\ &= \operatorname{co}\Big(\bigcup\nolimits_{j \in J} \operatorname{gph}(\varphi_j - p_j)^*\Big) \subset X^* \times \mathbb{R}. \end{aligned} \tag{7.98}$$

Observe that for the convex system $\sigma(0)$ under consideration, the corresponding SSC reads as $\sup_{t \in T} \varphi_t(\widehat{x}) < 0$ for some $\widehat{x} \in X$ and that $\widehat{x}$ is a strong Slater point for $\sigma(0)$ if and only if

$$\sup_{(j,u^*) \in T} \big\{ \langle u^*, \widehat{x} \rangle - \varphi_j^*(u^*) \big\} < 0.$$

The next result provides calculating the coderivative of the solution map (7.97) to the original infinite convex system (7.94) in terms of its initial data.

Proposition 7.30 (Calculating Coderivatives for Infinite Convex Systems). *Take* $\bar{x} \in \mathcal{F}(0)$ *for the solution map* (7.97) *to the convex system* (7.94)*. Then we have* $p^* \in D^*\mathcal{F}(0, \bar{x})(x^*)$ *if and only if*

$$\big(p^*, -x^*, -\langle x^*, \bar{x} \rangle\big) \in \operatorname{cl}^* \operatorname{cone}\Big(\bigcup_{j \in J} \big(\{-\delta_j\} \times \operatorname{gph} \varphi_j^*\big)\Big). \tag{7.99}$$

Proof. It follows directly from its linear counterpart in Proposition 7.24. △

Now we are ready to present the major result of this subsection proving an evaluation of the exact Lipschitzian bound for the feasible solution map (7.97) for infinite convex inequality systems.

Theorem 7.31 (Evaluation of the Coderivative Norm for Infinite Convex Systems). *For any* $\bar{x} \in \mathcal{F}(0)$ *from* (7.97)*, we have the relationships*

$$\sup\left\{\|u^*\|^{-1}\,\Big|\big(u^*,\langle u^*,\bar{x}\rangle\big)\in \text{cl}^*\text{co}\Big(\bigcup_{j\in J}\text{gph}\,\varphi_j^*\Big)\right\}\le \|D^*\mathcal{F}(0,\bar{x})\|\le \text{lip}\mathcal{F}(0,\bar{x}).$$

If furthermore either the set $\bigcup_{j\in J}\text{dom}\,\varphi_j^*$ *is bounded in* X^* *or the space* X *is reflexive, then the above inequalities hold as equalities.*

Proof. It follows from Theorem 7.29 applied to the linear system (7.95) with block perturbations by employing the above linearization procedure and the coderivative calculation given in Proposition 7.30. △

The next example shows that the boundedness assumption, which looks quite natural in the linear setting, may fail for very simple convex systems.

Example 7.32 (Failure of the Boundedness Assumption for Infinite Convex Inequality Systems). Consider the following single convex inequality involving one-dimensional decision and parameter variables:

$$x^2\le p \quad\text{with}\quad x,p\in\mathbb{R}.$$

The linearized system associated with it reads as follows:

$$\Big\{ux\le\frac{u^2}{4}+p,\ u\in\mathbb{R}\Big\},$$

and thus the boundedness assumption of Theorem 7.31 fails.

7.4 Metric Regularity of Infinite Convex Systems

In this section, we develop another approach to well-posedness of infinite convex constraint systems concentrating mainly on their metric regularity. The study of well-posedness in Chapter 3 reveals that, although metric regularity of general multifunctions is equivalent to the Lipschitz-like property of their inverses, the former is unnatural (fails as a rule), while the latter holds under unrestrictive qualification conditions for broad classes of set-valued mappings known as parametric *parametric variational systems (PVS)*; see Section 3.3. The situation is *parametric constraint systems (PCS)* different for parametric *constraint* systems, where both metric regularity and Lipschitzian properties can be studied in parallel and are satisfied under similar (symmetric) constraint qualifications; cf. Section 3.3 and [522, Section 4.3]. The infinite constraint systems considered in Sections 7.1 and 7.3 belong to the latter category, and so their metric regularity and Lipschitzian stability can be studied and characterized in a parallel way.

In fact, full characterizations of metric regularity for the infinite linear and convex inequality systems considered in Sections 7.1 and 7.3 can be derived from the *equalities* for their exact Lipschitzian bounds, which are reciprocal to the exact bounds of metric regularity. However, the aforementioned calculation of the exact Lipschitzian bound in Theorem 7.31 (which extends the previous ones for linear systems) is justified under the imposed *boundedness* assumption, which is rather restrictive (as shown in Example 7.32) while cannot be removed in the given proof unless the decision space is reflexive.

The new approach to characterizing metric regularity of infinite convex systems developed below is completely different from the one employed in the previous sections of this chapter. It first concerns the study of metric regularity of general multifunctions with closed and convex graphs for which we establish formulas for the *precise calculation* of the *exact regularity bound* in arbitrary Banach spaces without imposing any qualification conditions while with involving ε-*coderivatives*. Our approach to these issues is based on reducing metric regularity of such mappings to the unconstrained minimization of *DC* (difference of convex) functions. In this way we obtain regularity criteria for general convex-graph multifunctions and then apply them to metric regularity of infinite convex systems. It allows us not only to cover the case of infinite convex inequalities in arbitrary Banach spaces without imposing the aforementioned boundedness assumption but also to include additional linear equality and convex geometric constraints into consideration.

7.4.1 DC Optimization Approach to Metric Regularity

Recall in accordance with (3.2) in Definition 3.1, a set-valued mapping $F\colon X \rightrightarrows Y$ between metric spaces is *metrically regular* around $(\bar{x},\bar{y}) \in \operatorname{gph} F$ with modulus $\mu > 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \le \mu \operatorname{dist}\big(y; F(x)\big) \quad \text{for any} \quad x \in U \quad \text{and} \quad y \in Y.$$

The *exact regularity bound* $\operatorname{reg} F(\bar{x},\bar{y})$ of F around $(\bar{x},\bar{y})$ is the infimum of all such moduli μ. It is easy to observe directly from the definition that the metric regularity (3.2) is amount to saying that $(\bar{x},\bar{y})$ is a *local minimizer* of the following unconstrained optimization problem:

$$\text{minimize} \quad \mu \operatorname{dist}\big(y; F(x)\big) - \operatorname{dist}\big(x; F^{-1}(y)\big) \tag{7.100}$$

over $(x,y) \in X \times Y$. Throughout this and the next subsections, we consider, unless otherwise stated, multifunctions F between *arbitrary Banach* spaces with *closed* and *convex* graphs. Observe that (7.100) is a *DC minimization problem*. Problems of this type are briefly studied in Section 6.1 and in much more details in Section 7.5 while from different prospectives.

To proceed, we need to recall some notions and facts from convex analysis and DC optimization. Given a convex function $\varphi\colon X \to \overline{\mathbb{R}}$ and $\varepsilon \ge 0$, the ε-*subdifferential* of φ at $\bar{x} \in \operatorname{dom}\varphi$ is defined by

$$\partial_\varepsilon\varphi(\bar{x}) := \big\{x^* \in X^* \big|\ \langle x^*, x-\bar{x}\rangle \le \varphi(x) - \varphi(\bar{x}) + \varepsilon,\ x \in X\big\}, \tag{7.101}$$

which reduces to the subdifferential of convex analysis for $\varepsilon = 0$; this construction is also known as the *approximate subdifferential* of φ at $\bar{x}$ if $\varepsilon > 0$. We put $\partial_\varepsilon\varphi(\bar{x}) := \emptyset$

if $\bar{x} \notin \operatorname{dom}\varphi$. Note that (7.101) for $\varepsilon > 0$ is different from the ε-enlargement $\widehat{\partial}_\varepsilon \varphi(\bar{x})$ of the regular subdifferential from (1.34) in the case of convex functions under consideration; see Proposition 1.25. The following *ε-subdifferential sum rule* is well known in convex analysis:

$$\partial_\varepsilon(\varphi_1 + \varphi_2)(\bar{x}) = \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon \\ \varepsilon_1,\varepsilon_2 \geq 0}} \Big[\partial_{\varepsilon_1}\varphi_1(\bar{x}) + \partial_{\varepsilon_2}\varphi_2(\bar{x})\Big] \tag{7.102}$$

provided that one of the functions φ_i is continuous at $\bar{x} \in \operatorname{dom}\varphi_1 \cap \operatorname{dom}\varphi_2$; see Exercise 7.93 for more discussions.

Given a convex set $\Omega \subset X$, we have the collection of (*convex*) *ε-normals*

$$N_\varepsilon(\bar{x};\Omega) := \partial_\varepsilon \delta(\bar{x};\Omega) = \big\{x^* \in X^* \big|\ \langle x^*, x - \bar{x}\rangle \leq \varepsilon \ \text{ for all } \ x \in \Omega\big\}, \quad \varepsilon \geq 0,$$

which can be equivalently represented in the form

$$N_\varepsilon(\bar{x};\Omega) = \big\{x^* \in X^* \big|\ \sigma_\Omega(x^*) \leq \langle x^*, \bar{x}\rangle + \varepsilon\big\}, \tag{7.103}$$

where σ_Ω stands for the *support function* of Ω defined by

$$\sigma_\Omega(x^*) := \sup\big\{\langle x^*, x\rangle \big|\ x \in \Omega\big\}, \quad x^* \in X^*.$$

Again note that convex ε-normals in (7.103) are different as $\varepsilon > 0$ from regular ε-normals in $\widehat{N}_\varepsilon(\bar{x};\Omega)$ defined in (1.6) for general (including convex) sets.

The *ε-coderivative* of a set-valued mapping $F\colon X \rightrightarrows Y$ at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by the usual scheme via ε-normals to the graph

$$D^*_\varepsilon F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^* \big|\ (x^*, -y^*) \in N_\varepsilon\big((\bar{x},\bar{y}); \operatorname{gph} F\big)\big\} \tag{7.104}$$

for $\varepsilon \geq 0$ with $D^*_0 F(\bar{x},\bar{y}) = D^*F(\bar{x},\bar{y})$. The *$\varepsilon$-coderivative norm* is given by

$$\|D^*_\varepsilon F(\bar{x},\bar{y})\| := \sup\big\{\|x^*\| \ \big|\ x^* \in D^*_\varepsilon F(\bar{x},\bar{y})(y^*),\ y^* \in \mathbb{B}_{Y^*}\big\}. \tag{7.105}$$

If F is metrically regular around $(\bar{x},\bar{y})$, we get from Theorem 3.3(ii), by observing that this part holds in any Banach space, that $D^*F^{-1}(\bar{y},\bar{x})(0) = \{0\}$, and thus arrive at the norm representation via the unit sphere S_{X^*}:

$$\|D^*F^{-1}(\bar{y},\bar{x})\| = \sup\big\{\|y^*\| \ \big|\ y^* \in D^*F^{-1}(\bar{y},\bar{x})(x^*),\ x^* \in S_{X^*}\big\}. \tag{7.106}$$

The following two results from DC programming in Banach spaces involving ε-subgradients of convex functions (7.101) are important in the proof of the main theorem in the next subsection.

Lemma 7.33 (Necessary and Sufficient Conditions for Global DC Minimizers). *Let $\varphi_1, \varphi_2 : X \to \overline{\mathbb{R}}$ be convex functions. Then $\bar{x}$ is a global minimizer of the unconstrained DC program given by*

$$\textit{minimize} \;\; \varphi_1(x) - \varphi_2(x) \;\; \textit{over} \;\; x \in X \tag{7.107}$$

if and only if $\partial_\varepsilon \varphi_2(\bar{x}) \subset \partial_\varepsilon \varphi_1(\bar{x})$ *for all* $\varepsilon \geq 0$.

Note that the *necessity* of the obtained subdifferential inclusion with $\varepsilon = 0$ for *local* minimizers of (7.107) is established in Proposition 6.3 as a consequence of *upper* subdifferential conditions in unconstrained optimization; see more discussions and references in Exercise 7.94(i,ii). The next result provides a *sufficient* condition of this type for *local* minimizers of (7.107); see Exercise 7.94(iii,iv) for the proof and discussions.

Lemma 7.34 (Sufficient Conditions for Local DC Minimizers). *Let* $\varphi_1, \varphi_2 : X \to \overline{\mathbb{R}}$ *be convex functions, and let* φ_2 *be continuous at the point* $\bar{x} \in \operatorname{dom} \varphi_1 \cap [\operatorname{int}(\operatorname{dom} \varphi_2)]$. *Then* $\bar{x}$ *is a local minimizer of* (7.107) *if there is* $\varepsilon_0 > 0$ *such that* $\partial_\varepsilon \varphi_2(\bar{x}) \subset \partial_\varepsilon \varphi_1(\bar{x})$ *for all* $\varepsilon \in [0, \varepsilon_0]$.

7.4.2 Metric Regularity of Convex-Graph Multifunctions

Now we are ready to establish the main result on calculating the exact regularity bound of closed- and convex-graph multifunctions via their ε-coderivatives at the reference points. The next theorem presents two limiting formulas for calculating this bound in general Banach spaces.

Theorem 7.35 (ε-Coderivative Formulas for the Exact Regularity Bound). *Given a point* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, *assume that* $\bar{y} \in \operatorname{int}(\operatorname{rge} F)$. *Then we have*

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \|D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})\|, \tag{7.108}$$

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \Big[\sup \Big\{ \frac{1}{\|x^*\|} \;\Big|\; x^* \in D^*_\varepsilon F(\bar{x}, \bar{y})(y^*),\; y^* \in S_{Y^*} \Big\} \Big]. \tag{7.109}$$

Proof. Since $\bar{y} \in \operatorname{int}(\operatorname{rge} F)$, it follows from the Robinson-Ursescu theorem in Banach spaces (see Corollary 3.6 and Exercise 3.49) that F is metrically regular around $(\bar{x}, \bar{y})$, i.e., there are $\eta, \mu > 0$ such that

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \leq \mu \operatorname{dist}\big(y; F(x)\big) \;\; \text{for all } (x, y) \in B_\eta(\bar{x}, \bar{y}). \tag{7.110}$$

Consider now the convex functions φ_1, φ_2 on $X \times Y$ defined by

$$\varphi_1(x, y) := \operatorname{dist}\big(y; F(x)\big) \;\; \text{and} \;\; \varphi_2(x, y) := \operatorname{dist}\big(x; F^{-1}(y)\big) \tag{7.111}$$

and deduce from the covering property of F equivalent to metric regularity that there is $r > 0$ such that $B_{2r}(\bar{y}) \subset F(\bar{x} + \mathbb{B}_X)$. Combining this with the construction of φ_2 in (7.111) provides the estimate

$$\varphi_2(x, y) \leq \|x - \bar{x}\| + 1 \;\; \text{whenever} \;\; y \in B_{2r}(\bar{y}),$$

which tells us that φ_2 is upper bounded around $(\bar{x}, \bar{y})$, and thus it is locally Lipschitzian around this point due to the well-known result of convex analysis; see, e.g., [757, Corollary 2.2.13]. Implementing our approach to metric regularity, we conclude that $(\bar{x}, \bar{y})$ is a local minimizer of the DC program:

$$\text{minimize } \mu\varphi_1(x, y) - \varphi_2(x, y) \text{ subject to } (x, y) \in X \times Y, \tag{7.112}$$

and consequently it is a global minimizer of the DC function

$$\big(\mu\varphi_1 + \delta(\cdot; B_\eta(\bar{x}, \bar{y}))\big)(x, y) - \varphi_2(x, y) \text{ over } (x, y) \in X \times Y. \tag{7.113}$$

Applying Lemma 7.33 to the DC program (7.113) gives us the inclusion

$$\partial_\varepsilon \varphi_2(\bar{x}, \bar{y}) \subset \partial_\varepsilon \big(K\varphi_1 + \delta(\cdot; B_\eta(\bar{x}, \bar{y}))\big)(\bar{x}, \bar{y}) \text{ for all } \varepsilon \geq 0.$$

Since the function $\delta((\cdot, \cdot); B_\eta(\bar{x}, \bar{y}))$ is continuous at $(\bar{x}, \bar{y})$, it follows from the ε-subdifferential sum rule (7.102) that the latter inclusion reduces to

$$\begin{aligned} \partial_\varepsilon \varphi_2(\bar{x}, \bar{y}) &\subset \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon \\ \varepsilon_1,\varepsilon_2 \geq 0}} \Big[\partial_{\varepsilon_1}(K\varphi_1)(\bar{x}, \bar{y}) + \partial_{\varepsilon_2}\delta\big(\cdot; B_\eta(\bar{x}, \bar{y})\big)(\bar{x}, \bar{y})\Big] \\ &= \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon \\ \varepsilon_1,\varepsilon_2 \geq 0}} \Big[\partial_{\varepsilon_1}(K\varphi_1)(\bar{x}, \bar{y}) + \frac{\varepsilon_2}{\eta}\mathbb{B}_{X^* \times Y^*}\Big] \end{aligned} \tag{7.114}$$

due to the fact that $\partial_\varepsilon \delta(\cdot; B_r(x))(x) = \frac{\varepsilon}{r}\mathbb{B}_{X^*}$ for all $\varepsilon \geq 0$ and $r > 0$.

Let us next calculate the ε-subdifferentials of the functions $K\varphi_1$ and φ_2 from (7.111) at $(\bar{x}, \bar{y})$ by using their Fenchel conjugates (7.30) and the obvious ε-subdifferential representation for any convex function $\varphi\colon X \to \overline{\mathbb{R}}$:

$$\partial_\varepsilon \varphi(\bar{x}) = \big\{x^* \in X^* \big|\ \varphi^*(x^*) \leq \langle x^*, \bar{x}\rangle - \varphi(\bar{x}) + \varepsilon\big\}, \quad \varepsilon \geq 0.$$

In this way we get that $(x^*, y^*) \in \partial_{\varepsilon_1}(\mu\varphi_1)(\bar{x}, \bar{y})$ if and only if

$$(\mu\varphi_1)^*(x^*, y^*) \leq \langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle + \varepsilon_1, \tag{7.115}$$

which ensures in turn by elementary transformations that

$$\begin{aligned} (\mu\varphi_1)^*(x^*, y^*) &= \sup_{x,y}\Big(\langle x^*, x\rangle + \langle y^*, y\rangle - \mu\,\text{dist}\big(y; F(x)\big)\Big) \\ &= \sup_{x,y}\Big(\langle x^*, x\rangle + \langle y^*, y\rangle - \inf_u\big(\mu\|y - u\| + \delta(u; F(x))\big)\Big) \\ &= \sup_{u,x,y}\Big(\langle x^*, x\rangle + \langle y^*, y - u\rangle + \langle y^*, u\rangle - \mu\|y - u\| - \delta\big(u; F(x)\big)\Big) \\ &= \sup_{u,x,y}\Big(\langle x^*, x\rangle + \langle y^*, u\rangle - \delta\big(u; F(x)\big) + \langle y^*, y\rangle - \mu\|y\|\Big) \\ &= \sigma_{\text{gph}F}(x^*, y^*) + \delta\big(y^*; \mu\mathbb{B}_{Y^*}\big). \end{aligned}$$

By using (7.103) and (7.115), the latter implies that

$$\partial_{\varepsilon_1}(\mu\varphi_1)(\bar{x},\bar{y}) = N_{\varepsilon_1}\big((\bar{x},\bar{y});\operatorname{gph} F\big) \cap \big(X^* \times \mu\mathbb{B}_{Y^*}\big). \tag{7.116}$$

Similarly, by taking into account the form of φ_2 in (7.111), we arrive at

$$\partial_{\varepsilon}\varphi_2(\bar{x},\bar{y}) = N_{\varepsilon}\big((\bar{x},\bar{y});\operatorname{gph} F\big) \cap \big(\mathbb{B}_{X^*} \times Y^*\big). \tag{7.117}$$

Thus the inclusion in (7.114) reduces to the following:

$$\begin{aligned} N_{\varepsilon}\big((\bar{x},\bar{y});\operatorname{gph} F\big) \cap \big(\mathbb{B}^* \times Y^*\big) \subset & \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon \\ \varepsilon_1,\varepsilon_2\ge 0}} N_{\varepsilon_1}\big((\bar{x},\bar{y});\operatorname{gph} F\big) \cap \big(X^* \times \mu\mathbb{B}^*\big) \\ & + \frac{\varepsilon_2}{\eta}\mathbb{B}_{X^*\times Y^*}. \end{aligned} \tag{7.118}$$

To justify the equality in (7.108), let us fix $\varepsilon > 0$ and pick any $(x^*, y^*) \in \mathbb{B}_{X^*} \times Y^*$ satisfying $y^* \in D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})(x^*)$, which means that $(-x^*, y^*) \in N_{\varepsilon}((\bar{x},\bar{y});\operatorname{gph} F)$. It follows from (7.118) that there exist a number $\varepsilon_1 \in [0,\varepsilon]$ and ε_1-normals $(u^*, v^*) \in N_{\varepsilon_1}((\bar{x},\bar{y});\operatorname{gph} F)$ satisfying the estimates $\|v^*\| \le \mu$ and $\|y^* - v^*\| \le (\varepsilon-\varepsilon_1)\eta^{-1}$. Hence, we get the inequalities

$$\|y^*\| \le \|v^*\| + (\varepsilon-\varepsilon_1)\eta^{-1} \le \mu + \varepsilon\eta^{-1}.$$

Observe from (7.105) that the function $\varepsilon \mapsto \|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\|$ is nondecreasing, which implies therefore the relationships

$$\lim_{\varepsilon\downarrow 0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\| = \inf_{\varepsilon>0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\| \le \inf_{\varepsilon>0}\Big(\mu + \varepsilon\eta^{-1}\Big).$$

Letting $\mu \downarrow \operatorname{reg} F(\bar{x};\bar{y})$ above gives us the estimate

$$\lim_{\varepsilon\downarrow 0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\| \le \operatorname{reg} F(\bar{x},\bar{y}). \tag{7.119}$$

It follows from (7.119) that the equality in (7.108) is trivial if $\operatorname{reg} F(\bar{x},\bar{y}) = 0$. Considering further the case of $\operatorname{reg} F(\bar{x},\bar{y}) > 0$, we deduce from the definition of the exact regularity bound that $(\bar{x},\bar{y})$ is *not* a local minimizer of the DC problem (7.112) when $0 < \mu < \operatorname{reg} F(\bar{x},\bar{y})$. Then Lemma 7.34 allows us to find sequences $\varepsilon_k \downarrow 0$ and $(x_k^*, y_k^*) \in \partial_{\varepsilon_k}\varphi_2(\bar{x},\bar{y})$ such that $(x_k^*, y_k^*) \notin \partial_{\varepsilon_k}(\mu\varphi_1)(\bar{x},\bar{y})$ as $k \in \mathbb{N}$. Combining this with (7.116) and (7.117) implies that

$$\|x_k^*\| \le 1 \;\text{ and }\; \|y_k^*\| > \mu \;\text{ for all }\; k \in \mathbb{N}. \tag{7.120}$$

Since $B_{2r}(\bar{y}) \subset F(\bar{x} + \mathbb{B}_X)$ as mentioned, (7.117) and (7.120) yield

$$\begin{aligned} \varepsilon_k \ge & \sup_{(x,y)\in\operatorname{gph}F}\Big(\langle x_k^*, x-\bar{x}\rangle + \langle y_k^*, y-\bar{y}\rangle\Big) \\ \ge & \sup\nolimits_{y\in B_{2r}(\bar{y})}\big(\langle y_k^*, y-\bar{y}\rangle\big) - \|x_k^*\| \ge 2r\|y_k^*\| - \|x_k^*\| \ge 2r\mu - \|x_k^*\|. \end{aligned} \tag{7.121}$$

By $\varepsilon_k \downarrow 0$ as $k \to \infty$, we have $\|x_k^*\| \geq 2r\mu - \varepsilon_k \geq r\mu$ for sufficiently large k. Suppose without loss of generality that $\|x_k^*\| \geq r\mu$ for all $k \in \mathbb{N}$, and define

$$\widetilde{y}_k^* := y_k^* \|x_k^*\|^{-1}, \;\; \widetilde{x}_k^* := -x_k^* \|x_k^*\|^{-1}, \;\; \text{and} \;\; \widetilde{\varepsilon}_k := \varepsilon_k \|x_k^*\|^{-1}.$$

Then $\|\widetilde{x}_k^*\| = 1$, $\widetilde{\varepsilon}_k \downarrow 0$, and $\widetilde{y}_k^* \in D^*_{\widetilde{\varepsilon}_k} F^{-1}(\bar{y}, \bar{x})(\widetilde{x}_k^*)$. We get from (7.120) that

$$\sup\left\{\|y^*\| \;\middle|\; y^* \in D^*_{\widetilde{\varepsilon}_k} F^{-1}(\bar{y}, \bar{x})(y^*), \;\; x^* \in S_{X^*}\right\} \geq \|\widetilde{y}_k^*\| = \|y_k^*\| \cdot \|x_k^*\|^{-1} > \mu.$$

Letting $k \to \infty$ and $\mu \uparrow \operatorname{reg} F(\bar{x}, \bar{y})$ tells us that

$$\lim_{\varepsilon \downarrow 0} \sup\left\{\|y^*\| \middle| \;\; y^* \in D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})(x^*), \;\; x^* \in S_{X^*}\right\} \geq \operatorname{reg} F(\bar{x}; \bar{y}),$$

which yields the equality in (7.108) by using (7.119).

