ε-Optimality and ε-Lagrangian Duality for a Nonconvex Programming Problem with an Infinite Number of Constraints [1]

T.Q. Son[2], J.J. Strodiot[3], and V.H. Nguyen[3]

Communicated by J.P. Crouzeix

Abstract. In this paper ε-optimality conditions are given for a nonconvex programming problem which has an infinite number of constraints. The objective function and the constraint functions are supposed to be locally Lipschitz on a Banach space. In a first part, we introduce the concept of regular ε-solution and propose a generalization of the Karush-Kuhn-Tucker conditions. These conditions are up to ε and obtained by weakening the classical complementarity conditions. Furthermore they are satisfied without assuming any constraint qualification. Then we prove that these conditions are also sufficient for ε-optimality when the constraints are convex and the objective function ε-semiconvex. In a second part, we define quasi saddle-points associated with an ε-Lagrangian functional and we investigate their relationships with generalized KKT conditions. In particular, we formulate a Wolfe-type dual problem which allows us to present ε-duality theorems and relationships between KKT conditions and regular ε-solutions for the dual. Finally we apply these results to two important infinite programming problems: the cone-constrained convex problem and the semi-definite programming problem.

Key Words. Karush-Kuhn-Tucker conditions up to ε, approximate solution, quasi saddle-point, ε-Lagrange duality.

[1]The first author gratefully acknowledges the support of the Optimization and Applied Mathematics Program of the Belgian CIUF-CUD (CUI-UNV02 and CUI-MT03) during his three month stay at the Department of Mathematics, University of Namur, Namur, Belgium, in 2006. The authors wish to thank the referees whose comments and suggestions allow them to improve the paper by enlarging the perspectives of applications.

[2]Lecturer, Nhatrang Teacher College, Nhatrang, Vietnam.

[3]Professor, Department of Mathematics, University of Namur (FUNDP), Namur, Belgium.

1 Introduction

An infinite programming problem is an infinite dimensional optimization problem whose feasible set is described by infinitely many constraints. In this paper we consider the following infinite program:

$$(P)\begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & f_t(x) \le 0,\ t \in T, \\ & x \in C, \end{cases}$$

where T is an arbitrary (possibly infinite) index set, the functions f, $f_t : X \to \mathbb{R}$, $t \in T$, are locally Lipschitz functions on a Banach space X, and C is a nonempty closed subset of X.

The reader will find many applications of this problem in different fields such as Chebyshev approximation, robotics, mathematical physics, engineering design, optimal control, transportation problems, cooperative games, robust optimization, etc. There are also significant applications in statistics. A large list of applications as well as an introduction to the theoretical basis and to the numerical methods can be found in the recent survey by López and Still (Ref. 1) (see also the numerous references quoted in this paper). Let us also mention the books (Refs. $2-4$), and the recent references on generalized semi-infinite programming problems (Refs. $5-7$).

The aim of this paper is to present Karush-Kuhn-Tucker type conditions for ε-optimality without assuming constraint qualification hypothesis and to provide some corresponding duality results for problem (P). Characterizing approximate solutions of an optimization problem is essential because, numerically, only approximate solutions can be obtained. The most common definition of an approximate minimum of a function f from X to $\mathbb{R}$ is that of an ε-minimum, i.e., of a point z such that $f(z) \le f(x) + \varepsilon$ for all $x \in X$ where $\varepsilon > 0$ is some tolerance. It is clear that this definition has a global character and is quite well suitable for approximate minima of a convex function. For nonconvex functions, it is crucial to use local concepts as the following one: a point z is an ε-quasiminimum of f if z is a local minimum of the function $x \mapsto f(x) + \sqrt{\varepsilon}\,\|x - z\|$. It is very easy to observe that for such a point z, there exists a ball B around z with radius less than $\sqrt{\varepsilon}$ such that $f(z) \le f(x) + \varepsilon$ for all $x \in B$. When the function f is locally Lipschitz from $\mathbb{R}^n$ into $\mathbb{R}$, an ε-quasiminimum of f satisfies the properties: there exists $g \in \partial^c f(z)$ such that $\|g\| \le \sqrt{\varepsilon}$ and $f^\circ(z; d) \ge -\sqrt{\varepsilon}$ for any normalized direction in $\mathbb{R}^n$. Here $\partial^c f(z)$ and $f^\circ(z; d)$ denote the Clarke generalized gradient of f at z and the Clarke generalized derivative of f at z in the direction d (Ref. 8) (see also Ref. 9 for other properties of an ε-quasiminimum). Let us observe that the first property is also used as a stopping criterion in several bundle algorithms (see, e.g., Ref. 10).

When T is infinite and when all the functions f, f_t, $t \in T$, are locally Lipschitz on an Euclidean space, Zheng and Yang (Ref. 11) propose exact optimality conditions for (P) based on variational analysis. So, Lagrange multiplier rules are provided in terms of epi-coderivatives when a Slater-type constraint qualification is satisfied. In the case of convex constraints, many exact optimality conditions have been proposed (see, e.g., Refs. 12 − 14 for a convex objective function, and Ref. 15 for a semiconvex function). These conditions, called Karush-Kuhn-Tucker conditions (KKT conditions in short), are derived under the main assumption that the system

$$\sigma := \{f_t(x) \leq 0, t \in T, x \in C\}$$

is a Farkas Minkowski system. This assumption is a constraint qualification condition weaker than the well-known Slater's condition (see, e.g., Ref. 16, Proposition 2.1).

When T is finite, many papers present necessary conditions and saddle-point characterizations for approximate optimization (see, e.g., Refs. 17 − 22). In the last three papers the problem is convex and the ε-subdifferential of a convex function is suitable to obtain the approximate optimality conditions. Indeed, it is well known that z is an ε-minimum of the convex function f if and only if 0 belongs to the ε-subdifferential of f at z. In Refs. 17 − 19, the problem is no more convex and the approximate conditions are obtained thanks to the Ekeland Variational Principle (Ref. 23). In particular, this principle entails the existence of a pair (x^*, λ^*) that satisfies approximately the KKT conditions, without assuming any constraint qualification hypothesis. In Refs. 18 − 19, the generalized gradient of Clarke is used while in Ref. 17, it is the limiting subdifferential of Mordukhovich that is used. Furthermore, these three papers give characterizations of quasi saddle-points.

The aim of this paper is to generalize the theory of approximate optimality conditions and the characterization of quasi saddle-points to the case where the set T is infinite. First we present necessary conditions for approximate solutions under a Slater-type constraint qualification hypothesis. This is done by applying the calculus rule giving the generalized gradient of the function $F = \sup_{t\in T} f_t$ with respect to the generalized gradient of the functions f_t. This result allows us to define a generalized KKT pair up to ε, and to examine the existence and the properties of such KKT pairs without considering any constraint qualification hypothesis. This is obtained by applying a result due to Loridan (Ref. 19, Theorem 5.2) to the problem: minimize $f(x)$ subject to $F(x) \leq 0$ and $x \in C$. Afterwards, we particularize our results to two important examples of infinite programming problems. The first one is the cone-constrained convex problem where the constraints are of the form: $g(x) \in -K$ with K a closed convex cone. The second one is a particular case of the previous problem. This is the so-called semi-definite problem and it is a problem whose constraints require the positive semidefiniteness of a certain matrix.

In the second part of this paper, we extend the ε-Lagrangian functional defined in Ref. 19 to the case of infinitely many constraints. In particular we prove that under semiconvexity properties, any generalized KKT pair up to ε is a quasi saddle-point of the ε-Lagrangian functional, and conversely, that any quasi saddle-point of the ε-Lagrangian functional gives rise to an approximate solution. These two theorems generalize two results obtained by Loridan (Ref. 19). To end this paper we introduce a dual problem associated with the infinite problem (P), and we study the duality properties and existing relationships between generalized KKT pairs, quasi saddle-points, and approximate solutions of the dual problem.

The paper is divided in five sections. After this introduction, a section is devoted to preliminaries and to calculus rules concerning the Clarke generalized gradient of the function $F = \sup_{t\in T} f_t$. This rule allows us to derive KKT-type conditions for problem (P). In Section 3, we define several concepts of approximate solution and we recall the generalized KKT conditions up to ε introduced by Loridan when T is finite. In Section 4, we introduce generalized KKT conditions up to ε when T is infinite and we study their relationships with approximate solutions. We also examine the particular cases of cone-constrained convex problems and semi-definite problems. Finally, in Section 5, we extend the ε-Lagrangian duality to the infinite problem (P) and its Wolfe-type dual (D).

