MATH. SCAND. 54 (1984), 121-129

FONCTIONS CONVEXES DE PREMIÈRE CLASSE

JEAN SAINT RAYMOND

Resumé.

On démontre ici que si X est un convexe borné complet dans un e.l.c.s., toute fonction convexe f de première classe sur X est fortement convexe, c'est-à-dire vérifie, pour toute probabilité μ de Radon sur X, l'inégalité

$$\int f(x)\,d\mu(x) \geqq f(r\mu)$$

si $r\mu$ est le barycentre de μ.

Abstract.

It is proved that if X is a bounded complete convex subset of a Hausdorff locally convex space, every convex function f of first class on X is strongly convex, i.e. satisfies, for all Radon probability measure μ on X,

$$\int f(x)\,d\mu(x) \geqq f(r\mu)$$

where $r\mu$ denotes the barycenter of μ.

Le résultat que nous démontrons ici étend aux fonctions convexes un résultat antérieur de G. Choquet, relatif aux fonctions affines de première classe sur un convexe compact, qui affirme que ces fonctions vérifient le calcul barycentrique.

On remarque pour commencer que l'intégrale ci-dessus a un sens dans $\mathbf{R} \cup \{+\infty\}$ puisque f est minorée sur X, comme il résulte de la démonstration de [2]. En effet, même si f n'est pas borélienne, la fonction $\tilde{f}$ construite sur le cône positif P de l^1 est de première classe sur l'espace polonais P, donc borélienne.

On étudie maintenant le cas où la mesure μ est atomique.

THÉORÈME 1. *Soit X un convexe borné complet dans un* e.l.c.s. *Alors toute fonction convexe f de première classe sur X est dénombrablement convexe.*

Reçu le 7. février, 1983.

S'il n'en était pas ainsi, il existerait une suite (x_n) dans X et $\lambda^0=(\lambda_n^0)$ dans la base $\Delta=\{\lambda \in P \mid \|\lambda\|=1\}$ du cône P tels que

$$f\left(\sum_0^\infty \lambda_n^0 x_n\right) > \sum_{n=0}^\infty \lambda_n^0 f(x_n) .$$

Quitte à ajouter une constante à f, qui est minorée sur X, on peut supposer f positive. Si on pose $\alpha_n=\lambda_n^0 f(x_n)$, la série à termes positifs (α_n) est convergente, de somme inférieure à $f(\sum_0^\infty \lambda_n^0 x_n)$. On définit une fonction sous-linéaire $\bar{f}$ sur le cône P par

$$\begin{cases} \bar{f}(o) = 0 \\ \bar{f}(\lambda) = \|\lambda\| . f\left(\sum_0^\infty \frac{\lambda_n}{\|\lambda\|} x_n\right) . \end{cases}$$

Puisque l'application

$$b \colon \lambda \mapsto \sum_0^\infty \frac{\lambda_n}{\|\lambda\|} x_n$$

est continue de $P \setminus \{0\}$ dans X, $\bar{f}$ est de première classe sur $P \setminus \{0\}$, donc aussi sur P.

On définit maintenant une application φ sur le compact $K=\{0,1\}^{\mathbf{N}}$ par $\varphi(\xi)=\bar{f}(\xi.\lambda^0)$ où $\xi.\lambda^0=(\xi_n.\lambda_n^0)_n$.

Puisque l'application $\xi \mapsto \xi.\lambda^0$ est continue de K dans P, φ est de première classe de Baire sur K. Alors l'ensemble

$$A = \left\{\xi \in K \mid \varphi(\xi) \leqq \sum_0^\infty \xi_n \alpha_n\right\}$$

est un G_δ de K puisque la fonction $\xi \mapsto \varphi(\xi)-\sum_0^\infty \xi_n\alpha_n$ est de première classe, et dense dans K puisque si ξ est la fonction caractéristique d'une partie finie J de $\mathbf{N}$, on a, en notant (e_n) la base canonique de l^1:

$$\varphi(\xi) = \bar{f}\left(\sum_{n\in J} \lambda_n^0 e_n\right) \leqq \sum_{n \in J} \lambda_n^0 \bar{f}(e_n) = \sum_{n\in J} \lambda_n^0 f(e_n) = \sum_0^\infty \xi_n\alpha_n .$$