It remains to prove formula (7.109). By the arguments similar to those following (7.121), we arrive at the relationships

$$D^*_\varepsilon F(\bar{x}, \bar{y})(y^*) \cap r\mathbb{B}_{X^*} = \emptyset \;\; \text{for all} \;\; 0 < \varepsilon < r \;\; \text{and} \;\; y^* \in S_{Y^*}. \tag{7.122}$$

Pick any $(x^*, y^*) \in X^* \times S_{Y^*}$ such that $x^* \in D^*_\varepsilon F(\bar{x}, \bar{y})(y^*)$ for some $0 < \varepsilon < r$. Define further $\widehat{x}^* := -x^*\|x^*\|^{-1}$, $\widehat{y}^* := -y^*\|x^*\|^{-1}$, and $\widehat{\varepsilon} := \varepsilon\|x^*\|^{-1}$. This ensures that $\widehat{x}^* \in S_{X^*}$, $\|\widehat{y}^*\| = \|x^*\|^{-1}$, and $\widehat{y}^* \in D^*_{\widehat{\varepsilon}} F^{-1}(\bar{y}, \bar{x})(\widehat{x}^*)$. Observe from (7.122) that $\widehat{\varepsilon} \leq \varepsilon r^{-1}$, and thus we have

$$\|x^*\|^{-1} = \|\widehat{y}^*\| \leq \|D^*_{\widehat{\varepsilon}} F^{-1}(\bar{y}, \bar{x})\| \leq \|D^*_{\varepsilon r^{-1}} F^{-1}(\bar{y}, \bar{x})\|.$$

This together with (7.108) yields the inequality "$\geq$" in (7.109) by letting $\varepsilon \downarrow 0$.

To justify the converse inequality in (7.109), note first that it obviously holds when $\operatorname{reg} F(\bar{x}, \bar{y}) = 0$. If $\operatorname{reg} F(\bar{x}, \bar{y}) > 0$, we get from the equality in (7.108) and the norm definition in (7.105) that there exists a sufficiently small number $0 < s < \operatorname{reg} F(\bar{x}, \bar{y})$ ensuring the validity of the condition

$$D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})(x^*) \cap s\mathbb{B}_{Y^*} = \emptyset \;\; \text{for all} \;\; 0 < \varepsilon < s \;\; \text{and} \;\; x^* \in S_{X^*}.$$

The arguments similar to those after (7.122) give us the estimate

$$\|D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})\| \leq \sup\left\{\frac{1}{\|x^*\|} \middle| x^* \in D^*_{\varepsilon s^{-1}} F(\bar{x}, \bar{y})(y^*), \, y^* \in S_{Y^*}\right\}. \tag{7.123}$$

Indeed, pick $(y^*, x^*) \in Y^* \times S_{X^*}$ with $y^* \in D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})(x^*)$ and get $\|y^*\| > s$. Then for $\widetilde{x}^* := x^*\|y^*\|^{-1}$ and $\widetilde{y}^* := y^*\|y^*\|^{-1}$, we have $\|\widetilde{x}^*\|^{-1} = \|y^*\|$, $\widetilde{y}^* \in S_{Y^*}$, and $\widetilde{x}^* \in D^*_{\frac{\varepsilon}{\|y^*\|}} F(\bar{x}, \bar{y})(\widetilde{y}^*) \subset D^*_{\varepsilon s^{-1}} F(\bar{x}, \bar{y})(\widetilde{y}^*)$, which yields (7.123). Combining finally (7.108) with (7.123) justifies the inequality "$\leq$" in (7.109) and thus completes the proof of the theorem. △

The following consequence of Theorem 7.35 and the classical Brøndsted-Rockafellar density theorem of convex analysis (see, e.g., [638, Theorem 3.17]) establish a precise formula for the exact regularity bound of a closed convex multifunction F between Banach spaces by using the coderivative of F^{-1} instead of its ε-counterparts while involving points around the reference one.

Corollary 7.36 (Calculating the Exact Regularity Bound via Coderivatives at Nearby Points). *In the setting of Theorem 7.35, we have*

$$\operatorname{reg} F(\bar{x},\bar{y}) = \lim_{\varepsilon\downarrow 0}\Big[\sup\Big\{\|D^*F^{-1}(y,x)\| \text{ over } (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\}\Big]. \tag{7.124}$$

Proof. To verify the inequality "$\geq$" in (7.124), observe from (7.110) that for any $\mu > \operatorname{reg} F(\bar{x},\bar{y})$ and any sufficiently small $\varepsilon > 0$, we get

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \leq \mu \operatorname{dist}\big(y; F(x)\big) \text{ for all } (x,y)\in B_\varepsilon(\widetilde{x},\widetilde{y})$$

whenever $(\widetilde{x},\widetilde{y})\in B_\varepsilon(\bar{x},\bar{y})$. It follows from (7.108) that

$$\mu \geq \lim_{\eta\downarrow 0}\|D^*_\eta F^{-1}(\widetilde{y},\widetilde{x})\| \geq \|D^*F^{-1}(\widetilde{y},\widetilde{x})\| \text{ for all } (\widetilde{x},\widetilde{y})\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y}).$$

This clearly implies the estimate

$$\mu \geq \lim_{\varepsilon\downarrow 0}\Big[\sup\Big\{\|D^*F^{-1}(y,x)\| \;\Big|\; (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\}\Big].$$

Letting there $\mu\downarrow \operatorname{reg} F(\bar{x},\bar{y})$, we arrive at the inequality "$\geq$" in (7.124).

To prove the converse inequality in (7.124), take an arbitrary $\varepsilon > 0$, and observe from Theorem 7.35 that $\operatorname{reg} F(\bar{x},\bar{y}) \leq \|D^*_{\varepsilon^2}F^{-1}(\bar{y},\bar{x})\|$. This allows us to find $(x^*,y^*)\in X^*\times Y^*$ satisfying the condition $y^*\in D^*_{\varepsilon^2}F^{-1}(\bar{y},\bar{x})(x^*)$, i.e., $(-x^*,y^*)\in N_{\varepsilon^2}((\bar{x},\bar{y});\operatorname{gph} F)$. We have furthermore that

$$\|x^*\|\leq 1 \text{ and } \|y^*\|+\varepsilon \geq \operatorname{reg} F(\bar{x},\bar{y}). \tag{7.125}$$

By the Brøndsted-Rockafellar theorem, there are $(x_\varepsilon,y_\varepsilon)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})$ and $(-x^*_\varepsilon,y^*_\varepsilon)\in N((x_\varepsilon,y_\varepsilon);\operatorname{gph} F)$ satisfying $\|x^*_\varepsilon - x^*\|\leq\varepsilon$ and $\|y^*_\varepsilon - y^*\|\leq\varepsilon$. Thus we get $\|x^*_\varepsilon\|\leq\|x^*\|+\varepsilon\leq 1+\varepsilon$ and $\|y^*\|\leq\|y^*_\varepsilon\|+\varepsilon$, and thus

$$\|y^*\|\leq (1+\varepsilon)\|D^*F^{-1}(y_\varepsilon,x_\varepsilon)\|+\varepsilon.$$

Combining this with (7.125) yields the estimate

$$\operatorname{reg} F(\bar{x},\bar{y}) \leq (1+\varepsilon)\sup\Big\{\|D^*F^{-1}(y,x)\| \;\Big|\; (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\}+2\varepsilon,$$

which ensures the inequality "$\leq$" in (7.124) while letting $\varepsilon\downarrow 0$. △

The next consequence of Theorem 7.35 concerns calculating the exact covering bound of closed- and convex-graph multifunctions. This is indeed a major result of this section, which accumulates the previous developments.

Corollary 7.37 (Calculating the Exact Covering Bound for Convex-Graph Multifunctions). *Given a point* $(\bar{x},\bar{y}) \in \operatorname{gph} F$ *with* $\bar{y} \in \operatorname{int}(\operatorname{rge} F)$*, the exact covering bound of* F *at* $(\bar{x},\bar{y})$ *is calculated by*

$$\operatorname{cov} F(\bar{x},\bar{y}) = \lim_{\varepsilon\downarrow 0}\Big[\inf_{x^*\in X^*}\inf_{y^*\in S_{Y^*}}\Big(\|x^*\| + \frac{\sigma_{\operatorname{gph} F-(\bar{x},\bar{y})}(x^*,y^*)}{\varepsilon}\Big)\Big].$$

Proof. Define $\Omega := \operatorname{gph} F - (\bar{x},\bar{y})$. Since the number $\operatorname{cov} F(\bar{x},\bar{y})$ is the reciprocal of $\operatorname{reg} F(\bar{x},\bar{y})$, it suffices to show that

$$\operatorname{reg} F(\bar{x},\bar{y}) = \lim_{\varepsilon\downarrow 0}\Big[\sup_{x^*\in X^*}\sup_{y^*\in S_{Y^*}}\Big(\|x^*\| + \frac{\sigma_\Omega(x^*,y^*)}{\varepsilon}\Big)^{-1}\Big] =: \alpha. \quad (7.126)$$

By (7.109), we find sequences $\varepsilon_k \downarrow 0$ and $(x_k^*,y_k^*) \in X^* \times S_{Y^*}$ such that $x_k^* \in D^*_{\varepsilon_k}F(\bar{x},\bar{y})(y_k^*)$, which amounts to $\sigma_\Omega(x_k^*,-y_k^*) \le \varepsilon_k$ due to (7.103), and that $\|x_k^*\|^{-1} \to \operatorname{reg} F(\bar{x},\bar{y})$ as $k\to\infty$. Hence

$$\begin{aligned}\sup_{x^*\in X^*}\sup_{y^*\in S_{Y^*}}\Big(\|x^*\| + \frac{\sigma_\Omega(x^*,y^*)}{\sqrt{\varepsilon_k}}\Big)^{-1} &\ge \Big(\|x_k^*\| + \frac{\sigma_\Omega(x_k^*,-y_k^*)}{\sqrt{\varepsilon_k}}\Big)^{-1}\\ &\ge \Big(\|x_k^*\| + \sqrt{\varepsilon_k}\Big)^{-1},\end{aligned}$$

which yields the inequality "$\le$" in (7.126) by passing to the limit as $k\to\infty$.

Conversely, if the right-hand side of (7.126) is 0, the equality in (7.126) is obvious. Otherwise, we find sequences $\widetilde{\varepsilon}_k \downarrow 0$ and $(\widetilde{x}_k^*,\widetilde{y}_k^*) \in X^*\times S_{Y^*}$ with

$$\beta < \Big(\|\widetilde{x}_k^*\| + \frac{\sigma_\Omega(\widetilde{x}_k^*,\widetilde{y}_k^*)}{\widetilde{\varepsilon}_k}\Big)^{-1} \to \alpha \ \text{ as } \ k\to\infty \quad (7.127)$$

for some $\beta > 0$. It follows that $\sigma_\Omega(\widetilde{x}_k^*,\widetilde{y}_k^*) \le \widetilde{\varepsilon}_k\beta^{-1}$ for all $k\in\mathbb{N}$, which gives us $\widetilde{x}_k^* \in D^*_{\widehat{\varepsilon}_k}F(\bar{x},\bar{y})(-\widetilde{y}_k^*)$ with $\widehat{\varepsilon}_k := \widetilde{\varepsilon}_k\beta^{-1} \to 0$ by (7.103). Hence, we have

$$\begin{aligned}&\Big(\|\widetilde{x}_k^*\| + \frac{\sigma_\Omega(\widetilde{x}_k^*,\widetilde{y}_k^*)}{\widetilde{\varepsilon}_k}\Big)^{-1} \le \|\widetilde{x}_k^*\|^{-1}\\ &\le \sup\Big\{\|x^*\|^{-1}\,\Big|\, x^*\in D^*_{\widehat{\varepsilon}_k}F(\bar{x},\bar{y})(y^*),\ y^*\in S_{Y^*}\Big\}.\end{aligned} \quad (7.128)$$

Substituting the regularity formula (7.109) into (7.128) and using (7.127), we arrive at $\alpha \le \operatorname{reg} F(\bar{x},\bar{y})$ and thus complete the proof of the corollary. △

Finally in this subsection, let us introduce an additional condition, which helps us to remove $\varepsilon > 0$ in the exact bound formula (7.108) and get the precise equality (7.130) for calculating the exact regularity bound of closed- and convex-graph multifunctions between arbitrary Banach spaces as in case (3.8) of set-valued mapping between finite-dimensional spaces. Note that assumption (8.84) holds in the SIP setting of Subsection 7.4.3 and also when $\dim Y < \infty$, while X is an arbitrary Banach space.

Theorem 7.38 (Calculating the Exact Regularity Bound via the Basic Coderivative Norm). *In the setting of Theorem* 3.8*, assume in addition that*

$$\Lambda(S_{Y^*}) \subset S_{Y^*}, \tag{7.129}$$

where the set $\Lambda(S_{Y^*})$ *is defined sequentially by*

$$\Lambda(S_{Y^*}) := \Big\{ y^* \in Y^* \Big| \ \exists\, \varepsilon_k \downarrow 0,\ \ y_k^* \in S_{Y^*} \text{ such that } D^*_{\varepsilon_k} F(\bar{x}, \bar{y})(y_k^*) \neq \emptyset \\ \text{and } \ y^* \ \text{ is a weak}^* \text{ cluster point of } \ y_k^* \Big\}.$$

Then the exact regularity bound is calculated by

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \|D^* F^{-1}(\bar{y}, \bar{x})\|. \tag{7.130}$$

If furthermore $\operatorname{reg} F(\bar{x}, \bar{y}) > 0$*, we get the improved formula*

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \sup \Big\{ \|x^*\|^{-1} \Big|\ x^* \in D^* F(\bar{x}, \bar{y})(y^*),\ \ y^* \in S_{Y^*} \Big\}. \tag{7.131}$$

Proof. Note that the equality in (7.130) is trivial when $\operatorname{reg} F(\bar{x}, \bar{y}) = 0$. Otherwise, it follows from (7.109) that there are sequences $\varepsilon_k \downarrow 0$ and $x_k^* \in D^*_{\varepsilon_k} F(\bar{x}, \bar{y})(y_k^*)$ such that $\|x_k^*\| > 0$, $\|y_k^*\| = 1$, and

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \lim_{k \to \infty} \|x_k^*\|^{-1}. \tag{7.132}$$

Since the sequence $\{x_k^*\}$ is bounded by (7.132), we get from (7.129) and Alaoglu-Bourbaki theorem that there is a subnet $(x_\alpha^*, y_\alpha^*, \varepsilon_\alpha)$ of $(x_k^*, y_k^*, \varepsilon_k)$ weak* converging to some $(\bar{x}^*, \bar{y}^*, 0) \in X^* \times S_{Y^*} \times \mathbb{R}$. Note further that

$$\langle \bar{x}^*, x - \bar{x} \rangle - \langle \bar{y}^*, y - \bar{y} \rangle = \lim_{\alpha} \langle x_\alpha^*, x - \bar{x} \rangle - \langle y_\alpha^*, y - \bar{y} \rangle \leq \limsup_{\alpha} \varepsilon_\alpha = 0$$

for all $(x, y) \in \operatorname{gph} F$, which yields $\bar{x}^* \in D^* F(\bar{x}, \bar{y})(\bar{y}^*)$. Moreover, the classical uniform boundedness principle tells us that $\|\bar{x}^*\| \leq \liminf_{\alpha} \|x_\alpha^*\|$. This together with (7.132) ensures the validity of the inequalities

$$\operatorname{reg} F(\bar{x}, \bar{y}) \leq \frac{1}{\|\bar{x}^*\|} \leq \sup \Big\{ \frac{1}{\|x^*\|} \Big| x^* \in D^* F(\bar{x}, \bar{y})(y^*),\ \|y^*\| = 1 \Big\}. \tag{7.133}$$

Combining the latter with (7.109) yields (7.131). Furthermore, observe that $\widehat{x}^* := \bar{x}^* \|\bar{x}^*\|^{-1} \in S_{X^*}$ and $\widehat{y}^* := \bar{y}^* \|\bar{x}^*\|^{-1} \in D^* F^{-1}(\bar{y}, \bar{x})(\widehat{x}^*)$. Hence we get from (7.133) and (7.106) the relationships

$$\operatorname{reg} F(\bar{x}, \bar{y}) \leq \|\widehat{y}^*\| = \|\bar{x}^*\|^{-1} \leq \|D^* F^{-1}(\bar{y}, \bar{x})\|,$$

which together with (7.108) yield (7.130) and thus complete the proof. △

It is obvious that assumption (7.129) automatically holds when Y is finite-dimensional. More subtle, it also holds under the validity of the condition

$$\text{cl}^*\Big\{y^* \in S_{Y^*}\Big|\, \sigma_\Omega(x^*, y^*) < \infty,\ x^* \in X^*\Big\} \subset S_{Y^*}, \tag{7.134}$$

with $\Omega := \text{gph}\, F - (\bar{x}, \bar{y})$ due to the proper/strict inclusion

$$\begin{aligned}\Lambda(S_{Y^*}) &\subset \text{cl}^*\Big[\bigcup_{\varepsilon\geq 0}\Big\{y^* \in S_{Y^*}\Big|\, D^*_\varepsilon F(\bar{x}, \bar{y})(y^*) \neq \emptyset\Big\}\Big]\\ &= \text{cl}^*\Big\{y^* \in S_{Y^*}\Big|\, \sigma_\Omega(x^*, y^*) < \infty,\ x^* \in X^*\Big\}.\end{aligned} \tag{7.135}$$

7.4.3 Applications to Infinite Convex Constraint Systems

Here we develop applications of the results obtained in Subsection 7.4.2 to the special class of set-valued mappings $F : X \rightrightarrows Y := Z \times l^\infty(T)$ given by

$$F(x) := \begin{cases}\Big\{(z, p) \in Y\Big|\, Ax = z,\ f_t(x) \leq p_t,\ t \in T\Big\} & \text{if } x \in C,\\ \emptyset \quad \text{otherwise,}\end{cases} \tag{7.136}$$

which describes, in particular, sets of feasible solutions in parameterized SIPs with infinitely many inequality as well as equality and geometric constraints.

The data of (7.136) are as follows: $A : X \to Z$ is a bounded linear operator between two Banach spaces; the functions $f_t : X \to \overline{\mathbb{R}}$ are l.s.c. and convex for all t from the arbitrary index set T; and C is a closed and convex subset of X with nonempty interior. These assumptions clearly imply that F in (7.136) is closed- and convex-graph multifunction, and so we can implement the results on metric regularity at $(x, (z, p)) \in \text{gph}\, F$ obtained above to the infinite constraint system (7.136) provided the validity of the underlying condition

$$(z, p) \in \text{int}(\text{rge}\, F). \tag{7.137}$$

Note that this condition clearly implies that $z \in \text{int}(AX)$, which ensures that A is an open mapping, and hence it must be surjective.

Throughout this section, we denote $f(x) := \sup_{t\in T} f_t(x)$ and suppose that the space $Z \times l^\infty(T)$ is equipped with the maximum product norm

$$\|(z, p)\| = \max\Big\{\|z\|, \|p\|\Big\} \text{ for all } z \in Z,\ p \in l^\infty(T).$$

As mentioned above, F is metrically regular around $(x, (z, p)) \in \text{gph}\, F$ if and only if condition (7.137) holds. This motivates us to introduce a qualification condition via the initial data of (7.136), which ensures the validity of (7.137) and extend the usual strong SSC typically employed for infinite linear and convex inequality systems to the more general constraint case of (7.136).

Definition 7.39 (Bounded Strong Slater Condition). *We say that the infinite system* (7.136) *satisfies the* BOUNDED STRONG SLATER CONDITION (BOUNDED SSC) *at* $(z, p) \in Z \times l^\infty(T)$ *if there is* $\widehat{x} \in \operatorname{int} C$ *such that the function* f *is bounded from above around* $\widehat{x}$*, that* $A\widehat{x} = z$*, and that*

$$\sup_{t\in T}[f_t(\widehat{x}) - p_t] < 0. \tag{7.138}$$

Note that the Slater-type notion introduced in Definition 7.39 is generally different for infinite linear and convex systems from the strong Slater condition studied and applied in Sections 7.1 and 7.3. In the particular case of $C = X$, $Z = \{0\}$, and $f_t(x) = \langle a_t^*, x\rangle - b_t$ with $(a_t^*, b_t) \in X^* \times \mathbb{R}$ considered in Section 7.1, our bounded SSC is clearly weaker than the usual SSC provided that the coefficient set $\{a_t^* |\ t \in T\}$ is bounded in X^*, which is the underlying assumption therein. The following example demonstrates that it may be *strictly weaker* even in the one-dimensional case of $X = \mathbb{R}$.

Example 7.40 (Bounded from Above Linear Constraint Functions with Unbounded Coefficients). Let $X = \mathbb{R}$, $Z = \{0\}$, $T = (0, 1)$, and $f_t(x) = -\frac{1}{t}x + t$ in (7.136). Note that

$$f_t(x) = -\frac{1}{t}x + t = -\frac{1}{t}x - t + 2t \le -2\sqrt{x} + 2t \quad \text{for all} \quad x > 0,\ t \in T.$$

Taking $\widehat{x} = 4$ and $\bar{x} = 1$, we observe that $f_t(\widehat{x}) < -2$, $f_t(\bar{x}) \le 0$, and the supremum function f is bounded from above around $\widehat{x}$. However, the coefficient set $\left\{ -\frac{1}{t} \middle|\ t \in T\right\}$ is obviously unbounded.

The next proposition shows that the bounded SSC introduced is a sufficient condition for the validity of (7.137) while being in fact "almost necessary" for this, up to the upper boundedness of the supremum function f.