2 Preliminaries

Throughout this paper we assume that X is a Banach space, T is a compact topological space and $f : X \to \mathbb{R}$ is a locally Lipschitz function. We also assume that the constraint functions $f_t : X \to \mathbb{R}$, $t \in T$, are locally Lipschitz with respect to x uniformly in t, i.e., that for each $x \in X$, there exists a neighborhood U of x and a constant $K > 0$ such that

$$|f_t(u) - f_t(v)| \leq K \|u - v\| \quad \forall\, u,\, v \in U \text{ and } \forall\, t \in T.$$

We also suppose that the function $t \mapsto f_t(x)$ is upper semicontinuous on T for every $x \in X$. We also set, for every $x \in X$,

$$F(x) = \max_{t\in T} f_t(x) \qquad \text{and} \qquad T(x) = \{t \in T \mid f_t(x) = F(x)\}.$$

It is easy to prove that F is finite and locally Lipschitz on X. Finally let us recall the following definitions and basic concepts associated with a locally Lipschitz function $g : X \to \mathbb{R}$. First the Clarke generalized directional derivative (Ref. 8) of g at x in the direction $d \in X$ is denoted $g^\circ(x; d)$ and defined by:

$$g^\circ(x; d) = \limsup_{\substack{h\to 0\\ \lambda\downarrow 0}} \frac{g(x + h + \lambda d) - g(x + h)}{\lambda}.$$

The Clarke generalized gradient (Ref. 8) of g at x is defined by

$$\partial^c g(x) = \{x^* \in X^* \mid x^*(d) \leq g^\circ(x; d) \text{ for all } d \in X\},$$

where X^* denotes the dual of X. When g is convex, $\partial^c g(x)$ coincides with the subdifferential $\partial g(x)$ in the sense of convex analysis. In the sequel, we will use the same notation $\partial g(x)$ for g convex and for g locally Lipschitz. The function g is called quasidifferentiable (Ref. 24) (or regular (Ref. 8)) at x if, for each $d \in X$, the classical directional derivative $g'(x; d)$ exists and coincides with $g^\circ(x; d)$.

The generalized gradient of the max-function $F(x) = \max_{t\in T} f_t(x)$ is given in the next theorem which is a non-integral version of Theorem 2 in Ref. 25. Since the proof is based on a similar theorem but for the subdifferential of a maximum of convex functions, first we recall this result.

Lemma 2.1 *Let U be an open set of X, let S be a compact set, let $\{g_s\}_{s\in S}$ be a family of convex functions on X, and let $v_0 \in U$. Let also $G = \sup_{s\in S} g_s$. Assume that the function $s \mapsto g_s(v)$ is upper semicontinuous on S for each $v \in X$ and that the function g_s is continuous at v_0 for each $s \in S$.*
If (i) X is separable or if (ii) the function $s \mapsto g_s(v)$ is continuous on S for each $v \in X$, then

$$\partial G(v_0) = \overline{co}\, \{\cup\, \partial g_s(v_0) \,|\, s \in S(v_0)\}\,,$$

where $S(v_0) = \{s \in S \,|\, g_s(v_0) = G(v_0)\}$ and $\overline{co}(\cdot)$ denotes the closed convex hull with the closure taken in the weak* topology of the dual space X^*.

Proof. This result is due to Ioffe and Tihomirov (Ref. 26, Theorem 3, p.201) when (i) holds, and to Valadier (Ref. 27, Theorem 2.15, p.69) when (ii) holds. □

Theorem 2.1 *If at least one of the following holds:*

(a) X is separable; or

(b) T is metrizable and $\partial f_t(x)$ is lower semicontinuous (w^*) in t for each $x \in X$,

then, for any $x \in X$ and any $v \in X$, $F^\circ(x; v) \le \max_{t\in T(x)} f_t^\circ(x; v)$ and

$$\partial F(x) \subset \overline{co}\, \{\cup\, \partial f_t(x) \,|\, t \in T(x)\}\,. \tag{1}$$

Moreover, if the functions f_t are quasidifferentiable for every $t \in T$, then F is quasidifferentiable and the equality holds in (1).

Proof. Let $x \in X$. Using the first part of the proof of Theorem 2 in Ref. 25, we directly obtain the inequality

$$F^\circ(x; v) \le \max_{t\in T(x)} f_t^\circ(x; v) \quad \text{for all } v.$$

Note that the function $t \mapsto f_t^\circ(x; v)$ is upper semicontinuous and $T(x)$ is compact, so that the notation "max" is justified. Now let $\zeta \in \partial F(x)$. The previous inequality implies

$$\max_{t\in T(x)} \hat{f}_t(v) \ge \langle v, \zeta\rangle \quad \text{for all } v,$$

where $\hat{f}_t(v) = f_t^\circ(x; v)$. Since each $\hat{f}_t$ is convex in v and $\hat{f}_t(0) = 0$, we can conclude that ζ belongs to the subdifferential at 0 of the function G defined for each v, by $G(v) = \max_{t\in T(x)} \hat{f}_t(v)$. On the other hand, for every t, $\hat{f}_t$ is continuous at $v = 0$, and for every v, the function $t \mapsto \hat{f}_t(v)$ is upper semicontinuous. So when X is separable, Lemma 2.1 can be applied to obtain that

$$\partial G(0) = \overline{\text{co}} \left\{\cup\, \partial \hat{f}_t(0) \,|\, t \in S(0)\right\}, \tag{2}$$

where, for each t, $S(0) = \{t \in T(x) \,|\, \hat{f}_t(0) = G(0)\}$. When condition (b) holds, the function $t \mapsto \hat{f}_t(v)$ is continuous for every v, and Lemma 2.1 can also be applied to obtain (2). But this gives the announced result because $S(0) = T(x)$ and $\partial \hat{f}_t(0) = \partial f_t(x)$. □

When X is a finite dimensional space, the set $\{\cup\, \partial f_t(x) \,|\, t \in T(x)\}$ is compact, and consequently its convex hull is always closed.

Let D be a nonempty closed subset of X. Then the normal cone to D at $x \in D$ (Ref. 8, p.51) is defined by

$$N_D(x) = \{\, x^* \in X^* \,|\, x^*(v) \leq 0 \text{ for all } v \in T_D(x)\},$$

where $T_D(x) = \{v \in X \,|\, d_D^\circ(x; v) = 0\}$ denotes the tangent cone to D at x and d_D the distance function to D.

When D is convex, $N_D(x)$ coincides with the normal cone in the sense of convex analysis:

$$N_D(x) := \{x^* \in X^* \mid x^*(y - x) \leq 0 \text{ for all } y \in D\}.$$

Next, we also need to recall some optimality conditions for problem (P). In that purpose, let

$$A = \{x \in C \,|\, f_t(x) \leq 0 \text{ for all } t \in T\},$$

where C is a closed subset of X. Then A is closed and (Ref. 8, Corollary p.52),

$$x \text{ local minimum of } f \text{ over } A \quad \Rightarrow \quad 0 \in \partial f(x) + N_A(x). \tag{3}$$

In order to derive KKT conditions, we have to express the normal cone $N_A(x)$ in terms of the generalized gradient of the functions f_t and the normal cone to C at x. More precisely, let $F = \max_{t \in T} f_t$. By Theorem 6 of Ref. 28, we obtain that

$$N_A(x) \subset N_C(x) + \mathbb{R}_+ \partial F(x), \tag{4}$$

provided that the following constraint qualification condition holds when $F(x) = 0$:

$$\exists\, d \in T_C(x) \quad \text{such that} \quad F^\circ(x; d) < 0.$$

From (3), (4) and Theorem 2.1, we easily obtain the following result.

Proposition 2.1 *Let $x \in A$ and let $I(x) = \{t \in T \,|\, f_t(x) = 0\}$. Assume that the hypotheses of Theorem 2.1 are satisfied. If the following constraint qualification condition holds:*

$$(\mathcal{U}) \qquad \exists\, d \in T_C(x) \quad \textit{such that} \quad \forall\, t \in I(x) \quad f_t^\circ(x; d) < 0,$$

then

$$x \textit{ local minimum of } (P) \quad \Rightarrow \quad 0 \in \partial f(x) + \mathbb{R}_+ \overline{co}\, \{\cup\, \partial f_t(x) \,|\, t \in I(x)\} + N_C(x).$$

Let us denote by $\mathbb{R}^{(T)}$ the linear space of generalized finite sequences $\lambda = (\lambda_t)_{t\in T}$ with $\lambda_t \in \mathbb{R}$ for each $t \in T$, and with only finitely many λ_t different from zero. With $\lambda = (\lambda_t) \in \mathbb{R}^{(T)}$, we associate its supporting set $T(\lambda) := \{t \in T \mid \lambda_t \neq 0\}$. We also denote by $\mathbb{R}_+^{(T)} := \{\lambda = (\lambda_t) \in \mathbb{R}^{(T)} \mid \lambda_t \geq 0, t \in T\}$ the nonnegative cone of $\mathbb{R}^{(T)}$. It is easy to see that this cone is convex.

