

Représentation du cône polaire des fonctions convexes et applications

Guillaume Carlier^a, Thomas Lachand-Robert^b

^a Université Bordeaux I, MAB, UMR CNRS 5460 et Université Bordeaux IV, GRAPE, UMR CNRS 5113, avenue L. Duguit, 33608, Pessac, France

^b Lab. de mathématiques, UMR CNRS 5127, Université de Savoie, 73376 Le Bourget-du-Lac cedex, France

Reçu le 25 février 2002 ; accepté après révision le 23 juillet 2002

Note présentée par Pierre-Louis Lions.

Résumé

Suivant une idée de Y. Brenier, nous donnons une représentation du cône polaire de l'ensemble K des gradients des fonctions convexes, à l'aide des applications qui conservent la mesure. Cette représentation peut être reformulée en termes de mesures doublement stochastiques et admet une caractérisation géométrique.

Nous en déduisons une équation d'Euler–Lagrange et des résultats de régularité pour certains problèmes de minimisation posés dans l'ensemble K . *Pour citer cet article : G. Carlier, T. Lachand-Robert, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 571–576.*
© 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Representation of the polar cone of convex functions and applications

Abstract

Following Y. Brenier, we give a representation of the polar cone of the set K of the gradient of convex functions, implying the set of measure-preserving maps. This can also be formulated in terms of doubly stochastic measures, and has a geometrical characterization.

We deduce an Euler–Lagrange equation and regularity results for some minimization problems in the set K . *To cite this article: G. Carlier, T. Lachand-Robert, C. R. Acad. Sci. Paris, Ser. I 335 (2002) 571–576.*

© 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Abridged English version

1. Decomposition with gradients of convex functions

Let $\Omega \subset \mathbb{R}^N$ be convex and bounded. We denote by $\langle \cdot, \cdot \rangle$ the scalar product in the Hilbert space $L^2(\Omega)^N$ and

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

Its polar cone is defined by

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

Adresse e-mail : carlier@montesquieu.u-bordeaux.fr (G. Carlier).

URL : <http://www.lama.univ-savoie.fr/~lachand>.

The polar cone is a standard tool in the calculus of variations. It can be used in minimization problems on K , for instance the projection problem:

$$\begin{aligned} &\text{For } q_0 \in L^2(\Omega)^N \text{ given, find } \underline{q} \in K \text{ such that} \\ &\|\underline{q} - q_0\|_{L^2(\Omega)^N} = \inf_{q \in K} \|q - q_0\|_{L^2(\Omega)^N}. \end{aligned} \tag{1}$$

Here we can consider a *dual problem* which is a minimization problem on K^- (and a projection problem for (1)). Minimizers of the dual problem play the role of generalized Lagrange multipliers.

For this particular cone K , there exist several representations of K^- : see particularly [4]. The one presented here use a very different approach; it is very much in the spirit of classical calculus of variations, allowing use of well-known methods. It also gives a more intuitive representation of elements of K^- , as well as some geometrical interpretations, as expected when dealing with convex functions. For the proofs and additional details, we refer to [3].

Let us first define the set of measure-preserving maps:

$$S := \{s : \Omega \rightarrow \Omega : s\# dx = dx\}.$$

It is a semi-group for the composition law, whose neutral element is denoted by id . Here dx is the Lebesgue measure and $s\# dx$ is the measure defined by $\int_{\Omega} \varphi(x) s\# dx = \int_{\Omega} \varphi(s(x)) dx$ for any $\varphi \in C^0(\overline{\Omega})$.

This set S is of common use for mass-transportation problems, and Yann Brenier [1] proved that any vector field $q \in L^2(\Omega)^N$ can be written in the form $q = \nabla u \circ s$, with u convex and $s \in S$. We will now give a different decomposition. Let us first write a characterization of K^- :

THEOREM 1. – *The set K^- is closed and convex; it is the closure (in $L^2(\Omega)^N$) of the cone generated by $S - \text{id}$:*

$$K^- = \overline{\text{cone}}(S - \text{id}) = \mathbb{R}_+ \overline{\text{co}}(S - \text{id}).$$

This theorem follows directly from a property proved by Yann Brenier (see [1] and [3]):

PROPERTY 1. – *Let $q \in X$. Then $q \in K$ if and only if $\langle q, s - \text{id} \rangle \leq 0, \forall s \in S$.*

Indeed, this property can be written as $K = (S - \text{id})^- = (\mathbb{R}_+ \overline{\text{co}}(S - \text{id}))^-$. Taking the polar cone of each term gives the theorem.