L'application $\xi \mapsto \xi^*=1-\xi$ est un homéomorphisme de K sur lui-même; donc $A^*=\{\xi^* \mid \xi \in A\}$ est un G_δ dense de K. Il en résulte que $A \cap A^* \neq \emptyset$. Et si $\xi \in A \cap A^*$, on a $\xi \in A$ et $\xi^* \in A$, donc, puisque $\lambda^0=\xi\lambda^0+\xi^*\lambda^0$:

$$\begin{aligned} \varphi(1) = \bar{f}(\lambda^0) &\leqq \bar{f}(\xi.\lambda^0)+\bar{f}(\xi^*\lambda^0) = \varphi(\xi)+\varphi(\xi^*) \\ &\leqq \sum_0^\infty \xi_n\alpha_n+\sum_0^\infty (1-\xi_n)\alpha_n = \sum_0^\infty \alpha_n \end{aligned}$$

contrairement à l'hypothèse faite sur λ^0 que

$$\tilde{f}(\lambda^0) = f\left(\sum_0^\infty \lambda_n^0 x_n\right) > \sum_0^\infty \lambda_n^0 f(x_n) = \sum_0^\infty \alpha_n .$$

Et ceci achève la démonstration.

Pour passer au cas général, on étudie maintenant le cas où X est compact, et μ diffuse.

LEMME 2. *Soient X un convexe compact, μ une mesure diffuse sur X, f une fonction convexe de première classe sur X, et $\delta > 0$. Il existe un convexe fermé X' de X, distinct de X, μ' mesure sur X', inférieure à μ tels que $\mu = 0$ ou $\|\mu - \mu'\| > 0$ et*

$$\int_X f(x)\, d(\mu - \mu') \geqq \|\mu - \mu'\| (f(y) - \delta)$$

si on note y le barycentre de $(\mu - \mu')/\|\mu - \mu'\|$.

Si $\mu \neq 0$, on pose $S = \operatorname{supp} \mu$ et $Y = \overline{\operatorname{conv} S} \subset X$, d'où $Y \neq \emptyset$. Puisque f est de première classe, il existe un ouvert U non vide de Y tel que $f(U)$ ait un diamètre inférieur à δ. Il résulte du théorème de Krein-Milman que U contient un point $a = \sum_1^n \lambda_i y_i$, où $\lambda_i > 0$, $\sum_1^n \lambda_i = 1$, $y_i \in \mathscr{E}(Y)$. Les points extrémaux y_i ont des voisinages respectifs T_i qui sont des tranches $\{y \in Y \mid l_i(y) < \theta_i\}$ telles que

$$[z_i \in \bar{T}_i,\ i = 1, 2, \ldots, n] \Rightarrow \sum_1^n \lambda_i z_i \in U .$$

De plus, puisque $\mathscr{E}(Y)$ est inclus dans S, $T_i \cap S$ est un ouvert non vide de S, et $\mu(T_i) > 0$. Comme μ est diffuse, $\mu(\{y_i\}) = 0$; il existe donc des tranches T'_i, voisinages de y_i, définies par $T'_i = \{y \in Y \mid l'_i(y) < \theta'_i\}$ avec

$$\begin{cases} y_i \in T'_i \subset \bar{T}'_i \subset T_i \\ 0 < \dfrac{1}{\lambda_i} \mu(\bar{T}'_i) < \inf\limits_{k \leqq n} \dfrac{1}{\lambda_k} \mu(T_k) . \end{cases}$$

On choisit alors ϱ tel que

$$\sup_i \frac{1}{\lambda_i} \mu(\bar{T}'_i) < \varrho < \inf_k \frac{1}{\lambda_k} \mu(T_k) .$$