In order to simplify the writing, we introduce the following notation:

$$\sum_{t\in T} \lambda_t z_t = \begin{cases} \sum_{t\in T(\lambda)} \lambda_t z_t & \text{if } \; T(\lambda) \neq \emptyset, \\ 0 & \text{if } \; T(\lambda) = \emptyset, \end{cases}$$

for $\lambda \in \mathbb{R}^{(T)}$ and $\{z_t\}_{t\in T} \subset Z$, Z being a real linear space.

When X is a finite dimensional space, the subset $\operatorname{co}\{\cup\,\partial f_t(x) \,|\, t \in I(x)\}$ is closed, which gives the following optimality condition: if $(\mathcal{U})$ is satisfied and if x is an optimal solution to (P), then there exists $\lambda \in \mathbb{R}_+^{(T)}$ such that

$$0 \in \partial f(x) + \sum_{t\in T} \lambda_t\, \partial f_t(x) + N_C(x) \quad \text{ and } \quad f_t(x) = 0 \;\forall t \in T(\lambda). \tag{5}$$

When the functions f_t, $t \in T$, are convex, condition $(\mathcal{U})$ coincides with the classical Slater condition: $\exists\, x_0 \in C$ such that $F(x_0) < 0$. Recently, another constraint qualification hypothesis weaker than the Slater condition (see Ref. 16, Proposition 2.1) has been introduced for characterizing the normal cone to A at $x \in A$. However, in order to facilitate its introduction, let us recall that if $g : X \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous convex function defined on X, then its conjugate function is the function $g^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined for all $x^* \in X^*$ by

$$g^*(x^*) = \sup\{x^*(x) - g(x) \mid x \in \operatorname{dom} g\},$$

where $\operatorname{dom} g$ denotes the domain of g. The epigraph of g is defined by

$$\operatorname{epi} g = \{(x, r) \in X \times \mathbb{R} \mid x \in \operatorname{dom} g, g(x) \leq r\},$$

and the subdifferential of g at $x \in \operatorname{dom} g$ is defined by

$$\partial g(x) = \{x^* \in X^* \mid g(y) - g(x) \geq x^*(y - x) \text{ for all } y \in \operatorname{dom} g\}.$$

When the functions f_t, $t \in T$, are convex, we associate with the system $\sigma := \{f_t(x) \leq 0 \text{ for all } t \in T, x \in C\}$, the cone K defined by

$$K = \operatorname{cone}\{\bigcup_{t\in T} \operatorname{epi} f_t^*\} + \operatorname{epi} \delta_C^*,$$

where $\operatorname{cone}\{\cdot\}$ denotes the convex cone generated by the set $\{\cdot\}$.

Definition 2.1 (*Refs.* $12 - 14$). *The system σ is a Farkas Minkowski system ("σ is FM" in short) if the cone K is weak* closed.*

When problem (P) is convex and σ is FM, it follows from Theorem 1 and Theorem 3 of Ref. 13, that x is an optimal solution to (P) if and only if there exists $\lambda \in \mathbb{R}_+^{(T)}$ such that the optimality condition (5) is satisfied.

Finally in order to obtain sufficient conditions, we have to introduce the following definition:

Definition 2.2 *Let $\varepsilon > 0$ and let A be a closed subset of X. The function g is said to be ε-semiconvex at $x \in A$ with respect to A if:*

(a) g is locally Lipschitz at x,

(b) g is quasidifferentiable at x, and

(c) If $x + d \in A$ and $g'(x; d) + \sqrt{\varepsilon}\|d\| \geq 0$, then $g(x + d) + \sqrt{\varepsilon}\|d\| \geq g(x)$.

Remark 2.1 Let us mention that a convex function on X is ε-semiconvex with respect to X for any $\varepsilon \geq 0$ (Refs. 19, 29). Furthermore, when $\varepsilon = 0$, this concept coincides with the semiconvexity defined by Mifflin in Ref. 24.

3 Approximate Solutions

First we consider the abstract problem

$$(P) \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & x \in A, \end{cases}$$

where $f : X \to \mathbb{R}$ is locally Lipschitz and A is a nonempty closed subset of X. For this problem we define three kinds of approximate solutions (see Ref. 19):

Definition 3.1 *Let $\varepsilon > 0$, a point $x_\varepsilon \in A$ is said to be an*

(a) *ε-solution for (P) if $f(x_\varepsilon) \leq f(x) + \varepsilon$ for all $x \in A$,*

(b) *ε-quasisolution for (P) if $f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\,\|x - x_\varepsilon\|$ for all $x \in A$,*

(c) *regular ε-solution for (P) if it is an ε-solution and an ε-quasisolution for (P).*

Remark 3.1 If x_ε is an ε-quasisolution for (P), then there exists a ball B around x_ε with radius equal to $\sqrt{\varepsilon}$ such that $f(x_\varepsilon) \leq f(x) + \varepsilon$ for all $x \in B \cap A$. In this case, we can say that x_ε is a locally ε-solution for (P).

In the previous definitions of an approximate solution for (P), we have imposed that this solution is feasible. Sometimes this is a too restrictive assumption. Hence, for every $\varepsilon > 0$, we associate with problem (P) the problem (P_ε) defined by

$$(P_\varepsilon) \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & x \in A_\varepsilon, \end{cases}$$

where $A_\varepsilon := \{x \in C \mid f_t(x) \leq \sqrt{\varepsilon} \text{ for all } t \in T\}$. The set A_ε is nonempty and closed. The following definition (Ref. 19) will play an important role in the next sections.

Definition 3.2 *A point $x_\varepsilon \in X$ is said to be an almost regular ε-solution for (P) if x_ε satisfies the following conditions:*

(a) $x_\varepsilon \in A_\varepsilon$,

(b) $f(x_\varepsilon) \leq f(x) + \varepsilon$ *for all* $x \in A$,

(c) $f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\,\|x - x_\varepsilon\|$ *for all* $x \in A$.

A point $x_\varepsilon \in X$ is said to be an almost ε-quasisolution if conditions (a) *and* (c) *hold.*

When T is finite, Loridan (Ref. 19) showed the existence of an almost regular ε-solution for (P) satisfying generalized KKT conditions up to ε.

Theorem 3.1 (Ref. 19, Theorem 5.2). *Let $\varepsilon > 0$ and assume that $T = \{t_1, \ldots, t_m\}$ is finite. Then there exist an almost regular ε-solution x_ε for (P), and real numbers $\lambda_i \geq 0$, $i = 1, \ldots, m$, such that*

(1) $\lambda_i = 0$ *if* $f_{t_i}(x_\varepsilon) \leq 0$,

(2) $\lambda_i > 0$ *for all* $i \in I(\varepsilon) = \{i \,|\, 0 < f_{t_i}(x_\varepsilon) \leq \sqrt{\varepsilon}\}$,

(3) $0 \in \partial f(x_\varepsilon) + \sum_{i \in I(\varepsilon)} \lambda_i \partial f_{t_i}(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*$,

where B^* denotes the closed unit ball in X^*.

4 Generalized KKT Conditions up to ε

In order to define generalized KKT conditions up to ε when T is infinite, we examine the necessary optimality conditions associated with an ε-quasisolution.

Theorem 4.1 *Let $\varepsilon \geq 0$ and let x_ε be an ε-quasisolution for (P). Assume that all the assumptions of Theorem 2.1 are satisfied. If the constraint qualification condition $(\mathcal{U})$ holds and the convex hull of $\{\cup \partial f_t(x_\varepsilon) \,|\, t \in I(x_\varepsilon)\}$ is weak* closed, then there exists $\lambda \in \mathbb{R}_+^{(T)}$ such that*

$$0 \in \partial f(x_\varepsilon) + \sum_{t \in T} \lambda_t \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^* \quad \textit{and} \quad f_t(x_\varepsilon) = 0 \textit{ for all } t \in T(\lambda). \tag{6}$$

Proof. It suffices to apply Proposition 2.1 and to observe that

$$\partial\left(f + \sqrt{\varepsilon}\|\cdot - x_\varepsilon\|\right)(x_\varepsilon) \subset \partial f(x_\varepsilon) + \sqrt{\varepsilon}\, B^*.$$

□

In the particular case where the functions f_t, $t \in T$, are convex, the constraint qualification "σ is FM" is sufficient to obtain condition (6).

