Mathematical Programming 67 (1994) 225–245

Optimality conditions in mathematical programming and composite optimization

Jean-Paul Penot

University of Pau – URA 1204 CNRS, Laboratoire de Mathématiques Appliquées (I.P.R.A.), Av. de l'Université, 64000 Pau, France

Received 2 October 1990; revised manuscript received 24 March 1994

Abstract

New second order optimality conditions for mathematical programming problems and for the minimization of composite functions are presented. They are derived from a general second order Fermat's rule for the minimization of a function over an arbitrary subset of a Banach space. The necessary conditions are more accurate than the recent results of Kawasaki (1988) and Cominetti (1989); but, more importantly, in the finite dimensional case they are twinned with sufficient conditions which differ by the replacement of an inequality by a strict inequality. We point out the equivalence of the mathematical programming problem with the problem of minimizing a composite function. Our conditions are especially important when one deals with functional constraints. When the cone defining the constraints is polyhedral we recover the classical conditions of Ben-Tal–Zowe (1982) and Cominetti (1990).

Keywords: Composite functions; Compound derivatives; Compound tangent sets; Fermat rule; Mathematical programming; Multipliers; Optimality conditions; Second order conditions

The art of mathematical programming is intimately tied with the question of optimality conditions. Since most algorithms yield critical points (or Kuhn–Tucker points) rather than local minimizers, the exact meaning of the optimality conditions which describe such points is crucial, even for numerical needs. From the theoretical point of view it is largely recognized that the successive enlargements of mathematical programming (from equality constraints to equality and inequality constraints [13,14], from a finite number of such constraints [1,2,7–9,16,20,42–43,45], to constraints defined by a cone or an inclusion [3–6,10–12,17–25,28–38,44,46]) have lead not only to a better understanding of the optimality conditions but also to a much wider field of applications. In particular semi-infinite programming, calcul of variations with constraints, optimal control with distributed parameters are amenable to the framework that we choose ([28–30,44] for instance). The problem we treat is of the form:

SSDI 0025-5610(94)00030-W

$$(\mathscr{M}) \qquad \text{minimize} \quad j(x) \quad \text{subject to } x \in B \cap k^{-1}(C)\ ,$$

where $j: X \to \mathbb{R}$, $k: X \to Z$ are twice differentiable at some point a of $F := B \cap k^{-1}(C)$, B and C being closed convex subsets of the Banach spaces X and Z respectively.

Our necessary conditions involve multipliers depending not only on first order tangent directions as in [4] for instance, but also on second order tangent like vectors. They refine conditions given recently by Kawasaki [21] and Cominetti [10,11] in this sense that they are more exacting: the new "strange" term we introduce is greater than theirs. More importantly, to these necessary conditions can be associated sufficient conditions (in the finite dimensional case) which differ only by the replacement of an inequality by a strict inequality.

Such a desirable feature justifies the introduction of a new geometrical concept of tangent set of order two which is taylored to fit the data. We call it the compound tangent set of order two. The introduction of this set has been motivated by some expressions in the work [19] of A.D. Ioffe which show the way to find adapted tools. It is clear that a close study of this set has to be done in each particular application. Let us observe that such a notion has analytical counterparts and is also needed for the study of composite functions of the form $f = g \circ h$ with h twice differentiable and g convex lower semicontinuous (l.s.c.) (but not necessarily finite).

There is no great surprise in this fact since we show that the problem $(\mathscr{M})$ and the problem

$$(\mathscr{C}) \qquad \text{minimize} \quad f = g \circ h\ ,$$

with g and h as above, are equivalent problems provided suitable rewritings are performed. Moreover the necessary conditions on one hand and the sufficient conditions on the other hand for each of these two problems correspond exactly under these rewritings. This last fact is not a consequence of the previous equivalence since even in finite dimensions a gap remains between necessary conditions and sufficient conditions: as usual an inequality has to be replaced by a strict inequality. Thus none of the two approaches is superior to the other one and one has the following diagram showing the implications between necessary conditions and sufficient conditions for problems $(\mathscr{M})$ and $(\mathscr{C})$:

$$\begin{array}{ccc} \mathrm{CS}(\mathscr{M}) & \Leftrightarrow & \mathrm{CS}(\mathscr{C}) \\ \downarrow & & \downarrow \\ \text{solution of } (\mathscr{M}) & \Leftrightarrow & \text{solution of } (\mathscr{C}) \\ \downarrow & & \downarrow \\ \mathrm{CN}(\mathscr{M}) & \Leftrightarrow & \mathrm{CN}(\mathscr{C}) \end{array}$$

It seems that up to now the equivalence between problems ($\mathscr{M}$) and ($\mathscr{C}$) was not clearly recognized, although the reduction procedure of [18] and exact penalization [16] are well-known links between the two problems (but require qualification conditions and involve quite different tools). Our use of indicator functions is crucial in this matter and justifies our attention to functions g taking the value $+\infty$. This application to problem ($\mathscr{C}$) is treated in Section 4.

The first section deals with an abstract Fermat's rule for a general minimization problem with constraints. Here too the striking new feature is the sufficient condition and its strong analogy with the necessary condition. The second section is devoted to our geometrical tool, the compound tangent set and to its analytical counterpart. We believe that these notions have their own interests and can be used in other fields such as sensitivity analysis and optimal control theory.

The main results are given in Section 3 under a familiar constraint qualification condition (see [6,22,27,34,35,40,46]) extending the famous Mangasarian–Fromovitz constraint qualification condition (see [16] for instance) to the infinite dimensional framework of ($\mathscr{M}$).

Applications to semi-infinite programming and sensitivity analysis will be given elsewhere.

1. Second order Fermat's rule

Let us first consider the problem

$$(\mathscr{P}) \qquad \text{minimize} \quad f(x): \quad x \in F$$

of minimizing a function over an arbitrary feasible subset F of a normed vector space (n.v.s.) X. In fact here X could be an arbitrary topological vector space. We suppose f is twice differentiable at some point $a \in F$ in the following broad sense: there exist a continuous linear map $f'(a): X \to \mathbb{R}$ and a continuous bilinear map $f''(a): X \times X \to \mathbb{R}$ such that

$$r(x) := f(a+x) - f(a) - f'(a)x - \tfrac{1}{2} f''(a)xx$$

satisfies $\lim_{(t,u) \to (0,v)} t^{-2} r(tu) = 0$ for each $v \in X$.

The (superior) *tangent cone* to F at $a \in F$ (or contingent cone) is the set

$$F'(a) := \limsup_{t \to 0+} t^{-1}(F - a) .$$

The (superior) *second order tangent set* to F at $a \in F$ in the direction $v \in X$ is the set

$$F''(a, v) := \limsup_{t \to 0+} 2t^{-2}(F - a - tv) .$$

Thus $F'(a) = F''(a, 0)$ and $F''(a, v)$ is the set of limits of sequences (w_n) such that for some sequence (t_n) of $\mathbb{P} := (0, \infty)$ with limit 0 one has for each $n \in \mathbb{N}$ $a + t_n v + \tfrac{1}{2} t_n^2 w_n \in F$.

These sets are usually larger than the inferior tangent sets (or *incident sets*)

$$T^{\mathrm{i}}(F, a) := \liminf_{t\to 0+} t^{-1}(F-a)$$

and

$$T^{\mathrm{ii}}(F, a, v) := \liminf_{t\to 0+} 2t^{-2}(F-a-tv)$$

respectively. When F is convex $F'(a) = T^{\mathrm{i}}(F, a)$ is convex but $F''(a, v)$ may be strictly larger than $T^{\mathrm{ii}}(F, a, v)$. This fact happens frequently with subsets of functional spaces. But it already appears in two dimensional situations as the following example shows. On the other hand it must be added, because it is important for a comparison with the work of R. Cominetti [11], that the computation of $F''(a, v)$ is often more difficult than the computation of $T^{\mathrm{ii}}(F, a, v)$ (see [21] for instance).

1.1 Example. Let (r_n) be a sequence of $\mathbb{P} := (0, +\infty)$ such that for some $c \in (0, 1)$, $r_{n+1} \leqslant c r_n$ for each n and let $F = \{(x_1, x_2) \in \mathbb{R}^2 : \forall n \in \mathbb{N}, x_2 \geqslant (r_n + r_{n+1})x_1 - r_n r_{n+1}\}$. Then one can check that for $a = (0, 0)$, $v = (1, 0)$, $w = (0, 2)$ one has $v \in F'(a)$, $w \in F''(a, v)$ but $w \notin T^{\mathrm{ii}}(F, a, v)$.

1.2 Theorem. *Suppose $f: X \to \mathbb{R}$ is twice differentiable at a and that f attains on $F \subset X$ a local minimum at a. Then*

(a) $f'(a)v \geqslant 0$ *for each* $v \in F'(a)$;

(b) $$\tfrac{1}{2} f''(a)vv + \liminf_{\substack{(t,u)\to(0+,v)\\ a+tu\in F}} f'(a)t^{-1}(u-v) \geqslant 0$$

for each $v \in F'(a) \cap \ker f'(a)$.

