Séminaire Choquet.
Initiation à l'analyse

Gustave Choquet

Le problème des moments

Séminaire Choquet. Initiation à l'analyse, tome 1 (1962), exp. n° 4, p. 1-10
<http://www.numdam.org/item?id=SC_1962__1__A5_0>

Séminaire CHOQUET
(Initiation à l'Analyse)
1re année, 1961/62, n° 4

5 avril 1962

LE PROBLÈME DES MOMENTS

par Gustave CHOQUET

Ce problème a intéressé de très nombreux analystes pendant les cinquante dernières années. C'est STIELTJES qui baptisa le problème en 1894-95, dans un Mémoire riche d'idées nouvelles, où il introduisait, à côté d'applications des fractions continues, la notion d'intégrale qui porte maintenant son nom et qui, convenablement généralisée, devait conduire à la notion de mesure de Radon.

Voici de quoi il s'agit :

Si μ désigne une mesure de Radon positive sur R, telle que tout polynôme $t \to P(t)$ soit μ-intégrable, on appelle <u>moment d'ordre</u> n de μ le nombre $k_n = \int t^n \, d\mu(t)$.

Si alors (a_n) $(n = 0 , 1 , 2 , \ldots)$ désigne une suite donnée de nombres réels, la question se pose de savoir s'il existe une telle mesure $\mu \geqslant 0$ sur R pour laquelle $a_n = k_n$ pour tout n, et sous quelles conditions cette mesure μ éventuelle est unique.

Ce cas particulier du problème des moments s'appelle <u>problème de Hamburger</u>.

Un autre cas particulier, celui traité par STIELTJES, s'obtient en imposant à la mesure μ d'être portée par la demi-droite R_+.

Des recherches analogues, mais moins générales, de ČEBYŠEV, HEINE, A. MARKOV avaient précédé celles de STIELTJES.

D'autres les suivirent : celles de HAMBURGER, NEVANLINNA, M. RIESZ, CARLEMAN, HAUSDORFF, etc.

On peut formuler le problème de Hamburger d'une manière mieux adaptée à nos habitudes actuelles. En effet, se donner la suite (a_n) équivaut à se donner une forme linéaire T sur l'espace vectoriel $\mathcal{P}$ des polynômes à une variable $t \to P(t)$, ainsi définie :

$$\text{si} \quad P(t) = \sum \alpha_n t^n , \qquad T(P) = \sum \alpha_n a_n \quad .$$

Si le problème a une solution μ, T est évidemment positive, en ce sens que $T(P)$ est positif pour tout polynôme P tel que $P(t) \geqslant 0$ pour tout t. En effet, on a

$$T(P) = \int P(t)\, d\mu(t) \quad .$$

Le problème de Hamburger peut donc s'énoncer ainsi : Est-ce que toute forme linéaire positive T sur l'espace vectoriel $\mathcal{P}$ peut être identifiée à une mesure de Radon positive sur R, et dans quels cas celle-ci est-elle unique ?

Sous cette forme on aperçoit de nombreuses généralisations possibles. Soit V un espace vectoriel de fonctions numériques continues sur un espace localement compact X, et soit T une forme linéaire sur V, positive en ce sens que $T(f) \geqslant 0$ pour toute f partout positive sur X. A quelles conditions existe-t-il une mesure de Radon $\mu \geqslant 0$ sur X telle que toute $f \in V$ soit μ-intégrable et telle que $T(f) = \int f\, d\mu$ pour toute $f \in V$; et dans quelles conditions μ est-elle unique ?

Un cas particulier important de ce problème s'obtiendrait en prenant pour X un fermé quelconque de R^p, et pour V l'espace vectoriel des restrictions à X des polynômes à p variables sur R^p.

Nous aborderons à peine ici la question de l'unicité. Il serait cependant fort intéressant d'examiner jusqu'à quel point on peut utiliser dans le cadre général où nous nous placerons, les outils analytiques en usage dans les travaux classiques, tels que les fractions continues, les fonctions de variable complexe, les fractions rationnelles.

