Published in
SIAM Journal on Optimization 13 (1) (2002) pp. 212–227.

TUBULARITY AND ASYMPTOTIC CONVERGENCE OF PENALTY TRAJECTORIES IN CONVEX PROGRAMMING .

T. CHAMPION

Abstract. In this paper, we give a sufficient condition for the asymptotic convergence of penalty trajectories in convex programming with multiple solutions. We show that, for a wide class of penalty methods, the associated optimal trajectory converges to a particular solution of the original problem, characterized through a minimization selection principle. Our main assumption for this convergence result is that all the functions involved in the convex program are tubular. This new notion of regularity, weaker than that of quasianalyticity, is defined and studied in detail.

Keywords. Convex Programming, Penalty Methods, Asymptotic Analysis, Tubular Functions.

MSC 2000. 49M37,65K10,90C25, 90C30.

1. Introduction

Let us consider a general convex program

$$(\mathrm{CP}_0) \qquad \operatorname{Inf}\{\Phi_0(x)\,:\,x\in C\}$$

where Φ_0 is convex and the constraint C is a convex subset of $\mathbb{R}^N$ which can be written in the form

$$C := \left\{x\in\mathbb{R}^N\,:\,\Phi_i(x)\le 0,\quad i=1,\dots,M\right\}.$$

with continuous convex functions Φ_i. In order to handle this kind of constraints, for numerical computations or theoretical study, it has become classical to approximate this problem by means of a penalization method . Given a penalty function $\theta:\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$, we associate to (CP_0) a family $(\mathrm{CP}_r)_{r>0}$ of approximating problems given by

$$(\mathrm{CP}_r) \qquad \operatorname{Inf}\left\{\Phi_0(x)+\alpha(r)\sum_{i=1}^{M}\theta\left(\frac{\Phi_i(x)}{r}\right)\,:\,x\in\mathbb{R}^N\right\}$$

Laboratoire d'Analyse, de Calcul Scientifique Industriel et d'Optimisation de Montpellier, Département de Mathématiques, case courrier 051, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex, France. e-mail: champion@math.univ-montp2.fr.

where $\alpha :]0,+\infty[\to]0,+\infty[$ is a rescaling function. Suitable assumptions (see §2) on the functions Φ_i, θ and α guarantee that the optimal values $v(\mathrm{CP}_r)$ converge to $v(\mathrm{CP}_0)$ as r goes to 0. Our work addresses the asymptotic behaviour of the net of optimal solutions $(x_r)_{r>0}$ of the approximating problems (CP_r) as r goes to 0. One of the fundamental properties of usual penalty methods is that the penalty trajectory $(x_r)_{r>0}$ is bounded as r goes to 0 and that every cluster point of this net is an optimal solution of the initial problem (CP_0). Here, we are particularly interested in the case where (CP_0) has more than one solution (so it has infinitely many, since the optimal set is convex). In this case, the penalty trajectory may have several cluster points as r goes to 0, see §5.3.

The convergence of the whole trajectory to a single solution may be of practicle interest in numerical computations: when the trajectory does not converge, it may have a bad oscillating behaviour as r tends to 0. In linear programming, it is known that the penalty trajectory converges to a particular solution (related to the penalty function used) for some penalty methods. For example, this particular solution is called the analytic center for the logarithmic barrier method (see MacLinden [18], Sonnevend [24] or [6] for comments), whereas it is called the absolute minimizer for the exponential penalty (see Cominetti and San Martin [9]). In [6], Auslender, Cominetti and Haddou presented a general analysis for the asymptotic convergence of penalty trajectories in linear programming. For more general convex programs, the asymptotic convergence to the analytic center for the logarithmic barrier was obtained for analytic functions Φ_i by Monteiro and Zhou [20], while the convergence to the absolute minimizer for the exponential penalty was proven for quasianalytic functions Φ_i by Alvarez [1]. Recently, Cominetti [8] proposed a unified approach to this problem of asymptotic convergence for a wide class of penalty functions.

In this work, we use the notion of nonlinear averages introduced in this context in [8]. It allows us to isolate particular solutions of (CP_0), the θ-centers, which are local solutions (in the sense of Dal Maso-Modica [10]) of an auxiliary optimization problem. This notion of particular solution generalizes those discussed above in linear programming and that defined in [8]. We provide with sufficient conditions which ensure that there is a unique θ-center, and that the penalty trajectory converges to this optimal solution. The main assumption we make to show the convergence of the penalty trajectory is the tubularity of the functions Φ_i. The notion of tubularity for a function Φ, introduced in §5, is a regularity condition on the behaviour of Φ in the neighbourhood of any nontrivial segment on which it is constant. The tubularity condition generalizes that of quasianalyticity, and may be of independent interest for the study of convex (or non convex) problems with multiple solutions.

In section §2, we recall the basic results and state the assumptions needed in the rest of the paper. In §3, we define the θ-center of the convex mathematical program (CP_0), while §4 is devoted to the definition and the study of tubular functions. This notion then allows us to show our main convergence result (theorem 5.1). Finally, we discuss of the possible extension of this convergence result to more general penalty methods, and show that when the hypotheses of theorem 5.1 are not fulfilled the approximating net (x_r) may fail to converge.

2. Penalty methods in convex programming

Let us consider the convex programming problem (CP_0) given by

$$(\mathrm{CP}_0) \qquad \operatorname{Inf}\{\Phi_0(x) \,:\, x \in C\}$$

where the feasible set C is a nonempty closed convex subset of $\mathbb{R}^N$ of the form

$$C := \left\{x \in \mathbb{R}^N \,:\, \Phi_i(x) \le 0\,,\ \ i = 1, \dots, M\right\}.$$

In the sequel, we will make the following assumptions on the functions Φ_i.

$$(H_0) \qquad \begin{cases} \Phi_0 : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\} \text{ is a closed proper convex function,} \\ \text{for } 1 \le i \le M,\ \Phi_i : \mathbb{R}^N \to \mathbb{R} \text{ are continuous convex functions,} \\ \text{the set } S(\mathrm{CP}_0) \text{ of the optimal solutions of } (\mathrm{CP}_0) \text{ is } \textit{nonempty} \text{ and } \textit{compact}. \end{cases}$$

We follow [6] and [8] and consider the class of penalty methods for (CP_0) which consist in approximating (CP_0) by the family $(\mathrm{CP}_r)_{r>0}$ of optimization problems

$$(\mathrm{CP}_r) \qquad \operatorname{Inf}\left\{\Phi_0(x) + \alpha(r)\sum_{i=1}^{M} \theta\left(\frac{\Phi_i(x)}{r}\right) \,:\, x \in \mathbb{R}^N\right\}$$

where the positive parameter r is intended to go to 0. We shall assume that the functions $\alpha :]0, +\infty[\to]0, +\infty[$ and $\theta : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ satisfy

$$(H_1) \qquad \begin{cases} \theta \text{ is increasing and convex on } \operatorname{dom}(\theta) =]-\infty, \eta[\,,\ \eta \in [0, +\infty], \\ \lim\limits_{t \to \eta^-} \theta(t) = +\infty\,, \quad \theta_\infty(-1) = 0, \quad \theta_\infty(1) > 0\,, \\ \lim\limits_{r \to 0^+} \alpha(r) = 0\,, \quad \theta_\infty(1).\liminf\limits_{r \to 0^+} \alpha(r)/r = +\infty\,. \end{cases}$$

As noted in [6], many penalty methods of the type $(\mathrm{CP}_r)_{r>0}$ with a function θ satisfying (H_1) appear in the literature. We refer to [6] for an extensive list.

Examples. The Logarithmic Barrier Method is obtained for the choice $\theta_1(t) = -\log(-t)$, the Inverse Barrier Method for $\theta_2(t) = -1/t$ (both with $\eta = 0$ and $\alpha(r) = r$), while the Exponential Penalty Method is obtained with $\theta_1(t) = \exp(t)$, $\eta = +\infty$ and $\alpha(r) = r$.

In the case of *interior* penalty methods (when $\eta = 0$), we shall assume that Slater's condition holds, that is: there exists x in $\operatorname{dom}(\Phi_0)$ such that $\Phi_i(x) < 0$ for all i in $\{1, \dots, M\}$.

We recall the following result from [8], which states that the penalty method defined above is a good approximation scheme for solving (CP_0).