Proposition 7.41 (Bounded Strong Slater Condition and Metric Regularity). *Let* $(z, p) \in \operatorname{rge} F$ *for the infinite system* (7.136)*. Then the bounded SSC for* F *at* (z, p) *implies the validity of* (7.137)*. Conversely, if* (7.137) *holds, then there is* $\widehat{x} \in \operatorname{int} C$ *such that* $A\widehat{x} = z$ *and that* (7.138) *is satisfied.*

Proof. To verify the first part, suppose that the bounded SSC holds for F at (z, p). Then there are $\widehat{x} \in \imath C$ and $\varepsilon > 0$ such that the supremum function f is upper bounded around $\widehat{x}$ with $A(\widehat{x}) = z$ and $f^p(\widehat{x}) < -\varepsilon$, where

$$f^p(\cdot) := \sup_{t\in T}\left\{f_t(\cdot) - p_t\right\} \quad \text{for} \quad p \in l^\infty(T).$$

Note that the function $f^p(\cdot)$ is obviously a proper, l.s.c., convex, and upper bounded around $\widehat{x}$. We know from convex analysis that in this case it is continuous at $\widehat{x}$. Since A is surjective and $\widehat{x} \in \operatorname{int} C$, the classical open mapping theorem allows us to find $0 < s \le \frac{\varepsilon}{2}$ such that $B_s(z) \subset A(B_r(\widehat{x}) \cap C)$ for $r > 0$. Picking any

$(z', p') \in B_s(z, p)$, there exists $x \in B_r(\widehat{x}) \cap C$ with $Ax = z'$ and so that for each $t \in T$, we have

$$\begin{aligned} f^{p'}(x) &\le f^p(x) + s \le f^p(x) - f^p(\widehat{x}) + s + f^p(\widehat{x}) \\ &\le f^p(x) - f^p(\widehat{x}) + s - \varepsilon \le f^p(x) - f^p(\widehat{x}) - \varepsilon/2 \le 0 \end{aligned}$$

when r is sufficiently small. This yields $(z', p') \in \operatorname{rge} F$, which implies in turn that the inclusion $B_s(z, p) \subset \operatorname{rge} F$ holds.

To justify the necessity part, observe that $(z, (p_t - \varepsilon)_{t \in T}) \in \operatorname{rge} F$ for some $\varepsilon > 0$ if $(z, p) \in \operatorname{int}(\operatorname{rge} F)$. Hence there is $\widehat{x} \in X$ such that $A\widehat{x} = z$ and $f_t(\widehat{x}) - p_t \le -\varepsilon$ as $t \in T$, which thus completes the proof. △

Now we proceed with calculating the exact regularity bound for the constraint system (7.136) at $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$ based on the results of Subsection 7.4.2. It follows from Theorem 7.35 that $\operatorname{reg} F(\bar{x}, (\bar{z}, 0))$ can be calculated via the norms of ε-coderivatives. The next result, which is certainly of its own interest, accomplishes an important step in this direction.

Theorem 7.42 (Explicit Form of ε-Coderivatives for Infinite Convex Systems). *Let F be the infinite constraint system (7.136), and let $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$. Then for each $\varepsilon \ge 0$, we have the ε-coderivative representation*

$$D^*_\varepsilon F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z \times l^\infty(T))^*}\big) = \big\{x^* \big| (x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M\big\}, \tag{7.139}$$

where $x^* \in X^*$ and M is defined, with $C_0 := C \cap \operatorname{dom} f$, by

$$M := \bigcup_{z^* \in \mathbb{B}_{Z^*}} \operatorname{cl}^*\Big[(1 - \|z^*\|)\operatorname{co}\Big(\bigcup_{t \in T} \operatorname{epi} f_t^*\Big) + \operatorname{epi} \delta^*(\cdot; C_0)\Big] + (A^* z^*, \langle z^*, \bar{z}\rangle).$$

Proof. To verify the inclusion "⊂" in (7.139), pick $(z^*, p^*) \in S_{(Z \times l^\infty(T))^*}$ and $x^* \in D^*_\varepsilon F(\bar{x}, (\bar{z}, 0))(z^*, p^*)$. Then we have $\|z^*\| + \|p^*\| = 1$ and

$$\langle x^*, x - \bar{x}\rangle - \langle z^*, z - \bar{z}\rangle - \langle p^*, p\rangle \le \varepsilon \;\text{ for all }\; (x, z, p) \in \operatorname{gph} F,$$

which can be equivalently represented by

$$\begin{aligned} &\langle x^* - A^* z^*, x - \bar{x}\rangle - \langle p^*, p\rangle \le \varepsilon \\ &\text{if } (x, p) \in C_0 \times l^\infty(T),\ f_t(x) - \langle \delta_t, p\rangle \le 0,\ t \in T, \end{aligned} \tag{7.140}$$

via the Dirac measure $\delta_t \in (l^\infty(T))^*$ at t. It follows from the extended Farkas lemma in Proposition 7.3 that (7.140) reads as

$$\begin{aligned} &\big(p^*, x^* - A^* z^*, \langle x^* - A^* z^*, \bar{x}\rangle + \varepsilon\big) \\ &\in \operatorname{cl}^*\Big[\operatorname{cone}\Big\{\bigcup_{t \in T} \{\delta_t\} \times \operatorname{epi} f_t^*\Big\} + \{0\} \times \operatorname{epi} \delta^*(\cdot; C_0)\Big]. \end{aligned} \tag{7.141}$$

Hence there exist nets $\{\lambda_\nu\}_{\nu \in \mathcal{N}} \subset \mathbb{R}^{(T)}_+$, $\{(v^*_\nu, s_\nu)\}_{\nu \in \mathcal{N}} \subset \operatorname{epi} \delta^*(\cdot; C_0)$, and $\{(u^*_{t\nu}, r_{t\nu})\}_{\nu \in \mathcal{N}} \subset \operatorname{epi} f_t^*$ for each $t \in T$ such that

$$\big(p^*, x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(\delta_t, u^*_{t\nu}, r_{t\nu}) + (0, v^*_\nu, s_\nu)\Big].$$

Observe from the latter equality that $p^* = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\delta_t$. Thus we have

$$\begin{aligned}&\limsup_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu} \ge \sup_{\|p\|\le 1}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}p_t\\ &= \sup\nolimits_{\|p\|\le 1}\langle p^*, p\rangle = \|p^*\| \ge \langle p^*, e\rangle = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\end{aligned} \tag{7.142}$$

with $e \in l^\infty(T)$ satisfying $e_t = 1$ for all $t \in T$. This yields

$$1 - \|z^*\| = \|p^*\| = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}. \tag{7.143}$$

If $\|z^*\| = 1$, we get from the above the relationships

$$\begin{aligned}&\langle x^* - A^*z^*, x - \bar{x}\rangle - \varepsilon = \langle x^* - A^*z^*, x\rangle - (\langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon)\\ &= \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\langle u^*_{t\nu}, x\rangle + \langle v^*_\nu, x\rangle\Big] - \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}r_{t\nu} - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\big(\langle u^*_{t\nu}, x\rangle - f_t(x) - r_{t\nu} + f(x)\big) + \langle v^*_\nu, x\rangle - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\big(f^*_t(u^*_{t_\nu}) - r_{t\nu} + f(x)\big) + \delta^*(\cdot; C_0)(v^*_\nu) - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}f(x) = 0 \quad\text{for any}\quad x \in C_0.\end{aligned}$$

It follows from (7.30) that $(x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon) \in \operatorname{epi}\delta^*(\cdot; C_0)$; so

$$\begin{aligned}(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) &\in \operatorname{epi}\delta^*(\cdot; C_0) + (A^*z^*, \langle A^*z^*, \bar{x}\rangle)\\ &= \operatorname{epi}\delta^*(\cdot; C_0) + (A^*z^*, \langle z^*, \bar{z}\rangle) \subset M.\end{aligned}$$

If $\|z^*\| < 1$, it doesn't restrict the generality due to (7.143) to suppose that $\sum_{t\in T}\lambda_{t\nu} > 0$ for all $\nu \in \mathcal{N}$ and to define $\widetilde{\lambda}_{t\nu} := \dfrac{\lambda_{t\nu}}{\sum_{t'\in T}\lambda_{t'\nu}}$ for each $t \in T$ and $\nu \in \mathcal{N}$. It tells us by the "w^*-lim" expression after formula (7.141) that

$$\begin{aligned}&(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle)\\ &= (1 - \|z^*\|)w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\widetilde{\lambda}_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle) \subset M.\end{aligned}$$

Thus we get $(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M$ and justify the inclusion "$\subset$" in (7.139).

To verify the converse inclusion in (7.139), pick any element $x^* \in X^*$ satisfying $(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M$. Hence we find a unit functional $z^* \in \mathbb{B}_{Z^*}$ as well as nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}^{(T)}_+$, $\{(v^*_\nu, s_\nu)\}_{\nu\in\mathcal{N}} \subset \operatorname{epi}\delta^*(\cdot; C_0)$, and $\{(u^*_{t\nu}, r_{t\nu})\}_{\nu\in\mathcal{N}} \subset \operatorname{epi} f^*_t$, $t \in T$, such that $\sum_{t\in T}\lambda_{t\nu} = 1$ and

$$(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) = (1 - \|z^*\|)w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu)\Big]$$
$$+(A^*z^*, \langle z^*, \bar{z}\rangle).$$

Defining $p^*_\nu := (1 - \|z^*\|)\sum_{t\in T}\lambda_{t\nu}\delta_t$, deduce that $\|p^*_\nu\| = 1 - \|z^*\|$ while arguing similarly to the proof of (7.142). It follows from the classical Alaoglu-Bourbaki theorem that there exists a subnet of p^*_ν (without relabeling), which weak* converges to some $p^* \in \mathbb{B}_{(l^\infty(T))^*}$. By using again the arguments as in the proof of (7.142), we get $\|p^*\| = 1 - \|z^*\|$ and then obtain (7.141). Due to the equivalence between (7.140) and (7.141), this justifies the inclusion "$\supset$" in (7.139) and thus completes the proof of the theorem. △

In the *coderivative* case of Theorem 7.42 (i.e., if $\varepsilon = 0$), we can equivalently modify the representation in (7.139) and provide its further specification.

Proposition 7.43 (Explicit Forms of the Coderivative for Infinite Convex Systems). *Let* $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$ *for the constraint system* (7.136). *Then we have the coderivative representation*

$$D^*F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times l^\infty(T))^*}\big) = \big\{x^* \in X^* \big|\ (x^*, \langle x^*, \bar{x}\rangle) \in L\big\} \tag{7.144}$$

with $L := \displaystyle\bigcup_{z^*\in\mathbb{B}_{Z^*}} \operatorname{cl}^*\Big[(1-\|z^*\|)\operatorname{co}\Big(\bigcup_{t\in T}\operatorname{gph} f^*_t\Big)+\operatorname{gph}\delta^*(\cdot; C_0)\Big]+(A^*z^*, \langle z^*, \bar{z}\rangle).$
Furthermore, the term $\operatorname{gph}\delta^*(\cdot; C_0)$ *above can be dropped if* $\bar{x} \in \operatorname{int} C_0$.

Proof. To verify the inclusion "$\subset$" in (7.144), for any $x^* \in D^*F(\bar{x}, (\bar{z}, 0))(z^*, p^*)$ with $\|z^*\| + \|p^*\| = 1$, we deduce from the proof of Theorem 7.42 the validity of inclusion (7.140) with $\varepsilon = 0$. This allows us to find nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}^{(T)}_+$, $\{\rho_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+$, $\{(v^*_\nu, s_\nu)\}_{\nu\in\mathcal{N}} \subset \operatorname{gph}\delta^*(\cdot; C_0)$, and $\{(u^*_{t\nu}, r_{t\nu})\}_{\nu\in\mathcal{N}} \subset \operatorname{gph} f^*_t$ for each $t \in T$ providing the limiting representation

$$\begin{aligned}\big(p^*, x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}(\delta_t, u^*_{t\nu}, r_{t\nu})&\\ +(0, v^*_\nu, s_\nu) + (0, 0, \rho_\nu).&\end{aligned} \tag{7.145}$$

Similarly to the proof of Theorem 7.42, suppose without loss of generality that $\sum_{t\in T}\lambda_{t\nu} = 1 - \|z^*\|$ for all $\nu \in \mathcal{N}$ and then get

$$r_{t\nu} = f^*_t(u^*_{t\nu}) \ge \langle u^*_{t\nu}, \bar{x}\rangle - f_t(\bar{x}) \ge \langle u^*_{t\nu}, \bar{x}\rangle \ \text{ and }\ s_\nu = \delta^*(\cdot; C_0)(v^*_\nu) \ge \langle v^*_\nu, \bar{x}\rangle.$$

This implies together with (7.145) the relationships

$$\langle x^* - A^*z^*, \bar{x}\rangle = \lim_{\nu\in\mathcal{N}} \Big[\sum_{t\in T} \lambda_{t\nu} r_{t\nu} + s_\nu + \rho_\nu\Big] \geq \limsup_{\nu\in\mathcal{N}} \Big[\sum_{t\in T} \lambda_{t\nu} \langle u^*_{t\nu}, \bar{x}\rangle + \langle v^*_\nu, \bar{x}\rangle$$
$$+\rho_\nu\Big] \geq \langle x^* - A^*z^*, \bar{x}\rangle + \limsup_{\nu\in\mathcal{N}} \rho_\nu,$$

which ensure that $\limsup_{\nu\in\mathcal{N}} \rho_\nu = 0$. Then it follows from (7.145) that

$$(x^*, \langle x^*, \bar{x}\rangle) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu) + (A^*z^*, \langle z^*, \bar{z}\rangle) \in L,$$

and thus we arrive at the inclusion "$\subset$" in (7.144). The verification of the opposite inclusion in (7.144) follows the lines in the proof of Theorem 7.42.

Finally, let $\bar{x} \in \operatorname{int} C_0$ and pick $x^* \in D^*F(\bar{x}, (\bar{z}, 0))(z^*, p^*)$ with $(z^*, p^*) \in S_{(Z\times l^\infty(T))^*}$. Using the notation from the proof of (7.144) above, we have

$$0 = \langle x^* - A^*z^*, \bar{x}\rangle - \langle x^* - A^*z^*, \bar{x}\rangle = \lim_{\nu\in\mathcal{N}} \Big[\sum_{t\in T} \lambda_{t\nu}(\langle u^*_{t\nu}, \bar{x}\rangle - r_{t\nu})$$
$$\langle v^*_\nu, \bar{x}\rangle - s_{t\nu}\Big] \leq -\limsup_{\nu\in\mathcal{N}} \sup_{x\in C_0} \big[\langle v^*_\nu, x\rangle - \langle v^*_\nu, \bar{x}\rangle\big] \leq -\limsup_{\nu\in\mathcal{N}} \eta\|v^*_\nu\|,$$

where $\eta > 0$ is such that $B_\eta(\bar{x}) \subset C_0$. This implies that $\limsup_{\nu\in\mathcal{N}} \|v^*_\nu\| = 0$, and so we can remove $\operatorname{gph} \delta^*(\cdot; C_0)$ in the representation of L in (7.144). △

The next major result provides a precise calculation of the exact regularity bound of the infinite constraint system (7.136) entirely via its initial data.

Theorem 7.44 (Exact Regularity Bound of Infinite Constraint Systems). *Given* $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$ *for the infinite system in* (7.136)*, assume that the bounded SSC from Definition* 7.39 *holds at* $(\bar{z}, 0)$*. Then the exact regularity bound of F at* $(\bar{x}, (\bar{z}, 0))$ *is calculated by*

$$\operatorname{reg} F(\bar{x}, (\bar{z}, 0)) = \lim_{\varepsilon\downarrow 0} \Big[\sup\big\{\|x^*\|^{-1} \,\big|\, (x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M\big\}\Big], \tag{7.146}$$

where M is defined in Theorem 7.42*. If in addition* $0 < \dim Z < \infty$*, then*

$$\begin{aligned} \operatorname{reg} F(\bar{x}, (\bar{z}, 0)) &= \|D^*F^{-1}((\bar{z}, 0), \bar{x})\| \\ &= \sup\Big\{\|x^*\|^{-1} \,\Big|\, (x^*, \langle x^*, \bar{x}\rangle) \in L\Big\}, \end{aligned} \tag{7.147}$$

where the set L is defined in Proposition 7.43.

Proof. It follows from Proposition 7.41 that $(\bar{z}, 0) \in \operatorname{int}(\operatorname{rge} F)$, i.e., the mapping F is metrically regular around $(\bar{x}, (\bar{z}, 0))$. Substituting the ε-coderivative expression from Theorem 7.42 into the exact bound formula (7.109) of Theorem 7.35, we arrive at the limiting representation (7.146).

Let us now justify the equalities in (7.147) under the finite dimensionality of Z. By Theorem 7.38 and Proposition 7.43, we need to check that (7.129) holds and that $\operatorname{reg} F(\bar{x}, (\bar{z}, 0)) > 0$. To proceed, take any $\varepsilon > 0$ and $(z^*, p^*) \in S_{(Z\times l^\infty(T))^*}$

satisfying $D_\varepsilon F(\bar{x},(\bar{z},0))(z^*,p^*)\neq\emptyset$. By the same arguments as in the proofs of (7.141) and (7.143), we get the inclusion

$$p^*\in\big(1-\|z^*\|\big)\,\mathrm{cl}^*\mathrm{co}\,\big\{\delta_t\big|\ t\in T\big\}.$$

It shows that the set $\mathrm{cl}^*\{(z^*,p^*)\in S_{(Z\times l^\infty(T))^*}|\ D^*_\varepsilon F(\bar{x},(\bar{z},0))(z^*,p^*)\neq\emptyset\}$ is contained in the following one:

$$\mathrm{cl}^*\bigcup_{z^*\in\mathbb{B}_{Z^*}}\Big[\big\{z^*\big\}\times\big(1-\|z^*\|\big)\,\mathrm{cl}^*\mathrm{co}\,\big\{\delta_t\big|\ t\in T\big\}\Big]. \tag{7.148}$$

Further, we deduce from the proof of (7.143) that $\mathrm{cl}^*\mathrm{co}\,\{\delta_t|\ t\in T\}\subset S_{(l^\infty(T))^*}$. Since $\dim Z<\infty$, the latter implies that the set in (7.148) is a subset of $S_{(Z\times l^\infty(T))^*}$, which ensures the validity of (7.129).

It remains to verify that $\mathrm{reg}\,F(\bar{x},(\bar{z},0))>0$. We can easily see that

$$D^*F^{-1}\big((\bar{z},0),\bar{x}\big)(x^*)\supset\big\{(z^*,0)\in Z^*\times(l^\infty(T))^*\big|\ A^*z^*=x^*\big\}.$$

Since the operator A is surjective, we clearly have $\|(A^*)^{-1}\|>0$. This allows us to conclude that $\|D^*F^{-1}((\bar{z},0),\bar{x})\|>0$, which yields $\mathrm{reg}\,F(\bar{x},(\bar{z},0))>0$ by Theorem 7.35 and thus completes the proof. △

It immediately follows from Theorem 7.38 that the exact bound formula

$$\mathrm{reg}\,F\big(\bar{x},(\bar{z},0)\big)=\big\|D^*F^{-1}\big((\bar{z},0),\bar{x}\big)\big\| \tag{7.149}$$

holds also in the case of $\dim Z=0$. Recall that the Lipschitzian counterpart of (7.149) is proved for infinite linear inequality systems in Corollary 7.16 and for infinite convex inequality systems in Theorem 7.31 (with no equality and geometric constraints) under the boundedness assumptions therein. As discussed above, these assumptions are essentially stronger than the bounded SSC imposed in Theorem 7.44; see Example 7.40.

A natural question arising from Theorem 7.44 is whether the exact regularity bound expression (7.149) holds for infinite-dimensional spaces Z. The following *counterexample* is constructed for the case of the classical Asplund space $Z=c_0$, which has been already used above (i.e., the space of sequences of real numbers converging to zero and endowed with the supremum norm).

Example 7.45 (Failure of the Exact Bound Formula for Countable Inequality and Equality Constraints in Asplund Spaces.) Let $X=Z=c_0$ and $T=\mathbb{N}$. Define a linear operator $A\colon X\to Z$ by $Ax:=(x_2,x_3,\ldots)$ for all $x=(x_1,x_2,\ldots)\in X$. It is easy to see that A is bounded and surjective. We form a set-valued mapping $F\colon c_0\rightrightarrows c_0\times l^\infty$ of type (7.136) by

$$F(x):=\big\{(z,p)\in Z\times l^\infty\big|\ Ax=z,\,x_1+x_n+1\le p_n,\,n\in\mathbb{N}\big\} \tag{7.150}$$

for any $x \in X$. Take $\bar{x} := \big(-\frac{1}{n}\big)_{n\in\mathbb{N}}$, $\bar{z} := A\bar{x}$, and $\widehat{x} := (-2, -\frac{1}{2}, -\frac{1}{3}, \ldots) \in X$. Observe that the bounded strong Slater condition of Definition 7.39 is satisfied at $\widehat{x}$ for (7.150) and that $\bar{x} \in F^{-1}(\bar{z}, 0)$. Defining further

$$x^k := \Big(-1, -\frac{1}{2}, \ldots, -\frac{1}{k-1}, \frac{1}{k}, -\frac{1}{k+1}, -\frac{1}{k+2}, \ldots\Big),$$
$$z^k := \Big(-\frac{1}{2}, \ldots, -\frac{1}{k-1}, \frac{2}{k}, -\frac{1}{k+1}, -\frac{1}{k+2}, \ldots\Big)$$

shows that $x^k \to \bar{x}$ and $z^k \to \bar{z}$ in c_0. Moreover, we have the equalities

$$\operatorname{dist}\big((z^k, 0); F(x^k)\big) = \max\Big\{\sup_n (x_1^k + x_n^k + 1)_+, \sup_n |\, x_{n+1}^k - z_n^k|\Big\} = \frac{1}{k}$$

with $\alpha_+ = \max\{0, \alpha\}$ as usual. It is easy to calculate the inverse mapping value $F^{-1}(z^k, 0) = \{(a, z_1^k, z_2^k, \ldots) \in c_0 |\, a \le -\frac{2}{k} - 1\}$, which gives us

$$\operatorname{dist}\big(x^k; F^{-1}(z^k, 0)\big) = \max\Big\{\Big(x_1^k + \frac{2}{k} + 1\Big)_+, \sup_n |x_{n+1}^k - z_n^k|\Big\} = \frac{2}{k}.$$

It follows from the distance expressions above that $\operatorname{reg} F(\bar{x}, (\bar{z}, 0)) \ge 2$. Thus the exact bound formula (7.149) fails if we show that

$$\|x^*\| \ge 1 \ \text{ for all } \ x^* \in D^*F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times l^\infty)^*}\big). \tag{7.151}$$

To verify (7.151), employ the explicit coderivative form from Proposition 7.43 that gives some us $z^* \in \mathbb{B}_{Z^*}$ with

$$(x^*, \langle x^*, \bar{x}\rangle) \in \operatorname{cl}^*\Big[(1 - \|z^*\|)\operatorname{co}\big\{(\delta_1 + \delta_n, -1)\big|\; n \in \mathbb{N}\big\}\Big] + \big(A^*z^*, \langle z^*, \bar{z}\rangle\big),$$

where $\delta_n \in c_0^*$ and $\langle \delta_n, x\rangle = x_n$ for all $x \in c_0$ and $n \in \mathbb{N}$. Hence there is a net $(\lambda_\nu)_{\nu\in\mathcal{N}} \subset \mathbb{R}^{(\mathbb{N})}$ such that $\sum_{n\in\mathbb{N}} \lambda_{n\nu} = 1 - \|z^*\|$ for all $\nu \in \mathcal{N}$ and that

$$(x^*, \langle x^*, \bar{x}\rangle) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \lambda_{n\nu}(\delta_1 + \delta_n, -1) + (A^*z^*, \langle z^*, \bar{z}\rangle),$$

which readily implies the limiting relationships

$$0 = \lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \lambda_{n\nu}(-\langle \delta_1 + \delta_n, \bar{x}\rangle - 1) = \lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \frac{\lambda_{n\nu}}{n}. \tag{7.152}$$

Since $c_0^* = l_1$, we write z^* in the form $(z_1^*, z_2^*, \ldots) \in l_1$ and observe that $A^*z^* = (0, z_1^*, z_2^*, \ldots) \in l_1$. Thus for any $\varepsilon > 0$, there is $k \in \mathbb{N}$ sufficiently large and such that $\sum_{n=k+1}^{\infty} |z_n^*| \le \varepsilon$, which ensures that $\|A^*z^* - \widehat{z}_k^*\| \le \varepsilon$ with $\widehat{z}_k^* := (0, z_1^*, \ldots, z_k^*, 0, 0, \ldots) \in l_1$. Define further $\widehat{x}_k^*$ by

$$\widehat{x}_k^* := w^*\text{-}\lim_{\nu \in \mathcal{N}} \sum_{n \in \mathbb{N}} \lambda_{n\nu}(\delta_1 + \delta_n) + \widehat{z}_k^*,$$

take $e^k := (1, \text{sign}(z_1^*), \ldots, \text{sign}(z_k^*), 0, \ldots) \in c_0$, and get $\|e^k\| = 1$ with

$$\begin{aligned} \|\widehat{x}_k^*\| \geq \langle \widehat{x}_k^*, e^k \rangle &= \lim_{\nu \in \mathcal{N}} \sum_{n \in \mathbb{N}} \lambda_{n\nu}(e_1^k + e_n^k) + \sum_{n=1}^{k} z_n^* e_{n+1}^k \\ &\geq \lim_{\nu \in \mathcal{N}} \sum_{n \in \mathbb{N}} \lambda_{n\nu} + \sum_{n=1}^{k} |z_n^*| - \limsup_{\nu \in \mathcal{N}} \sum_{n=1}^{k+1} \lambda_{n\nu}. \end{aligned} \tag{7.153}$$

It follows from the equations in (7.152) that

$$0 \leq \limsup_{\nu \in \mathcal{N}} \sum_{n=1}^{k+1} \lambda_{n\nu} \leq (k+1) \limsup_{\nu \in \mathcal{N}} \sum_{n \in \mathbb{N}} \frac{\lambda_{n\nu}}{n} = 0.$$

Combining this with (7.153) gives us the estimates

$$\|\widehat{x}_k^*\| \geq 1 - \|z^*\| + \sum_{n=1}^{k} |z_n^*| \geq 1 - \|z^*\| + \|z^*\| - \varepsilon = 1 - \varepsilon.$$

It is clear furthermore that $\|x^* - \widehat{x}_k^*\| = \|A^* z^* - \widehat{z}_k^*\| \leq \varepsilon$. Thus we arrive at

$$\|x^*\| \geq \|\widehat{x}_k^*\| - \|x^* - \widehat{x}_k^*\| \geq 1 - \varepsilon - \varepsilon = 1 - 2\varepsilon \quad \text{for all} \quad \varepsilon > 0,$$

yielding $\|x^*\| \geq 1$ and (7.151). This confirms the failure of (7.149).