Pour un exposé classique de ces méthodes, nous renvoyons à l'excellente monographie de Shohat et Tamarkin ([1]).

1. Prolongement d'une forme linéaire positive.

PROPOSITION 1. - *Soient* E *un espace vectoriel sur* R, E' *un sous-espace vectoriel de* E, *et* P *un cône convexe de* E. *Si* $E' + P = E$, *toute forme linéaire* f' *sur* E, *qui est* $\geqslant 0$ *sur* $E' \cap P$, *se prolonge à* E *en une forme linéaire* f *qui est* $\geqslant 0$ *sur* $E \cap P$.

Pour que ce prolongement f *soit unique, il faut et il suffit que, pour tout* $x \in E$, *on ait*

([1]) SHOHAT (J. A.) and TAMARKIN (J. D.). - The problems of moments, 2nd edition. - New-York, American mathematical Society, 1950 (Mathematical Surveys, 1).

$$\sup_{\substack{a' \leqslant x \\ a' \in E'}} f'(a') = \inf_{\substack{x \leqslant b' \\ b' \in E'}} f'(b')$$

(<u>où la relation de pré-ordre</u> $\leqslant$ <u>est celle définie dans</u> E <u>par le cône convexe</u> P).

<u>Démonstration</u>. - On utilise le théorème de Zorn, après avoir ordonné l'ensemble des prolongements de f' comme suit :

Soit A l'ensemble des couples (F, φ), où F est un sous-espace vectoriel de E contenant E', et φ une forme linéaire sur F, qui prolonge f' et est $\geqslant 0$ sur $F \cap P$.

On pose $(F_1, \varphi_1) \prec (F_2, \varphi_2)$ si $F_1 \subset F_2$ et si φ_1 est la restriction de φ_2 à F_1.

Cet ensemble ordonné est évidemment inductif ; si (F, φ) en est un élément maximal, on veut montrer que $F = E$. Ceci revient (en changeant les notations) à montrer que le prolongement est possible lorsque E' est un hyperplan de E. Soit $x \in E \setminus E'$. Par hypothèse $E' + P = E$, ce qui entraîne $-E' - P = -E$, ou encore $E' - P = E$.

Il existe donc $a', b' \in E'$ tels que

$$a' \leqslant x \leqslant b' \quad .$$

Posons

$$\alpha = \sup_{\substack{a' \leqslant x \\ a' \in E'}} f'(a') \qquad \text{et} \quad \beta = \inf_{\substack{x \leqslant b' \\ b' \in E'}} f'(b') \quad .$$

Les relations $a' \leqslant x$ et $x \leqslant b'$ entraînent $a' \leqslant b'$, donc, puisque f' est croissante $f'(a') \leqslant f'(b')$, donc aussi $\alpha \leqslant \beta$.

Or si f est un des prolongements cherchés, la relation $a' \leqslant x \leqslant b'$ entraîne

$$f'(a') = f(a') \leqslant f(x) \leqslant f(b') = f'(b')$$

donc aussi

$$\alpha \leqslant f(x) \leqslant \beta \quad .$$

Inversement, montrons que, pour tout nombre k tel que $\alpha \leqslant k \leqslant \beta$, il existe une solution f au problème, telle que $f(x) = k$. En effet, tout point u de E s'écrit d'une façon unique

$$u = x' - \lambda x , \qquad \text{où} \quad x' \in E' \quad \text{et} \quad \lambda \in R \quad .$$

Posons $f(u) = f'(x') - \lambda k$; il est évident que f est une forme linéaire qui prolonge f' ; montrons que $f(u) \geqslant 0$ si $u \in P$. En effet, la condition $u \geqslant 0$ s'écrit $x' \geqslant \lambda x$.

