On the equality of Clarke and Greenberg-Pierskalla differentials*

Simone Cerreia-Vioglio[a] Fabio Maccheroni[a] Massimo Marinacci[a] Luigi Montrucchio[b]

[a]Department of Decision Sciences and IGIER, Università Bocconi

[b]Collegio Carlo Alberto, Università di Torino

June 2021

Abstract

We study the relations for quasiconcave and continuous functions between the two important notions of Clarke and Greenberg-Pierskalla differentiability. As an application, we generalize the classic Roy's identity of consumer theory.

1 Introduction

Since the seminal studies of de Finetti [21] and Fenchel [22], quasiconvex analysis has been the subject of active research.[1] Starting with the paper of de Finetti [21], this field has been deeply influenced by economic theory. In keeping with this tradition, our purpose here is to relate two notions of differentiability that have been proven useful in both quasiconvex analysis and economic theory: Greenberg-Pierskalla differentiability and Clarke differentiability.[2] Specifically, we study a continuous and quasiconcave function g defined on an open convex subset of a normed vector space X. Our main result, Theorem 1, shows that if g is Clarke differentiable at x and $0 \notin \partial_C g(x)$, then the Greenberg-Pierskalla superdifferential at x coincides with the weak*-closure of the cone generated by the Clarke superdifferential at x. In particular, Corollary 1 shows that the closure is no longer necessary when the Clarke directional derivative is finite. As a further consequence, Corollary 4 establishes a differential characterization of quasiconcavity a la Arrow and Enthoven [3] for Clarke differentiable functions.

The next part of the paper, Section 4, studies a few related issues. In particular, Theorem 2 shows that continuous and quasiconcave functions are directionally Lipschitzian, a nice regularity property. This theorem builds on a result of Crouzeix on the monotonicity of quasiconcave functions with respect to a specific cone. As an application of Theorems 1 and 2, in Theorem 3 we extend to our quasiconcave setting the classic max-formula of convex analysis (an envelope theorem in the economics jargon). In this part, we also formulate a version of our main result that replaces Clarke superdifferentials with other types of superdifferentials. In the last section, Section 5, we derive general versions of the classic Roy's identity of consumer theory: this was actually one of the original motivations of the present paper.

*We thank for the financial support the European Research Council (grants SDDM-TEA and INDIMACRO) and the Ministero dell'università e della ricerca (grant 2017CY2NCA). First Draft: November 2015.

[1]We refer the reader to Penot [32] for a survey. See also Crouzeix [13], [15], and [16], Martinez-Legaz [28] and [29], and Penot and Volle [34].

[2]A stark difference with convex analysis is the presence of different types of dualities as well as different notions of differentiability. See, for example, Komlosi [27], Penot [31], and Cerreia-Vioglio et al. [9].

1.1 Related literature

There are already few papers dealing with the relation between the Greenberg-Pierskalla differential and the cone generated by the Clarke differential. Penot [31, Propositions 25 and 26] implicitly proved the equality under the assumption of Clarke differentiability: this would yield our Corollary 1. Differently, our main results cover the more general class of Clarke-Rockafellar differentiable functions. Daniilidis et al. [19, Proposition 7 and Corollary 12] provided a result on an inclusion, but the Clarke subdifferential is replaced by a larger class of abstract subdifferentials, defined axiomatically according to [5]. This kind of extension will be discussed in Section 4.3.

More recently Kabgani and Soleimani-damaneh in [26] studied the relationship between convexificators and the Greenberg-Pierskalla subdifferentials and proved that the two cones coincide, under the hypotheses that X is finite dimensional and the functional g is locally Lipschitz and Clarke regular at the point x: see [26, Theorem 3.7-(iii)]. Such a result has to be compared with our Corollary 1 in which we show that such an equality still holds under weaker assumptions.

2 Preliminaries

Throughout the paper, $g : C \to \mathbb{R}$ denotes a real-valued function defined on a nonempty convex subset C of a normed space X, with norm dual X^*. An exception is Section 5, in which the vector space X is placed in duality with a space X^*.

2.1 Greenberg-Pierskalla

An important notion of differential in quasi-convex analysis is due to Greenberg and Pierskalla [23].

Definition 1 *The* Greenberg-Pierskalla superdifferential *of a function $g : C \to \mathbb{R}$ at a point $x \in C$ is*

$$\partial_{GP} g(x) = \{\xi \in X^* : y \in C \text{ and } \langle y, \xi\rangle \leq \langle x, \xi\rangle \Longrightarrow g(y) \leq g(x)\}.$$

This notion has an ordinal nature:[3] if $f : \operatorname{Im} g \to \mathbb{R}$ is strictly increasing, then

$$\partial_{GP} g(x) = \partial_{GP}(f \circ g)(x).$$

So, it is a notion well suited for quasiconcave functions. Next we list some of its properties.

1. $\partial_{GP} g(x)$ is a convex cone, not necessarily containing the origin 0.[4]
2. $0 \in \partial_{GP} g(x) \Longleftrightarrow \partial_{GP} g(x) = X^* \Longleftrightarrow x$ is a global maximizer, i.e., $g(x) \geq g(y)$ for every $y \in C$.
3. $\partial_{GP} g(x) \cap -\partial_{GP} g(x) = \emptyset \Longleftrightarrow 0 \notin \partial_{GP} g(x)$.
4. If $\partial_{GP} g(x) \neq \emptyset$ for all $x \in C$, then $g : C \to \mathbb{R}$ is quasiconcave.
5. $\partial_{GP} g(x) \neq \emptyset$ if C is open and g is quasiconcave and lower semicontinuous at x.

Points 1-3 are elementary, while point 4 will be shown later (Corollary 4). Point 5 is easily obtained through a separation theorem.

Recall that a function is *radially continuous* if its restrictions over linear segments are all continuous.

[3] A property is ordinal when it is satisfied by a function and by all its strictly increasing transformations.

[4] A subset A of a vector space is a *cone* if $\xi \in A$ and $\lambda > 0$ imply $\lambda\xi \in A$.

Lemma 1 *Let C be open and g either lower semicontinuous or radially continuous. If $\xi \in \partial_{GP} g(x)$, then for all $y \in C$,*

$$\langle y, \xi \rangle < \langle x, \xi \rangle \implies g(y) \leq g(x). \tag{1}$$

Moreover, if $\xi \neq 0$, then the converse also holds.

Proof The main implication trivially follows by definition of $\partial_{GP} g(x)$. As for the converse, assume that $\xi \neq 0$. We only need to show that if $\langle y, \xi \rangle = \langle x, \xi \rangle$, then $g(y) \leq g(x)$. Since $\xi \neq 0$, there exists some $z \in X$ such that $\langle z, \xi \rangle > 0$. Consider the sequence of points $y_n = y - t_n z$ with $t_n \downarrow 0$. We have $y_n \in C$ for n sufficiently large. Clearly, $\langle y_n, \xi \rangle = \langle y, \xi \rangle - t_n \langle z, \xi \rangle < \langle x, \xi \rangle$. It follows that $g(y_n) \leq g(x)$. If g is lower semicontinuous, by passing to the limit we have $g(y) \leq \liminf_n g(y_n) \leq g(x)$, which is the desired property. Similarly, if g is radially continuous then, by passing to the limit, $g(y_n) = g(y - t_n z) \leq g(x)$ yields again $g(y) \leq g(x)$. ■

Point 3 implies that $\partial_{GP} g(x) \cup \{0\}$ is a pointed convex cone provided $0 \notin \partial_{GP} g(x)$. The next property is then easily proved.

Proposition 1 *Under the hypotheses of Lemma 1, if $0 \notin \partial_{GP} g(x)$ then*

$$\partial_{GP} g(x) \cup \{0\} = \left\{ \xi \in X^* : \forall y \in \tilde{L}_x, \langle y - x, \xi \rangle \geq 0 \right\}.$$

where $\tilde{L}_x$ is the strict superlevel set $(g > g(x))$.

In words, $-\partial_{GP} g(x) \cup \{0\}$ is the (weak*-closed) normal cone to the convex set $\tilde{L}_x = (g > g(x))$.

The Greenberg-Pierskalla superdifferential has the following dual version: the *Greenberg-Pierskalla subdifferential* of a function $g : C \to \mathbb{R}$ at a point $x \in C$ is

$$\partial^{GP} g(x) = \{\xi \in X^* : y \in C \text{ and } \langle y, \xi \rangle \geq \langle x, \xi \rangle \Longrightarrow g(y) \geq g(x)\}.$$

There is a simple duality relation:

$$\partial^{GP} g(x) = -\partial_{GP}(-g)(x). \tag{2}$$

Remark In the literature there are many variants of the Greenberg-Pierskalla subdifferential. All of them are closely related to normal cones of superlevel sets. For instance, property (1) defines the so called star-superdifferential (see [31] and [35]). Another variant is Crouzeix's tangential cone (see [14]).

2.2 Clarke

If $g : C \to \mathbb{R}$ is continuous and C is open, the *Clarke lower (directional) derivative* $g^{\downarrow}(x; \cdot) : X \to [-\infty, \infty]$ of g at x is defined by:

$$g^{\downarrow}(x; y) = \lim_{\varepsilon \downarrow 0} \liminf_{\substack{x' \to x \\ t \downarrow 0}} \sup_{\|y - y'\| < \varepsilon} \frac{g(x' + ty') - g(x')}{t} \qquad \forall y \in X. \tag{3}$$

The function $g^{\downarrow}$ permits to introduce an important notion of superdifferential.[5]

[5] The directional derivative $g^{\downarrow}(x; \cdot)$ is related to the Clarke tangent cone of hypo g at the point $(x, g(x))$. It was introduced by Rockafellar [37] to generalize an earlier notion given by Clarke [10] and [11]. For this reason, the Clarke-Rockafellar terminology is often used. The function $g^{\downarrow}$ may assume several expressions. In particular, (3) relies on the continuity of g (see [40, Proposition 3.2.3]). Comprehensive studies can be found in [12] and [40].