A more general condition than (6) can be obtained by replacing in this condition, $f_t(x_\varepsilon) = 0$ by $f_t(x_\varepsilon) \geq 0$.

Definition 4.1 *A pair of vectors $(x_\varepsilon, \lambda) \in A_\varepsilon \times \mathbb{R}_+^{(T)}$ is a generalized KKT pair up to ε if*

$$0 \in \partial f(x_\varepsilon) + \sum_{t \in T} \lambda_t \, \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^* \quad \text{and} \quad f_t(x_\varepsilon) \geq 0 \text{ for all } t \in T(\lambda).$$

The pair is called strict if $f_t(x_\varepsilon) > 0$ for all $t \in T(\lambda)$, which is equivalent to $\lambda_t = 0$ if $f_t(x_\varepsilon) \leq 0$.

This generalized KKT condition being weaker than condition (6), allows us to obtain the existence of a generalized KKT pair up to ε without assuming a constraint qualification condition. It is the subject of the next theorem which can be considered as a generalization of Theorem 3.1.

Theorem 4.2 *Let $\varepsilon > 0$. Assume that the assumptions of Theorem 2.1 are satisfied, and that for every $x \in A_\varepsilon$, the strong closure of the subset $co\{\cup \partial f_t(x) \,|\, t \in T(x)\}$ is weak* closed. (In particular, this assumption is satisfied when X is a reflexive Banach space). Then there exist an almost regular ε-solution x_ε for (P) and $\lambda \in \mathbb{R}_+^{(T)}$ such that (x_ε, λ) is a strict generalized KKT pair up to ε.*

Proof. Let us define $F(x) = \sup_{t\in T} f_t(x)$ for all $x \in X$. Then problem (P) can be written as

$$\min f(x) \quad \text{subject to } F(x) \leq 0 \text{ and } x \in C.$$

Applying Theorem 3.1 to this problem, there exist $\alpha \geq 0$ and an almost regular $\varepsilon/4$–solution x_ε, such that

$$\begin{aligned}
&\alpha = 0 \quad \text{if } F(x_\varepsilon) \leq 0,\\
&\alpha > 0 \quad \text{if } 0 < F(x_\varepsilon) \leq \sqrt{\varepsilon/4},\\
&0 \in \partial f(x_\varepsilon) + \alpha\, \partial F(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon/4}\, B^*.
\end{aligned}$$

It is obvious that x_ε is an almost regular ε-solution for (P).
If $\alpha = 0$, then $F(x_\varepsilon) \leq 0$, and so $f_t(x_\varepsilon) \leq 0$ for all $t \in T$. Now we have

$$0 \in \partial f(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon/4}\, B^*,$$

which entails that $(x_\varepsilon,\, 0) \in A_\varepsilon \times \mathbb{R}_+^{(T)}$ is a strict generalized KKT pair up to ε for (P). If $\alpha > 0$, we make the following reasoning. By the second assumption, the equality

$$\partial F(x_\varepsilon) = \overline{\mathrm{co}}\, \{\cup\, \partial f_t(x_\varepsilon) \,|\, t \in T(x_\varepsilon)\}$$

is also valid when the closure is taken for the strong topology. Then for any $h \in \partial F(x_\varepsilon)$, there must exist $\bar{h} \in \mathrm{co}\, \{\cup\, \partial f_t(x_\varepsilon) \,|\, t \in T(x_\varepsilon)\}$ such that $\|h - \bar{h}\|_* \le \alpha^{-1} \sqrt{\varepsilon/4}$. (Here $\|\cdot\|_*$ denotes the dual norm).

Furthermore, by definition of the convex hull, $\bar{h} = \sum_{t \in T} \bar{\lambda}_t\, g_t$ where $\bar{\lambda} \in \mathbb{R}_+^{(T)}$ with $T(\bar{\lambda}) \subset T(x_\varepsilon)$, $\sum_{t \in T} \bar{\lambda}_t = 1$, and $g_t \in \partial f_t(x_\varepsilon)$ for all $t \in T(\bar{\lambda})$.

Combining the previous results, and starting from

$$0 \in \partial f(x_\varepsilon) + \alpha \partial F(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon/4}\, B^*,$$

we deduce the existence of $\bar{\lambda} \in \mathbb{R}_+^{(T)}$, with $T(\bar{\lambda}) \subset T(x_\varepsilon)$, and such that

$$\begin{aligned} 0 &\in \partial f(x_\varepsilon) + \textstyle\sum_{t \in T} \alpha \bar{\lambda}_t\, \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + \alpha(h - \bar{h}) + \sqrt{\varepsilon/4}\, B^* \\ &\subset \partial f(x_\varepsilon) + \textstyle\sum_{t \in T} \alpha \bar{\lambda}_t\, \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + 2\sqrt{\varepsilon/4}\, B^*. \end{aligned}$$

Setting $\lambda := \alpha \bar{\lambda}$, we can write

$$0 \in \partial f(x_\varepsilon) + \sum_{t \in T} \lambda_t\, \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*.$$

If $f_{t_0}(x_\varepsilon) \le 0$, since $F(x_\varepsilon) > 0$, we can conclude that $t_0 \notin T(x_\varepsilon)$ and, consequently, that $t_0 \notin T(\bar{\lambda}) = T(\lambda)$. In other words, $\lambda_{t_0} = 0$, and $(x_\varepsilon,\, \lambda) \in A_\varepsilon \times \mathbb{R}_+^{(T)}$ is a strict generalized KKT pair up to ε for (P). $\square$

When the Banach space X is reflexive, it is well known that the weak* topology coincides with the weak topology on the dual space X^*. So it is very easy to deduce that the second assumption of Theorem 4.2 is satisfied in that situation.

Theorem 4.2 is a generalization of Theorem 3.1 (due to Loridan) to the case of an infinite number of locally Lipschitz constraints. The existence of a strict generalized KKT pair $(x_\varepsilon,\, \lambda)$ is obtained satisfying $\lambda_t = 0$ when $f_t(x_\varepsilon) \le 0$ and $\lambda_t > 0$ for all t belonging to a finite subset of $\{t \in T \,|\, f_t(x_\varepsilon) > 0\}$ (see Theorem 3.1 (1), (2)).

Conversely, we can derive the following sufficient condition for an almost ε-quasi solution for (P).

Theorem 4.3 *Assume that C is convex and that the functions f_t, $t \in T$, are convex. Let $\varepsilon \ge 0$ and let $(x_\varepsilon, \lambda) \in A_\varepsilon \times \mathbb{R}_+^{(T)}$ be a generalized KKT pair up to ε. If f is ε-semiconvex at x_ε with respect to C, then*

$$f(x_\varepsilon) \le f(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\| \text{ for all } x \in C \text{ such that } f_t(x) \le f_t(x_\varepsilon)\ \forall t \in T(\lambda).$$

In particular, x_ε is an almost ε-quasisolution for (P).