Si $\lambda > 0$, ceci s'écrit $x \leqslant \frac{x'}{\lambda}$, d'où

$$\beta \leqslant f'(\frac{x'}{\lambda}) , \qquad \text{d'où} \quad f'(x') \geqslant \lambda\beta \geqslant \lambda k \quad .$$

Si $\lambda < 0$, ceci s'écrit $\frac{x'}{\lambda} \leqslant x$, d'où

$$f'(\frac{x'}{\lambda}) \leqslant \alpha , \qquad \text{d'où} \quad f'(x') \geqslant \lambda\alpha \geqslant \lambda k \quad .$$

Si $\lambda = 0$, la relation $x' \geqslant 0$ entraîne bien $f'(x') \geqslant 0$. Dans les trois cas, on a bien $f(u) = f'(x') - \lambda k \geqslant 0$.

Si pour tout $x \in E \setminus E'$ les nombres α et β associés à x sont égaux, on a forcément $f(x) = \alpha = \beta$, donc le prolongement f est unique. Par contre, s'il existe un x pour lequel $\alpha \neq \beta$, on peut commencer par prolonger f' à l'espace engendré par E' et x en donnant à $f(x)$ n'importe quelle valeur de l'intervalle $[\alpha , \beta]$; donc il n'y a pas unicité du prolongement.

<u>Application</u>. - Soit X un espace localement compact et soit V, un sous-espace vectoriel de $\mathcal{C}(X)$. Nous dirons que V est <u>adapté</u> si

1° Il n'existe aucun point de X en lequel toutes les f de V s'annulent ;

2° Le cône $V_+ = V \cap \mathcal{C}_+(X)$ engendre V ;

3° Toute $f \in V_+$ est dépassée par une autre $g \in V_+$ en ce sens que $g \geqslant f$ et que, pour tout scalaire $k > 0$, il existe un compact $K \subset X$ tel que $g(x) \geqslant kf(x)$ hors de K.

On peut alors énoncer

PROPOSITION 2. - Si V est un espace vectoriel adapté de fonctions continues sur X, pour toute forme linéaire T sur V qui est $\geqslant 0$ sur V_+, il existe au moins une mesure de Radon $\mu \geqslant 0$ sur X telle que toute f de V soit μ-intégrable et vérifie $\mu(f) = T(f)$.

Démonstration. - Avec les notations de la proposition 1, prenons :

E = le sous-espace de $\mathcal{C}(X)$, constitué par les fonctions qui sont minorées et majorées sur X par des éléments de V ;

$E' = V$;

$P = E_+ = E \cap \mathcal{C}_+(X)$.

La condition $E = E' + P$ est bien satisfaite ; donc T se prolonge en une forme linéaire U sur E, qui est $\geqslant 0$ sur E_+.

Or, pour tout point $x \in X$, il existe une $f \in V$ avec $f(x) \neq 0$, donc aussi avec $f \in V_+$ avec $f(x) > 0$. Il en résulte que, pour tout compact $K \subset X$, il existe une $g \in V_+$ qui est > 0 sur K (utiliser le théorème de Borel-Lebesgue). Donc $E_+ \supset \mathcal{K}_+(X)$, donc aussi $E \supset \mathcal{K}(X)$.

Ainsi la restriction de U à $\mathcal{K}(X)$ est une mesure de Radon $\mu \geqslant 0$ sur X.

Montrons que toute $f \in E$ est μ-intégrable et que $\mu(f) = U(f)$; il suffit pour cela de montrer que, pour toute $f \in V_+$ et pour tout $\varepsilon > 0$, on peut trouver $f_1 \in \mathcal{K}_+(X)$ et $f_2 \geqslant 0$ telles que

$$f = f_1 + f_2 \quad \text{et} \quad U(f_2) < \varepsilon \quad .$$

Or, comme V est adapté, il existe $g \in V_+$ qui dépasse f ; pour tout $k > 0$, il existe un compact K tel que $g \geqslant kf$ hors de K ; soit alors $\varphi \in \mathcal{K}(E)$ avec $0 \leqslant \varphi \leqslant 1$ et $\varphi = 1$ sur K.