The next example shows that the formula (7.149) for calculating the exact regularity bound fails when $\dim Z = \infty$ even for constraint systems (7.136) with a *single* convex inequality, while both spaces X and Z are Asplund.

Example 7.46 (Failure of the Exact Bound Formula for Single Inequality and Infinite-Dimensional Equality Constraints). Let $X = Z = c_0$ and $T = \{1\}$. Define the linear operator $A : X \to Z$ as in Example 7.45, and consider $F : X \rightrightarrows Z \times \mathbb{R}$ given by

$$F(x) := \big\{(z, p) \in Z \times \mathbb{R} \big| \; Ax = z, \;\; f(x) \leq p\big\} \;\; \text{for any } x \in X,$$

where $f(x) := \sup\{x_1 + x_n + 1 |\; n \in \mathbb{N}\}$ with $\text{dom}\, f = X$. Then we have

$$\text{dist}\big((z^k, 0); F(x^k)\big) = k^{-1} \;\; \text{and} \;\; \text{dist}\big(x^k; F^{-1}(z^k, 0)\big) = 2k^{-1}$$

in the notation of Example 7.45, and so $\text{reg}\, F(\bar{x}, (\bar{z}, 0)) \geq 2$. Also

$$\text{epi}\, f^* = \text{cl}^*\text{co}\,\big\{(\delta_1 + \delta_n, -1) \big|\; n \in \mathbb{N}\big\} + \{0\} \times \mathbb{R}_+, \tag{7.154}$$

which follows from the well-known formula for general supremum functions:

$$\text{epi}\, f^* = \text{cl}^*\text{co} \bigcup_{t \in T} \Big(\text{epi}\, f_t^*\Big). \tag{7.155}$$

Picking now any $x^* \in D^*F(\bar{x}, (\bar{z}, 0))(S_{(Z\times\mathbb{R})^*})$ and using Theorem 7.42 together with representation (7.155), we arrive at

$$(x^*, \langle x^*, \bar{x}\rangle) \in \text{cl}^*\Big[(1 - \|z^*\|)\,\text{co}\,\big\{(\delta_1 + \delta_n, -1)\big|\, n \in \mathbb{N}\big\}\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle).$$

As in Example 7.45, this gives us $\|x^*\| \geq 1$, and thus (7.149) fails.

The following result provides efficient conditions, which ensure the validity of the major regularity formula (7.149) when $\dim Z = \infty$. The given proof is different from that of (7.147) in Theorem 7.44 with $\dim Z < \infty$. In particular, it doesn't rely on condition (7.129) that may not hold. Indeed, even in the simplest setting of $T = \emptyset$, the left-hand side of (7.129) is $\text{cl}^* S_{Z^*}$, which is obviously not a subset of S_{Z^*} when $\dim Z = \infty$.

Theorem 7.47 (Exact Bound Formula for Finite Inequality and Infinite Equality Constraints). *In the case of arbitrary Banach spaces X and Z in* (7.136), *assume that the index set T is finite, that*

$$f_t(x) = \langle a_t^*, x\rangle - b_t \;\; \textit{for all} \;\; x \in X,\;\; t \in T \;\; \textit{with} \;\; (a_t^*, b_t) \in X^* \times \mathbb{R},$$

and that, given $(\bar{x}, (\bar{z}, 0)) \in \text{gph}\, F$, the constraint mapping F satisfies the standard Slater condition at $(\bar{z}, 0)$ with $\bar{x} \in C$. Then formula (7.149) *holds.*

Proof. Letting $T := \{1, \ldots, k\}$, observe that $\text{dom}\, f = X$ and so $C_0 = C$ in the notation of Theorem 7.42. Since we obviously have

$$\text{epi}\, f_n^* = (a_n^*, b_n) + \{0\} \times \mathbb{R}_+ \;\;\text{and}\;\; \{0\} \times \mathbb{R}_+ + \text{epi}\,\delta^*(\cdot; C) \subset \text{epi}\,\delta^*(\cdot; C)$$

for any $z^* \in \mathbb{B}_{Z^*}$ and $n \in \{1, \ldots, k\}$, it follows that

$$\big(1 - \|z^*\|\big)\text{co}\big\{\text{epi}\, f_t^*\big| \in T\big\} + \text{epi}\,\delta^*(\cdot; C_0) = \big(1 - \|z^*\|\big)\text{co}\big\{(a_n^*, b_n)\big| 1 \leq n \leq k\big\} + \text{epi}\,\delta^*(\cdot; C).$$

The latter set is clearly weak* closed in $X^* \times \mathbb{R}$, and hence the set M in Theorem 7.42 is represented by

$$M = \bigcup_{z^* \in \mathbb{B}_{Z^*}} \Big\{(1 - \|z^*\|)\text{co}\big\{(a_n^*, b_n)\big| 1 \leq n \leq k\big\} + \text{epi}\,\delta^*(\cdot; C) + (A^*z^*, \langle z^*, \bar{z}\rangle)\Big\}.$$

Invoking now the result in the first part of Theorem 7.44, we find sequences of $x_m^* \in X^*$, $\lambda^m \in \mathbb{R}_+^k$, $(v_m^*, s_m) \in \text{epi}\,\delta^*(\cdot; C)$, and $z_m^* \in \mathbb{B}_{Z^*}$ for all $m \in \mathbb{N}$ such that $\sum_{n=1}^k \lambda_n^m = 1 - \|z_m^*\|$ and

$$\Big(x_m^*, \langle x_m^*, \bar{x}\rangle + m^{-1}\Big) = \sum_{n=1}^{k} \lambda_n^m (a_n^*, b_n) + (v_m^*, s_m) \\ +(A^* z_m^*, \langle z_m^*, \bar{z}\rangle) \tag{7.156}$$

with the upper estimate of the regularity bound

$$\text{reg}\, F\big(\bar{x}, (\bar{z}, 0)\big) \le \|x_m^*\|^{-1} + o(1) = \Big\| \sum_{n=1}^{k} \lambda_n^m a_n^* + v_m^* + A^* z_m^* \Big\|^{-1} + o(1).$$

It follows from considering the second components in (7.156) that

$$\begin{aligned} \frac{1}{m} = \langle x_m^*, \bar{x}\rangle + \frac{1}{m} - \langle x_m^*, \bar{x}\rangle &= \sum_{n=1}^{k} \lambda_n^m b_n + s_m - \sum_{n=1}^{k} \lambda_n^m \langle a_n^*, \bar{x}\rangle - \langle v_m^*, \bar{x}\rangle \\ &\ge \sum_{n=1}^{k} \lambda_n^m (b_n - \langle a_n^*, \bar{x}\rangle) + s_m - \langle v_m^*, \bar{x}\rangle \ge \sum_{n=1}^{k} \lambda_n^m (b_n - \langle a_n^*, \bar{x}\rangle) \ge 0. \end{aligned}$$

Since $\|\lambda^m\| \le 1$, we suppose that $\lambda^m \to \lambda \in \mathbb{R}_+^k$ as $m \to \infty$ and thus deduce from the above that $\sum_{n=1}^{k} \lambda_n (b_n - \langle a_n^*, \bar{x}\rangle) = 0$ by passing to the limit as $m \to \infty$. Defining further the sequences of

$$\varepsilon_m := \sum_{n=1}^{k} |\lambda_n^m - \lambda_n|, \;\; \eta_m := \sum_{n=1}^{k} \lambda_n + \|z_m^*\|, \;\; \widehat{x}_m^* := \sum_{n=1}^{k} \lambda_n a_n^* + A^* z_m^*,$$

note that $\varepsilon_m = o(1)$ and $\eta_m = 1 - o(1)$. Then Proposition 7.43 tells us that

$$\eta_m^{-1} \widehat{x}_m^* \in D^* F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times\mathbb{R}^k)^*}\big).$$

Moreover, the same arguments as in the proof of the second part of Proposition 7.43 show that $\|w_m^*\| \to 0$. It follows therefore that

$$\|x_m^* - \widehat{x}_m^*\| = \Big\| \sum_{n=1}^{k} (\lambda_n^m - \lambda_n) a_n^* + w_m^* \Big\| \le \varepsilon_m \sup_{1\le n\le k} \|a_n^*\| + \|w_m^*\| = o(1),$$

which implies together with the above estimates of $\text{reg}\, F\big(\bar{x}, (\bar{z}, 0)\big)$ that

$$\begin{aligned} \text{reg}\, F\big(\bar{x}, (\bar{z}, 0)\big) &\le \big(\|\widehat{x}_m^*\| + o(1)\big)^{-1} + o(1) \le \big(\eta_m \|\eta_m^{-1} \widehat{x}_m^*\| + o(1)\big)^{-1} + o(1) \\ &\le \Big[(1 - o(1)) \inf\Big\{\|x^*\| \,\Big|\, (x^*, \langle x^*, \bar{x}\rangle) \in L\Big\} + o(1)\Big]^{-1} + o(1). \end{aligned}$$

Letting $m \to \infty$ therein, we arrive at

$$\text{reg}\, F\big(\bar{x}, (\bar{z}, 0)\big) \le \sup\Big\{\|x^*\|^{-1} \,\Big|\, (x^*, \langle x^*, \bar{x}\rangle) \in L\Big\},$$

which yields (7.149) and thus completes the proof of the theorem. △

7.5 Value Functions in DC Semi-infinite Optimization

In this section, we continue investigating SIPs in general Banach (and partly in Asplund) spaces while considering now the minimization of *DC objectives* subject to infinite convex inequality constraints with arbitrary index sets. As mentioned earlier, the abbreviation "DC" stands for the *difference of convex* functions, which have been recognized as a convenient form for representing various remarkable classes of problems important in optimization and its applications. Our main attention is paid here to the study of subdifferential properties of (nonconvex) *marginal/value functions* in parametric versions of such SIPs. Based on these developments, we present applications to sensitivity analysis and necessary optimality conditions in DC SIPs considered in both nonparametric and parametric settings as well as to bilevel semi-infinite programs with fully convex data in Banach and Asplund spaces.

7.5.1 Optimality Conditions for DC Semi-infinite Programs

Consider first *nonparametric* SIPs with DC objectives and infinite convex constraints and obtain necessary optimality conditions for them (necessary and sufficient for fully convex problems) under weakest qualification conditions. These results of their own interest are instrumental to derive in what follows subdifferential formulas for value functions in parametric versions of such SIPs with subsequent applications to optimality conditions and Lipschitzian stability under perturbations. In this subsection, we study the problem

$$\begin{cases} \text{minimize} \;\; \vartheta(x) - \theta(x) \;\; \text{subject to} \\ \vartheta_t(x) \le 0, \quad t \in T, \quad \text{and} \quad x \in \Theta, \end{cases} \tag{7.157}$$

where T is an arbitrary index set, where $\Theta \subset X$ is a closed and convex subset of a Banach space X, and where the functions $\vartheta, \theta, \vartheta_t \colon X \to \overline{\mathbb{R}}$ are l.s.c. and convex. Being oriented toward minimization, we impose by convention that $\infty - \infty := \infty$ along with the standard operations involving ∞ and $-\infty$. The set of feasible solutions to (7.157) is denoted by

$$\Xi := \Theta \cap \big\{x \in X \big|\; \vartheta_t(x) \le 0 \;\text{ for all }\; t \in T\big\}. \tag{7.158}$$

Using the infinite product notation $\mathbb{R}^T$, $\mathbb{R}^{(T)}$, and $\mathbb{R}^{(T)}_+$ from Subsection 7.1.1, define $\operatorname{supp}\lambda := \{t \in T \,|\; \lambda_t \neq 0\}$ for any $\lambda \in \mathbb{R}^{(T)}$, and observe that

$$\lambda u := \sum_{t \in T} \lambda_t u_t = \sum_{t \in \operatorname{supp}\lambda} \lambda_t u_t \;\text{ whenever }\; u \in \mathbb{R}^T.$$

Recall next definition (7.30) of the Fenchel conjugate, and introduce the following *dual-space* qualification condition, which plays a crucial role in deriving necessary optimality conditions of the KKT type for (7.157).

Definition 7.48 (Closedness Qualification Condition). *We say that the triple* $(\vartheta, \vartheta_t, \Theta)$ *in problem* (7.157) *satisfies the* CLOSEDNESS QUALIFICATION CONDITION (CQC) *if the set*

$$\operatorname{epi}\vartheta^* + \operatorname{cone}\Big\{\bigcup_{t\in T}\operatorname{epi}\vartheta_t^*\Big\} + \operatorname{epi}\delta^*(\cdot;\Theta)$$

*is weak** *closed in the product space* $X^*\times\mathbb{R}$.

Note that the introduced CQC is not a "constraint qualification" since it involves not only constraint but also cost functions, namely, the *plus* part ϑ of the cost in (7.157). The closest constraint qualification to CQC is the following one, where the cost term $\operatorname{epi}\vartheta^*$ in Definition 7.48 is omitted: the set

$$\operatorname{cone}\Big\{\bigcup_{t\in T}\operatorname{epi}\vartheta_t^*\Big\} + \operatorname{epi}\delta^*(\cdot;\Theta)\ \text{ is weak}^*\text{ closed} \tag{7.159}$$

in $X^*\times\mathbb{R}$. This condition known as the convex *Farkas-Minkowski constraint qualification* (convex FMCQ) reduces to the Farkas-Minkowski property (7.49) for linear infinite systems of type (7.48). The reader can check that FMCQ (7.159) implies CQC in the following two cases: either ϑ is continuous at some feasible point $x\in\Xi$ in (7.158), or the convex conic hull $\operatorname{cone}(\operatorname{dom}\vartheta-\Xi)$ is a closed subspace of X. It has been well recognized in semi-infinite programming that *dual* qualification conditions of the CQC and Farkas-Minkowski type for infinite convex systems strictly improve *primal* ones of the Slater type; see Exercise 7.98 and the corresponding commentaries in Section 7.7.

To proceed, we recall some needed results of convex analysis summarized in the following two lemmas. The first one contains relationship between epigraphical duality and subdifferential calculus.

Lemma 7.49 (Epigraphical and Subdifferential Sum Rules). *Let the functions* $\varphi_1,\varphi_2: X\to\overline{\mathbb{R}}$ *be l.s.c. and convex, and let* $\operatorname{dom}\varphi_1\cap\operatorname{dom}\varphi_2\neq\emptyset$. *Then the following conditions are equivalent:*

(i) *The set* $\operatorname{epi}\varphi_1^*+\operatorname{epi}\varphi_2^*$ *is weak** *closed in* $X^*\times\mathbb{R}$.

(ii) *The conjugate epigraphical rule holds*

$$\operatorname{epi}(\varphi_1+\varphi_2)^* = \operatorname{epi}\varphi_1^*+\operatorname{epi}\varphi_2^*.$$

Furthermore, we have the subdifferential sum rule

$$\partial(\varphi_1+\varphi_2)(\bar{x}) = \partial\varphi_1(\bar{x})+\partial\varphi_2(\bar{x})$$

provided that the aforementioned equivalent conditions are satisfied.

The next result presents an appropriate extension of the Farkas lemma to the case of epigraphical convex systems.

Lemma 7.50 (Generalized Farkas Lemma for Epigraphical Systems). *Given* $\alpha \in \mathbb{R}$*, the following conditions are equivalent:*

(i) $\vartheta(x) \ge \alpha$ *for all* $x \in \Xi$*;*

(ii) $\big(0, -\alpha\big) \in \operatorname{cl}^*\Big(\operatorname{epi}\vartheta^* + \operatorname{cone}\Big[\bigcup_{t \in T} \operatorname{epi}\vartheta_t^*\Big] + \operatorname{epi}\delta^*(\cdot;\Theta)\Big)$.

Now we are ready to establish necessary optimality conditions for the DC program under consideration in (7.157). Given $\bar{x} \in \Xi \cap \operatorname{dom}\theta$, define the set of *active constraint multipliers* by

$$A(\bar{x}) := \big\{\lambda \in \mathbb{R}_+^{(T)} \,\big|\, \lambda_t \vartheta_t(\bar{x}) = 0 \;\text{ for all }\; t \in \operatorname{supp}\lambda\big\}. \tag{7.160}$$

Theorem 7.51 (Necessary Optimality Conditions for DC Semi-infinite Programs). *Let* $\bar{x} \in \Xi \cap \operatorname{dom}\vartheta$ *be a local minimizer for problem* (7.157) *satisfying the CQC requirement. Then we have the inclusion*

$$\partial\theta(\bar{x}) \subset \partial\vartheta(\bar{x}) + \bigcup_{\lambda \in A(\bar{x})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta). \tag{7.161}$$

Proof. There are two possible cases regarding $\bar{x} \in \Xi \cap \operatorname{dom}\vartheta$: either $\bar{x} \notin \operatorname{dom}\theta$ or $\bar{x} \in \operatorname{dom}\theta$. In the first case, we have $\partial\theta(\bar{x}) = \emptyset$, and hence (7.161) holds automatically. Considering the remaining case of $\bar{x} \in \operatorname{dom}\theta$, find by the subdifferential definition of convex analysis such $x^* \in X^*$ that

$$\theta(x) - \theta(\bar{x}) \ge \langle x^*, x - \bar{x}\rangle \;\text{ for all }\; x \in X.$$

This implies that the reference local minimizer $\bar{x}$ for (7.157) is also a local minimizer for the following *convex* SIP:

$$\begin{cases} \text{minimize } \;\widetilde{\vartheta}(x) := \vartheta(x) - \langle x^*, x - \bar{x}\rangle - \theta(\bar{x}) \\ \text{subject to } \;\vartheta_t(x) \le 0, \quad t \in T, \quad \text{and } \; x \in \Theta. \end{cases} \tag{7.162}$$

Since (7.162) is convex, its local minimizer $\bar{x}$ is its global solution, i.e.,

$$\widetilde{\vartheta}(\bar{x}) \le \widetilde{\vartheta}(x) \;\text{ for all }\; x \in \Xi.$$

Then Lemma 7.50 tells us that the latter is equivalent to the inclusion

$$\big(0, -\widetilde{\vartheta}(\bar{x})\big) \in \operatorname{cl}^*\Big(\operatorname{epi}\widetilde{\vartheta}^* + \operatorname{cone}\Big[\bigcup_{t \in T} \operatorname{epi}\vartheta_t^*\Big] + \operatorname{epi}\delta^*(\cdot;\Theta)\Big).$$

Observing from the structure of $\widetilde{\vartheta}$ in (7.162) that $\operatorname{epi}\widetilde{\vartheta}^* = (-x^*, \theta(\bar{x}) - \langle x^*, \bar{x}\rangle) + \operatorname{epi}\vartheta^*$, we get therefore the relationship

$$\begin{aligned}&\big(0,-\widetilde{\vartheta}(\bar{x})\big)\in\big(-x^*,\theta(\bar{x})-\langle x^*,\bar{x}\rangle\big)\\&+\text{cl}^*\Big(\text{epi}\,\vartheta^*+\text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\vartheta_t^*\Big]+\text{epi}\,\delta^*(\cdot;\Theta)\Big).\end{aligned}\tag{7.163}$$

Furthermore, the assumed CQC ensures that (7.163) is equivalent to

$$\begin{aligned}&\big(x^*,-\widetilde{\vartheta}(\bar{x})-\theta(\bar{x})+\langle x^*,\bar{x}\rangle\big)\\&\in\Big(\text{epi}\,\vartheta^*+\text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\vartheta_t^*\Big]+\text{epi}\,\delta^*(\cdot;\Theta)\Big).\end{aligned}\tag{7.164}$$

Now applying the useful representation

$$\text{epi}\,\varphi^*=\bigcup_{\varepsilon\ge0}\Big\{\big(x^*,\langle x^*,x\rangle+\varepsilon-\varphi(x)\big)\Big|\;x^*\in\partial_\varepsilon\varphi(x)\Big\},\tag{7.165}$$

which is valid for all $x\in\text{dom}\,\varphi$, to the conjugate functions ϑ^*, ϑ_t^*, and $\delta^*(\cdot;\Theta)$ with taking into account the structure of the positive cone $\mathbb{R}_+^{(T)}$ in (7.3) and noting that $-\widetilde{\vartheta}(\bar{x})-\theta(\bar{x})+\langle x^*,\bar{x}\rangle=\langle x^*,\bar{x}\rangle-\vartheta(\bar{x})$, we find

$$\varepsilon,\varepsilon_t,\gamma\ge0,\;\; u^*\in\partial_\varepsilon\vartheta(\bar{x}),\;\;\lambda\in\mathbb{R}_+^{(T)},\;\; u_t^*\in\partial_{\varepsilon_t}\vartheta_t(\bar{x}),\;\;\text{and}\;\; v^*\in\partial\delta_\gamma(\bar{x};\Theta)$$

satisfying the following two equalities:

$$\begin{cases}x^*=u^*+\displaystyle\sum_{t\in T}\lambda_t u_t^*+v^*,\\ \langle x^*,\bar{x}\rangle-\vartheta(\bar{x})=\langle u^*,\bar{x}\rangle+\varepsilon-\vartheta(\bar{x})+\displaystyle\sum_{t\in T}\lambda_t\Big[\langle u_t^*,\bar{x}\rangle+\varepsilon_t-\langle\vartheta_t^*,\bar{x}\rangle\Big]\\ +\langle v^*,\bar{x}\rangle+\gamma-\delta(\bar{x};\Theta).\end{cases}$$

Since $\bar{x}\in\Theta$, the first equality above allows us reducing the second one to

$$\varepsilon+\sum_{t\in T}\lambda_t\varepsilon_t-\sum_{t\in T}\lambda_t\vartheta_t(\bar{x})+\gamma=0.\tag{7.166}$$

The feasibility of $\bar{x}$ for problem (7.157) and the choice of $(\varepsilon,\lambda_t,\gamma)$ yield

$$\varepsilon\ge0,\;\gamma\ge0,\;\lambda_t\ge0,\;\;\text{and}\;\;\lambda_t\vartheta_t(\bar{x})\le0\;\;\text{for all}\;\;t\in T,$$

and therefore we get from (7.166) that in fact $\varepsilon=0$, $\gamma=0$, $\lambda_t\vartheta_t(\bar{x})=0$, and $\lambda_t\varepsilon_t=0$ for all $t\in T$. Furthermore, the latter implies that $\varepsilon_t=0$ for all $t\in\text{supp}\,\lambda$. Hence we obtain the inclusions

$$u^*\in\partial\vartheta(\bar{x}),\;\;u_t^*\in\partial\vartheta_t(\bar{x}),\;\;\text{and}\;\;v^*\in\partial\delta(\bar{x};\Theta)=N(\bar{x};\Theta),$$

which allow us to conclude from the above that

$$x^*\in\partial\vartheta(\bar{x})+\sum_{t\in\text{supp}\,\lambda}\lambda_t\partial\vartheta_t(\bar{x})+N(\bar{x};\Theta)\;\;\text{with}\;\;\lambda_t\vartheta_t(\bar{x})=0\;\;\text{for}\;\;t\in\text{supp}\,\lambda.$$

This justifies (7.161) and thus completes the proof of the theorem. △

Let us present two useful consequences of Theorem 7.51 concerning subdifferential/normal cone *calculus* for infinite convex systems.

