Journal of Convex Analysis
Volume 15 (2008), No. 3, 535–546

Representation of the Polar Cone of Convex Functions and Applications

G. Carlier
CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris, France
carlier@ceremade.dauphine.fr

T. Lachand-Robert
Laboratoire de Mathématiques, UMR CNRS 5127, Université de Savoie, 73376 Le Bourget-du-Lac, France

Dedicated to the memory of Thomas Lachand-Robert.

Received: May 25, 2006
Revised manuscript received: October 23, 2006

Using a result of Y. Brenier [1], we give a representation of the polar cone of monotone gradient fields in terms of measure-preserving maps, or bistochastic measures. Some applications to variational problems subject to a convexity constraint are given.

Keywords: Convexity constraint, Euler-Lagrange equation, measure-preserving maps, bistochastic measures

1. Introduction

Minimization problems in the set of convex functions, of the following form:

$$\inf_{\substack{u\in H^1(\Omega)\\ u \text{ convex}}} \int_\Omega g(x,u(x),\nabla u(x))\,\mathrm{d}x \tag{1}$$

arise in several fields of applied mathematics (see [2], [5]). Even for a very simple and well-behaved g, the convexity constraint raises a number of mathematical difficulties. Even the simplest case, that is the projection problem:

$$\inf_{\substack{u\in H^1(\Omega)\\ u \text{ convex}}} \int_\Omega |\nabla u - q_0|^2 \tag{2}$$

(where $q_0 \in L^2(\Omega)^N$ is given) has yet to be investigated in detail, in particular regarding regularity issues. When dealing with this type of minimization problems on a closed cone (here the cone of convex functions), it is natural to try to identify the polar cone in order to write an Euler-Lagrange equation. A first proposal in this direction has been given in [7]; unfortunately the corresponding equation is difficult to handle in general.

In the present paper we propose a different representation of the polar cone of the set of convex functions (more precisely: of the polar cone of the set of gradients of convex

ISSN 0944-6532 / \$ 2.50 © Heldermann Verlag

functions). It makes use of the set S of measure-preserving maps (see the precise definition and statements in the next section). We prove that every vector field in $L^2(\Omega)^N$ can be decomposed as the sum of the gradient of a convex function and an element of the closed cone generated by $(S - \mathrm{id})$ (Section 3). This is to be compared to the well-kown polar decomposition of Y. Brenier [1], which is a composition instead of a sum. This result is actually a direct consequence of a result of Y. Brenier (see [1] and Proposition 2.1). Another way to express the same decomposition is to make use of the so-called *bistochastic measures*, a special case of Young measures (Section 4). Some applications to variational problems of the form (1) are given in Sections 5 (Euler-Lagrange equations) and 6 (regularity).

2. Representation of the polar cone of gradients of convex functions

In the whole paper, Ω is some bounded open convex subset of $\mathbb{R}^N$, $N \geq 1$. We denote by $\langle \cdot, \cdot \rangle$ the usual scalar product in the Hilbert space $L^2(\Omega)^N$. We consider

$$K := \{q \in L^2(\Omega)^N : \ q = \nabla u, \text{ for some convex function } u\}.$$

It is a closed convex cone in $L^2(\Omega)^N$. Its polar cone is defined as usual:

$$K^- := \{p \in L^2(\Omega)^N : \langle p, q \rangle \leq 0, \ \forall q \in K\}.$$

We make use of the set of measure-preserving maps:

$$S := \{s \text{ measurable} : \Omega \to \Omega : \ s\sharp \mathrm{d}x = \mathrm{d}x\}.$$

Here $\mathrm{d}x$ denotes Lebesgue's measure and $s\sharp\mathrm{d}x$ is the measure defined by

$$\int_\Omega \varphi(x) s\sharp \mathrm{d}x = \int_\Omega \varphi(s(x)) \mathrm{d}x$$

for every bounded continuous function φ.

Let us start with a characterization of K:

Proposition 2.1 (Brenier). *Let $q \in L^2(\Omega)^N$; then $q \in K$ if and only if*

$$\langle q, s - \mathrm{id} \rangle \leq 0, \quad \textit{for all } s \in S.$$

The proof is already given in [1], we recall it here briefly for the sake of completeness.

Proof. Assume first that $q \in K$, $q = \nabla u$ with u convex. Let $s \in S$; by convexity, one has $\nabla u(x) \cdot (s(x) - x) \leq u(s(x)) - u(x)$, hence:

$$\langle \nabla u, s - \mathrm{id} \rangle \leq \int_\Omega u(s(x)) \mathrm{d}x - \int_\Omega u(x) \mathrm{d}x = 0.$$

Conversely, assume that q satisfies the inequalities of the Proposition and let us prove that $q \in K$. Let $(x_0, x_1, \ldots, x_n = x_0) \in \Omega^{n+1}$ be a cycle of distinct Lebesgue points of q, let $\varepsilon > 0$ be such that the balls $B(x_i, \varepsilon)$ are disjoint and included in Ω. Define then the measure preserving map s_ε by

$$s_\varepsilon(x) = \begin{cases} x & \text{if } x \in \Omega \setminus \bigcup_{i=0}^{n-1} B(x_i, \varepsilon) \\ x + (x_{i+1} - x_i) & \text{if } x \in B(x_i, \varepsilon),\ i = 0, \dots, n-1. \end{cases}$$