On a $f = f\varphi + f(1 - \varphi)$; or, $f(1-\varphi) \leqslant \frac{g(1-\varphi)}{k} \leqslant \frac{g}{k}$, donc $U(f(1-\varphi)) \leqslant \frac{1}{k} Ug$.

On peut choisir k assez grand pour que $U(g) \leqslant k\varepsilon$; la décomposition cherchée est alors fournie par $f_1 = f\varphi$ et $f_2 = f(1-\varphi)$.

La mesure μ sera unique lorsque, pour toute $\psi \subset \mathcal{K}(X)$, on aura

$$\sup_{f \leqslant \psi} T(f) = \inf_{\psi \leqslant g} T(g) \qquad (\text{où } f, g \in V) \quad .$$

Exemple. - Soit X un fermé de R^p et soit V l'espace vectoriel des restrictions à X des polynômes réels définis dans R^p. Cet espace V est évidemment adapté, donc la proposition 2 est applicable.

Cas particuliers.

1° Revenons au problème de Hamburger ; la proposition 2 montre que, avec les notations initiales, il aura une solution lorsque, pour tout polynôme $\sum a_n t^n$ partout positif sur R, on a $\sum a_n \alpha_n \geqslant 0$.

Nous allons transformer cette condition :

LEMME. - Tout polynôme $P \geqslant 0$ sur R est une somme de carrés de polynômes.

En effet, soit p le degré de P et soit C_p le cône des polynômes de degré $\leqslant p$, qui sont $\geqslant 0$ sur R.

Comme C_p est de dimension finie, tout élément de C_p est somme finie d'éléments extrémaux de C_p ; soit Q un tel élément extrémal.

Si $Q > 0$, on a $Q = k + Q'$ où $k > 0$ et $Q' \in C_+$; comme Q est proportionnel à k, Q est bien un carré. Si Q a des racines, chacune est paire, donc $Q = M^2 . Q'$, où $Q' > 0$; comme Q est extrémal, Q' l'est aussi, donc Q' est une constante, et Q est bien encore un carré.

En résumé, on a bien $P = \sum M_i^2$ où les M_i sont des polynômes.

Pour exprimer que la forme linéaire T sur $\mathcal{P}$ est positive, il suffit donc d'exprimer que $T(M^2) \geqslant 0$ pour tout polynôme M. Posons alors $M(t) = \sum b_n t^n$.

Les conditions $T(M^2) \geqslant 0$ s'écrivent donc

$$\sum \alpha_{i+j}\, b_i\, b_j \geqslant 0 \qquad \text{quels que soient les } b_n \quad .$$

Ceci revient encore à dire que la fonction $(i, j) \to \alpha_{i+j}$ définie sur $N \times N$ est de type positif, ce qui s'exprime encore par le fait que chacun des déterminants

$$\Delta_n = |\alpha_{i+j}|_{i,j=0}^{n} \text{ est } \geqslant 0 \quad .$$

2° Le cas de Stieltjes s'obtient en prenant $X = [0, \infty[$.

Ici le cône C_p des polynômes de degré $\leqslant p$ qui sont positifs sur X a des éléments extrémaux de la forme $M^2_{(t)}$ et $tM^2(t)$; la condition $T \geqslant 0$ s'exprime au moyen des déterminants Δ_n et Δ'_n :

$$\Delta_n = |\alpha_{i+j}|_0^n \geqslant 0 \qquad \text{et} \quad \Delta'_n = |\alpha_{i+j+1}|_0^n \geqslant 0 \quad .$$

2. Exemple de non-unicité pour les problèmes de Hamburger et Stieltjes.

Pour tout entier $n \in \mathbb{N}$, posons

$$I_n = \int_0^\infty t^n \exp(-t^{1/4}) \sin t^{1/4}\, dt \quad .$$

Posons $t = u^4$; l'intégrale devient

$$I_n = 4\mathfrak{I}\left(\int_0^\infty u^{4n+3} \exp u(i-1)\, du\right) \quad .$$

Sous cette forme on vérifie aisément (intégration par parties) que chacun des I_n est nul.