This yields the promised decomposition:

THEOREM 2. – *Let $q_0 \in L^2(\Omega)^N$. There exists a unique $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ and a unique convex function \underline{u} (up to an additive constant) such that $q_0 = \nabla \underline{u} + \underline{p}$ and $\langle \underline{p}, \nabla \underline{u} \rangle = 0$.*

This is a classical result; $\nabla \underline{u} = \underline{q}$ is the solution of (1) and \underline{p} is the projection of q_0 on K^- .

2. A geometrical representation

Let us give some links between the abstract decomposition given in Theorem 2 and the geometrical properties associated with convex functions.

We recall that any convex function u has a closed epigraph (in $\Omega \times \mathbb{R}$): $\text{epi } u = \{(x, t) \in \Omega \times \mathbb{R}; t \geq u(x)\}$ (see, for instance, [5] for this property and the following ones). As any convex set with dimension $N + 1$, that one can be partitioned in relatively open convex faces with dimensions varying from 0 to $N + 1$. The only face with dimension $N + 1$ is the interior of the epigraph $\{(x, t); t > u(x)\}$. The projection of all other faces define a partition of Ω into relatively open convex subsets, which we call “faces of u ”. Faces of dimension N are open convex subsets where the gradient of u is well-defined and constant. For the other

faces, the gradient does not always exist, but the subdifferential ∂u is constant. Hence the “faces of u ” are in some sense the “level sets” of ∂u .

Let us note $F_u(x)$ the convex face of u containing x . It is easy to prove that $y \in \overline{F_u}(x)$ if and only if $u(y) = u(x) + \xi \cdot (y - x), \forall \xi \in \partial u(x)$.

THEOREM 3. – *Let $q_0 \in L^2(\Omega)^N, q \notin K$, so that in the decomposition $q_0 = \nabla \underline{u} + \underline{p}$ given by Theorem 2, we have $\underline{p} \neq 0$. We have then $\underline{p} = \lambda(\tau - \text{id})$ with $\lambda > 0$ and $\tau \in \overline{\text{co}} S$. Then for almost every $x \in \Omega$ we have*

$$\underline{u}(\tau(x)) = \underline{u}(x) + \nabla \underline{u}(x) \cdot (\tau(x) - x), \tag{2}$$

meaning that $\tau(x) \in \overline{F_{\underline{u}}}(x)$.

3. Doubly stochastic measures

Another way to express that $\tau \in \overline{\text{co}} S$ is to write it as the *moment of a doubly stochastic measure* γ on $\Omega \times \Omega$, that is, a Radon-measure such that $\iint f(x) d\gamma(x, y) = \iint f(y) d\gamma(x, y) = \int f$ for any bounded continuous function f on Ω . Such a measure has a disintegration $\{\mu_x\}_{x \in \Omega}$, so we have:

$$\tau(x) = \int_{\Omega} y d\mu_x(y), \quad \text{and} \quad \forall f \in C^0(\overline{\Omega}), \quad \int_{\Omega} \left(\int_{\Omega} f(y) d\mu_x(y) \right) dx = \int_{\Omega} f(x) dx. \tag{3}$$

Theorem 3 can be expressed as follows: for almost every $x, \text{Supp}(\mu_x) \subset \partial \underline{u}^*(\nabla \underline{u}(x))$, or more explicitly:

$$\underline{u}(y) = \underline{u}(x) + \nabla \underline{u}(x) \cdot (y - x), \quad \mu_x\text{-a.e. } y \in \Omega. \tag{4}$$