Corollary 7.52 (Subdifferential Sum Rule Involving Convex Infinite Constraints). *Let $\bar{x} \in \Xi$ with $\theta(\bar{x}) = 0$ and $\vartheta(\bar{x}) < \infty$, and let $(\vartheta, \vartheta_t, \Theta)$ satisfy all the assumptions of Theorem 7.51. Then we have the equality*

$$\partial\big(\vartheta + \delta(\cdot;\Xi)\big)(\bar{x}) = \partial\vartheta(\bar{x}) + \bigcup_{\lambda\in A(\bar{x})}\Big[\sum_{t\in\operatorname{supp}\lambda}\partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta).$$

Proof. The inclusion "$\supset$" in the claimed sum rule can be derived directly from the definitions. To verify the opposite inclusion therein, pick an arbitrary subgradient $x^* \in \partial(\vartheta + \delta(\cdot;\Xi))(\bar{x})$ with $\bar{x} \in \Xi \cap \operatorname{dom}\vartheta$, and get

$$\vartheta(x) - \vartheta(\bar{x}) \ge \langle x^*, x - \bar{x}\rangle \quad \text{whenever} \quad x \in \Xi,$$

which means by the construction of Ξ in (7.158) that $\bar{x}$ is a (global) minimizer for the following *DC program* with infinite constraints:

$$\begin{cases} \text{minimize} \;\; \vartheta(x) - \widetilde{\theta}(x) \;\; \text{with} \;\; \widetilde{\theta}(x) := \langle x^*, x - \bar{x}\rangle + \vartheta(\bar{x}) \\ \text{subject to} \;\; \vartheta_t(x) \le 0 \;\; \text{for all} \;\; t \in T, \;\; \text{and} \;\; x \in \Theta. \end{cases} \tag{7.167}$$

Applying Theorem 7.51 to problem (7.167) and taking into account the structure of the linear function $\widetilde{\theta}$ therein, we get from (7.161) that

$$\partial\widetilde{\theta}(\bar{x}) = \big\{x^*\big\} \subset \partial\vartheta(\bar{x}) + \bigcup_{\lambda\in A(\bar{x})}\Big[\sum_{t\in\operatorname{supp}\lambda}\partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta),$$

which justifies the claimed inclusion and thus completes the proof. △

The next corollary provides a calculation of the normal cone to the feasible constraint set Ξ in terms of its initial data of (7.12) and the set of active constraint multipliers (7.160).

Corollary 7.53 (Normal Cone Calculation for Convex Infinite Constraints). *Assume that ϑ_t and Θ satisfy the assumptions of Theorem 7.51 with CQC specified as FMCQ (7.159). Then for any $\bar{x} \in \Xi$, we have*

$$N(\bar{x};\Xi) = \bigcup_{\lambda\in A(\bar{x})}\Big[\sum_{t\in\operatorname{supp}\lambda}\partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta).$$

Proof. Follows from Corollary 7.52 by letting $\vartheta(x) \equiv 0$ therein. △

The final result of this subsection concerns the *convex SIP*, which is a specification of (7.157) with $\theta \equiv 0$. We show that in this case the necessary conditions of Theorem 7.51 are also sufficient for (global) optimality.

Theorem 7.54 (Necessary and Sufficient Optimality Conditions for Convex SIPs). *Let $\bar{x} \in \Xi$ be a feasible solution to problem (7.157) with $\theta \equiv 0$ and $\vartheta(\bar{x}) < \infty$, and let the assumptions of Theorem 7.51 be satisfied. Then $\bar{x}$ is optimal to this problem if and only if there is $\lambda \in \mathbb{R}_+^{(T)}$ such that the following generalized KKT condition holds:*

$$0 \in \partial\vartheta(\bar{x}) + \bigcup_{\lambda \in A(\bar{x})} \Big[\sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x}) \Big] + N(\bar{x}; \Theta). \tag{7.168}$$

Proof. The necessary of (7.168) for optimality in this problem follows immediately from Theorem 7.51 with $\theta(x) \equiv 0$. To justify the sufficiency part, suppose that (7.168) holds with some $\lambda \in A(\bar{x})$; the latter implies, in particular, that $\partial\vartheta_t(\bar{x}) \neq \emptyset$ whenever $t \in \operatorname{supp}\lambda$. Then we find $x^* \in X^*$ satisfying the inclusions $-x^* \in N(\bar{x}; \Theta)$ and

$$x^* \in \partial\vartheta(\bar{x}) + \sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x}) \subset \partial\Big(\vartheta + \sum_{t \in T} \lambda_t \vartheta_t\Big)(\bar{x}).$$

This tells by the construction of convex subgradients that

$$\vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) \geq \vartheta(\bar{x}) + \sum_{t \in T} \lambda_t \vartheta_t(\bar{x}) + \langle x^*, x - \bar{x} \rangle \geq 0 \tag{7.169}$$

for all $x \in X$. Since $\lambda_t \vartheta_t(\bar{x}) = 0$ for all $t \in T$ by $\lambda \in A(\bar{x})$ in (7.160) and since $-x^* \in N(\bar{x}; \Theta)$, we get from (7.169) and the normal cone structure that

$$\vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) - \vartheta(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle \geq 0 \;\text{ for all }\; x \in \Theta,$$

which yields by (7.158) and (7.160) the inequality

$$\vartheta(x) \geq \vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) \geq \vartheta(\bar{x}) \;\text{ whenever }\; x \in \Xi$$

and thus verifies the claimed global optimality of $\bar{x}$. △

7.5.2 Regular Subgradients of Value Functions for DC SIPs

Let us now consider the *parametric* version of the DC semi-infinite program (7.157) formalized, with a bit different notation, as

$$\text{minimize}_y \; \varphi(x, y) - \psi(x, y) \;\text{ subject to } y \in F(x) \cap G(x), \tag{7.170}$$

where the moving (parameterized by x) constraint sets are given by

$$F(x) := \big\{ y \in Y \big|\, (x, y) \in \Omega \big\}, \tag{7.171}$$

$$G(x) := \big\{y \in Y \big|\ \varphi_t(x, y) \le 0,\ t \in T\big\}. \tag{7.172}$$

In what follows, we assume, unless otherwise stated, that the spaces X and Y are Banach, that T is an arbitrary index set, that the functions $\varphi, \psi, \varphi_t \colon X \times Y \to \overline{\mathbb{R}}$ are l.s.c. and convex, and that the set Ω is closed and convex.

The main object of our study in the rest of this section is the (optimal) *value function* in (7.170) defined by

$$\mu(x) := \inf\big\{\varphi(x, y) - \psi(x, y) \big|\ y \in F(x) \cap G(x)\big\}, \tag{7.173}$$

which is nonconvex unless $\psi \equiv 0$. The value function (7.173) belongs to the general class of marginal functions whose subdifferential properties have been studied in Section 4.1; see also the corresponding commentaries in Section 4.6. However, the results obtained therein are expressed in terms of the coderivative of the constraint mapping in (7.170), while the major goal of our study here is to derive subdifferential results for (7.173) expressed entirely via the *initial data* of (7.170) with taking into account the infinite inequality constraint nature of (7.172) and the DC structure of the cost function in (7.170).

In this subsection, we concentrate on evaluating the *regular subdifferential* of (7.173), which is defined in Banach spaces exactly as in finite dimensions (1.33). The results obtained are of their own interest, while they also can be considered, together with similar calculations for the ε-enlargements (1.34), as approximating tools for evaluating the limiting (both basic and singular) subdifferentials of the value function, which are the most valuable applications to DC semi-infinite optimization and Lipschitzian stability of (7.170). The necessary optimality conditions for the nonparametric DC version (7.157) obtained in Subsection 7.5.1 play a significant role in our subdifferential device. For brevity, we confine ourselves here to considering only regular subgradients of (7.173) while leaving the ε-case as an exercise for the reader.

In the next theorem and further results below, we use the notation

$$M(x) := \big\{y \in F(x) \cap G(x) \big|\ \mu(x) = \varphi(x, y) - \psi(x, y)\big\}, \tag{7.174}$$

$$\Gamma := \Omega \cap \big\{(x, y) \in X \times Y \big|\ \varphi_t(x, y) \le 0\ \text{ for all }\ t \in T\big\}, \tag{7.175}$$

$$\begin{aligned}\Lambda(\bar{x}, \bar{y}, y^*) := \Big\{\lambda \in \mathbb{R}^{(T)}_+ \Big|\ & y^* \in \partial_y\varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}) \\ & + N_y\big((\bar{x}, \bar{y}); \Omega\big),\ \lambda_t\varphi_t(\bar{x}, \bar{y}) = 0\ \text{ as }\ t \in \operatorname{supp}\lambda\Big\},\end{aligned} \tag{7.176}$$

where $N_y((\bar{x}, \bar{y}); \Omega)$ stands for the subdifferential of the indicator function $y \mapsto \delta((\bar{x}, y); \Omega)$ at $\bar{y}$; the notation $N_x((\bar{x}, \bar{y}); \Omega)$ below is similar.

Theorem 7.55 (Upper Estimate for Regular Subgradients of Value Functions in DC SIPs). *Let* dom $M \neq \emptyset$*, and let CQC from Definition 7.48 be satisfied for the* *triple* $(\varphi, \varphi_t, \Omega)$ *in* (7.170). *Then, given any* $(\bar{x}, \bar{y}) \in \operatorname{gph} M \cap \operatorname{dom} \partial\psi$ *and* $\gamma > 0$, *we have the inclusion*

$$\widehat{\partial}\mu(\bar{x}) \subset \bigcap_{(x^*,y^*)\in\partial\psi(\bar{x},\bar{y})} \Big\{\partial_x\varphi(\bar{x},\bar{y}) - x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial_x\varphi_t(\bar{x},\bar{y})\Big]\Big\}$$
$$+N_x\big((\bar{x},\bar{y});\Omega\big) + \gamma\mathbb{B}^*.$$

Proof. Fix $(\bar{x}, \bar{y}) \in \operatorname{gph} M \cap \operatorname{dom} \partial\psi$, $u^* \in \widehat{\partial}\mu(\bar{x})$, and $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$. Picking an arbitrary positive number γ and using the definition of regular subgradients, find $\eta > 0$ such that

$$\mu(x) - \mu(\bar{x}) - \langle u^*, x - \bar{x}\rangle + \gamma\|x - \bar{x}\| \ge 0 \ \text{ if } \ x \in \bar{x} + \eta\mathbb{B}. \tag{7.177}$$

Since $\mu(\bar{x}) = \varphi(\bar{x}, \bar{y}) - \psi(\bar{x}, \bar{y})$ by $\bar{y} \in M(\bar{x})$ and since $\mu(x) \le \varphi(x, y) - \psi(x, y)$ for all $(x, y) \in \Gamma$, we get from (7.177) and $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$ that

$$\begin{aligned} 0 &\le \varphi(x,y) - \varphi(\bar{x},\bar{y}) - \psi(x,y) + \psi(\bar{x},\bar{y}) - \langle u^*, x-\bar{x}\rangle + \gamma\|x-\bar{x}\| \\ &\le \varphi(x,y) - \varphi(\bar{x},\bar{y}) - \langle u^* + x^*, x-\bar{x}\rangle - \langle y^*, y-\bar{y}\rangle + \gamma\|x-\bar{x}\| \end{aligned}$$

for $(x, y) \in \Omega \cap [(\bar{x} + \eta\mathbb{B}) \times Y]$ with $\varphi_t(x, y) \le 0$, $t \in T$. Consider the function

$$\vartheta(x,y) := \varphi(x,y) - \varphi(\bar{x},\bar{y}) - \langle u^* + x^*, x-\bar{x}\rangle - \langle y^*, y-\bar{y}\rangle + \gamma\|x-\bar{x}\|,$$

which is l.s.c. and convex on $X \times Y$. It follows from (7.177) and the construction of ϑ that $(\bar{x}, \bar{y})$ is a solution to the following *nonparametric convex* SIP:

$$\begin{cases} \text{minimize } \ \vartheta(x,y) \ \text{ with respect to both } (x,y) \text{ subject to} \\ \varphi_t(x,y) \le 0 \ \text{ as } \ t \in T, \quad (x,y) \in \Omega \cap \big[(\bar{x} + \eta\mathbb{B}) \times Y\big]. \end{cases} \tag{7.178}$$

The technical Lemma 7.56, which is presented for convenience after the proof of the theorem, tells us the CQC requirement on $(\varphi, \varphi_t, \Omega)$ imposed in this theorem yields the validity of CQC for (7.178). Applying now the optimality conditions from Theorem 7.54 to (7.178) gives us $\lambda \in \mathbb{R}^{(T)}_+$ such that

$$\begin{aligned} &0 \in \partial\vartheta(\bar{x},\bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial\varphi_t(\bar{x},\bar{y}) + N\big((\bar{x},\bar{y}); \Omega \cap [(\bar{x}+\eta\mathbb{B}) \times Y]\big) \\ &\text{with } \ \lambda_t\varphi_t(\bar{x},\bar{y}) = 0 \ \text{ for all } \ t \in \operatorname{supp}\lambda. \end{aligned}$$

Since $(\bar{x}, \bar{y}) \in \imath[(\bar{x} + \eta\mathbb{B}) \times Y]$, it follows from the classical subdifferential rule of convex analysis and the construction of ϑ that

$$\partial\vartheta(\bar{x},\bar{y}) = \partial\varphi(\bar{x},\bar{y}) + (-u^* - x^*, -y^*) + (\gamma\mathbb{B}^*) \times \{0\}.$$

Thus we get by (i)$\Rightarrow$(iii) in Lemma 7.49 applied to the indicator functions $\delta((\bar{x}, \bar{y}); \Omega)$ and $\delta((\bar{x}, \bar{y}); (\bar{x} + \eta\mathbb{B}) \times Y)$ that

$$N\big((\bar{x},\bar{y});\Omega\cap[(\bar{x}+\eta\mathbb{B})\times Y]\big)=N\big((\bar{x},\bar{y});\Omega\big).$$

Substituting this into the above optimality condition for (7.178) with taking into account the well-known relationships

$$\partial\varphi(\bar{x},\bar{y})\subset\partial_x\varphi(\bar{x},\bar{y})\times\partial_y\varphi(\bar{x},\bar{y})\ \text{ and }\ \partial\varphi_t(\bar{x},\bar{y})\subset\partial_x\varphi_t(\bar{x},\bar{y})\times\partial_y\varphi_t(\bar{x},\bar{y})$$

ensures the fulfillment of the two inclusions

$$u^*\in\partial_x\varphi(\bar{x},\bar{y})-x^*+\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x},\bar{y})+N_x\big((\bar{x},\bar{y});\Omega\big)+\gamma\mathbb{B}^*,$$
$$y^*\in\partial_y\varphi(\bar{x},\bar{y})+\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_y\varphi_t(\bar{x},\bar{y})+N_y\big((\bar{x},\bar{y});\Omega\big)$$

with $\lambda_t\varphi_t(\bar{x},\bar{y})=0$, $t\in\operatorname{supp}\lambda$. This verifies by (7.176) the claimed estimate of $\widehat{\partial}\mu(\bar{x})$ by the construction in (7.176) and Lemma 7.56 justified below. △

Lemma 7.56 (Relationships Between Parametric and Nonparametric CQC). *The validity of CQC for* $(\varphi,\varphi_t,\Omega)$ *imposed in Theorem 7.55 yields the fulfillment of this condition for the nonparametric problem* (7.178).

Proof. In the notation of Theorem 7.55, take $(\bar{x},\bar{y})\in\operatorname{gph}M\cap\operatorname{dom}\partial\psi$ with $(\bar{x},\bar{y})\in\operatorname{dom}\varphi\cap\Gamma$, and define the convex and continuous function

$$\xi(x,y):=-\varphi(\bar{x},\bar{y})-\langle u^*+x^*,x-\bar{x}\rangle-\langle y^*,y-\bar{y}\rangle+\gamma\|x-\bar{x}\|$$

on $X\times Y$ that gives us the representation $\vartheta=\varphi+\xi$. Substituting the latter into the assumed CQC for $(\varphi,\varphi_t,\Omega)$ and using the epigraphical rule from Lemma 7.49 with taking into account that the continuity of $\delta(\cdot;(\bar{x}+\eta\mathbb{B}^*)\times Y)$ at the interior point $(\bar{x},\bar{y})$, we conclude that the corresponding set in the CQC property for (7.178) reduces to

$$\operatorname{epi}\varphi^*+\operatorname{cone}\Big[\bigcup_{t\in T}\operatorname{epi}\varphi_t^*\Big]+\operatorname{epi}\delta^*(\cdot;\Omega)+\operatorname{epi}\big[\xi+\delta\big(\cdot;(\bar{x}+\eta\mathbb{B})\times Y\big)\big]^*.$$

On the other hand, by Lemma 7.49, the CQC requirement for $(\varphi,\varphi_t,\Omega)$ yields

$$\operatorname{epi}\big(\varphi+\delta(\cdot;\Gamma)\big)^*=\operatorname{epi}\varphi^*+\operatorname{cone}\Big[\bigcup_{t\in T}\operatorname{epi}\varphi_t^*\Big]+\operatorname{epi}\delta^*(\cdot;\Omega).$$

Substituting this equality into the aforementioned CQC set for $(\varphi,\varphi_t,\Omega)$, we express the latter set as follows:

$$\operatorname{epi}\big(\varphi+\delta(\cdot;\Gamma)\big)^*+\operatorname{epi}\big[\xi+\delta\big(\cdot;(\bar{x}+\eta\mathbb{B})\times Y\big)\big]^*,$$

which in turn reduces to the form

$$\operatorname{epi}\big[\varphi+\delta(\cdot;\Gamma)+\xi+\delta(\cdot;(\bar{x}+\eta\mathbb{B})\times Y)\big]^*$$

by using Lemma 7.49 and the continuity of the function $\xi + \delta\big(\cdot; (\bar{x} + \eta\mathbb{B}) \times Y\big)$ at $(\bar{x}, \bar{y}) \in \operatorname{dom}\big(\varphi + \delta(\cdot; \Gamma)\big)$. The latter set is weak* closed in $X^* \times Y^* \times \mathbb{R}$ as the epigraph of the conjugate function to $\varphi + \delta(\cdot; \Gamma) + \xi + \delta(\cdot; (\bar{x} + \eta\mathbb{B}) \times Y)$. Thus we are done with the proof of this lemma. △

As a consequence of Theorem 7.55, we derive necessary optimality conditions for the parametric DC program (7.170) that are *upper subdifferential* conditions according to the terminology of Section 6.1. Indeed, they involve *all* the upper subgradients of the concave function $-\psi$ at the reference point, which reduce to subgradients of the convex function ψ in the cost of (7.170).

Corollary 7.57 (Upper Subdifferential Conditions for Parametric DC SIPs). *Given a parameter value* $\bar{x} \in \operatorname{dom} M$ *in* (7.174)*, let* $\bar{y}$ *be a* (*global*) *optimal solution to the parametric DC program*

$$\text{minimize} \;\; \varphi(\bar{x}, y) - \psi(\bar{x}, y) \;\; \text{subject to} \;\; y \in F(\bar{x}) \cap G(\bar{x}) \tag{7.179}$$

with F *and* G *from* (7.171) *and* (7.172)*, respectively, under the standing assumptions made. Suppose in addition that* $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ *for the value function* (7.173) *under the CQC property for* $(\varphi, \varphi_t, \Omega)$*. Then for each* $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$ *and* $\gamma > 0$*, there are* $u^* \in X^*$ *and* $\lambda \in \mathbb{R}_+^{(T)}$ *from* (7.3) *such that*

$$\begin{aligned}
&u^* + x^* \in \partial_x\varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_x\varphi_t(\bar{x}, \bar{y}) + N_x\big((\bar{x}, \bar{y}); \Omega\big) + \gamma\mathbb{B}^*,\\
&y^* \in \partial_y\varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_y\varphi_t(\bar{x}, \bar{y}) + N_y\big((\bar{x}, \bar{y}); \Omega\big),\\
&\lambda_t\varphi_t(\bar{x}, \bar{y}) = 0 \;\; \text{for all} \;\; t \in \operatorname{supp}\lambda.
\end{aligned}$$

Proof. Follows directly from the upper estimate in Theorem 7.55 due to $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ and the construction of the KKT multiplier set in (7.176). △

The most restrictive and not easily verifiable assumption in Corollary 7.57 is that of $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$. In the next subsection, we derive improved necessary optimality conditions for (7.170) while replacing the restrictive requirement on $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ by more natural and verifiable assumptions in the case of Asplund spaces. This comes as a consequence of upper estimates for basic and singular subgradients of the DC value function (7.173) in more general settings.