Donc, si ν désigne la mesure de densité $\exp(-t^{1/4}) \sin t^{1/4}$ sur R_+ et 0 sur R_- , les mesures positives ν_+ et ν_- ont mêmes moments de tout ordre, donc définissent la même forme linéaire positive T sur $\mathcal{P}$ bien qu'étant différentes.

3. Étude du cône des formes linéaires positives T sur un espace adapté V .

Soit V un sous-espace vectoriel adapté de $\mathcal{C}(X)$, où X est localement compact. Désignons par V^*_+ le cône convexe des formes linéaires sur V , qui sont $\geqslant 0$ sur V_+ . Ce cône est complet pour la topologie faible sur V^* ; et si V a une base algébrique dénombrable (par exemple, si V est un espace vectoriel de polynômes), il est métrisable.

Lorsque V^*_+ est métrisable, on sait que chacun de ses éléments est barycentre d'une mesure positive portée par ses génératrices extrémales.

PROPOSITION 3. - Pour tout sous-espace vectoriel adapté V de $\mathcal{C}(X)$, tout élément extrémal T de V^*_+ est représentable par une mesure $k_{(a)}\, \varepsilon_a$ sur X . L'ensemble A_T des points a ainsi associés à T est fermé dans X et les valeurs prises en deux points quelconques de A_T par les f de V sont proportionnelles.

<u>Toute mesure</u> $\mu \geqslant 0$ <u>sur</u> X <u>représentant</u> T <u>est portée par</u> A_T .

En effet, soit $\mu \geqslant 0$ une mesure représentant T . Pour toute mesure μ' telle que $0 \leqslant \mu' \leqslant \mu$, le fait que T soit extrémale entraîne que $\mu'(f) = kT(f)$ pour toute $f \in V$.

Pour tout $a \in S\mu$, support de μ , on a donc, par un passage à la limite évident $f(a) = \lambda(a)\ T(f)$ pour toute $f \in V$. Comme pour tout $a \in X$, il existe une $f \in V$ telle que $f(a) \neq 0$, on a $\lambda(a) \neq 0$; d'où $T = k(a)\ \varepsilon_a$ (où $(k(a) = 1/\lambda(a)$).

Le reste de l'énoncé est évident.

L'ensemble A_T se réduit à un seul point lorsque, quels que soient $x , y \in X$, avec $x \neq y$, $f(x)$, $f(y)$ ne sont pas proportionnels, autrement dit s'il existe une $f \in V$ qui s'annule en x mais non en y . C'est le cas, par exemple, si V est l'espace des restrictions des polynômes à un fermé X de R^p .

<u>Remarques</u>.

1° Lorsque V_+^* est métrisable, le théorème de représentation intégrale pour les cônes faiblement complets métrisables, montre que toute $T \in V_+^*$ est représentable par une mesure $\mu \geqslant 0$ sur X , portée par l'ensemble des points-frontière a de X (à savoir les points a tels que la mesure ε_a soit un élément extrémal de V_+^* ; cet ensemble est un G_δ).

2° Il est inexact, même lorsque X est compact, que toute mesure ε_a soit un élément extrémal de V_+^* (cf. l'espace V des fonctions continues sur le disque plan $x^2 + y^2 \leqslant 1$, et harmoniques à l'intérieur).

3° Lorsque V est l'espace des restrictions des polynômes à un fermé X de R^p , tout point $a \in X$ est un point-frontière. En effet, soit μ une mesure positive sur X telle que $f(a) = \mu(f)$ pour tout polynôme f ; il existe un polynôme $f \geqslant 0$ dans R^p , avec $f(a) = 0$ et $f(x) > 0$ pour $x \neq a$. La relation $\mu(f) = 0$ montre alors que μ est portée par le point a , d'où le résultat.

4. Étude des mesures associées à une $T \geqslant 0$.

PROPOSITION 4. - <u>Soit</u> V <u>un sous-espace vectoriel</u> adapté <u>de</u> $\mathcal{C}(X)$. <u>Pour toute</u> $T \in V_+^*$, <u>l'ensemble</u> B_T <u>des</u> $\mu \geqslant 0$ <u>qui représentent</u> T <u>est un</u> convexe compact <u>de</u> $\mathfrak{M}_+(X)$.