We have

$$\langle q, s_\varepsilon - \mathrm{id} \rangle \leq 0$$

by assumption. This yields

$$\sum_{i=0}^{n-1} \frac{1}{|B(x_i, \varepsilon)|} \int_{B(x_i,\varepsilon)} q(x) \cdot (x_{i+1} - x_i) \, \mathrm{d}x \leq 0. \tag{3}$$

Letting ε go to zero, we get then

$$\sum_{i=0}^{n-1} q(x_i) \cdot (x_{i+1} - x_i) \leq 0.$$

In other words the set $A := \{(x, q(x)),\ x \text{ Lebesgue point of } q\}$ is cyclically monotone, so that from Rockafellar [9] there exists a convex function u such that $A \subset \partial u$; in particular $q = \nabla u$ almost everywhere. □

As a consequence of Proposition 2.1 one may characterize the polar cone of K:

Corollary 2.2. *The polar cone K^- is the closure (in L^2) of the convex cone generated by $(S - \mathrm{id})$*

$$K^- = \overline{\mathrm{cone}}(S - \mathrm{id}).$$

Proof. Proposition 2.1 can be stated as:

$$K = (S - \mathrm{id})^- = (\overline{\mathrm{cone}}(S - \mathrm{id}))^-.$$

Taking the polar cones of those two closed convex cones gives the desired result. □

One may ask what is the difference between the two cones $\overline{\mathrm{cone}}(S - \mathrm{id})$ and $\mathbb{R}_+ \overline{\mathrm{co}}(S - \mathrm{id})$. Indeed, the latter is much easier to charaterize, as we shall see in the next sections. It has also some interesting properties:

Proposition 2.3. *Let $p \in \mathbb{R}_+ \overline{\mathrm{co}}(S - \mathrm{id})$, $p \neq 0$, so that $p = \lambda(\tau - \mathrm{id})$ with $\lambda > 0$ and $\tau \in \overline{\mathrm{co}}\, S \setminus \{\mathrm{id}\}$. Then, for any convex function $u \in H^1(\Omega)$, one has*

$$\int_\Omega u(\tau(x)) \, \mathrm{d}x \leq \int_\Omega u(x) \, \mathrm{d}x. \tag{4}$$

Additionnally, for any convex function $u \in H^1(\Omega)$ such that $\langle p, \nabla u \rangle = 0$, one has:

$$\int_\Omega u(\tau(x)) \, \mathrm{d}x = \int_\Omega u(x) \, \mathrm{d}x \tag{5}$$

$$u(\tau(x)) - u(x) = \nabla u(x) \cdot (\tau(x) - x) \quad \text{a.e. } x \in \Omega. \tag{6}$$

Notice that (6) means that the convex function u is actually affine on the segment $[x, \tau(x)]$. In particular, $p \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ and $\langle p, \nabla u \rangle = 0$ for some strictly convex u implies $p = 0$.

Let us postpone the proof for a while and continue the discussion on the links between $\overline{\text{cone}}(S - \text{id})$ and $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$. One has obviously

$$\mathbb{R}_+ \overline{\text{co}}(S - \text{id}) \subset \overline{\text{cone}}(S - \text{id}) = \overline{\mathbb{R}_+ \overline{\text{co}}(S - \text{id})}, \tag{7}$$

but it turns out that $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ is not closed and therefore there is no equality in the inclusion above. Indeed, consider the set

$$\ker \text{div} := \{p \in (L^2(\Omega))^N : \langle p, \nabla v \rangle = 0, \forall v \in H^1(\Omega)\}.$$

(That is, divergence-free vector fields and whose normal component at the boundary vanishes, whenever regular.)

Corollary 2.4. *One has*

$$\ker \text{div} \subset \overline{\text{cone}}(S - \text{id})$$
$$\ker \text{div} \cap \mathbb{R}_+ \overline{\text{co}}(S - \text{id}) = \{0\}.$$

In particular, $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ *is not closed.*

Proof of Proposition 2.3 and Corollary 2.4. Since $\tau \in \overline{\text{co}}\, S$, there exists some sequence of the form $\sum_{i=1}^n \alpha_i^n s_i^n$ with $s_i^n \in S$, $\alpha_i^n \geq 0$, $\sum_i \alpha_i^n = 1$, converging in L^2 and almost everywhere to τ.

For any $u \in H^1(\Omega)$ convex, we have $u\big(\sum_i \alpha_i^n s_i^n(x)\big) \leq \sum_i \alpha_i^n u(s_i^n(x))$, for almost every $x \in \Omega$. Hence

$$\int_\Omega \left[u\Big(\sum_{i=1}^n \alpha_i^n s_i^n(x)\Big) - u(x) \right] \mathrm{d}x \leq \sum_{i=1}^n \alpha_i^n \int_\Omega (u(s_i^n(x)) - u(x))\, \mathrm{d}x = 0,$$

taking into account that each s_i^n is measure preserving. In the limit we get (4) by Fatou's Lemma.