Pour chaque i, la fonction

$$\varphi_i(x, s) = \left[\frac{(\theta_i - l_i(x))_+}{(\theta_i - l_i(x))_+ (l'_i(x) - \theta'_i)_+} \right]^{s/1-s}$$

est continue de $Y \times]0, 1[$ dans $[0, 1]$, vaut 1 si $x \in \bar{T}'_i$ et 0 si $x \notin T_i$. Donc

$$\psi_i(s) = \int_Y \varphi_i(x, s)\, d\mu(x)$$

est continue sur $]0, 1[$, et

$$\lim_{s \to 0} \psi_i(s) = \mu(T_i) > \lambda_i \varrho > \lim_{s \to 1} \psi_i(s) = \mu(\bar{T}'_i) .$$

Il existe donc $s_i \in\,]0, 1[$ tel que $\psi_i(s_i) = \lambda_i \varrho$. On pose alors $d\mu_i(x) = \varphi_i(x, s_i) d\mu(x)$. Il en résulte que μ_i est portée par $\bar{T}_i$, $\|\mu_i\| = \lambda_i \varrho$ et $\mu - \mu_i$ est portée par $Y - T'_i$. Si on pose maintenant

$$\begin{cases} \mu' = \mu - \sum_{i=1}^{n} \mu_i \\ z_i = r\left(\dfrac{\mu_i}{\|\mu_i\|}\right) \in \bar{T}_i \\ X' = \bigcap_{i=1}^{n} (Y \setminus T'_i) \subset X \end{cases}$$

on a $\|\mu - \mu'\| = \sum_1^n \|\mu_i\| = \sum_1^n \lambda_i \varrho = \varrho > 0$, $X \setminus X' \supset Y \setminus X' \neq \emptyset$. De plus, si on définit sur Y,

$$\check{f}(y) = \liminf_{\substack{z \to y \\ z \in Y}} f(z)$$

$\check{f}$ est convexe et s.c.i, minore f, et

$$\forall\, y \in U \quad \check{f}(y) \geqq \inf_{z \in U} f(z) \geqq f(y) - \delta .$$

Alors

$$\int_X f(x)\, d\mu_i(x) = \int_{\bar{T}_i} f(x)\, d\mu_i(x) \geqq \int_{\bar{T}_i} \check{f}(x)\, d\mu_i(x) \geqq \|\mu_i\| \check{f}(z_i)$$

puisque $\check{f}$ vérifie les inégalités barycentriques.

$$\int_X f(x)\, d(\mu - \mu') = \sum_1^n \int_X f(x)\, d\mu_i(x) \geqq \sum_1^n \lambda_i \varrho \check{f}(z_i) \geqq \varrho \check{f}\left(\sum_1^n \lambda_i z_i\right)$$
$$\geqq \varrho\left[f\left(\sum_1^n \lambda_i z_i\right) - \delta\right]$$

car $z_i \in \bar{T}_i$ et $\sum_1^n \lambda_i z_i \in U$. Et puisque $y = \sum_1^n \lambda_i z_i$ est le barycentre de $(\mu - \mu')/\|\mu - \mu'\|$, on obtient le résultat cherché.

THÉORÈME 3. *Soient X un convexe compact et μ une probabilité de Radon diffuse sur X. Si f est une fonction convexe de première classe sur X, on a*

$$\int_X f(x)\,d\mu(x) \geqq f(r\mu)$$

si $r\mu$ est le barycentre de μ.