Theorem 2.1. *Suppose that* (H_0) *and* (H_1) *hold. Then for* $r > 0$ *sufficiently small, the optimal set* $S(CP_r)$ *is nonempty and compact. This holds for any positive* r *if one has* $\theta_\infty(1) = +\infty$. *Moreover, each selection* $x_r \in S(CP_r)$ *stays bounded as* r *tends to* 0, *any cluster point* x_0 *of such a net* $(x_r)_{r>0}$ *belongs to* $S(CP_0)$ *and*

$$\lim_{r\to 0}\left[\Phi_0(x_r) + \alpha(r)\sum_{i=1}^{M}\theta\left(\frac{\Phi_i(x_r)}{r}\right)\right] \quad = \quad \mathrm{Inf}\,\{\Phi_0(x)\,:\, x \in C\}$$

Remark. This result also holds when the constraints Φ_i (for $i \in \{1,\ldots,M\}$) are only assumed to be l.s.c. on $\mathbb{R}^N$. The continuity assumption for these functions will only reveal necessary in the proof of the selection property (see theorem 5.2).

The following easy lemma ensures the uniqueness of the optimal solution to (CP_r) under a further condition on the functions Φ_i.

Lemma 2.2. *Suppose that* (H_0), (H_1) *hold and that* θ *is strictly convex. Assume that the functions* Φ_i *are such that*

(H_2) if the function $z \mapsto (\Phi_0(z),\ldots,\Phi_M(z))$ is constant on $[x,y]$ then $x = y$.

Then *for every positive* r, (CP_r) *has at most one solution.*

Remark. In the case of linear programming, (H_2) is satisfied as soon as the kernel of the linear part of the affine operator $(\Phi_0,\ldots,\Phi_M)$ is reduced to $\{0\}$, which is obviously the case when $S(\mathrm{CP}_0)$ is compact.

3. Non-linear averages and θ-centers

We now want to identify those optimal solutions $x_0 \in S(\mathrm{CP}_0)$ which can be obtained as cluster points (when $r \to 0$) of a selection $x_r \in S(\mathrm{CP}_r)$. Following Cominetti [8], we are thus led to introduce a notion of particular (or viscosity) solution of (CP_0), the notion of θ-center. To this end, we first recall the notion of θ-average (see [8]), which can be viewed as an asymptotic nonlinear average for vectors in $\mathbb{R}^m_-$.

Proposition and Definition 3.1. *Let* $\theta\,:\,]-\infty,0[\to\mathbb{R}$ *be increasing and convex, then for any* $m \geq 1$, *there exists a unique continuous function* $A^m_\theta : \mathbb{R}^m_- \to \mathbb{R}$, *which we call the* **$\theta$-average**, *such that for any* $y \in\,]-\infty,0[^m$ *one has*

$$A^m_\theta(y) = \limsup_{r\to 0^+} r\theta^{-1}\left(\frac{1}{m}\sum_{i=1}^{m}\theta\left(\frac{y_i}{r}\right)\right).$$

Moreover, A_θ^m is positively homogeneous, convex, symmetric, componentwise nondecreasing and satisfies

$$\forall y \in]-\infty, 0]^m \qquad \frac{1}{m}\sum_{i=1}^m y_i \le A_\theta^m(y) \le \max_{1\le i\le m} y_i.$$

We refer to [8] for the proof of this proposition and for further properties of the θ-average.

Examples. The computation of the θ-averages for the examples of the preceding section leads to

$$1)(Logarithmic Barrier) \qquad \forall y \in \mathbb{R}_-^m \qquad A_{\theta_1}^m(y) = -\left(\prod_{1\le i\le m}(-y_i)\right)^{\frac{1}{m}},$$

$$2)(Inverse Barrier) \qquad \forall y \in]-\infty, 0[^m \qquad A_{\theta_2}^m(y) = \left(\frac{1}{m}\sum_{1\le i\le m}\frac{1}{y_i}\right)^{-1},$$

$$3)(Exponential Penalty) \qquad \forall y \in \mathbb{R}_-^m \qquad A_{\theta_3}^m(y) = \max_{1\le i\le m} y_i\,.$$

Notice that the limsup in the definition of the θ-average over $]-\infty, 0[^m$ need not be a limit: it is still an open problem whether it is a limit or not in the general case. However, it is a limit for the above examples as well as for the other examples of penalty functions θ given in [6]. In the proof of our main result, we will have to assume that this also holds, i.e.

$$(H_3) \qquad \forall m \ge 1 \qquad \forall y \in]-\infty, 0[^m \qquad A_\theta^m(y) = \lim_{r\to 0^+} r\theta^{-1}\left(\frac{1}{m}\sum_{i=1}^m \theta\left(\frac{y_i}{r}\right)\right).$$

We next define the notion of θ-center of (CP_0): as theorem 5.2 shows, this is the viscosity solution (in the sense of [4]) associated to the penalty function θ.

Definition 3.2. *An optimal solution $x_* \in S(CP_0)$ is a* **θ-center** *of (CP_0) if for any $J \subset \{1, \dots, M\}$, x_* is an optimal solution of*

$$(\mathrm{CP}_{0,\theta,J}) \qquad \mathrm{Inf}\left\{A_\theta^{|J|}((\Phi_j(x))_{j\in J}) \,:\, x \in S(\mathrm{CP}_0) \text{ s.t. } \forall i \notin J\,,\ \Phi_i(x) = \Phi_i(x_*)\right\}.$$

Examples. We give two examples in the setting of linear programming: for every $i \in \{0, \dots, M\}$, Φ_i is affine and $x \mapsto (\Phi_0(x), \dots, \Phi_M(x))$ is injective (otherwise $S(\mathrm{CP}_0)$ is not compact).

1) In the case $\theta = \theta_1$, i.e. for the logarithmic barrier method, there exists a unique θ-center, usually called the analytic center. Indeed, set $I = \{i \,:\, 1 \le i \le M\,, \forall x \in S(\mathrm{CP}_0)\,,\ \Phi_i(x) = 0\}$: either $I = \{1, \dots, M\}$, in which case $S(\mathrm{CP}_0)$ is a singleton (so it is reduced to the unique θ-center), or $I \ne \{1, \dots, M\}$, then since the non-linear average $A_{\theta_1}^{M-|I|}$ is strictly convex on $]-\infty, 0[^{M-|I|}$, the analytic center of (CP_0) is the unique

optimal solution of the auxiliary problem

$$\operatorname{Inf}\left\{ A_{\theta_1}^{M-|I|}((\Phi_i(x))_{i\notin I}) = -\left(\prod_{i\notin I}(-\Phi_i(x))\right)^{\frac{1}{M-|I|}} \ : \ x \in S(\mathrm{CP}_0) \right\}$$

2) When $\theta = \theta_3$, i.e. for the exponential penalty method, the θ-center is also called the centroid, or absolute minimizer. Its existence and uniqueness is proven in [9]. In this case, the above definition reads: $x_* \in S(\mathrm{CP}_0)$ is the centroid if for any $J \subset \{1, \dots, M\}$, x_* is an optimal solution of

$$\operatorname{Inf}\left\{ \max_{j\in J} \Phi_j(x) \ : \ x \in S(\mathrm{CP}_0) \text{ such that } \forall i \notin J \quad \Phi_i(x) = \Phi_i(x_*) \right\}.$$

The notion of θ-center defined above is equivalent to that given in [8] under the more restrictive assumptions therein. With the hypotheses made in [8], the θ-center is shown to exist and be unique, while our hypotheses don't a priori imply neither the existence nor the uniqueness of a θ-center. However, theorems 5.1 and 5.2 below ensure the existence of at least one θ-center when the functions Φ_i are tubular. We now give a condition on the functions A_θ^m (for $m \geq 1$) under which the θ-center is uniquely defined.

