COLLOQUIUM MATHEMATICUM

VOL. LVII 1989 FASC. 2

QUELQUES REMARQUES SUR LE PROBLÈME DE LAGRANGE DU CALCUL DES VARIATIONS

PAR

JEAN-NOËL CORVELLEC (MONTRÉAL, QUÉBEC)

Dans cette note on présente certaines remarques concernant le problème de Lagrange du calcul des variations pour une fonctionnelle

$$I(y) = \int_{x_1}^{x_2} L(x, y(x), y'(x))dx,$$

où $L(\cdot, y(\cdot), y'(\cdot))$ n'est pas nécessairement défini pour $x = x_1, x_2$ (si y est une fonction admissible); on montre que si L est de classe C^1 sur son domaine, et si $L(x, \cdot, \cdot)$ est convexe, alors l'équation d'Euler–Lagrange est une condition nécessaire pour qu'une function absolument continue $\hat{y}$ minimise I, qui est suffisante si

$$\frac{\partial L}{\partial y}(\cdot, \hat{y}(\cdot), \hat{y}'(\cdot))$$

est sommable sur $[x_1, x_2]$. On donne deux exemples d'application de ce résultat, dont le problème du brachistochrone, qui est ainsi résolu de façon très simple. Je voudrais remercier le Professeur A. Granas qui m'a incité à m'intéresser à cette question dans le cadre de son séminaire à l'Université de Montréal.

1. Énoncé du résultat. On considère le problème classique du calcul des variations suivant: minimiser la fonctionnelle

$$I(y) = \int_{x_1}^{x_2} L(x, y(x), y'(x))dx \qquad \left(x_1, x_2 \in \boldsymbol{R},\ x_1 < x_2;\ y = (y^1, \ldots, y^m),\ (') = \frac{d}{dx}\right)$$

avec la contrainte

$$(1) \qquad y(x) \in A \subset \boldsymbol{R}^m \quad \text{pour tout } x \in (x_1, x_2),$$

et la condition aux limites

$$(2) \qquad y(x_i) = y_i \text{ fixé}, \quad i = 1, 2.$$

On suppose que A est ouvert et convexe, et les hypothèses sur le lagrangien L sont:

(H1) L: $(x_1, x_2) \times A \times \boldsymbol{R}^m \to \boldsymbol{R}$ est de classe C^1;

(H2) $(y, y') \mapsto L(x, y, y')$ est convexe sur $A \times \boldsymbol{R}^m$ pour chaque $x \in (x_1, x_2)$.

On minimise sur la classe Ω des fonctions absolument continues (AC)

$$y\colon [x_1, x_2] \to \boldsymbol{R}^m$$

qui vérifient (1), (2), et telles que

$$L(\cdot, y(\cdot), y'(\cdot)) \in L^1(x_1, x_2).$$

Notons que par l'hypothèse (H2), Ω (qui est supposée non vide — en particulier $y_i \in \bar{A}$, $i = 1, 2$) et I: $\Omega \to \boldsymbol{R}$ sont convexes. On note

$$L_y = \left(\frac{\partial L}{\partial y^i}\right)_{1 \leqslant i \leqslant m} \quad \text{et} \quad L_{y'} = \left(\frac{\partial L}{\partial y'^i}\right)_{1 \leqslant i \leqslant m}$$

les vecteurs des dérivées partielles de $L(x, y, y')$, et on montre le résultat suivant:

THÉORÈME. *Avec les notations et les hypothèses précédentes, soit $\tilde{y} \in \Omega$ tel que $\tilde{y}' \in L^\infty_{\text{loc}}(x_1, x_2)^m$. On a alors:*

(a) *Si $\tilde{y}$ réalise le minimum de I sur Ω, alors $\tilde{y}$ est solution de l'équation (vectorielle) d'Euler–Lagrange pour I:*

$$(*) \qquad L_y(x, y(x), y'(x)) = \frac{d}{dx} L_{y'}(x, y(x), y'(x))$$

pour presque tout $x \in (x_1, x_2)$.