Proof. Let $\varepsilon \geq 0$, let $(x_\varepsilon, \lambda) \in A_\varepsilon \times \mathbb{R}_+^{(T)}$ be a generalized KKT pair up to ε, and let $x \in C$ such that $f_t(x) \leq f_t(x_\varepsilon)$ for all $t \in T(\lambda)$. Then $f_t(x_\varepsilon) \geq 0$ for all $t \in T(\lambda)$, and there exist $u \in \partial f(x_\varepsilon)$, $u_t \in \partial f_t(x_\varepsilon)$, $t \in T(\lambda)$, $w \in N_C(x_\varepsilon)$ and $v \in B^*$ such that

$$0 = u + \sum_{t \in T} \lambda_t u_t + w + \sqrt{\varepsilon}\, v. \tag{7}$$

Furthermore, by definition of $\partial f_t(x_\varepsilon)$, $N_C(x_\varepsilon)$ and B^*, we have the following properties for all $t \in T(\lambda)$

$$\begin{aligned} &u_t(x - x_\varepsilon) \leq f_t(x) - f_t(x_\varepsilon) \leq 0, \\ &w(x - x_\varepsilon) \leq 0 \text{ and } v(x - x_\varepsilon) \leq \|x - x_\varepsilon\|. \end{aligned}$$

Combining these properties with (7), we obtain

$$\begin{aligned} u(x - x_\epsilon) + \sqrt{\varepsilon}\, \|x - x_\epsilon\| &\geq u(x - x_\varepsilon) + \sqrt{\varepsilon}\, v(x - x_\varepsilon) \\ &= -\sum_{t \in T} \lambda_t u_t(x - x_\varepsilon) - w(x - x_\varepsilon) \geq 0. \end{aligned} \tag{8}$$

On the other hand, since f is quasidifferentiable at x_ε, the property $u \in \partial f(x_\varepsilon)$ implies that $f'(x_\varepsilon; x - x_\varepsilon) \geq u(x - x_\varepsilon)$ for all $x \in C$. So from (8), we can deduce that

$$f'(x_\varepsilon; x - x_\varepsilon) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\| \geq 0.$$

But this implies, by definition of an ε-semiconvex function with respect to C, that

$$f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\|.$$

Finally, since A is contained in the set of points $x \in C$ such that $f_t(x) \leq f_t(x_\varepsilon)$ for all $t \in T(\lambda)$, the vector x_ε is also an almost ε-quasisolution for (P). □

When $\varepsilon = 0$, an (almost) ε-quasisolution for (P) is in fact a solution for (P). In that case, Theorems 4.1 and 4.3 allow us to find again the optimality results concerning semiconvex programs presented in Ref. 15, Theorem 5.1.

As a first example of an infinite programming problem of the form (P), we consider the following optimization problem

$$(P_1) \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & g(x) \in -K, \\ & x \in C, \end{cases}$$

where f and C are as above, g is a mapping from X into a separated locally convex topological vector space Y, and K is a closed convex cone in Y. Here we assume that g is continuous and K-convex in the sense that

$$g\big(\xi x + (1 - \xi) y\big) - \xi g(x) - (1 - \xi) g(y) \in -K$$

for every $x, y \in X$ and every $\xi \in [0, 1]$. Note that, in this case, the feasible set of (P_1) is a closed convex subset of X. When f is convex, problem (P_1) has been studied in particular in Refs. 30 − 32. We also assume that f is bounded from below on the feasible set of (P_1).

For writing problem (P_1) under the form of (P), we have to use the dual cone of K, denoted K^+ and defined by

$$K^+ := \{y \in Y^* \mid y(k) \geq 0 \text{ for all } k \in K\}.$$

This set is weak* closed and it is easy to see that

$$g(x) \in -K \quad \Leftrightarrow \quad (\lambda g)(x) \leq 0 \text{ for all } \lambda \in K^+,$$

where $(\lambda g)(x)$ stands for $\lambda(g(x))$. So problem (P_1) is equivalent to the infinite programming problem

$$(P_2) \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & (\lambda g)(x) \leq 0 \text{ for all } \lambda \in K^+, \\ & x \in C. \end{cases}$$

Since $\lambda g : X \to \mathbb{R}$ is a convex function for all $\lambda \in K^+$, problem (P_2) has the same structure as (P), and we can apply the previous theory to (P_2) to derive results for (P_1). This is the aim of the following propositions. However in order to use Theorem 4.2 for problem (P_2) we need to have a compact set of indices, which is not the case with the cone K^+. But it is very easy to see that if Y is a normed space and S^* denotes the closed unit ball of Y^*, then the set $K^+ \cap S^*$ is weak* compact and

$$g(x) \in -K \quad \Leftrightarrow \quad (\lambda g)(x) \leq 0 \text{ for all } \lambda \in K^+ \cap S^*.$$

So in (P_2) we can replace K^+ by the weak* compact set $K^+ \cap S^*$. We denote by (P_3) the corresponding problem. Finally we say that x_ε is an almost regular ε-solution for (P_1) if it is an almost regular ε-solution for (P_3), i.e., if for all $\lambda \in K^+ \cap S^*$ and all $x \in C$ such that $g(x) \in -K$, one has

$$(\lambda g)(x_\varepsilon) \leq \sqrt{\varepsilon}, \quad f(x_\varepsilon) \leq f(x) + \varepsilon, \quad \text{and} \quad f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\|.$$

Proposition 4.1 *Assume that X is a reflexive Banach space and Y is a normed space. Let also $\varepsilon > 0$. Then there exist an almost regular ε-solution x_ε for (P_1) and $\bar{\lambda} \in K^+$ such that $(x_\varepsilon, \bar{\lambda})$ is a strict generalized KKT pair up to ε, i.e.,*

$$0 \in \partial f(x_\varepsilon) + \partial(\bar{\lambda} g)(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^* \quad \text{and} \quad (\bar{\lambda} g)(x_\varepsilon) > 0. \tag{9}$$

Proof. Since $K^+ \cap S^*$ is weak* compact and the mapping $\lambda \mapsto (\lambda g)(x)$ from Y^* to $\mathbb{R}$ is weak* continuous for each x, it follows from Theorem 4.2 applied to problem (P_3), that there exist an almost regular ε-solution x_ε for (P_3) and a vector $\mu \in \mathbb{R}_+^{(K^+ \cap S^*)}$ such that (x_ε, μ) is a strict generalized KKT pair up to ε, i.e.,

$$0 \in \partial f(x_\varepsilon) + \sum_{\lambda \in T} \mu_\lambda \partial(\lambda g)(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*, \; (\lambda g)(x_\varepsilon) > 0 \text{ for all } \lambda \in T(\mu), \tag{10}$$

where $T(\mu) := \{\lambda \in K^+ \cap S^* \mid \mu_\lambda > 0\}$.
Furthermore, for each $\lambda \in K^+$, the function λg is continuous and convex. So we have

$$\sum_{\lambda \in T} \mu_\lambda \partial(\lambda g)(x_\varepsilon) = \partial \left(\sum_{\lambda \in T} \mu_\lambda \lambda \right) g(x_\varepsilon). \tag{11}$$

Then $\bar{\lambda} = (\sum_{\lambda \in T} \mu_\lambda \lambda)$ is suitable. Indeed $\bar{\lambda} \in K^+$ because $T(\mu)$ is finite and K^+ is a convex cone. Furthermore, since $(\bar{\lambda} g)(x_\varepsilon) \geq 0$, it follows from (10) and (11) that (9) holds. □

The next proposition gives sufficient conditions to obtain an ε-quasisolution for (P_1), i.e., a point x_ε feasible for (P_1) and such that for all x feasible for (P_1), $f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\|$. The proof is omitted because it is similar to the one of Theorem 4.3. However, let us note that here we do not have to assume that X is reflexive and that the index set is compact. So Y can be any separated locally convex topological vector space.

Proposition 4.2 *Let $\varepsilon \geq 0$ and let $(x_\varepsilon, \bar{\lambda}) \in C \times K^+$ such that $g(x_\varepsilon) \in -K$ and*

$$0 \in \partial f(x_\varepsilon) + \partial (\bar{\lambda} g)(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^* \quad \text{and} \quad \bar{\lambda} g(x_\varepsilon) \geq 0.$$

If f is quasidifferentiable and ε-semiconvex at x_ε with respect to C, then x_ε is an ε-quasisolution for (P_1). □

These results (Propositions 4.1 and 4.2) are related to Theorem 4.1 of Ref. 31, where f is convex and $\varepsilon = 0$.