Introducing the set $S_{v,F}$ of $r \in \mathbb{R}$ such that there exists a sequence (t_n, v_n, r_n) in $\mathbb{P} \times X \times \mathbb{R}$ (where $\mathbb{P} = (0, \infty)$) with limit $(0, v, r)$, satisfying $a + t_n v_n \in F$, $f(a) + f'(a) t_n v_n + \tfrac{1}{2} t_n^2 r_n \leqslant f(a)$ for each $n \in \mathbb{N}$, condition (b) can be written

(b′) $$f''(a)vv \geqslant \sup S_{v,F}.$$

Let us give another interpretation of the preceding conditions using the second order epi-derivative of the indicator function i_F of F (where $i_F(x) = 0$ for $x \in F$, $i_F(x) = +\infty$ for $x \in X \setminus F$) given by

$$i''_F(a, a^*, v) = \liminf_{(t,u)\to(0+,v)} 2t^{-2}[i_F(a+tu) - i_F(a) - \langle a^*, tu\rangle]$$

for $a^* \in X^*$, $v \in X$. Condition (a) can be written $-f'(a) \in \partial i_F(a)$ for $\partial i_F(a) := N(F, a) := F'(a)^0$, and condition (b) reads $f''(a)vv + i''_F(a, a^*, v) \geq 0$ for $a^* = -f'(a)$ or $f''_F(a, 0, v) \geq 0$ for $f_F = f + i_F$, a very natural condition.

The proof we give below does not use the (easily established) calculus rule $f''_F(a, 0, v) = f''(a)vv + i''_F(a, -f'(a), v)$ and the optimality condition $f''_F(a, 0, v) \geq 0$. It relies on a simple direct argument.

Proof. The first assertion is immediate and well-known. The second one is a consequence of the relation $t^{-1}f'(a)v = 0$ and of the following inequalities valid for $t \in \mathbb{P} := (0, \infty)$ and $u - v$ small enough with $a + tu \in F$:

$$0 \leq t^{-2}(f(a+tu) - f(a)) = t^{-1}f'(a)u + \tfrac{1}{2}f''(a)uu + \epsilon(u, t)$$

with $\epsilon(t, u) \to 0$ as $t \to 0_+$, $u \to v$. □

The following statement is an immediate consequence.

1.3 Corollary. *Suppose $f: X \to \mathbb{R}$ is twice differentiable at a and that f attains on $F \subset X$ a local minimum at a. Then*

(a) $f'(a)v \geq 0$ *for each* $v \in F'(a)$,

(b) $f''(a)vv + f'(a)w \geq 0$ *for any* $v \in F'(a) \cap \ker f'(a)$, *and any* $w \in F''(a, v)$.

A similar result was given in [10,11] with $F'(a)$, $F''(a, v)$ replaced by the smaller sets $T^{i}(F, a)$, $T^{ii}(F, a, v)$, so that our conditions are more exacting, even when F is convex. For instance, let F be as in Example 1.1, with $r_n = c^n$, $c \in (0, 1)$ and let $f(x_1, x_2) = x_2 - bx_1^2$ with $b \in \mathbb{R}$. For $a = (0, 0)$, $v = (1, 0)$, $w = (w_1, w_2)$ we have $w \in F''(a, v)$ whenever $w_2 \geq 2$ but $w \notin T^2(F, a, v)$ when $w_2 < 4$. Therefore the necessary condition of Corollary 1.3 yields $-2b + w_2 \geq 0$ for $w_2 \geq 2$, hence $b \leq 1$ and this condition is sufficient as $F \subset \{(x_1, x_2): x_2 - x_1^2 \geq 0\}$ whereas the corresponding condition with $F''(a, v)$ replaced by $T^{ii}(F, a, v)$ cannot give more than $b \leq 2$ and the condition is not sufficient.

This corollary seems to be easier to apply than Theorem 1.2 but it rules out situations in which Theorem 1.2 allows to conclude, as the following example shows. Moreover Theorem 1.2 is closer to the sufficient condition which follows.

1.4 Example. Let $F = \{(x_1, x_2) \in \mathbb{R}^2: x_1^6 = x_2^4\}$, $X = \mathbb{R}^2$, $a = 0$, so that $F'(a) = \mathbb{R} \times \{0\}$. The first order optimality condition is $D_1f(0, 0) = 0$. Since $F''(a, 0) = \mathbb{R} \times \{0\}$ and $F''(a, v)$ is empty for each $v \in F'(a) \setminus \{0\}$ the second order optimality condition of the preceding corollary does not give anything else than $D_1f(0, 0) = 0$, whereas Theorem 1.2 yields $D_1f(0, 0) = 0$, $D_2f(0, 0) = 0$ and $D_1^2f(0, 0) \geq 0$. Let us note that the conditions $Df(0, 0) = 0$, $D_1^2f(0, 0) > 0$ are sufficient to ensure that $(0, 0)$ is a local minimizer.

The preceding example suggests the following corollary.

1.5 Corollary. *Suppose $f: X \to \mathbb{R}$ is twice differentiable at a and attains on F a local minimum at a. Then for each $v \in F'(a)$ with $f'(a)v = 0$, for each $s \in]0, 1[$ and for each w in*

$$T^{1,s}(F, a, v) = \{w \in X: \exists (t_i, w_i) \to (0_+, w): a + t_i v + t_i^{1+s} w_i \in F, \forall i \in \mathbb{N}\}$$

one has $f'(a)w \geqslant 0$.

Proof. In fact if we had $f'(a)w < 0$ we would have

$$\tfrac{1}{2} f''(a)vv + \liminf_i f'(a) t_i^{-1}(v + t_i^s w_i - v) = -\infty . \qquad \square$$

1.6 Corollary. *Suppose f attains on F a local minimum at a and suppose $f'(a) = 0$, which is the case when the closed convex hull $\overline{\text{co}}(F'(a))$ of $F'(a)$ is X. Then for each $v \in F'(a)$ one has $f''(a)vv \geqslant 0$.*

Now let us turn to sufficient optimality conditions.

1.7 Theorem. *Suppose X is finite dimensional. Suppose $f: X \to \mathbb{R}$ is twice differentiable at a and that*

(a) *$f'(a)v \geqslant 0$ for each $v \in F'(a)$;*

(b) $$\tfrac{1}{2} f''(a)vv + \liminf_{\substack{(t,u) \to (0_+, v) \\ a + tu \in F}} f'(a) t^{-1}(u - v) > 0$$

for each $v \in F'(a) \setminus \{0\} \cap \ker f'(a)$.
Then f attains on F a strict local minimum at a.

Proof. Suppose on the contrary there exists a sequence (a_n) in F with $\lim a_n = a$, $f(a_n) \leqslant f(a)$; $a_n \neq a$ for each n. Let $t_n = \|a_n - a\|$, $v_n = t_n^{-1}(a_n - a)$. Without loss of generality we may suppose that the sequence (v_n) of unit vectors of X has a limit v. Then we get $f'(a)v \leqslant 0$, $v \in F'(a)$ and by (a), $f'(a)v = 0$.

Moreover

$$\liminf_n \, [t_n^{-1} f'(a) v_n + \tfrac{1}{2} f''(a) v_n v_n] = \liminf_n \, t_n^{-2} [f(a + t_n v_n) - f(a)] \leqslant 0 ,$$

a contradiction with our assumption, as $f''(a)v_n v_n \to f''(a)vv$ and $t_n^{-1} f'(a) v_n = f'(a) t_n^{-1}(v_n - v)$. $\square$

Although the preceding condition is extremely simple, it seems that it has not yet been pointed out. Its usefulness depends on the possibility of couching the required inequality in simple terms associated with the problem at hand. We will turn to such a task for the mathematical programming problem after introducing adapted tools.

2. Compound tangent sets and compound derivatives of order two

It is common sense to observe that when a problem has some structure, it is fruitful to use tools which are adapted to this structure, even when they appear to be complex at the first glance. The ones we introduce here are inspired by some expressions in [19]: they take into account the occurrence of two spaces and of a mapping between them.