Definition 2 *The* Clarke superdifferential *of a function $g : C \to \mathbb{R}$ at a point $x \in C$ is defined by*

$$\partial_C g(x) = \left\{ \xi \in X^* : \langle y, \xi \rangle \geq g^{\downarrow}(x; y) \text{ for all } y \in X \right\}. \tag{4}$$

The set $\partial_C g(x)$ is a, possibly empty, weak*-closed convex subset of X^* (see [37, Theorem 4 and p. 276]). The dual notion of *Clarke upper derivative* $g^{\uparrow}(x; \cdot) : X \to [-\infty, \infty]$ at $x \in C$ is defined by

$$g^{\uparrow}(x; y) = \lim_{\varepsilon \downarrow 0} \limsup_{\substack{x' \to x \\ t \downarrow 0}} \inf_{\|y - y'\| < \varepsilon} \frac{g(x' + ty') - g(x')}{t} \qquad \forall y \in X.$$

The *Clarke subdifferential* is then $\partial^C g(x) = \left\{ \xi \in X^* : \forall y \in X, \langle y, \xi \rangle \leq g^{\uparrow}(x; y) \right\}$. We have the simple relations:

$$g^{\uparrow}(x; y) = -(-g)^{\downarrow}(x; y) \quad \text{and} \quad \partial^C g(x) = -\partial_C(-g)(x). \tag{5}$$

3 Main result

The next theorem, our main result, relates the Greenberg-Pierskalla and Clarke superdifferentials.

Theorem 1 *Let C be open and g continuous and quasiconcave. Then,*

$$\partial_C g(x) \subseteq \partial_{GP} g(x) \cup \{0\} \qquad \forall x \in C.$$

Moreover,

$$\overline{\text{cone}}^* \partial_C g(x) = \partial_{GP} g(x) \cup \{0\} \tag{6}$$

provided $0 \notin \partial_C g(x) \neq \emptyset$.[6]

The additional hypothesis for (6) is needed: for the quasiconcave function $g(t) = \sqrt[3]{t}$ we have $\partial_C g(0) = \emptyset$ and $\partial_{GP} g(0) = (0, \infty)$, while for the function $g(t) = -t^+$ we have $0 \in \partial_C g(0) = [-1, 0]$ and $\partial_{GP} g(0) = \mathbb{R}$ and so $\overline{\text{cone}}^* \partial_C g(x)$ is strictly included in $\partial_{GP} g(x) \cup \{0\}$.

Theorem 1 has been formulated through the superdifferential $\partial_C g(x)$. Later, it will be shown that under our assumptions the Clarke superdifferential and subdifferential coincide (Corollary 5). So, (6) can be replaced by the equivalent relation $\overline{\text{cone}}^* \partial^C g(x) = \partial_{GP} g(x) \cup \{0\}$.

Proof We divide the proof in three steps.

(i) We first show that

$$\partial_C g(x) \subseteq \partial_{GP} g(x) \cup \{0\} \qquad \forall x \in C.$$

If $\xi = 0$, this inclusion is trivially true. So, let $0 \neq \xi \in \partial_C g(x)$ and $y \in C$ such that $\langle y, \xi \rangle < \langle x, \xi \rangle$. That is, $\langle y - x, \xi \rangle = \alpha < 0$. By (4), $g^{\downarrow}(x; y - x) \leq \alpha < 0$.

By the definition (3), there exist sequences $x_n \to x$, $z_n \to y - x$, and $t_n \downarrow 0$ such that $g(x_n + t_n z_n) - g(x_n) < 0$ and $t_n < 1$ for n large enough. Setting $y_n = z_n + x_n$, we have $y_n \to y$ and $g(t_n y_n + (1 - t_n) x_n) < g(x_n)$. Since g is quasiconcave, we have

$$g(x_n) > g(t_n y_n + (1 - t_n) x_n) \geq \min \{ g(y_n), g(x_n) \}.$$

Hence, $g(x_n) > g(y_n)$. By continuity, letting $n \to \infty$, we get $g(x) \geq g(y)$. Lemma 1 then implies that $\xi \in \partial_{GP} g(x)$, as desired.

[6] Given a nonempty subset K of X^*, we define by $\text{cone}\, K$ the set of vectors $\lambda \xi$ where $\lambda \geq 0$ and $\xi \in K$. The set $\overline{\text{cone}}^* K$ is its weak*-closure. This notion of generated cone is an abuse of terminology because, compared to our notion of cone, it allows for $\lambda = 0$. Yet, when K is a closed cone, then $K = \text{cone}\, K$, which is the relevant case for us.

(ii). Point (i) implies $\operatorname{cone} \partial_C g(x) \subseteq \partial_{GP} g(x) \cup \{0\}$. By Proposition 1, $\partial_{GP} g(x) \cup \{0\}$ is a weak*-closed convex cone. Therefore, $\overline{\operatorname{cone}}^* \partial_C g(x) \subseteq \partial_{GP} g(x) \cup \{0\}$.

(iii). We need to show (6), that is, $\overline{\operatorname{cone}}^* \partial_C g(x) = \partial_{GP} g(x) \cup \{0\}$. Suppose, by contradiction, that $\overline{\operatorname{cone}}^* \partial_C g(x)$ is strictly included in $\partial_{GP} g(x) \cup \{0\}$, namely, there exists $0 \neq \bar{\xi} \in \partial_{GP} g(x)$ such that $\bar{\xi} \notin \overline{\operatorname{cone}}^* \partial_C g(x)$. By the Strong Separating Hyperplane Theorem (see, e.g., [39, Theorem 3.4]), there exist $y_1 \in X$ and $\gamma \in \mathbb{R}$ such that

$$\langle y_1, \bar{\xi} \rangle < \gamma < \langle y_1, \xi \rangle \qquad \forall \xi \in \overline{\operatorname{cone}}^* \partial_C g(x).$$

As $0 \in \overline{\operatorname{cone}}^* \partial_C g(x)$, it follows that $\gamma < 0$. Moreover, as $\gamma < \langle y_1, \lambda \xi \rangle$ holds for all $\lambda > 0$ and $\xi \in \overline{\operatorname{cone}}^* \partial_C g(x)$, we have that $\langle y_1, \xi \rangle \geq 0$ holds over $\overline{\operatorname{cone}}^* \partial_C g(x)$. Hence,

$$\langle y_1, \bar{\xi} \rangle < \gamma < 0 \leq \langle y_1, \xi \rangle \qquad \forall \xi \in \overline{\operatorname{cone}}^* \partial_C g(x).$$

In particular,

$$\langle y_1, \bar{\xi} \rangle < \gamma < 0 \leq \langle y_1, \xi \rangle \qquad \forall \xi \in \partial_C g(x).$$

On the other hand, the set $\partial_C g(x)$ is nonempty, weak*-closed and convex, with $0 \notin \partial_C g(x)$. By the Strong Separating Hyperplane Theorem (see, e.g., [39, Theorem 3.4]), there exist a vector $\bar{y} \in X$ and $\eta \in \mathbb{R}$ such that $\langle \bar{y}, \xi \rangle \geq \eta > 0$ for all $\xi \in \partial_C g(x)$.

Setting $z = y_1 + \varepsilon \bar{y}$, we have $\langle z, \xi \rangle = \langle y_1, \xi \rangle + \varepsilon \langle \bar{y}, \xi \rangle \geq \varepsilon \eta$ for all $\xi \in \partial_C g(x)$. Consequently, for $\varepsilon > 0$ sufficiently small, we have

$$\langle z, \bar{\xi} \rangle < \gamma < 0 < \varepsilon \eta \leq \langle z, \xi \rangle \qquad \forall \xi \in \partial_C g(x).$$

By [37, Theorem 4 and p. 276], $g^{\downarrow}(x; z) = \inf \{\langle z, \xi \rangle : \xi \in \partial_C g(x)\}$. Hence, $\langle z, \bar{\xi} \rangle < \gamma < 0 < \delta \leq g^{\downarrow}(x; z)$, where $\delta = \varepsilon \eta$. In view of definition (3), it holds

$$g^{\downarrow}(x; z) \leq \lim_{\varepsilon \downarrow 0} \liminf_{t \downarrow 0} \sup_{\|z - z'\| < \varepsilon} \frac{g(x + t z') - g(x)}{t}.$$

Therefore, there exists a sequence $z_n \to z$ and $t_n \downarrow 0$ such that, eventually,

$$\frac{g(x + t_n z_n) - g(x)}{t_n} > \frac{\delta}{2}.$$

Hence, it eventually holds

$$\langle z_n, \bar{\xi} \rangle < \gamma < 0 < \delta/2 \leq \frac{g(x + t_n z_n) - g(x)}{t_n}. \tag{7}$$

But, this is a contradiction. Indeed, since $g(x + t_n z_n) > g(x)$, it follows that $\langle x + t_n z_n, \bar{\xi} \rangle > \langle x, \bar{\xi} \rangle$, as $\bar{\xi}$ is a Greenberg-Pierskalla superdifferential at x. Namely, $\langle z_n, \bar{\xi} \rangle > 0$, which contradicts (7). We conclude that $\overline{\operatorname{cone}}^* \partial_C g(x) = \partial_{GP} g(x) \cup \{0\}$, i.e., (6) holds. ■

The dual of Theorem 1 for quasiconvex functions can be easily established via the duality relations (2) and (5). More interestingly, next we present an important corollary in which the closure of the cone is no longer necessary.

Corollary 1 *Let C be open and g continuous and quasiconcave. If at a point $x \in C$ the Clarke lower derivative $g^{\downarrow}(x; y)$ of g is finite for all $y \in X$, then*

$$\operatorname{cone} \partial_C g(x) = \partial_{GP} g(x) \cup \{0\}$$

provided $0 \notin \partial_C g(x)$. In particular,

$$\partial_{GP} g(x) = \{\lambda \xi : \lambda > 0 \text{ and } \xi \in \partial_C g(x)\}.$$

Locally Lipschitz functions around x are an important class of functions satisfying the finiteness condition for the lower derivative. For these functions, directional derivatives can be expressed via the original Clarke directional derivatives g^0 and g_0, that is, $g^{\uparrow}(x;y) = g^0(x;y)$ and $g^{\downarrow}(x;y) = g_0(x;y)$, where $g^0(x;y) = \limsup_{x'\to x,\ t\downarrow 0} [g(x'+ty) - g(x')]/t$.