If we assume $\langle p, \nabla u \rangle = 0$, then we get:

$$\int_\Omega \big[u(\tau(x)) - u(x) - \nabla u(x) \cdot (\tau(x) - x) \big] \mathrm{d}x \leq 0.$$

Since u is convex, the integrand is nonnegative. Hence we obtain (5–6).

It is clear that $\ker \text{div} \subset \overline{\text{cone}}(S - \text{id})$, since from the very definitions $p \in \ker \text{div}$ and $q \in K$ implies $\langle p, q \rangle = 0$.

Now assume $p \in \ker \text{div} \cap \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$, $p \neq 0$. Then $\langle p, \nabla u \rangle = 0$ for all u, so (6) holds also for any u convex. Since this equation means that u is affine on the segment $[x, \tau(x)]$, there is clearly a contradiction whenever u is strictly convex. □

The non-closedeness of $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ makes difficult the manipulation of $\overline{\text{cone}}(S - \text{id})$ in general. One has to notice the following, however:

Proposition 2.5. *Let $p \in \overline{\text{cone}}(S - \text{id})$. Then either $p \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$, or, for any sequences $(\lambda_n) \subset \mathbb{R}_+$, $(\tau_n) \subset \overline{\text{co}}\, S$ such that $\lambda_n(\tau_n - \text{id}) \to p$ strongly in $L^2(\Omega)^N$, one has $\lambda_n \to +\infty$ and $\tau_n \to \text{id}$.*

Proof. Since $p \in \overline{\text{cone}}(S - \text{id})$, the closure of $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$, it can be written as the limit of a sequence (p_n) such that $p_n = \lambda_n(\tau_n - \text{id})$, with $\lambda_n \in \mathbb{R}_+$, $\tau_n \in \overline{\text{co}}\, S$.

If λ_n does not converge to $+\infty$, then we may extract a subsequence and assume that it converges to some limit $\lambda \in \mathbb{R}_+$. If $\lambda = 0$, then $p_n \to 0$ since $\overline{\text{co}}\, S$ is bounded: so $p = 0 \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$. If $\lambda > 0$, then $\tau_n = \text{id} + \lambda_n^{-1} p_n$ converges to $\tau = \text{id} + \lambda^{-1} p$, and $\tau \in \overline{\text{co}}(S)$ since this is a closed set. Now $p = \lambda(\tau - \text{id}) \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$.

If $\lambda_n \to +\infty$ on the other hand, one get $\tau_n - \text{id} = \lambda_n^{-1} p + o(1) \to 0$ as $n \to \infty$. □

Remark 2.6. We have seen that $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ is dense in K^- and that $\mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ is not closed in L^2, a natural question at this point is then : how big is $K^- \setminus \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$? To partially answer this question, let us remark that $\mathbb{R}_+ \overline{\text{co}}(S - \text{id}) \subset L^\infty$ and that K^- contains many non bounded elements: for instance all the (not necessarily bounded) vector fields in ker div , but also ∇v for every $v \in H_0^1(\Omega)$ with $v \geq 0$, since $\int \nabla v \cdot \nabla u = -\langle \Delta u, v \rangle$.

One can also construct non bounded elements of K^- as follows: let $f \geq 0$ be a scalar function in $L^2 \setminus L^\infty$. Denoting by S_f the set of maps $\sigma : \Omega \to \Omega$ preserving the measure $f(x)\,\mathrm{d}x$, one actually has $(\sigma - \text{id})f \in L^2$ and $\in K^-$ for every $\sigma \in S_f$. For instance if Ω is the ball and f is further assumed to be radial then $(R - \text{id})f \in K^-$ for every rotation R.

Remark 2.7. In view of Corollary 2.4, it is also natural to wonder whether one has $K^- =$ ker div $+ \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$? The previous identity is equivalent to ker div $+ \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ being closed in L^2. Again, the answer is unfortunately negative as the following counter-example shows. Assume that Ω is the unit ball of $\mathbb{R}^2$ and define $v(x) := (1 - |x|)^{2/3}$, we already know that $\nabla v \in K^-$. Now, if ∇v admitted a decomposition of the form $\nabla v = p + \lambda(\tau - \text{id})$ with $(p, \lambda, \tau) \in$ ker div $\times \mathbb{R}_+ \times \overline{\text{co}}(S)$ then for every $\phi \in W^{1,1}(\Omega)$ the quantity $\int_\Omega \nabla v . \nabla \phi$ would be finite. Taking for instance $\phi(x) := (1 - |x|)^{1/4}$ yields the desired contradiction.

Remark 2.8. One should finally notice that, even for a $p \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$, the writing $p = \lambda(\tau - \text{id})$ is not unique. This is obvious if $p = 0$, where one can choose $\lambda = 0$ and any τ, or $\tau = \text{id}$ and any λ. But it is also true if $p \neq 0$. Indeed, for any given $\tau \in \overline{\text{co}}\, S \setminus \{\text{id}\}$, let us consider $I_\tau := \{r \geq 0; \tau_r := r\tau + (1 - r)\text{id} \in \overline{\text{co}}\, S\}$. Since $\overline{\text{co}}\, S$ is closed, convex, and bounded, I_τ is an interval of the form $[0, R]$, with $R \geq 1$. Now we have $p = \lambda(\tau - \text{id}) = \frac{\lambda}{r}(\tau_r - \text{id})$ for any r in this interval.