2.1 Definition. Let X, Y be n.v.s., let D and E be subsets of X and Y respectively, let $h:X\to Y$ be a mapping which is differentiable at $\bar{x}\in D$, with $\bar{y}:=h(\bar{x})\in E$. Then for $v\in D'(\bar{x})$ the second order compound tangent set to E at $(\bar{y}, v)$ (with respect to h and D) is the set

$$E''_{h,D}(\bar{y}, v) = \limsup_{(t,u)\to^{D}(0+,v)} 2t^{-2}(E-\bar{y}-th'(\bar{x})u) .$$

In other words, $w\in E''_{h,D}(\bar{y}, v)$ iff there exists a sequence $((t_n, v_n, w_n))$ with limit $(0, v, w)$ such that $\bar{x}+t_nv_n\in D$ and $\bar{y}+t_nh'(\bar{x})v_n+\frac{1}{2}t_n^2w_n\in E$ for each $n\in\mathbb{N}$. Denoting by $\mathbb{S}$ the set of sequences of $\mathbb{P}:=(0,\infty)$ with limit 0 and, for $s:=(s_n)\in\mathbb{S}$, $v\in X$, denoting by $D(s, v)$ the set of sequences $\vec{v}:=(v_n)$ with limit v such that $\bar{x}+s_nv_n\in D$ for each $n\in\mathbb{N}$, we have $w\in E''_{h,D}(\bar{y}, v)$ iff there exist $s:=(s_n)\in\mathbb{S}$, $\vec{v}=(v_n)\in D(s, v)$ with

$$w\in E''_{h,D,s}(\bar{y}, \vec{v}) := \liminf_n 2s_n^{-2}(E-\bar{y}-s_nh'(\bar{x})v_n) .$$

We set

$$E''_{h,D,s}(\bar{y}, v) := \bigcup\{E''_{h,D,s}(\bar{y}, \vec{v}) \colon \vec{v}\in D(s, v)\}.$$

We observe that these definitions (which contain some abuse of notations) depend on $\bar{x}$ rather than on $\bar{y}$ and on the derivative $h'(\bar{x})$ of h at $\bar{x}$ and not on h itself. When there is no risk of confusion we simplify this notation by dropping some of its elements. In particular when $D=X$ we omit X.

The (second order) compound tangent set $E''_{h,D}(\bar{y}, v)$ is always larger than $E''(\bar{y}, h'(\bar{x})v)$ when $\bar{x}+tv\in D$ for $t>0$ small enough. When the rank of $h'(\bar{x})$ is low they coincide, as the following lemma shows.

2.2 Lemma. *When $h'(\bar{x})(X)\subseteq\mathbb{R}h'(\bar{x})(v)$ one has $E''_h(\bar{y}, v)=E''(\bar{y}, h'(\bar{x})v)$.*

Proof. The assertion is obvious if $h'(\bar{x})=0$. Suppose $h'(\bar{x})(X)=\mathbb{R}h'(\bar{x})(v)$ with $h'(\bar{x})(v)\neq\{0\}$. Let $w\in E''_h(\bar{y}, v)$ and let $s:=(s_n)\in\mathbb{S}$, $(v_n)\in D(s, v)$, $(w_n)\to w$ be such that $\bar{y}+s_nh'(\bar{x})v_n+\frac{1}{2}s_n^2w_n\in E$ for each $n\in\mathbb{N}$. We can write $h'(\bar{x})v_n=r_nh'(\bar{x})v$ for some $r_n\in\mathbb{R}$ and (r_n) must converge to 1. Then for $t_n:=r_ns_n$ we have $(r_n^{-2}w_n)\to w$ and $\bar{y}+t_nr_n^{-1}h'(\bar{x})v+\frac{1}{2}t_n^2r_n^{-2}w_n\in E$ for each $n\in\mathbb{N}$ so that $w\in E''(\bar{y}, v)$. □

The following example shows that $E''_h(\bar{y}, v)$ may be much larger than $E''(\bar{y}, h'(\bar{x})v)$; this fact will be crucial for the optimality conditions we have in view.

2.3 Example. Suppose E, $\bar{y}$, v are such that, for some $r>0$ and some sequences (t_n), (v_n) with limits 0_+ and v respectively the ball $B(\bar{y}+t_nh'(\bar{x})v_n, \frac{1}{2}rt_n^2)$ with center $\bar{y}+t_nh'(\bar{x})v_n$ and radius $\frac{1}{2}rt_n^2$ is contained in E for n large enough. Then $E''_h(\bar{y}, v)$ contains the ball $B(0, r)$. However $E''(\bar{y}, v)$ may be empty.

Let us take for instance $X=\mathbb{R}^2$, $Y=\mathbb{R}^3$, $\bar{x}=0$, $\bar{y}=0$, $v=(1, 0)$ and let us define h and E by $h(x_1, x_2)=(x_1, x_2, 0)$, $E=\{(y_1, y_2, y_3): y_1\geqslant 0, y_2\geqslant f(y_1)\}$, where $f(t)=t^{\alpha+1}$ with $\alpha\in(0, 1)$. Let $\beta\in(0, \alpha)$. Given any $(t_n)\in\mathbb{S}$ let us set $v_n:=(1, t_n^\beta)$. Then, for each $r>0$, $B(\bar{y}+t_nh'(\bar{x})v_n, \frac{1}{2}rt_n^2)\subseteq E$ for n large enough, so that $E''_h(\bar{y}, v)=Y$. However $E''(\bar{y}, v)$ is empty.

Another case in which $E''_h(\bar{y}, v)$ is large is given in the following proposition.

2.4 Proposition. *Suppose $h'(\bar{x})$ is an open mapping from X onto Y. Then if $E''_h(y, v)$ is nonempty it is the whole space Y.*

Proof. Let $w\in E''_h(\bar{y}, v)$ and let $s:=(s_n)\in\mathbb{S}$, $(v_n)\to v$, $(w_n)\to w$ be such that $\bar{y}+s_nh'(\bar{x})v_n+\frac{1}{2}s_n^2w_n\in E$ for each $n\in\mathbb{N}$. Given $z\in Y$ let $u\in X$ be such that $h'(\bar{x})u=w-z$. Since $h'(\bar{x})$ is open we can find a sequence (u_n) with limit u such that $h'(\bar{x})u_n=w_n-z$. Then $v'_n:=v_n+\frac{1}{2}s_nu_n$ converges to v and

$$\bar{y}+s_nh'(\bar{x})v'_n+\tfrac{1}{2}s_n^2z=\bar{y}+s_nh'(\bar{x})v_n+\tfrac{1}{2}s_n^2w_n\in E$$

for each $n\in\mathbb{N}$ so that $z\in E''_h(\bar{y}, v)$. □

Let us now consider convexity issues which are important for the sequel. We remark that when D and E are convex the set $E''_{h,D}(\bar{y}, v)$ is not always convex. However we can select a convex subset by choosing a fixed sequence, as the following lemma shows. Its proof is an easy consequence of the definitions.

2.5 Lemma. *When D and E are convex, for each $v\in D'(\bar{x})$ and each $s\in\mathbb{S}$ the set*

$$E''_{h,D,s}(\bar{y}, v):=\bigcup\{E''_{h,D,s}(\bar{y}, \bar{v}): \bar{v}\in D(s, v)\}$$

is convex.

When $Y=Z\times\mathbb{R}$ and E is the epigraph of some function $\varphi: Z\to\mathbb{R}^\bullet:=\mathbb{R}\cup\{\infty\}$ we observe that $E''_{h,D}(\bar{y}, v)$ and $E''_{h,D,s}(\bar{y}, v)$ are epigraphs because they are closed and stable by addition of elements of $\{0\}\times\mathbb{R}_+$. This leads us to give the following definition close to a proposal in [19].

2.6 Definition. Given $g: Y\to\mathbb{R}^\bullet\cup\{0\}$, $h: X\to Y$ differentiable at some $\bar{x}\in D\subset X$, $v\in D'(\bar{x})$, $\bar{y}:=h(\bar{x})\in Y$ with $g(\bar{y})<\infty$, $\bar{y}^*\in Y^*$, $w\in Y$, $A:=h'(\bar{x})$, the compound second order derivative of g at $\bar{y}$, $\bar{y}^*$, v, w with respect to h and D is defined by:

$$g''_{h,D}(\bar{y}, \bar{y}^*, v, w) = \liminf_{\substack{(t,u,y) \to (0_+, v, w) \\ \bar{x}+tu \in D}} 2t^{-2}[g(\bar{y}+tAu+\tfrac{1}{2}t^2y) - g(\bar{y}) - \langle \bar{y}^*, tAu+\tfrac{1}{2}t^2y\rangle] .$$

When $D=X$ we write g''_h instead of $g''_{h,X}$. When $D=x+\mathbb{R}v$, or $x+\mathbb{R}_+v$ one sees easily that

$$g''_{h,D}(\bar{y}, \bar{y}^*, v, w) = g''(\bar{y}, \bar{y}^*, h'(\bar{x})v, w) ,$$

where the mixed second derivative of g is given by

$$g''(\bar{y}, \bar{y}^*, z, w) = \liminf_{(t,y) \to (0_+, w)} 2t^{-2}[g(\bar{y}+tz+\tfrac{1}{2}t^2y) - g(\bar{y}) - \langle \bar{y}^*, tz+\tfrac{1}{2}t^2y\rangle] ,$$

what somewhat justifies our notation. The preceding equality also occurs when $h'(\bar{x}) = 0$ or when $h'(\bar{x})$ has rank one with $h'(\bar{x})v \neq 0$. On the other hand, when $A := h'(\bar{x})$ is an open mapping from X onto Y we have

$$g''_h(\bar{y}, \bar{y}^*, v, w) = g''_h(\bar{y}, \bar{y}^*, v, 0)$$

for any w in Y, a property which is similar to what occurs when g is twice differentiable at $\bar{y}$, in which case we have, with $\bar{y}^* = g'(\bar{y})$,

$$g''_h(\bar{y}, \bar{y}^*, v, w) = g''(\bar{y}) \cdot h'(\bar{x})v \cdot h'(\bar{x})v = g''(\bar{y}, \bar{y}^*, Av) ,$$

what justifies our choice. Let us observe that we have $g''_{h,D}(\bar{y}, \bar{y}^*, v, w) \geq 0$ whenever g is convex and $\bar{y}^* \in \partial g(\bar{y})$, a desirable feature for a second order derivative. The analogy of the proofs of some of the preceding assertions with the proofs of Lemma 2.2 and Proposition 2.4 stems from the relationships between the geometrical and the analytical concepts we introduced. This analogy is strengthened by the following two results. The proof of the first one is immediate.

