Appl Math Optim 32:213–234 (1995)

Applied Mathematics & Optimization

© 1995 Springer-Verlag New York Inc.

# New Results in Subdifferential Calculus with Applications to Convex Optimization*

G. Romano

Dipartimento di Scienza delle Costruzioni, Università di Napoli Federico II,
P.le Tecchio 80, Napoli 80125, Italia

**Abstract.** Chain and addition rules of subdifferential calculus are revisited in the paper and new proofs, providing local necessary and sufficient conditions for their validity, are presented. A new product rule pertaining to the composition of a convex functional and a Young function is also established and applied to obtain a proof of Kuhn–Tucker conditions in convex optimization under minimal assumptions on the data. Applications to plasticity theory are briefly outlined in the concluding remarks.

**Key Words.** Convex analysis, Subdifferentials, Chain rules, Optimization.

**AMS Classification.** Primary 4902, Secondary 49J52.

## Introduction

Subdifferential calculus is nowadays a well-developed chapter of nonsmooth analysis which is recognized for its many applications to optimization theory. The very definition of subdifferential and the basic results concerning the addition and the chain rule of subdifferential calculus were first established in the early sixties by Rockafellar [10] with reference to convex functions on $\mathcal{R}^n$. A comprehensive treatment of the subject has been provided by himself in the later book on convex analysis [11]. The theory was developed further by Moreau [7] in the context of linear topological vector spaces and applied to problems of unilateral mechanics [8].

* The financial support of the Italian Ministry for University and Scientific and Technological Research is gratefully acknowledged.

A summary of basic mathematical results can also be found in the book by Laurent [5] and in an introductory chapter of the book by Ekeland and Temam [3].

In the early seventies different attempts were initiated to extend the range of validity of subdifferential calculus to nonconvex functions, mainly by Rockafellar and his school. In this context saddle functions were considered by McLinden [6]. Significant advances were made by Clarke [1], [2], who set up a definition of subdifferential for arbitrary lower semicontinuous functions on $\mathscr{R}^n$ and extended the validity of the rules of subdifferential calculus to this nonconvex context. His results were later developed and extended by Rockafellar [12], [13], who has also provided a nice exposition of the state of art, up to the beginning of the eighties, in [14].

A review of the main results and applications in different areas of mathematical physics can be found in a recent book by Panagiotopoulos [9]. A different treatment of the subject is presented in the book by Ioffe and Tihomirov [4], who introduce the notion of regular local convexity to deal with the nonconvex case.

A careful review of all these contributions to subdifferential calculus leads however to the following considerations.

The results provided up to now to establish the validity of the addition and of the chain rule for subdifferentials appear to rely upon sufficient but largely not necessary assumptions. In fact a number of important situations, in which the results do hold true, are beyond the target of existing theorems. On the other hand the author has realized the lack of a chain rule concerning the very important case of convex functionals which are expressed as the composition of a monotone convex function and another convex functional.

The first observation in this respect was made with reference to positively homogeneous convex functionals of order greater than one or, more generally, to convex functionals which are composed by a Young function and a sublinear functional (gauge-like functionals in Rockafellar's terminology).

The theorems presented in this paper are intended to contribute to the filling of these gaps; progress is provided in two directions.

The first concerns the chain rule pertaining to the composition of a convex functional and a differentiable operator. We have addressed the question of finding a necessary and sufficient condition for its validity. The theorem provided here shows that this task can be accomplished to within a closure operation; the proof is straightforward and relies on a well-known lemma of convex analysis concerning sublinear functionals.

The result obtained must be considered as optimal; a simple counterexample reveals indeed that there is no hope of dropping the closure operation. On the contrary, to establish a perfect equality (one not requiring closures) in the chain-rule formula, classical treatments were compelled to set undue restrictions on the range of validity of the result. In this respect it has to be remarked that classical conditions were *global* in character, in the sense that validity of chain and addition rules were ensured at all points. The new results provided here are instead based upon *local* conditions which imply validity of the rules only at the very point where subdifferentials have to be evaluated. It follows that classical conditions can be verified *a priori* while the new conditions must be checked *a posteriori* at the point of interest.

The second contribution consists in establishing a new chain-rule formula concerning functionals which are formed by the composition of a monotone convex function and a convex functional. A natural application of these results can be exploited in convex optimization problems. It is shown in fact that the Kuhn and Tucker multipliers theory can be immediately derived from the above theorems and the existence proof can be performed under assumptions less stringent than the classical Slater conditions [18].

Computation of the subdifferentials involved in the proof requires considering the following two special cases of the new chain rules of subdifferential calculus contributed here. In the first case we have to deal with the composition of the indicator of the zero and of an affine functional. In the second one we must consider a functional formed by the composition of the indicator of nonpositive reals and of a convex functional. Both cases were not covered by previous results.

## 1. Local Convexity and Subdifferentials

Let $(X, X')$ be a pair of locally convex topological vector spaces (l.c.t.v.s.) placed in separating duality by a bilinear form $\langle \cdot, \cdot \rangle$ and let $f\colon X \mapsto \mathscr{R} \cup \{+\infty\}$ be an extended real-valued functional with a nonempty effective domain:

$$\operatorname{dom} f = \{x \in X \,|\, f(x) < +\infty\}.$$

The one-sided directional derivative of $f$ at the point $x \in \operatorname{dom} f$, along the vector $h \in X$, is defined by the limit

$$df(x; h) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}[f(x + \varepsilon h) - f(x)].$$

The derivative of $f$ at $x$ is then the extended real-valued functional $p\colon X \mapsto \{-\infty\} \cup \mathscr{R} \cup \{+\infty\}$ defined by

$$p(h) \stackrel{\text{def}}{=} df(x; h),$$

which is easily seen to be positively homogeneous in $h$.

The functional $f$ is said to be **locally convex** at $x$ when $p$ is sublinear in $h$, that is,

$$\begin{cases} p(\alpha h) = \alpha p(h), & \forall \alpha \geq 0 \quad \text{(positive homogeneity)}, \\ p(h_1) + p(h_2) \geq p(h_1 + h_2), & \forall h_1, h_2 \in X \quad \text{(subadditivity)}. \end{cases}$$

The epigraph of $p$ is then a convex cone in $X \times \mathscr{R}$.

A locally convex functional $f$ is said to be **locally subdifferentiable** at $x$ if its one-sided derivative $p$ is a proper sublinear functional, i.e., if it is nowhere $-\infty$. In fact, denoting by $\bar{p}$ the closure of $p$ defined by the limit formula

$$\bar{p}(h) = \liminf_{z \to h} p(z), \qquad \forall h \in X,$$

a well-known result of convex analysis ensures that the proper lower semicontinuous (l.s.c.) sublinear functional $\bar{p}$, turns out to be the support functional of a nonempty closed convex set $K^*$, that is,

$$\bar{p}(h) = \sup\{\langle x^*, h\rangle : x^* \in K^*\},$$

with

$$K^* = \{x^* \in X' : p(h) \geq \langle x^*, h\rangle, \forall h \in X\}.$$

The **local subdifferential** of the functional $f$ is then defined by

$$\partial f(x) \overset{\text{def}}{=} K^*.$$

A relevant special case, which will be referred to in the sequel, occurs when the one-sided derivative of $f$ at $x$, turns out to be l.s.c. so that $p = \bar{p}$. The functional $f$ is then said to be **regularly locally subdifferentiable** at $x$.

When the functional $f$ is differentiable at $x \in X$ the local subdifferential is a singleton and coincides with the usual differential.