3. Decomposition of vector fields

Given $q_0 \in L^2(\Omega)^N$ we consider the problem (2), which may also be stated as follows: given $q_0 \in L^2(\Omega)^N$, solve

$$\inf_{q \in K} \|q - q_0\|_{L^2}^2 . \tag{8}$$

There exists a unique solution $\underline{q} = \nabla \underline{u}$ of this projection problem. Moreover the characterization of this solution is given by:

Theorem 3.1. *Let $q_0 \in L^2(\Omega)^N$. Then there exist a unique (up to a constant) $\underline{u} \in H^1$, convex, a unique $\underline{p} \in K^- = \overline{\mathrm{cone}}(S - \mathrm{id})$, such that:*

$$q_0 = \nabla \underline{u} + \underline{p} \tag{9}$$

$$\langle \nabla \underline{u}, \underline{p} \rangle = 0. \tag{10}$$

Here $\underline{q} = \nabla \underline{u}$ solves (8), while $\underline{p}$ solves the dual problem (projection on K^-).

Moreover, if $\underline{p} \neq 0$ (i.e. $q_0 \notin K$), then either

$$\underline{p} = \lim_n \lambda_n(\tau_n - \mathrm{id}) \quad \text{where } \lambda_n \in \mathbb{R}_+ \to +\infty, \tau_n \in \overline{\mathrm{co}}(S) \to \mathrm{id}, \tag{11}$$

or $\underline{p} = \lambda(\tau - \mathrm{id})$ with $\lambda > 0$, $\tau \in \overline{\mathrm{co}}\, S$ and

$$\int_\Omega \underline{u}(\tau(x))\, \mathrm{d}x = \int_\Omega \underline{u}(x)\, \mathrm{d}x \tag{12}$$

$$\underline{u}(\tau(x)) - \underline{u}(x) = \nabla \underline{u}(x) \cdot (\tau(x) - x) \quad \text{a.e. } x \in \Omega. \tag{13}$$

Proof. The minimization problem (8) can also be written as:

$$\inf_{q \in L^2(\Omega)^N} \sup_{p \in K^-} L(p, q) \tag{14}$$

where $L(p,q) := \frac{1}{2}\|q - q_0\|^2 + \langle p, q \rangle$.

It is well-known that L admits a unique saddle point $(\underline{q}, \underline{p}) = (\nabla \underline{u}, \underline{p})$ in $K \times K^-$ (see for instance [6]) where $\underline{q} = \nabla \underline{u}$ is the solution of (14) and $\underline{p}$ is the solution of the dual problem:

$$\sup_{p \in K^-} \inf_{q \in L^2(\Omega)^N} L(p, q). \tag{15}$$

Since

$$\inf_{q \in L^2(\Omega)^N} L(p,q) = -\frac{1}{2}\|p - q_0\|^2 + \frac{1}{2}\|q_0\|^2,$$

the solution of (15), that is $\underline{p}$, is the projection of q_0 on $K^- = \overline{\mathrm{cone}}(S - \mathrm{id})$, *i.e.* the solution to

$$\inf_{p \in \overline{\mathrm{cone}}(S - \mathrm{id})} \|p - q_0\|^2. \tag{16}$$

The saddle-point $(\underline{q}, \underline{p})$ is characterized by (9–10) since K, K^- are polar cones.

Conversely if (9-10) hold with $\underline{u}$ convex and $\underline{p} \in K^-$ then $\nabla \underline{u}$ (respectively $\underline{p}$) is the projection of q_0 onto K (respectively onto K^-).

If $\underline{p} \neq 0$, we deduce the remaining properties from Proposition 2.3 and Proposition 2.5. □

If, in the previous satetement, $\underline{p} = \lambda(\tau - \mathrm{id})$ with $\lambda > 0$ and $\tau \in \overline{\mathrm{co}}\, S \setminus \{\mathrm{id}\}$, then $\underline{u}$ is affine between x and $\tau(x)$ for almost every x; in particular if x is such that $(x, \underline{u}(x))$ does not belong to a face of dimension ≥ 1 of the epigraph of $\underline{u}$, then $\underline{p}(x) = 0$. When $\underline{p} \in \overline{\mathrm{cone}}(S - \mathrm{id}) \setminus \mathbb{R}_+ \overline{\mathrm{co}}(S - \mathrm{id})$ (a situation that, as already noted, may unfortunately occur), we don't have a tractable representation of $\underline{p}$ so that we don't have such a nice property as (13).

4. Bistochastic measures

In this section, we investigate a little bit further what elements of $\overline{\mathrm{co}}\, S$ look like (the closure is still taken in L^2). This can conveniently be done in terms of *bistochastic measures*.