On construit par récurrence, éventuellement transfinie, une suite décroissante de convexes compacts X_α et une suite décroissante μ_α de mesures portées par X_α, indexées par les ordinaux, avec

$$X_0 = X;\ \mu_0 = \mu$$

$$\|\mu_{\alpha+1}\| < \|\mu_\alpha\| \text{ ou } \mu_\alpha = 0$$

$$\int f(x)\,d(\mu_\alpha - \mu_{\alpha+1}) \geqq \|\mu_\alpha - \mu_{\alpha+1}\| \cdot (f(y) - \delta)$$

$$\text{où } y_\alpha = r\left(\frac{\mu_\alpha - \mu_{\alpha+1}}{\|\mu_\alpha - \mu_{\alpha+1}\|}\right)$$

$$\text{Pour } \tau \text{ de 2e espèce } X_\tau = \bigcap_{\alpha<\tau} X_\alpha,\ \mu_\tau = \lim_{\alpha\nearrow\tau} \mu_\alpha\,.$$

On construit $\mu_{\alpha+1}$ à partir de μ_α en utilisant le Lemme 2, et si τ est un ordinal limite, puisque

$$\alpha \leqq \beta < \tau \Rightarrow \|\mu_\alpha - \mu_\beta\| = \|\mu_\alpha\| - \|\mu_\beta\| \leqq \|\mu_\alpha\| - \inf_{\gamma<\tau} \|\mu_\gamma\|$$

les mesures $(\mu_\alpha)_{\alpha<\tau}$ convergent en norme vers une mesure μ_τ, portée par $\bigcap_{\alpha<\tau} X_\alpha = X_\tau$, quand α croît vers τ.

La suite $\|\mu_\alpha\|$ étant décroissante, il existe un ordinal dénombrable γ tel que $\|\mu_\gamma\| = \|\mu_{\gamma+1}\|$, donc $\mu_\gamma = 0$. On a alors, en norme,

$$\mu = \sum_{\alpha<\gamma} (\mu_\alpha - \mu_{\alpha+1})$$

$$1 = \|\mu\| = \sum_{\alpha<\gamma} \|\mu_\alpha - \mu_{\alpha+1}\|$$

d'où

$$\int_X f(x)\,d\mu(x) = \sum_{\alpha>\gamma} \int f(x)\,d(\mu_\alpha - \mu_{\alpha+1}) \geqq \sum_{\alpha<\gamma} \|\mu_\alpha - \mu_{\alpha+1}\| (f(y) - \delta)$$

$$\geqq \sum_{\alpha<\gamma} \|\mu_\alpha - \mu_{\alpha+1}\| \cdot f(y_\alpha) - \delta$$

et puisque

$$r\mu = \sum_{\alpha<\gamma} \|\mu_\alpha - \mu_{\alpha+1}\| \cdot y_\alpha$$

le Théorème 1 entraîne que

$$\int_X f(x)\,d\mu(x) \geqq \sum_{\alpha<\gamma} \|\mu_\alpha - \mu_{\alpha+1}\| \cdot f(y_\alpha) - \delta \geqq f(r\mu) - \delta$$

d'où le résultat, puisque δ est arbitraire.

THÉORÈME 4. *Soient X un convexe borné complet dans un* e.l.c.s. *et μ une probabilité de Radon sur X. Pour toute fonction convexe f de première classe, on α*

$$\int_X f(x)\,d\mu(x) \geqq f(r\mu)$$

si $r\mu$ désigne le barycentre de μ.

La mesure μ est la somme d'une mesure atomique $\mu_1 = \sum_1^\infty \lambda_n \varepsilon_{y_n}$ et d'une mesure de Radon diffuse μ_2, qui est portée par le réunion d'une suite (S_n) de compacts deux à deux disjoints. Si on note $v_n = \mu_{2|S_n}$, $X_n = \overline{\mathrm{conv}}\, S_n$ et $z_n = r(v_n/\|v_n\|)$, les convexes X_n sont compacts puisque X est complet et portent les mesures diffuses v_n. Il résulte alors du Théorème 3 que

$$\int_{X_n} f(x)\,dv_n(x) \geqq \|v_n\| \cdot f(z_n)$$

et puisque

$$1 = \|\mu\| = \|\mu_1\| + \|\mu_2\| = \sum_1^\infty |\lambda_n| + \sum_1^\infty \|v_n\|$$