Proof From [37, Corollary 2, p. 275] it follows that in this case the superdifferential $\partial_C g(x)$ is nonempty weak*-compact and convex. It remains to check that the cone generated by $\partial_C g(x)$ is weak*-closed. To show this, we prove a more general fact. We show that if D is a nonempty and weak*-compact subset of X^* which does not contain 0, then $\operatorname{cone} D$ is weak*-closed. Thus, consider a net $\{\xi_\alpha\}_{\alpha\in A} \subseteq \operatorname{cone} D$ such that $\xi_\alpha \to \xi$ in X^*. By definition of $\operatorname{cone} D$, there exists $\{\lambda_\alpha\}_{\alpha\in A} \subseteq [0,\infty)$ and $\{\bar{\xi}_\alpha\}_{\alpha\in A} \subseteq D$ such that $\xi_\alpha = \lambda_\alpha \bar{\xi}_\alpha$ for all $\alpha \in A$. Since D is weak*-compact, it follows that there exists a subnet $\{\bar{\xi}_{\alpha_\beta}\}_{\beta\in B}$ and $\bar{\xi} \in D$ such that $\bar{\xi}_{\alpha_\beta} \to \bar{\xi}$. Since $0 \notin D$, then $\bar{\xi} \neq 0$ and there exists $\hat{z} \in X$ such that $\langle \hat{z}, \bar{\xi}\rangle > 0$. Since also $\xi_{\alpha_\beta} \to \xi$ and $\bar{\xi}_{\alpha_\beta} \to \bar{\xi}$, we have that

$$\lambda_{\alpha_\beta}\left\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\right\rangle = \left\langle \hat{z}, \lambda_{\alpha_\beta}\bar{\xi}_{\alpha_\beta}\right\rangle = \left\langle \hat{z}, \xi_{\alpha_\beta}\right\rangle \to \langle \hat{z}, \xi\rangle \text{ and } \left\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\right\rangle \to \left\langle \hat{z}, \bar{\xi}\right\rangle > 0.$$

The latter fact implies that $\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\rangle$ is eventually strictly positive. This yields that eventually λ_{α_β} can be written as $\lambda_{\alpha_\beta}\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\rangle / \langle \hat{z}, \bar{\xi}_{\alpha_\beta}\rangle$, proving that

$$\lambda_{\alpha_\beta} = \frac{\lambda_{\alpha_\beta}\left\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\right\rangle}{\left\langle \hat{z}, \bar{\xi}_{\alpha_\beta}\right\rangle} \to \frac{\langle \hat{z}, \xi\rangle}{\left\langle \hat{z}, \bar{\xi}\right\rangle} \overset{\text{def}}{=} \bar{\lambda}.$$

Since each λ_{α_β} is nonnegative, we have that $\bar{\lambda} \geq 0$. Moreover, we also have that

$$\lambda_{\alpha_\beta}\bar{\xi}_{\alpha_\beta} = \xi_{\alpha_\beta} \to \xi \text{ and } \lambda_{\alpha_\beta}\bar{\xi}_{\alpha_\beta} \to \bar{\lambda}\bar{\xi}.$$

Since the limit is unique, we can conclude that $\xi = \bar{\lambda}\bar{\xi} \in \operatorname{cone} D$. Since $0 \notin \partial_C g(x)$, the rest of the statement trivially follows. ■

An immediate consequence of this corollary is the following description of the cone $\partial_{GP} g(x) \cup \{0\}$ when the function is concave. Here ∂_{FM} denotes the usual Fenchel-Moreau superdifferential of convex analysis.

Corollary 2 *Let C be open and g continuous and concave. If $0 \notin \partial_{FM} g(x)$, then*

$$\operatorname{cone} \partial_{FM} g(x) = \partial_{GP} g(x) \cup \{0\}.$$

In particular, $\partial_{GP} g(x) = \{\lambda\xi : \lambda > 0 \text{ and } \xi \in \partial_{FM} g(x)\}$.

Proof The function g is locally Lipschitz, so Corollary 1 applies. Moreover, the Clarke superdifferential coincides with the superdifferential $\partial_{FM} g(x)$ of convex analysis. ■

When g is strictly differentiable (in the full limit sense) at x, $\partial_C g(x)$ is a singleton (and in particular weak*-compact). The only element of $\partial_C g(x)$ is denoted by $\nabla g(x)$.[7] The next corollary is a direct consequence of Corollary 1. Later, a more general result will be proved. In Proposition 3, we will show that (8) holds even if C fails to be open and g is merely Gateaux differentiable (cf. also (22) below).

Corollary 3 *Let C be open and g continuous and quasiconcave. If g is strictly differentiable at $x \in C$, with $\nabla g(x) \neq 0$, then*

$$\partial_{GP} g(x) = \{\lambda \nabla g(x) : \lambda > 0\}. \tag{8}$$

[7] See Rockafellar [36, Proposition 4 and p. 340].

The next corollary presents a differential characterization of quasiconcavity that generalizes the one established by Arrow and Enthoven [3] for differentiable functions.[8]

Corollary 4 *Let C be open and g continuous. If $0 \notin \partial_C g(x) \neq \emptyset$ for every $x \in C$, then the following statements are equivalent:*

(i) g is quasiconcave;

(ii) for each $x \in C$ and $\xi \in \partial_C g(x)$,

$$g(y) > g(x) \implies \langle y, \xi \rangle \geq \langle x, \xi \rangle ; \tag{9}$$

(iii) $\partial_{GP} g(x) \neq \emptyset$ for all $x \in C$.

For another characterization through the quasimonotonicity property of the superdifferential, we refer to Aussell et al. [5].

Proof (i) implies (ii). Let $\xi \in \partial_C g(x)$ and $x \in C$. By Theorem 1, we have that $\xi \in \partial_{GP} g(x)$. In view of Lemma 1, condition (9) is just the property that characterizes the nonnull elements of $\partial_{GP} g(x)$.

By Lemma 1 and since $0 \notin \partial_C g(x) \neq \emptyset$ for all $x \in C$, (ii) trivially implies (iii). It remains to prove that (iii) implies (i). For any $\lambda \in \mathbb{R}$, denote by $\tilde{L}_\lambda$ the open superlevel set $(g > \lambda)$. We can assume $\emptyset \neq \tilde{L}_\lambda \neq C$ which are trivial cases. Pick a vector $\xi_x \in \partial_{GP} g(x)$ for every $x \in C$ and consider the open half-space $H^+(x) = \{y \in X : \langle y - x, \xi_x \rangle > 0\}$. Clearly, $\tilde{L}_\lambda \subseteq C \cap \left[\bigcap_{x \in C \setminus \tilde{L}_\lambda} H^+(x)\right]$. On the other hand $x \notin H^+(x)$ for all $x \in C$. Hence, $\tilde{L}_\lambda = C \cap \left[\bigcap_{x \in C \setminus \tilde{L}_\lambda} H^+(x)\right]$ and so $\tilde{L}_\lambda$ is convex. We conclude that g is quasiconcave. ■

4 Extensions and refinements

4.1 Directionally Lipschitzian functions

In general, the relation between the two directional derivatives $g^\uparrow$ and $g^\downarrow$ is rather difficult to establish. Yet, Rockafellar [36] and [37] isolates a class of functions, termed *directionally Lipschitzian*, for which this relation is easier to study.

Definition 3 *A continuous function $g : C \to \mathbb{R}$ is* directionally Lipschitzian at $x \in C$ with respect to $y \in X$ *if*

$$\inf_{\varepsilon > 0} \limsup_{x' \to x,\ t \downarrow 0} \sup_{\|y - y'\| < \varepsilon} \frac{g(x' + ty') - g(x')}{t} < \infty. \tag{10}$$

We say that g is *directionally Lipschitzian at x* if there is at least a direction y along which g is directionally Lipschitzian at x. It is well known that if g is directionally Lipschitzian at x with respect to $y = 0$, then g is locally Lipschitz around x.

Rockafellar [37, Theorems 3 and 6] shows, *inter alia*, that $g^\uparrow(x; y) = -g^\downarrow(x; -y)$ for all $y \in X$ and so $\partial_C g(x) = \partial^C g(x)$, whenever g is directionally Lipschitzian at x. Moreover, he establishes several criteria that guarantee that a function g is directionally Lipschitzian at a point x (cf. [37, Proposition 4]).

Thanks to an elegant result of Crouzeix [17, Theorem 3.1], we can show that quasiconcave functions are directionally Lipschitzian. For completeness, we first provide a proof of Crouzeix's result (see also Crouzeix et al. [18, Proposition 4]).

[8] See also Komlosi [27, Theorem 10.4] for a characterization in terms of Dini derivatives, as well as Hadjisavvas [24] for further characterizations related to the quasiconcavity of the subdifferential operators.

Lemma 2 *Let C be open and g continuous and quasiconcave. If $x_0 \notin \operatorname{argmax}_C g$, then there exist an open convex neighborhood $V(x_0)$ of x_0 and a nonempty, open, and convex cone K such that $0 \notin K$ and for each $x, y \in V(x_0)$*

$$y - x \in K \Longrightarrow g(y) \geq g(x).$$

Proof By hypothesis, there exists a point $x_1 \in C$ and a scalar λ such that $g(x_1) > \lambda > g(x_0)$. Since g is continuous, there are two open balls of ray η contained in C, $B_\eta(x_1)$ and $B_\eta(x_0)$, such that $g(x) > \lambda$ for all $x \in B_\eta(x_1)$ and $g(x) < \lambda$ for all $x \in B_\eta(x_0)$. Since $B_\eta(x_0) \subseteq C$, there exists $\bar{t} > 0$ such that $x_2 \stackrel{\text{def}}{=} x_0 + \bar{t}(x_0 - x_1) \in B_\eta(x_0) \subseteq C$. Define the set

$$K = \{d : \exists t > 0, x_2 + td \in B_\eta(x_1)\}.$$

The set K is a nonempty, open, and convex cone. In fact, since

$$x_2 + td \in B_\eta(x_1) \Longleftrightarrow d \in t^{-1}[B_\eta(x_1) - x_2],$$

