Published in
Math. Program. 99 (2004), no. 3, Ser. A, pp. 487–498.

DUALITY GAP IN CONVEX PROGRAMMING

T. CHAMPION

ABSTRACT. In this paper, we consider general convex programming problems and give a sufficient condition for the equality of the infimum of the original problem and the supremum of its ordinary dual. This condition may be seen as a continuity assumption on the constraint sets (i.e. on the sublevel sets of the constraint function) under linear perturbation. It allows us to generalize as well as treat previous results in a unified framework. Our main result is in fact based on a quite general constraint qualification result for quasiconvex programs involving a convex objective function proven in the setting of a real normed vector space.

Keywords. Duality gap, convex programming, quasiconvex programming, class $\mathcal{F}$, weakly analytic functions.
MSC 2000. 90C25,90C26,90C30,90C31.

1. Introduction

Let (P_0) be the following ordinary convex program

$$(P_0) \qquad \mathrm{Inf}\,\{f(x) \,:\, f_1(x) \leq 0, \ldots, f_k(x) \leq 0\}$$

where the fonctions $f, f_i : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semi-continuous. In the sequel, we shall assume that $\mathrm{Inf}(P_0) < +\infty$, which means that the constraint set intersects the domain of f. It is classical to associate with problem (P_0) the following dual convex problem

$$(D) \qquad \mathrm{Sup}\left\{ \inf\{f(x) + \sum_{i=1}^{k} \lambda_i f_i(x) \,:\, x \in \mathbb{R}^N\} \,:\, \lambda_1 \geq 0, \ldots, \lambda_k \geq 0 \right\}.$$

Then it can be checked that $\mathrm{Inf}(P_0) \geq \mathrm{Sup}(D)$, and when $\mathrm{Inf}(P_0) \in \mathbb{R}$ the difference $\delta := \mathrm{Inf}(P_0) - \mathrm{Sup}(D)$ between these two optimal values is non-negative. Notice that if $\mathrm{Inf}(P_0) = -\infty$, one obviously has $\mathrm{Sup}(D) = -\infty$ so that we may set $\delta := 0$ as well. When the difference δ is positive, it is said that there is a ***duality gap*** between (P_0) and (D). It is well known that the existence of a duality gap is closely related to the study of the approximated problems (P_ε) given by

$$(P_\varepsilon) \qquad \mathrm{Inf}\,\{f(x) \,:\, f_1(x) \leq \varepsilon, \ldots, f_k(x) \leq \varepsilon\}$$

where ε is a positive parameter. Indeed, if we denote $v : [0, +\infty[\to \overline{\mathbb{R}}$ the value function which associates to any $t \geq 0$ the value $v(t) := \mathrm{Inf}(P_t)$, then v is convex, non-increasing on $[0, +\infty[$ and duality theory yields $\lim\limits_{\varepsilon \to 0, \varepsilon > 0} v(\varepsilon) = \mathrm{Sup}(D)$. As a consequence, the existence of a duality gap between (P_0) and (D) is equivalent to the fact that v is not continuous at 0. We refer to [2], [3] and [8] for more details on duality theory and the link with the continuity of v.

The question of finding conditions on the functions f, f_i which ensure that there is no duality gap (i.e. $\delta = 0$) is not only important for the theoretical understanding of the duality gap phenomenum but also for numerical purposes. Indeed, in numerical lagrangian methods for finding the infimum of (P_0), it is important that there is no duality gap since these methods compute the supremum of (D). In this paper, we shall focus our attention on the duality gap issue without adressing the existence of Lagrange multipliers. We give a new condition, which we denote property $\mathcal{D}$, which ensures that there is no duality gap between (P_0) and (D). In fact, property $\mathcal{D}$ may in some sense be considered as a characterization of the absence of duality gap: indeed theorem 2.6.i yields that there is no duality gap whenever this property is satisfied together with a mild assumption on the objective function f, while theorem 2.6.ii implies that there exists smooth objective functions f for which there is a positive duality gap when property $\mathcal{D}$ is not satisfied. This condition is quite general since the semi-continuity result for the value function v that we prove is in fact obtained in the setting of quasiconvex constraint functions f_i defined on a real normed vector space which may be infinite dimensional (see theorem 2.6). In the setting of convex programming in finite dimensions, this result improves previous works by Auslender [1], Kummer [4], Li [5] and Rockafellar [7]. Our result also allows to recover that the usual Slater constraint qualification as well as the compactness hypothesis on the constraint set are sufficient conditions for there is no duality gap. These issues are discussed more precisely after theorem 4.1 in §4.

The paper is organized as follows: we define property $\mathcal{D}$ in §2 in the general setting of quasiconvex functions over a normed vector space, and we provide some examples and prove the main stability result (theorem 2.6). Section 3 is devoted to the convex setting and finite dimensional setting, in which we can prove finer existence results for functions satisfying property $\mathcal{D}$. Finally the main result on the duality gap phenomenum in convex programming (theorem 4.1) is given and discussed in §4, where the results of the two first sections are commented.