Proposition 3.3. *Assume that (H_0), (H_1), (H_2) and the following (H_4) hold for $m \geq 1$.*

$$(H_4) \qquad \forall x, y \in \mathbb{R}_-^m \qquad \max_{1\leq i\leq m} x_i \neq \max_{1\leq i\leq m} y_i \Rightarrow A_\theta^m(\frac{x+y}{2}) < \max\{A_\theta^m(x), A_\theta^m(y)\}.$$

Then (CP_0) *has at most one θ-center.*

Proof. By contradiction, suppose that x and y are two distinct θ-centers of (CP_0). Then the set $I \subset \{1, \dots, M\}$ of indices i for which Φ_i is not constant over $[x, y]$ is nonempty, otherwise (H_2) implies that $x = y$. Since x and y are both θ-centers of (CP_0), they are both optimal solutions of

$$(\mathrm{CP}_{0,\theta,I}) \qquad \operatorname{Inf}\left\{ A_\theta^{|I|}((\Phi_i(z))_{i\in I}) \ : \ z \in S(\mathrm{CP}_0) \text{ such that } \forall j \notin I \quad \Phi_j(z) = \Phi_j(x) \right\}.$$

Let $z = (x + y)/2$ be the middle of $[x, y]$, then we claim that for any i in I

$$(1) \qquad \Phi_i(z) < \max\{\Phi_i(x), \Phi_i(y)\}.$$

By contradiction, assume that (1) is false, then, since Φ_i is convex, we obtain

$$\Phi_i(z) = \max\{\Phi_i(x), \Phi_i(y)\}.$$

We infer from the definition of z that Φ_i is constant on $[x, y]$, which contradicts the definition of I. Hence, one either has $\max_{i\in I} \Phi_i(z) < \max_{i\in I} \Phi_i(x)$ or $\max_{i\in I} \Phi_i(z) < \max_{i\in I} \Phi_i(y)$. Without loss of generality, we assume that the first inequality holds. Then (H_4) yields

$$A_\theta^{|I|}((\Phi_i((x + y)/4))_{i\in I}) < \max\{A_\theta^{|I|}((\Phi_i(x))_{i\in I}), A_\theta^{|I|}((\Phi_i(z))_{i\in I})\}.$$

Since $A_\theta^{|I|}$ is convex and $A_\theta^{|I|}((\Phi_i(x))_{i\in I}) = A_\theta^{|I|}((\Phi_i(y))_{i\in I})$, this implies

$$A_\theta^{|I|}((\Phi_i((x+y)/4))_{i\in I}) < A_\theta^{|I|}((\Phi_i(x))_{i\in I}).$$

But for $j \notin I$, one has $\Phi_j((x+y)/4) = \Phi_j(x)$, so the above inequality contradicts the optimality of x for $(\mathrm{CP}_{0,\theta,I})$. **Q.E.D.**

Notice that for the usual penalty functions (e.g. for θ_1, θ_2 and θ_3), hypothesis (H_4) is satisfied, so that under condition (H_2), the θ-center is uniquely determined by definition 3.2.

4. Tubularity and related notions

In [8], Cominetti shows that when all the functions Φ_i are quasianalytic and under conditions $(H_0) - (H_4)$ and some more hypotheses on the non-linear averages A_θ^m, the penalty trajectory $(x_r)_{r>0}$ converges towards the unique θ-center of (CP_0) as r goes to 0. We recall that quasianalyticity is defined as follows.

Definition 4.1. *A function $\Phi : E \to \mathbb{R} \cup \{+\infty\}$ is* **quasianalytic** *if whenever $x \neq y$ are such that Φ is finite and constant on $[x, y]$, then Φ is constant on the whole line passing through x and y.*

For example, every convex analytic function or strictly convex function is quasianalytic. However, simple functions such as convex piecewise affine functions or finite suprema of quadratic forms are not in general quasianalytic. This motivates the introduction of a weaker property than quasianalyticity for which the convergence of the penalty trajecory towards the θ-center still holds. Before giving the definition of this weaker property, we recall the notion of tubularity for subsets of $\mathbb{R}^N$. In the study of the l^∞-projection on a closed convex subset of $\mathbb{R}^N$, Huotari and Marano were led to define the notion of total tubularity for a convex set (also called property P, see [15], [17] and [19]) as a sufficient condition for the convergence of the Polya algorithm. Here we shall use the term tubular instead of totally tubular.

Definition 4.2. *Let d belong to $\mathbb{R}^N \setminus \{0\}$. A closed convex set C is* **d-tubular** *if for all x in C such that $x + td$ belongs to C for some positive t, there exists a neighbourhood V of x in C and a positive ε such that $z + \varepsilon d \in C$ for all z in $V \cap C$.*

A closed convex subset C of $\mathbb{R}^N$ is **tubular** *if it is d-tubular for any d in $\mathbb{R}^N \setminus \{0\}$.*

In terms of local recession vectors (see 6.33 in [23]), the above definition reads: C is d-tubular if d is a local recession vector for C at x whenever $x \in C$ is such that $x + td$ belongs to C for some positive t.

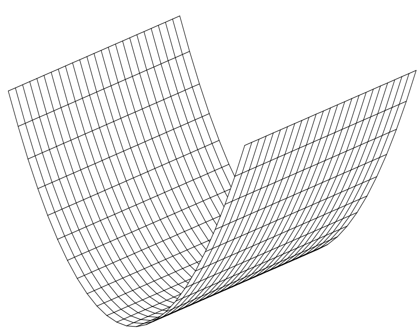
Examples. Any convex polyhedron and any cylinder of convex base in $\mathbb{R}^3$ is tubular. One can also prove that any convex subset of $\mathbb{R}^2$ is tubular (see proposition 4.5). Notice that simple convex sets may fail to be tubular, as shown in proposition 4.4 for the convex cone of $\mathbb{R}^{N+1}$ obtained as the epigraph of $x \mapsto \|x\|_{N,2} = \left(\sum_{i=1}^N x_i^2\right)^{\frac{1}{2}}$ (for $N \geq 2$). We refer to [17] for further comments and examples.

Let us introduce the sufficient condition for the convergence theorem 5.1, which is a generalization of the notion of tubularity to functions.

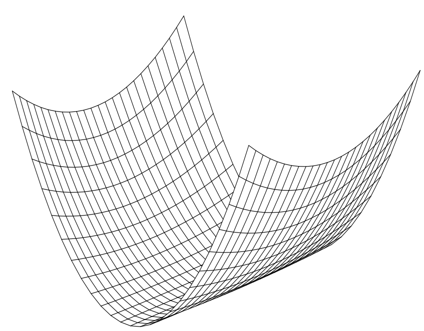
Definition 4.3. *A closed proper function $\Phi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is* **d-tubular** *(with $d \neq 0$) if whenever $x \in \mathbb{R}^N$ and $t > 0$ are such that Φ is finite and constant on $[x, x + td]$, then there exists a neighbourhood V of x and a positive ε such that for any z in V the function $s \mapsto \Phi(z + sd)$ is nonincreasing on $[0, \varepsilon]$ whenever $\Phi(z) \leq \Phi(x) + \varepsilon$.*

The function Φ is **tubular** *if it is d-tubular for any $d \neq 0$.*

Example. The tubularity property is satisfied by $\Psi_1(x, y) = y^2$ since it is always constant in the direction $(1, 0)$, whereas $\Psi_2(x, y) = (x^2 + 12)y^2$ (which is convex on $[-2, 2] \times \mathbb{R}$) is constant on $[-2, 2] \times \{0\}$ but is increasing on $]0, 2[\times \mathbb{R}^*$ in the direction $(1, 0)$.



$\Psi_1(x, y) = y^2$ is tubular



$\Psi_2(x, y) = (x^2 + 12)y^2$ is not tubular

The following proposition establishes some links between tubularity, quasianalyticity and tubularity of the epigraph.

Proposition 4.4. *i. If Φ is convex and quasianalytic,* then *it is tubular.*

ii. A closed convex set C is d-tubular if and only if its indicator function δ_C is d-tubular.

iii. If a continuous convex function Φ is d-tubular then its epigraph is $(d, 0)$-tubular. Conversely if the epigraph of a closed convex function Φ is $(d, 0)$-tubular then Φ is d-tubular.

iv. A closed proper convex function $\Phi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is tubular if and only if whenever Φ is constant on $[x, y]$, there exists a neighbourhood V of x and a positive ε such that

$$\forall z \in \operatorname{dom}(\partial\Phi) \cap V \ \ s.t.\ \ \Phi(z) \leq \Phi(x) + \varepsilon \quad \forall \xi \in \partial\Phi(z) \quad \langle \xi, y - x\rangle \leq 0.$$

v. For any $N \geq 2$, the continuous convex function $\Psi : x \mapsto \|x\|_{N,2}$ is tubular, but its epigraph is not tubular.

Proof of Proposition 4.4. *i.* Suppose that Φ is quasianalytic and that $x \in \mathbb{R}^N$ and $t > 0$ are such that Φ is finite and constant on $[x, x + td]$ for some $d \neq 0$. Then Φ is constant on the line $(x, x + d)$, and since it is convex, Φ is constant on any line $(z, z + d)$ for z in the domain of Φ.

ii. This is straightforward from the definitions.

iii. Let Φ be continuous and d-tubular. Let (x, r) in $\operatorname{epi}(\Phi)$ and $t > 0$ be such that $(x, r) + t(v, 0)$ belongs to $\operatorname{epi}(\Phi)$.

Suppose first that there exists a positive s such that $(x, r) + s(d, 0)$ is in the interior of $\operatorname{epi}(\Phi)$. Then there exists an open subset V of $\mathbb{R}^{N+1}$ such that $(x, r) + s(d, 0) \in V \subset \operatorname{epi}(\Phi)$. Therefore, for any (z, r') in the neighbourhood $(V - s(d, 0)) \cap \operatorname{epi}(\Phi)$ of (x, r) in $\operatorname{epi}(\Phi)$, one has $(z, r') + s(d, 0) \in V \subset \operatorname{epi}(\Phi)$. This shows that $\operatorname{epi}(\Phi)$ is $(d, 0)$-tubular in this case.