il résulte du Théorème 1 que

$$\int f(x)\,d\mu(x) = \sum_1^\infty \lambda_n f(y_n) + \sum_1^\infty \int_{X_n} f(x)\,dv_n(x) \geqq \sum_1^\infty \lambda_n f(y_n) + \sum_1^\infty \|v_n\| f(z_n)$$
$$\geqq f\left(\sum_1^\infty \lambda_n y_n + \sum_1^\infty \|v_n\| \cdot z_n\right)$$

qui est l'inégalité cherchée, puisque

$$r\mu = \sum_1^\infty \lambda_n y_n + \sum_1^\infty \|v_n\| \cdot z_n\,.$$

Et ceci achève la démonstration.

THÉORÈME 5. *Soient X un convexe compact métrisable et f une fonction convexe de première classe sur X. Alors f est limite simple d'une suite de fonctions convexes continues (f_n).*

Puisque f est de première classe de Baire et minorée par m, il existe une suite g_n de fonctions continues minorées par m qui converge vers $f^{\cdot}$ simplement. Si on pose $g_n' = \sup_{p \geqq n} g_p + 1/n$ et $g_n'' = \inf_{p \geqq n} g_p$, g_n' est s.c.i., supérieure à $f + 1/n$, et g_n'' est s.c.s., minorée par m et inférieure à f. Si K_n désigne le compact de

$$X \times \mathbf{R} : \{(x,t) \mid m \leqq t \leqq g_n''(x)\} ,$$

F_n le fermé de $X \times \mathbf{R}$ $\{(x,t) \mid t \geqq g_n'(x)\}$, et $C_n = \overline{\mathrm{conv}}\, F_n$, on a $K_n \cap C_n = \emptyset$. En effet, l'application r de $\mathscr{M}^1(F_n)$ dans $X \times \mathbf{R}$ est affine et propre et son image contient F_n. Donc $r(\mathscr{M}^1(F_n))$ est un convexe fermé de $X \times \mathbf{R}$ contenant F_n, donc C_n. Si $\bar{\omega}$ est une probabilité sur F_n de résultante (x,t) et de projection μ sur X, on a $\forall (y,s) \in F_n$ $s \geqq g_n'(y) \geqq f(y) + 1/n$ donc

$$t \geqq \int_X \left(f(y) + \frac{1}{n}\right) d\mu(y) \geqq f(r\mu) + \frac{1}{n}$$

$$t \geqq f(x) + \frac{1}{n}$$

ce qui prouve que $C_n \subset \{(x,t) \mid t > f(x)\} \subset K_n^c$.

Le théorème de Hahn–Banach permet alors de séparer chaque point de K_n du convexe C_n par une forme linéaire $t - l_n(x)$. La compacité de K_n entraîne alors l'existence d'une famille finie $(l_1^{(n)}, \ldots, l_{k_n}^{(n)})$ de formes affines continues telles que

$$\forall x \in X \qquad g_n''(x) < \sup_{j \leqq k_n} l_j^{(n)}(X) < g_n'(X) .$$

Les fonctions convexes continues $f_n = \sup_{j \leqq k_n} l_j^{(n)}$ répondent à la question puisque

$$\liminf_{n \to \infty} f_n \geqq \liminf_{n \to \infty} g_n'' = f$$

et

$$\limsup_{n \to \infty} f_n \leqq \limsup_{n \to \infty} g_n' = f$$

ce qui termine la démonstration.