A very important instance of problem (P_1) is the following semi-definite programming problem:

$$(SDP) \begin{cases} \text{Minimize} & f(x) \\ \text{subject to} & G(x) \preceq 0, \\ & x \in C, \end{cases}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz continuous function, C is a nonempty closed convex subset of $\mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathcal{S}^p$ is a continuous mapping. Here $\mathcal{S}^p$ denotes the space of $p \times p$ symmetric matrices and the notation $A \preceq 0$ ($A \succeq 0$) means that the matrix A is negative semidefinite (positive semidefinite). We assume that the space $\mathcal{S}^p$ is equipped with the scalar product $A \bullet B = \text{trace}\,(AB)$. In order to establish the relationship existing between problems (SDP) and (P_1), we note that the constraint $G(x) \preceq 0$ is equivalent to $G(x) \in -K$ where $K = \mathcal{S}^p_+$, the convex cone of positive semidefinite $p \times p$ symmetric matrices, and that the dual cone K^+ coincides with $\mathcal{S}^p_+$. With these notations, for any $\Lambda \in K^+$, the function ΛG becomes a function from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$(\Lambda G)(x) = \Lambda \bullet G(x) \qquad \forall x \in \mathbb{R}.$$

If the constraint mapping is affine, i.e., $G(x) = F_0 + \sum_{i=1}^n F_i x_i$ where $F_0, F_1, \ldots, F_n \in \mathcal{S}^p$ are $n+1$ given matrices, then the subdifferential $\partial(\Lambda G)(x)$ is equal to the vector $(\Lambda \bullet F) = (\Lambda \bullet F_1, \ldots, \Lambda \bullet F_n)$. In that case Propositions 4.1 and 4.2 become:

Proposition 4.3 *Let $\varepsilon > 0$. Then there exist an almost regular ε-solution x_ε for (SDP) and a matrix $\bar{\Lambda} \in \mathcal{S}_+^p$ such that $(x_\varepsilon, \bar{\Lambda})$ is a strict generalized KKT pair up to ε, i.e.,*

$$0 \in \partial f(x_\varepsilon) + \bar{\Lambda} \bullet F + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*, \quad \bar{\Lambda} \bullet \left[F_0 + \sum_{i=1}^{n} F_i x_{\varepsilon,i} \right] \geq 0, \qquad (12)$$

where $x_{\varepsilon,i}$, $i = 1, \cdots, n$, denote the components of x_ε.
Conversely, let x_ε be feasible for (SDP) and $\bar{\Lambda} \in \mathcal{S}_+^p$ such that (12) is satisfied. If f is quasidifferentiable and ε-semiconvex at x_ε with respect to C, then x_ε is an ε-quasisolution for (SDP).

5 Quasi Saddle-Points and ε-Lagrangian Duality

In this section we associate with Problem (P) the following Lagrangian function:

$$L(x, \lambda) = f(x) + \sum_{t \in T} \lambda_t f_t(x) \text{ for all } (x, \lambda) \in X \times \mathbb{R}_+^{(T)}.$$

For every $\lambda \in \mathbb{R}_+^{(T)}$, the function $L(\cdot, \lambda)$ is locally Lipschitz on X. For problems (P_1) and (SDP), the corresponding Lagrangian functions become

$$L(x, \lambda) = f(x) + (\lambda g)(x) \quad \text{for all } (x, \lambda) \in X \times K^+$$

and

$$L(x, \Lambda) = f(x) + \Lambda \bullet G(x) \quad \text{for all } (x, \Lambda) \in \mathbb{R}^n \times \mathcal{S}_+^p,$$

respectively.

When the number of constraints is finite (i.e., when T is finite), Loridan introduced in Ref. 19 an ε-Lagrangian function and the corresponding concept of saddle-point. Here our aim is to propose an extension of these notions to Problem (P) with an infinite number of constraints.

Let $\varepsilon \geq 0$. The ε-Lagrangian function associated with Problem (P) is a function L_ε defined on $X \times \mathbb{R}_+^{(T)} \times X \times \mathbb{R}_+^{(T)}$ by setting for all $(x, \lambda, z, \mu) \in X \times \mathbb{R}_+^{(T)} \times X \times \mathbb{R}_+^{(T)}$

$$L_\varepsilon(x, \lambda, z, \mu) = L(x, \lambda) + \sqrt{\varepsilon}\, \|x - z\| - \sqrt{\varepsilon}\, \|\lambda - \mu\|_1,$$

where $\|\lambda - \mu\|_1 = \sum_{t \in T} |\lambda_t - \mu_t|$. The corresponding quasi saddle-point is defined as follows.

Definition 5.1 *The point $(\bar{x}, \bar{\lambda}) \in C \times \mathbb{R}_+^{(T)}$ is a quasi saddle-point for the ε-Lagrangian L_ε if*

$$L_\varepsilon(\bar{x}, \lambda, \bar{x}, \bar{\lambda}) \leq L_\varepsilon(\bar{x}, \bar{\lambda}, \bar{x}, \bar{\lambda}) \leq L_\varepsilon(x, \bar{\lambda}, \bar{x}, \bar{\lambda}) \text{ for all } (x, \lambda) \in C \times \mathbb{R}_+^{(T)}.$$

Equivalently:

$$L(\bar{x}, \lambda) - \sqrt{\varepsilon}\, \|\lambda - \bar{\lambda}\|_1 \leq L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda}) + \sqrt{\varepsilon}\, \|x - \bar{x}\| \text{ for all } (x, \lambda) \in C \times \mathbb{R}_+^{(T)}.$$

For problem (P_1) with Y a normed space, the left-hand side of the previous inequality can be written as

$$L(\bar{x}, \lambda) - \sqrt{\varepsilon}\, \|\lambda - \bar{\lambda}\|_* \leq L(\bar{x}, \bar{\lambda}) \quad \text{ for all } \lambda \in K^+,$$

where the norm $\|\cdot\|_*$ is taken in the dual space Y^*. For problem (SDP) this inequality becomes

$$L(\bar{x}, \Lambda) - \sqrt{\varepsilon}\, \|\Lambda - \bar{\Lambda}\| \leq L(\bar{x}, \bar{\Lambda}) \quad \text{ for all } \Lambda \in \mathcal{S}^p_+,$$

where the norm $\|\cdot\|$ is taken in the space of symmetric matrices $\mathcal{S}^p$.

In the following two theorems we study the relationships between generalized KKT pairs and quasi saddle-points for problem (P). They are extensions of Theorems 6.1 and 6.2 from Ref. 19 to the infinite programming problem (P). However, in order to present these results, we need the following lemma.

Lemma 5.1 *Let $z \in C$ and $\lambda \in \mathbb{R}^{(T)}_+$. Assume that C is convex, the functions f and f_t, $t \in T$ are quasidifferentiable at z and $L(\cdot, \lambda)$ is ε-semiconvex at z with respect to C. If the following condition holds*

$$0 \in \partial f(z) + \sum_{t \in T} \lambda_t\, \partial f_t(z) + N_C(z) + \sqrt{\varepsilon}\, B^*,$$

then $L(x, \lambda) + \sqrt{\varepsilon}\, \|x - z\| \geq L(z, \lambda)$ for all $x \in C$.

Proof. By assumption there exist $u \in \partial f(z)$, $u_t \in \partial f_t(z)$, for all $t \in T(\lambda)$, $w \in N_C(z)$ and $v \in B^*$ such that

$$0 = u + \sum_{t \in T} \lambda_t\, u_t + w + \sqrt{\varepsilon}\, v.$$

Then for all $x \in C$, we have, f and f_t, $t \in T$, being quasidifferentiable at z, and C being convex, the following properties:

$$\begin{aligned}
&f'(z\,;\, x - z) = f^\circ(z\,;\, x - z) \geq u(x - z),\\
&f'_t(z\,;\, x - z) = f^\circ_t(z\,;\, x - z) \geq u_t(x - z) \quad \text{ for all } t \in T(\lambda),\\
&w(x - z) \leq 0.
\end{aligned}$$

So, for all $x \in C$ we obtain that

$$f'(z\,;\, x-z) + \sum_{t \in T} \lambda_t f'_t(z\,;\, x-z) \geq (u + \sum_{t \in T} \lambda_t u_t + w)(x-z) = -\sqrt{\varepsilon}\, v(x-z) \geq -\sqrt{\varepsilon}\, \|x-z\|,$$

i.e.,

$$L(\cdot, \lambda)'(z\,;\, x - z) \geq -\sqrt{\varepsilon}\, \|x - z\| \quad \text{ for all } x \in C.$$

Since $L(\cdot, \lambda)$ is ε-semiconvex at z with respect to C, we deduce from the previous inequality that

$$L(x, \lambda) + \sqrt{\varepsilon}\, \|x - z\| \geq L(z, \lambda) \quad \text{ for all } x \in C,$$

i.e., what we have to prove. □

Theorem 5.1 *Let $x_\varepsilon \in A_\varepsilon$ and $\bar\lambda \in \mathbb{R}_+^{(T)}$ such that $(x_\varepsilon, \bar\lambda)$ is a strict generalized KKT pair up to ε. Then*

(a) $L(x_\varepsilon, \bar\lambda) - L(x_\varepsilon, \lambda) \geq -\sqrt{\varepsilon}\, \|\bar\lambda - \lambda\|_1$ *for all $\lambda \in \mathbb{R}_+^{(T)}$;*