(b) *Si $\tilde{y}$ est solution de $(*)$ (p.p. $x \in (x_1, x_2)$) et*

$$L_y(\cdot, \tilde{y}(\cdot), \tilde{y}'(\cdot)) \in L^1(x_1, x_2)^m,$$

alors $\tilde{y}$ réalise le minimum de I sur Ω.

Dans le cas où $A = \boldsymbol{R}^m$, L: $[x_1, x_2] \times \boldsymbol{R}^m \times \boldsymbol{R}^m \to \boldsymbol{R}$ est de classe C^1 et $L(x, \cdot, \cdot)$ est convexe pour tout $x \in [x_1, x_2]$, on en déduit immédiatement le

COROLLAIRE. *Une fonction $\tilde{y} \in \Omega$ telle que $\tilde{y}' \in L^\infty(x_1, x_2)^m$ minimise I si et seulement si*

$$L_y(x, \tilde{y}(x), \tilde{y}'(x)) = \frac{d}{dx} L_{y'}(x, \tilde{y}(x), \tilde{y}'(x))$$

presque partout sur $[x_1, x_2]$.

Remarque. Si une fonction AC $\tilde{y}$ vérifie les hypothèses de la partie (b) du théorème, et si $L(x, \cdot, \cdot)$ est *strictement* convexe, alors $\tilde{y}$ est l'unique fonction minimisante de I; en pratique, on peut donc utiliser ce résultat dans des situations où A n'est pas (au moins a priori) nécessairement *ouvert*.

2. Trois lemmes préliminaires. Pour démontrer ce théorème, on se sert de deux lemmes élémentaires essentiellement fondés sur l'hypothèse de convexité (H2). Notons d'abord que, pour y, $z \in \Omega$, la dérivée directionnelle de I en y dans la direction $z-y$

$$I'(y, z-y) \stackrel{\text{df}}{=} \lim_{\lambda \to 0^+} \lambda^{-1}(I(y+\lambda(z-y))-I(y))$$

est bien définie (dans $\boldsymbol{R} \cup \{-\infty\}$); en effet, I étant convexe, la fonction

$$\lambda \mapsto \lambda^{-1}(I(y+\lambda(z-y))-I(y)), \qquad \lambda \in (0, 1],$$

est croissante.

Avec les notations et les hypothèses précédentes, on a

LEMME 1. $\big(I(y) = \min_{z\in\Omega} I(z)\big) \Leftrightarrow \big(I'(y, z-y) \geqslant 0 \ \ \forall z \in \Omega\big)$.

Preuve. Soient $y, z \in \Omega$, et $\lambda_0 \in (0, 1]$; par la convexité de I on a

$$\begin{aligned} I'(y, z-y) &= \inf_{\lambda \in (0,1]} \lambda^{-1}(I(y+\lambda(z-y))-I(y)) \\ &\leqslant \lambda_0^{-1}(I(y+\lambda_0(z-y))-I(y)) \leqslant I(z)-I(y), \end{aligned}$$

d'où le résultat.

LEMME 2. *Soient* $y, z \in \Omega$. *Alors*

$$\begin{aligned} I'(y, z-y) = \int_{x_1}^{x_2} [L_y(x, y(x), y'(x))\cdot(z(x)-y(x)) \\ + L_{y'}(x, y(x), y'(x))\cdot(z'(x)-y'(x))]\,dx, \end{aligned}$$

où $\cdot$ *représente le produit scalaire dans* $\boldsymbol{R}^m$.

Preuve. Soient $y, z \in \Omega$ et $\{\lambda_n\} \subset (0, 1]$ une suite décroissante vers 0. Pour chaque $n \geqslant 1$, on définit (p.p. $x \in (x_1, x_2)$) la fonction φ_n par

$$\varphi_n(x) = \lambda_n^{-1}\big[L(x, (y+\lambda_n(z-y))(x), (y'+\lambda_n(z'-y'))(x))-L(x, y(x), y'(x))\big].$$

Par l'hypothèse (H2), $\{\varphi_n\} \subset L^1(x_1, x_2)$ est une suite décroissante ($\varphi_1(x) \geqslant \varphi_2(x) \geqslant \ldots$ (p.p.)) qui converge (par (H1)) vers

$$\varphi(x) = L_y(x, y(x), y'(x))\cdot(z(x)-y(x))+L_{y'}(x, y(x), y'(x))\cdot(z'(x)-y'(x)) \text{ (p.p.)}.$$

Il résulte alors du théorème de convergence monotone (ou du théorème de convergence dominée) de Lebesgue que

$$I'(y, z-y) \stackrel{\text{df}}{=} \lim_{n\to\infty} \int_{x_1}^{x_2} \varphi_n(x)dx = \int_{x_1}^{x_2} \varphi(x)dx,$$

ce qu'on voulait montrer.