the set K can be rewritten as $K = \cup_{\alpha>0}\alpha\Gamma$ where Γ is the open ball $B_\eta(x_1) - x_2 = B_\eta(x_1 - x_2)$. Moreover, since $x_2 \in B_\eta(x_0)$ and $B_\eta(x_0) \cap B_\eta(x_1) = \emptyset$, we have that $0 \notin K$. Define $V(x_0) = (x_2 + K) \cap B_\eta(x_0)$. Since K is open and convex, so is $V(x_0)$. Moreover, since $x_1 - x_0 \in K$, $x_0 \in B_\eta(x_0)$, and $td \in K$ for all $t > 0$ and for all $d \in K$, we have that $x_0 = x_2 + \bar{t}(x_1 - x_0) \in x_2 + K$, yielding that $x_0 \in V(x_0) \neq \emptyset$. Next, consider $x, y \in V(x_0)$ with $y - x \in K$. Since $y \in V(x_0)$, it follows that $y = x_2 + d \in B_\eta(x_0)$ for some $d \in K$. Since $d \in K$, we have that $x_2 + \tilde{t}d \in B_\eta(x_1)$ for some $\tilde{t} > 0$. Note that $\tilde{t} > 1$.[9] If we define $z_2 = x_2 + \tilde{t}d \in B_\eta(x_1)$ and $s = 1 - 1/\tilde{t}$, then

$$y = x_2 + d = sx_2 + (1 - s)z_2. \tag{11}$$

Since $y - x \in K$, we have that $y = x + \hat{d}$ for some $\hat{d} \in K$. Since $\hat{d} \in K$, we have that there exists $\check{t} > 0$ such that $z_1 \stackrel{\text{def}}{=} x_2 + \check{t}\hat{d} \in B_\eta(x_1)$. By (11), it follows that

$$y = sz_1 + (1 - s)z_2 - s\check{t}\hat{d} = z - \tau(y - x) \tag{12}$$

where $z = sz_1 + (1 - s)z_2 \in B_\eta(x_1)$ and $\tau = s\check{t} > 0$. By rearranging (12), we obtain that

$$y = \frac{1}{1+\tau}z + \frac{\tau}{1+\tau}x.$$

Since g is quasiconcave and $z \in B_\eta(x_1)$ and $x \in B_\eta(x_0)$, it follows that $g(y) \geq \min\{g(z), g(x)\} = g(x)$. ■

Theorem 2 *If C is open and g is continuous and quasiconcave, then g is directionally Lipschitzian at each point $x \notin \operatorname{argmax}_C g$.*

The argmax condition cannot be avoided. For the quasiconcave function $g(x) = -\sqrt{|x|}$, at the maximizer 0 we have $g^{\uparrow}(0; 0) = -\infty$. Thus $\partial^C g(0) = \emptyset$, while $0 \in \partial_C g(0)$. Consequently, $\partial_C g(0) \neq \partial^C g(0)$.

Proof Condition (10) is verified if there exist scalars $\varepsilon, \delta_1, \delta_2 > 0$ such that

$$\sup_{t \in (0,\delta_2)} \sup_{\|x - x'\| < \delta_1} \sup_{\|y - y'\| < \varepsilon} \frac{g(x' + ty') - g(x')}{t} < \infty. \tag{13}$$

Let $x \notin \operatorname{argmax}_C g$. By the previous lemma there exists an open neighborhood $B_\eta(x)$ of ray η, and an open convex cone K, such that if $x_1, x_2 \in B_\eta(x) \subseteq C$ and $x_2 - x_1 \in K$ then $g(x_2) \geq g(x_1)$. Since K is nonempty

[9] In fact, since $x_2 + \tilde{t}d \in B_\eta(x_1)$, $y = x_2 + d \in B_\eta(x_0)$, and $B_\eta(x_0) \cap B_\eta(x_1) = \emptyset$, clearly, $\tilde{t} \neq 1$. Since $x_2 \in B_\eta(x_0)$, if $\tilde{t} \in (0, 1)$, then $x_2 + \tilde{t}d = (1 - \tilde{t})x_2 + \tilde{t}(x_2 + d) \in B_\eta(x_0)$, a contradiction.

and open, consider a vector $y \in -K\setminus\{0\}$. As K is open, there exists $\varepsilon > 0$ such that $\|y - y'\| < \varepsilon$ implies $y' \in -K\setminus\{0\}$. Therefore, under the conditions

$$\|x - x'\| < \eta/3 \quad , \quad 0 < t < \min\left\{\frac{\eta}{3\varepsilon}, \frac{\eta}{3\|y\|}\right\} \quad , \quad \|y - y'\| < \varepsilon$$

we have

$$\|(x' + ty') - x\| = \|(x' - x) + t(y' - y) + ty\| \leq \|x' - x\| + t\|y' - y\| + t\|y\| < \frac{\eta}{3} + \frac{\eta}{3} + \frac{\eta}{3}.$$

Consequently, $x' + ty' \in B_\eta(x)$. Moreover, $(x' + ty') - x' \in -K$ and so $g(x' + ty') - g(x') \leq 0$. We can infer that (13) holds by setting $\delta_1 = \eta/3$ and $\delta_2 = \min\{\eta/3\varepsilon, \eta/3\|y\|\}$. So, g is directionally Lipschitzian with respect to each vector $y \in -K\setminus\{0\}$. ■

As a corollary of this theorem, we get the result announced in the previous section: the equivalence between the Clarke superdifferential and subdifferential for continuous quasiconcave functions.

Corollary 5 *Let C be open and g continuous and quasiconcave. We have $\partial_C g(x) = \partial^C g(x)$ and*

$$\overline{\text{cone}}^*\left[\partial^C g(x)\right] = \partial_{GP} g(x) \cup \{0\}$$

provided $0 \notin \partial_C g(x) \neq \emptyset$.

Proof Since $0 \notin \partial_C g(x)$, we have that x is not a global maximizer. By Theorem 2, this implies that g is directionally Lipschitzian at x. By [37, Theorem 6], this implies that $\partial_C g(x) = \partial^C g(x)$. By Theorem 1, the rest of the statement follows. ■

Next result about quasiaffine functions is another straightforward consequence of Theorem 2.

Corollary 6 *Let C be open and g continuous and quasiaffine. Then*

$$\partial_{GP} g(x) = \partial^{GP} g(x)$$

provided $0 \notin \partial_C g(x) \neq \emptyset$.

Proof As g is quasiconcave, by Corollary 5 we have $\partial_{GP} g(x) \cup \{0\} = \overline{\text{cone}}^*\left[\partial^C g(x)\right]$. Since g is quasiconvex as well, by Theorem 1 we have $\overline{\text{cone}}^*\left[\partial^C g(x)\right] = \partial^{GP} g(x) \cup \{0\}$, which leads to the desired equality. ■

4.2 A min-formula

The next result, an application of our main results (Theorems 1 and 2), extends to our quasiconcave setting the classic max-formula of convex analysis.[10] To this end, define the *lower envelope* $\bar{g} : C \to \mathbb{R}$ of a collection $\{g_i\}_{i \in I}$ of real-valued functions over C by

$$\bar{g}(x) = \inf_{i \in I} g_i(x) \qquad \forall x \in C$$

and set $M(x) = \operatorname{argmin}_{i \in I} g_i(x) = \{i \in I : \bar{g}(x) = g_i(x)\}$.

Theorem 3 *Let C be open, $x_0 \in C$, and I a compact space. If $\{g_i\}_{i \in I} \subseteq \mathbb{R}^C$ is such that*

1. *g_i is quasiconcave on C for all $i \in I$,*

[10] See, e.g., [33, Proposition 3.42] and [40, Theorem 2.4.18].

2. *the map* $I \times C \ni (i, x) \mapsto g_i(x)$ *is continuous,*

3. $0 \notin \overline{\text{co}}^* \left(\cup_{i \in M(x_0)} \partial_C g_i(x_0)\right)$,

4. $g_i^{\downarrow}(x_0; y)$ *is finite for all* $y \in X$ *and for all* $i \in M(x_0)$,

then the lower envelope $\bar{g}$ *is continuous, quasiconcave, and*

$$\partial_{GP}\bar{g}(x_0) \cup \{0\} = \overline{\text{co}}^* \left(\cup_{i \in M(x_0)} \partial_{GP} g_i(x_0)\right).$$

In words, under mild conditions the Greenberg-Pierskalla superdifferential, at a point x_0, of the lower envelope of a family of quasiconcave functions is the closed convex hull of the union of the Greenberg-Pierskalla superdifferentials of the functions attaining the minimum at x_0.

In economics, a result of this kind is called an envelope theorem. It is typically stated as a characterization of the derivative of the value function in a parameterized optimization problem

$$\max_a f(a, \theta) \quad \text{sub}\, a \in A$$

with objective function $f : A \times \Theta \to \mathbb{R}$. The object of interest is the marginal behavior of the value function $v : \Theta \to (-\infty, \infty]$ defined by $v(\theta) = \sup_{a \in A} f(a, \theta)$ for all $\theta \in \Theta$. To frame our Theorem 3 as an envelope theorem, assume that Θ is an open and convex set of a normed space X and that A is a compact topological space. Denote by $\sigma : \Theta \rightrightarrows A$ the solution correspondence defined by $\sigma(\theta) = \text{argmax}_{a \in A} f(a, \theta)$ for all $\theta \in \Theta$. If f is jointly continuous and quasiconvex in the second argument, with $0 \notin \overline{\text{co}}^* \left(\cup_{a \in \sigma(\theta_0)} \partial^C f_a(\theta_0)\right)$ and $f_a^{\uparrow}(\theta_0; y)$ finite for all $y \in X$ and all $a \in \sigma(\theta_0)$,[11] then

$$\partial^{GP} v(\theta_0) \cup \{0\} = \overline{\text{co}}^* \left(\cup_{a \in \sigma(\theta_0)} \partial^{GP} f_a(\theta_0)\right). \tag{14}$$

In the important case when A is convex and f is strictly quasiconcave in the first argument, σ becomes a function and so this equality takes the simpler form

$$\partial^{GP} v(\theta_0) \cup \{0\} = \overline{\partial^{GP} f_{\sigma(\theta_0)}(\theta_0)}^*$$

By Proposition 3 below, when v is Gateaux differentiable at θ_0, and so is f_a for some $a \in \sigma(\theta_0)$ with both $\nabla v(\theta_0)$ and $\nabla f_a(\theta_0)$ nonzero, then (14) amounts to the existence of a scalar $\gamma > 0$ such that $\nabla v(\theta_0) = \gamma \nabla f_a(\theta_0)$. It is then routine to prove that $\gamma = 1$, so that

$$\nabla v(\theta_0) = \nabla f_a(\theta_0)$$

This is the basic differential form of the envelope theorem a la Danskin (1966), which can thus be viewed as a special case of our min-formula.