For a convex functional $f$: $X \mapsto \mathscr{R} \cup \{+\infty\}$, the difference quotient in the definition of a one-sided directional derivative does not increase as $\varepsilon$ decreases to zero [4], [11]. Hence the limit exists at every point $x \in \operatorname{dom} f$ along any direction $h \in X$ and the following formula holds:

$$df(x; h) = \inf_{\varepsilon > 0} \frac{1}{\varepsilon}[f(x + \varepsilon h) - f(x)].$$

A simple computation shows that the directional derivative of $f$ is convex as a function of $h$ and hence sublinear.

Moreover the definition of local subdifferential turns out to be equivalent to

$$x^* \in \partial f(x) \Leftrightarrow f(y) - f(x) \geq \langle x^*, y - x\rangle, \qquad \forall y \in X,$$

which is the usual definition of subdifferential in convex analysis [11].

## 2. Classical Subdifferential Calculus

Let $f_1, f_2$: $X \mapsto \mathscr{R} \cup \{+\infty\}$ and $f$: $Y \mapsto \mathscr{R} \cup \{+\infty\}$ be convex functionals and let $L$: $X \mapsto Y$ be a continuous linear operator. From the definition of local subdifferential it follows easily that

$$\partial(\lambda f)(x) = \lambda\, \partial f(x), \qquad \lambda \geq 0,$$

$$\partial(f_1 + f_2)(x) \supseteq \partial f_1(x) + \partial f_2(x),$$

$$\partial(f \circ L)(x) \supseteq L'\, \partial f(Lx),$$

where $L'$ denotes the dual of $L$.

As remarked in [3] equality in the last two relations is far from being always realized. The aim of subdifferential calculus has thus primarily consisted in providing conditions sufficient to ensure that the converse of the last two inclusions does hold true. In convex analysis this task has been classically accomplished by the following kind of results [3]–[5], [11].

**Theorem 2.1** (Additivity). *If* $f_1, f_2\colon X \mapsto \mathcal{R} \cup \{+\infty\}$ *are convex and at least one of them is continuous at a point of* $\operatorname{dom} f_1 \cap \operatorname{dom} f_2$, *then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x), \qquad \forall x \in X.$$

**Theorem 2.2** (Chain-Rule). *Given a continuous linear operator* $L\colon X \mapsto Y$ *and a convex functional* $f\colon Y \mapsto \mathcal{R} \cup \{+\infty\}$ *which is continuous at a point of* $\operatorname{dom} f \cap \operatorname{Im} L$, *it results that*

$$\partial(f \circ L)(x) = L' \,\partial f(Lx), \qquad \forall x \in X.$$

The chain-rule equality above can be equivalently written with the more familiar notation

$$\partial(f \circ L)(x) = \partial f(Lx) \circ L, \qquad \forall x \in X.$$

A generalization of the previous results can be performed to get a chain rule involving a locally convex functional and a nonlinear differentiable operator.

Given a nonlinear differentiable operator $A\colon X \mapsto Y$ and a functional $f\colon Y \mapsto \mathcal{R} \cup \{+\infty\}$ which is locally convex at $y_0 = A(x_0)$, we have to prove the following equality:

$$\partial(f \circ A)(x_0) = \partial f[A(x_0)] \circ dA(x_0) = [dA(x_0)]' \,\partial f[A(x_0)],$$

where $dA(x_0)$ is the derivative of the operator $A$ at $x_0 \in X$.

The task can be accomplished by first providing conditions sufficient to guarantee the validity of the chain-rule identity for one-sided directional derivatives:

$$d(f \circ A)(x_0; x) = df[A(x_0); dA(x_0)x], \qquad \forall x \in X,$$

which is easily seen to hold trivially when $A$ is an affine operator. Then, setting $L \overset{\text{def}}{=} dA(x_0)$, we consider the sublinear functionals

$$p(y) \overset{\text{def}}{=} df(A(x_0); y) \quad \text{and} \quad q(x) \overset{\text{def}}{=} d(f \circ A)(x_0; x).$$

The identity above ensures that $q = p \circ L$; further, observing that, by definition,

$$\partial p(0) = \partial f(A(x)) \quad \text{and} \quad \partial q(0) = \partial(p \circ L)(0) = \partial(f \circ A)(x),$$

the equality to be proved can be rewritten as

$$\partial(p \circ L)(0) = L' \,\partial p(0).$$

This result can be inferred from the chain-rule theorem concerning convex functionals by assuming that the sublinear functional $p$ is continuous at a point of $\operatorname{dom} p \cap \operatorname{Im} L$.

In this respect we remark that it has been shown in [4] that, assuming the functional $f$ to be *regularly locally convex* at $A(x) \in Y$, that is locally convex and uniformly differentiable in all directions at $A(x)$, its derivative $p$ turns out to be continuous in the whole space $Y$. Therein it is also proved that a convex functional is regularly locally convex at a point if and only if it is continuous at that point. An analogous generalization can be performed for the addition formula of subdifferential calculus.

A different and more general treatment of the nonconvex case has been developed, on the basis of Clarke's [1], [2] contributions, by Rockafellar [12], [13]. According to his approach the validity of the chain rule was proved by assuming that the operator $A$ is strictly differentiable at $x \in X$, that $f$ is finite, directionally Lipschitzian, and subdifferentially regular at $A(x)$ and that the interior of the domain of the one-sided derivative of $f$ at $x \in X$ has a nonempty intersection with the range of $dA(x)$.

Reference is made to the quoted papers for a precise assessment of definitions and proofs.

## 3. New Results

As illustrated above, all the contributions provided to subdifferential calculus until now have directed their efforts in the direction of finding conditions directly sufficient to ensure the validity of the equality sign in the relevant relations. This approach has led to the formulation of very stringent conditions which rule out a number of significant situations.

In the next subsection we propose an alternative approach to the assessment of the chain rule pertaining to the composition of a convex functional and a differentiable operator. Further we derive the addition rule as a special case of this chain rule.

In the second subsection we present the proof of a new product-rule formula of subdifferential calculus which deals with the composition of a monotone convex function and a convex functional.

These results are applied in the last subsection to assess the existence of Kuhn and Tucker multipliers in convex optimization problems, under assumptions less stringent than the classical Slater conditions [18] (see also [11], and [14]).

### *3.1. Classical Addition and Chain-Rule Formulas*

The new approach to classical rules of subdifferential calculus consists in splitting the procedure into two steps. It has in fact been realized that getting the equality at once in the related relations requires too stringent assumptions and follows less deep insight into the problem.

The classical chain rule requires the equality of the subdifferential of a composite function, which is a closed convex set, to the image of the subdifferential of a convex function through a linear operator. Since in general the image of a closed convex set fails to be closed too, it is natural to look first for conditions apt to provide equality of the former set to the closure of the latter one, leaving to a subsequent step the answer about the closedness of the latter set.

The first step is performed by means of the following result.