La preuve de (a) se fait analoguement à la méthode classique de dérivation de l'équation d'Euler–Lagrange comme condition nécessaire pour un extremum (local en général), et utilise le lemme bien connu:

LEMME 3. *Soit* $h \in L^1(a, b)$ *une fonction (à valeurs réelles) telle que*

$$\int_a^b h(x)\eta'(x)dx = 0$$

pour toute fonction AC $\eta: [a, b] \to \mathbf{R}$ *telle que* $\eta' \in L^\infty(a, b)$ *et* $\eta(a) = \eta(b) = 0$. *Alors il existe une constante* c *telle que* $h(x) = c$ *presque partout sur* $[a, b]$. (Pour une preuve du lemme 3, voir [1], p. 42.)

3. Preuve du théorème. On fait la preuve dans le cas $m = 1$, la généralisation à m quelconque dans N étant élémentaire.

(a) Soit $\tilde{y} \in \Omega$ tel que $\tilde{y}' \in L^\infty_{\text{loc}}(x_1, x_2)$, on suppose que

$$I(\tilde{y}) = \min_{z \in \Omega} I(z).$$

Soient $a, b \in \mathbf{R}$ tels que $x_1 < a < b < x_2$, et $\eta: [x_1, x_2] \to \mathbf{R}$ une fonction AC telle que

$$\eta(x) = 0 \quad \text{si } x \in [x_1, x_2] \backslash (a, b) \text{ et } \eta' \in L^\infty(x_1, x_2).$$

Il existe $r > 0$ suffisamment petit tel que $y_+ = \tilde{y} + r\eta$ et $y_- = \tilde{y} - r\eta$ appartiennent à Ω; comme

$$I'(\tilde{y}, y_+ - \tilde{y}) = -I'(\tilde{y}, y_- - \tilde{y}),$$

il résulte du lemme 1 que $I'(\tilde{y}, r\eta) = 0$, soit, d'après le lemme 2:

$$\int_a^b [L_y(x, \tilde{y}(x), \tilde{y}'(x))\eta(x) + L_{y'}(x, \tilde{y}(x), \tilde{y}'(x))\eta'(x)]\,dx = 0. \tag{3}$$

D'après (H1) et les hypothèses sur $\tilde{y}$, la fonction

$$x \mapsto \int_a^x L_y(t, \tilde{y}(t), \tilde{y}'(t))dt$$

est AC sur $[a, b]$; on peut donc intégrer par parties dans (3) pour obtenir

$$\int_a^b [-\int_a^x L_y(t, \tilde{y}(t), \tilde{y}'(t))dt + L_{y'}(x, \tilde{y}(x), \tilde{y}'(x))]\,\eta'(x)dx = 0,$$

on en déduit, par le lemme 3, que pour une certaine constante c on a

$$\int_a^x L_y(t, \tilde{y}(t), \tilde{y}'(t))dt + c = L_{y'}(x, \tilde{y}(x), \tilde{y}'(x)) \text{ p.p. } x \in [a, b]. \tag{4}$$

La fonction $L_{y'}(\cdot, \tilde{y}(\cdot), \tilde{y}'(\cdot))$ est donc égale presque partout sur $[a, b]$ à une fonction AC; en identifiant ces deux fonctions sur $[a, b]$, et en effectuant la différentiation, on obtient

$$L_y(x, \tilde{y}(x), \tilde{y}'(x)) = \frac{d}{dx} L_{y'}(x, \tilde{y}(x), \tilde{y}'(x)) \text{ p.p. } x \in [a, b].$$

Comme $[a, b] \subset (x_1, x_2)$ est quelconque, on conclut que $\tilde{y}$ vérifie (*), p.p. $x \in (x_1, x_2)$.