2. The property $\mathcal{D}$ in quasiconvex programming

In the following, E denotes a real normed vector space. We recall that a function $h : E \to \mathbb{R} \cup \{+\infty\}$ is quasiconvex if its sublevels $\{h \leq t\}$ are convex for any $t \in \mathbb{R}$. We denote by $\overline{A}$ the closure of a subset A of E.

Definition 2.1. *Let $h : E \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous proper quasiconvex function. Let t belong to $h(E) \setminus \{+\infty\}$, then h satisfies property $\mathcal{D}(t)$ whenever*

$$\mathcal{D}(t) \qquad \bigcap_{s>t} \overline{\{h \leq s\} + F} \;=\; \overline{\{h \leq t\} + F}$$

for any closed subspace F of E.

In fact, it is easy to check that h satisfies condition $\mathcal{D}(t)$ whenever

$$\forall F \text{ closed subspace of } E \qquad \bigcap_{s>t} \overline{\{h \leq s\} + F} \;\subset\; \overline{\{h \leq t\} + F}.$$

For example, any l.s.c. quasiconvex function $h : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ satisfies property $\mathcal{D}(t)$ for all $t \in h(\mathbb{R}) \setminus \{+\infty\}$: indeed, the only test subspaces to be considered are $\{0\}$ and $\mathbb{R}$. Other simple examples are the distance functions and indicator functions of closed convex subsets of E, as well as continuous affine forms over E:

Lemma 2.2. *Let $C \subset E$ be nonempty, closed and convex, then the convex functions $x \mapsto \delta_C(x)$ and $x \mapsto d(x, C)$ satisfy property $\mathcal{D}(t)$ for any non-negative t.*

Proof. We recall that the indicator function $x \mapsto \delta_C(x)$ is given by $\delta_C(x) = 0$ for x in C and $\delta_C(x) = +\infty$ otherwise. Then since it only takes the two values 0 and $+\infty$, one has

$$\forall t \in [0, +\infty[\qquad \overline{\{\delta_C \leq t\} + F} \;=\; \overline{\{\delta_C \leq 0\} + F}$$

for any closed subspace F of E.

On the other hand, let F be a closed subspace of E, let t belong to $[0, +\infty[$ and assume that x belongs to $\bigcap_{s>t} \overline{\{d(., C) \leq s\} + F}$. Then there exist two sequences $(x_n)_n$ and $(\xi_n)_n$ such that

$d(x_n, C) \leq t + \frac{1}{n}$ and $\xi_n \in F$ for any n as well as $x_n + \xi_n \to x$. Now for any $n \in \mathbb{N}$ there exists y_n such that $d(y_n, C) \leq t$ as well as $\|x_n - y_n\| \leq \frac{2}{n}$. Since $\|(x_n + \xi_n) - (y_n + \xi_n)\| \to 0$, we infer $y_n + \xi_n \to x$ and thus x belongs to $\overline{\{d(., C) \leq t\} + F}$. □

Lemma 2.3. *Let $h : E \to \mathbb{R}$ be a non-constant continuous affine form, then h satisfies $\mathcal{D}(t)$ for any t in $\mathbb{R}$.*

Proof. Let F be a closed subspace of E. Then either F is included in the kernel of the linear part of h, in which case $\{h \leq t\} + F = \{h \leq t\}$ for any $t \in \mathbb{R}$, or F is not included in that kernel in which case $\{h \leq t\} + F = E$ for any $t \in \mathbb{R}$. In both cases property $\mathcal{D}(t)$ is easily checked. □

Among those functions h that also satisfy property $\mathcal{D}(t)$ for any $t \in h(E) \setminus \{+\infty\}$, one may think to weakly coercive functions, as the following lemma shows.

Lemma 2.4. *Let $h : E \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. proper quasiconvex function such that $\{h \leq t\}$ is weakly sequentially compact for any $t \in \mathbb{R}$.* Then *h satisfies $\mathcal{D}(t)$ for any t in $h(E) \setminus \{+\infty\}$.*

Proof. Let t in $h(E)\setminus\{+\infty\}$ and F be a closed subspace of E. Let x belong to $\bigcap_{s>t} \overline{\{h \leq s\} + F}$, then there exist two sequences $(x_n)_n$ and $(\xi_n)_n$ such that: $x_n + \xi_n \to x$, $x_n \in \{h \leq t + \frac{1}{n}\}$ and $\xi_n \in F$ for any $n \in \mathbb{N}$. We can extract from $(x_n)_n$ a converging subsequence $(x_{n_k})_k$ which weakly converges to some $z \in \{h \leq t\}$, and then $(\xi_{n_k})_k$ weakly converges to $x - z \in F$, thus x belongs to $\{h \leq t\} + F$. □

If E is finite dimensional, then lemma 2.3 yields that any affine function h satisfies $\mathcal{D}(t)$ for any t in $h(\mathbb{R}^N)$, and more generally we shall see in §3 that this also holds true for any analytic convex function $h : \mathbb{R}^N \to \mathbb{R}$. Let us give an example of a continuous convex function which does not satisfy this property. We define the convex continuous function g on $\mathbb{R}^2$ by

$$g(x, y) \;:=\; \sqrt{x^2 + y^2} - x. \tag{2.1}$$

The function g is that given in Duffin's example of a duality gap (see [2]). Then $0 \in g(\mathbb{R}^2)$ but g does not satisfy $\mathcal{D}(0)$. Indeed, take $F = \mathbb{R} \times \{0\}$, then one has

$$\bigcap_{s>0} \overline{\{g \leq s\} + F} = \mathbb{R}^2 \neq \mathbb{R} \times \{0\} = \overline{\{g \leq 0\} + F}.$$

We notice on the above example that the functions $(x, y) \mapsto \sqrt{x^2 + y^2}$ and $(x, y) \mapsto -x$ satisfy $\mathcal{D}(0)$ (the first is the distance to $\{(0, 0)\}$, the second is affine) whereas their sum does not.