(b) *If C is convex, if the functions f and f_t, $t \in T$, are quasidifferentiable at x_ε and if $L(\cdot, \bar\lambda)$ is ε-semiconvex at x_ε with respect to C, then $(x_\varepsilon, \bar\lambda)$ is a quasi saddle-point for the ε-Lagrangian L_ε on $C \times \mathbb{R}_+^{(T)}$.*

Proof. First we prove that for all $\lambda \in \mathbb{R}_+^{(T)}$, we have (a), i.e.,

$$\sum_{t \in T} (\bar\lambda_t - \lambda_t)\, f_t(z_\varepsilon) \geq -\sqrt{\varepsilon} \sum_{t \in T} |\bar\lambda_t - \lambda_t|. \tag{13}$$

Let $t \in T$. If $f_t(x_\varepsilon) \leq 0$, then by assumption $\bar\lambda_t = 0$ and

$$(\bar\lambda_t - \lambda_t)\, f_t(x_\varepsilon) = -\lambda_t\, f_t(x_\varepsilon) \geq 0. \tag{14}$$

If $0 < f_t(x_\varepsilon) \leq \sqrt{\varepsilon}$, then

$$(\bar\lambda_t - \lambda_t)\, f_t(x_\varepsilon) \geq -|\bar\lambda_t - \lambda_t|\, f_t(x_\varepsilon) \geq -|\bar\lambda_t - \lambda_t|\, \sqrt{\varepsilon}. \tag{15}$$

So (13) directly follows from (14) and (15) and the first part is proven.
To prove the second part, Lemma 5.1 is applied with $z = x_\varepsilon$ and $\lambda = \bar\lambda$ to give

$$L(x, \bar\lambda) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\| \geq L(x_\varepsilon, \bar\lambda) \quad \text{for all } x \in C.$$

But this inequality and the one proven in the first part imply that $(z_\varepsilon, \bar\lambda)$ is a quasi saddle-point. □

Since a convex function is quasidifferentiable and ε-semiconvex, the next corollary is a direct consequence of Theorem 5.1.

Corollary 5.1 *If C is convex and if the functions f and f_t, $t \in T$, are convex, then any strict generalized KKT pair up to ε is a quasi saddle-point for the ε-Lagrangian on $C \times \mathbb{R}_+^{(T)}$.*

Conversely we do not obtain that any quasi saddle-point $(z_\varepsilon, \bar\lambda)$ is a generalized KKT pair because the condition $f_t(z_\varepsilon) \geq 0$ is not necessarily satisfied for each $t \in T(\bar\lambda)$. This fact was already observed by Loridan (Ref. 19) in the case of a finite number of constraints. However we obtain the following result.

Theorem 5.2 *If $(x_\varepsilon, \bar\lambda)$ is a quasi saddle-point for the ε-Lagrangian L_ε on $C \times \mathbb{R}_+^{(T)}$, then x_ε satisfies the following statements:*

(a) $f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\|$ *for all $x \in C$ verifying $f_t(x) \leq f_t(x_\varepsilon)$ for all $t \in T(\bar\lambda)$;*

(*b*) $f_t(x_\varepsilon) \le \sqrt{\varepsilon}$ *for all* $t \in T$*;*

(*c*) $0 \in \partial f(x_\varepsilon) + \sum_{t\in T} \bar{\lambda}_t \, \partial f_t(x_\varepsilon) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*$*;*

(*d*) $\bar{\lambda}_t > 0$ *implies* $-\sqrt{\varepsilon} \le f_t(x_\varepsilon)$*.*

Proof. From the definition of a quasi saddle-point, we have

$$f(x_\varepsilon) + \sum_{t\in T} \bar{\lambda}_t f_t(x_\varepsilon) \le f(x) + \sum_{t\in T} \bar{\lambda}_t f_t(x) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\| \quad \text{for all } x \in C.$$

Hence, for all $t \in T(\bar{\lambda})$ and for all $x \in C$ verifying $f_t(x) \le f_t(x_\varepsilon)$, we obtain (a).
In order to prove (b), let $t_0 \in T$ (arbitrary) and let us consider the vector $\lambda = (\lambda_t)_{t\in T}$ defined by

$$\lambda_{t_0} = 1 + \bar{\lambda}_{t_0} \quad \text{and} \quad \lambda_t = \bar{\lambda}_t \quad \text{if } t \ne t_0.$$

It is clear that $\lambda \in \mathbb{R}_+^{(T)}$ and $t_0 \in T(\lambda)$. Moreover

$$T(\lambda) = T(\bar{\lambda}) \quad \text{if } \bar{\lambda}_{t_0} > 0 \quad \text{and} \quad T(\lambda) = T(\bar{\lambda}) \cup \{t_0\} \quad \text{if } \bar{\lambda}_{t_0} = 0.$$

Since $(x_\varepsilon, \bar{\lambda})$ is a quasi saddle-point of L_ε, we obtain the following inequality

$$f(x_\varepsilon) + \sum_{t\in T} \lambda_t f_t(x_\varepsilon) - \sqrt{\varepsilon} \sum_{t\in T} |\bar{\lambda}_t - \lambda_t| \le f(x_\varepsilon) + \sum_{t\in T} \bar{\lambda}_t f_t(x_\varepsilon). \tag{16}$$

Hence $f_{t_0}(x_\varepsilon) = \sum_{t\in T} (\lambda_t - \bar{\lambda}_t)\, f_t(x_\varepsilon) \le \sqrt{\varepsilon}\, \sum_{t\in T} |\lambda_t - \bar{\lambda}_t| = \sqrt{\varepsilon}$ and (b) is obtained.
In order to prove (c), we consider the second inequality defining the quasi saddle-point $(x_\varepsilon, \bar{\lambda})$

$$L(x_\varepsilon, \bar{\lambda}) \le L(x, \bar{\lambda}) + \sqrt{\varepsilon}\, \|x - x_\varepsilon\| \quad \text{for all } x \in C,$$

which gives rise to the necessary optimality condition

$$0 \in \partial_x L(x_\varepsilon, \bar{\lambda}) + N_C(x_\varepsilon) + \sqrt{\varepsilon}\, B^*,$$

where $\partial_x L(x_\varepsilon, \bar{\lambda})$ denotes the generalized gradient of $L(\cdot, \bar{\lambda})$ at x_ε. But this is nothing else than (c). Finally, to prove (d), assume that $\bar{\lambda}_t > 0$ for some $t \in T(\bar{\lambda})$ and consider a vector $\lambda = (\lambda_u)_{u\in T}$ such that

$$0 < \lambda_t < \bar{\lambda}_t \quad \text{and} \quad \lambda_u = \bar{\lambda}_u \quad \text{if } u \ne t.$$

Then $\lambda \in \mathbb{R}_+^{(T)}$ and $T(\lambda) = T(\bar{\lambda})$. Hence, it follows from (16) that

$$\lambda_t f_t(x_\varepsilon) - \sqrt{\varepsilon}(\bar{\lambda}_t - \lambda_t) \le \bar{\lambda}_t f_t(x_\varepsilon),$$

i.e.,$-\sqrt{\varepsilon}\,(\bar{\lambda}_t - \lambda_t) \le (\bar{\lambda}_t - \lambda_t)\, f_t(x_\varepsilon)$. Since $\bar{\lambda}_t - \lambda_t > 0$, we obtain $-\sqrt{\varepsilon} \le f_t(x_\varepsilon)$ and the proof is complete. □

In the last part of this section we associate with problem (P) the following dual problem of Wolfe-type:

$$(D) \begin{cases} \text{Maximize} & L(y, \lambda) \\ \text{s.t.} & 0 \in \partial f(y) + \sum_{t\in T} \lambda_t \, \partial f_t(y) + N_C(y) + \sqrt{\varepsilon}\, B^* \\ & y \in C, \ \lambda \in \mathbb{R}_+^{(T)}. \end{cases}$$

First a weak duality result is given in the next proposition.