2.7 Lemma. *If g is the indicator function i_E of E (given by $i_E(y) = 0$ if $y \in E$, $+\infty$ otherwise) then $g''_{h,D}(\bar{y}, 0, v, \cdot)$ is the indicator function of $E''_{h,D}(\bar{y}, v)$.*

2.8 Lemma. *If $E \subset Y := Z \times \mathbb{R}$ is the epigraph of some $g : Z \to \mathbb{R} \cup \{\infty\}$, if $h := (k, j) : X \to Z \times \mathbb{R}$ with $j'(\bar{x}) = \bar{z}^* \circ k'(\bar{x})$ for some $\bar{z}^* \in Z^*$ then for $\bar{y} := h(\bar{x})$, $\bar{z} := k(\bar{x})$, $v \in X$ the set $E''_h(\bar{y}, v)$ is the epigraph of $g''_k(\bar{z}, \bar{z}^*, v, \cdot) + \langle \bar{z}^*, \cdot \rangle$. In particular, if $h = (k, 0)$ then $E''_h(\bar{y}, v)$ is the epigraph of $g''_k(\bar{z}, 0, v, \cdot)$.*

Proof. One has $(w, r) \in E''_h(\bar{y}, v)$ iff for some sequences (s_n), (v_n), (w_n), (r_n) with limits 0_+, v, w, r respectively one has

$$g(\bar{z}+s_n h'(\bar{x})v_n + \tfrac{1}{2}s_n^2 w_n) \leq g(\bar{y}) + s_n j'(\bar{x})v_n + \tfrac{1}{2}s_n^2 r_n$$

iff

$$g''_k(\bar{z}, \bar{z}^*, v, w) + \langle \bar{z}^*, w \rangle \leq r . \qquad \square$$

3. Mathematical programming under a constraint qualification condition

In the sequel we tackle problems $(\mathscr{M})$ and $(\mathscr{C})$, with

$$(\mathscr{M}) \qquad \text{minimize} \quad j(x) \qquad \text{subject to } x \in B \cap k^{-1}(C),$$

$$(\mathscr{C}) \qquad \text{minimize} \quad f(x) := g(h(x)) \qquad \text{subject to } x \in D,$$

where $j: X \to \mathbb{R}$, $k: X \to Z$ are twice differentiable at some point a of $F := B \cap k^{-1}(C)$, B and C being closed convex subsets of the Banach spaces X and Z respectively and where $h: X \to Y$ is twice differentiable, $g: Y \to \mathbb{R} \cup \{+\infty\}$ is a closed proper convex function and D is a closed convex subset of X.

3.1 Theorem. *Problems $(\mathscr{M})$ and $(\mathscr{C})$ are equivalent.*

Proof. Let us first observe that $(\mathscr{M})$ can be rewritten in the form of $(\mathscr{C})$. Setting $D = B$, $Y = Z \times \mathbb{R}$, $h(x) = (k(x), j(x))$, $g(z, r) = i_C(z) + r$, where i_C is the indicator function of C, we have $g(h(x)) = j(x) + i_C(k(x))$ for $x \in X$, so that the objective functions of $(\mathscr{M})$ and $(\mathscr{C})$ coincide and the two problems have the same sets of solutions.

Conversely, problem $(\mathscr{C})$ can be transformed into the form of problem $(\mathscr{M})$. To see this, let us set

$$Z = Y \times Y \times \mathbb{R}, \qquad \hat{X} = X \times Y \times \mathbb{R}, \qquad B = D \times Y \times \mathbb{R}, \qquad C = \{0\} \times E_g,$$

where $E_g = \{(y, r) \in Y \times \mathbb{R}: r \geqslant g(y)\}$ is the epigraph of g and let $j: X \times Y \times \mathbb{R} \to \mathbb{R}$, $k: X \times Y \times \mathbb{R} \to Z$ be given by

$$j(x, y, r) = r,$$

$$k(x, y, r) = (h(x) - y, y, r).$$

Then $(x, y, r) \in \hat{X}$ is feasible for $(\mathscr{M})$ iff $x \in D$, $y = h(x)$, $r \geqslant g(h(x))$, so that

$$\inf\{ j(x, y, r) : (x, y, r) \in B \cap k^{-1}(C)\} = \inf\{g(h(x)) : x \in D\}$$

and $(\mathscr{C})$ and $(\mathscr{M})$ have the same values. Moreover $\bar{x}$ is a solution to $(\mathscr{C})$ iff $a := (\bar{x}, \bar{y}, \bar{r})$, with $\bar{y} := h(\bar{x})$, $\bar{r} := f(\bar{x})$ is a solution to $(\mathscr{M})$. □

To treat problem $(\mathscr{M})$ it will be convenient to introduce the following notations. Given a solution a of $(\mathscr{M})$ we denote by $L(a)$ the set of Lagrange–Karush–Kuhn–Tucker multipliers for $(\mathscr{M})$ at a:

$$L(a) = \{z^* \in Z^*: z^* \in N_{k(a)}C,\ 0 \in j'(a) + z^* \circ k'(a) + N_a B\}$$

where $N_a B$ $(= N(B, a)) = B'(a)^0 = \{x^* \in X^*: \forall v \in B'(a)\ \langle x^*, v\rangle \leqslant 0\}$ is the normal cone to B at a and $N_{k(a)}C$ has a similar meaning. Let us abbreviate the notation for the compound tangent set to $E := C\times(-\infty, j(a)]$ at $((k(a), j(a)), v)$ w.r.t. $h := (k, j)$ and B into:

$$S_{v,B} := \limsup_{\substack{(t,u)\to(0+,v)\\ a+tu\in B}} 2t^{-2}[C\times(j(a)-\mathbb{R}_+)-(k,j)(a)-t(k,j)'(a)(u)]\ .$$

We remark that the set $S_{v,F}$ introduced after Theorem 1.2 corresponds to the case $j=f$, $B=F$, $C=\{0\}$, $k=0$, identifying $\{0\}\times\mathbb{R}$ with $\mathbb{R}$.

The following lemma is crucial for the treatment of problem $(\mathscr{M})$. In it we use the following notion. The mapping $k: X\to Z$ is said to be *metrically regular* at a with respect to (w.r.t.) B and C if there exist $c>0$ and a neighborhood U of a such that

$$d(x, B\cap k^{-1}(C))\leqslant c d(k(x), C)\quad \text{for each } x\in U\cap B\ .$$

3.2 Lemma. *Suppose k is metrically regular at a w.r.t. B and C. Let $u\in B'(a)$, $v\in B'(a)\cap k'(a)^{-1}(C'(k(a)))$ with $j'(a)v=0$ and let $(z, r)\in S_{v,B}$ be such that $w := k''(a)vv+k'(a)u-z\in C'(k(a))$. Then*

$$j''(a)vv+j'(a)u-r\geqslant 0\ .$$

Proof. Let $(t_n, v_n, (z_n, r_n))_{n\in\mathbb{N}}$ be a sequence in $\mathbb{P}\times X\times(Z\times\mathbb{R})$ with limit $(0, v, (z, r))$ as in Definition 2.1, with $D := B$, $h := (k, j)$: for each $n\in\mathbb{N}$ we have $a+t_nv_n\in B$, $h(a)+t_nh'(a)v_n+\frac{1}{2}t_n^2(z_n, r_n)\in E$. As $u\in B'(a)=T^{\mathrm{i}}(B, a)$, $w\in C'(k(a))=T^{\mathrm{i}}(C, k'(a))$ we can find sequences (u_n), (w_n) in X and Z with limits u and w respectively such that $a+\frac{1}{2}t_nu_n\in B$ and $k(a)+\frac{1}{2}t_nw_n\in C$ for each $n\in\mathbb{N}$. We can write

$$w_n=k''(a)v_nv_n+k'(a)u_n-z_n'$$

for some sequence (z_n') with limit z. By convexity of B and C we have, for $v_n' := (1-t_n)v_n+\frac{1}{2}t_nu_n$,

$$a+t_nv_n'=(1-t_n)(a+t_nv_n)+t_n(a+\tfrac{1}{2}t_nu_n)\in B\ ,$$

$$\begin{aligned}c_n&=k(a)+t_nk'(a)v_n'+\tfrac{1}{2}t_n^2(k''(a)v_nv_n+(1-t_n)z_n-z_n')\\&=(1-t_n)(k(a)+t_nk'(a)v_n+\tfrac{1}{2}t_n^2z_n)+t_n(k(a)+\tfrac{1}{2}t_nw_n)\in C\ .\end{aligned}$$