Property $\mathcal{D}$ may be seen as a stability condition for the sublevel sets of h under linear perturbation. The following result gives another caracterization which will reveal more easy to handle in the proof of theorem 2.6.

Proposition 2.5. *Let $h : E \to \mathbb{R} \cup \{+\infty\}$ be an l.s.c. quasiconvex function, and let $t \in h(E) \setminus \{+\infty\}$. The following are equivalent*

1. *h satisfies $\mathcal{D}(t)$;*
2. *for any closed convex set $C \subset E$ such that $\{h \leq s\} \cap C \neq \emptyset$ for any $s > t$, the distance $d(\{h \leq t\}, C)$ is equal to 0.*

In the above statement, the gap $d(A, B)$ between two subsets A and B of E is defined by $d(A, B) := \inf\{d(x, B) : x \in A\}$.

Proof. We first prove (1) ⇒ (2) by contradiction. Assume that $C \subset E$ is convex, $\{h \leq s\} \cap C \neq \emptyset$ for any $s > t$ and $d(\{h \leq t\}, C) > 0$. Then the Hahn-Banach theorem yields the existence of a continuous linear form l such that

$$\sup\{l(x) : x \in C\} \;<\; \inf\{l(x) : x \in \{h \leq t\}\}. \tag{2.2}$$

If we denote by H the kernel of l, this yields

$$\overline{C+H} \;\cap\; \overline{\{h \le t\}+H} \;=\; \emptyset. \tag{2.3}$$

On the other hand, one has

$$\overline{C+H} \;\cap\; \overline{\{h \le s\}+H} \;\neq\; \emptyset \tag{2.4}$$

for any $s > t$. Moreover, let us note that

$$\forall x \in E \quad \forall A \subset E \qquad l(x) \in \overline{l(A)} \;\Rightarrow\; x + H \subset \overline{A+H}. \tag{2.5}$$

Indeed, take $y \in E$ such that $l(y) = 1$ and let $(\alpha_n)_n$ be a sequence in $l(A)$ such that $\alpha_n \to l(x)$. Since for any $\alpha \in \mathbb{R}$ one has $\alpha y + H = \{l = \alpha\}$, we infer that $\alpha_n y$ belongs to $A + H$ for any n so that $l(x)y$ belongs to $\overline{A+H}$ and thus $x + H = l(x)y + H$ is included in $\overline{A+H}$.

Let $\overline{x} \in E$ be such that $l(\overline{x}) = \sup\{l(x) : x \in C\}$. We notice that $\overline{x} + H = \{l = l(\overline{x})\}$ and we get from (2.5) that $\overline{x} + H$ is included in $\overline{C+H}$. Let $s > t$, then the set $\overline{\{h \le s\}+H}$ is convex and contains $\overline{\{h \le t\}+H}$, so we infer from (2.2) and (2.4) that $l(\overline{x})$ belongs to $l(\overline{\{h \le s\}+H})$. As a consequence of (2.5), $\overline{x} + H$ is included in $\overline{\{h \le s\}+H}$ for any $s > t$, so that

$$(\overline{x}+H) \;\subset\; \left(\overline{C+H} \;\cap\; \bigcap_{s>t} \overline{\{h \le s\}+H}\right).$$

This together with (2.3) obviously contradicts that h satisfies $\mathcal{D}(t)$.

We now turn to (2) $\Rightarrow$ (1). Assume that F is a closed subspace of E such that

$$x \;\in\; \bigcap_{s>t} \overline{\{h \le s\}+F} \;\setminus\; \overline{\{h \le t\}+F} \tag{2.6}$$

for some x. Then there exists $r > 0$ such that $B(x,r) + F \;\cap\; \overline{\{h \le t\}+F}$ is empty, so that $d(C, \{h \le t\}) \ge \frac{r}{2} > 0$ where $C := \overline{B(x, \frac{r}{2}) + F}$. But (2.6) yields that $C \cap \{h \le s\} \neq \emptyset$ for any $s > t$, which contradicts (2). □

The following result links property $\mathcal{D}$ and the existence of a duality gap in mathematical programming.