7.5.3 Limiting Subgradients of Value Functions for DC SIPs

We begin with the constructive evaluation of the basic subdifferential (1.24) of the value function (7.173) and obtain two independent results in this direction under different assumptions and with completely different proofs. Recall from Section 1.5 (see also [522] for more details) that the basic subdifferential of any $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X at $\bar{x} \in \operatorname{dom}\varphi$ is defined by

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x},\, \varepsilon\downarrow 0} \widehat{\partial}_\varepsilon\varphi(x) \tag{7.180}$$

via the sequential outer limit of the ε-subdifferential mappings $\widehat{\partial}_\varepsilon\varphi\colon X \rightrightarrows X^*$ of φ at points nearby. If φ is l.s.c. around $\bar{x}$ and the space X is Asplund, then $\varepsilon > 0$ can be equivalently omitted in (7.180); see [522, Theorem 2.34].

For the first result, we need the following condition on the *minus* term ψ in (7.173), which allows us to derive a tight upper estimate of $\partial\mu(\bar{x})$.

Definition 7.58 (Inner Subdifferential Stability). *We say that a convex function* $\psi\colon X \to \overline{\mathbb{R}}$ *is* INNER SUBDIFFERENTIALLY STABLE *at* $\bar{x} \in \operatorname{dom}\psi$ *if*

$$\operatorname*{Lim\,inf}_{x \xrightarrow{\operatorname{dom}\psi} \bar{x}} \partial\psi(x) \neq \emptyset, \tag{7.181}$$

where Lim inf *stands for the Painlevé-Kuratowski inner limit* (1.20) *with the usage of the weak** *sequential convergence on* X^*.

Note that (7.181) reduces to a singleton in the case of general Banach spaces if ψ is Gâteaux differentiable on a neighborhood of $\bar{x}$ and its Gâteaux derivative operator $d\psi\colon X \to X^*$ is continuous with respect to the weak* topology of X^*. The next proposition relaxes the smoothness assumption *around* $\bar{x}$ provided that the closed unit ball $\mathbb{B}^*$ in X^* is *weak** *sequentially compact*. This latter property holds for general classes of Banach spaces X, in particular; for those admitting an equivalent norm Gâteaux differentiable at nonzero points (Gâteaux smooth), for weak Asplund spaces that includes every Asplund space and every weakly compactly generated space, every reflexive and every separable space, etc.; see, e.g., [255] for more details.

Proposition 7.59 (Sufficient Conditions for Inner Subdifferential Stability). *Let* X *be a Banach space such that the closed unit ball* $\mathbb{B}^*$ *is weak** *sequentially compact in* X^*, *and let* ψ *be convex, continuous, and Gâteaux differentiable at* $\bar{x} \in \operatorname{int}(\operatorname{dom}\psi)$. *Then* ψ *is inner subdifferentially stable at* $\bar{x}$.

Proof. Take any sequence $x_k \to \bar{x}$ as $k \to \infty$ and suppose that it entirely belongs to some neighborhood $U \subset \operatorname{dom}\psi$ of $\bar{x}$. It follows from the continuity of the convex function ψ at $\bar{x}$ that it is actually Lipschitz continuous around $\bar{x}$, and hence its subdifferential mapping $\partial\psi(\cdot)$ is bounded in X^* by the Lipschitz constant of ψ; see Exercises 1.69(i) and 7.102. This implies by using the weak* sequential compactness of the dual ball B^* that every subset of the set

$$V^* := \left\{x^* \in X^* \middle|\ \exists\, x \in U \ \text{ with } \ x^* \in \partial\psi(x)\right\}$$

contains a subsequence converging in the weak* topology of X^*. Then picking any sequence of subgradients $x_k^* \in \partial\psi(x_k)$, we suppose without loss of generality that there is $x^* \in X^*$ such that $x_k^* \xrightarrow{w^*} x^*$ as $k \to \infty$. It follows from convex subdifferential definition (1.35) with $\varepsilon = 0$ that $x^* \in \partial\psi(\bar{x})$. Since ψ is continuous and Gâteaux differentiable at $\bar{x}$, we have from standard convex analysis that

$\partial\psi(\bar{x}) = \{d\psi(\bar{x})\}$, and therefore $x_k^* \xrightarrow{w^*} d\psi(\bar{x})$ as $k \to \infty$. This clearly verifies the inner subdifferential stability (7.181) of ψ at $\bar{x}$. △

It is not hard to give various examples of functions, which are not Gâteaux differentiable at the reference point while being inner subdifferentially stable at it. Such functions can be constructed by the following scheme. Take a closed and convex subset Ω of a Gâteaux smooth space X, a point $\bar{x} \in \mathrm{bd}\,\Omega$, and a function $\theta(x)$ that is convex, continuous, and Gâteaux differentiable on an open set containing $\bar{x}$. Then define $\psi\colon X \to \overline{\mathbb{R}}$ as $\psi(x) := \theta(x)$ on Ω and as $\psi(x) := \emptyset$ otherwise. It follows from Definition 7.58 and Proposition 7.59 that $\mathrm{Lim\,inf}\,\psi$ in (7.181) reduces to $\{d\theta(\bar{x})\}$, and thus we have the inner subdifferential stability of ψ at $\bar{x}$. Observe that

$$\partial\psi(\bar{x}) = d\theta(\bar{x}) + N(\bar{x};\Omega)$$

by the subdifferential sum rule from Lemma 7.49 due the assumed continuity of θ. Taking into account our convention on $\infty - \infty = \infty$, we get a boundary domain point $\bar{x} \in \mathrm{bd}(\mathrm{dom}\,\psi)$, which is a local minimizer for the DC function $\varphi - \psi$ provided that $\mathrm{dom}\,\varphi \subset \mathrm{dom}\,\psi$.

Now we are ready to establish the aforementioned tight upper estimate of basic subgradients of the value function (7.173) under the inner subdifferential stability of ψ in (7.170). This result requires also the inner semicontinuity property (1.20) of the argminimum mapping $M(\cdot)$ from (7.174).

Theorem 7.60 (Basic Subgradients of DC Value Functions Under Inner Subdifferential Stability). *Given $(\bar{x},\bar{y}) \in \mathrm{gph}\,M$ in (7.170), suppose that $M(\cdot)$ is inner semicontinuous, that ψ is inner subdifferentially stable, and that CQC holds for $(\varphi,\varphi_t,\Omega)$ at this point. Then for any fixed $(x^*,y^*) \in \mathop{\mathrm{Lim\,inf}}_{(x,y)\xrightarrow{\mathrm{dom}\,\psi}(\bar{x},\bar{y})} \partial\psi(x,y)$, we have the inclusion*

$$\partial\mu(\bar{x}) \subset \partial_x\varphi(\bar{x},\bar{y}) - x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\mathrm{supp}\,\lambda} \lambda_t\partial_x\varphi_t(\bar{x},\bar{y})\Big] + N_x\big((\bar{x},\bar{y});\Omega\big)$$

with the set of KKT multipliers $\Lambda(\bar{x},\bar{y},y^*)$ defined in (7.176).

Proof. Fix the pair (x^*,y^*) from the theorem formulation, and pick an arbitrary subgradient $u^* \in \partial\mu(\bar{x})$. Then definition (7.180) gives us sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$, and $u_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ with $u_k^* \xrightarrow{w^*} u^*$ as $k \to \infty$. Fixing $k \in \mathbb{N}$ and using ε_k-subgradient construction (1.34) for u_k^*, we find $\eta_k > 0$ such that

$$\langle u_k^*, x - x_k\rangle \le \mu(x) - \mu(x_k) + 2\varepsilon_k\|x - x_k\| \ \text{ if } \ x \in x_k + \eta_k\mathbb{B}. \qquad (7.182)$$

The inner semicontinuity of $M(\cdot)$ at $(\bar{x},\bar{y})$ allows us to find a sequence of $y_k \in M(x_k)$ that contains a subsequence converging to $\bar{y}$; we suppose that $y_k \to \bar{y}$ for all $k \to \infty$. By the choice of (x^*,y^*), there is a sequence of subgradients $(x_k^*,y_k^*) \in \partial\psi(x_k,y_k)$ with $(x_k^*,y_k^*) \xrightarrow{w^*} (x^*,y^*)$ as $k \to \infty$. It follows from (7.174) and (7.182) that

$$\begin{aligned}&\langle u_k^*, x-x_k\rangle \le \varphi(x,y)-\psi(x,y)-\varphi(x_k,y_k)+\psi(x_k,y_k)\\&+2\varepsilon_k\big(\|x-x_k\|+\|y-y_k\|\big)\le \varphi(x,y)-\varphi(x_k,y_k)-\langle x_k^*, x-x_k\rangle-\langle y_k^*, y-y_k\rangle\\&+2\varepsilon_k\big(\|x-x_k\|+\|y-y_k\|\big)\ \text{ for all }\ (x,y)\in\Gamma\cap\big((x_k,y_k)+\eta_k\mathbb{B}\big).\end{aligned}$$

The latter implies in turn that the inequality

$$\begin{aligned}\langle u_k^*+x_k^*, x-x_k\rangle+\langle y_k^*, y-y_k\rangle \le \varphi(x,y)-\varphi(x_k,y_k)\\ +2\varepsilon_k\big(\|x-x_k\|+\|y-y_k\|\big)\end{aligned}$$

valid for all such (x, y) can be written as the ε-subdifferentials inclusion

$$(u_k^*+x_k^*, y_k^*)\in\widehat{\partial}_{2\varepsilon_k}\big(\varphi+\delta(\cdot;\Gamma)\big)(x_k,y_k)\ \text{ for all }\ k\in\mathbb{N}.$$

Passing now to the limit as $k\to\infty$ and taking into account the weak* convergence $(u_k^*+x_k^*, y_k^*)\xrightarrow{w^*}(u^*+x^*, y^*)$, we get from definition (7.180) that

$$(u^*+x^*, y^*)\in\partial\big(\varphi+\delta(\cdot;\Gamma)\big)(\bar{x},\bar{y}).\tag{7.183}$$

Since the function $\varphi+\delta(\cdot;\Gamma)$ is convex on $X\times Y$, the basic subdifferential in (7.183) reduces to the one of convex analysis. Thus applying to (7.183) the subdifferential sum rule for infinite systems from Corollary 7.52, which holds under the imposed CQC, gives us the inclusion

$$\partial\big(\varphi+\delta(\cdot;\Gamma)\big)(\bar{x},\bar{y})\subset\partial\varphi(\bar{x},\bar{y})+\bigcup_{\lambda\in A(\bar{x},\bar{y})}\Big[\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial\varphi_t(\bar{x},\bar{y})\Big]+N\big((\bar{x},\bar{y});\Omega\big)$$

with $A(\bar{x},\bar{y})=\{\lambda\in\mathbb{R}_+^{(T)}\mid \lambda_t\varphi_t(\bar{x},\bar{y})=0\ \text{ for all }\ t\in\operatorname{supp}\lambda\}$. Substituting it into (7.183) and taking into account the aforementioned relationships between the full and partial subdifferentials of convex functions, we arrive at

$$\begin{cases}u^*\in\partial_x\varphi(\bar{x},\bar{y})-x^*+\displaystyle\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x},\bar{y})+N_x\big((\bar{x},\bar{y});\Omega\big),\\ y^*\in\partial_y\varphi(\bar{x},\bar{y})+\displaystyle\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_y\varphi_t(\bar{x},\bar{y})+N_y\big((\bar{x},\bar{y});\Omega\big)\end{cases}$$

for some $\lambda\in A(\bar{x},\bar{y})$. This completes the proof of the theorem. △

As discussed above, the inner subdifferential stability of the minus term ψ in (7.170) required in Theorem 7.60 is a rather restrictive requirement. In the next theorem, we replace it by a much more flexible assumption on ψ that holds, in particular, for any *continuous* convex functions. The upper estimate for basic subgradients of (7.173) obtained under the following assumption is less precise in comparison with Theorem 7.60 while being sufficient for the majority of applications including those in this book.

Definition 7.61 (Subdifferential Boundedness). *We say that a convex function* $\psi\colon X \to \overline{\mathbb{R}}$ *is* SUBDIFFERENTIALLY BOUNDED *around* $\bar{x} \in \mathrm{dom}\,\psi$ *if for any sequences* $\varepsilon_k \downarrow 0$ *and* $x_k \overset{\mathrm{dom}\,\psi}{\longrightarrow} \bar{x}$ *as* $k \to \infty$ *there is a sequence of* $x_k^* \in \partial_{\varepsilon_k}\psi(x_k)$, $k \in \mathbb{N}$, *such that the set* $\{x_k^*|\ k \in \mathbb{N}\}$ *is bounded in* X^*.

As mentioned, this property holds for a broad class of convex functions.

Proposition 7.62 (Sufficient Condition for Subdifferential Boundedness of Convex Functions). *Let* $\psi\colon X \to \overline{\mathbb{R}}$ *be a convex function continuous at* $\bar{x} \in \mathrm{int}(\mathrm{dom}\,\psi)$. *Then* ψ *is subdifferentially bounded around this point.*

Proof. As well known in convex analysis (see Exercise 7.102), the continuity of a convex function ψ at the reference point $\bar{x} \in \mathrm{int}(\mathrm{dom}\,\psi)$ yields that ψ is locally Lipschitzian around $\bar{x}$. On the other hand, the local Lipschitz continuity of any (not only convex) function ensures the uniform boundedness of subgradients around the point in question; see Exercise 1.69. Furthermore, $\partial\psi(x) \subset \partial_\varepsilon\psi(x)$ for any $\varepsilon > 0$. Taking now arbitrary sequences $\varepsilon_k \downarrow 0$ and $x_k \overset{\mathrm{dom}\,\psi}{\longrightarrow} \bar{x}$ as $k \to \infty$, we have $x_k^* \in \partial_{\varepsilon_k}\psi(x_k)$ for any sequence of subgradients $x_k^* \in \partial\psi(x_k)$. This justifies the subdifferential boundedness of ψ. △

The following theorem provides a result independent of Theorem 7.60. Its proof involves the classical Brøndsted-Rockafellar theorem on subdifferential density in convex analysis, which is a predecessor and convex counterpart of the fundamental Ekeland's variational principle.

Theorem 7.63 (Basic Subgradients of Value Functions in DC Programs Under Subdifferential Boundedness). *Suppose that for both spaces* X *and* Y *the dual unit balls are sequentially weak* compact, that the argminimum mapping* (7.24) *is inner semicontinuous at some point* $(\bar{x}, \bar{y}) \in \mathrm{gph}\,M$, *that* ψ *in* (7.173) *is subdifferentially bounded around* $(\bar{x}, \bar{y})$, *and that CQC holds for* $(\varphi, \varphi_t, \Omega)$. *Then we have the upper estimate*

$$\partial\mu(\bar{x}) \subset \partial_x\varphi(\bar{x}, \bar{y}) + \bigcup_{(x^*,y^*)\in\partial\psi(\bar{x},\bar{y})} \Big\{ -x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\mathrm{supp}\,\lambda} \lambda_t \partial_x\varphi_t(\bar{x}, \bar{y}) \Big]\Big\} + N_x\big((\bar{x}, \bar{y}); \Omega\big).$$

Proof. Pick any $u^* \in \partial\mu(\bar{x})$, and similar to the proof of Theorem 7.60, find sequences $\varepsilon_k \downarrow 0$, $x_k \overset{\mu}{\to} \bar{x}$, and $u_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ satisfying $u_k^* \overset{w^*}{\to} u^*$ as $k \to \infty$. Then we get $\eta_k \downarrow 0$ such that inequality (7.182) holds and, by the assumed inner semicontinuity of $M(\cdot)$, obtain a sequence of $y_k \in M(x_k)$ converging to $\bar{y}$ as $k \to \infty$. Select further $\nu_k > 0$ with $2\sqrt{\nu_k} < \eta_k$ and, by taking into account that $\nu_k \downarrow 0$ and $(x_k, y_k) \to (\bar{x}, \bar{y})$ and by employing the subdifferential boundedness of ψ, find a sequence of $(x_k^*, y_k^*) \in \partial_{\nu_k}\psi(x_k, y_k)$ such that the set $\{(x_k^*, y_k^*) \in X^* \times Y^*|\ k \in \mathbb{N}\}$ is bounded. It follows from the structure of the ε-subdifferential mapping (7.101) that $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$. Similar to the proof of Theorem 7.60, we derive from (7.182) the inequality

$$\langle u_k^* + x_k^*, x - x_k\rangle + \langle y_k^*, y - y_k\rangle - \nu_k \le \varphi(x,y) - \varphi(x_k,y_k) \\ +2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big)$$

held for all $(x, y) \in \Gamma \cap \big((x_k, y_k) + \eta_k\mathbb{B}\big)$. This implies that

$$(u_k^* + x_k^*, y_k^*) \in \partial_{\nu_k}\vartheta_k(x_k, y_k), \quad k \in \mathbb{N}, \tag{7.184}$$

in terms of the ε-subdifferentials (with $\varepsilon := \nu_k$) of the convex l.s.c. functions $\vartheta_k(\cdot)$ given in the summation form

$$\begin{aligned}\vartheta_k(x,y) :&= \varphi(x,y) + \delta\big((x,y); \Gamma \cap [(x_k,y_k) + \eta_k\mathbb{B}]\big) \\ &-\varphi(x_k,y_k) + 2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big).\end{aligned} \tag{7.185}$$

Applying now to the elements in (7.184) the Brøndsted-Rockafellar density theorem, we find pairs $(\widetilde{x}_k, \widetilde{y}_k) \in \operatorname{dom}\vartheta_k$ and $(\widetilde{x}_k^*, \widetilde{y}_k^*) \in \partial\vartheta_k(\widetilde{x}_k, \widetilde{y}_k)$ satisfying for all $k \in \mathbb{N}$ the following inequalities:

$$\begin{aligned}&\|\widetilde{x}_k - x_k\| + \|\widetilde{y}_k - y_k\| \le \sqrt{\nu_k} \ \text{ and} \\ &\|\widetilde{x}_k^* - (u_k^* + x_k^*)\| + \|\widetilde{y}_k^* - y_k^*\| \le \sqrt{\nu_k}.\end{aligned} \tag{7.186}$$

They imply by the constructions above and the choice of ν_k that

$$\begin{aligned}&\langle \widetilde{x}_k^*, x - \widetilde{x}_k\rangle + \langle \widetilde{y}_k^*, y - \widetilde{y}_k\rangle \le \vartheta_k(x,y) - \vartheta_k(\widetilde{x}_k,\widetilde{y}_k) \le \varphi(x,y) - \varphi(\widetilde{x}_k,\widetilde{y}_k) \\ &+2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big) - 2\varepsilon_k\big(\|\widetilde{x}_k - x_k\| + \|\widetilde{y}_k - y_k\|\big) \\ &\le \varphi(x,y) - \varphi(\widetilde{x}_k,\widetilde{y}_k) + 2\varepsilon_k\big(\|x - \widetilde{x}_k\| + \|y - \widetilde{y}_k\|\big)\end{aligned}$$

for all $(x, y) \in \Gamma \cap \big((x_k, y_k) + \eta_k\mathbb{B}\big)$, which yields in turn the inclusions

$$(\widetilde{x}_k^*, \widetilde{y}_k^*) \in \widehat{\partial}_{2\varepsilon_k}\big(\varphi + \delta(\cdot;\Gamma)\big)(\widetilde{x}_k, \widetilde{y}_k), \quad k \in \mathbb{N}. \tag{7.187}$$

It easily follows from the convergence $(x_k, y_k) \to (\bar{x}, \bar{y})$, $(u_k^* + x_k^*, y_k^*) \xrightarrow{w^*} (u^* + x^*, y^*)$ and from the norm estimates in (7.186) that

$$(\widetilde{x}_k, \widetilde{y}_k) \to (\bar{x}, \bar{y}) \ \text{ and } \ (\widetilde{x}_k^*, \widetilde{y}_k^*) \xrightarrow{w^*} (u^* + x^*, y^*) \ \text{ as } \ k \to \infty.$$

Thus passing to the limit in (7.187) as $k \to \infty$ and using construction (7.180) of the basic subdifferential, we arrive at inclusion (7.183) as in the proof of Theorem 7.60, where the basic subdifferential agrees with the subdifferential of convex analysis for the convex function $\varphi + \delta(\cdot;\Gamma)$. Proceeding finally as in the proof of Theorem 7.60 by employing the subdifferential sum rule from Corollary 7.52, we complete the proof of the theorem. △

Our next results concern the singular subdifferential $\partial^\infty\mu(\bar{x})$ of the DC value function (7.173). According to (1.38) and Exercise 1.68, the singular subdifferential of any l.s.c. $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X is defined by

$$\partial^\infty\varphi(\bar{x}) := \mathop{\rm Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda,\, \varepsilon \downarrow 0}} \lambda\widehat{\partial}_\varepsilon\varphi(x) \tag{7.188}$$

via the sequential outer limit, where $\varepsilon > 0$ can be omitted if X is Asplund.

Theorem 7.64 (Singular Subgradients of Value Functions in DC Programs). *Suppose that the assumptions of Theorem 7.63 are satisfied with replacing CQC for* $(\varphi, \varphi_t, \Omega)$ *by the corresponding FMCQ* (7.159) *for* (φ_t, Ω) *in* (7.170)*. Assume in addition that* $\Gamma \subset \operatorname{dom}\varphi$ *for the set of feasible solutions* (7.175)*. Then we have the upper estimate*

$$\partial^\infty\mu(\bar{x}) \subset \bigcup_{\lambda \in \Lambda^\infty(\bar{x},\bar{y})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_x\varphi_t(\bar{x}, \bar{y})\Big] + N_x\big((\bar{x}, \bar{y}); \Omega\big), \tag{7.189}$$

where the set of singular multipliers is defined by

$$\Lambda^\infty(\bar{x}, \bar{y}) := \Big\{\lambda \in \mathbb{R}^{(T)}_+ \Big|\ 0 \in \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_y\varphi_t(\bar{x}, \bar{y}) + N_y\big((\bar{x}, \bar{y}); \Omega\big),$$
$$\lambda_t\varphi_t(\bar{x}, \bar{y}) = 0 \ \textit{for all}\ \ t \in \operatorname{supp}\lambda\Big\}.$$

Proof. Pick any $u^* \in \partial^\infty\mu(\bar{x})$, and by (7.188), find sequences

$$\lambda_k \downarrow 0,\ \ \varepsilon_k \downarrow 0,\ \ x_k \xrightarrow{\mu} \bar{x},\ \ u_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)\ \text{ with }\ \lambda_k u_k^* \xrightarrow{w^*} u^* \text{ as }\ k \to \infty.$$

Following the proof of Theorem 7.63, select sequences

$$\nu_k \downarrow 0\ \text{ as }\ k \to \infty, \quad y_k \in M(x_k), \ \text{ and }\ (x_k^*, y_k^*) \in \partial_{\nu_k}\psi(x_k, y_k), \quad k \in \mathbb{N},$$

such that $\{(x_k^*, y_k^*)\}$ weak* converges in $X^* \times Y^*$ to some $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$. Further, the application of the Brøndsted-Rockafellar theorem to the function $\vartheta_k(x, y)$ from (7.185) gives us sequences of $(\widetilde{x}_k, \widetilde{y}_k) \in \operatorname{dom}\vartheta_k$ and $(\widetilde{x}_k^*, \widetilde{y}_k^*) \in \partial\vartheta_k(\widetilde{x}_k, \widetilde{y}_k)$ satisfying the estimates in (7.186) and the subdifferential inclusions (7.187) for all $k \in \mathbb{N}$. Using the convexity of $\varphi + \delta(\cdot; \Gamma)$ and the assumption on $\Gamma \subset \operatorname{dom}\varphi$ allows us to rewrite (7.187) as

$$\langle\widetilde{x}_k^*, x - \widetilde{x}_k\rangle + \langle\widetilde{y}_k^*, y - \widetilde{y}_k\rangle \le \varphi(x, y) - \varphi(\widetilde{x}_k, \widetilde{y}_k) + 2\varepsilon_k\big(\|x - \widetilde{x}_k\| + \|y - \widetilde{y}_k\|\big)$$

for all $(x, y) \in \Gamma$ and $k \in \mathbb{N}$. This implies, by picking any $\gamma > 0$ and employing the lower semicontinuity of φ around $(\bar{x}, \bar{y})$, that

$$\begin{aligned}&\lambda_k\big[\langle\widetilde{x}_k^*, x - \widetilde{x}_k\rangle + \langle\widetilde{y}_k^*, y - \widetilde{y}_k\rangle\big] \le \lambda_k\big[\varphi(x, y) - \varphi(\widetilde{x}_k, \widetilde{y}_k) + 2\varepsilon_k\big(\|x - \widetilde{x}_k\| \\ &+\|y - \widetilde{y}_k\|\big)\big] \le \lambda_k\big[\varphi(x, y) - \varphi(\bar{x}, \bar{y}) + \gamma + 2\varepsilon_k\big(\|x - \widetilde{x}_k\| + \|y - \widetilde{y}_k\|\big)\big]\end{aligned}$$

for all $(x, y) \in \Gamma$ and all $k \in \mathbb{N}$ sufficiently large. Passing now to the limit as $k \to \infty$ and taking into account that the sequence $\{\widetilde{y}_k^*\}$ is bounded in Y^*, that $\lambda_k \downarrow 0$, and that $\lambda_k\widetilde{x}_k^* \xrightarrow{w^*} u^*$ by (7.186), we get the relationship

$$\langle u^*, x - \bar{x}\rangle \le 0 \ \text{ for all } \ (x, y) \in \Gamma,$$

which can be rewritten as $(u^*, 0) \in N((\bar{x}, \bar{y}); \Gamma)$. Applying the normal cone calculus for infinite systems from Corollary 7.53 gives us

$$(u^*, 0) \in \bigcup_{\lambda \in A(\bar{x},\bar{y})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y}) \Big] + N\big((\bar{x}, \bar{y}); \Omega\big)$$

with $A(\bar{x}, \bar{y}) = \{\lambda \in \mathbb{R}^{(T)}_+ |\ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0,\ t \in \operatorname{supp}\lambda\}$. The latter yields (7.189) and thus completes the proof of the theorem. △

The next theorem presents applications of the upper estimates for both basic and singular subdifferentials of the value function (7.173) established in Theorems 7.63 and 7.64 to derive efficient conditions ensuring the local Lipschitz continuity of (7.173) via the initial data as well as necessary optimality conditions for local optimality in the parametric DC semi-infinite program under consideration. The obtained results essentially use the *Asplund* property of the *parameter* space X; this is not required for the decision space Y.