Then, for some constant $c>0$, we have for each n large enough

$$\begin{aligned}&d(a+t_nv_n', B\cap k^{-1}(C))\leqslant cd(k(a+t_nv_n'), C)\leqslant c\|c_n-k(a+t_nv_n')\|\\&\leqslant c\|k(a)+t_nk'(a)v_n'+\tfrac{1}{2}t_n^2k''(a)v_nv_n-k(a+t_nv_n')\|+\tfrac{1}{2}\|t_n^2((1-t_n)z_n-z_n')\|\\&\leqslant \epsilon(t_n)t_n^2\end{aligned}$$

for some $\epsilon(\cdot)$ with $\epsilon(t)\to 0$ as $t\to 0$ since $(v_n)\to v$, $(v_n')\to v$, $(z_n)\to z$, $(z_n')\to z$. It follows that there exists a sequence (v_n'') with $t_n^{-1}\|v_n''-v_n'\|\leqslant 2\epsilon(t_n)$ such that $a+t_nv_n''\in B\cap k^{-1}(C)$ for each $n\in\mathbb{N}$. Then by Theorem 1.2 and the inclusion $j(a)+t_nj'(a)v_n+\frac{1}{2}t_n^2r_n\in(-\infty, j(a)]$ we get

$$0 \leqslant j''(a)vv + 2 \liminf_n t_n^{-1} j'(a)(v_n'' - v)$$

$$\leqslant j''(a)vv + 2 \liminf_n t_n^{-1} j'(a) v_n'$$

$$\leqslant j''(a)vv + 2 \liminf_n (1-t_n) t_n^{-1} j'(a) v_n + \lim_n j'(a) u_n$$

$$\leqslant j''(a)vv - r + j'(a)u . \qquad \square$$

We need a form of the Lagrange multiplier rule which is close to the classical statement of [26] but uses a weaker qualification condition. It could be derived from [5] but we prefer to present a proof based on the following form of the famous Farkas lemma.

3.3 Lemma ([22], [34] Theorem 2.3, [35] Corollary 2.4, [46]). *Let P and Q be closed convex cones of the Banach spaces X and Z respectively and let $A: X \to Z, f: X \to \mathbb{R}$ be linear and continuous such that $f(x) \geqslant 0$ for each $x \in P \cap A^{-1}(Q)$. Then if $Z_0 = A(P) - Q$ is a closed vector subspace of Z there exists some $z^* \in Q^0$ such that*

$$0 \in f + z^* \circ A + P^0 .$$

In fact this statement can easily be reduced to the usual case $Z = A(P) - Q$. The following form of the Lagrange multiplier theorem is adapted to our needs.

3.4 Corollary (Lagrange multiplier theorem). *With the data of the preceding lemma suppose that for some $m \in \mathbb{R}$ and some $b \in Z_0$ one has*

$$f(x) \geqslant m \quad \text{for } x \in R := \{x \in P: Ax - b \in Q\} .$$

Then there exists $z^* \in Q^0$ such that for each $x \in P$

$$f(x) + \langle z^*, Ax - b \rangle \geqslant m .$$

Proof. Since $b \in A(P) - Q$, R is nonempty. Moreover for $x_0 \in R$ and any $v \in R_\infty := P \cap A^{-1}(Q)$ we have $x_0 + v \in R$ so that $f(v) \geqslant 0$, R_∞ being a cone and f being bounded from below by $m - f(x_0)$ on R_∞. Now for any $(x, t) \in P \times (0, \infty)$, such that $Ax - bt \in Q$ we have $t^{-1}x \in R$ hence $f(t^{-1}x) \geqslant m$ and $f(x) - mt \geqslant 0$. Therefore $f(x) - mt \geqslant 0$ for any $(x, t) \in P \times \mathbb{R}_+$ such that $Ax - bt \in Q$. Since $A(P) - \mathbb{R}_+ b - Q = Z_0$ Lemma 3.3 yields some $z^* \in Q^0$, $x^* \in P^0$, $r \in \mathbb{R}_+$ such that

$$f(x) - mt + \langle z^*, Ax - bt \rangle + \langle x^*, x \rangle = rt$$

for each $(x, t) \in P \times \mathbb{R}_+$. Taking $t = 1$ we get the announced inequality. $\square$

Let us observe that the preceding inequality implies that $f + z^* \circ A \in -P^0$ as this linear functional is bounded below on P.

3.5 Theorem. *Let a be a local solution to problem* ($\mathscr{M}$). *Suppose k is metrically regular at a w.r.t. B and C and*

$$\text{(R)} \qquad k'(a)B'(a) - C'(k(a)) = Z .$$

Then the set $L(a)$ *of Lagrange–Karush–Kuhn–Tucker multipliers is nonempty and for each* v *in the critical cone*

$$K(a) := \{v \in B'(a) : k'(a)v \in C'(k(a)), j'(a)v = 0\}$$

and each $(z, r) \in S_{v,B}$ *one can find some* $z^* \in N(C, k(a))$ *such that for* $l = j + z^* \circ k$ *one has*

$$-l'(a) \in N(B, a) ,$$

$$l''(a)vv \geqslant r + \langle z^*, z\rangle .$$

In other words: for each *critical vector* v, i.e. each $v \in K(a)$, one has

$$j''(a)vv \geqslant \sup\{r + \min_{z^* \in L(a)} \langle z^*, z - k''(a)vv\rangle : (z, r) \in S_{v,B}\} .$$

Proof. Let $(z, r) \in S_{v,B}$ and let us set $P = B'(a)$, $Q = C'(k(a))$, $f = j'(a)$, $A = k'(a)$, $b = z - k''(a)vv$, $m = r - j''(a)vv$, so that, in view of Lemma 3.2, Corollary 3.4 yields $z^* \in C'(k(a))^0 = N(C, k(a))$ such that $j'(a) + z^* \circ k'(a) \in -P^0 = -N(B, a)$ and for each $u \in B'(a)$ one has

$$j'(a)u + \langle z^*, k'(a)u - z + k''(a)vv\rangle \geqslant r - j''(a)vv .$$

Taking $u = 0$ we get the result. □

Remark. If instead of condition (R) we suppose the weaker condition

$$(\mathrm{R}_0) \qquad Z_0 := k'(a)B'(a) - C'(a) \text{ is a closed vector subspace of } Z$$

then for any $v \in K(a)$ the conclusion $l''(a)vv \geqslant r + \langle z^*, z\rangle$ holds for any $(z, r) \in S_{v,B}$ such that $z - k''(a)vv \in k'(a)B'(a) - C'(a)$.

3.6 Corollary. *Let a be a local solution to problem* ($\mathscr{M}$). *Suppose*

$$(\mathrm{R}') \qquad k'(a)\mathbb{R}_+(B - a) - \mathbb{R}_+(C - k(a)) = Z$$

holds or, more generally, suppose (R) *holds and k is metrically regular at a w.r.t. B and C. Let* $v \in K(a)$. *Then for any convex subset H of* $S_{v,B}$ *one can find* $z^* \in N(C, k(a))$ *such that for* $l = j + z^* \circ k$ *one has*

$$0 \in l'(a) + N(B, a) ,$$

$$l''(a)vv \geqslant \sup\{r + \langle z^*, z\rangle : (z, r) \in H\} .$$

In particular, for any convex subset T of $(B\times C)''(a_k, v_k)$, *where* $a_k=(a, k(a))$, $v_k=(v, k'(a)v)$ *one can find* $z^*\in N(C, k(a))$ *such that* $0\in l'(a)+N(B, a)$ *and*

$$l''(a)vv \geqslant \sup\{\langle z^*, z\rangle - l'(a)w\colon (w, z)\in T\}.$$

Proof. The fact that (R′) implies that k is metrically regular w.r.t. B and C is well known (see [40,34,35] for instance). Thus the first assertion follows from the classical minsup theorem of Moreau, the set of multipliers of ($\mathscr{M}$) being weak* compact under assumption (R) (see [35] Corollary 3.7 for instance).