4.1. Bistochastic measures and $\overline{\mathrm{co}}\, S$

Let a sequence of elements of $\mathrm{co}\, S$ of the form $\tau_n := \sum_{i=1}^{k_n} \alpha_i^n s_i^n$ ($\alpha_i^n \geq 0$, $\sum_i \alpha_i^n = 1$) converge in L^2 to some limit τ. For each n and each $x \in \Omega$ one may define the probability measure

$$\mu_x^n := \sum_{i=1}^{k_n} \alpha_i^n \delta_{s_i^n(x)}$$

so that for every x one has

$$\tau_n(x) = \int_\Omega y \, \mathrm{d}\mu_x^n(y). \tag{17}$$

One may also define the nonnegative Radon measure γ_n on $\Omega \times \Omega$ by

$$\int_\Omega \int_\Omega f(x,y)\, d\gamma_n(x,y) := \int_\Omega \left(\int_\Omega f(x,y) \, \mathrm{d}\mu_x^n(y) \right) \mathrm{d}x, \quad \forall f \in C_0^0(\Omega \times \Omega).$$

Since each s_i^n is volume-preserving, one has for every continuous bounded function f on Ω :

$$\int_\Omega \int_\Omega f(x)\, d\gamma_n(x,y) = \int_\Omega \int_\Omega f(y)\, d\gamma_n(x,y) = \int_\Omega f.$$

This expresses that the two projections (marginals) of γ_n equal Lebesgue measure on Ω; hence γ_n is a so-called *bistochastic measure*. Note that bistochastic measures are special cases of Young measures [11]. It is standard then to prove that the limit τ can also be represented by the measure given by the weak$*$ limit γ of (γ_n) (or some subsequence if necessary). More precisely, γ is also a bistochastic measure which therefore admits a disintegration given by a measurable family of probability measures $\{\mu_x\}_{x\in\Omega}$:

$$\int_\Omega \int_\Omega f(x,y) d\gamma(x,y) := \int_\Omega \left(\int_\Omega f(x,y) \, \mathrm{d}\mu_x(y) \right) \mathrm{d}x, \quad \forall f \in C_0^0(\Omega \times \Omega),$$

and

$$\tau(x) = \int_\Omega y \, \mathrm{d}\mu_x(y), \quad \text{a.e. } x \in \Omega. \tag{18}$$

Note finally for further use that since γ is bistochastic, one has for every continuous bounded function f on Ω:

$$\int_\Omega \left(\int_\Omega f(y) \, \mathrm{d}\mu_x(y) \right) \mathrm{d}x = \int_\Omega f(x) \mathrm{d}x. \tag{19}$$

It is also interesting to notice that

$$\forall v \text{ convex}, \quad \int_\Omega v(\tau(x)) \, \mathrm{d}x \leq \int_\Omega v(x) \, \mathrm{d}x. \tag{20}$$

This comes from (18) and Jensen's inequality. We see in (5) that $\underline{u}$ saturates the inequality (20) with respect to τ.

4.2. The decomposition with bistochastic measures

Let us see now how to use bistochastic measures for the projection problem (8). As already noticed $\nabla\underline{u}$ is the projection of q_0 onto K if and only if q_0 can be decomposed as in (9–10).

We are interested in the nontrivial case $q_0 \notin K$, denoting by $\underline{p}$ the projection of q_0 on K^-, we then have $\underline{p} \neq 0$. Let us assume that $\underline{p} \in \mathbb{R}_+ \overline{\mathrm{co}}(S - \mathrm{id})$ (which, unfortunately, is not the case for any $q_0 \in L^2(\Omega)^N$) then $\underline{p} = \lambda(\tau - \mathrm{id})$ with $\lambda > 0$ and $\tau \in \overline{\mathrm{co}}\, S$. In this case, (12–13) hold.

Using bistochastic measures we see that there exists a measurable family $\{\mu_x\}_{x\in\Omega}$ of probability measures on Ω, satisfying (19) and such that (18) hold (where τ now is given by Theorem 3.1). Using (19), (5) and (18), we get

$$\int_\Omega \left(\int_\Omega \underline{u}(y)\ \mathrm{d}\mu_x(y) \right) \mathrm{d}x = \int_\Omega \underline{u} = \int_\Omega \underline{u}(\tau(x))\mathrm{d}x = \int_\Omega \underline{u} \left(\int_\Omega y\ \mathrm{d}\mu_x(y) \right) \mathrm{d}x.$$

Since $\underline{u}$ is convex this implies that for a.e. $x \in \Omega$:

$$\underline{u} \left(\int_\Omega y\ \mathrm{d}\mu_x(y) \right) = \underline{u}(\tau(x)) = \int_\Omega \underline{u}(y)\ \mathrm{d}\mu_x(y).$$

This means that $\underline{u}$ is essentially affine on $\mathrm{Supp}(\mu_x)$. In particular if $\underline{u}$ is differentiable at x then $\mathrm{Supp}(\mu_x) \subset \partial\underline{u}^*(\nabla\underline{u}(x))$, or equivalently:

$$\underline{u}(y) = \underline{u}(x) + \nabla\underline{u}(x) \cdot (y - x), \quad \mu_x\text{-a.e. } y \in \Omega. \tag{21}$$