Proof By Berge's Maximum Theorem (see, e.g., [1, Theorem 17.31]), $\bar{g}$ is continuous and the nonempty- and compact-valued correspondence $x \mapsto M(x)$ is upper hemicontinuous. It is also routine to prove that $\bar{g}$ is quasiconcave.

Fix $x_0 \in C$ and consider $\xi \in \cup_{i \in M(x_0)} \partial_{GP} g_i(x_0)$, that is, $\xi \in \partial_{GP} g_i(x_0)$ for some $i \in M(x_0)$. Since $\bar{g}(x_0) = g_i(x_0)$, for each $y \in C$ we have that

$$\langle y, \xi \rangle \leq \langle x_0, \xi \rangle \implies \bar{g}(y) \leq g_i(y) \leq g_i(x_0) = \bar{g}(x_0),$$

yielding that $\xi \in \partial_{GP}\bar{g}(x_0)$, that is, $\cup_{i \in M(x_0)} \partial_{GP} g_i(x_0) \subseteq \partial_{GP}\bar{g}(x_0)$. By Lemma 1, $\partial_{GP}\bar{g}(x_0) \cup \{0\}$ is a closed and convex cone. This implies that $\overline{\text{co}}^* \left(\cup_{i \in M(x_0)} \partial_{GP} g_i(x_0)\right) \subseteq \partial_{GP}\bar{g}(x_0) \cup \{0\}$.

[11] Here f_a is the section of f at a.

We next prove the opposite inclusion. Set $D = \overline{\text{co}}^* \left(\cup_{i\in M(x_0)} \partial_{GP} g_i (x_0)\right)$ and, by contradiction, assume that $(\partial_{GP}\bar{g}(x_0) \cup \{0\}) \setminus D \neq \emptyset$. Let $\bar{\xi} \in (\partial_{GP}\bar{g}(x_0) \cup \{0\}) \setminus D \neq \emptyset$. Since $0 \in D$, we have that $\bar{\xi} \neq 0$ and so $\bar{\xi} \in \partial_{GP}\bar{g}(x_0)$. Following a similar argument to the one contained in the proof of Theorem 1, by the Strong Separating Hyperplane Theorem and since D is a closed convex cone, there exists an element $\hat{y} \in X$ such that

$$\langle \hat{y}, \xi \rangle \geq 0 > \langle \hat{y}, \bar{\xi} \rangle \qquad \forall \xi \in D.$$

By Theorem 1 and since $0 \notin \partial_C g_i (x_0)$ for all $i \in M(x_0)$, we have that $\partial_C g_i (x_0) \subseteq \partial_{GP} g_i (x_0) \subseteq D$ for all $i \in M(x_0)$ and, in particular,

$$\langle \hat{y}, \xi \rangle \geq 0 > \langle \hat{y}, \bar{\xi} \rangle \qquad \forall \xi \in \partial_C g_i (x_0), \forall i \in M(x_0). \tag{15}$$

By [37, p. 276] and since $g_i^{\downarrow}(x_0; y)$ is finite for all $y \in X$ and for all $i \in M(x_0)$, we have that $\partial_C g_i (x_0) \neq \emptyset$ for all $i \in M(x_0)$. Since $0 \notin \overline{\text{co}}^* \left(\cup_{i\in M(x_0)} \partial_C g_i (x_0)\right) \neq \emptyset$, the Strong Separating Hyperplane Theorem implies the existence of a vector $\bar{y} \in X$ and $\eta \in \mathbb{R}$ such that $\langle \bar{y}, \xi \rangle \geq \eta > 0$ for all $\xi \in \partial_C g_i (x_0)$ and for all $i \in M(x_0)$. Define $\bar{x} = \hat{y} + \varepsilon \bar{y}$ with $\varepsilon > 0$. We have that $\langle \bar{x}, \xi \rangle = \langle \hat{y}, \xi \rangle + \varepsilon \langle \bar{y}, \xi \rangle \geq \varepsilon\eta > 0$ for all $\xi \in \partial_C g_i (x_0)$ and for all $i \in M(x_0)$. Moreover, in view of (15), if $\varepsilon > 0$ is sufficiently small, we have also

$$\langle \bar{x}, \xi \rangle \geq \varepsilon\eta > 0 > \langle \bar{x}, \bar{\xi} \rangle \quad \forall \xi \in \partial_C g_i (x_0), \forall i \in M(x_0). \tag{16}$$

By Theorem 2 and since $0 \notin \partial_C g_i (x_0)$ for all $i \in M(x_0)$, we have that each g_i is directionally Lipschitzian at x_0. By [37, Theorem 3] and since $g_i^{\downarrow}(x_0; y)$ is finite for all $y \in X$ and for all $i \in M(x_0)$, we have that

$$g_i^{\downarrow}(x_0; \bar{x}) = \liminf_{\substack{x' \to x_0 \\ t \downarrow 0}} \frac{g_i (x' + t\bar{x}) - g_i (x')}{t} \quad \forall i \in M(x_0).$$

This implies that for each $i \in M(x_0)$

$$\liminf_{t\downarrow 0} \frac{g_i (x_0 + t\bar{x}) - g_i (x_0)}{t} \geq g_i^{\downarrow}(x_0; \bar{x}) = \inf_{\xi \in \partial_C g_i (x_0)} \langle \bar{x}, \xi \rangle \geq \varepsilon\eta > 0. \tag{17}$$

Since C is open, there exists $\delta > 0$ such that $x_0 + t\bar{x} \in C$ for all $t \in [0, \delta]$. In view of (17), it follows that for each $i \in M(x_0)$ there exists $t_i \in (0, \delta)$ such that

$$g_i (x_0 + t\bar{x}) - g_i (x_0) \geq \frac{\varepsilon\eta}{2} t > 0 \quad \forall t \in (0, t_i]. \tag{18}$$

The next claim proves that such t_i values can be chosen to be uniform. This will be instrumental for what follows.

Claim. There exists $\bar{t} \in (0, \delta)$ *such that* $g_i (x_0 + t\bar{x}) - g_i (x_0) > 0$ *for all* $i \in M(x_0)$ *and for all* $t \in (0, \bar{t}]$.
Proof of the Claim Set $\varphi_i (t) = g_i (x_0 + t\bar{x}) - g_i (x_0)$ for all $t \in [0, \delta]$ and for all $i \in M(x_0)$. Consider the function $T : M(x_0) \to \mathbb{R}$ given by

$$T(i) = \sup \{t \in (0, \delta) : \varphi_i (t) > 0\} \quad \forall i \in M(x_0).$$

By (18) and since each φ_i is continuous and quasiconcave, $\{t \in (0, \delta) : \varphi_i (t) > 0\}$ is an open interval which includes $(0, t_i]$ for all $i \in M(x_0)$. By definition of T and since each φ_i is continuous and each $t_i \in (0, \delta)$, this implies that $T(i) > t_i > 0$ and

$$t \in (0, T(i)) \implies \varphi_i (t) > 0 \quad \forall i \in M(x_0). \tag{19}$$

We are left to show that T is lower semicontinuous. In fact, by Weierstrass' Theorem, this will yield that there exists $i_\star \in M(x_0)$ such that $T(i) \geq T(i_\star) > t_{i_\star} > 0$ for all $i \in M(x_0)$. By (19), this will allow us to conclude that $\varphi_i (t) > 0$ for all $0 < t \leq t_{i_\star}$ and for all $i \in M(x_0)$, as claimed.

To establish the lower semicontinuity of T, we will prove that $\{i \in M(x_0) : T(i) > k\}$, with $k \in \mathbb{R}$, is open. If either $k \geq \delta$ or $k \leq 0$, then the set is either empty or $M(x_0)$ and, in particular, open in the relative topology. Otherwise, consider $k \in (0, \delta)$ and assume that the set is nonempty. Consider also $i_0 \in M(x_0)$ such that $T(i_0) > k$. Set $\varepsilon > 0$ such that $\hat{t} = T(i_0) - \varepsilon > k > 0$. By (19) and since $\hat{t} \in (0, T(i_0))$, it follows that $\varphi_{i_0}(\hat{t}) > 0$. Since the map $(i,t) \mapsto \varphi_i(t)$ is continuous, the condition $\varphi_i(\hat{t}) > 0$ continues to hold for all $i \in U(i_0)$ where $U(i_0)$ is a neighborhood of i_0 in $M(x_0)$. By definition of T, it follows that $T(i) \geq \hat{t} > k$ for all $i \in U(i_0)$, proving that $U(i_0) \subseteq \{i \in M(x_0) : T(i) > k\}$. Since i_0 and k were arbitrarily chosen, the latter set is open and T lower semicontinuous. □

By [1, Theorem 17.16] and since the correspondence $x \mapsto M(x)$ is upper hemicontinuous at x_0 and $M(x_0)$ is compact, if we consider the net $t \mapsto (x_0 + t\bar{x}, i_t)$ in the graph of M, there exists a subnet $\{i_\beta\}$ of $\{i_t\}$ and a point $i_0 \in M(x_0)$ such that $i_\beta \to i_0 \in M(x_0)$ and $t_\beta \to 0$.

By the above Claim and since $i_0 \in M(x_0)$, we have that $g_{i_0}(x_0 + \bar{t}\bar{x}) > g_{i_0}(x_0)$. Since $(i,x) \mapsto g_i(x)$ is continuous and $i_\beta \to i_0$, there exists $\bar{\beta}$ such that

$$g_{i_{\bar{\beta}}}(x_0 + \bar{t}\bar{x}) > g_{i_0}(x_0)$$

and $0 < t_{\bar{\beta}} < \bar{t}$.