Proposition 5.1 *Let x and (y, λ) be feasible points for problems (P) and (D) respectively. If C is convex, if the functions f and f_t, $t \in T$ are quasidifferentiable at y and if $L(\cdot, \lambda)$ is ε-semiconvex on C, then*

$$f(x) - L(y, \lambda) \geq -\sqrt{\varepsilon}\, \|x - y\|.$$

Proof. Applying Lemma 5.1 to the dual feasible point (y, λ), we obtain the following inequality

$$L(x, \lambda) - L(y, \lambda) \geq -\sqrt{\varepsilon}\, \|x - y\| \quad \text{for all } x \in C. \tag{17}$$

On the other hand, x being feasible for (P), we have that

$$L(x, \lambda) = f(x) + \sum_{t \in T(\lambda)} \lambda_t\, f_t(x) \leq f(x). \tag{18}$$

Combining (17) and (18), we deduce the required inequality. □

Proposition 5.2 *Let $(x_\varepsilon, \bar{\lambda})$ be a quasi saddle-point for the ε-Lagrangian L_ε on $C \times \mathbb{R}_+^{(T)}$. If C is convex, if the functions f and f_t, $t \in T$ are quasidifferentiable on C, and if $L(\cdot, \lambda)$ is ε-semiconvex on C for each $\lambda \in \mathbb{R}_+^{(T)}$, then $(x_\varepsilon, \bar{\lambda})$ is an ε-quasisolution for (D).*

Proof. First observe that by Theorem 5.2 (c), the pair $(x_\varepsilon, \bar{\lambda})$ is feasible for the dual problem (D). Then let (y, λ) be an arbitrary feasible point for (D). We have to prove that

$$L(x_\varepsilon, \bar{\lambda}) \geq L(y, \lambda) - \sqrt{\varepsilon}\, \|x_\varepsilon - y\| - \sqrt{\varepsilon}\, \|\bar{\lambda} - \lambda\|_1. \tag{19}$$

Since $y \in C$, $\lambda \in \mathbb{R}_+^{(T)}$ and $0 \in \partial f(y) + \sum_{t \in T(\lambda)} \lambda_t\, \partial f_t(y) + N_C(y) + \sqrt{\varepsilon}\, B^*$, it follows from Lemma 5.1 that

$$L(x_\varepsilon, \lambda) + \sqrt{\varepsilon}\, \|x_\varepsilon - y\| \geq L(y, \lambda). \tag{20}$$

On the other hand, by definition of a quasi saddle-point, we have

$$L(x_\varepsilon, \bar{\lambda}) \geq L(x_\varepsilon, \lambda) - \sqrt{\varepsilon}\, \|\bar{\lambda} - \lambda\|_1. \tag{21}$$

Hence, combining (20) and (21), we immediately derive (19). □

Using Theorem 5.1, we obtain the following corollary.

Corollary 5.2 *Let $(x_\varepsilon, \bar{\lambda})$ be a strict generalized KKT pair up to ε. If C is convex, if the functions f and f_t, $t \in T$ are quasidifferentiable on C and if $L(\cdot, \lambda)$ is ε-semiconvex on C for each $\lambda \in \mathbb{R}_+^{(T)}$, then $(x_\varepsilon, \bar{\lambda})$ is an ε-quasisolution for (D).*

References

1. López, M., and Still, G., Semi-infinite Programming, European Journal of Operational Research, Vol. 180, pp. 491-518, 2007.

2. Bonnans, J., and Shapiro, A., Perturbation Analysis of Optimization Problems, Springer Series in Operations Research, Springer-Verlag, New York, 2000.

3. Goberna, M., and López, M., Linear Semi-Infinite Optimization, John Wiley and Sons, Chichester, England, 1998.

4. Polak, E., Optimization. Algorithms and Consistent Approximations, Applied Mathematical Sciences, Springer-Verlag, New York, 1997.

5. Stein, 0., and Still, G., On Optimality Conditions for Generalized Semi-infinite Programming Problems, Journal of Optimization Theory and Applications, Vol. 104, pp. 443-458, 2000.

6. Stein, 0., and Still, G., On Generalized Semi-infinite Optimization and Bilevel Optimization, European Journal of Operational Research, Vol. 142, pp. 444-462, 2002.

7. Stein, G., Solving Generalized Semi-infinite Programs by Reduction to Simpler Problems, Optimization, Vol. 53, pp. 19-38, 2004.

8. Clarke, F.H., Optimization and Nonsmooth Analysis, Wiley-Interscience, New York, 1983.

9. Bihain, A., Optimization of Upper Semidifferentiable Functions, Journal of Optimization Theory and Applications, Vol. 44, pp. 545-568, 1984.

10. Hiriart-Urruty, J.B., and Lemaréchal, C., Convex Analysis and Minimization Algorithms, A Series of Comprehensive Studies in Mathematics, Springer-Verlag, New York, 1993.

11. Zheng, X.Y., and Yang, X., Lagrange Multipliers in Nonsmooth Semi-infinite Optimization Problems, Mathematics of Operations Research, Vol. 32, pp. 168-181, 2007.

12. Dinh, N., Goberna, M.A., and López, M.A., From Linear to Convex Systems: Consistency, Farkas Lemma and Applications, Journal of Convex Analysis, Vol. 13, pp. 113-133, 2006.

13. Dinh, D., Goberna, M.A., López, M.A., and Son, T.Q., New Farkas-type Constraint Qualifications in Convex Infinite Programming, ESAIM: Control, Optimisation and Calculus of Variations, Vol. 13, pp. 580-597, 2007.

14. Goberna, M., Jeyakumar, V., and López, M., Necessary and Sufficient Constraint Qualifications for Solvability of Systems of Infinite Convex Inequalities, Nonlinear Analysis, Vol. 68, pp. 1184-1194, 2008 doi:10.1016/j.na.2006.12.014.

15. Son, T.Q., and Dinh, N., Characterizations of Optimal Solution Sets of Convex Infinite Programs, TOP, 2008, doi 10.1007/s11750-008-0039-2.

16. Jeyakumar, V., Dinh, N., and Lee, G.M., A New Closed Cone Constraint Qualification for Convex Optimization, Applied Mathematical Report AMR04/8, University of New South Wales, Sidney, Australia, 2004.

17. Dutta, J., Necessary Optimality Conditions and Saddle Points for Approximate Optimization in Banach Spaces, TOP, Vol. 13, pp. 127-143, 2005.

18. Hamel, A., An ε-Lagrange Multiplier Rule for a Mathematical Programming Problem on Banach Spaces, Optimization, Vol. 49, pp. 137-149, 2001.

19. Loridan, P., Necessary Conditions for ε-Optimality, Mathematical Programming Study, Vol. 19, pp. 140-152, 1982.

20. Scovel, C., Hush, D., and Steinwart, I., Approximate Duality, Los Alamos National Laboratory, Technical Report La-UR-05-6755, September 2005.

21. Strodiot, J.J., Nguyen, V.H., and Heukemes, N., ε-Optimal Solutions in Nondifferentiable Convex Programming and Some Related Questions, Mathematical Programming, Vol. 25, pp. 307-328, 1983.

22. Yokoyama, K., ε-Optimality Criteria for Convex Programming Problems Via Exact Penalty Functions, Mathematical Programming, Vol. 56, pp. 233-243, 1992.

23. Ekeland, I., On the Variational Principle, Journal of Mathematical Analysis and Applications, Vol. 47, pp. 324-353, 1974.

24. Mifflin, R., Semismooth and Semiconvex Functions in Constrained Optimization, SIAM Journal on Control and Optimization, Vol. 15, pp. 959-972, 1977.

25. Clarke, F.H., Generalized Gradients of Lipschitz Functionals, Advances in Mathematics, Vol. 40, pp. 52-67, 1981.

26. Ioffe, A.D., and Tihomirov, V.M., Theory of Extremal Problems, Studies in Mathematics and Its Applications, North-Holland, Amsterdam, 1979.

27. Valadier, M., Contribution à l'Analyse Convexe, Thèse de Doctorat d'Etat, Université de Montpellier, France, 1970.

28. Hiriart-Urruty, J.B., On Optimality Conditions in Nondifferentiable Programming, Mathematical Programming, Vol. 14, pp. 73-86, 1978.

29. Liu, J.C., ε-Duality Theorem of Nondifferentiable Nonconvex Multiobjective Programming, Journal of Optimization Theory and Applications, Vol. 69, pp. 153-167, 1991.

30. Dinh, N., Jeyakumar, V., and Lee, G.M., Sequential Lagrangian Conditions for Convex Programs with Applications to Semidefinite Programming, Journal of Optimization Theory and Applications, Vol. 125, pp. 85-112, 2005.

31. Jeyakumar, V., Lee, G.M., and Dinh, N., New Sequence Lagrange Multiplier Conditions Characterizing Optimality Without Constraint Qualification for Convex Programs, SIAM Journal on Optimization, Vol. 14, pp. 534-547, 2003.

32. Jeyakumar, V., and Wolkowicz, H., Generalization of Slater's Constraint Qualification for Infinite Convex Programs, Mathematical Programming, Vol. 57, pp. 85-101, 1992.