4. Applications

As an application of this representation of K^- , we consider the problem (1), or a more general one:

$$\inf_{u \text{ convex}} \int_{\Omega} g(x, \nabla u(x)) dx. \tag{5}$$

We can now write an Euler–Lagrange equation for this problem, and deduce a continuity property for the solutions:

THEOREM 4. – *Let $g : \Omega \times \mathbb{R}^N \rightarrow \mathbb{R}$ be continuous and continuously differentiable with respect to its second variable. If \underline{u} solves (5), then there exists $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ such that*

$$\frac{\partial g}{\partial q}(x, \nabla \underline{u}) + \underline{p} = 0. \tag{6}$$

If $\underline{p} \neq 0$, then $\underline{p} = \lambda(\tau - \text{id})$ with $\lambda > 0$ and $\tau \in \overline{\text{co}} S$ is such that \underline{u} is affine between x and $\tau(x)$ for almost every x .

Moreover if $\frac{\partial g}{\partial q}(x, q)$ is strictly monotone with respect to q for every x (i.e., $(\frac{\partial g}{\partial q}(x, q_1) - \frac{\partial g}{\partial q}(x, q_2)) \cdot (q_1 - q_2) > 0$ for all $(q_1, q_2) \in \mathbb{R}^N$ with $q_1 \neq q_2$), then \underline{u} is of class C^1 on Ω . That implies that τ is continuous and (2) holds for any x .

In particular, we see that the projection problem (1) has a continuous solution $\underline{q} = \nabla \underline{u}$ whenever q_0 is continuous.

The proof of this regularity property makes a direct use of the Euler–Lagrange equation (6) and of the integral representation τ in (3): we refer to [3]. Hence it is a very different approach from the direct methods used in [2] to prove similar results.

1. Décomposition avec les gradients des fonctions convexes

Soit $\Omega \subset \mathbb{R}^N$ un ouvert borné et convexe. Nous notons $\langle \cdot, \cdot \rangle$ le produit scalaire dans l’espace de Hilbert $L^2(\Omega)^N$ et

$$K := \{q \in L^2(\Omega)^N : q = \nabla u, \text{ avec } u \text{ convexe}\}.$$

Son cône polaire est défini par

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

Le cône polaire est un outil classique du calcul des variations ; lorsqu’on s’intéresse à un problème de minimisation dans K , comme par exemple le problème de projection :

Pour $q_0 \in L^2(\Omega)^N$ donné, trouver $\underline{q} \in K$ tel que

$$\|\underline{q} - q_0\|_{L^2(\Omega)^N} = \inf_{q \in K} \|q - q_0\|_{L^2(\Omega)^N}, \tag{1}$$

on peut utiliser un *problème dual* qui est encore un problème de minimisation, mais sur le cône polaire (et qui est une projection lorsque le problème primal en est un). Le ou les minimiseurs de ce problème dual jouent le rôle d’un « multiplicateur de Lagrange généralisé ».

Dans le cas particulier qui nous intéresse, où K est l’ensemble des gradients des fonctions convexes, plusieurs tentatives ont été faites pour donner une représentation commode de K^- : voir notamment [4]. Celle que nous présentons ici est basée sur une approche très différente ; celle-ci est particulièrement conforme à l’esprit du calcul des variations, permettant d’utiliser des méthodes qui ont fait leur preuves depuis longtemps. Mais elle donne également une représentation (assez) « intuitive » des éléments de K^- ainsi qu’une connection avec l’aspect géométrique naturellement associé aux fonctions convexes. Pour les démonstrations et différents détails supplémentaires, nous renvoyons à [3].

Pour définir cette caractérisation, nous introduisons l’ensemble des applications qui préservent la mesure :

$$S := \{s : \Omega \rightarrow \Omega : s\# dx = dx\}.$$

Il s’agit d’un semi-groupe de composition dont on notera id l’élément neutre. Ici dx est la mesure de Lebesgue et $s\# dx$ est la mesure définie par $\int_{\Omega} \varphi(x) s\# dx = \int_{\Omega} \varphi(s(x)) dx$ pour tout $\varphi \in C^0(\overline{\Omega})$.