**Theorem 3.1** (New Proof of the Classical Chain Rule). *Let $A: X \mapsto Y$ be a nonlinear operator which is differentiable at a point $x_0 \in X$ with derivative $dA(x_0): X \mapsto Y$ linear and continuous. Let further $f: Y \mapsto \mathscr{R} \cup \{+\infty\}$ be a functional which is locally subdifferentiable at $A(x_0) \in Y$ and assume that $f \circ A: X \mapsto \mathscr{R} \cup \{+\infty\}$ is locally subdifferentiable at $x_0 \in X$. Then we have that*

$$\partial(f \circ A)(x_0) = \overline{\partial f[A(x_0)] \circ [dA(x_0)]} = \overline{[dA(x_0)]' \partial f[A(x_0)]}$$

*if and only if*

$$\bar{q}(x) = \bar{p}(Lx), \qquad \forall x \in X,$$

*where $q(\cdot) \overset{\text{def}}{=} d(f \circ A)(x_0; \cdot)$, $p(\cdot) \overset{\text{def}}{=} df[A(x_0); \cdot]$, and $L \overset{\text{def}}{=} dA(x_0)$, a superimposed bar denoting the closure.*

*Proof.* $f$ being locally subdifferentiable at $A(x_0) \in Y$, its directional derivative $p: Y \mapsto \mathscr{R} \cup \{+\infty\}$ is a proper sublinear functional, so that

$$\bar{p}(y) = \sup\{\langle y^*, y\rangle | y^* \in K^*\},$$

where

$$K^* = \partial p(0) \overset{\text{def}}{=} \{y^* \in Y' | p(y) \geq \langle y^*, y\rangle, \forall y \in Y\}$$

is a nonempty, closed convex set. Then we have

$$\bar{p}(Lx) = \sup\{\langle y^*, Lx\rangle | y^* \in K^*\} = \sup\{\langle x^*, x\rangle | x^* \in L'K^*\}.$$

Similarly, $f \circ A$ being locally subdifferentiable at $x_0 \in X$ its directional derivative $q: X \mapsto \mathscr{R} \cup \{+\infty\}$ is a proper sublinear functional, so that

$$\bar{q}(x) = \sup\{\langle x^*, x\rangle | x^* \in C^*\},$$

where

$$C^* = \{x^* \in X' | q(x) \geq \langle x^*, x\rangle, \forall x \in X\}$$

is a nonempty, closed convex set in $X'$.

Comparison of the two expressions above leads directly to the following conclusion:

$$\bar{q}(x) = \bar{p}(Lx) \quad \text{if and only if } C^* = \overline{L'K^*}.$$

The statement of the theorem is then inferred by observing that

$$\begin{aligned}\partial f[A(x_0)] &= \{y^* \in Y' | df[A(x_0); y] \geq \langle y^*, y\rangle, \forall y \in Y\}\\ &= \{y^* \in Y' | p(y) \geq \langle y^*, y\rangle, \forall y \in Y\} = \partial p(0) = K^*\end{aligned}$$

and

$$\begin{aligned}\partial(f \circ A)(x_0) &= \{x^* \in X' | d(f \circ A)(x_0; x) \geq \langle x^*, x\rangle, \forall x \in X\}\\ &= \{x^* \in X' | q(x) \geq \langle x^*, x\rangle, \forall x \in X\} = \partial q(0) = C^*\end{aligned}$$

and the proof is complete. □

A useful variant is stated in the following:

**Corollary 3.2.** *Let $A$: $X \mapsto Y$ be a nonlinear operator which is differentiable at a point $x_0 \in X$ with derivative $dA(x_0)$: $X \mapsto Y$ linear and continuous. Let further $f$: $Y \mapsto \mathscr{R} \cup \{+\infty\}$ be a functional which is locally subdifferentiable at $A(x_0) \in Y$ and assume that the following identity holds:*

$$d(f \circ A)(x_0; x) = df[A(x_0); dA(x_0)x], \qquad \forall x \in X.$$

*Then we have*

$$\partial(f \circ A)(x_0) = \overline{\partial f[A(x_0)] \circ [dA(x_0)]} = \overline{[dA(x_0)]' \, \partial f[A(x_0)]}$$

*if and only if*

$$\overline{(p \circ L)}(x) = \bar{p}(Lx), \qquad \forall x \in X,$$

*where $p(y) \stackrel{\text{def}}{=} df[A(x_0); y]$ and $L \stackrel{\text{def}}{=} dA(x_0)$, a superimposed bar denoting the closure.*

*Proof.* The result is directly inferred from the theorem above by noting that the assumed identity amounts to requiring that $q = p \circ L$. □

It has to be remarked that the chain rule for one-sided directional derivatives assumed in the statement of the corollary holds trivially for every affine operator $A$. Moreover the necessary and sufficient condition is fulfilled when the sublinear functional $p$ is closed.

The next result shows that the addition rule for subdifferentials can be directly derived by applying the result provided by the chain-rule theorem.

**Theorem 3.3** (New Proof of the Classical Addition Rule). *We consider the functionals $f_i$: $X \mapsto \mathscr{R} \cup \{+\infty\}$ with $i = 1, \ldots, n$ and assume that they are locally subdifferentiable at $x_0 \in X$ with $p_i(x) \stackrel{\text{def}}{=} df_i(x_0; x)$. The following addition rule then holds:*

$$\partial\left(\sum_{i=1}^{n} f_i\right)(x_0) = \overline{\sum_{i=1}^{n} \partial f_i(x_0)}$$

*if and only if*

$$\left(\overline{\sum_{i=1}^{n} p_i}\right)(x) = \sum_{i=1}^{n} \bar{p}_i(x), \qquad \forall x \in X.$$

*Proof.* Let $A$: $X \mapsto X^n$ be the iteration operator defined as

$$Ax = |x_i|, \qquad x_i = x, \qquad i = 1, \ldots, n.$$

The dual operator $A'$: $X' \mapsto (X^n)'$ meets the identity

$$\langle A'|x_i^*|, x\rangle = \langle |x_i^*|, Ax\rangle = \sum_{i=1}^{n} \langle x_i^*, x\rangle = \left\langle \sum_{i=1}^{n} x_i^*, x\right\rangle, \qquad \forall x \in X,$$

and hence is the addition operator

$$A'|x_i^*| = \sum_{i=1}^{n} x_i^*.$$

Defining the functional $f$: $X^n \mapsto \mathscr{R} \cup \{+\infty\}$ as $f(|x_i|) \overset{\text{def}}{=} \sum_{i=1}^{n} f_i(x_i)$, we have $(f \circ A)(x) = \sum_{i=1}^{n} f_i(x)$ and hence

$$\partial(f \circ A)(x_0) = \partial\left(\sum_{i=1}^{n} f_i\right)(x_0).$$

On the other hand,

$$df(Ax_0; |x_i|) = \lim_{\alpha \downarrow 0} \frac{1}{\alpha}[f(Ax_0 + \alpha|x_i|) - f(Ax_0)]$$

$$= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \sum_{i=1}^{n} [f_i(x_0 + \alpha x_i) - f_i(x_0)] = \sum_{i=1}^{n} df_i(x_0; x_i).$$

By the definition of local subdifferential we then get

$$|x_i^*| \in \partial f(Ax_0) \Leftrightarrow x_i^* \in \partial f_i(x_0)$$

so that

$$A' \, \partial f(Ax_0) = \sum_{i=1}^{n} \partial f_i(x_0).$$

$$p(|x_i|) \stackrel{\text{def}}{=} df(Ax_0; |x_i|) = \sum_{i=1}^{n} df_i(x_0; x_i) \stackrel{\text{def}}{=} \sum_{i=1}^{n} p_i(x_i),$$

$$\overline{(p \circ A)}(x) = \left(\overline{\sum_{i=1}^{n} p_i}\right)(x),$$

$$\bar{p}(Ax) = \sum_{i=1}^{n} \bar{p}_i(x),$$

the proof follows from the result contributed in the chain-rule theorem above. □

**Corollary 3.4.** *We consider the functionals $f_i$: $X \mapsto \mathscr{R} \cup \{+\infty\}$ with $i = 1, \ldots, n$, and assume that they are regularly locally subdifferentiable at $x_0 \in X$. The following addition rule then holds*:

$$\partial\left(\sum_{i=1}^{n} f_i\right)(x_0) = \overline{\sum_{i=1}^{n} \partial f_i(x_0)}.$$

*Proof.* The result follows at once by Theorem 4.3, observing that

$$p_i \ (i = 1, \ldots, n) \quad \text{closed} \Rightarrow \sum_{i=1}^{n} p_i \quad \text{closed}.$$ □

We now derive a special case of the chain-rule formula which is referred to later when dealing with the existence of Kuhn and Tucker vectors in convex optimization.