On the other hand, suppose that $[(x, r), (x, r) + t(d, 0)]$ is included in the boundary of $\operatorname{epi}(\Phi)$. Then this means that $r = \Phi(x)$ and that Φ is constant on the segment $[x, x + td]$. Since Φ is d-tubular, there exist a neighbourhood U of x and a positive ε such that for any z in U for which $\Phi(z) \leq \Phi(x) + \varepsilon$, the function $s \mapsto \Phi(z + sd)$ is nonincreasing on $[0, \varepsilon]$. Let U' be a neighbourhood of x such that $U' + sd \subset U$ for some positive s. Then for any (z, r') in the neighbourhood $U' \times]\Phi(x) - \varepsilon, \Phi(x) + \varepsilon[\cap \operatorname{epi}(\Phi)$ of (x, r) in $\operatorname{epi}(\Phi)$, $(z, r') + s(d, 0)$ belongs to $\operatorname{epi}(\Phi)$. This concludes the proof of the first part of the claim.

Suppose now that epigraph of Φ is $(d, 0)$-tubular for some $d \neq 0$. Let x and $t > 0$ be such that Φ is finite and constant on $[x, x + td]$. Then $(x, \Phi(x))$ and $(x + td, \Phi(x))$ belong to $\operatorname{epi}(\Phi)$, so there exists a neighbourhood V of x and a positive ε such that

$$\forall z \in V \ \text{ s.t. } \ \Phi(z) \in]\Phi(x) - \varepsilon, \Phi(x) + \varepsilon[\quad (z + \varepsilon d, \Phi(z)) \in \operatorname{epi}(\Phi).$$

Let $\eta > 0$ be such that $B(x, \eta) \subset V \cap \{y : \Phi(y) > \Phi(x) - \varepsilon\}$. Then it is easy to check that for every z in $B(x, \eta)$ such that $\Phi(z) \leq \Phi(x) + \varepsilon$, the function $s \mapsto \Phi(z + sd)$ is nonincreasing on $[0, \varepsilon]$.

iv. It is a straightforward consequence of Proposition 8.50 (equivalence between (d) and (f)) of [23].

v. The function Ψ is tubular since it is never constant on a nontrivial segment. To show that its epigraph is not tubular, we use characterization *iv* above to prove that its indicator function is not tubular. We are then led to show that there exist x and y in $\text{epi}(\Psi)$ such that for any neighbourhood V of x

$$\exists z \in \text{epi}(\Psi) \cap V \quad \exists \xi \in N_{\text{epi}(\Psi)}(z) \quad \langle \xi, y - x \rangle > 0.$$

Suppose that $N \geq 2$ and $x \neq 0$. We notice that $N_{\text{epi}(\Psi)}(x, \Psi(x)) = \mathbb{R}_+(\frac{x}{\|x\|_2}, -1)$. As the segment $[(x, \Psi(x)), (0,0)]$ is included in $\text{epi}(\Psi)$, one is led to consider the expression

$$\left\langle (\frac{z}{\|z\|_2}, -1), (-x, -\Psi(x)) \right\rangle = -\frac{1}{\|z\|_2} \langle z, x \rangle + \|x\|_{N,2}$$

for $z \neq 0$ in a neighbourhood of x. The Cauchy-Schwartz inequality implies that this is positive for any z which is not colinear to x. Such a z exists in any neighbourhood of x since $N \geq 2$, thus the epigraph of Ψ is not tubular. **Q.E.D.**

Remark. We deduce from Proposition 4.4 *ii.* and *iv.* the following characterization of tubular sets: a closed convex set C is tubular if and only if: whenever x and y belong to C, there exists a neighbourhood V of x such that

$$\forall z \in C \cap V \quad \forall \xi \in N_C(z) \quad \langle \xi, y - x \rangle \leq 0.$$

We now prove that in low dimension, every closed convex subset of $\mathbb{R}^N$ is tubular. Notice that the result below is sharp since proposition 4.4 provides an example of a non tubular convex subset of $\mathbb{R}^N$ for any N greater than 3.

Proposition 4.5. *Every closed convex subset of* $\mathbb{R}$ *and* $\mathbb{R}^2$ *(or of any vector space of dimension less than two) is tubular. As a consequence, any l.s.c. proper convex function* $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *is tubular.*

Proof. For $N = 1$, the proof is straightforward since a convex subset of $\mathbb{R}$ is an interval.

Let C be a closed convex subset of $\mathbb{R}^2$ and x, y belong to C. Without loss of generality, we may assume that $x = (0,0)$ and $y = (1,0)$. We first show that $C \cap \mathbb{R} \times [0, +\infty[$ is tubular.

If $C \cap \mathbb{R} \times]0, +\infty[$ is empty, there is nothing to prove (it reduces to the one-dimensional case). On the contrary, assume that some $z = (z_1, z_2)$ belongs to this set. Let δ denote the distance from x to the line (y, z) and set $\varepsilon = \min(\delta, z_2)/2$. We claim that for any v in $C \cap \mathbb{R} \times [0, +\infty[\cap B(x, \varepsilon)$, $v + \varepsilon(y - x)$ belongs to C. Indeed, we infer from the definition of ε that there exists $t_0 > \varepsilon$ such that $v + t_0(y - x)$ belongs to $[z, y]$. Since C is convex, $v + \varepsilon(y - x)$ belongs to C. This proves our claim.

The same arguments yield a positive η such that for any v in $C \cap \mathbb{R} \times] - \infty, 0] \cap B(x, \eta)$, $v + \eta(y - x)$ belongs to C.We now set $\gamma = \min(\varepsilon, \eta)$, then the neighbourhood $V = B(x, \gamma)$ of x and the positive γ are such that for any $z \in C \cap B(x, \gamma)$, $z + t(y - x)$ belongs to C. This completes the proof.

We may apply characterization *iii* of proposition 4.4 to get that any closed proper convex function $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is tubular, but we can also deduce this directly from the definition. Indeed, assume that Φ is constant and finite on $[x, x+td]$ (with $d \neq 0$ and $t > 0$). Without loss of generality, we may assume that $d > 0$, then since Φ is convex, one has

$$\forall s < r \leq t \qquad \frac{\Phi(x+rd) - \Phi(x+sd)}{(r-s)d} \leq \frac{\Phi(x+td) - \Phi(x)}{td} = 0$$

so that Φ is non increasing on $]-\infty, x+td]$. **Q.E.D.**

Remark. Notice that in the proof of the tubularity of a closed convex C subset of $\mathbb{R}^2$, the trick is to write C as the union of its intersection with the two half plans $\mathbb{R} \times [0, +\infty[$ and $\mathbb{R} \times]-\infty, 0]$. In $\mathbb{R}^3$ one would need infinitely many half plans, which is the reason why this proof can't be adapted to dimensions higher than 3.

It is noticed in *v.* of the proposition 4.4 that the euclidian norm $x \mapsto \|x\|_2$ is tubular. It may be checked from calculus, but it is also a consequence of the strict convexity of its sublevels, as the next proposition shows. We also prove below that the set of proper tubular convex functions is stable under composition by an increasing convex function.

Proposition 4.6. *i. If $\Phi : \mathbb{R}^N \to \mathbb{R} \cup +\infty$ is a closed proper convex function whose sublevel sets are (void or) strictly convex, then it is tubular.*

ii. If $\Phi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is a closed proper tubular convex function and $\theta : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is an increasing convex function such that $\theta \circ \Phi \not\equiv +\infty$, then $\theta \circ \Phi$ is tubular.

Proof. *i.* Suppose that $\Phi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is a closed proper convex function whose level sets are strictly convex, and assume that Φ is finite and constant over $[x, y]$, with $x \neq y$. As the set $\{z \in \mathbb{R}^N : \Phi(z) \leq \Phi(x) = \Phi(y)\}$ is strictly convex, there exists an open ball B centered at the middle $(x+y)/2$ of $[x, y]$ included in this set. We claim that the neighbourhood $B - (y-x)/2$ of x and $\varepsilon = 1/2$ have the desired property. To show this, we first notice that

$$\Phi(x) = \Phi(y) = \Phi(\frac{x+y}{2}) = \text{ Min}\,\{\Phi(z) : z \in \mathbb{R}^N\}.$$

Otherwise, there exists $z \in \mathbb{R}^N$ such that $\Phi(z) < \Phi(x)$. If we set $h(t) = \Phi(tz + (1-t)(x+y)/2)$, then h is convex and we have

$$h(1) < h(0) \qquad \text{and} \qquad h(t) \leq h(0) = \Phi(x)$$

for any negative t such that $tz + (1-t)(x+y)/2$ belongs to B. This obviously contradicts the convexity of h.