By definition of $i_{\bar{\beta}}$ and (16) and since $t_{\bar{\beta}} \in (0, \bar{t})$ and $\bar{\xi} \in \partial_{GP}\bar{g}(x_0)$, we have that

$$\left\langle x_0 + t_{\bar{\beta}}\bar{x}, \bar{\xi} \right\rangle \leq \left\langle x_0, \bar{\xi} \right\rangle \implies g_{i_{\bar{\beta}}}\left(x_0 + t_{\bar{\beta}}\bar{x}\right) = \bar{g}\left(x_0 + t_{\bar{\beta}}\bar{x}\right) \leq \bar{g}(x_0) = g_{i_0}(x_0). \tag{20}$$

We have two cases:

1. $i_{\bar{\beta}} \notin M(x_0)$. We have that $g_{i_{\bar{\beta}}}(x_0) > \bar{g}(x_0) = g_{i_0}(x_0)$. Since $t \mapsto g_{i_{\bar{\beta}}}(x_0 + t\bar{x})$ is quasiconcave, for each $t \in [0, \bar{t}]$ we have that

$$g_{i_{\bar{\beta}}}(x_0 + t\bar{x}) \geq \min\left\{g_{i_{\bar{\beta}}}(x_0), g_{i_{\bar{\beta}}}(x_0 + \bar{t}\bar{x})\right\} > g_{i_0}(x_0),$$

   and, in particular, a contradiction with (20).

2. $i_{\bar{\beta}} \in M(x_0)$. By the above Claim and since $t_{\bar{\beta}} \in (0, \bar{t})$, this implies that $g_{i_{\bar{\beta}}}\left(x_0 + t_{\bar{\beta}}\bar{x}\right) > g_{i_{\bar{\beta}}}(x_0) = \bar{g}(x_0) = g_{i_0}(x)$, a contradiction with (20). ■

We close with a couple of remarks. First, as the proof shows, the inclusion $\cup_{i \in M(x_0)} \partial_{GP} g_i(x_0) \subseteq \partial_{GP}\bar{g}(x_0)$ always holds even without imposing assumptions 1-4. Second, assumption 3 cannot be removed. Let $n \geq 2$ and consider a continuously differentiable quasiaffine function $\rho : \mathbb{R}^n \to \mathbb{R}$ such that $\rho(0) = 0$ and $\nabla\rho(0) \neq 0$. Consider also the collection $\{g_1, g_2\}$ where $g_1 = \rho$ and $g_2 = -\rho$. In this case we have $\bar{g} = -|\rho|$. Since $0 \in \text{co}(\nabla\rho(0), -\nabla\rho(0)) = \overline{\text{co}}^*\left(\cup_{i \in M(0)} \partial_C g_i(0)\right)$, if we consider $x_0 = 0$, then all of the above assumptions are satisfied with the exception of the third one. On the one hand, since $0 \in \operatorname{argmax}_C \bar{g}$, we can conclude that $\partial_{GP}\bar{g}(0) = \mathbb{R}^n$. On the other hand, by Corollary 3 we have $\overline{\text{co}}^*(\partial_{GP} g_1(0) \cup \partial_{GP} g_2(0)) = \text{span}(\nabla\rho(0))$, which is a one-dimensional linear space, yielding that $\partial_{GP}\bar{g}(0) \cup \{0\} \neq \overline{\text{co}}^*(\partial_{GP} g_1(0) \cup \partial_{GP} g_2(0))$.

4.3 More superdifferentials

What happens if in our main result, Theorem 1, the Clarke superdifferential is replaced by other types of superdifferentials? The answer to this question is rather simple (and, more or less, known) for the inclusion part of Theorem 1. Indeed, the superdifferential ∂_C is quite large and so the inclusion will continue to hold for many other superdifferentials.

In the next proposition we consider this inclusion for the lower Dini superdifferential ∂_{D-}. We focus on the two superdifferentials ∂_C and ∂_{D-} because they are the largest ones among the classical superdifferentials. Recall that (see [6] and [4]):

$$\begin{array}{ccl} & & \subseteq \quad \partial_C \\ & \nearrow & \\ \partial_{LS} \subseteq \partial_F \subseteq \partial_H & & \\ & \searrow & \\ & & \subseteq \quad \partial_{D+} \subseteq \partial_{D-} \end{array}$$

where:

- ∂_{LS} is the *Lipschitz smooth* superdifferential;
- ∂_F is the *Frechet superdifferential*:

$$\partial_F g(x) = \left\{ \xi \in X^* : \limsup_{\|y\| \to 0} \frac{g(x+y) - g(x) - \langle y, \xi \rangle}{\|y\|} \le 0 \right\};$$

- ∂_H is the *Hadamard* superdifferential:

$$\partial_H g(x) = \left\{ \xi \in X^* : \langle y, \xi \rangle \ge \limsup_{t \downarrow 0,\ y' \to y} \frac{g(x+ty') - g(x)}{t} \right\};$$

- ∂_{D+} and ∂_{D-} denote the *upper and lower* (respectively) *Dini* superdifferentials, for instance,

$$\partial_{D-} g(x) = \left\{ \xi \in X^* : \langle y, \xi \rangle \ge \liminf_{t \downarrow 0} \frac{g(x+ty) - g(x)}{t} \right\}.$$

In the next proposition ∂_* denotes an abstract superdifferential $\partial_* : X \rightrightarrows X^*$.

Proposition 2 *Let C be open and g continuous and quasiconcave. If either $\partial_* \subseteq \partial_C$ or $\partial_* \subseteq \partial_{D-}$ hold for a superdifferential ∂_*, then*

$$\partial_* g(x) \subseteq \partial_{GP} g(x) \cup \{0\} \qquad \forall x \in C.$$

Proof If $\partial_* \subseteq \partial_C$, then the claim is true by Theorem 1. Suppose now $\partial_* \subseteq \partial_{D-}$. Let $0 \neq \xi \in \partial_* g(x)$ and $y \in C$ such that $\langle y - x, \xi \rangle = \alpha < 0$. Since $\xi \in \partial_{D-} g(x)$, it follows

$$\liminf_{t \downarrow 0} \frac{g(x + t(y-x)) - g(x)}{t} \le \langle y - x, \xi \rangle < 0.$$

Consequently, there is a sequence $t_n \downarrow 0$ for which $g(t_n y + (1-t_n) x) < g(x)$. By the quasiconcavity of g, we have

$$g(x) > g(t_n y + (1-t_n) x) \ge \min\{g(y), g(x)\},$$

yielding that $g(x) > g(y)$. By Lemma 1, we can conclude that $\xi \in \partial_{GP} g(x)$. ■

In contrast, the extension of equality (6) to other superdifferentials is less obvious.

5 Affine Gateaux differentiability and Roy's identity

5.1 Affine Gateaux differentiability

Roy's identity is one of the key formulas of consumer theory. In this section we show that our equivalence results for the Clarke and Greenberg-Pierskalla differentials, in particular Corollary 3, permit to prove a general version of this classical formula. To this end, we first extend Corollary 3 to functions g that are just Gateaux differentiable, rather than strictly differentiable, and whose convex domain C is no longer assumed to be open.

This relaxed condition on domains requires a notion of "affine Gateaux derivative", which builds upon a notion used in economic theory.[12] Throughout this section, we consider two vector spaces (X, X^*) in duality, and a convex set C of X.

Definition 4 *A function $g : C \to \mathbb{R}$ is called* affinely Gateaux differentiable *at a point $x \in C$ if there is some $\xi \in X^*$ such that*

$$\lim_{t \downarrow 0} \frac{g(x + t(y - x)) - g(x)}{t} = \langle y - x, \xi \rangle$$

holds for all $y \in C$. The element ξ is called an affine (Gateaux) gradient *and denoted by $\nabla^a g(x)$.*

The gradient $\nabla^a g(x)$ is, in general, not unique (when it exists). In particular, it is easy to see that if ξ is an affine gradient at a point x, then $\xi' \in X^*$ is also an affine gradient at the same point if and only if

$$\xi - \xi' \in (C - C)^{\perp}.$$

That is, $\nabla^a g(x)$ is actually an equivalence class $[\nabla^a g(x)]$ of the quotient space $X^*/(C - C)^{\perp}$. In this regard, recall that $(C - C)^{\perp}$ is a weak*-closed vector subspace of X^* called the *annihilator* of $C - C$. Observe further that $(C - C)^{\perp} = (C - x)^{\perp}$ holds for any fixed element x of C. Note that the present approach becomes relevant when C is not open, otherwise $(C - C)^{\perp} = \{0\}$ when C is open and we get the familiar Gateaux differential.

The space generated by the set C,

$$V = \operatorname{span}(C - C)$$

plays an important role in our arguments. Of course, $(C - C)^{\perp} = V^{\perp} = \bar{V}^{\perp}$ and so $X^*/(C - C)^{\perp} = X^*/V^{\perp}$.

Thanks to the space V we can give another representation for the equivalence classes $[\xi] \in X^*/V^{\perp}$. Given $\xi \in X^*$, denote by $\xi_V : V \to \mathbb{R}$ its restriction to the space $V \subseteq X$, namely, $\xi_V(x) = \langle x, \xi \rangle$ for all $x \in V$. It is then easy to check that

$$\xi' \in [\xi] \Longleftrightarrow \xi'_V = \xi_V$$

which leads to the identification $[\xi] \leftrightarrow \xi_V$.

Remark More is true when X is a normed vector space and X^* is its norm dual: the Banach space $X^*/V^{\perp}$ is then isometrically isomorphic to the dual space V^* (see, e.g., [30, Theorem 1.10.16]).

A similar argument applies to the Greenberg-Pierskalla superdifferential because of the following implication:

$$\xi \in \partial_{GP} g(x) \Longrightarrow \xi + (C - C)^{\perp} \subseteq \partial_{GP} g(x).$$

The cone $\partial_{GP} g(x)$ can thus be partitioned into equivalence classes. Let us denote by $[\partial_{GP} g(x)]$ the partitioned cone.

[12] See Cerreia-Vioglio et al. [8] and the references therein.

Our last observation pertains the concept of *(relative) internal point*. A vector $x \in X$ is said to be *internal* to a convex set C if, for each $v \in V$, there exists $\varepsilon > 0$ such that $x + tv \in C$ for all $|t| < \varepsilon$. The set of internal points of C forms its *relative algebraic interior*.[13]

We can now state the announced generalization of Corollary 3. Observe that the quasiconcavity of the function is no longer assumed here.