Theorem 2.6. *Let $h : E \to \mathbb{R} \cup \{+\infty\}$ be an l.s.c. quasiconvex function and let t belong to $h(E) \setminus \{+\infty\}$. If $f : E \to \mathbb{R} \cup \{+\infty\}$ is a convex function, we set*

$$\forall s \ge t \qquad v(s) \;:=\; \inf\{f(x) \;:\; h(x) \le s\}.$$

Then *one either has*

i. *h satisfies condition $\mathcal{D}(t)$, in which case $\lim_{s \to t, s > t} v(s) = v(t)$ for any l.s.c. convex function $f : E \to \mathbb{R} \cup \{+\infty\}$ such that f is finite and continuous at at least one point of $\{h \le t\}$.*

ii. *h does not satisfy condition $\mathcal{D}(t)$, in which case there exists a continuous linear form $f : E \to \mathbb{R}$ for which $\lim_{s \to t, s > t} v(s) < v(t)$.*

Proof. i. Assume first that h satisfies $\mathcal{D}(t)$, and let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. convex function which is continuous at some point $\tilde{x} \in \{x : h(x) \le t\}$. We aim to show that the value function $v(.)$ is l.s.c. at t. We may assume without loss of generality that $v(t) > -\infty$. We proceed by contradiction: suppose that

$$\liminf_{s \to t, s > t} v(s) \;\le\; \alpha \;<\; v(t) \tag{2.7}$$

for some $\alpha \in \mathbb{R}$. Then if we set $\delta := (v(t)-\alpha)/3$, hypothesis (2.7) yields that $\{f \leq \alpha+\delta\} \cap \{h \leq s\} \neq \emptyset$ for any $s > t$. Then applying proposition 2.5 yields

$$d(\{f \leq \alpha+\delta\}, \{h \leq t\}) = 0.$$

As a consequence, there exist two sequences $(x_n)_n$ and $(y_n)_n$ such that $x_n - y_n \to 0$ and $h(x_n) \leq t$ and $f(y_n) \leq \alpha + \delta$ for any $n \in \mathbb{N}$. Since f is continuous at $\tilde{x}$, there exists $r > 0$ such that f is bounded by some number M over the ball $B(\tilde{x}, 2r)$. For any $n \in \mathbb{N}$, we set

$$z_n \quad := \quad \frac{\|x_n - y_n\|}{\|x_n - y_n\| + r}\,\tilde{x} + \frac{r}{\|x_n - y_n\| + r}\,x_n, \tag{2.8}$$

and we notice that z_n also satisfies

$$z_n \quad = \quad \frac{\|x_n - y_n\|}{\|x_n - y_n\| + r}\left(\tilde{x} + \frac{r}{\|x_n - y_n\|}(x_n - y_n)\right) + \frac{r}{\|x_n - y_n\| + r}\,y_n. \tag{2.9}$$

Since $\tilde{x}$ and x_n belong to $\{h \leq t\}$, we deduce from (2.8) and the quasiconvexity of h that $h(z_n) \leq t$ for any $n \in \mathbb{N}$. We then deduce from (2.9) and the convexity of f that

$$\limsup_{n\to+\infty} f(z_n) \;\leq\; \limsup_{n\to+\infty}\left(\frac{\|x_n - y_n\|}{\|x_n - y_n\| + r}\,M + \frac{r}{\|x_n - y_n\| + r}\,(\alpha+\delta)\right) \;=\; \alpha + \delta$$

so that $v(t) \leq \alpha + \delta \leq \frac{2}{3}\alpha + \frac{1}{3}v(t)$, which is a contradiction.

ii. Assume now that h does not satisfy $\mathcal{D}(t)$. This means that there exist a closed subspace F of E and a vector z such that

$$z \in \bigcap_{s>t} \overline{\{h \leq s\} + F} \setminus \overline{\{h \leq t\} + F}.$$

Then the Hahn-Banach theorem yields the existence of a continuous linear form $f : E \to \mathbb{R}$ such that

$$f(z) \quad < \quad \alpha \;:=\; \inf\left\{f(x) : x \in \overline{\{h \leq t\} + F}\right\}.$$

As a consequence, F is included in the kernel of f, so that

$$v(s) \;=\; \inf\{f(x)\,:\, x \in \{h \leq s\}\} = \inf\left\{f(x)\,:\, x \in \overline{\{h \leq s\} + F}\right\} \;\leq\; f(z)$$

for any $s > t$. Therefore $\limsup_{s\to t, s>t} v(s) \leq f(z) < \alpha = v(t)$, which concludes the proof. ☐

We notice that the hypothesis f *is continuous at at least one point of* $\{h \leq 0\}$ in theorem 2.6.i can't be weakened in general. Indeed, define the convex functions $h, f : \mathbb{R} \to \mathbb{R}$ by $h(x) = x$, and $f(x) = +\infty$ for $x < 0$, $f(0) = 1$ and $f(x) = 0$ for $x > 0$. Then h obviously satisfies $\mathcal{D}(0)$ but $\lim_{s\to 0, s>0} v(s) = 0 < 1 = v(0)$.

3. The property $\mathcal{D}$ in the convex case

As the following proposition shows, for a convex function h the property $\mathcal{D}(t)$ only really makes sense for the special case $t = \min(h)$.