Let z belong to $B - (y-x)/2$, then the function $s \mapsto \Phi(z + s(y-x))$ is convex on $[0, \varepsilon]$ and attains its minimum at ε, so that it is nonincreasing on $[0, \varepsilon]$. This concludes the proof.

ii. Assume that $\Phi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is a closed proper tubular convex function and $\theta : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is an increasing convex function. If $\theta \circ \Phi$ is finite and constant on a segment $[x, x+td]$, then since θ is increasing, Φ is also constant on $[x, x+td]$. As a consequence, there exists a neighbourhood V of x and a positive ε such that for any $z \in V$ the function $s \mapsto \Phi(z+sd)$ is nonincreasing on $[0, \varepsilon]$ whenever $\Phi(z) \leq \Phi(x) + \varepsilon$. Since θ is increasing, $s \mapsto \theta \circ \Phi(z+sd)$ is also nonincreasing on $[0, \varepsilon]$ for any z in V such that $\theta \circ \Phi(z) \leq \theta(\varepsilon + \Phi(x))$. Therefore $\theta \circ \Phi$ is tubular. **Q.E.D.**

The following proposition, provides different ways to build tubular sets and functions.

Proposition 4.7. *i. If* $C_1, \ldots, C_M$ *are closed tubular convex subsets of* $\mathbb{R}^N$ then $\bigcap_{i=1}^{M} C_i$ *is tubular.*

ii. If $\Phi_1, \ldots, \Phi_M$ *are continuous tubular convex functions* then $\sup_{1 \leq i \leq M} \Phi_i$ *is tubular.*

iii. Let $\Phi_1, \ldots, \Phi_M : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *be l.s.c. proper convex functions.* Then *the l.s.c. proper convex function* $(x_1, \ldots, x_M) \longmapsto \sum_{i=1}^{M} \Phi_i(x_i)$ *is tubular on* $\mathbb{R}^M$.

Proof. i. This is a simple consequence of definition 4.2.

ii. Let $\Phi_1, \ldots, \Phi_M$ be continuous convex functions and set $\Phi = \sup_i \Phi_i$, then $\mathrm{epi}(\Phi) = \bigcap_i \mathrm{epi}(\Phi_i)$. ¿From *iii.* of proposition 4.4, we know that each $\mathrm{epi}(\Phi_i)$ is $(d, 0)$-tubular for every $d \neq 0$. It is easily checked that this also holds true for the finite intersection $\bigcap_i \mathrm{epi}(\Phi_i)$. Applying proposition 4.4 again yields that Φ is tubular.

iii. We restrain ourselves to the case $M = 2$, the proof being easily adapted to the general case. We infer from proposition 4.5 that as Φ_1 and Φ_2 are proper l.s.c. convex functions on $\mathbb{R}$, they are tubular.

We set $\Psi(x, y) = \Phi_1(x) + \Phi_2(y)$. Assume that Ψ is constant on $[(x, y), (x, y) + t(d_1, d_2)]$ for some positive t. We must find a neighbourhood V of (x, y) and a positive ε such that for any $z \in V$ the function $s \mapsto \Psi(z + s(d_1, d_2))$ is nonincreasing on $[0, \varepsilon]$ whenever $\Psi(z) \leq \Psi(x, y) + \varepsilon$. If $d_1 = d_2 = 0$, there is nothing to prove.

Suppose first that $d_1 = 0$ or $d_2 = 0$. Without loss of generality, we assume that the first holds. Then Φ_2 is constant on $[y, y + td_2]$, so that there exists a neighbourhood V of y and a positive ε such that $s \mapsto \Phi_2(z + sd_2)$ is nonincreasing on $[0, \varepsilon]$ whenever $z \in V$ and $\Phi_2(z) \leq \Phi_2(y) + \varepsilon$. Recalling the definition of Ψ, it is easy to check that for any (z_1, z_2) in the neighbourhood $\{x' : \Phi_1(x') > \Phi_1(x) - \varepsilon/2\} \times V$ of (x, y) and such that $\Psi(z_1, z_2) \leq \Psi(x, y) + \varepsilon/2$, the function $s \mapsto \Psi(z + s(0, d_2))$ is non increasing on $[0, \varepsilon/2]$. This concludes the proof in this case.

Now, suppose that $d_1 \neq 0$ and $d_2 \neq 0$. Without loss of generality, we may assume that both d_1 and d_2 are positive. Since $\Phi_1(x + std_1) = \Psi(x, y) - \Phi_2(y + std_2)$ for s in $[0, 1]$, we deduce that Φ_1 is linear on $[x, x + td_1]$. By symetry, Φ_2 is also linear on $[y, y + td_2]$. As a consequence, Ψ is constant on any segment included in $[x, x+td_1] \times [y, y+td_2]$ with direction (d_1, d_2). We claim that the neighbourhood $V =]x - \frac{d_1}{4}, x + \frac{d_1}{4}[\times]y - \frac{d_2}{4}, y + \frac{d_2}{4}[$ of (x, y) and $\varepsilon = 1/2$ have the desired property. Indeed, let (z_1, z_2) belong to V, then the convex function $s \mapsto \Psi((z_1, z_2) + s(d_1, d_2))$ is constant on $[1/4, 3/4]$. Reasonning as in the proof of proposition 4.5, we infer that this function is non increasing on $]-\infty, 1/2]$. This completes the proof. **Q.E.D.**

Corollary 4.8. *Any closed convex polyhedron as well as any convex piecewise affine function is tubular. Moreover, every finite sup of convex analytic functions on $\mathbb{R}^N$ is tubular.*

Remark. The notion of tubularity is not stable with respect to the addition. Indeed, the functions $\Phi_1(x, y) = (x^2 + 12)y^2 + y$ (which is never constant on a nontrivial segment of $\mathbb{R}^2$) and $\Phi_2(x, y) = -y$ are convex and tubular on $[-2, 2] \times \mathbb{R}$, but their sum is no more tubular (see the example after definition 4.3).

5. Asymptotic convergence to the θ-center

5.1. **The convergence result.** We turn to the main result of this paper, which is a generalization of previous similar results in [6] and [8] on the selection of a particular solution of (CP_0) by a penalty method.

Theorem 5.1. *Assume that (H_0), ..., (H_4) hold, and that the function Φ_i is tubular for any i in $\{0, \dots, M\}$. Then the net $(x_r)_{r>0}$ of the optimal solutions of the penalized problems (CP_r) converges as r tends to 0 towards the unique θ-center of (CP_0).*

Remark. To use this selection result, it would be of practical interest to be able to associate to any given function A^m (having the properties of non-linear averages described in definition 3.1) a penalty function θ such that $A^m = A^m_\theta$. It is still an open problem whether this is possible in the general case. For example, for the function $A^m(x) := -(\frac{1}{m}\sum_{i=1}^m(-x_i)^p)^{\frac{1}{p}}$ (where $p \in]0, 1[$), one may choose $\theta_p(t) := -(-t)^p$.

Before proving theorem 5.1, we illustrate it with two examples. In the case of the Logarithmic penalty method, MacLinden [18] shows the convergence of the penalty trajectory towards the analytic center (i.e. the θ-center when $\theta = \theta_1$ is the log penalty) under the strict complementarity assumption. In the following example, this assumption does not hold while theorem 5.1 ensures the convergence of the penalty trajectory.

Example. Set $\Phi_0(x, y) = x^2 + \max\{0, -1 - y\}$, $\Phi_1(x, y) = x$ and $\Phi_2(x, y) = y$. Then $(H_0) - (H_4)$ are fullfilled (with $\theta = \theta_1$ is the Log penalty function) and the functions

Φ_i are tubular (see lemma 4.7) so that the penalty trajectory converges to the unique θ-center $(0,-1)$. However, every optimal solution of problem (CP_0) is of the form $(0,\alpha)$ for α in $[-1,0]$, and the unique solution of the dual problem is $(0,0)$, so that the strict complementarity assumption does not hold.

When $(H_0)-(H_4)$ hold, the penalty trajectory may converge to an optimal solution different from the θ-center if the functions Φ_i are not tubular (as the following example shows), or may not converge at all (see §5.3).