**A Special Case.** Let $A$: $X \mapsto Y$ be a continuous affine operator, that is,

$$A(x) = L(x) + c$$

with $L$: $X \mapsto Y$ linear and continuous and $c \in Y$. Let further $f$: $Y \mapsto \mathscr{R} \cup \{+\infty\}$ be the convex indicator of the point $\{A(x_0)\}$:

$$f(y) = \text{ind}_{\{A(x_0)\}}(y), \qquad \forall y \in Y.$$

The chain rule for one-sided directional derivatives holds true since $A$ is affine. Moreover the functionals

$$p(y) \stackrel{\text{def}}{=} df[A(x_0); y] = \text{ind}_{\{0\}}(y),$$

$$(p \circ L)(x) \stackrel{\text{def}}{=} df[A(x_0); Lx] = \text{ind}_{\{0\}}(Lx) = \text{ind}_{\{KerL\}}(x)$$

turn out to be sublinear, proper, and closed.

On the basis of the corollary to the chain-rule theorem provided above we may then state that

$$\partial(f \circ A)(x_0) = \overline{L'Y'} = \overline{\text{Im}\, L'}.$$

The particular case when $Y = \mathscr{R}$ will be of special interest in the sequel. In this case we may write

$$A(x) = \langle a^*, x\rangle + c \qquad \text{with} \quad a^* \in X', c \in \mathscr{R}.$$

Note that now $L = a^*\colon X \mapsto \mathscr{R}$ and $L'\colon \mathscr{R} \mapsto X'$ with $Lx = \langle a^*, x\rangle$ and $L'\alpha = \alpha a^*$.

It follows that $\operatorname{Im} L' = \operatorname{Lin}\{a^*\}$ is a closed subspace and hence

$$\partial(f \circ A)(x_0) = \operatorname{Im} L' = \operatorname{Lin}\{a^*\} = L'\,\partial f[A(x_0)] = \partial f[A(x_0)]L = \mathscr{R}a^*,$$

which is the formula of future interest.

Two significant examples are reported hereafter to enlighten the meaning of the conditions required for the validity of the chain-rule formula.

**Examples.** The first example shows that, when the necessary and sufficient condition for the validity of the chain-rule formula is not satisfied, the two convex sets involved in the formula can in fact be quite different from one another.

Let $f$ be the convex indicator of a circular set in $\mathscr{R}^2$ centered at the origin and let $(x_0, 0)$ be a point on its boundary. The one-sided directional derivative of $f$ at $(x_0, 0)$ is the proper sublinear functional $p\colon \mathscr{R}^2 \mapsto \mathscr{R}$ given by

$$p(x,y) = \begin{cases} 0 & \text{for } x < 0 \text{ and at the origin,} \\ +\infty, & \text{elsewhere.} \end{cases}$$

Denoting the orthogonal projector on the axis $\mathscr{R}_y$ by $L = L'$ we have

$$(p \circ L) = \operatorname{ind}_{\{\mathscr{R}_x\}} \qquad \text{and then} \quad \partial(p \circ L)(0,0) = \mathscr{R}_y.$$

On the other hand,

$$\partial p(0,0) = \mathscr{R}_x^+ \qquad \text{so that} \quad L'\,\partial p(0,0) = L'\mathscr{R}_x^+ = (0,0).$$

The second example provides a situation in which all the assumptions set forth in the corollary are met but still the two convex sets fail to be equal since the second one is nonclosed. Let $K^*$ be the hyperbolic convex set in $\mathscr{R}^2$ defined by

$$K^* \overset{\text{def}}{=} \{(x^*, y^*) \in \mathscr{R}^2 | x^* y^* \geq 1\}$$

and let $p$ be its support functional:

$$p(x,y) \overset{\text{def}}{=} \sup\{\langle x^*, x\rangle + \langle y^*, y\rangle | (x^*, y^*) \in K^*\}.$$

Denoting the orthogonal projector on the axis $\mathscr{R}_y$ again by $L = L'$ we then have

$$(p \circ L)(x,y) = \begin{cases} 0 & \text{on} \quad \mathscr{R}_x \times \mathscr{R}_y^-, \\ +\infty & \text{elsewhere.} \end{cases}$$

Hence $K^* = \partial(p \circ L)(0,0) = \mathscr{R}_y^+$ but $L'\,\partial p(0,0) = L'K^* = \mathscr{R}_y^+ - (0,0)$ which is open.

### 3.2. *A New Product-Rule Formula*

We present here the proof of a new product-rule formula of subdifferential calculus which deals with the composition of a monotone convex function and a convex functional.

The original interest of the author for this kind of product rule arose in connection with subdifferential relations involving gauge-like functionals [11] which are composed by a monotone convex Young function and a sublinear Minkowsky functional. The new product-rule formula turns out to be of the utmost interest in dealing with minimization problems involving convex constraints expressed in terms of level sets of convex functionals.

A new approach to the Kuhn and Tucker theory of convex optimization can be founded upon these results and is carried out in the next subsection.

Two introductory lemmas, which the main theorem resorts to, are preliminarily reported hereafter.

**Lemma 3.5.** *Let $I = [\lambda_1, \lambda_2]$ be an interval belonging to the nonnegative real line and let $C$ be a weakly compact convex set in $X$. Then the set $IC$ is convex and closed if either*

(a) $0 \notin C$, *or*
(b) *$I$ is compact (i.e., bounded).*

*Proof.* We first prove that $C$ being convex, the set $IC$ is convex too.

If $\bar{x}_1, \bar{x}_2 \in IC$ and $\alpha_1, \alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = 1$, we have

$$\alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 = \alpha_1 m_1 x_1 + \alpha_2 m_2 x_2 \quad \text{with } m_1, m_2 \in I, \quad x_1, x_2 \in C.$$

Now, by the convexity of $C$ [11, Theorem 3.2],

$$\alpha_1 m_1 x_1 + \alpha_2 m_2 x_2 \in \alpha_1 m_1 C + \alpha_2 m_2 C = (\alpha_1 m_1 + \alpha_2 m_2) C \subseteq IC,$$

the last inclusion holding true since $\alpha_1 m_1 + \alpha_2 m_2 \in I$, by the convexity of $I$.

To prove the weak closedness of $IC$, we consider a weak limit point $z$ of $IC$ and a sequence $\{a_k x_k\}$, with $a_k \in I$ and $x_k \in C$, converging weakly to $z$:

$$\langle x^*, a_k x_k \rangle \mapsto \langle x^*, z \rangle, \qquad \forall x^* \in X'.$$

$C$ being weakly compact in $X$, we may assume that the sequence $\{x_k\}$ is weakly convergent to a point $x \in C$.