Remark 4.1. At least formally, one may characterize $\underline{u}$, the solution of (2), in terms of *balayages*, a well-known notion in potential theory and probabilities (see Meyer [8]). Indeed if K is a convex *compact* subset of $\mathbb{R}^n$, the polar cone of convex functions for the duality between $C^0(K, \mathbb{R})$ functions and Radon measures in K (a duality that is not adapted to our variational problem) can be represented using Cartier-Fell-Meyer Theorem [4]. This theorem states that if a signed measure with Jordan decomposition $\lambda - \mu$ is in the polar cone of $C^0(K, \mathbb{R})$ convex functions (so that λ and μ have the same mass and the same moment), then there exists a measurable family $\{\nu_x\}_{x\in K}$ of probability measures on K such that the barycenter of ν_x is x for all x and for every $f \in C^0(K, \mathbb{R})$:

$$\int_K f(x)\ \mathrm{d}\mu(x) = \int_K \left(\int_K f(y)\,\mathrm{d}\nu_x(y) \right) \mathrm{d}\lambda(x). \tag{22}$$

If $\underline{u} \in C^0(K, \mathbb{R})$ is convex and $\int_K \underline{u}\ \mathrm{d}\mu = \int_K \underline{u}\, d\lambda$ then it is easy to see that $\underline{u}$ is affine on $\mathrm{Supp}(\nu_x)$. Let us also note that if we specify $K := \overline{\Omega}$, then there is no reason for the measures λ and μ above to vanish on $\partial\Omega$. Formally one would like to use this representation to characterize $\underline{u}$, solution of (2), by:

$$\int_\Omega \langle \nabla\underline{u} - q_0, \nabla v \rangle = \int_{\overline{\Omega}} v\, \mathrm{d}(\mu - \lambda),\ \forall v \in H^1,\ \int_{\overline{\Omega}} \underline{u}\, \mathrm{d}(\mu - \lambda) = 0 \tag{23}$$

where λ and μ satisfy (22). We refer to Rochet and Choné [5] for the use of balayages in variational problems subject to a convexity constraint. Even though what one formally obtains with balayages and our representation using bistochastic measures look similar, one has to take the divergence of our representation to obtain Rochet and Choné's one. In other words, there is no obvious link with the kernel $\{\mu_x\}$ we obtain and the kernel $\{\nu_x\}$ formally obtained from Cartier-Fell-Meyer Theorem.

5. Euler-Lagrange Equations

We may use now Proposition 2.1 to derive the Euler-Lagrange equation of problems of the form (1). In the following, we assume that g is differentiable with respect to u and its third argument q and satisfies some suitable growth conditions. Since we consider a minimization problem in H^1, we can add constant functions to the solutions; we deduce that any solution $\underline{u}$ of (1) satisfies:

$$\int_\Omega \frac{\partial g}{\partial u}(x, \underline{u}, \nabla \underline{u})\, \mathrm{d}x = 0. \tag{24}$$

This implies in particular that there exists a solution ψ to the Laplace equation with homogeneous Neumann boundary condition:

$$\begin{cases} \Delta\psi = \dfrac{\partial g}{\partial u}(x, \underline{u}, \nabla \underline{u}) & \text{in } \Omega, \\ \dfrac{\partial \psi}{\partial n} = 0 & \text{on } \partial\Omega. \end{cases} \tag{25}$$

The variational inequalities associated to (1) can be written as:

$$\int_\Omega \left(\frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) - \nabla\psi \right) \cdot \nabla(v - \underline{u})\, \mathrm{d}x \geq 0, \quad \text{for all convex } v \in H^1(\Omega),$$

or equivalently

$$\nabla\psi - \frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) \in \overline{\mathrm{cone}}(S - \mathrm{id}),$$

$$\int_\Omega \left(\nabla\psi - \frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) \right) \cdot \nabla \underline{u}\, \mathrm{d}x = 0.$$

This yields:

$$\frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) = \nabla\psi - \underline{p} \tag{26}$$

for some $\underline{p} \in \overline{\mathrm{cone}}(S - \mathrm{id})$ with $\langle \underline{p}, \nabla \underline{u} \rangle = 0$. This implies in particular:

$$\frac{\partial g}{\partial u}(x, \underline{u}, \nabla \underline{u}) - \mathrm{div}\Big(\frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) + \underline{p} \Big) = 0. \tag{27}$$

In the special case where $\underline{p}$ is of the form $\underline{p} = \lambda(\tau - \mathrm{id})$ with $\lambda > 0$ and $\tau \in \overline{\mathrm{co}}\, S \setminus \{\mathrm{id}\}$ (again, this is not always true and almost impossible to check a priori), then for almost every x such that $(x, \underline{u}(x))$ does not belong to a face of dimension ≥ 1 of the epigraph of $\underline{u}$, then $\underline{p}(x) = 0$. In this case, we get the following alternative (or complementary slackness condition) for the solutions of (1): for a.e. $x \in \Omega$,

- either there is a line segment in the graph of $\underline{u}$ containing $(x, \underline{u}(x))$,
- or

$$\frac{\partial g}{\partial q}(x, \underline{u}, \nabla \underline{u}) = \nabla\psi.$$

6. Regularity

In this final section, given $q_0 \in L^2(\Omega)^N$, we consider the projection problem (2). We recall that its solution $\underline{u}$ is characterized by the conditions

$$q_0 = \nabla \underline{u} + \underline{p} \quad \text{with } \underline{p} \in \overline{\text{cone}}(S - \text{id}),\ \langle \nabla \underline{u}, \underline{p} \rangle = 0.$$

Proposition 6.1. *Let us assume that* $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$.