Example. Set $\Phi_0(x,y)=(x^2+3)y^2+\delta_C(x,y)$ with $C=[-1,1]\times\mathbb{R}$, $\Phi_1(x,y)=x-6$ and $\Phi_2(x,y)=y$. Then hypotheses $(H_0)-(H_4)$ are fullfilled (with $\theta=\theta_1$), but Φ_0 is not tubular on C. Let (x_r,y_r) denote the unique solution of (CP_r), then for $r>0$ sufficiently small, (x_r,y_r) belongs to the interior of C and the optimality conditions read

$$2x_r y_r^2=\frac{r}{x_r-6}\quad \textit{and}\quad 2y_r(x_r^2+3)=\frac{r}{y_r}.$$

As a consequence, $(x_r,y_r)_r$ converges to $(-\frac{1}{2},0)$, whereas the θ-center is $(-1,0)$.

Theorem 5.1 is in fact a straightforward consequence of the following selection result. Theorem 5.2 below characterizes those optimal solutions of (CP_0) which can be obtained as limit points (as $r\to 0$) of nets $(x_r)_r$ of optimal approximate solutions. Notice that when (H_4) does not hold, there may be several θ-centers, so the following theorem does not imply the convergence of the penalty trajectories.

Theorem 5.2. *Suppose that (H_0), (H_1) and (H_3) hold. Also assume that the function Φ_i is tubular for any i in $\{0,\dots,M\}$. Let $x_0\in S(CP_0)$ be a cluster point of $(x_r)_{r>0}$ as r goes to 0, where $x_r\in S(CP_r)$ for all r.* Then *x_0 is a θ-center of (CP_0).*

We shall need the following lemma in the proof of theorem 5.2.

Lemma 5.3. *Let x_0 and x^* belong to $S(CP_0)$, then there exists $t\in]0,1[$ such that for any i in $I=\{1\le i\le M\,:\,\Phi_i(x_0)=\Phi_i(x^t)\}$ the function Φ_i is constant on $[x_0,x^t]$, where we have set $x^t:=tx_0+(1-t)x^*$.*

Proof. For $1\le i\le M$, we set $s_i=\max\{s\in[0,1/2]\,:\,\Phi_i(x_0)=\Phi_i(x^s)\}$, where x^s denotes $sx_0+(1-s)x^*$. Notice that if $s\in]0,s_i[$ is such that $\Phi_i(x^s)=\Phi_i(x_0)$, then Φ_i is constant on $]0,s_i[$.

If for every $1\le i\le M$, one has $s_i=0$, then $t=1/4$ has the desired property. Indeed, $I=\{1\le i\le M\,:\,\Phi_i(x_0)=\Phi_i(x^t)\}$ is empty for this choice of t.

Otherwise, we set $t=\dfrac{1}{2}\min\{s_i\,:\,s_i>0\}$. Then if i belongs to $I=\{1\le i\le M\,:\,\Phi_i(x_0)=\Phi_i(x^t)\}$, the function Φ_i is constant on $]x_0,x^t[$ since $t\in]0,s_i[$. This concludes the proof. **Q.E.D.**

Proof of theorem 5.2. Let $x_0 \in S(\text{CP}_0)$ be a cluster point of $(x_r)_{r>0}$ as r goes to 0, where $x_r \in S(\text{CP}_r)$ for all r. To simplify the notations, we assume that the whole net converges to x_0 as r tends to 0. Let x^* belong to $S(\text{CP}_0)$ and $I \subset \{1, \dots, M\}$ such that for all i in I, $\Phi_i(x_0) = \Phi_i(x^*)$. Then we must check that

$$A_\theta^{M-|I|}((\Phi_j(x_0))_{j\notin I}) \leq A_\theta^{M-|I|}((\Phi_j(x^*))_{j\notin I}).$$

We apply lemma 5.3 to get a real $t \in]0,1[$ for which the function Φ_j is constant on $[x_0, x^t]$ for every $j \in J = \{1 \leq i \leq M \,:\, \Phi_i(x_0) = \Phi_i(x^t)\}$ where we have set $x^t = tx_0 + (1-t)x^*$. Notice that as x_0 and x^t both belong to $S(\text{CP}_0)$, the function Φ_0 is constant on the segment $[x_0, x^t]$. The optimality condition for x_r reads

$$0 \in \partial\Phi_0(x_r) + \alpha(r) \sum_{i=1}^{M} \partial \left(\theta \left(\frac{\Phi_i(.)}{r} \right) \right) (x_r).$$

Since the functions Φ_0 and $\theta(\Phi_j(.)/r)$ are tubular (see lemma 4.6) and constant over $[x_0, x^t]$ for j in J, we infer from proposition 4.4 *iv* that for r small enough,

$$\sum_{i\notin J} \langle \xi_i^r, x^t - x_0 \rangle \geq 0$$

for some vectors ξ_i^r in $\partial(\theta(\Phi_i(.)/r))(x_r)$. Using the convexity of the functions $\theta(\Phi_i(.)/r)$, we deduce from the previous inequality that

$$\sum_{i\notin J} \theta\left(\frac{\Phi_i(x_r)}{r}\right) \leq \sum_{i\notin J} \theta\left(\frac{\Phi_i(x_r + x^t - x_0)}{r}\right).$$

We notice that for i in $I \setminus J$, we have $\Phi_i(x^t) < \Phi_i(x^*) = \Phi_i(x_0)$. As a consequence, for r small enough, we deduce from the strict monotonicity of θ and the continuity of the functions Φ_i that

$$\sum_{i\notin J, i\notin I} \theta\left(\frac{\Phi_i(x_r)}{r}\right) \leq \sum_{i\notin J, i\notin I} \theta\left(\frac{\Phi_i(x_r + x^t - x_0)}{r}\right).$$

We can now choose real numbers δ_i and δ_i^t such that $\delta_i < \Phi(x_0) \leq 0$ and $\Phi_i(x^t) < \delta_i^t < 0$ for $i \notin (I \cup J)$. Indeed, for $i \notin (I \cup J)$, $\Phi_i(x^t)$ belongs to $]-\infty, \max\{\Phi_i(x_0), \Phi_i(x^*)\}[$, so that $\Phi_i(x^t) < 0$. For $i \in J \setminus I$, we notice that $\Phi_i(x_0) = \Phi_i(x^t) \leq \Phi_i(x^*)$ and we choose $\delta_i = \delta_i^t < \Phi_i(x_0)$. With these notations, we conclude from the monotonicity of θ that

$$\sum_{i\notin I} \theta\left(\frac{\delta_i}{r}\right) \leq \sum_{i\notin I} \theta\left(\frac{\delta_i^t}{r}\right).$$

Notice that δ and δ^t both belong to $]-\infty, 0[^{M-|I|}$. We thus divide the above inequality by $M - |I|$, compose it by the increasing function θ^{-1}, then divide by r and let r go to 0 (that's where (H_3) is needed). This leads to

$$A_\theta^{M-|I|}(\delta) \leq A_\theta^{M-|I|}(\delta^t)$$

Then, letting δ_i (resp. δ_i^t) go to $\Phi_i(x_0)$ (resp. $\Phi_i(x^t)$), we get

$$A_\theta^{M-|I|}((\Phi_j(x_0))_{j\notin I}) \leq A_\theta^{M-|I|}((\Phi_j(x^t))_{j\notin I}).$$

As $A_\theta^{M-|I|}$ is convex and componentwise non-decreasing, and since $x^t = tx_0 + (1-t)x^*$ with $0 < t < 1$, this proves our claim. **Q.E.D.**

5.2. **Extensions to other penalty methods.** In the proof of the selection result theorem 5.2, the hypotheses (H_0) and (H_1) are mainly assumed in order to ensure that the conclusion of theorem 2.1 holds. To be more precise, in theorem 5.2 we can take the following hypotheses (H_0') and (H_1') instead of (H_0) and (H_1), with

$$(H_0') \qquad \begin{cases} \Phi_0 : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\} \text{ is a closed proper convex function,} \\ \text{for } 1 \leq i \leq M,\ \Phi_i : \mathbb{R}^N \to \mathbb{R} \text{ are continuous convex functions.} \end{cases}$$

$$(H_1') \qquad \begin{cases} \theta \text{ is increasing and convex on } \operatorname{dom}(\theta) =]-\infty, \eta[\,,\ \eta \in [0,+\infty], \\ \alpha \text{ is positive on }]0,+\infty[\text{ and the conclusions of theorem 2.1 hold.} \end{cases}$$

Notice that the boundedness of a selection $(x_r)_r$ of approximate solutions $x_r \in S(\mathrm{CP}_r)$ as well as the optimality of any cluster point x_0 of such a family is contained in (H_1'), so that $S(\mathrm{CP}_0)$ is non empty. We can then get the following convergence result, which is a simple extension of theorem 5.1 to this setting.