Under assumption (a) we then infer that $x \neq 0$ so that there is an $\bar{x}^*$ such that

$$\langle \bar{x}^*, x_k \rangle \mapsto \langle \bar{x}^*, x \rangle > 0.$$

For a sufficiently large $k$, $\langle \bar{x}^*, x_k \rangle \geq \xi > 0$, and hence the sequence $\{a_k\}$ cannot be unbounded. In fact otherwise $\langle x^*, a_k x_k \rangle \geq a_k \xi \mapsto +\infty$, contrary to the assumption that $a_k x_k \overset{w}{\mapsto} z$.

Under assumption (b) the boundedness of the sequence $\{a_k\}$ is a trivial consequence of the boundedness of $I$.

In both cases we may then assume that $a_k \mapsto a \in I$ and $x_k \overset{w}{\mapsto} x \in C$. As a consequence we get that $\{a_k x_k\} \overset{w}{\mapsto} ax$ and hence $z = ax \in IC$. □

**Lemma 3.6.** *Let $f$: $X \mapsto \mathscr{R}$ be a continuous nonconstant convex functional. Denoting its zero level set by $N$, if there is a vector $x_- \in N$ such that $f(x_-) < 0$ then*

$$\text{int } N = N_- \overset{\text{def}}{=} \{x \in X | f(x) < 0\},$$

$$\text{bnd } N = N_0 \overset{\text{def}}{=} \{x \in X | f(x) = 0\},$$

*and both sets turn out to be nonempty.*

*Proof.* Since $f(x_-) < 0$, by the continuity of $f$ a neighborhood $\mathscr{N}(x_-)$ exists such that $f(x) < 0$, $\forall x \in \mathscr{N}(x_-)$. Hence $\mathscr{N}(x_-) \subset N$ so that $x_- \in \text{int } N$.

Further, $f$ being nonconstant and negative at $x_-$, by convexity there will be an $x_0 \in X$ such that $f(x_0) = 0$. Let $S(x_0; x_-) \subset N$ be the segment joining $x_0$ and $x_-$ and let $L(x_0; x)$ be the line generated by $S(x_0; x_-)$ (see Figure 1(a)).

Setting $f_L(t) = f[\hat{x}(t)]$ with $\hat{x}(t) = (1 - t)x_0 + tx_-$, $t \in \mathscr{R}$, we have

$$f_L(0) = 0 \quad \text{and} \quad f_L(1) < 0.$$

Hence, by convexity (see Figure 1(b)),

$$f_L(t) < 0 \quad \text{for} \quad 0 < t \leq 1 \quad \text{and} \quad f_L(t) > 0 \quad \text{for} \quad t < 0.$$

We may then conclude that

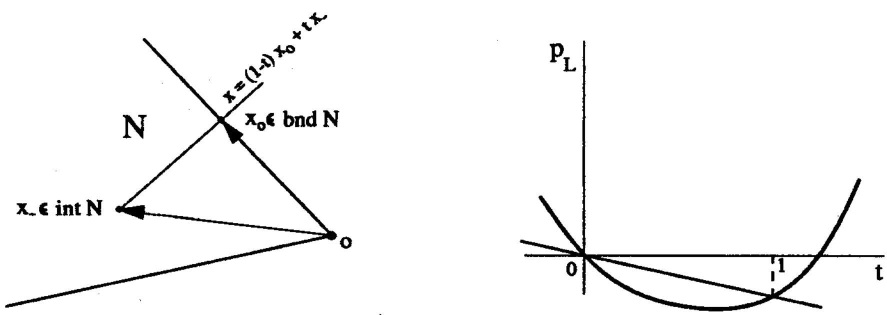
$$N_- \subseteq \text{int } N \quad \text{and} \quad N_0 \subseteq \text{bnd } N$$

and the relations

$$\text{int } N = N \setminus \text{bnd } N \subseteq N \setminus N_0 = N_-,$$

$$\text{bnd } N = N \setminus \text{int } N \subseteq N \setminus N_- = N_0$$

yield the converse inclusions. □




(a) (b)

**Figure 1.** (a) The zero level set of $f(x)$ and (b) the graph of $f_L(t)$.

The main theorem providing the new product rule can now be stated.

**Theorem 3.7** (The New Product Rule). *Let* $m\colon \mathscr{R} \cup \{+\infty\} \mapsto \mathscr{R} \cup \{+\infty\}$ *be a monotone convex function with* $m(+\infty) = +\infty$ *and let* $k\colon X \mapsto \mathscr{R} \cup \{+\infty\}$ *be a proper convex functional continuous at* $x \in X$. *Then, if* $x$ *is not a minimum point of* $k$ *and* $m$ *is subdifferentiable at* $k(x)$, *setting* $f = m \circ k$, *results in*

$$\partial f(x) = \partial m[k(x)]\, \partial k(x).$$

*Proof.* The proof is carried out in two steps. First we provide a representation formula for the closure of the one-sided directional derivative of $f$; then recourse to the two preliminary lemmas will yield the result.

To provide the representation formula, given a director $h \in X - \{0\}$, we define the convex real function $\chi\colon \mathscr{R}^+ \mapsto \mathscr{R}$ as the restriction of $k$ to the half-line starting at $x$ and directed along $h$: $\chi(\alpha) \overset{\text{def}}{=} k(x + \alpha h)$ so that $\chi'(0) \overset{\text{def}}{=} dk(x; h)$. In investigating the behavior of $df(x; \cdot)$ it is basic to consider the zero level set of $dk(x; \cdot)$.

First we observe that the continuity of $k$ at $x$ implies [4] the continuity of the sublinear function $dk(x; h)$ as a function of $h$. Its zero level set $N = \{h \in X | dk(x; h) \leq 0\}$ is then a closed convex cone.

Since by assumption $x$ is not a minimum point for $k$, the preliminary lemma, Lemma 3.6, states that the interior and the boundary of $N$ are not empty, being $dk(x; \cdot) < 0$ in $\operatorname{int} N$ and $dk(x; \cdot) = 0$ on $\operatorname{bnd} N$. The derivative $df(x; h)$ of the product functional $f = m \circ k$ can then be immediately computed along the directions $h \in \operatorname{int} N$ and $h \notin N$. In fact if $dk(x; h) = \chi'(0)$ does not vanish, $\alpha \downarrow 0$ implies that definitively either $\chi(\alpha) \downarrow \chi(0)$ if $\chi'(0) > 0$ or $\chi(\alpha) \uparrow \chi(0)$ if $\chi'(0) < 0$ (see Figure 2).

Hence, denoting the right and left derivatives of $m$ at the point $k(x) = \chi(0)$ by $m'_+$ and $m'_-$, it will be seen that

$$\begin{aligned}
df(x; h) &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha}[f(x + \alpha h) - f(x)] \\
&= \lim_{\alpha \downarrow 0} \frac{1}{\alpha}[m[k(x + \alpha h)] - m[k(x)]] \\
&= \lim_{\alpha \downarrow 0} \frac{1}{\alpha}[m[\chi(\alpha)] - m[\chi(0)]] \\
&= \lim_{\alpha \downarrow 0} \frac{m[\chi(\alpha)] - m[\chi(0)]}{\chi(\alpha) - \chi(0)} \cdot \frac{\chi(\alpha) - \chi(0)}{\alpha} \\
&= \lim_{\chi(\alpha) \downarrow \chi(0)} \frac{m[\chi(\alpha)] - m[\chi(0)]}{\chi(\alpha) - \chi(0)} \cdot \lim_{\alpha \downarrow 0} \frac{\chi(\alpha) - \chi(0)}{\alpha} \\
&= m'_+ \chi'(0)
\end{aligned}$$

if $\chi'(0) > 0$. Apparently $df(x; h) = m'_- \chi'(0)$ if $\chi'(0) < 0$.