1. *if* q_0 *is continuous at* $x_0 \in \Omega$ *then* $\underline{u}$ *is differentiable at* x_0,
2. *if* $q_0 \in C^0(\Omega, \mathbb{R}^N)$ *then* $\underline{u} \in C^1(\Omega, \mathbb{R})$ *and if* $q_0 \in C^{0,\alpha}(\Omega, \mathbb{R}^N)$ *for some* $\alpha \in (0,1)$ *then* $\underline{u} \in C^{1,\alpha}(\Omega, \mathbb{R})$.

Proof. If $\underline{p} = 0$, there is nothing to prove, by assumption we therefore have $\underline{p} = \lambda(\tau - \text{id})$ with $\lambda > 0$ and $\tau \in \overline{\text{co}}\, S \setminus \{\text{id}\}$. Using the representation (18) of elements of $\overline{\text{co}}\, S$ as moments of bistochastic measures, we deduce that there exists a measurable family of probability measures $\{\mu_x\}_{x \in \Omega}$ such that:

$$\tau(x) = \int_\Omega y \, \mathrm{d}\mu_x(y), \quad \text{a.e. } x \in \Omega. \tag{28}$$

1. Assume by contradiction that q_0 is continuous at $x_0 \in \Omega$ and $\partial \underline{u}(x_0)$ is not reduced to a point. With no loss of generality we may assume that $\underline{u}(x_0) = 0$ and that 0 belongs to the relative interior of $\partial \underline{u}(x_0)$. By a standard convex analysis argument, there exist $k \leq N+1$, $(\alpha_1, \cdots, \alpha_k) \in (0,1)^k$ such that $\sum_{\imath=1}^k \alpha_i = 1$, sequences $(x_n^i)_n \in \Omega^{\mathbb{N}}$, for $i = 1, \cdots, k$ such that $\underline{u}$ is differentiable at every x_n^i, $q_0(x_n^i) = \nabla \underline{u}(x_n^i) + \lambda(\tau(x_n^i) - x_n^i)$, the measure $\mu_{x_n^i}$ is well-defined at x_n^i and:

$$\lim_n x_n^i = x_0, \quad \lim_n \nabla \underline{u}(x_n^i) = \gamma_i \neq 0,\ i = 1, \cdots, k, \quad \text{and} \quad \sum_{\imath=1}^k \alpha_i \gamma_i = 0$$

We then have:

$$q_0(x_n^i) = \nabla \underline{u}(x_n^i) + \lambda \left(\int_\Omega y d\mu_{x_n^i}(y) - x_n^i \right) \tag{29}$$

hence:

$$\sum_{i=1}^k \alpha_i \gamma_i \cdot q_0(x_n^i) = \sum_{i=1}^k \alpha_i \gamma_i \cdot \nabla \underline{u}(x_n^i) + \lambda \sum_{i=1}^k \alpha_i \left(\int_\Omega y \cdot \gamma_i d\mu_{x_n^i}(y) - \gamma_i \cdot x_n^i \right) \tag{30}$$

Using the fact that q_0 is continuous at x_0 and passing to the limit yields:

$$0 > -\sum_{i=1}^k \alpha_i |\gamma_i|^2 = \lambda \lim_n \sum_{i=1}^k \alpha_i \left(\int_\Omega y \cdot \gamma_i d\mu_{x_n^i}(y) \right) \tag{31}$$

Moreover, $\text{Supp}(\mu_{x_n^i}) \subset \partial \underline{u}^*(\nabla \underline{u}(x_n^i))$ so that for $\mu_{x_n^i}$-a.e. y one has:

$$\underline{u}(y) = \underline{u}(x_n^i) + \nabla \underline{u}(x_n^i) \cdot (y - x_n^i) \tag{32}$$

which can also be written as

$$\gamma_i \cdot y = \underline{u}(y) + \varepsilon_n^i(y) \tag{33}$$

with $\varepsilon_n^i(y) = \nabla\underline{u}(x_n^i) \cdot x_n^i - \underline{u}(x_n^i) + (\gamma_i - \nabla\underline{u}(x_n^i)) \cdot y$ which tends to 0 as n tends to ∞. Using (33), and the fact that $\underline{u} \geq \underline{u}(x_0) = 0$ we get:

$$\lim_n \int_\Omega y \cdot \gamma_i d\mu_{x_n^i}(y) \geq 0 \quad \text{for all } i = 1, \cdots, k$$

and the latter contradicts (31).