Theorem 5.4. *Assume that (H_0'), (H_1'), $(H_2), \ldots, (H_4)$ hold, and that the function Φ_i is tubular for any i in $\{0, \ldots, M\}$. Then the net $(x_r)_{r>0}$ of the optimal solutions of the penalized problems (CP_r) converges as r tends to 0 towards the unique θ-center of (CP_0).*

For example, the above result applies to the non linear algorithm studied in [7]. This algorithm is based on the penalty scheme (CP_r) and generates a bounded sequence of approximate optimal solutions (x_r) whose cluster points are optimal solutions of (CP_0). For this algorithm, hypothesis (H_1) is not satisfied (because it is associated with a function α such that $\alpha(r)/r$ is bounded near 0), whereas hypothesis (H_1') follows from lemma 6 therein.

5.3. **A non convergence result.** It is possible, under some further hypotheses on the penalty function θ, to build a continuous convex function $\Phi_0 : \mathbb{R}^2 \to \mathbb{R}$ such that (H_0) holds with the affine constraints $\Phi_1(x) = x_1 - 1$ and $\Phi_2(x) = x_2 - 1$ and for which the net $(x_r)_{r>0}$ of the optimal solutions of $(CP_r)_{r>0}$ does not converge as r goes to 0. Notice that for Φ_1 and Φ_2 defined as above, the hypothesis (H_2) is clearly satisfied. Our example is defined as follows: Φ_0 is given as the supremum of a denumerable family of affine functions ϕ_n, which corresponds to defining its epigraph as a denumerable intersection of half spaces. The difficulty is to define ϕ_n so that it is associated to a point x^n and a real number r_n

for which $x^n = x_{r_n}$ (the optimal solution of (CP_{r_n})), $\phi_n = \Phi_0$ in a neighbourhood of x^n, the net x^nhas at least to cluster points and r_n goes to 0 as $n \to +\infty$. Our construction requires the following hypothesis on θ:

$$(H_5) \qquad \begin{cases} \theta \text{ is differentiable on }]-\infty, K[\text{ for some negative } K \\ \text{and } t \mapsto |t|\theta'(t) \text{ is non decreasing on }]-\infty, K[. \end{cases}$$

The hypothesis of differentiability on θ is not really restrictive since most penalty functions studied in the litterature (and in particular any such function cited in [6]) are at least of class C^1 on their domain. The monotonicity of $t \mapsto |t|\theta'(t)$ is more technical, but simple calculations show that the penalty functions θ_1, θ_2 and θ_3 of the preceding sections have this property. As a consequence, theorem 5.5 applies to the Exponential Penalty Method and the Logarithmic Barrier Method. We also refer to a recent work by Gilbert *et al.* [14], where the particular case of the Logarithmic Barrier Method is considered: they give an example of a $\mathcal{C}^\infty$-smooth function Φ_0 such that for the constraint $\Phi_1(x_1, x_2) := x_2 \geq 0$, the penalty trajectory does not converge as r goes to 0. Notice that such a behaviour is impossible for an analytic function Φ_0, since analyticity implies tubularity.

Theorem 5.5. *Assume that (H_5) holds, and that (H_1) is valid with $\alpha(r) = r$. Then there exists a continuous convex function $\Phi : \mathbb{R}^2 \to \mathbb{R}$ such that the net $(x_r)_{r>0}$ of the optimal solutions of the family of problems $(CP_r)_{r>0}$ with*

$$(\mathrm{CP}_r) \qquad \mathrm{Inf}\left\{\Phi(x_1, x_2) + r\theta\left(\frac{x_1 - 1}{r}\right) + r\theta\left(\frac{x_2 - 1}{r}\right) \ : \ (x_1, x_2) \in \mathbb{R}^2\right\}$$

has at least two cluster points as r goes to 0.

Proof. We shall define Φ through its epigraph $\mathrm{epi}(\Phi)$. We first assume that there exist two sequences $(a^n)_{n\geq 1}$ and $(r_n)_{n\geq 1}$ of elements of $]0, 1/3]\times]0, 1[\times]1, +\infty[$ and $]0, +\infty[$ respectively such that

(i) $(r_n)_n$ is decreasing and $\lim_{n\to+\infty} r_n = 0$,

(ii) $\forall n \geq 1 \quad a_1^{2n} = \frac{1}{3} - \frac{1}{4n}$ and $a_1^{2n+1} = \frac{1}{4n}$, $\quad a_2^n < a_2^{n+1} < 1$, $\quad \frac{a_2^n - 1}{r_n} \leq \frac{3K}{2}$ and $a_3^n > a_3^{n+1} > 1$,

(iii) $\forall n \geq 1 \quad \forall k \neq n \quad a^k \in D(a^n, r_n)$.

where $D(x, r)$ is the open subset of $\mathbb{R}^3$ given by

$$D(x, r) = \left\{z \in \mathbb{R}^3 \ : \ z_3 > x_3 - \theta'\left(\frac{x_1 - 1}{r}\right)(z_1 - x_1) - \theta'\left(\frac{x_2 - 1}{r}\right)(z_2 - x_2)\right\}$$

when (x_1, x_2) belongs to $]-\infty, 1[^2$ and $r > 0$ are such that $\max\left\{\frac{x_1-1}{r}, \frac{x_2-1}{r}\right\} < K$.

Let $\mathcal{C} := \bigcap_{n\geq 1} \overline{D(a^n, r_n)}$, then $\mathcal{C}$ is a closed convex subset of $\mathbb{R}^3$ and it is the epigraph of a continuous convex function $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Indeed, let $x = (x_1, x_2) \in \mathbb{R}^2$, we must show that there exists $\phi(x)$ in $\mathbb{R}$ such that $\{(x_1,x_2,x_3) : x_3 \geq \phi(x)\} = \mathcal{C} \cap \{(x_1,x_2,x_3) : x_3 \in \mathbb{R}\}$. ¿From the definition of $\mathcal{C}$ we infer that if $(x_1,x_2,x_3) \in \mathcal{C}$ then $\{(x_1,x_2,y) : y \geq x_3\} \subset \mathcal{C}$. Moreover, we deduce from (ii) and (H_5) that for all $n \geq 1$

$$a_3^n - \theta'\left(\frac{a_1^n - 1}{r_n}\right)(x_1 - a_1^n) - \theta'\left(\frac{a_2^n - 1}{r_n}\right)(x_2 - a_2^n) \leq |a_3^n| + \theta'\left(\frac{3K}{2}\right)[|x_1 - a_1^n| + |x_2 - a_2^n|]$$

and as $(a^n)_n$ is bounded, there exists x_3 in $\mathbb{R}$ such that (x_1,x_2,x_3) belongs to $\mathcal{C}$, and we can thus define $\phi(x) := \min\{x_3 : (x_1,x_2,x_3) \in \mathcal{C}\}$.

¿From the definition of $\mathcal{C}$, we have that for any $n \geq 1$, $\phi(a_1^n, a_2^n) = a_3^n$ and

$$\left(-\theta'\left(\frac{a_1^n - 1}{r}\right), -\theta'\left(\frac{a_2^n - 1}{r}\right)\right) \in \partial\phi(a_1^n, a_2^n).$$

Let $m = \min\{\phi(x) : x \in [-1,2]^2\}$, then we claim that the function $\Phi : (x_1,x_2) \mapsto \max\{\phi(x_1,x_2), x_1^2 + x_2^2 - 5 + m\}$ has the desired properties. Indeed, Φ is continuous, convex and coercive on $\mathbb{R}^2$, so that (H_0) is satisfied. Moreover, Φ is equal to ϕ on $[-1,2]^2$, so we infer from the previous remarks that for any $n \geq 1$, the optimal solution x_{r_n} of (CP_{r_n}) is (a_1^n, a_2^n). As a consequence, the net $(x_r)_{r>0}$ has at least two cluster points as r goes to 0, namely $(0,l)$ and $(\frac{1}{3}, l)$, where $l = \lim_{n\to+\infty} a_2^n$. This proves the claim.

It remains to prove that there exist two sequences $(a^n)_{n\geq 1}$ and $(r_n)_{n\geq 1}$ for which $(i)\ldots(iii)$ hold. We build such sequences by induction on n.

Since θ is convex and increasing, θ' is positive on $]-\infty, K[$. We set $a^1 = (\frac{1}{3}, \frac{1}{2}, 1+\varepsilon)$ with $\varepsilon = \frac{1}{9}\theta'(2K)$ (notice that ε is positive) and $r_1 = -\frac{1}{3K}$. Then $(0,1,1)$ and $(1,1,1)$ belong to $D(a^1, r_1)$.