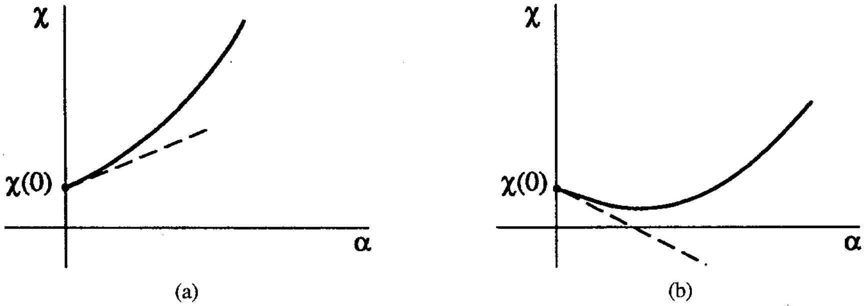



**Figure 2.** (a) $\chi'(0) > 0$ and (b) $\chi'(0) < 0$.

A more detailed discussion has to be made when $h \in \text{bnd}\, N$, so that $\chi'(0) = 0$. In this case, as shown in Figure 3, the convexity of $k$ implies that either $\chi(\alpha)$ goes to $\chi(0)$ with a strict monotonic descent or it attains the value $\chi(0)$ for some $\alpha > 0$ and then remains definitively constant.

In both cases $df(x; h) = 0$ if the right derivative $m'_+$ is finite. In fact, in the case of Figure 3(a), the formula $df(x; h) = m'_+ \, dk(x; h)$ holds with $dk(x; h) = 0$; in the case of Figure 3(b) the conclusion is trivial.

We may then conclude that

$$df(x; \cdot) = \begin{cases} 0 & \text{on bnd } N, \\ m'_- \, dk(x; \cdot) \le 0 & \text{in int } N, \\ m'_+ \, dk(x; \cdot) \ge 0 & \text{outside } N, \end{cases}$$

so that the following formula holds:

$$df(x; h) = \sup_{\lambda \in I} \lambda \, dk(x; h) \qquad \text{with} \quad I = [m'_-, m'_+].$$

An indecisive situation occurs instead when $m'_+ = +\infty$ since, in the case of Figure 3(a), $df(x; h) = +\infty$.

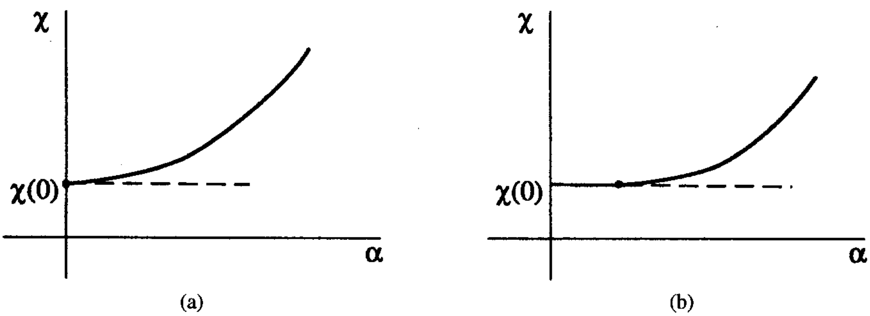



**Figure 3.** Graphs of $\chi(\alpha)$ for $\chi'(0) = 0$. (a) Monotonic descent and (b) definitive constancy.

Noticing that $\partial m[k(x)] = I = [m'_-, m'_+]$, the assumed subdifferentiability of $m$ at $k(x)$ ensures that $m'_- < +\infty$. Hence we get

$$df(x;\cdot) = \begin{cases} 0 \text{ or } +\infty & \text{on bnd } N, \\ m'_- \, dk(x;\cdot) \le 0 & \text{in int } N, \\ m'_+ \, dk(x;\cdot) = +\infty & \text{outside } N. \end{cases}$$

To resolve the indecisive situation on bnd $N$ we observe that, $dk(x;\cdot)$ being continuous in $X$ and vanishing on bnd $N$, the restriction of $df(x;\cdot)$ to int $N$ can be extended by continuity to zero on bnd $N$.

As a consequence the closure of $df(x;\cdot)$ will vanish on bnd $N$, being equal to $df(x;\cdot)$ elsewhere:

$$\overline{df(x;\cdot)} = \begin{cases} 0 & \text{on bnd } N, \\ m'_- \, dk(x;\cdot) \le 0 & \text{in int } N, \\ m'_+ \, dk(x;\cdot) = +\infty & \text{outside } N. \end{cases}$$

From the analysis above we infer then the general validity of the formula

$$\overline{df(x;h)} = \sup_{\lambda \in I} \lambda \, dk(x;h) \qquad \text{with} \quad I = [m'_-, m'_+]$$

holding whether $m'_+$ is finite or not.

To get the product rule we finally remark that, by the continuity of $dk(x;\cdot)$,

$$dk(x;h) = \sup\{\langle x^*, h\rangle | x^* \in \partial k(x)\}$$

so that the formula above may be rewritten as

$$\overline{df(x;h)} = \sup_{\lambda \in I} \{\lambda \sup\{\langle x^*, h\rangle | x^* \in \partial k(x)\}\} = \sup\{\langle x^*, h\rangle | x^* \in I\,\partial k(x)\}.$$

The set $I\,\partial k(x) = \partial m[k(x)]\,\partial k(x)$ being convex by Lemma 3.5, we then get

$$\partial f(x) = \overline{\partial m[k(x)]\,\partial k(x)}.$$

Finally we observe that, by the continuity of $k$ at $x$, the convex set $\partial k(x)$ is nonempty, closed, and weakly compact in $X'$ [7, Proposition 10.c.]; further it does not contain the origin since $x$ is not a minimum point for $k$. By Lemma 3.5 we may then infer the closure of the set $\partial m[k(x)]\,\partial k(x)$ and the proof is complete. □

Typical shapes of the monotone convex function $m$ in the case when $m'_+ = +\infty$ are shown in Figure 4 depending on whether $m'_- > 0$ or $m'_- = 0$. The latter case reveals that a significant special choice for $m$ is the convex indicator of the nonpositive real axis. This is in fact the choice to be made in discussing convex optimization problems.

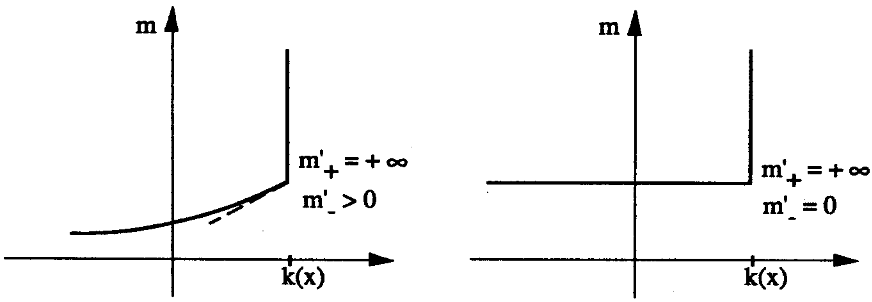



**Figure 4.** Graphs of $m$ when $m'_+ = +\infty$.

### *3.3. Applications to Convex Optimization*

Given a proper convex functional $f$: $X \mapsto \mathscr{R} \cup \{+\infty\}$ we consider the following convex optimization problem:

$$\inf\{f(x) | x \in C\},$$

where $C$ is the feasible set, defined by

$$C = C_g \cap C_h$$

with

$$C_g = \left\{ \bigcap C_i | i = 1, \ldots, n_1 \right\},$$

$$C_h = \left\{ \bigcap C_j | j = 1, \ldots, n_2 \right\},$$

$$C_i = \{x \in X | g_i(x) \leq 0\},$$

$$C_j = \{x \in X | h_j(x) = 0\}.$$

In order that the optimization problem above be meaningful, we have to assume that the intersection between the feasible set and the domain of the objective functional is not empty, i.e., $\operatorname{dom} f \cap C \neq \varnothing$.