(b) Soient $\tilde{y} \in \Omega$ tel que $\tilde{y}' \in L^\infty_{\text{loc}}(x_1, x_2)$ et $z \in \Omega$. Pour simplifier l'écriture, on note $\tilde{L}_y(x)$ pour $L_y(x, \tilde{y}(x), \tilde{y}'(x))$ et $\tilde{L}_{y'}(x)$ pour $L_{y'}(x, \tilde{y}(x), \tilde{y}'(x))$. L'hypothèse (H2) implique que la fonction ψ définie (p.p. $x \in (x_1, x_2)$) par

$$\psi(x) = L(x, z(x), z'(x)) - L(x, \tilde{y}(x), \tilde{y}'(x)) - [\tilde{L}_y(x)(z(x) - \tilde{y}(x)) + \tilde{L}_{y'}(x)(z'(x) - \tilde{y}'(x))]$$

est positive sur (x_1, x_2). Soit $\{\lambda_n\} \subset (0, \frac{1}{2}(x_2 - x_1))$ une suite décroissante vers 0; pour chaque $n \geqslant 1$, on définit la fonction

$$\psi_n = \psi \chi_{[x_1 + \lambda_n, x_2 - \lambda_n]},$$

produit de ψ et de la fonction caractéristique de l'intervalle $[x_1 + \lambda_n, x_2 - \lambda_n]$. $\{\psi_n\} \subset L^1(x_1, x_2)$ est une suite croissante de fonctions qui converge p.p. $x \in (x_1, x_2)$ vers $\psi(x)$, et on conclut comme dans la preuve du lemme 2 que

$$\int_{x_1}^{x_2} \psi(x)dx = \lim_{n \to \infty} \int_{x_1}^{x_2} \psi_n(x)dx = \lim_{n \to \infty} \int_{x_1 + \lambda_n}^{x_2 - \lambda_n} \psi(x)dx;$$

comme

$$\int_{x_1}^{x_2} [L(x, z(x), z'(x)) - L(x, \tilde{y}(x), \tilde{y}'(x))]\, dx \in \mathbf{R},$$

on a en particulier

$$I'(\tilde{y}, z - \tilde{y}) = \lim_{n \to \infty} \int_{x_1 + \lambda_n}^{x_2 - \lambda_n} [\tilde{L}_y(x)(z(x) - \tilde{y}(x)) + \tilde{L}_{y'}(x)(z'(x) - \tilde{y}'(x))]\, dx,$$

et en intégrant par parties

$$(5) \quad I'(\tilde{y}, z - \tilde{y}) = \lim_{n \to \infty} \Big\{ \int_{x_1 + \lambda_n}^{x_2 - \lambda_n} [- \int_{x_1 + \lambda_n}^{x} \tilde{L}_y(t)dt + \tilde{L}_{y'}(x)](z'(x) - \tilde{y}'(x))dx + \Big(\int_{x_1 + \lambda_n}^{x_2 - \lambda_n} \tilde{L}_y(t)dt\Big)(z(x_2 - \lambda_n) - \tilde{y}(x_2 - \lambda_n))\Big\}.$$

Si $\tilde{y}$ satisfait (*) (ou, de façon équivalente, (4) sur $[x_1 + \lambda_n, x_2 - \lambda_n]$, $n \geqslant 1$) et si $\tilde{L}_y \in L^1(x_1, x_2)$, on obtient avec (5) que $I'(\tilde{y}, z - \tilde{y}) = 0$, et puisque $z \in \Omega$ est quelconque, on conclut d'après le lemme 1 que

$$I(\tilde{y}) = \min_{z \in \Omega} I(z),$$

et le théorème est prouvé.