In order to prove the last assertion it suffices to observe that for any $(w, z)\in(B\times C)''(a_k, v_k)$ one has $(z-k'(a)w, -j'(a)w)\in S_{v,B}$: in fact taking a sequence $((t_n, w_n, z_n))$ in $\mathbb{P}\times X\times Z$ with limit $(0, w, z)$ such that for $v_n:=v+\frac{1}{2}t_nw_n$ one has $a+t_nv_n\in B$, $k(a)+t_nk'(a)v_n+\frac{1}{2}t_n^2(z_n-k'(a)w_n)\in C$ one sees that $r_n:=-j'(a)w_n$ is such that $(r_n)\to -j'(a)w$ and $t_nj'(a)v_n+\frac{1}{2}t_n^2r_n=0$. □

In order to facilitate the application of Corollary 3.6 let us give some possible choices for H and T.

(a) For any $s:=(s_n)\in\mathbb{S}$ we can take $H=E''_{v,B,s}(\bar{y}, v)$, with $E=C\times(-\infty, j(a)]$, $\bar{y}=(k(a), j(a))$, which is a convex subset of $S_{v,B}$. This choice explains the prominent role of sequences in [19].

(b) Let z_0^* be a fixed element of the set $L(a)$ of Lagrange multipliers at a and let $Q:=\{z\in C\colon \langle z_0^*, z\rangle=0\}$. Then for any convex subset Q_0 of $Q''_{k,B}(k(a), v)$ the set $H:=\{(w, -\langle z_0^*, w\rangle)\colon w\in Q_0\}$ is easily seen to be a convex subset of $S_{v,B}$ so that for some $z^*\in L(a)$ one has

$$l''(a)vv\geqslant \sup\{\langle z^*-z_0^*, w\rangle\colon w\in Q_0\}.$$

(c) Given $s=(s_n)$ in $\mathbb{S}$ one can take for T the set

$$T_s=\liminf_n 2s_n^{-2}(B\times C-(a, k(a))-s_n(v, k'(a)v))$$

which is convex and contained in $(B\times C)''(a_k, v_k)$. This set contains the intersection $T^{\text{ii}}(B, a, v)\times T^{\text{ii}}(C, k(a), k'(a)v)$ of all the T_s's for s in $\mathbb{S}$. Thus Corollary 3.6 is stronger than the main result of [11].

(d) In particular, for any convex subset T_C of $C''(k(a), k'(a)v)$ we can take $T=B^{\text{ii}}(a, v)\times T_C$.

(e) When $v\in\mathbb{R}_+(B-a)$ is such that $k'(a)v\in C'(k(a))$ and $j'(a)v=0$, we can take for T the set $\{0\}\times T_C$ where T_C is any convex subset of $C''(k(a), k'(a)v)$ (in particular $T_C=\{0\}$ when $k'(a)v\in\mathbb{R}_+(C-k(a))$ or more generally when $0\in C''(k(a), k'(a)v)$ which is the case when C is polyhedral). Thus, taking $T=\{(0, 0)\}$, we recover the results of [3,4] and [41]. The following example shows that Corollary 3.6 applies in cases these results cannot be used.

Example. Let us take $C=E$, $k=h$, where E and h are as in Example 2.3. Then $C''(k(a), k'(a)v)$ is empty, but as $C''_k(k(a), v)=Z$ one can easily ensure that $S_{v,B}$ is nonempty. Taking for C the cone generated by $E\times\{1\}$ and $k=(h, 1)$, we may suppose C is a convex cone.

Now let us turn to sufficient optimality conditions for the mathematical programming problem $(\mathscr{M})$.

3.7 Theorem. *Suppose X is finite dimensional, j and k are twice differentiable at a and the set L(a) of Lagrange–Kuhn–Tucker multipliers at a is nonempty. If for each* $v\in K(a):=B'(a)\cap k'(a)^{-1}(C'(a))\cap j'(a)^{-1}(0)$ *and each* $(z, r)\in S_{v,B}$ *one can find* $z^*\in L(a)$ *such that, for* $l=j+z^*\circ k$ *the condition*

$$l''(a)vv>r+\langle z^*, z\rangle$$

is satisfied, then a is a strict local minimizer of j on $F=B\cap k^{-1}(C)$.

In particular a is a local strict minimizer when for any $v\in K(a)$ one has

$$j''(a)vv>\sup\{r+\inf_{z^*\in L(a)}\langle z^*, z-k''(a)vv\rangle\colon (z, r)\in S_{v,B}\}\ .$$

Proof. Suppose on the contrary that for some sequence (a_n) in $F\setminus\{a\}$ with limit a we have $j(a_n)\leqslant j(a)$. Let $t_n=\|a_n-a\|$, $v_n=t_n^{-1}(a_n-a)$; without loss of generality we may suppose (v_n) converges to some unit vector $v\in F'(a)$ with $j'(a)v\leqslant 0$. Let $z^*\in L(a)$. We have $j'(a)v+\langle z^*, k'(a)v\rangle\geqslant 0$ since $v\in B'(a)$, $j'(a)+z^*\circ k'(a)\in -N(B, a)$ and $\langle z^*, k'(a)v\rangle\leqslant 0$ since $z^*\in N(C, k(a))$ and $k'(a)v\in C'(k(a))$. Therefore $j'(a)v=0$. Taking $z=k''(a)vv$, $r=j''(a)vv$, so that $(z, r)\in S_{v,B}$ by definition of this set and choosing $z^*\in L(a)$ such that

$$(j+z^*\circ k)''(a)vv>r+\langle z^*, z\rangle$$

we get a contradiction. □

Let us observe that the sufficient condition of Theorem 3.7:

$$j''(a)vv>\sup\{r+\inf_{z^*\in L(a)}\langle z^*, z-k''(a)vv\rangle\colon (z, r)\in S_{v,B}\}$$

for $v\in K(a)$ corresponds to the necessary condition of Theorem 3.5 through the replacement of an inequality by a strict inequality. Such a situation is looked for by optimizers since it already prevails in the unconstrained case and as it cannot be improved.

4. Application to composite minimization

The study of functions of the form $f=g\circ h$ with $h:X\to Y$ of class C^2 between two Banach spaces, $g:Y\to\mathbb{R}\cup\{\infty\}$ convex lower semicontinuous (l.s.c.) and proper has been a topic of wide interest during the last few years. Here we intend to show that the results of the preceding sections apply to the problem $(\mathscr{C})$ of minimizing f on a closed convex subset D of X. Such a constraint is not taken into account in the existing literature but it is clear that it does appear in situations of interest in which the variables are usually subject to natural restrictions such as nonnegativity.

Let us impose the following qualification condition which reduces to a condition introduced by R.T. Rockafellar for the finite dimensional case, with $D=X$ and g piecewise linear-quadratic [42]

$$(\mathrm{R}_D)\qquad \mathbb{R}_+(\operatorname{dom} g-\bar{y})-h'(\bar{x})(\mathbb{R}_+(D-\bar{x}))=Y,$$

where $\bar{x}\in D$ is a solution to $(\mathscr{C})$ and $\bar{y}=h(\bar{x})$. This condition is obviously satisfied when g is finite everywhere.

Let us set

$$M_D(\bar{x})=\{y^*\in\partial g(\bar{y}):\ -y^*\circ h'(\bar{x})\in N(D,\bar{x})\}$$

where $\partial g(\bar{y})$ is the subdifferential of g at $\bar{y}$.

The following statement which is the main result of [39] will be deduced from Theorem 3.5. Let us note that it is shown there that conversely this result implies the necessary condition of Theorem 3.5. Here $g'(\bar{y},\cdot)$ denotes the contingent derivative of g at $\bar{y}$ given by

$$g'(\bar{y},y):=\inf\{r:\ (y,r)\in E_g'(\bar{y},g(\bar{y}))\},$$

where E_g is the epigraph of g.

4.1 Theorem. *Let $\bar{x}$ be a (local) minimizer of $f=g\circ h$ on D. Suppose (R_D) holds. Then the set $M_D(\bar{x})$ is nonempty and for each $v\in D'(\bar{x})$ such that $g'(\bar{y},h'(\bar{x})v)=0$ and for each $y\in Y$ one can find some $\bar{y}^*\in M_D(\bar{x})$ such that*

$$0\leqslant\langle\bar{y}^*,h''(\bar{x})vv-y\rangle+\liminf_{(t,u,z)\to^D(0_+,v,y)} 2t^{-2}(g(\bar{y}+th'(\bar{x})u+\tfrac{1}{2}t^2z)-g(\bar{y}))$$

where $(t,u,z)\to^D(0_+,v,y)$ means $(t,u,z)\to(0_+,v,y)$ with $\bar{x}+tu\in D$.