Here $g_i$: $X \mapsto \mathscr{R}$ are $n_1$ continuous convex functionals and $h_j$: $X \mapsto \mathscr{R}$ are $n_2$ continuous affine functionals, that is, $h_j(x) = \langle a_j^*, x \rangle + c$ with $a_j^* \in X'$ and $c \in \mathscr{R}$. Without loss of generality the functionals $g_i$ and $h_j$ can be assumed to be nonconstant; further it is natural to assume that each of the convex functionals $g_i$ do assume negative values.

The following preliminary result is easily proved.

**Lemma 3.8.** *Let $g$: $X \mapsto \mathscr{R}$ be a nonconstant continuous convex functional. Denoting its zero level set by $N$, if a vector $x_- \in N$ exists such that $g(x_-) < 0$ we have that*

$$\partial(\operatorname{ind}_{\{\mathscr{R}^-\}} \circ g)(x) = \partial \operatorname{ind}_{\{\mathscr{R}^-\}}[g(x)]\, \partial g_i(x), \qquad \forall x \in N.$$

*Proof.* By Lemma 3.6 it follows that

$$\operatorname{int} N = N_- \stackrel{\text{def}}{=} \{x \in X \,|\, g(x) < 0\},$$

$$\operatorname{bnd} N = N_0 \stackrel{\text{def}}{=} \{x \in X \,|\, g(x) = 0\},$$

and both sets turn out to be nonempty.

Now, if $x \in N_0$ the properties ensuring the validity of the new product-rule formula proved in Section 3.2 are fulfilled. On the other hand, if $x \in N_-$, by the continuity of $g$ there is a neighborhood of $x$ in which $g$ is negative. The formula above then follows by observing that in this case $\partial \operatorname{ind}_{\{\mathscr{R}^-\}}[g(x)] = \{0\}$. □

We are now ready to discuss the convex optimization problem considered above which can be conveniently reformulated as

$$\inf \psi(x) \qquad \text{with} \quad \psi(x) = f(x) + \sum_{i=1}^{n} \operatorname{ind}_{\{\mathscr{R}^-\}}[g_i(x)] + \sum_{j=1}^{m} \operatorname{ind}_{\{0\}}\big[h_j(x)\big].$$

Convex analysis tells us that

$$x_0 = \operatorname{argmin} \psi(x) \Leftrightarrow 0 \in \partial\psi(x_0)$$

or explicitly

$$0 \in \partial\left[f(x_0) + \sum_{i=1}^{n} \operatorname{ind}_{\{\mathscr{R}^-\}}[g_i(x_0)] + \sum_{j=1}^{m} \operatorname{ind}_{\{0\}}\big[h_j(x_0)\big]\right].$$

Under the validity of the addition rule of subdifferential calculus the extremum condition becomes

$$0 \in \partial f(x_0) + \sum_{i=1}^{n} \partial(\operatorname{ind}_{\{\mathscr{R}^-\}} \circ g_i)(x_0) + \sum_{j=1}^{m} \partial(\operatorname{ind}_{\{0\}} \circ h_j)(x_0).$$

Here we apply the result contributed above in Lemma 3.8 to compute the subdifferentials related to inequality constraints:

$$\partial(\operatorname{ind}_{\{\mathscr{R}^-\}} \circ g_i)(x_0) = \partial \operatorname{ind}_{\{\mathscr{R}^-\}}[g_i(x_0)]\, \partial g_i(x_0).$$

The new proof of the chain rule provided in Section 3.1 allows us to carry out computation of the subdifferentials related to equality constraints:

$$\partial(\operatorname{ind}_{\{0\}} \circ h_j)(x_0) = \partial \operatorname{ind}_{\{0\}}\big[h_j(x_0)\big]\, \partial h_j(x_0).$$

Finally we observe that

$$\partial \operatorname{ind}_{\{\mathscr{R}^-\}}[g_i(x_0)] = \mathscr{N}_{\{\mathscr{R}^-\}}[g_i(x_0)],$$

$$\partial \operatorname{ind}_{\{0\}}\big[h_j(x_0)\big] = \mathscr{R}, \quad \text{and} \quad \partial h_j(x_0) = a_j^*,$$

where $\mathcal{N}_{\{\mathscr{R}^-\}}[g_i(x_0)]$ is the normal cone to $\mathscr{R}^-$ at the point $g_i(x_0)$. It turns out to be equal to $\{0\}$ when $g_i(x_0) < 0$ and to $\mathscr{R}^+$ when $g_i(x_0) = 0$.

The extremum condition above can thus be restated explicitly in terms of Kuhn and Tucker complementarity relations:

$$\begin{cases} \lambda_i \in \mathscr{R}^+, \quad g_i(x_0) \in \mathscr{R}^-, \quad \lambda_i g_i(x_0) = 0, \quad i = 1,\dots,n, \\ \mu_j \in \mathscr{R}, \quad j = 1,\dots,m, \end{cases}$$

and of the related stationarity condition:

$$0 \in \partial f(x_0) + \sum_{i=1}^{n} \lambda_i \, \partial g_i(x_0) + \sum_{j=1}^{m} \mu_j a_j^*.$$

The corresponding Lagrangian is given by

$$L(x, \lambda_i, \mu_j) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x) - \sum_{i=1}^{n} \operatorname{ind}_{\{\mathscr{R}^+\}}(\lambda_i) - \sum_{j=1}^{m} \operatorname{ind}_{\{\mathscr{R}\}}(\mu_j),$$

where the last inessential term has been added for formal symmetry.

The Kuhn and Tucker conditions above are easily seen to be equivalent to the existence of a saddle point for the Lagrangian.

Classically the existence of Kuhn and Tucker multipliers is ensured by the fulfillment of Slater's conditions [11], [18]:

$$\exists \bar{x} \in X \quad \text{such that} \quad f(\bar{x}) < +\infty \quad \text{and} \quad \begin{cases} g_i(\bar{x}) < 0, & i = 1,\dots,n, \\ h_j(\bar{x}) = 0, & j = 1,\dots,m, \end{cases}$$

i.e., by assuming that the intersection between the domain of the objective functional and the interior of the set $C_g$ is nonempty.

According to the treatment developed in this paper the existence of Kuhn and Tucker multipliers can in fact be assessed under far less stringent conditions; these amount in the obvious minimal requirement that the optimization problem is well posed (i.e., the intersection between the domain of the objective functional and the feasible set is nonempty) and in the further assumption that, at the optimal point, the property ensuring the validity of the addition rule is satisfied.