Proposition 3 *Let $g : C \to \mathbb{R}$ be affinely Gateaux differentiable at an internal point x of C with $\partial_{GP} g(x) \neq \emptyset$. If $[\nabla^a g(x)] \neq 0$, then*

$$[\partial_{GP} g(x)] = \{\lambda [\nabla^a g(x)] : \lambda > 0\}. \tag{21}$$

Whenever $\bar{V} = X$, the equivalence classes $[\xi]$ are singletons and so (21) reduces to the familiar

$$\partial_{GP} g(x) = \{\lambda \nabla^a g(x) : \lambda > 0\}. \tag{22}$$

Proof The assumption $[\nabla^a g(x)] \neq 0$ implies that x is not a maximizer. Let $\bar{\xi} \in \partial_{GP} g(x)$ and $\bar{\xi}_V : V \to \mathbb{R}$ denote its restriction to V. Since x is not a maximizer, we have that $\bar{\xi}_V \neq 0$. Consider $z \in \ker \bar{\xi}_V \subseteq V$. Since x is internal, there exists $\delta > 0$ such that $y_+ \stackrel{\text{def}}{=} x + \delta z \in C$ and $y_- \stackrel{\text{def}}{=} x - \delta z \in C$. Since C is convex, we also have that $x + t(y_+ - x), x + t(y_- - x) \in C$ for all $t \in (0,1)$. Moreover, we have that

$$\langle x + t(y_+ - x), \bar{\xi} \rangle = \langle x, \bar{\xi} \rangle + \delta t \bar{\xi}_V(z) = \langle x, \bar{\xi} \rangle$$

and

$$\langle x + t(y_- - x), \bar{\xi} \rangle = \langle x, \bar{\xi} \rangle - \delta t \bar{\xi}_V(z) = \langle x, \bar{\xi} \rangle.$$

Therefore, the inequality $\langle x + t(y_+ - x), \bar{\xi} \rangle \leq \langle x, \bar{\xi} \rangle$ implies that for each $t \in (0,1)$

$$\frac{g(x + t(y_+ - x)) - g(x)}{t} \leq 0.$$

By passing to the limit, this yields that

$$\delta \nabla^a g(x)_V(z) = \langle y_+ - x, \nabla^a g(x) \rangle = \lim_{t \downarrow 0} \frac{g(x + t(y_+ - x)) - g(x)}{t} \leq 0,$$

proving that $\nabla^a g(x)_V(z) \leq 0$. The same argument, applied to y_-, instead leads to $\nabla^a g(x)_V(z) \geq 0$, that is, $z \in \ker \nabla^a g(x)_V$. Since z was arbitrarily chosen, we obtain that $\ker \bar{\xi}_V \subseteq \ker \nabla^a g(x)_V$.

By [1, Theorem 5.91], there exists $\bar{\lambda} \in \mathbb{R}$ such that $\nabla^a g(x)_V = \bar{\lambda} \bar{\xi}_V$, that is, $[\nabla^a g(x)] = \bar{\lambda} [\bar{\xi}]$. Since $[\nabla^a g(x)] \neq 0$, we have that $\bar{\lambda} \neq 0$. We next show that $\bar{\lambda} > 0$. By contradiction, assume that $\bar{\lambda} < 0$. Since $\bar{\xi}_V \neq 0$, it follows that there exists $z \in V$ such that $\bar{\xi}_V(z) < 0$ and $\nabla^a g(x)_V(z) > 0$. By hypothesis, there exists $\delta > 0$ sufficiently small such that $y \stackrel{\text{def}}{=} x + \delta z \in C$. Since C is convex, we have that $x + t(y - x) \in C$ and $\langle x + t(y - x), \bar{\xi} \rangle \leq \langle x, \bar{\xi} \rangle$ for all $t \in (0,1)$. Consequently, we have that $[g(x + t(y - x)) - g(x)]/t \leq 0$ for all $t \in (0,1)$ and

$$0 < \delta \nabla^a g(x)_V(z) = \langle y - x, \nabla^a g(x) \rangle = \lim_{t \downarrow 0} \frac{g(x + t(y - x)) - g(x)}{t} \leq 0,$$

a contradiction. We conclude that $\bar{\lambda} > 0$ and $[\bar{\xi}] = \lambda [\nabla^a g(x)]$, with $\lambda = \bar{\lambda}^{-1} > 0$. Since $\bar{\xi}$ was arbitrarily chosen, this implies that

$$[\partial_{GP} g(x)] \subseteq \{\lambda [\nabla^a g(x)] : \lambda > 0\}.$$

At the same time, since $[\partial_{GP} g(x)]$ is a cone, we have $[\nabla^a g(x)] = \bar{\lambda} [\bar{\xi}] \in [\partial_{GP} g(x)]$ and, in particular, $\{\lambda [\nabla^a g(x)] : \lambda > 0\} \subseteq [\partial_{GP} g(x)]$. This yields our result. ■

[13] It is also known as the intrinsic core of C. See for instance [25].

5.2 The Roy identity

Given a price system $p \in \mathbb{R}^n_{++}$ and a wealth level $w > 0$, a consumer with a continuous utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ confronts the budget set $B(p,w) = \left\{x \in \mathbb{R}^n_+ : \langle x, p\rangle \leq w\right\}$ of affordable consumption bundles.

The consumer optimization problem is

$$\max_x u(x) \quad \operatorname{sub} x \in B(p,w).$$

Since the budget set is compact and the utility function is continuous, this problem admits solution for all price-wealth pairs $(p,w) \in \mathbb{R}^n_{++} \times \mathbb{R}_{++}$.

The *demand correspondence* $d : \mathbb{R}^n_{++} \times \mathbb{R}_{++} \rightrightarrows \mathbb{R}^n_+$, given by $d(p,w) = \operatorname{argmax}_{B(p,w)} u$, associates to each price-wealth pair (p,w) the set of all optimal consumption bundles. The *indirect utility function* $v : \mathbb{R}^n_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ given by $v(p,w) = \max_{B(p,w)} u$ associates to each price-wealth pair (p,w) the utility level attained at the optimal consumption bundle. The next classic result relates these two notions.

Proposition 4 (Antonelli-Roy) *Suppose $u : \mathbb{R}^n_+ \to \mathbb{R}$ is continuous. If the indirect utility function v is Gateaux differentiable at $(p,w) \in \mathbb{R}^n_{++} \times \mathbb{R}_{++}$, with $\partial v(p,w)/\partial w \neq 0$, then $d(p,w)$ is a singleton and*

$$d(p,w) = -\frac{1}{\frac{\partial v(p,w)}{\partial w}} \nabla_p v(p,w). \tag{23}$$

Relation (23), called *Roy's identity*, permits to compute demand functions from indirect utility functions.[14] Our aim is to extend it to an infinite dimensional setting.

To this end, in the duality (X, X^*) we interpret a closed convex cone $K \subseteq X$ as the set of bundles of goods and its conjugate cone $K^* \subseteq X^*$ as the cone of prices. Yet, in this setting K^* may be too large, usually without internal points. Therefore, we consider a suitable convex subcone $P \subseteq K^*$, as it will be discussed momentarily.

Using the wealth cone $\mathbb{R}_{++}$, we can define the budget set $B(p,w) = \{x \in K : \langle x, p\rangle \leq w\}$ for each $(p,w) \in P \times \mathbb{R}_{++}$. Given a utility function $u : K \to \mathbb{R}$, we can also define the indirect utility $v : P \times \mathbb{R}_{++} \to \mathbb{R} \cup \{+\infty\}$ by $v(p,w) = \sup_{B(p,w)} u$.

Proposition 5 *Given a $u : K \to \mathbb{R}$, let $P \subseteq K^*$ be a convex cone such that $(P-P)^{\perp} = \{0\}$ and $(p,w) \in P \times \mathbb{R}_{++}$ a vector with p internal to P. If the indirect utility function v is real-valued and affinely Gateaux differentiable at (p,w), with $\partial v(p,w)/\partial w \neq 0$, then $d(p,w)$ is, when nonempty,*[15] *a singleton and*

$$d(p,w) = -\frac{1}{\frac{\partial v(p,w)}{\partial w}} \nabla^a_p v(p,w). \tag{24}$$

Proof Suppose $d(p,w) \neq \emptyset$. The first step is to show that $\hat{x} \in d(p,w)$ implies

$$(-\hat{x}, 1) \in \partial^{GP} v(p,w). \tag{25}$$

That is, we must verify the implication

$$-\langle \hat{x}, p' - p\rangle + 1 \cdot (w' - w) \geq 0 \Longrightarrow v(p', w') \geq v(p,w) \tag{26}$$

for all $(p', w') \in P \times \mathbb{R}_{++}$. But, the left-hand side of (26) is equivalent to

$$w - \langle \hat{x}, p\rangle \leq w' - \langle \hat{x}, p'\rangle.$$

[14] First stated by Antonelli (1886) as equation 24 in his work (in both the original and translated version) and then rediscovered by Roy (1942), under whose name it came to be known.

[15] In our setting, the nonemptiness of the demand correspondence $d : P \times \mathbb{R}_{++} \rightrightarrows K$ is no longer ensured without adding further conditions, an issue that, however, we do not consider here.

Since $\hat{x} \in d(p,w) \subseteq B(p,w)$, this implies that $\langle \hat{x}, p' \rangle \leq w'$, yielding that $v(p',w') \geq u(\hat{x}) = v(p,w)$ and, in particular, (25).

By hypothesis, v is affinely Gateaux differentiable at (p,w). Hence, there exists the gradient

$$\nabla^a v(p,w) = \nabla_p^a v(p,w) \times \{\bar{\lambda}\} \in X \times \mathbb{R}$$

where $\bar{\lambda} = \partial v(p,w)/\partial w$. Since $(P-P)^{\perp} = \{0\}$, $[\nabla^a v(p,w)]$ is a singleton. Given that p is internal to P, (p,w) is internal to $P \times \mathbb{R}_{++}$. As $\bar{\lambda} \neq 0$, we have $\nabla^a v(p,w) \neq 0$. We can then invoke Proposition 3, namely,

$$\partial^{GP} v(p,w) = \{\lambda \nabla^a v(p,w) : \lambda > 0\}.$$

On the other hand, we proved that $\hat{x} \in d(p,w)$ implies $(-\hat{x}, 1) \in \partial^{GP} v(p,w)$. Hence, $-\hat{x} = \lambda \nabla_p^a v(p,w)$ and $1 = \lambda \partial v(p,w) \backslash \partial w$ for some $\lambda > 0$. This proves (24) as well as the fact that $d(p,w)$ is a singleton. ■

The standard Roy's identity (23) is the special case of (24) when $K = \mathbb{R}_+^n$ and $P = \mathbb{R}_{++}^n$, with the additional condition that the utility function is continuous, so that $d(p,w) \neq \emptyset$ and v is real-valued.