Remarque. L'hypothèse "$\tilde{y}' \in L^\infty_{\text{loc}}$" du théorème n'est pas essentielle, et peut être négligée dans certaines applications. Au vu de la démonstration, ce qui est précisément requis est:

$$(6) \qquad L_y(\cdot, \tilde{y}(\cdot), \tilde{y}'(\cdot)),\ L_{y'}(\cdot, \tilde{y}(\cdot), \tilde{y}'(\cdot)) \in L^1_{\text{loc}}(x_1, x_2).$$

Dans le cas du problème du brachistochrone (convexe), par exemple, le calcul des dérivées partielles de L montre que (6) est vérifié pour $\tilde{y} = y$ quelconque dans Ω.

Sous réserve d'obtenir (6), on peut également affaiblir l'hypothèse (H1) en supposant que L vérifie des conditions de Carathéodory, c'est-à-dire:

(i) $x \mapsto L(x, y, y')$ est mesurable sur (x_1, x_2), pour tout $(y, y') \in A \times \boldsymbol{R}^m$;

(ii) $(y, y') \mapsto L(x, y, y')$ est de classe C^1 sur $A \times \boldsymbol{R}^m$, p.p. $x \in (x_1, x_2)$.

En effet, le lemme 2 demeure valide sous l'hypothèse $\{$(i), (ii)$\}$.

4. Exemples. Le résultat présenté dans cette note a été suggéré par un papier de Kosmol [2], où l'on peut voir que le problème classique du brachistochrone (sans vitesse initiale) peut être résolu de façon simple en le transformant en un problème convexe équivalent; ceci se fait au moyen d'un simple changement de variable. En plus de cet exemple, on donne ici celui d'un autre problème qui se ramène à un problème convexe par changement de variable. Pour ces deux exemples, comparer avec [1], sections 3.12 et 3.10 respectivement.

4.1. *Problème du brachistochrone.* Il s'agit de minimiser la fonctionnelle

$$I_0(y) = \int_0^a y(x)^{-1/2}(1 + y'(x)^2)^{1/2}\,dx, \qquad a > 0,$$

sur

$$\Omega_0 = \{y \in \mathrm{AC}([0, a], \boldsymbol{R})\colon\ y > 0 \text{ sur } (0, a],\quad y(0) = 0,\ y(a) = y_a > 0 \text{ et } I_0(y) < +\infty\}.$$

$L_0(y, y') = y^{-1/2}(1 + y'^2)^{1/2}$ n'est pas convexe, mais si on effectue le changement de variable $s^2 = y$, alors la fonction L, définie sur $(0, +\infty) \times \boldsymbol{R}$ par

$$L(s, s') = (s^{-2} + 4s'^2)^{1/2} = L_0(s^2, 2ss'),$$

est strictement convexe; il est facile de s'en rendre compte en notant que

$$L(s, s') = \|(s^{-1}, 2s')\|,$$

où $\|\cdot\|$ est la norme euclidienne dans $\boldsymbol{R}^2$. Le problème est alors équivalent à celui de minimiser la fonctionnelle

$$I(s) = \int_0^a (s(x)^{-2} + 4s'(x)^2)^{1/2}\,dx$$

sur

$$\Omega = \{s \in \mathrm{AC}([0, a], \boldsymbol{R})\colon\ s > 0 \text{ sur } (0, a],\ s(0) = 0,\quad s(a) = y_a^{1/2} \text{ et } I(s) < +\infty\},$$

c'est-à-dire que la correspondance $y \mapsto y^{1/2}$ est une bijection de Ω_0 sur Ω, et que $I_0(y) = I(y^{1/2})$ pour tout $y \in \Omega_0$.

On montre facilement que si $y \in \Omega_0$ est solution de l'équation d'Euler–Lagrange pour I_0 (p.p. $x \in (0, a)$), alors y est de classe C^2 sur $(0, a]$, et puisque L est de classe C^2, on a

$$\frac{d}{dx}(L_0(y, y') - y'(L_{0y'}(y, y'))) = y'\left(L_{0y}(y, y') - \frac{d}{dx}L_{0y'}(y, y')\right) = 0$$

sur $(0, a]$,

et y est donc solution de l'équation

$$L_0(y(x), y'(x)) - y'(x)L_{0y'}(y(x), y'(x)) = c_1, \quad c_1 = \text{const}, \ x \in (0, a],$$