Proposition 3.1. *Let $h : E \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function. Then for any t in $]\inf(h), +\infty[$, h satisfies condition $\mathcal{D}(t)$.*

Proof. Let t belong to $]\inf(h),+\infty[$ and F be a closed subspace of E. To check that

$$\bigcap_{s>t}\overline{\{h\le s\}+F}\ \subset\ \overline{\{h\le t\}+F},$$

we fix some z in $\bigcap_{s>t}\overline{\{h\le s\}+F}$: there exist two sequences $(x_n)_{n\in\mathbb{N}}$ and $(\xi_n)_{n\in\mathbb{N}}$ such that $(x_n+\xi_n)_{n\in\mathbb{N}}$ converges to z, with $h(x_n)\le t+1/n$ and $\xi_n\in F$ for any $n\ge 1$. Let now y in E be such that $h(y)<t$, we then set for any $n\ge 1$:

$$\alpha_n:=\frac{t-h(y)}{t+\frac{1}{n}-h(y)}\quad\text{and}\quad\tilde{x}_n:=\alpha_n x_n+(1-\alpha_n)y.$$

Then α_n belongs to $]0,1[$ and we infer from the convexity of h that $h(\tilde{x}_n)\le t$ for any $n\ge 1$. Since $(\alpha_n)_n$ converges to 1 as n goes to $+\infty$ and since

$$\|\tilde{x}_n+\alpha_n\xi_n-z\|\le\alpha_n\|x_n+\xi_n-z\|+(1-\alpha_n)\|y-z\|$$

we conclude that $(\tilde{x}_n+\alpha_n\xi_n)_n$ converges to z, so that z belongs to $\overline{\{h\le t\}+F}$. □

Proposition 3.1 above does not hold for quasiconvex functions. Let us indeed define the continuous quasiconvex function $\tilde{g}$ on $\mathbb{R}^2$ by

$$\tilde{g}(x,y):=\begin{cases} g(x,y) & \text{if } y\ge 0,\\ g(x,0)=-2x & \text{if } x,y\le 0,\\ \max(-x,y) & \text{if } x\ge 0,\ y\le 0,\end{cases}$$

where g is defined by (2.1). Then $\tilde{g}(1,-1)=-1$ but $\tilde{g}$ does not satisfy $\mathcal{D}(0)$ since

$$\bigcap_{s>0}\overline{\{\tilde{g}\le s\}+\mathbb{R}\times\{0\}}\ =\mathbb{R}^2\ \ne\ \mathbb{R}\times]-\infty,0]\ =\ \overline{\{\tilde{g}\le 0\}+\mathbb{R}\times\{0\}}.$$

It has already been noticed that property $\mathcal{D}(t)$ is not stable under the addition, the following proposition shows that it is stable under the *max* operation under mild assumptions.

Proposition 3.2. *Let $t\in\mathbb{R}$ and $f_1,f_2:E\to\mathbb{R}\cup\{+\infty\}$ be proper l.s.c. convex functions such that $\{f_1<t\}\cap\{f_2\le t\}\ne\emptyset$ and f_1 is continuous at some point y in $\{f_1<t\}\cap\{f_2\le t\}$.* Then *if f_2 satisfies $\mathcal{D}(t)$, the function $h:=\max\{f_1,f_2\}$ also satisfies condition $\mathcal{D}(t)$.*

Proof. Let $l:E\to\mathbb{R}$ be a continuous linear form. We claim that

$$\lim_{s\to t,s>t}\inf\{l(x)\ :\ f_1(x)\le s,f_2(x)\le s\}\ =\ \inf\{l(x)\ :\ f_1(x)\le t,f_2(x)\le t\}.\tag{3.1}$$

For any $s\ge t$, we set $v(s):=\inf\{l\ :\ f_1,f_2\le s\}$. We first claim that $\lim_{s\to t}v(s)=\lim_{s\to t}w(s)$ where $w(s):=\inf\{l\ :\ f_1\le t,f_2\le s\}$ for any $s\ge t$. Indeed, let $(x_s)_{s>t}$ be a family such that $l(x_s)\le v(s)+s-t$ and $f_1(x_s),f_2(x_s)\le s$ for any $s>t$. Let $y\in\{f_1<t\}\cap\{f_2\le t\}$ and set

$$\forall s>t\qquad\qquad \alpha_s:=\frac{t-f_1(y)}{s-f_1(y)}.$$

Then $\alpha_s x_s+(1-\alpha_s)y$ belongs to $\{f_1\le t,f_2\le s\}$ for any $s>t$ and

$$\lim_{s\to t}v(s)\le\lim_{s\to t}w(s)\le\lim_{s\to t}l(\alpha_s x_s+(1-\alpha_s)y)\le\lim_{s\to t}v(s)$$

which proves the claim. We now notice that

$$w(s)=\inf\{l(x)+\delta_{\{f_1\le t\}}(x)\ :\ f_2(x)\le s\}$$

for any $s\ge t$. Since by hypothesis the function $l+\delta_{\{f_1\le t\}}$ is continuous at some point y in $\{f_2\le t\}$, we may apply theorem 2.6.i which yields that $\lim_{s\to t,s>t}w(s)=w(t)=v(t)$, which concludes the proof of (3.1). Since (3.1) holds for any continuous linear form l, theorem 2.6.ii yields that h satisfies $\mathcal{D}(t)$. □