Proof. Let us transform $(\mathscr{C})$ into problem $(\mathscr{M})$ as in Theorem 3.1, changing X into $\hat{X}=X\times Y\times\mathbb{R}$, and taking $Z=Y\times Y\times\mathbb{R}$, $B=D\times Y\times\mathbb{R}$, $C=\{0\}\times E_g$, $j(x,y,r)=r$, $k(x,y,r)=(h(x)-y,y,r)$, $a=(\bar{x},\bar{y},\bar{r})=(\bar{x},h(\bar{x}),f(\bar{x}))$. It is easy to see that condition (R_D) is equivalent to

$$(\mathrm{R}')\qquad \mathbb{R}_+(C-k(a))-k'(a)(\mathbb{R}_+(B-a))=Z.$$

Let us identify the set $L(a)$ of Lagrange–Kuhn–Tucker multipliers of $(\mathscr{M})$. Given $z^*=(y^*, y_1^*, r^*)$ in $N(C, k(a))$ such that $-(j'(a)+z^*\circ k'(a))\in N(B, a)=N(D, \bar{x})\times\{(0,0)\}$, we have $y_1^*=y^*$, $r^*=-1$ and $(y^*, -1)\in N(E_g, (\bar{y}, \bar{r}))$ i.e. $y^*\in\partial g(\bar{y})$, and $-y^*\circ h'(\bar{x})\in N(D,\bar{x})$, so that $y^*\in M_D(\bar{x})$. Conversely for any $y^*\in M_D(\bar{x})$ one sees easily that $(y^*, y^*, -1)$ is a multiplier for problem $(\mathscr{M})$.

Now let $v\in D'(\bar{x})$ with $g'(\bar{y}, h'(\bar{x})v)=0$ and let $y\in Y$, $r\in\mathbb{R}$ with

$$r\geqslant \liminf_{(t,u,z)\to^{D}(0+,v,y)} 2t^{-2}(g(\bar{y}+th'(\bar{x})u+\tfrac{1}{2}t^2z)-g(\bar{y}))\ .$$

We can find a sequence $((t_n, v_n, y_n, r_n))$ in $\mathbb{P}\times X\times Y\times\mathbb{R}$ with limit $(0, v, y, r)$ such that $\bar{x}+t_nv_n\in D$ and $g(\bar{y}+t_nh'(\bar{x})v_n+\frac{1}{2}t_n^2y_n)\leqslant g(\bar{y})+\frac{1}{2}t_n^2r_n$ for each $n\in\mathbb{N}$. Let $\hat{v}:=(v, h'(\bar{x})v, 0)$, so that $\hat{v}\in B'(a)$, $k'(a)\hat{v}\in C'(k(a))$, $j'(a)\hat{v}=0$ and let $z=(0, y, r)$. Let us show that $(z, 0)$ belongs to $S_{\hat{v},B}$. In fact $\hat{v}_n:=(v_n, h'(\bar{x})v_n, 0)\to\hat{v}$, $z_n=(0, y_n, r_n)\to z$ with

$$a+t_n\hat{v}_n\in B=D\times Y\times\mathbb{R}\ ,$$

$$(0, \bar{y}, \bar{r})+t_nk'(a)\hat{v}_n+\tfrac{1}{2}t_n^2(0, y_n, r_n)\in C\ ,$$

$$j(a)+j'(a)t_n\hat{v}_n+\tfrac{1}{2}t_n^2 0\leqslant j(a)$$

for each $n\in\mathbb{N}$, since

$$(\bar{y}, \bar{r})+t_n(h'(\bar{x})v_n, 0)+\tfrac{1}{2}t_n^2(y_n, r_n)\in E_g$$

by our choice of (t_n, v_n, y_n, r_n).

Then it follows from Theorem 3.5 that we can associate to $\hat{v}$ and $(z, 0)$ some $z^*=(y^*, y^*, -1)$ with $y^*\in\partial g(\bar{y})$, $-y^*\circ h'(\bar{x})\in N(D,\bar{x})$ such that

$$y^*\circ h''(\bar{x})vv=j''(a)\hat{v}\hat{v}+z^*\circ k''(a)\hat{v}\hat{v}\geqslant\langle z^*, z\rangle=\langle y^*, y\rangle-r\ .$$

Therefore

$$0\leqslant r+\langle y^*, h''(\bar{x})vv-y\rangle\ . \qquad \square$$

Taking $D=X$ in Theorem 4.1 we get the following consequence.

4.2 Corollary. *Suppose $\bar{x}$ is a local minimizer of $f=g\circ h$ on X and* (R_D) *holds with $D=X$. Then the set*

$$M(\bar{x})=\{y^*\in\partial g(\bar{y}): y^*\circ h'(\bar{x})=0\}$$

is nonempty and for each $v\in X$ with $f'(\bar{x}, v)=0$ and each $y\in Y$ one can find some $\bar{y}^*\in M(\bar{x})$ such that

$$0\leqslant\langle\bar{y}^*, h''(\bar{x})vv\rangle+g_h''(\bar{y}, \bar{y}^*, v, y)\ .$$

Let us observe that when g is twice differentiable the preceding inequality is independent of y and amounts to the classical condition

$$0 \leqslant g'(\bar{y})h''(\bar{x})vv + g''(\bar{y})h'(\bar{x})vh'(\bar{x})v = f''(\bar{x})vv .$$

It is shown in [39] that one can give a sufficient condition which differs from the condition of Theorem 4.1 by the replacement of the inequality by a strict inequality (and does not require $M_D(\bar{x})$ if one uses $y = h''(\bar{x})vv$ only). Although the proof of [39] Proposition 1.2 is direct and simple while the following proof is rather tedious, we now show that this sufficient condition is a consequence of Theorem 3.7. Along with the proof of Theorem 4.1 it will give an opportunity to show the concrete use of compound tangent sets.

4.3 Theorem. *Let $\bar{x} \in D$ be such that $M_D(\bar{x})$ is nonempty and such that for any $v \in D'(\bar{x})$ with $g'(\bar{y}, h'(\bar{x})v) = 0$ and any $y \in Y$ one can find some $\bar{y}^* \in M_D(\bar{x})$ for which*

$$0 < \langle \bar{y}^*, h''(\bar{x})vv - y \rangle + \liminf_{(t,u,z) \to^{D} (0+,v,y)} 2t^{-2}(g(\bar{y} + th'(\bar{x})u + \tfrac{1}{2}t^2 z) - g(\bar{y})) .$$

Suppose X is finite dimensional. Then $\bar{x}$ is a local minimizer of f on D.

Proof. Let us rewrite $(\mathscr{C})$ as a mathematical programming problem as in the proof of Theorem 4.1 of which we keep the notations.

Let $\hat{v} = (v, w, s) \in B'(a)$ with $k'(a)\hat{v} \in C'(k(a)), j'(a)\hat{v} = 0$. This means that $v \in D'(\bar{x})$, $w = h'(\bar{x})v$, $s = 0$, $g'(\bar{y}, h'(\bar{x})v) = g'(\bar{y}, w) \leqslant 0$. Now for any $y^* \in M_D(\bar{x})$ we have

$$g'(\bar{y}, h'(\bar{x})v) \geqslant \langle y^*, h'(\bar{x})v \rangle \geqslant 0$$

since $-y^* \circ h'(\bar{x}) \in N(D, \bar{x})$ and $v \in D'(\bar{x})$. Therefore $g'(\bar{y}, h'(\bar{x})v) = 0$: v is critical for $(\mathscr{C})$.

Now let us show that for any $(z, r) = (y', y'', q, r) \in S_{\hat{v},B}$ we have

$$g''_{h,D}(\bar{y}, 0, v, y' + y'') \leqslant q - r .$$

Taking $y = y' + y''$ and setting $z^* = (\bar{y}^*, \bar{y}^*, -1)$ with $\bar{y}^* \in M_D(\bar{x})$ given by our assumption, that will show $\langle z^*, z \rangle + r < \langle \bar{y}^*, h''(\bar{x})vv \rangle = l''(a)vv$. Since $(z, r) \in S_{\hat{v},B}$ we can find a sequence $(t_n, \hat{v}_n, z_n, r_n)$ in $\mathbb{P} \times \hat{X} \times Z \times \mathbb{R}$ with limit $(0, \hat{v}, z, r)$, such that $a + t_n \hat{v}_n \in B$, $k(a) + t_n k'(a)\hat{v}_n + \frac{1}{2}t_n^2 z_n \in C$, $t_n j'(a)\hat{v}_n + \frac{1}{2}r_n t_n^2 \leqslant 0$. Setting $\hat{v}_n = (v_n, w_n, s_n)$, $z_n = (y'_n, y''_n, q_n)$ we get for each $n \in \mathbb{N}$: $\bar{x} + t_n v_n \in D$, $t_n s_n + \frac{1}{2}t_n^2 r_n \leqslant 0$ and

$$(0, \bar{y}, \bar{r}) + t_n(h'(a)v_n - w_n, w_n, s_n) + \tfrac{1}{2}t_n^2(y'_n, y''_n, q_n) \in \{0\} \times E_g$$

or

$$w_n = h'(a)v_n + \tfrac{1}{2}t_n y'_n ,$$

$$g(\bar{y} + t_n w_n + \tfrac{1}{2}t_n^2 y''_n) \leqslant g(\bar{y}) + t_n s_n + \tfrac{1}{2}t_n^2 q_n ,$$

so that

$$\liminf_n 2t_n^{-2}[g(\bar{y} + t_n h'(a)v_n + \tfrac{1}{2}t_n^2(y'_n + y''_n)) - g(\bar{y})] \leqslant \lim_n (q_n - r_n) = q - r$$

and the announced inequality. □

Conversely it can be shown (see [39] Theorem 4.4) that the preceding sufficient condition implies the sufficient condition of Theorem 3.7 for problem $(\mathscr{M})$.