On peut démontrer plus généralement le théorème suivant:

THÉORÈME 6. *Soient X un convexe compact et f une fonction convexe de première classe sur X. Il existe un filtre sur l'ensemble des fonctions convexes continues sur X qui converge quasi-uniformément vers f.*

La démonstration est essentiellement la même. La fonction f est limite quasiuniforme d'un filtre $\mathscr{F}$ sur $\mathscr{C}(X)$ puisqu'elle est de première classe (cf. [3]). Pour chaque ensemble A du filtre $\mathscr{F}$ et chaque entier p, on considère les fonctions sur X:

$$u_{A,p}(x) = \inf_{g \in A} (g(x) - 1/p)$$

$$v_{A,p}(x) = \sup_{g \in A} (g(x) + 1/p) .$$

La fonction $u_{A,p}$ est s.c.s. et inférieure à $f - 1/p$. La fonction $v_{A,p}$ est s.c.i. et supérieure à $f + 1/p$. La fonction $\check{v}_{A,p}$ définie par

$$\check{v}_{A,p}(x) = \inf\left\{\int_X v_{A,p}(y)\, d\mu(y) \;\middle|\; \mu \in \mathscr{M}^1(X) \text{ et } r\mu = x\right\}$$

est convexe, s.c.i. et inférieure à $v_{A,p}$. De plus, on a

$$\check{v}_{A,p}(x) \geqq \inf\left\{\int_X f(y)\, d\mu(y) + \frac{1}{y} \;\middle|\; u \in \mathscr{M}^1(X) \text{ et } r\mu = x\right\}$$

$$\geqq f(x) + \frac{1}{p} > u_{A,p}(x) .$$

On en déduit, comme dans le théorème précédent, l'existence d'une famille finie $(l_1, \ldots, l_k)$ de fonctions affines continues sur X telle que, pour tout x de X:

$$u_{A,p}(x) < \sup_{j \leqq k} l_j(x) < \check{v}_{A,p}(x) \leqq v_{A,p}(x) .$$

Alors, si on pose $\varphi_{A,p} = \sup_{j \leqq k} l_j$ et

$$C(A,p) = \{\varphi_{B,q} \mid B \in \mathscr{F},\ B \subset A,\ q \in \mathbf{N},\ q \geqq p\}$$

les

$$(C(A,p))_{\substack{A \in \mathscr{F} \\ p \in \mathbf{N}}}$$

forment la base d'un filtre $\mathscr{G}$ sur les fonctions convexes continues. Et $\mathscr{G}$ converge quasi-uniformément vers f puisque si F est fermé de X, et $\varepsilon > 0$, il existe $p > 2/\varepsilon$, $A \in \mathscr{F}$ et U ouvert de F, non vide si $F \neq \varnothing$, tels que

$$\sup_{g \in A, x \in U} |f(x) - g(x)| \leqq \frac{\varepsilon}{2}.$$

Alors, si $q \geqq p$

$x \in U$ et $B \subset A$

$$\begin{cases} u_{B,q}(x) \geqq f(x) - \frac{\varepsilon}{2} - \frac{1}{q} \geqq f(x) - \varepsilon \\ \check{v}_{B,q}(x) \leqq v_{B,q}(x) \leqq f(x) + \frac{\varepsilon}{2} + \frac{1}{q} \leqq f(x) + \varepsilon \end{cases}$$

d'où $|\varphi_{B,q}(x) - f(x)| \leqq \varepsilon$ ce qui achève la démonstration, puisque $\sup_{\varphi \in C(A,p), x \in U} |\varphi(X) - f(X)| \leqq \varepsilon$.

REFERENCES

1. G. Choquet, *Remarques à propos de la démonstration d'unicité de P.-A. Meyer*, *Séminaire de Théorie du Potentiel*, eds. M. Brelot, G. Choquet, et J. Deny, 6e année: 1961/62. Exposés 6-13. Secrétariat Mathématique, Paris, 1962.
2. J. Saint Raymond, *Fonctions convexes sur une convexe borné complet*, Bull. Sci. Math. 102 (1978), 331–336.
3. J. Saint Raymond, *Fonctions de première classe*, *Séminaire d'Initiation à l'Analyse*, 22ᵉ année 1982/83. Publications mathématiques de l'Université Pierre et Marie Curie, N° 59, Paris, 1983.

ÉQUIPE D'ANALYSE
UNIVERSITÉ PARIS VI
4, PLACE JUSSIEU
75230 PARIS-CEDEX 05
FRANCE