soit

$$(**) \qquad y(x)(1 + y'(x)^2) = c^2, \quad x \in (0, a], \ c = c_1^{-1} > 0.$$

En introduisant le paramètre τ par l'intermédiaire de l'équation

$$y' = -\tan\left(\frac{\tau - \pi}{2}\right) = \frac{1 + \cos\tau}{\sin\tau},$$

on montre que (**) possède une solution $\tilde{y} \in \Omega_0$, un arc de cycloïde dont une représentation paramétrique est donnée par

$$x(\tau) = \frac{\tilde{c}^2}{2}(\tau - \sin\tau),$$

$$y(\tau) = \frac{\tilde{c}^2}{2}(1 - \cos\tau), \quad \tau \in [0, \tau_a],$$

où $\tilde{c}$ et $\tau_a \in (0, 2\pi)$ sont déterminés de façon unique par la condition $y(a) = y_a$.

On obtient alors que $\tilde{s} = \tilde{y}^{1/2}$ est solution sur $(0, a]$ de l'équation d'Euler–Lagrange pour I, $\tilde{s} \in \Omega$ ($I(\tilde{s}) = I_0(\tilde{y}) = \tilde{c}\tau_a$), est de classe C^∞ sur $(0, a]$, et

$$\int_0^a L_s(\tilde{s}(x), \tilde{s}'(x))dx = 2\left(\cos\frac{\tau_a}{2} - 1\right);$$

donc, par (b) du théorème, et puisque I est strictement convexe, $\tilde{s}$ est l'unique fonction minimisante de I, d'où $\tilde{y}$ est la solution du problème original.

4.2. *L'intégrale* $I_0 = \int yy'^2 dx$. On veut minimiser la fonctionnelle

$$I_0(y) = \int_{x_1}^{x_2} y(x)y'(x)^2 dx, \quad x_1 < x_2,$$

sur

$$\Omega_0 = \{y \in \mathrm{AC}([x_1, x_2], \boldsymbol{R}) \colon y \geqslant 0 \text{ sur } [x_1, x_2],$$
$$y(x_i) = y_i \geqslant 0, \ i = 1, 2, \text{ et } I_0(y) < +\infty\}.$$

Par le changement de variable $s^{2/3} = y$, on se ramène au problème de minimiser

$$I(s) = \int_{x_1}^{x_2} s'(x)^2 dx$$

sur

$$\Omega = \{s \in \mathrm{AC}([x_1, x_2], \boldsymbol{R}):\ s(x_i) = y_i^{3/2},\ i = 1, 2, \quad \text{et } I(s) < +\infty\}.$$

L'équation d'Euler–Lagrange pour I se réduit à $s' = c$, $c = \text{const}$, et une fonction $\tilde{s} \in \Omega$ qui vérifie cette équation sur $[x_1, x_2]$ est donnée par

$$\tilde{s}(x) = \tilde{c}x + \tilde{d}, \qquad x \in [x_1, x_2],$$

où

$$\tilde{c} = (y_2^{3/2} - y_1^{3/2})(x_2 - x_1)^{-1} \quad \text{et} \quad \tilde{d} = (y_1^{3/2} x_2 - y_2^{3/2} x_1)(x_2 - x_1)^{-1}.$$

$L(s') = s'^2$ étant de classe C^∞ et strictement convexe sur $\boldsymbol{R}$, et $\tilde{s}$ étant de classe C^∞ sur $[x_1, x_2]$, $\tilde{s}$ est l'unique fonction minimisante de I; comme $y \mapsto y^{3/2}$ est injective de Ω_0 dans Ω, et $I_0(y) = \frac{4}{9} I(y^{3/2})$ pour tout $y \in \Omega_0$, on en conclut que $\tilde{y} = \tilde{s}^{2/3} \in \Omega_0$ est l'unique point minimisant de I_0.

TRAVAUX CITÉS

[1] L. Cesari, *Optimization – Theory and Applications*, Springer-Verlag, 1983.

[2] P. Kosmol, *Bemerkungen zur Brachistochrone*, Abh. Math. Sem. Univ. Hamburg 54 (1984), pp. 91–94.

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE „A"
MONTRÉAL, P.Q.
H3C 3J7 CANADA

Reçu par la Rédaction le 18.7.1987