The graphical sketches in Figure 5 exemplify the different assumptions about the feasible set $C = C_1 \cap C_2$ in the special case $n_1 = 2$ and $n_2 = 0$. Slater's condition is easily seen to be a straightforward consequence of the classical theorems on addition rule for subdifferentials [11], [14]. The new condition is based on the results provided in the present paper. Validity of the addition rule cannot however be imposed *a priori* but has to be verified *a posteriori* at the extremal point. In this respect it has to be pointed out that when the optimal point $x_0$ lies on the

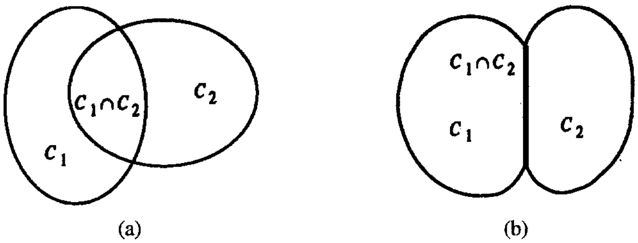



**Figure 5.** Assumptions about the feasible set. (a) Slater's condition and (b) new requirement.

boundary of the set $C_g$ the simple sufficient condition provided by Corollary 3.4 results in a special requirement on the local shape of $C_g$ around $x_0$.

In fact, when $x_0$ belongs to the boundary of one of the sets $C_i$, the one-sided directional derivative $d(\mathrm{ind}_{\{\mathscr{R}^-\}} \circ g_i)(x_0; \cdot)$ will be l.s.c. if and only if the monotone convex function $\mathrm{ind}_{\{\mathscr{R}^-\}}$ is definitively constant toward zero along any direction $h$ such that $dg(x_0; h) = 0$, i.e., $h \in \mathrm{bnd}\, C_i$ (see Lemma 3.6, Theorem 3.7, and Figure 3). This means that there must be an $\alpha_0 > 0$ such that $g(x_0 + \alpha h) = 0$ when $\alpha_0 \geq \alpha \geq 0$. The boundary of $C_g$ must thus have a conical shape around $x_0$ as sketched in Figure 6.

We finally provide an example in which Slater's condition fails, but the existence of Kuhn–Tucker multipliers can still be assessed on the basis of the new results contributed above. To this end we consider a two-dimensional optimization problem for the convex function $f(x, y) = \frac{1}{2}(x^2 + y^2)$ under the following inequality constraints:

$$h_1 = x - 1 \leq 0, \qquad h_2 = -x + 1 \leq 0, \qquad h_3 = y - 2 \leq 0,$$

$$h_4 = -y + 2 \leq 0.$$

It is apparent that the feasible set does have an empty interior so that Slater's condition is not fulfilled. On the contrary the differentiability of the constraint functions ensures the validity of the addition rule so that the new requirement is satisfied.

The feasible set $C$ is depicted in Figure 7 and a set of Kuhn–Tucker multipliers at the optimal point $x = 1$, $y = 1$ is given by $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 0$, $\lambda_4 = 1$.

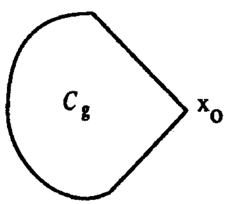



**Figure 6.** Local conical shape around the optimal point.

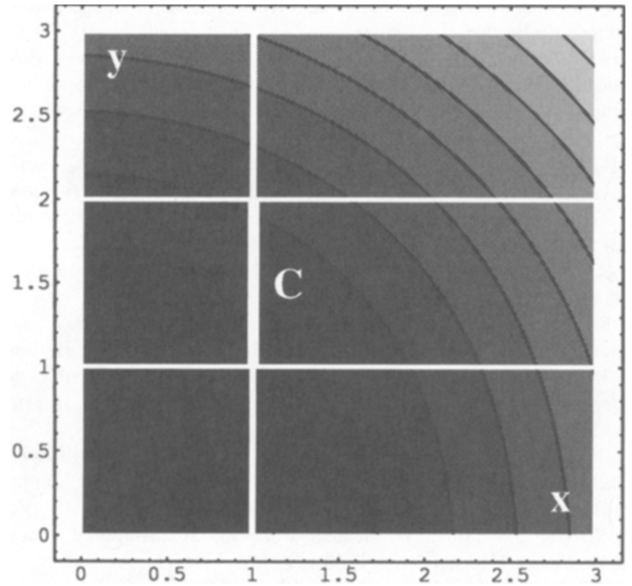



**Figure 7.** Feasible set and contour plot of the objective function.

## 4. Conclusions

The new approach to classical chain and addition rules of subdifferential calculus and the new product-rule formula presented in this paper have been shown to provide a useful and simple tool in the analysis of convex optimization problems. Kuhn–Tucker optimality conditions have been proved under minimal assumptions on the data. Further applications of the results contributed here can be envisaged in different areas of mathematical physics.

The original motivation for the study stemmed from problems in the theory of plasticity. In fact, starting from the classical normality rule of the plastic flow to the convex domain of admissible static states, the new product rule provides a simple and effective tool to derive the equivalent expression of the flow in terms of plastic multipliers and gradients of the yield modes. A comprehensive treatment of the subject can be found in two recent papers by the author and coworkers [15]–[17].

## References

1. Clarke FH (1973) Necessary conditions for nonsmooth problems in optimal control and the calculus of variations. Thesis, University of Washington, Seattle, WA
2. Clarke FH (1983) Optimization and Nonsmooth Analysis. Wiley, New York
3. Ekeland I, Temam R (1974) Analyse convexe et problèmes variationnels. Dunod, Paris
4. Ioffe AD, Tihomirov VM (1979) The Theory of Extremal Problems. Nauka, Moscow (English translation, North-Holland, Amsterdam)
5. Laurent PJ (1972) Approximation et Optimisation. Hermann, Paris
6. McLinden L (1973) Dual operations on saddle functions. Trans Amer Math Soc 179:363–381

7. Moreau JJ (1966) Fonctionelles convexes. Lecture notes, séminaire: équationes aux dérivées partielles, Collège de France
8. Moreau JJ (1973) On unilateral constraints, friction and plasticity. New Variational Techniques in Mathematical Physics, CIME Bressanone, Ed. Cremonese, Roma, pp 171–322
9. Panagiotopoulos PD (1985) Inequality Problems in Mechanics and Applications. Birkhäuser, Boston
10. Rockafellar RT (1963) Convex functions and dual extremum problems. Thesis Harvard, MA
11. Rockafellar RT (1970) Convex Analysis. Princeton University Press, Princeton, NJ
12. Rockafellar RT (1979) Directionally Lipschitzian functions and subdifferential calculus. Proc London Math Soc (3) 39:331–355
13. Rockafellar RT (1980) Generalized directional derivatives and subgradients of nonconvex functions. Canad J Math XXXII(2):257–280
14. Rockafellar RT (1981) The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions. Heldermann Verlag, Berlin
15. Romano G, Rosati L, Marotti de Sciarra F (1992) An internal variable theory of inelastic behaviour derived from the uniaxial rigid-perfectly plastic law. Internat J Engrg Sci 31(8):1105–1120
16. Romano G, Rosati L, Marotti de Sciarra F (1993) Variational principles for a class of finite-step elasto-plastic problems with non-linear mixed hardening. Comput Methods Appl Mech Engrg 109:293–314
17. Romano G, Rosati L, Marotti de Sciarra F (1993) A variational theory for finite-step elasto-plastic problems. Internat J Solids 30(17):2317–2334
18. Slater M (1950) Lagrange multipliers revisited: a contribution to non-linear programming. Cowles Commission Discussion Paper, Math 403

*Accepted 28 October 1993*