Recall that characterizing the local Lipschitz continuity of any l.s.c. function φ on an Asplund space presented in Exercise 4.34(ii) involves both the triviality condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ for the singular subdifferential and the SNEC property of φ at the reference point in the case of infinite dimensions. While the condition $\partial^\infty\mu(\bar{x}) = \{0\}$ for the value function (7.173) is straightforward from Theorem 7.64, it is not the case for SNEC, which is fully independent from the above triviality condition. Nevertheless, the following lemma of its own interest shows that for the general class of *marginal/value functions*, including the one in (7.173), the SNEC property holds under natural assumptions on the initial problem data.

Lemma 7.65 (SNEC Property of Marginal Functions). *Let*

$$\mu(x) := \inf\big\{\phi(x, y)\big|\ y \in \Phi(x)\big\}, \quad x \in X, \tag{7.190}$$

where X is Asplund, where the argminimum map

$$x \mapsto S(x) := \big\{y \in \Phi(x)\big|\ \phi(x, y) = \mu(x)\big\}$$

is inner semicontinuous at some point $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ and where ϕ is locally Lipschitzian around this point. Then (7.190) is SNEC at $\bar{x}$ provided that it is l.s.c. around $\bar{x}$ and that the mapping Φ therein is Lipschitz-like around $(\bar{x}, \bar{y})$.

Proof. To verify the SNEC property of (7.190) at $\bar{x}$, we use its subdifferential characterization presented in Exercise 2.50. Based on this, take any sequences $\lambda_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$, and $x_k^* \in \lambda_k\widehat{\partial}\mu(x_k)$ with $x_k^* \xrightarrow{w^*} 0$, and then show that $\|x_k^*\| \to 0$ along some subsequence. To proceed, employ the inner semicontinuity of $S(\cdot)$ at $(\bar{x}, \bar{y})$ and select a sequence of $y_k \in S(x_k)$ whose subsequence converges (with no relabeling) to $\bar{y}$. Take $\widetilde{x}_k^* \in \widehat{\partial}\mu(x_k)$ such that $x_k^* = \lambda_k\widetilde{x}_k^*$. Since $\widetilde{x}_k^*$ is a regular subgradient of φ at x_k, for any $\eta > 0$, there is $\gamma > 0$ such that

$$\langle \widetilde{x}_k^*, x - x_k \rangle \le \mu(x) - \mu(x_k) + \eta\|x - x_k\| \text{ whenever } x \in x_k + \gamma\mathbb{B}.$$

Considering the extended-real-valued function

$$\xi(x, y) := \phi(x, y) + \delta\big((x, y); \operatorname{gph}\Phi\big) \text{ for all } (x, y) \in X \times Y,$$

we easily conclude from the above that

$$\big\langle(\widetilde{x}_k^*, 0), (x - x_k, y - y_k)\big\rangle \le \xi(x, y) - \xi(x_k, y_k) + \eta\big(\|x - x_k\| + \|y - y_k\|\big)$$

whenever $(x, y) \in (x_k, y_k) + \gamma\mathbb{B}$, which means that $(\widetilde{x}_k^*, 0) \in \widehat{\partial}\xi(x_k, y_k)$.

Fix now an arbitrary sequence $\varepsilon_k \downarrow 0$ as $k \to \infty$. Since ξ is locally Lipschitzian around $(\bar{x}, \bar{y})$, while X and Y are Asplund, we apply the fuzzy sum rule from Exercise 2.42 to the summation function ξ at (x_k, y_k) and thus find, by taking into account the convergence above, sequences

$$\begin{aligned}&(x_{1k}, y_{1k}) \xrightarrow{\phi} (\bar{x}, \bar{y}),\ (x_{2k}, y_{2k}) \xrightarrow{\operatorname{gph}\Phi} (\bar{x}, \bar{y}) \text{ as } k \to \infty,\\ &(x_{1k}^*, y_{1k}^*) \in \widehat{\partial}\phi(x_{1k}, y_{1k}), \text{ and } (x_{2k}^*, y_{2k}^*) \in \widehat{N}\big((x_{2k}, y_{2k}); \operatorname{gph}\Phi\big)\end{aligned}$$

such that $\lambda_k\|(x_{1k}^*, y_{1k}^*)\| \to (0, 0)$ with the estimates

$$\|\widetilde{x}_k^* - x_{1k}^* - x_{2k}^*\| \le \varepsilon_k \text{ and } \|y_{1k}^* + y_{2k}^*\| \le \varepsilon_k \text{ as } k \in \mathbb{N}. \tag{7.191}$$

This implies that $\lambda_k\|y_{2k}^*\| \to 0$ as $k \to \infty$. Taking now into account that

$$\big(\lambda_k x_{2k}^*, \lambda_k y_{2k}^*\big) \in \widehat{N}\big((x_{2k}, y_{2k}); \operatorname{gph}\Phi\big) \Longleftrightarrow \lambda_k x_{2k}^* \in \widehat{D}^*\Phi(x_{2k}, y_{2k})(-\lambda_k y_{2k}^*)$$

and that Φ is Lipschitz-like around $(\bar{x}, \bar{y})$ with some modulus $\ell > 0$, we get from the coderivative estimate for Lipschitz-like mappings (see implication (a)$\Rightarrow$(b) of Exercise 3.41, which holds in any Banach space) that

$$\|\lambda_k x_{2k}^*\| \le \ell\|\lambda_k y_{2k}^*\| \text{ for large } k \in \mathbb{N}.$$

This clearly yields $\lambda_k\|x_{2k}^*\| \to 0$. Combining the latter with (7.191) and with $x_k^* = \lambda_k\widetilde{x}_k^*$, we conclude that $\|x_k^*\| \to 0$ as $k \to \infty$ and thus justify the SNEC property of μ at $\bar{x}$ claimed in the lemma. $\triangle$

Now we are ready to establish the aforementioned major theorem.

Theorem 7.66 (Lipschitz Continuity of Value Functions and Optimality Conditions for Parametric DC SIPs). *Let the parameter space X be Asplund in the assumptions of Theorem 7.64 and suppose in addition that*

$$\Big\{\bigcup_{\lambda\in\Lambda^\infty(\bar{x},\bar{y})}\Big[\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x}, \bar{y})\Big] + N_x\big((\bar{x}, \bar{y}); \Omega\big)\Big\} = \{0\}. \tag{7.192}$$

Then the value function $\mu(\cdot)$ is locally Lipschitzian around $\bar{x}$ provided that it is l.s.c. around this point (which is ensured by the inner semicontinuity of $M(\cdot)$ around

$(\bar{x}, \bar{y})$) in each of the following two cases: either **(a)** $\dim X < \infty$ *or* **(b)** *both* φ *and* ψ *are continuous at* $(\bar{x}, \bar{y})$, *and the constraint mapping* $x \mapsto F(x) \cap G(x)$ *is Lipschitz-like around* $(\bar{x}, \bar{y})$.

If furthermore CQC holds for $(\varphi, \varphi_t, \Omega)$, *then we have the following necessary optimality conditions for the (global) minimizer* $\bar{y}$ *of the DC program* (7.179)*: there are* $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$, $u^* \in X^*$, *and* $\lambda \in \mathbb{R}_+^{(T)}$ *satisfying*

$$\begin{cases} u^* + x^* \in \partial_x \varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_x \varphi_t(\bar{x}, \bar{y}) + N_x\big((\bar{x}, \bar{y}); \Omega\big), \\ y^* \in \partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}) + N_y\big((\bar{x}, \bar{y}); \Omega\big), \\ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0 \text{ for all } t \in \operatorname{supp}\lambda. \end{cases} \tag{7.193}$$

Proof. If (7.192) holds, then $\partial^\infty \mu(\bar{x}) = \{0\}$ by Theorem 7.64. Further, it is easy to derive directly from the definitions that the lower semicontinuity of $\mu(\cdot)$ around $\bar{x}$ follows from the inner semicontinuity of $M(\cdot)$ around $(\bar{x}, \bar{y})$. Thus the local Lipschitz continuity of $\mu(\cdot)$ around $\bar{x}$ is a consequence of Theorem 1.22 in the case (a) where X is finite-dimensional.

In case (b), recall that the continuity of the convex functions φ and ψ at $(\bar{x}, \bar{y})$ implies their Lipschitz continuity around this point, and thus $\mu(\cdot)$ is SNEC at $\bar{x}$ due to Lemma 7.65. This verifies the first part of the theorem.

To justify the second part on the necessary optimality conditions, observe that any $\bar{y} \in M(\bar{x})$ under the consideration in this theorem is a *global* solution to (7.179). It follows from the local Lipschitz continuity of μ around $\bar{x}$ that $\partial\mu(\bar{x}) \neq \emptyset$; see Exercise 2.32(ii). Thus using the upper estimate of $\partial\mu(\bar{x})$ in Theorem 7.63 under the assumed CQC for $(\varphi, \varphi_t, \Omega)$, we conclude that the set on the right-hand side of this estimate is nonempty as well. This yields the claimed necessary optimality conditions (7.193) by construction (7.176) of the KKT multiplier set $\Lambda(\bar{x}, \bar{y}, y^*)$. △

Note that, in contrast to the necessary optimality conditions of Corollary 7.57, the results of (7.193) give us *lower* subdifferential optimality conditions in the enhanced form (with $\gamma = 0$ instead of $\gamma > 0$ in Corollary 7.57) under different while easily verifiable assumptions. Note also that the results of Sections 7.1 and 7.3 provide *characterizations* of the Lipschitz-like property of the infinite constraint inequality system in (7.179) entirely via the functions φ_t for the cases of linear, block-perturbed, and convex structures.

Convex ($\psi \equiv 0$) and concave ($\varphi \equiv 0$) SIPs are particular cases of the DC programs under consideration, and so the obtained results for general DC SIPs can be directly applied to these important cases with the corresponding specifications. Furthermore, the convex case allows us to derive new results, which cannot be deduced from those for general DC SIPs obtained above. The next theorem establishes a *precise formula* (equality, not inclusion) for calculating the subdifferential of the convex value function in such SIPs.

Theorem 7.67 (Calculating Subgradients of Value Functions in Convex SIPs). *Consider the value function* $\mu(\cdot)$ *from* (7.173) *with* $\psi \equiv 0$, *and suppose that CQC* *holds for the convex triple* $(\varphi, \varphi_t, \Omega)$ *in general Banach spaces. Then* $\mu(\cdot)$ *is convex, and its subdifferential at* $\bar{x} \in \operatorname{dom}\mu$ *is calculated by*

$$\partial\mu(\bar{x}) = \Big\{x^* \in X^* \Big| \ (x^*, 0) \in \partial\varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x}, \bar{y})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y})\Big] + N\big((\bar{x}, \bar{y}); \Omega\big)\Big\} \ \text{for any} \ \ \bar{y} \in M(\bar{x}),$$

where $A(\bar{x}, \bar{y}) := \big\{\lambda \in \mathbb{R}_+^{(T)} \big| \ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0, \ t \in \operatorname{supp}\lambda\big\}$.

Proof. The convexity of the value function (7.173) with $\psi \equiv 0$ and all the convex data easily follows from its definition and the convexity assumptions. To verify first the inclusion "$\subset$" in the claimed formula for $\partial\mu(\bar{x})$, we proceed as in the proof of Theorem 7.55 by taking $\gamma = 0$ and $\eta = \infty$.

To justify the opposite inclusion, pick any x^* from the right-hand side therein and thus find $\lambda \in A(\bar{x}, \bar{y})$, $(u^*, v^*) \in \partial\varphi(\bar{x}, \bar{y})$, $(u_t^*, v_t^*) \in \partial\varphi_t(\bar{x}, \bar{y})$, and $(\widetilde{u}^*, \widetilde{v}^*) \in N((\bar{x}, \bar{y}); \Omega)$ satisfying the equality

$$(x^*, 0) = (u^*, v^*) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t (u_t^*, v_t^*) + (\widetilde{u}^*, \widetilde{v}^*).$$

It follows from the construction of $A(\bar{x}, \bar{y})$ that for the chosen pairs (u^*, v^*), (u_t^*, v_t^*), and $(\widetilde{u}^*, \widetilde{v}^*)$, we have the relationships

$$\begin{cases} \varphi(x, y) - \mu(\bar{x}) = \varphi(x, y) - \varphi(\bar{x}, \bar{y}) \geq \langle u^*, x - \bar{x}\rangle + \langle v^*, y - \bar{y}\rangle, \\ 0 \geq \lambda_t \varphi_t(x, y) - \lambda_t \varphi_t(\bar{x}, \bar{y}) \geq \lambda_t \langle u_t^*, x - \bar{x}\rangle + \lambda_t \langle v_t^*, y - \bar{y}\rangle, \quad t \in \operatorname{supp}\lambda, \\ 0 \geq \langle \widetilde{u}^*, x - \bar{x}\rangle + \langle \widetilde{v}^*, y - \bar{y}\rangle \ \text{ whenever } \ (x, y) \in \Gamma, \end{cases}$$

which imply together with the above equality that

$$\varphi(x, y) + \delta\big((x, y); \Gamma\big) - \mu(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle \ \text{ for all } \ (x, y) \in X \times Y.$$

The latter shows in turn that $\mu(x) - \mu(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle$ for all $x \in X$ and hence completes the proof of the theorem. △

7.5.4 Bilevel Semi-infinite Programs with Convex Data

In this subsection, we return to optimistic bilevel programs studied in Chapter 6 for the case of finitely many inequality constraints at the lower level described by smooth as well as by locally Lipschitzian functions on finite-dimensional spaces. Here we consider *fully convex* bilevel programs in arbitrary *Banach* spaces with *infinite* constraints and derive for them necessary optimality conditions, which cannot be deduced from the results of Chapter 6 even in the case finitely many constraints in $\mathbb{R}^n$. Developing the value function approach allows us to reduce the bilevel programs under consideration to single-level DC SIPs and then apply the results obtained above in Section 7.5.

Consider the optimistic bilevel program

$$\begin{cases} \text{minimize} \;\; f(x,y) \;\; \text{subject to} \\ y \in M(x) := \big\{ y \in G(x) \big| \; \varphi(x,y) = \mu(x) \big\}, \end{cases} \tag{7.194}$$

where $M(x)$ is the set of optimal solutions to the lower-level problem

$$\text{minimize } \varphi(x,y) \text{ subject to } y \in G(x) := \big\{ y \in Y \big| \; \varphi_t(x,y) \le 0, \; t \in T \big\}$$

with an arbitrary index set T, and where $\mu(\cdot)$ is the optimal value function of the parametric lower-level problem defined by

$$\mu(x) := \inf \big\{ \varphi(x,y) \big| \; y \in G(x) \big\}. \tag{7.195}$$

The standing assumption of this section is that the bilevel problem (7.194) is *fully convex* on the Banach spaces X, Y meaning that all the functions there are l.s.c. and convex with respect to both variables.

To evaluate subgradients of the value function (7.195) and derive necessary optimality conditions for (7.194), we proceed via penalization under partial calmness. Observe that all the results of Subsection 6.2.3 apply to problem (7.194) with no change. Based on them, we get that any partially calm feasible solution $(\bar{x}, \bar{y})$ to (7.194) is a local optimal solution to the single-level program:

$$\begin{cases} \text{minimize} \;\; \kappa^{-1} f(x,y) + \varphi(x,y) - \mu(x) \\ \text{subject to} \;\; \varphi_t(x,y) \le 0, \;\; t \in T, \end{cases} \tag{7.196}$$

where $\kappa > 0$ is the constant of partial calmness, provided that the upper-level objective f is continuous at $(\bar{x}, \bar{y})$. Let us first efficiently evaluate the convex subdifferential of the value function (7.195) in the lower-level program.

Theorem 7.68 (Subgradients of Value Functions in Convex Bilevel Programs). *Let $(\bar{x}, \bar{y})$ be a partially calm feasible solution to the fully convex bilevel program (7.194). Suppose that CQC holds for (φ, φ_t) and that f is continuous at $(\bar{x}, \bar{y})$. Then there is a number $\kappa > 0$ such that*

$$\partial\mu(\bar{x}) \times \{0\} \subset \kappa^{-1} \partial f(\bar{x}, \bar{y}) + \partial\varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x}, \bar{y})} \Big[\sum_{t \in \operatorname{supp} \lambda} \lambda_t \partial \varphi_t(\bar{x}, \bar{y}) \Big],$$

where the set $A(\bar{x}, \bar{y})$ of active constraint multipliers is defined in Theorem 7.67. In particular, we have the upper estimate

$$\partial\mu(\bar{x}) \subset \kappa^{-1} \partial_x f(\bar{x}, \bar{y}) + \partial_x \varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x}, \bar{y})} \Big[\sum_{t \in \operatorname{supp} \lambda} \lambda_t \partial_x \varphi_t(\bar{x}, \bar{y}) \Big].$$

Proof. The second inclusion in the theorem clearly follows from the first one; so we verify the latter. The assumptions made ensure that $(\bar{x}, \bar{y})$ a local minimizer of the penalized problem (7.196), which is a DC SIP of type (7.157) described by the l.s.c. convex functions

$$\vartheta(x, y) := \kappa^{-1} f(x, y) + \varphi(x, y), \quad \theta(x, y) := \mu(x), \ \vartheta_t(x, y) := \varphi_t(x, y)$$

with $\Theta = X \times Y$ in (7.11). Let us check that the imposed CQC for (φ, φ_t) yields the validity of CQC for (ϑ, ϑ_t). Using the structure of the feasible set

$$\Xi := \left\{(x, y) \in X \times Y \middle| \ \varphi_t(x, y) \le 0 \ \text{ for all } \ t \in T\right\}$$

in (7.196), the well-known conjugate representation from convex analysis

$$\text{epi}\,(\varphi_1 + \varphi_2)^* = \text{cl}^*(\text{epi}\,\varphi_1^* + \text{epi}\,\varphi_2^*), \tag{7.197}$$

which is valid for any l.s.c. convex functions such that $\text{dom}\,\varphi_1 \cap \text{dom}\,\varphi_2 \neq \emptyset$ with omitting the weak* closure if one of the functions is continuous at some point $\bar{x} \in \text{dom}\,\varphi_1 \cap \text{dom}\,\varphi_2$, and then employing the imposed CQC give us

$$\begin{aligned}\text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\vartheta_t^*\Big] &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\varphi^* + \text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\varphi_t^*\Big]\\ &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\big(\varphi + \delta(\cdot;\Xi)\big)^* = \text{epi}\,\big(\vartheta + \delta(\cdot;\Xi)\big)^*.\end{aligned}$$

Applying further (7.197) without the closure operation to the above sum function ϑ with the continuous term f implies that

$$\begin{aligned}\text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\vartheta_t^*\Big] &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\varphi^* + \text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\varphi_t^*\Big]\\ &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\big(\varphi + \delta(\cdot;\Xi)\big)^* = \text{epi}\,\big(\vartheta + \delta(\cdot;\Xi)\big)^*\end{aligned}$$

and thus allows us to conclude that the set

$$\text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T}\text{epi}\,\vartheta_t^*\Big] \ \text{ is weak* closed in } \ X^* \times Y^* \times \mathbb{R}.$$

This is exactly the CQC property needed for the application of Theorem 7.51 to (7.196). Employing the latter result and the subdifferential sum rule

$$\partial\vartheta(\bar{x}, \bar{y}) = \partial\big(\kappa^{-1} f + \varphi\big)(\bar{x}, \bar{y}) = \kappa^{-1}\partial f(\bar{x}, \bar{y}) + \partial\varphi(\bar{x}, \bar{y}),$$

which holds by the continuity of f, we arrive at the first inclusion claimed in the theorem and thus complete the whole proof. △

Next we establish the main result of this subsection providing necessary optimality conditions for the fully convex bilevel programs with an arbitrary (finite or infinite) number of inequality constraints.