The conclusion of proposition 3.2 may not hold when $\{f_1 < t\} \cap \{f_2 \le t\} = \emptyset$, even if the functions f_1 and f_2 are continuous on E. For example, consider the convex cones C_1 and C_2 of $\mathbb{R}^3$ given by

$$\begin{aligned} C_1 &:= \{(x,y,z) \,:\, x \ge 0, \|(y,z) - (x,0)\| \le x\}\,, \\ C_2 &:= \{(x,y,z) \,:\, x \ge 0, \|(y,z) - (-x,0)\| \le x\}\,. \end{aligned}$$

Then the distance functions $f_1 := d(C_1,.)$ and $f_2 := d(C_2,.)$ to these sets both satisfy condition $\mathcal{D}(0)$ and are both Lipschitz continuous on $\mathbb{R}^3$, but the function $h := \max\{f_1, f_2\}$ does not satisfy $\mathcal{D}(0)$ since

$$\bigcap_{s>0} \overline{\{h \le s\} + \mathbb{R} \times \{(0,0)\}} = \mathbb{R} \times \{0\} \times \mathbb{R} \neq \mathbb{R} \times \{(0,0)\} = \overline{\{h \le 0\} + \mathbb{R} \times \{(0,0)\}}.$$

We now turn more specifically to the finite dimensional case, and prove that a wide class of convex functions satisfy property $\mathcal{D}$ in that setting. Before this, we recall the definitions of the notion of weakly analytic (or quasianalytic) functions introduced in [4] and of the class of functions $\mathcal{F}$ introduced in [1].

Definition 3.3. *Let $h : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be an l.s.c. proper convex function. Then f is weakly analytic if the following holds: if f is constant on a segment $[x,y]$ for some $x \neq y$, then f is constant on the whole line (x,y).*

The function h belongs to the class of functions $\mathcal{F}$ if the following holds: for any $\rho > 0$, any sequence $(\alpha_n)_n$ of real numbers converging to some α and any sequence $(x_n)_n$ in $\mathbb{R}^N$ such that

$$\forall n \in \mathbb{N} \quad x_n \in \{h \le \alpha_n\}\,,\ \|x_n\| \to +\infty\,,\ \frac{x_n}{\|x_n\|} \to \overline{x} \in \{h_\infty = 0\}$$

there exists $n_0 \in \mathbb{N}$ such that

$$\forall n \ge n_0 \qquad x_n - \rho\overline{x} \in \{h \le \alpha_n\}.$$

In the above definition, h_∞ denotes the recession function of h. Of course, analytic convex functions are weakly analytic, so that any linear function and any convex quadratic function is weakly analytic. Moreover, any strictly convex function is obviously weakly analytic. Linear and convex quadratic functions also are good examples of functions belonging to the class $\mathcal{F}$, and this class of function is stable under the "max" operation. We refer to [1] and [4] for more examples of functions in both classes. We also point out that these two classes of functions are distinct as shown in [1].

Thanks to the notions introduced in definition 3.3, the following result enables to show that many functions satisfy property $\mathcal{D}$.

Proposition 3.4. *Let $f_i : \mathbb{R}^N \to \mathbb{R}$ be continuous convex functions such that for any i in* $\{1, \dots, k\}$ either *f_i is weakly analytic* or *f belongs to $\mathcal{F}$.* Then $h := \max\{f_i : i = 1, \dots, k\}$ *satisfies condition $\mathcal{D}(t)$ for any t in $h(\mathbb{R}^N)$.*

Proof. Thanks to proposition 3.1, it is sufficient to check that h satisfies $\mathcal{D}(t)$ for $t := \min(h)$ whenever h attains its minimum on $\mathbb{R}^N$ (otherwise there is nothing to prove). Assume that h attains $t := \min(h)$ on $\mathbb{R}^N$. Let I be the set of indices i in 1,...,k such that for some $x \in Argmin(h)$ one has $f_i(x) < t$. Then by convexity there exists x in $Argmin(h)$ such that $f_i(x) < t$ for any $i \in I$. If we set $h_I := \max\{f_i : i \in I\}$ and $h_J := \max\{f_j : j \notin I\}$, proposition 3.2 yields that it is sufficient to check that h_J satisfies property $\mathcal{D}(t)$. We may then assume that t is the minimum of h_J on $\mathbb{R}^N$, otherwise proposition 3.1 applies again. Repeating the preceding arguments, we may also assume that any function f_j with $j \notin I$ is constant and equal to t on $Argmin(h_J)$.

Let now F be a subspace of $\mathbb{R}^N$, and let z belong to $\bigcap_{s>t} \overline{\{h_J \leq s\} + F}$. Let then $(x_n)_{n\in\mathbb{N}}$ be such that for any integer $n \geq 1$, x_n is of minimal norm among those vectors $y \in \mathbb{R}^N$ such that there exists ξ in F for which

$$\|z - (y + \xi)\| \leq 1/n \quad and \quad h_J(y) \leq t + 1/n.$$

Let then $(\xi_n)_n$ be such that $\|z - (x_n + \xi_n)\| \leq 1/n$ for any $n \geq 1$. If $\|x_n\| \not\to +\infty$, then it is easily checked that z belongs to $\overline{\{h_J \leq t\} + F}$. We may thus assume that $\|x_n\| \to +\infty$ as well as $\frac{x_n}{\|x_n\|} \to \overline{x}$ for some $\overline{x}$ in $\mathbb{R}^N$. Since h_J attains its minimum t on $\mathbb{R}^N$, we infer that $h_{J\infty}(\overline{x}) = 0$, and then since any function f_j with $j \notin I$ is constant and equal to t on $Argmin(h_J)$ we also infer that $f_{j\infty}(\overline{x}) = 0$ for any such j. Moreover, we notice that $\frac{\xi_n}{\|x_n\|} \to -\overline{x}$ so that $\overline{x}$ belongs to F.