2. The previous argument actually proves that $\nabla\underline{u}$ is continuous on Ω provided q_0 is. This implies that τ is continuous on Ω too and

$$q_0(x) = \nabla\underline{u}(x) + \lambda(\tau(x) - x),\ \forall x \in \Omega. \tag{34}$$

In addition, since $\underline{u}$ is affine between x and $\tau(x)$, we also have:

$$\nabla\underline{u}(x) = \nabla\underline{u}(\tau(x)),\ \forall x \in \Omega. \tag{35}$$

Let $(x, y) \in \Omega \times \Omega$, using (34), one gets

$$\begin{aligned} &|\nabla\underline{u}(x) - \nabla\underline{u}(y)|^2 + \lambda(\tau(x) - \tau(y)) \cdot (\nabla\underline{u}(x) - \nabla\underline{u}(y)) \\ &= (q_0(x) - q_0(y) + \lambda(x - y)) \cdot (\nabla\underline{u}(x) - \nabla\underline{u}(y)). \end{aligned} \tag{36}$$

Using (35) yields

$$(\tau(x) - \tau(y)) \cdot (\nabla\underline{u}(x) - \nabla\underline{u}(y)) = (\tau(x) - \tau(y)) \cdot (\nabla\underline{u}(\tau(x)) - \nabla\underline{u}(\tau(y)) \geq 0.$$

With (36), we thus obtain:

$$|\nabla\underline{u}(x) - \nabla\underline{u}(y)| \leq |q_0(x) - q_0(y)| + \lambda|x - y|$$

If $q_0 \in C^{0,\alpha}(\Omega, \mathbb{R}^N)$ for some $\alpha \in (0, 1)$, the previous inequality implies that $\underline{u} \in C^{1,\alpha}(\Omega, \mathbb{R})$ and the proof is complete. □

Theorem 6.1 states two kinds of regularity results that are of quite different nature: the first assertion states a pointwise regularity result while the second states a global regularity result. Both statements are obtained under the assumption that $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$. This assumption is not satisfactory, since it is not always satisfied and at least very difficult to check. Concerning the global regularity result, we have not been able to remove it although we believe that the result remains true for a general $\underline{p} \in \overline{\text{cone}}(S - \text{id})$. Let us mention however that the proof above carries over when $\underline{p}$ is of the form $(\sigma - \text{id})f$ with $f \in C^0(\overline{\Omega}, \mathbb{R}_+)$ and σ in the closed convex hull of S_f defined in Remark 2.6 (but again this is a special case). We refer the reader to [3] for other C^1 regularity results obtained with totally different arguments.

The pointwise regularity part is more surprising (from the point of elliptic regularity theory for instance), and it does not hold in general (i.e. when $\underline{p} \notin \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$) as the next counter-example shows. Let $\Omega := (-3, 3) \times (-1, 1)$ and define for all $x = (x_1, x_2) \in \Omega$:

$$u_0(x) := \begin{cases} \max(|x_1 + 2|, |x_2|) - 1 & \text{if } \max(|x_1 + 2|, |x_2|) \leq 1 \\ \max(|x_1 - 2|, |x_2|) - 1 & \text{if } \max(|x_1 - 2|, |x_2|) \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Define $q_0 := \nabla u_0$ and $\underline{u}$ as the largest convex minorant of u_0. We then have $q_0 = \nabla\underline{u} + \nabla(u_0 - \underline{u})$. Since $u_0 - u \geq 0$ and $u_0 = u$ on $\partial\Omega$, $\nabla(u_0 - u) \in K^-$ (see Remark 2.6). On the other hand, it is easy to check that $\int_\Omega \nabla\underline{u} \cdot \nabla(u_0 - \underline{u}) = 0$. This proves that $\nabla\underline{u}$ is the projection of $q_0 := \nabla u_0$ on K. Let us remark now that $\underline{u}$ is singular on the segment $[(-2,0),(2,0)]$ whereas q_0 is continuous on this segment except at the points $(-2,0)$, $(-1,0)$, $(1,0)$ and $(2,0)$. This proves that the pointwise regularity statement is not true without additional assumption on the multiplier $\underline{p}$.

Acknowledgements. The authors thank Yann Brenier for bringing to their attention Proposition 2.1 and for stimulating discussions.

References

[1] Y. Brenier: Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. 44(4) (1991) 375–417.

[2] G. Buttazzo, V. Ferone, B. Kawohl: Minimum problems over sets of concave functions and related questions, Math. Nachr. 173 (1995) 71–89.

[3] G. Carlier, T. Lachand-Robert: Regularity of solutions for some variational problems subject to a convexity constraint, Comm. Pure Appl. Math. 54 (2001) 583–594.

[4] P. Cartier, J. M. G. Fell, P.-A. Meyer: Comparaison des mesures portées par un ensemble convexe compact, Bull. Soc. Math. France 92 (1964) 435–445.

[5] P. Choné, J.-C. Rochet: Ironing, sweeping and multidimensional screening, Econometrica 66 (1998) 783–826.

[6] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, North-Holland, Amsterdam (1972).

[7] P.-L. Lions: Identification du cône dual des fonctions convexes et applications, C. R. Acad. Sci., Paris, Sér. I, Math. 326 (1998) 1385–1390.

[8] P.-A. Meyer: Probabilités et Potentiel, Hermann, Paris (1966).

[9] R. T. Rockafellar: Characterization of the subdifferentials of convex functions, Pacific J. Math. 17 (1966) 497–510.

[10] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[11] L. C. Young: Lectures on the Calculus of Variations and Optimal Control, Chelsea, New York (1980).