Soit, en effet, $\mathcal{U}$ un ultrafiltre sur B_T. Par hypothèse, on a

$$T(f) = \int f \, d\mu \qquad \text{pour toute } \mu \in B_T \text{ et toute } f \in V_+ \quad .$$

Pour toute $\varphi \in \mathcal{K}_+(X)$, il existe $f \in V_+$ telle que $0 \leq \varphi \leq f$, ce qui entraîne $0 \leq \mu(\varphi) \leq \mu(f) = T(f)$.

Donc $\lim_{\mathcal{U}} \mu(\varphi)$ existe ; désignons-la par ν ; c'est une forme linéaire positive sur $\mathcal{K}(X)$, donc une mesure ≥ 0, et $\nu = \lim_{\mathcal{U}} \mu$.

On va montrer que $\nu \in B_T$, ce qui démontrera bien la compacité de B_T.

Soit $f \in V_+$, et soit $g \in V_+$, où g dépasse f.

Pour tout scalaire $k > 0$, il existe un compact $K \subset X$ hors duquel $g \geq kf$. Il existe une $\varphi \in \mathcal{K}(X)$ avec $0 \leq \varphi \leq 1$ et $\varphi = 1$ sur K.

Pour $\mu \in B_T$, on a $\mu(g) = T(g)$, d'où $\mu(g(1 - \varphi)) \leq T(g)$, d'où $\mu(f(1 - \varphi)) \leq \frac{T(g)}{k}$, puis $\mu(f\varphi) \geq \mu(f) - \frac{T(g)}{k} = T(f) - \frac{T(g)}{k}$. Comme $f\varphi \in \mathcal{K}(E)$, on a $\lim_{\mathcal{U}} \mu(f\varphi) = \nu(f\varphi)$. Donc $\nu(f) \geq \nu(f\varphi) \geq T(f) - \frac{T(g)}{k}$.

Comme k est arbitraire, $\nu(f) \geq T(f)$.

D'autre part, par semi-continuité, $\nu(f) \leq \lim_{\mathcal{U}} \mu(f) = T(f)$. Donc $\nu(f) = T(f)$, autrement dit $\nu \in B_T$.

Enfin la convexité de B_T est évidente.

Application. - On peut désormais utiliser les éléments extrémaux de B_T. Dans le cas où V est l'espace des polynômes sur un fermé X de R^p, il serait intéressant d'étudier ces éléments extrémaux et d'étudier leur rapport avec les mesures appelées aussi extrémales dans SHOHAT et TAMARKIN (p. 60).

5. Caractérisation des cas d'unicité.

V désignant toujours un sous-espace vectoriel adapté de $\mathcal{C}(X)$, et T une forme linéaire positive sur V, nous savons que dire que B_T est réduit à un seul élément, équivaut à dire que, pour toute $\psi \in \mathcal{K}(X)$, on a

$$\sup_{f \leq \psi} T(f) = \inf_{\psi \leq g} T(g) \qquad (\text{où } f, g \in V) \quad ,$$

ou encore que, pour tout $\varepsilon > 0$, on peut encadrer ψ par deux éléments f, g de V tels que $T(g - f) < \varepsilon$.

Soit μ un élément quelconque de B_T. L'espace V est évidemment un sous-espace vectoriel de L^1_μ. Le critère d'unicité qu'on vient de rappeler montre que, si $B_T = \{\mu\}$, V est partout dense dans L^1_μ. Il serait intéressant de déterminer, plus généralement, dans quels cas V est partout dense dans L^1_μ ; par exemple on peut voir aisément que ceci ne peut avoir lieu que si μ est un élément extrémal de B_T ; dans quels cas cette condition suffit-elle ?

Des questions analogues se posent pour les espaces L^p_μ (dans le cas où $V \subset L^p_\mu$).