Let now J_w (resp. $J_{\mathcal{F}}$) denote the set of indices in $j \in \{1, \ldots, k\} \setminus I$ such that f_j is weakly analytic (resp. $f_j \in \mathcal{F}$). Then by definition, the functions f_j for which $j \in J_w$ are constant in the direction $\overline{x}$, so that $f_j(x_n - \overline{x}) \leq t + 1/n$ for any such function. On the other hand, there exists $n_0 \in \mathbb{N}$ such that $f_j(x_n - \overline{x}) \leq t + 1/n$ for any $n \geq n_0$ and $j \in J_{\mathcal{F}}$. Now one has

$$\forall n \geq n_0 \qquad h_J(x_n - \overline{x}) \leq t + 1/n\,,\ \xi_n + \overline{x} \in F\,,\ \|z - ((x_n - \overline{x}) + (\xi_n + \overline{x}))\| \leq \frac{1}{n}\,.$$

We conclude as in [1] by noticing that

$$\|x_n - \overline{x}\| = \left\|(1 - \frac{1}{\|x_n\|})x_n + (\frac{x_n}{\|x_n\|} - \overline{x})\right\| \leq (1 - \frac{1}{\|x_n\|})\|x_n\| + \left\|\frac{x_n}{\|x_n\|} - \overline{x}\right\|$$

and since $\|x_n - \overline{x}\| \geq \|x_n\|$ (x_n is of minimal norm), this yields

$$\forall n \geq n_0 \qquad 1 \leq \left\|\frac{x_n}{\|x_n\|} - \overline{x}\right\|.$$

We then infer the contradiction $1 \leq 0$ by letting n go to $+\infty$. As a consequence $(x_n)_n$ is necessarily bounded and the proof is complete. □

4. Duality gap in convex programming

We now go back to the duality gap in convex programming: for any non-negative t, we consider the ordinary convex program (P_t) given by

$$(P_t) \qquad \mathrm{Inf}\left\{f(x)\ :\ x \in \mathbb{R}^N, f_1(x) \leq t, \ldots, f_k(x) \leq t\right\}$$

where the fonctions $f, f_i : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semi-continunous. In the following, we assume that $\mathrm{Inf}(P_0) > -\infty$. The dual convex problem associated to (P_0) is then given by

$$(D) \qquad \mathrm{Sup}\left\{\inf\{f(x) + \sum_{i=1}^{k} \lambda_i f_i(x)\ :\ x \in \mathbb{R}^N\}\ :\ \lambda_1 \geq 0, \ldots, \lambda_k \geq 0\right\}.$$

Convex duality theory then yields that there is no duality gap between (P_0) and (D) whenever the value function $v : \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$ is lower semi-continuous at 0, where v is associated to the family of problems $(P_t)_{t\geq 0}$ by $v(t) := \mathrm{Inf}(P_t)$ for any $t \geq 0$ (we refer to ch. 3 in [3]). As a straightforward corollary of theorem 2.6, we get the following result on the duality gap in convex programming.

Theorem 4.1. *Let $f_i : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be l.s.c. convex functions such that the set $\{h \leq 0\}$ is not empty, where $h := \max\{f_1, \ldots, f_k\}$.* Then *one either has*

i. *h satisfies condition $\mathcal{D}(0)$, in which case there is no duality gap between (P_0) and (D) whenever the l.s.c. convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is finite and continuous at at least one point of $\{h \le 0\}$.*

ii. *h does not satisfy condition $\mathcal{D}(0)$, in which case there exists an affine function $f : \mathbb{R}^N \to \mathbb{R}$ for which the duality gap between (P_0) and (D) is positive.*

The results obtained so far in sections 2 and 3 concerning property $\mathcal{D}$ then yield the following results and comments.