Theorem 7.69 (Necessary Optimality Condition for Fully Convex Bilevel SIPs). *Let $(\bar{x}, \bar{y})$ be a partially calm optimal solution to the fully convex bilevel program (7.194). Suppose that CQC holds for the lower-level program in (7.194), that the upper-level objective f is continuous at $(\bar{x}, \bar{y})$, and that $\partial\mu(\bar{x}) \neq \emptyset$ for the convex value function (7.195). Then for each $\widetilde{y} \in M(\bar{x})$ from the argminimum set in (7.194), there exist a number $\kappa > 0$ and multipliers $\lambda = (\lambda_t) \in \mathbb{R}_+^{(T)}$ and $\beta = (\beta_t) \in \mathbb{R}_+^{(T)}$ from the positive cone in (7.3) such that we have the following relationships:*

$$
\begin{aligned}
&0 \in \partial_x f(\bar{x}, \bar{y}) + \kappa\big[\partial_x\varphi(\bar{x}, \bar{y}) - \partial_x\varphi(\bar{x}, \widetilde{y})\big] + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial_x\varphi_t(\bar{x}, \bar{y}) \\
&\quad -\kappa \sum_{t\in\operatorname{supp}\beta} \beta_t \partial_x\varphi_t(\bar{x}, \widetilde{y}), \\
&0 \in \partial_y f(\bar{x}, \bar{y}) + \kappa\partial_y\varphi(\bar{x}, \bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial_y\varphi_t(\bar{x}, \bar{y}), \\
&0 \in \partial_y\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \beta_t \partial_y\varphi_t(\bar{x}, \widetilde{y}), \\
&\lambda_t\varphi_t(\bar{x}, \bar{y}) = \beta_t\varphi_t(\bar{x}, \widetilde{y}) = 0 \ \textit{for all} \ t \in T.
\end{aligned}
$$

Proof. Since $\partial\mu(\bar{x}) \neq \emptyset$, we take $x^* \in \partial\mu(\bar{x})$ and by Theorem 7.68 find $\kappa > 0$ and $\lambda \in \mathbb{R}_+^{(T)}$ satisfying the inclusion

$$
\kappa(x^*, 0) \in \partial f(\bar{x}, \bar{y}) + \kappa\partial\varphi(\bar{x}, \bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y}) \tag{7.198}
$$

with $\lambda_t\varphi_t(\bar{x}, \bar{y}) = 0$ as $t \in \operatorname{supp}\lambda$. On the other hand, picking $\widetilde{y} \in M(\bar{x})$ and applying to $x^* \in \partial\mu(\bar{x})$ the result of Theorem 7.67 give us $\beta \in \mathbb{R}_+^{(T)}$ such that

$$
x^* \in \partial_x\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \partial_x\varphi_t(\bar{x}, \widetilde{y}), \quad 0 \in \partial_y\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \partial_y\varphi_t(\bar{x}, \widetilde{y}),
$$

and $\beta_t\varphi_t(\bar{x}, \widetilde{y}) = 0$ for all $t \in \operatorname{supp}\beta$. Substituting this into (7.198) leads us to the claimed necessary optimality conditions. △

As an immediate consequence of Theorem 7.69, we get the following necessary optimality conditions for the bilevel SIP (7.194) involving only the reference optimal solution $(\bar{x}, \bar{y})$.

Corollary 7.70 (Specification of Necessary Optimality Conditions for Bilevel SIPs). *Let $(\bar{x}, \bar{y})$ be an optimal solution to (7.194) under the assumptions of Theorem 7.69. Then there are $\kappa > 0$ and $\lambda, \beta \in \mathbb{R}_+^{(T)}$ such that*

$$
\begin{aligned}
&0 \in \partial_x f(\bar{x}, \bar{y}) + \kappa\big[\partial_x \varphi(\bar{x}, \bar{y}) - \partial_x \varphi(\bar{x}, \bar{y})\big] + \sum_{t\in T}\Big[\big(\lambda_t - \kappa\beta_t\big)\partial_x \varphi_t(\bar{x}, \bar{y})\Big],\\
&0 \in \partial_y f(\bar{x}, \bar{y}) + \kappa\partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t\in T}\lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}),\\
&0 \in \partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t\in T}\beta_t \partial_y \varphi_t(\bar{x}, \bar{y}),\\
&\lambda_t \varphi_t(\bar{x}, \bar{y}) = \beta_t \varphi_t(\bar{x}, \bar{y}) = 0 \ \ \textit{for all} \ \ t \in T.
\end{aligned}
$$

Proof. Follows directly from Theorem 7.69 by putting $\widetilde{y} = \bar{y} \in M(\bar{x})$ in the necessary optimality conditions obtained therein. △

It has been well recognized in convex analysis that the subdifferentiability assumption $\partial\mu(\bar{x}) \neq \emptyset$ imposed in Theorem 7.69 and Corollary 7.70 is not restrictive. In particular, it holds in the Banach space setting of (7.195) under certain primal and dual qualification conditions; see Exercise 7.110.

7.6 Exercises for Chapter 7

Exercise 7.71 (Dual Description of Consistency for Infinite Linear Inequality Systems). Verify the equivalence in (7.7) by using convex separation. *Hint:* Compare it with the proof of [210, Theorem 3.1].

Exercise 7.72 (Interiority Conditions for Infinite Linear Systems). Prove the following statements for infinite inequality systems $\mathcal{F}$ in (7.1):

(i) If gph $\mathcal{F} \neq \emptyset$ and the set $\{a_t^*|\ t \in T\}$ is bounded, then int(gph $\mathcal{F}$) $\neq \emptyset$. *Hint:* Proceed similarly to the proof of implication (iii)⇒(v) in Theorem 7.2.

(ii) int(dom $\mathcal{F}$) $\neq \emptyset$ if gph $\mathcal{F} \neq \emptyset$ without the boundedness assumption.

Exercise 7.73 (Extended Farkas Lemma). Verify Proposition 7.3. *Hint:* Compare it with the proof in [210, Lemma 2.4].

Exercise 7.74 (Distance Function Representation of the Exact Lipschitzian Bound). Verify formula (7.26). *Hint:* Employ the equivalent between the Lipschitz-like property of F and the metric regularity one for F^{-1} established in Theorem 3.2(ii) with the exact bound relationship therein, and then proceed by using Definition 3.1(b) of the exact regularity bound for F^{-1}.

Exercise 7.75 (Existence of Best Approximations). Justify the existence of solutions to the optimization problem (7.28). *Hint:* Use the Alaoglu-Bourbaki theorem and the continuity of the mapping $x^* \mapsto \langle x^*, x\rangle$ in the weak* topology of X^*.

Exercise 7.76 (Fenchel Conjugates). Given a proper function $\varphi\colon X \to \overline{\mathbb{R}}$, verify the convexity and lower semicontinuity of the Fenchel conjugate (7.30).

Exercise 7.77 (Fenchel Conjugates for Suprema of Linear Functions). Prove the representations in (7.40). *Hint:* Compare it with [121] and [297].

Exercise 7.78 (Coderivative Calculation for Infinite Linear Inequality Systems). Calculate the coderivative for the general linear inequality system given in (7.48). *Hint:* Proceed as in the proof of Theorem 7.5.

Exercise 7.79 (Farkas-Minkowski Property for Infinite Linear Inequalities). Give sufficient conditions for the validity of the Farkas-Minkowski property (7.49) for the infinite linear system (7.48).

Exercise 7.80 (Equivalent Descriptions of the Strong Slater Condition for the Infinite Linear Inequality Systems). Formulate and prove a counterpart of Theorem 7.2 for the infinite linear constraint systems defined in (7.48).

Exercise 7.81 (Farkas-Minkowski Property from Strong Slater Condition).

(i) Verify that (7.58) implies the Farkas-Minkowski property in finite dimensions provided that the set $\text{co}\{-c_t^*, a_t^*, b_t)|\ t \in T\}$ is compact, and clarify whether the latter condition is essential for this statement.

(ii) Does it hold in infinite-dimensional spaces?

(iii) Does it hold in infinite dimensions if the set on the right-hand side of (7.58) is replaced by its weak* closure?

(iv) Does the strong Slater condition (7.57) for infinite linear systems always imply the Farkas-Minkowski property in finite-dimensional spaces?

Exercise 7.82 (Nonempty Graphical Interior for Infinite Linear Systems). Let X and P be arbitrary Banach spaces in (7.48).

(i) Show that SSC (7.57) and the boundedness of the set $\{(a_t^*, c_t^*)|\ t \in T\}$ in $X^* \times P^*$ imply that $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$.

(ii) Is either of these conditions necessary to have $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$?

(iii) Is either of these conditions essential to have $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$?

Exercise 7.83 (Lower Subdifferential Optimality Conditions in the KKT Form). Formulate and prove a lower subdifferential counterpart of Corollary 7.19.

Exercise 7.84 (Coderivatives of Block-Perturbed Infinite Linear Systems). Give a detailed proof of Proposition 7.24.

Exercise 7.85 (Characterization of SSC for Block-Perturbed Linear Systems). Give a detailed proof of the equivalence (iii)⇔(iv) in Proposition 7.25. *Hint:* Consider first the case of the maximum partition $\mathcal{J} = \mathcal{J}_{\max}$, and compare it with the proof in [298, Theorem 6.1].

Exercise 7.86 (Distance Function for Maximum Partition).

(i) Given a direct proof of assertion (7.91).

(ii) Prove that SSC for $\sigma(0)$ is equivalent to the inner/lower semicontinuity of $\mathcal{F}_{\max}$ (cf. [211, Theorem 5.1]), and deduce from it the property in (7.91).

Exercise 7.87 (Characteristic Set for Infinite Convex Inequalities). Obtain the characteristic set representation for convex inequality systems in (7.98) from that in (7.81) for block-perturbed linear systems.

Exercise 7.88 (Calculation of the Coderivative Norm for Convex Systems).

(i) Give an example when the equality holds in the setting of Theorem 7.31, while the set $\bigcup_{j \in J} \text{dom}\,\varphi_j^*$ is unbounded.

(ii) Is the reflexivity of X necessary for the equalities in Theorem 7.31?

(iii) Is the reflexivity of X essential for the equalities in Theorem 7.31?

Exercise 7.89 (Coderivative Criterion for Lipschitzian Stability of Convex Systems). Formulate and prove a convex counterpart of Proposition 7.25.

Exercise 7.90 (Metric Regularity from Lipschitzian Stability for Infinite Convex Inequality Systems). Derive a characterization of metric regularity for infinite convex inequality systems from the equality formula for the exact Lipschitzian bound obtained in Theorem 7.31.

Exercise 7.91 (Optimality Conditions for SIPs with Block-Perturbed Linear Constraints). Derive upper and lower subdifferential optimality conditions for minimizing extended-real-valued function subject to the infinite linear block-perturbed inequality constraints (7.77) in Banach and Asplund spaces, respectively.

Exercise 7.92 (Necessary Optimality Conditions for SIPs with Convex Inequality Constraints). Derive upper and lower subdifferential optimality conditions for minimizing extended-real-valued function subject to the infinite convex inequality constraints (7.94) in Banach and Asplund spaces, respectively.

Exercise 7.93 (Sum Rule for ε-Subgradients of Convex Functions). Given convex functions $\varphi_1, \varphi_2 \colon X \to \overline{\mathbb{R}}$ one of which is continuous at $\bar{x} \in \operatorname{dom} \varphi_1 \cap \operatorname{dom} \varphi_2$, justify the ε-subdifferential sum rule (7.102). *Hint:* Modify known proofs of the classical Moreau-Rockafellar theorem for the case of $\varepsilon > 0$ in (7.102); compare it, e.g., with the proof given in [757, Theorem 2.8.7].

Exercise 7.94 (Optimality Conditions in DC Programming). Consider the DC program defined in (7.107).

(i) Give a proof of the characterization of global minimizers in Lemma 7.33, and compare it with the one in [348].

(ii) Is the subdifferential inclusion formulated in Lemma 7.33 necessary for the local optimality of $\bar{x}$ in (7.107)?

(iii) Verify the sufficient condition for local minimizers in Lemma 7.34. *Hint:* Compare it with the proof in [235] given under the Lipschitz continuity of φ_2 around $\bar{x}$, and check that the latter assumption is equivalent to the continuity of φ_2 at $\bar{x}$.

(vi) Is the condition of Lemma 7.34 necessary for the local optimality in (7.107)?

Exercise 7.95 (Conditions for Calculating the Exact Regularity). Verify the relationships in (7.135), and show that the inclusion therein is generally strict.

Exercise 7.96 (Fenchel Conjugates for Suprema of Convex Functions).

(i) Given a direct proof of representation (7.154).

(ii) Verify formula (7.155) for the supremum of convex functions $f(x) := \sup_{t \in T} f_t(x)$. *Hint:* Deduce this, e.g., from [352, Vol. 2, Theorem 2.4.4].

Exercise 7.97 (Relationships Between CQC and FMCQ for Infinite Convex Systems). Consider the DC optimization problem (7.157), its feasible set Ξ (7.158), and the qualification conditions CQC (7.48) and FMCQ (7.159).

(i) Show that FMCQ$\Rightarrow$CQC if ϑ in (7.157) is continuous at some $x \in \Xi$.

(ii) Show that FMCQ$\Rightarrow$CQC if $\operatorname{cone}(\operatorname{dom} \vartheta - \Xi)$ is a closed subspace of X.

(iii) Give examples showing that CQC and FMCQ are generally independent.

Exercise 7.98 (Slater Constraint Qualification for Infinite Convex Systems). The convex inequality system $\{\vartheta_t(x) \le 0,\ \ t \in T \subset \mathbb{R}^m,\ \ x \in \mathbb{R}^n\}$ satisfies the *Slater qualification condition* (SCQ) if T is compact, the mapping $(t, x) \mapsto \vartheta_t(x)$ is continuous on $T \times \mathbb{R}^n$, and there is $x_0 \in \mathbb{R}^n$ such that $\vartheta_t(x_0) < 0$ for all $t \in T$.

(i) Show that SCQ$\Rightarrow$FMCQ if the set Ξ in (7.158) with $\Theta = \mathbb{R}^n$ is bounded.

(ii) Give an example of an infinite convex inequality system with $n = 2$ and $m = 1$ for which the converse implication in (i) is violated.

Exercise 7.99 (Conjugate Epigraphical and Subdifferential Sum Rules).

(i) Give a detailed proof of Lemma 7.49 and compare it with [131].

(ii) Construct an example showing that the subdifferential sum rule doesn't imply the epigraphical one therein.

(iii) Compare the equivalent epigraphical qualification conditions for the subdifferential sum rule given in Lemma 7.49 with other qualification conditions for this rule well recognized in convex and variational analysis in both finite and infinite dimensions; see [667, 757] and also the singular subdifferential condition (2.34) from Theorem 2.19 and Exercise 2.54(i).

Exercise 7.100 (Epigraphical Farkas Lemma).

(i) Give a detailed proof of Lemma 7.50 and compare it with [212].

(ii) Under which assumptions the weak* closure in Lemma 7.50(ii) can be replaced by the norm closure and when any closure operation can be omitted therein?

Exercise 7.101 (Epigraphs of Conjugate Functions via ε-Subdifferentials). Give a proof of representation (7.165) and compare it with [387].

Exercise 7.102 (Local Lipschitz Continuity of Convex Functions). Show that any convex function, which is continuous at some interior point of its domain, is locally Lipschitzian around this point.

Exercise 7.103 (Estimates for ε-Subgradients of Value Functions in DC SIPs). Derive a counterpart of Theorem 7.55 for ε-subgradients (1.34) of (7.173).

Exercise 7.104 (Basic Subgradients of DC Value Functions Under Extended Inner Semicontinuity). Using the definition of μ-inner semicontinuity given in Exercise 4.21, perform the following:

(i) Prove extended versions of Theorems 7.60, 7.63, and 7.64 with replacing the inner semicontinuity of the mapping $M(\cdot)$ therein by its μ-inner semicontinuity.

(ii) Construct examples showing the extensions obtained in this way are strictly better than the original formulations.

Exercise 7.105 (Closed-Graph Property of Subdifferential Mappings for l.s.c. Convex Functions on Banach Spaces).

(i) Let $\varphi\colon X \to \overline{\mathbb{R}}$ is a l.s.c. convex function on a Banach space. Prove that the graph of $x \mapsto \partial_\varepsilon \varphi(x)$ is closed in $X \times X^*$ for any $\varepsilon \geq 0$.

(ii) Show that $(x^*, y^*) \in \partial \psi(\bar{x}, \bar{y})$ in the proof of Theorem 7.63.

Exercise 7.106 (Relationships Between Subdifferential Upper Estimates for DC Value Functions). Let $\mu(\cdot)$ be the DC value function (7.173).

(i) Give an example showing that the upper estimate of $\partial \mu(\bar{x})$ from Theorem 7.60 may be better than the one in Theorem 7.63.

(ii) Investigate the possibilities to obtain upper estimates for $\partial \mu(\bar{x})$ by passing to the limit from that for regular subgradients in Theorem 7.55 in the case of Asplund (in particular, finite-dimensional) spaces and from the corresponding counterpart of Theorem 7.55 for the ε-enlargements $\widehat{\partial}_\varepsilon \mu(\cdot)$ in more general Banach space settings.

(iii) Clarify the same issues as in (ii) for the singular subdifferential $\partial^\infty \mu(\bar{x})$.

Exercise 7.107 (Lipschitz-Like Property of Feasible Solution Maps for Parameterized Versions of DC SIPs).

(i) Show that the Lipschitz-like property of the feasible solution map $x \mapsto F(x) \cap G(x)$ in the framework of Theorem 7.66 is essential for the validity of both stability and optimality conclusions of this theorem.

(ii) Based on characterizing the Lipschitz-like property of the infinite inequality systems in (7.172) obtained in Sections 7.1 and 7.3, impose appropriate assumptions on the constraint set Ω in (7.171) ensuring the feasible solution map $x \to F(x) \cap G(x)$ is Lipschitz-like at the reference point.

Exercise 7.108 (Upper Subdifferential Estimate for Value Functions in Convex SIPs). Give a detailed proof of the upper estimate of $\partial \mu(\bar{x})$ in Theorem 7.67.

Exercise 7.109 (Conjugate Epigraphical Representations). Verify representation (7.197), and show that the weak* closure can be omitted therein if one of the functions is continuous at some common point of the domains $\operatorname{dom} \varphi_i$, $i = 1, 2$. *Hint:* Compare it with the corresponding results and proofs in [116, 757].

Exercise 7.110 (Subdifferentiation of Value Functions for Convex Programs).

(i) Find appropriate qualification conditions ensuring subdifferentiability of value functions for convex programs with finitely many constraints in both finite and infinite dimensions. Are

Slater-type and subdifferential Mangasarian-Fromovitz constraint qualifications sufficient for this property?

(ii) Find appropriate qualification conditions ensuring subdifferentiability of value functions for convex SIPs in Banach spaces. *Hint:* Proceed first with dual constraint qualifications of the FMCQ and CQC types and then with primal ones of the Slater type; compare this with [210].

Exercise 7.111 (Value Functions and Optimality Conditions for Fully Convex SIPs with Upper-Level Constraints). Extend the results of Subsection 7.5.4 to bilevel SIPs with convex constraints at the upper level.

Exercise 7.112 (Comparison Between Lipschitzian and DC Approaches to Convex Bilevel Programs). Compare the necessary optimality conditions for fully convex bilevel programs containing finitely many inequality constraints that follow from Lipschitzian problems (see Theorems 6.21 and 6.23 and Exercise 6.46) with those obtained in Theorem 7.69 and Corollary 7.70 when the index set T is finite.

7.7 Commentaries to Chapter 7

Sections 7.1–7.3. *Semi-infinite programs* constitute a remarkable class of optimization problems that are intrinsically *infinite-dimensional* even in the case of linear inequality constraints on *finite-dimensional* decision variables. Their systematic study has started in the 1960s for SIPs with *linear* inequality systems and *compact* index sets being mainly motivated by applications to approximation theory, linear optimal control, and practical optimization models; see more information in [15, 298, 345] and their references. Then the study and applications have been extended to *convex* and also nonconvex while *differentiable* inequality systems over compact index sets as, e.g., in [96, 137, 394, 395, 396, 418, 442, 696, 783]. Note that the index set compactness was very essential in the obtained methods and results in these and related studies. More recently, further developments have been done for linear and convex systems with *arbitrary* index sets by using different techniques; see [139, 140, 141, 142, 210, 211, 212, 261, 299, 331, 464], among other publications. The major issues addressed in the SIP literature concerned well-posedness and ill-posedness properties, qualitative/topological and quantitative/Lipschitz-type stability analysis of parameterized feasible and optimal solution sets, necessary and sufficient optimality conditions, numerical methods, as well as various applications.

The material presented in Sections 7.1–7.3 is based on the author's joint papers with Cánovas, López, and Parra [140, 141, 142] dealing with robust Lipschitzian stability of parameterized infinite systems of linear, block-perturbed, and convex inequalities, necessary optimality conditions for minimizing nonsmooth functionals constrained by such systems, and some applications to water resource optimization. As seen above, methods and results of variational analysis and generalized differentiation presented in the previous chapters played a crucial role in these developments.

Section 7.4. This section is based on the author's joint paper with Nghia [548]. Note that, while the approach of [140] led us to complete qualitative and quantitative characterizations of the Lipschitz-like property of solution sets to linear infinite inequalities under adequate assumptions, its extension [142] to convex infinite inequalities via linear block perturbations and Fenchel duality ended up with a rather restrictive boundedness condition in the case of nonreflexive spaces; see Theorem 7.31 and Example 7.32. The latter condition was dismissed for a larger setting of perturbed infinite convex inequality and linear equality systems as a consequence of more general results on metric regularity of convex-graph multifunctions between arbitrary Banach spaces. The novel approach of [548] reduced the study of metric regularity for such mappings to the *unconstrained minimization* of *DC functions* and brought us to precise calculation of the exact regularity bounds of convex-graph multifunctions and infinite constraint systems via ε-coderivative and coderivative norms. Lemma 7.33 from global DC optimization was established by Hiriart-Urruty [348], while its local counterpart in Lemma 7.34 was obtained by Dür [235].

Corollary 7.37, summarizing the previous developments of this section, presents a major result of [548] allowing us to *precisely calculate* the *exact covering bound* of a general convex-graph multifunction between Banach spaces without additional assumptions. It implies, in particular, the regularity formula (7.149) for infinite convex constraint systems under the bounded SSC introduced in [548]. Note that another proof of (7.149) is given, in a different form under a certain uniform boundedness condition on the functions f_t, in the parallel study [373] based on the previous developments in [377] on *perfect regularity* for convex-graph multifunctions. However, there is a mistake in the proof of the aforementioned result in [373] due to the incorrect application on p. 1025 therein of the classical Sion's minimax theorem [691] whose assumptions fail to fulfill in the setting under consideration in [373].

Section 7.5. This section is mainly based on the author's joint paper with Dinh and Nghia [215] and is devoted to the subdifferentiation of the optimal value functions in DC SIPs with various applications. Note that the optimal value/marginal function for such problems is generally nonconvex, while evaluating its both basic and singular limiting subdifferentials gives us a crucial information concerning sensitivity analysis, optimality conditions, and their applications in finite and infinite dimensions. An important role in our analysis is played by the closedness qualification conditions from Definition 7.48, introduced and comprehensively studied by the same team [214] in the general LCTV space setting. In the latter paper the reader can find more discussions on the genesis of CQC and its relationships with the Farkas-Minkowski property as well as with other well-recognized constraint qualifications for finite and infinite convex systems of both primal and dual types; cf. also [116, 120, 121, 212, 213, 303, 479, 757] and the references therein. Lemma 7.49, taken from Burachik and Jeyakumar [131], provides probably the weakest conditions for the validity of the convex subdifferential sum rule in Banach spaces. Note that the equivalence between assertions (i) and (ii) in this result follows from the well-known formula (7.197). Lemma 7.50 established by Dinh et al. [212] is yet another extension of the classical Farkas lemma to infinite convex constraint systems; see the recent survey [213] on more results and discussions in this direction. Lemma 7.65 of its own interest is taken from the author's paper with Nam [532].

The last subsection of Section 7.5 implements the value function approach described in Chapter 6, together with the subdifferential results obtained above in this section, to the case of fully convex bilevel semi-infinite programs in Banach spaces indexed by arbitrary sets. Observe that in this way, we are able to significantly improve the results presented in Chapter 6, while specified to the fully convex setting, even for finitely many inequality constraints in finite-dimensional spaces.

Section 7.6. This section contains various exercises with different levels of difficulties concerning all the basic material presented in Chapter 7. As usual, we provide hints and references for the most difficult exercises. Similarly to the results of Chapter 6 on bilevel programs with finitely many constraints, relaxing the partial calmness assumption remains a challenging issue. It seems also that the pessimistic version of bilevel SIPs is Terra incognita in bilevel optimization.