Cet ensemble S est couramment utilisé dans les problèmes de transport de masse, et Yann Brenier a démontré que tout champ de vecteur $q \in L^2(\Omega)^N$ se décompose sous la forme $q = \nabla u \circ s$, avec u convexe et $s \in S$. Nous allons ici donner une décomposition différente. Pour cela, nous donnons d’abord une caractérisation de K^- :

THÉORÈME 1. – *L’ensemble K^- est fermé et convexe ; c’est l’adhérence (dans $L^2(\Omega)^N$) du cône engendré par $S - \text{id}$:*

$$K^- = \overline{\text{cône}(S - \text{id})} = \mathbb{R}_+ \overline{\text{co}}(S - \text{id}).$$

Ce théorème est une conséquence immédiate de la propriété suivante de Yann Brenier (voir [1] et [3]) :

PROPRIÉTÉ 1. – *Soit $q \in X$. Alors $q \in K$ si et seulement si $\langle q, s - \text{id} \rangle \leq 0, \forall s \in S$.*

En effet, nous pouvons abrégier cette proposition sous la forme $K = (S - \text{id})^- = (\mathbb{R}_+ \overline{\text{co}}(S - \text{id}))^-$. En prenant le cône polaire de part et d'autre, on obtient bien le résultat du théorème.

Nous en déduisons la décomposition annoncée :

THÉORÈME 2. – Soit $q_0 \in L^2(\Omega)^N$. Il existe un unique $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ et une unique fonction convexe \underline{u} (à l'addition d'une constante près) telles que $q_0 = \nabla \underline{u} + \underline{p}$ et $\langle \underline{p}, \nabla \underline{u} \rangle = 0$.

Cette décomposition est un résultat classique ; $\nabla \underline{u} = \underline{q}$ est la solution de (1) et \underline{p} est la projection de q_0 sur K^- .

2. Une représentation géométrique

Les fonctions convexes sont naturellement associées aux ensembles convexes et à de nombreuses propriétés géométriques. Nous cherchons ici à faire le lien en ces représentations classiques et la décomposition abstraite donnée dans le Théorème 2.

Nous rappelons que toute fonction convexe u a un épigraphe convexe fermé (dans $\Omega \times \mathbb{R}$) : $\text{epi } u = \{(x, t) \in \Omega \times \mathbb{R}; t \geq u(x)\}$ (voir par exemple [5] pour cette propriété et les suivantes). Comme tout ensemble convexe de dimension $N + 1$, celui-ci admet une partition en faces convexes relativement ouvertes de dimensions variant de 0 à $N + 1$. L'unique face de dimension $N + 1$ est l'intérieur de l'épigraphe $\{(x, t); t > u(x)\}$. Les projections des faces de dimensions inférieures définissent une partition de Ω en ensembles convexes relativement ouverts, que nous appellerons les « faces de u ». Les faces de dimensions N sont des sous-ensembles ouverts convexes où le gradient de u est bien défini et constant. Sur les autres faces, le gradient n'est pas forcément bien défini, mais par contre le sous-différentiel ∂u est constant. Ainsi les « faces de u » peut être vues comme les « lignes de niveau » de l'application ensembliste ∂u .

Nous noterons $F_u(x)$ la face convexe de u contenant x . Il est facile de montrer que $y \in \overline{F}_u(x)$ si et seulement si $u(y) = u(x) + \xi \cdot (y - x)$, $\forall \xi \in \partial u(x)$.

THÉORÈME 3. – Soit $q_0 \in L^2(\Omega)^N$, $q \notin K$, de sorte que dans la décomposition $q_0 = \nabla \underline{u} + \underline{p}$ donnée au Théorème 2, on a $\underline{p} \neq 0$. On a donc $\underline{p} = \lambda(\tau - \text{id})$ avec $\lambda > 0$ et $\tau \in \overline{\text{co}} S$. Alors pour presque tout $x \in \Omega$ on a

$$\underline{u}(\tau(x)) = \underline{u}(x) + \nabla \underline{u}(x) \cdot (\tau(x) - x), \tag{2}$$

c'est-à-dire que $\tau(x) \in \overline{F}_u(x)$.