- When one function f_i is coercive on $\mathbb{R}^N$ (which in finite dimension reads $f_i(x) \to +\infty$ whenever $\|x\| \to +\infty$), lemma 2.4 yields that the function h satisfies property $\mathcal{D}(0)$, so that theorem 4.1.i applies. Notice that in this case, the fact that there is no duality gap is in fact a simple consequence of the compactness of the constraint sets involved in the problems (P_t).
- When the functions f_i satisfy the Slater condition (i.e. $\{h < 0\}$ is not empty), then proposition 3.1 also allows to apply theorem 4.1. This result is classical, and in fact in this case one may even show that the dual problem (D) has at least one optimal solution: indeed the value function v is then defined with real values in a neighbourhood of 0, and since v is convex and nondecreasing it turns out that it is continuous at 0 and thus its subdifferential at 0 is not empty so that (D) has optimal solutions (we refer to ch.3 of [3] for more details). This result was improved by Li [5], where the absence of a duality gap and existence of an optimal solution to (D) is proven under the constraint qualification $\sup\{dist(\{h \le 0\}, \{h \le \varepsilon\})/\varepsilon : \varepsilon > 0\} < \infty$. This constraint qualification is stronger than property $\mathcal{D}(0)$ since this may be satisfied even if (D) has no optimal solution as noted below.
- When each of the functions f_i is either weakly analytic or belongs to $\mathcal{F}$, then once again h satisfies property $\mathcal{D}(0)$ as a consequence of proposition 3.4. The fact that there is no duality gap in this case generalizes previous results by Auslender [1], Kummer [4] and Rockafellar [7]. In [1], it is assumed that all the functions involved in (P_0) belong to $\mathcal{F}$ and that they have the same domain, as a counterpart the author not only proves that there is no duality gap but also that the infimum in (P_0) is attained. Notice that on the one hand there is no continuity assumption made upon the constraint functions f_i in [1] (while this is assumed in proposition 3.4), but on the other hand there is a regularity assumption made on every function f and f_i (that should all belong to $\mathcal{F}$). In [7], it is shown that there is no duality gap under the hypothesis that every function involved in (P_0) is continuous and faithfully convex (which is a little stronger than weakly analytic), and this result was improved in [4] where it is only assumed that the constraint functions f_i are weakly analytic. Proposition 3.4 thus allows to mix the classes $\mathcal{F}$ and that of weakly analytic functions and still ensures that there is no duality gap.
- If one takes the single constraint $h := g - 1$ where g is the Duffin's function given by (2.1), then f_1 satisfies $\mathcal{D}(0)$ (thanks to proposition 3.1) whereas one may observe that the diameter of the sets $\{f_1 \le \varepsilon\} \setminus \{f_1 \le 0\}$ is $+\infty$ for any positive ε, so that property $\mathcal{D}(0)$ does not imply the convergence of the sets $\{f_1 \le \varepsilon\}$ to $\{f_1 \le 0\}$.

Besides the duality gap, another important question concerning (P_0) and (D) is that of the existence of Lagrange multipliers, i.e. of optimal solutions to the dual problem (D). The following example, given in [6], shows that Lagrange multipliers may not exist even if there is no duality gap. Let the convex continuous functions f, f_1 and f_2 be defined on $\mathbb{R}^2$ by $f(x_1, x_2) := x_1$, $f_1(x_1, x_2) := x_2$ and $f_2(x_1, x_2) := x_1^2 - x_2$. The constrained problem we consider is

$$(P) \qquad \operatorname{Inf}\{f(x) \,:\, f_1(x) \le 0, f_2(x) \le 0\}$$

and its dual problem given by

$$(D) \qquad \text{Sup}\left\{L(\lambda_1,\lambda_2) := \inf\{f(x)+\lambda_1 f_1(x)+\lambda_2 f_2(x) \,:\, x\in\mathbb{R}^2\} \,:\, \lambda_1\geq 0, \lambda_2\geq 0\right\}.$$

Since the functions f_1 and f_2 are quasianalytic, the function $\max\{f_1,f_2\}$ satisfies property $\mathcal{D}(0)$ so that there is no duality gap between (P) and (D), and one can indeed check that $inf(P) = sup(D) = 0$. However, the supremum of (D) is not attained since $L(\lambda_1,\lambda_2) = -\infty$ whenever $\lambda_1 \neq \lambda_2$, whereas $L(0,0) = -\infty$ and $L(\lambda,\lambda) = -\frac{1}{4\lambda}$ for any $\lambda > 0$.

References

[1] A. Auslender. *Existence of optimal solutions and duality results under weak conditions.* Math. Prog., Ser A (88):45–59, 2000.

[2] J.M. Borwein and A.S. Lewis, *Convex analysis and nonlinear optimization.* Springer, New York, 2000.

[3] I. Ekeland and R. Temam. *Analyse convexe et problèmes variationnels.* Dunod Gauthiers-Villars, Paris, 1974.

[4] B. Kummer. *Stability and weak duality in convex programming without regularity.* Wiss. Z. Humboldt-Univ. Berlin Math.-Natur. Reihe, 30 (5):381–386, 1981.

[5] S. Li. *A constraint qualification for convex programming.* Acta Math. Appl. Sinica (English Ser.), 16 (4):362–365, 2000.

[6] R.T. Rockafellar. *Convex Analysis.* Princetown Univ. Press, Princetown, 1970.

[7] R.T. Rockafellar. *Ordinary convex programs without a duality gap.* J. Opt. Th Appl, 7 (3):143–148, 1965.

[8] R.T. Rockafellar and R.J-B. Wets. *Variational Analysis.* Springer-Verlag, Berlin, 1998.

T. Champion, Laboratoire d'Analyse Non Linéaire Appliquée, U.F.R. des Sciences et Techniques, Université de Toulon et du Var, Avenue de l'Université, BP 132, 83957 La Garde Cedex, FRANCE.

E-mail address: champion@univ-tln.fr