Let us also observe that although the condition of Theorem 4.3 is not as simple as the sufficient condition of [39] Proposition 1.2, it has a double interest: it shows the links between the optimality conditions for problems $(\mathscr{C})$ and $(\mathscr{M})$ and it shows that the necessary condition of Theorem 4.1 is close to a sufficient condition.

Acknowledgement

The author gratefully acknowledges the careful readings and the criticisms of L. Barbet, A.D. Ioffe and L. Thibault about the successive versions of the present paper.

References

[1] A. Auslender, "Stabilité en programmation non différentiable, dérivée directionnelle seconde inférieure," *C.R. Académie des Sciences Série I*, 293 (1981) 301–304.
[2] A. Auslender, "Stability in mathematical programming with nondifferentiable data," *SIAM Journal Control and Optimization* 22 (1984) 239–254.
[3] A. Ben-Tal, "Second-order and related extremality conditions in nonlinear programming," *Journal Optimization Theory and Applications* 31 (1980) 143–145.
[4] A. Ben-Tal and J. Zowe, "A unified theory of first and second order conditions for extremum problems in topological vector spaces," *Mathematical Programming Study* 19 (1982) 39–76.
[5] J.M. Borwein, "A Lagrange multiplier theorem and a sandwich theorem for convex relations," *Mathematical Scandinavica* 48 (2) (1981) 189–204.
[6] J.M. Borwein, "Stability and regular points of inequality systems," *Journal Optimization Theory and Applications* 48 (1986) 9–52.
[7] R.W. Chaney, "On sufficient conditions in nonsmooth optimization," *Mathematics Operations Research* 7(3) (1982) 463–475.
[8] R.W. Chaney, "A general sufficiency theorem for nonsmooth nonlinear programming," *Transactions of the American Mathematical Society* 276(1) (1983) 235–245.
[9] R.W. Chaney, "Second-order sufficiency conditions for nondifferentiable programming problems," *SIAM Journal Control and Optimization* 20(1) (1982) 20–33.
[10] R. Cominetti, *Analyse du second ordre de problèmes d'optimisation non différentiables*, Thesis, Univ. of Clermont-Ferrand (1989).
[11] R. Cominetti, "Metric regularity, tangent sets and second order optimality conditions," *Applied Mathematics and Optimization* 21 (1990) 265–287.
[12] P.H. Dien and P.H. Sach, "Further properties of the regularity of inclusion systems," *Nonlinear Analysis, Theory, Methods and Applications* 13 (11) (1989) 1251–1269.
[13] A. Ya. Dubovitskii and A.A. Milyutin, "Extremum problems with constraints," *Zh. Vychisl. Mat. i. Mat. Fiz.* 5 (1965) 395–453 = *USSR Computational Mathematics and Mathematical Physics* 5(3) (1965) 1–80.
[14] A. Ya. Dubovitskii and A.A. Milyutin, "Second variations in extremal problems with constraints," *Doklady Akademii Nauk S.S.S.R.* 160 (1965) = *Soviet Mathematics Doklady* 6 (1965) 12–16.
[15] J. Gwinner, "Results of Farkas type," *Numerical Functional Analysis and Optimization* 9 (5&6) (1987) 471–520.
[16] S.P. Han and O.L. Mangasarian, "Exact penalty functions in nonlinear programming," *Mathematical Programming* 17 (1979) 251–269.

[17] K.H. Hoffmann and H.J. Kornstaedt, "Higher order necessary conditions in abstract mathematical programming," *Journal of Optimization Theory and Applications* 26 (1978) 533–569.
[18] A.D. Ioffe, "Necessary and sufficient conditions for a local minimum 3: Second order conditions and augmented duality," *SIAM Journal Control and Optimization* 17 (1979) 266–288.
[19] A.D. Ioffe, "On some recent developments in the theory of second order optimality conditions," in: S. Dolecki ed., *Optimization*, Proc. Conf. Varetz, 1988, *Lecture Notes in Mathematics*, n° 1405 (Springer Verlag, Berlin, 1989) 55–68.
[20] A.D. Ioffe and V.M. Tikhomirov, *Theory of Extremal Problems* (Nauka, Moscow, 1974, North-Holland, Amsterdam, 1979).
[21] H. Kawasaki, "An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems," *Mathematical Programming* 41 (1988) 73–96.
[22] S. Kurcyusz, "On the existence and nonexistence of Lagrange multipliers in Banach spaces," *Journal of Optimization Theory and Applications* 20 (1976) 81–110.
[23] F. Lempio and J. Zowe, "Higher order optimality conditions," in: B. Korte ed., *Modern Applied Mathematics* (North-Holland, Amsterdam, 1982) 148–193.
[24] E.S. Levitin, A.A. Milyutin and N.P. Osmolovski, "Conditions of high order for a local minimum in problems with constraints," *Uspekhi Mat. Nauk* 33 (6) (1978) 85–148 = *Russian Mathematical Surveys* 33(6) (1978) 97–168.
[25] A. Linnemann, "Higher-order necessary conditions for infinite and semi-infinite optimization," *Journal Optimization Theory and Applications* 38 (4) (1982) 483–512.
[26] D.G. Luenberger, *Optimization by Vector Spaces Methods* (John Wiley, New York, 1969).
[27] J. Maguregui, "Regular multivalued functions and algorithmic applications," Ph. D. Thesis, University of Wisconsin-Madison (1976).
[28] H. Maurer, "First and second order sufficient optimality conditions in mathematical programming and optimal control," *Mathematical Programming Study* 14 (1981) 163–177.
[29] H. Maurer and J. Zowe, "First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems," *Mathematical Programming* 16 (1979) 98–110.
[30] Ph. Michel, "Problème des inégalités. Applications à la programmation et au contrôle optimal," *Bulletin de la Société Mathématique de France* 101 (1973) 413–439.
[31] Ph. Michel, "Problèmes d'optimisation définis par des fonctions qui sont sommes de fonctions convexes et de fonctions dérivables," *Journal de Mathématiques Pures et Appliquées* 53 (1974) 321–330.
[32] L.W. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems I. General theory," *Journal S.I.A.M. Control* 4 (3) (1966) 505–527.
[33] L.W. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems II, Applications," *Journal S.I.A.M. Control* 5(1) (1967) 90–137.
[34] J.-P. Penot, "On the existence of Lagrange multipliers in nonlinear programming in Banach spaces," in: A. Auslender, W. Oettli, J. Stoer ed., *Optimization and Optimal Control*, Proc. Conf. Oberwolfach (1980), *Lecture Notes in Control and Information Science n°* 30 (Springer Verlag, Berlin, 1981) 89–104.
[35] J.-P. Penot, "On regularity conditions in mathematical programming," *Mathematical Programming Study* 19 (1982) 167–199.
[36] J.-P. Penot, "A view of second order extremality conditions," in C. Lemaréchal ed., *Third Franco-German Conference in Optimization*, INRIA, July 1984, pp. 34–39.
[37] J.-P. Penot, "A geometric approach to higher order necessary conditions," manuscript, July 1984.
[38] J.-P. Penot, "Generalized higher order derivatives and higher order optimality conditions," preprint, Univ. of Santiago, Nov. 1984.
[39] J.-P. Penot, "Optimality conditions for composite functions," preprint Univ. of Pau 1990.
[40] S. Robinson, "Stability theory for systems of inequalities, part II: differentiable nonlinear systems," *SIAM Journal Numerical Analysis* 13 (1976) 497–513.
[41] S. Robinson, "Local structure of feasible sets in nonlinear programming. Part I. Regularity," in *Numerical Methods*, Caracas (1982), *Lecture Notes in Mathematics* n° 1005 (Springer Verlag, Berlin, 1983) 240–251.
[42] R.T. Rockafellar, "First and second-order epi-differentiability in nonlinear programming." *Transactions American Mathematical Society* 307 (1) (1988) 75–108.
[43] R.T. Rockafellar, "Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives," *Mathematics of Operation Research* 14 (3) (1989) 462–484.

[44] F. Tröltzsch, "On changing the spaces in Lagrange multiplier rules for the optimal control of non-linear operator equations," *Optimization* 16 (6) (1985) 877–885.

[45] C. Vîrsan, "Necessary conditions of extremality of high order," *Revue Roumaine de Mathématiques Pures et Appliquées* 28 (1973) 591–611.

[46] J. Zowe and S. Kurcyusz, "Regularity and stability for the mathematical programming problem in Banach spaces," *Applied Mathematics and Optimization* 5 (1) (1979) 46–62.