3. Mesures doublement stochastiques

Une autre façon d'exprimer le fait que $\tau \in \overline{\text{co}} S$ est de l'écrire comme le moment d'une mesure doublement stochastique γ sur $\Omega \times \Omega$, c'est-à-dire d'une mesure de Radon telle que $\int \int f(x) d\gamma(x, y) = \int \int f(y) d\gamma(x, y) = \int f$ pour toute fonction f continue et bornée sur Ω . Une telle mesure admet une désintégration $\{\mu_x\}_{x \in \Omega}$, et l'on a donc

$$\tau(x) = \int_{\Omega} y d\mu_x(y), \quad \text{et} \quad \forall f \in C^0(\overline{\Omega}), \quad \int_{\Omega} \left(\int_{\Omega} f(y) d\mu_x(y) \right) dx = \int_{\Omega} f(x) dx. \tag{3}$$

L'équivalent du Théorème 3 s'exprime alors sous la forme : pour presque tout x , $\text{Supp}(\mu_x) \subset \partial \underline{u}^*(\nabla \underline{u}(x))$, ou plus explicitement :

$$\underline{u}(y) = \underline{u}(x) + \nabla \underline{u}(x) \cdot (y - x), \quad \mu_x\text{-a.e. } y \in \Omega. \tag{4}$$

4. Applications

Comme application de cette représentation de K^- , nous considérons le problème (1), ou plus généralement :

$$\inf_{u \text{ convexe}} \int_{\Omega} g(x, \nabla u(x)) \, dx. \quad (5)$$

Nous pouvons écrire une équation d'Euler pour ce problème, et en déduire une propriété de continuité de la solution :

THÉORÈME 4. – Soit $g : \Omega \times \mathbb{R}^N \rightarrow \mathbb{R}$ continue et continûment différentiable par rapport à sa deuxième variable. Si \underline{u} est une solution de (5), alors il existe $\underline{p} \in \mathbb{R}_+ \overline{\text{co}}(S - \text{id})$ tel que

$$\frac{\partial g}{\partial q}(x, \nabla \underline{u}) + \underline{p} = 0. \quad (6)$$

Si $\underline{p} \neq 0$, alors $\underline{p} = \lambda(\tau - \text{id})$ avec $\lambda > 0$ et $\tau \in \overline{\text{co}}S$ est tel que \underline{u} est affine entre x et $\tau(x)$ pour presque tout x .

De plus, si $\frac{\partial g}{\partial q}(x, q)$ est strictement monotone par rapport à q pour tout x (i.e. $(\frac{\partial g}{\partial q}(x, q_1) - \frac{\partial g}{\partial q}(x, q_2)) \cdot (q_1 - q_2) > 0$ pour tout $(q_1, q_2) \in \mathbb{R}^N$ tel que $q_1 \neq q_2$), alors \underline{u} est de classe C^1 sur Ω . De ce fait τ est aussi continue et (2) a lieu pour tout x .

En particulier, dans le problème de projection (1), on voit que le projeté $\underline{q} = \nabla \underline{u}$ est continu si q_0 est continu.

La preuve de ce résultat de régularité utilise de manière directe l'équation d'Euler-Lagrange (6) et la représentation intégrale de τ dans (3) : voir [3]. Il s'agit donc d'une approche très différente des méthodes directes utilisées par exemple dans [2] pour obtenir des résultats similaires.

Références bibliographiques

- [1] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44 (4) (1991) 375–417.
- [2] G. Carlier, T. Lachand-Robert, Regularity of solutions for some variational problems subject to convexity constraint, *Comm. Pure Appl. Math.* 54 (2001) 583–594.
- [3] G. Carlier, T. Lachand-Robert, Representation of the polar cone of convex functions and applications, to appear.
- [4] P.-L. Lions, Identification du cône dual des fonctions convexes et applications, *C. R. Acad. Sci. Paris* (1998).
- [5] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.