Since $P - P$ is a linear subspace of X^*, to get the condition $(P-P)^{\perp} = \{0\}$, it is sufficient that P spans a space which is weak*-dense in X^*. Moreover, since $p \in P$ is internal, the cone P must have a nonempty relative algebraic interior.

To illustrate, consider the commodity space $L^p \equiv L^p(\Omega, \mu)$, where μ is a finite measure and $1 \leq p \leq \infty$, and the price space L^q.[16] In this case, a good choice are the cones $K = L_+^p$ and $P = L_{++}^{\infty} \subseteq L_+^q = K^*$, where $p \in L_{++}^{\infty}$ if and only if $p \geq \varepsilon 1_{\Omega}$ for some $\varepsilon > 0$. The span of L_{++}^{∞} is weak*-dense in L^q and any vector $p \in L_{++}^{\infty}$ is internal. We thus have the following corollary.

Corollary 7 *Let $u : L_+^p \to \mathbb{R}$ and $(p,w) \in L_{++}^{\infty} \times \mathbb{R}_{++}$. If the indirect utility function v is real-valued and affinely Gateaux differentiable at (p,w), with $\partial v(p,w)/\partial w \neq 0$, then $d(p,w)$ is, when nonempty, a singleton and formula (24) holds.*

In general, Proposition 3 can be applied in optimization problems that feature a differential link, similar to that in the Roy identity, between value functions and solutions. Consider for instance the infimal generators of quasiconcave functions. So, let $G : X \times \mathbb{R} \to \mathbb{R}$ be a functional such that $G(x, \cdot)$ is nondecreasing for each x. If Γ is a subset, not necessarily convex, of X and if C is an open convex set of X^*, we can define a value function v on C by

$$v(\xi) = \inf_{x \in \Gamma} G(x, \langle x, \xi \rangle) \quad \forall \xi \in C.$$

Since $\xi \mapsto G(x, \langle x, \xi \rangle)$ is quasiaffine for every $x \in \Gamma$, the value function v is quasiconcave. Under additional strong assumptions, generators like G appear in quasiconvex duality theory. In particular, the specification $G(x,t) = t - f(x)$ leads to concave Fenchel conjugation (see [9] for a detailed analysis).

The relationship with Roy's identity comes from the following, easily checked, inclusion

$$\operatorname{argmin}_{x \in \Gamma} G(x, \langle x, \xi \rangle) \subseteq \partial_{GP} v(\xi)$$

To exemplify, endow X^* with the dual norm, assume that C is open and let Γ be a subset of the unit sphere of X. By Proposition 3, if the value function v is real-valued and (affinely) Gateaux differentiable at ξ, with $\nabla v(\xi) \neq 0$, then $\operatorname{argmin}_{x \in \Gamma} G(x, \langle x, \xi \rangle)$ is, when nonempty, a singleton and

$$\hat{x}_{\xi} = \|\nabla v(\xi)\|^{-1} \nabla v(\xi).$$

[16] With a small abuse of notation, we denote by p both the exponent of L^p and the price functional. This should not generate any confusion.

References

[1] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, 3rd ed., Springer-Verlag, Berlin, 2006.

[2] G. B. Antonelli, *Sulla teoria matematica dell'economia politica*, Tipografia del Folchetto, Pisa, 1886 (reprinted in *Giornale degli Economisti*, 10, 233-263, 1951; trans. in J. S. Chipman et al., eds., *Preferences, Utility and Demand*, Harcourt Brace Jovanovich, New York, 1971).

[3] K. J. Arrow and A. C. Enthoven, Quasi-concave programming, *Econometrica*, 29, 779-800, 1961.

[4] D. Aussel, Subdifferential property of quasiconvex and pseudoconvex functions: unified approach, *Journal of Optimization Theory and Applications*, 97, 29-45, 1998.

[5] D. Aussel, J.-N. Corvellec, and M. Lassonde, Subdifferential characterization of quasiconvexity and convexity, *Journal of Convex Analysis*, 1, 195-201, 1994.

[6] D. Aussel, J.-N. Corvellec, and M. Lassonde, Mean value property and subdifferential criteria for lower semicontinuous functions, *Transactions of the American Mathematical Society*, 347, 4147-4161, 1995.

[7] J. M. Borwein and D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, *Transactions of the American Mathematical Society*, 303, 517-527, 1987.

[8] S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci, Stochastic dominance analysis without the independence axiom, *Management Science*, 63, 1097-1109, 2017.

[9] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Complete monotone quasiconcave duality, *Mathematics of Operations Research*, 36, 321-339, 2011.

[10] F. H. Clarke, Generalized gradients and applications, *Transactions of the American Mathematical Society*, 205, 247-262, 1975.

[11] F. H. Clarke, A new approach to Lagrange multipliers, *Mathematics of Operations Research*, 1, 165-174, 1976.

[12] F. H. Clarke, *Optimization and nonsmooth analysis*, SIAM, Philadelphia, 1990.

[13] J.-P. Crouzeix, Polaires quasi-convexes et dualité, *Compte Rendue Acad. Sci. Paris Series A*, 279, 955-958, 1974.

[14] J.-P. Crouzeix, *Contributions a l'étude des functions quasiconvex*, Ph.D. thesis, Université de Clermont-Ferrand II, Clermont-Ferrand, France, 1977.

[15] J.-P. Crouzeix, Conditions for convexity of quasiconvex functions, *Mathematics of Operations Research*, 5, 120-125, 1980.

[16] J.-P. Crouzeix, Duality between direct and indirect utility functions: differentiability properties, *Journal of Mathematical Economics*, 12, 149-165, 1983.

[17] J.-P. Crouzeix, Continuity and differentiability of quasiconvex functions, in *Handbook of generalized convexity and generalized monotonicity*, (N. Hadjisavvas, S. Komlosi, and S. Schaible, eds.), Springer, New York, 2005.

[18] J.-P. Crouzeix, A. Eberhard, and D. Ralph, A geometrical insight on pseudoconvexity and pseudomonotonicity, *Mathematical Programming*, 123, 61-83, 2010.

[19] A. Daniilidis, N. Hadjisavvas, and J-E. Martinez-Legaz, An appropriate subdifferential for quasiconvex functions, *SIAM Journal on Optimization*, 12, 407-420, 2001.

[20] J. M. Danskin, The theory of max-min, with applications, *SIAM Journal on Applied Mathematics,* 14, 641-664, 1966.

[21] B. de Finetti, Sulle stratificazioni convesse, *Annali di Matematica Pura e Applicata*, 30, 173-183, 1949.

[22] W. Fenchel, *Convex cones, sets, and functions*, mimeo, Princeton University, Princeton, 1953.

[23] H. P. Greenberg and W. P. Pierskalla, Quasi-conjugate functions and surrogate duality*, Cahiers du Centre d'Études de Recherche Opérationelle*, 15, 437-448, 1973.

[24] N. Hadjisavvas, Generalized convexity, generalized monotonicity and nonsmooth analysis, in *Handbook of generalized convexity and generalized monotonicity*, (N. Hadjisavvas, S. Komlosi, and S. Schaible, eds.), Springer, New York, 2005.

[25] R. B. Holmes, *Geometric functional analysis and its applications*, Springer-Verlag, New York, 1975.

[26] A. Kabgani, M. Soleimani-damaneh, The relationships between convexificators and Greenberg-Pierskalla subdifferentials for quasiconvex functions, *Numerical Functional Analysis and Optimization*, 38, 1548-1563, 2017.

[27] S. Komlosi, Generalized convexity and generalized derivatives, in *Handbook of generalized convexity and generalized monotonicity*, (N. Hadjisavvas, S. Komlosi, and S. Schaible, eds.), Springer, New York, 2005.

[28] J.-E. Martinez-Legaz, A generalized concept of conjugation*,* in *Optimization theory and algorithms*, (A. Auslander, J.-B. Hiriart-Urruty, and W. Oetti, eds.), Dekker, New York, 1983.

[29] J.-E. Martinez-Legaz, Duality between direct and indirect utility functions under minimal hypotheses, *Journal of Mathematical Economics*, 20, 199-209, 1991.

[30] R. E. Meggison, *An introduction to Banach space theory*, Springer-Verlag, New York, 1998.

[31] J.-P. Penot, Are generalized derivatives useful for generalized convex functions?, in *Generalized convexity, generalized monotonicity: recent results*, (J.-P. Crouzeix, J.-E. Martinez-Legaz, and M. Volle, eds.), Kluwer, Dordrecht, 1998.

[32] J.-P. Penot, What is quasiconvex analysis?, *Optimization*, 47, 35-110, 2000.

[33] J.-P. Penot, *Calculus without derivatives*, Springer, New York, 2013.

[34] J.-P. Penot and M. Volle, On quasi-convex duality, *Mathematics of Operations Research,* 15, 597-625, 1990.

[35] J.-P. Penot and C. Zalinescu, Elements of quasiconvex subdifferential calculus, *Journal of Convex Analysis*, 7, 243-269, 2000.

[36] R. T. Rockafellar, Directionally Lipschitzian functions and subdifferential calculus, *Proceedings of the London Mathematical Society*, 39, 331-355, 1979.

[37] R. T. Rockafellar, Generalized directional derivatives and subgradients of nonconvex functions, *Canadian Journal of Mathematics*, 32, 257-280, 1980.

[38] R. Roy, *De l'utilité: contribution à la théorie des choix*, Hermann, Paris, 1942.

[39] W. Rudin, *Functional analysis*, 2^{nd} ed., McGraw-Hill, New York, 1991.

[40] C. Zalinescu, *Convex analysis in general vector spaces*, World Scientific Publishing, New Jersey, 2002.