Now fix some $n \in \mathbb{N} \setminus \{0\}$ and assume that $a^1, \ldots, a^n$ belong to $]0,1/3]\times]0,1[\times]1,+\infty[$ and $0 < r_n < \ldots < r_1$ are such that (ii) and (iii) hold up to n. Then we claim that for $r \in]0, \min\{\frac{1}{n+1}, r_n\}[$ small enough, one has

1) $a_2^n < a(r)_2 < 1, \quad \dfrac{a(r)_2 - 1}{r} \leq \dfrac{3K}{2}$ and $a_3^n > a(r)_3 > 1\,,$
2) $\forall i \in \{1,\ldots,n\} \quad a(r) \in D(a^i, r_i) \quad$ and $\quad a^i \in D(a(r), r)\,,$
3) $(0,1,1) \in D(a(r), r) \quad$ and $\quad (1,1,1) \in D(a(r), r)\,.$

where $a(r) = (\beta, 1-\sqrt{r}, 1+\delta(r))$, with $\beta := 1/3 - 1/(2(n+1))$ if n is odd, $\beta := 1/(2n)$ otherwise and $\delta(r) = \frac{1}{12}\theta'\left(-\frac{1}{2r}\right)$. Notice that since $\theta_\infty(-1) = 0$, the limit of θ' at $-\infty$ is equal to 0, so that $\delta(r)$ tends to 0 as r decreases to 0. Let us check that if $r > 0$ is small enough, then $a = a(r)$ and r satisfy properties *1)*, *2)* and *3)*.

1) It is easily seen that $a(r)$ belongs to $]0,1/3[\times]0,1[\times]1,+\infty[$ for r small enough. Moreover, $a(r)$ and $\dfrac{a(r)_2-1}{r}$ respectively tend to $(\beta,1,1)$ and $-\infty$ as r goes to 0, so that for r small enough, condition *1)* is fullfilled for $a=a(r)$.

2) The set $\bigcap_{i=1}^{n} D(a^i,r_i)$ is a convex open subset of $\mathbb{R}^3$ containing the segment $[(0,1,1),(1,1,1)]$.

This implies that $a(r)$ belongs to $\bigcap_{i=1}^{n} D(a^i,r_i)$ for r small enough, which is the first part of condition *2)*. The second part of condition 2) for $i\in\{1,\dots,n\}$ reads

$$a_3^i > 1+\delta(r)-\theta'\left(\frac{\beta-1}{r}\right)(a_1^i-\beta)-\theta'(-\frac{1}{\sqrt{r}})(a_2^i-1+\sqrt{r})$$

$$\iff f(r)=a_3^i-1-\delta(r)+\theta'\left(\frac{\beta-1}{r}\right)(a_1^i-\beta)+\theta'(-\frac{1}{\sqrt{r}})(a_2^i-1+\sqrt{r})>0$$

Since $\beta<1$ and θ' tends to 0 at $-\infty$, we have $\lim_{r\to 0^+} f(r)=a_3^i-1>0$. As a consequence, condition 2) is satisfied for $r>0$ small enough.

3) It is sufficient to check that $(0,1,1)\in D(a(r),r)$. This amounts to

$$-\delta(r)-\beta\theta'\left(\frac{\beta-1}{r}\right)+\theta'(-\frac{1}{\sqrt{r}})\sqrt{r}>0.$$

Since θ' is non-decreasing on $]-\infty,K[$ and β belongs to $]0,1/3[$, $4\delta(r)$ is greater than $\beta\theta'\left(\frac{\beta-1}{r}\right)$. It is therefore sufficient to check that

$$-5\delta(r)+\theta'(-\frac{1}{\sqrt{r}})\sqrt{r}=-\frac{5}{12}\theta'\left(-\frac{1}{2r}\right)+\theta'(-\frac{1}{\sqrt{r}})\sqrt{r}>0.$$

We infer from (H_5) that

$$\frac{1}{\sqrt{r}}\theta'\left(-\frac{1}{\sqrt{r}}\right)\geq\frac{1}{2r}\theta'\left(-\frac{1}{2r}\right)>\frac{5}{12r}\theta'\left(-\frac{1}{2r}\right)$$

as soon as $0<r<\min\{\frac{1}{4},\frac{1}{K^2}\}$.

As a consequence, if we set $a^{n+1}:=a(r)$ and $r_{n+1}=r$ for r small enough, the families $(a^k)_{1\leq k\leq n+1}$ and $(r_k)_{1\leq k\leq n+1}$ satisfy *(ii)* and *(iii)* and $r_{n+1}\leq\frac{1}{n+1}$.

The induction on n thus yield two sequences $(a^n)_n$ and $(r_n)_n$ which satisfy *(i)*...*(iii)*. This concludes the proof. **Q.E.D.**

References

[1] F. Alvarez, *Absolute minimizer in convex programming by exponential penalty*, J. Conv. Anal., 7 (2000), pp. 197-202.

[2] G. Aronsson, *Extension of functions satisfying Lipschitz conditions*, Ark. Math., 6 (1967), pp. 551-561.

[3] G. Aronsson, *On the partial differential equation* $u_x^2u_{xx}+2u_xu_yu_{xy}+u_y^2u_{yy}=0$, Ark. Math., 7 (1968), pp. 395-425.

[4] H. Attouch, *Viscosity solutions of minimization problems*, SIAM J. Optim., 6 (3) (1996), pp. 769-806.

[5] H. Attouch and R. Cominetti, L^p *Approximation of variational problems in* L^1 *and* L^∞, Nonlinear Analysis, 36 (1999), pp. 373-399.

[6] A. Auslender, R. Cominetti and M. Haddou, *Asymptotic analysis for penalty and barrier methods in convex and linear programming*, Math. Oper. Res., 22 (1997) 1, pp. 43-62.

[7] R. Castillo and C. Gonzaga, *A nonlinear programming algorithm based on non-coercive penalty functions*, to appear in Math. Programming.

[8] R. Cominetti, *Nonlinear averages and convergence of penalty trajectories in convex programming*, Proceeding of the Workshop on Ill-posed Variational Problems and Regularization Techniques, Trier (1998), L.N. in Economics and Math. Systems 477, Springer-Verlag (1999), pp. 65-78.

[9] R. Cominetti and J. San Martin, *Asymptotic analysis for the exponential penalty trajectory in linear programming*, Math. Programming, 67 (1994), pp. 169-187.

[10] G. Dal Maso and L. Modica, *On the convergence of local minima*, Boll. Un. Mat. Ital. A (6), 1 (1982), pp. 55-61.

[11] J. Descloux, *Approximation in L^p and Chebyshev approximation*, J. Soc. Ind. Appl. Math., 11 (1963), pp. 1017-1026.

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

[13] A. Egger and R. Huotari, *The Polya algorithm on convex sets*, J. Approx. Theory, 56 (1989), pp. 212-216.

[14] J.Ch. Gilbert, C. Gonzaga and E. Karas, *Examples of ill-behaved central paths in convex optimization*, Rapport de Recherche INRIA (2000), submitted to Mathematical Programming.

[15] R. Huotari, *Tubular sets and multivariate Polya Algorithm*, J. Austrl. Math. Soc.(series A), 55 (1993), pp. 232-237.

[16] R. Huotari and W. Li, *Continuity of Metric Projection, Polya Algorithm, Strict Best Approximation and Tubularity of Convex Sets*, J. Math. Anal. Appl., 182 (1994), pp. 836-856.

[17] R. Huotari and M. Marano, *The Polya algorithm on tubular sets*, J. Comp. Appl. Math., 54 (1994), pp. 151-157.

[18] L. MacLinden, *An analogue of Moreau's proximation theorem with applications to the nonlinear complementary problem*, Pacific. J. Math., 88 (1980), pp. 101-161.

[19] M. Marano, *Strict approximation on closed convex sets*, Approx. Theory Appl., 6 (1990), pp. 99-109.

[20] R. Monteiro and F. Zhou, *On the existence and convergence of the central path for convex programming and some duality results*, Comput. Optim. Appl. , 10 (1998), pp. 51-77.

[21] G. Polya, *Sur un algorithme toujours convergent pour obtenir les polynômes de meilleure approximation de Tchebycheff*, C. R. Acad. Sci. Paris, 157 (1913), pp. 840-843.

[22] J. Rice, *Tchebycheff approximation in a compact metric space*, Bull. Amer. Math. Soc., 68 (1962), pp. 405-410.

[23] R. T. Rockafellar and R. J.- B. Wets, Variational Analysis, Springer-Verlag, Berlin (1998).

[24] G. Sonnevend, *An 'analytic center' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, Lecture Notes in Control and Information Sciences, Springer-Verlag, New York (1985), pp. 866-876.