1 Points extrêmaux

Définition 1. *Soit F une partie d'un espace vectoriel et X une partie de E. Un **sous-ensemble** S de X est dit **extrêmal dans** X si, étant donnés $x, y \in X$ et $t \in]0,1[$,*

$$tx + (1-t)y \in S \Rightarrow x \in S \text{ et } y \in S.$$

*Si la partie S est réduite à un point, ce **point** est dit **extrêmal dans** X.*

S'il n'y a pas d'ambiguité possible sur l'ensemble X, on ne dira pas "extrêmal dans X" mais seulement "extrêmal". On note $\mathcal{E}(X)$ l'ensemble des points extrêmaux de X.

Exercice : montrez que dans un convexe K un point $x \in K$ est extrêmal si, et seulement si, $K \backslash \{x\}$ est convexe.

Questions – ***(1)*** Un ensemble de points extrêmaux est-il extrêmal ?

(2) Quel est l'ensemble des points extrêmaux du convexe formé par les probabilités boréliennes sur un compact ?

Proposition 1. *L'ensemble des points extrêmaux d'un convexe compact de $\mathbb{R}^n$ est non vide* [1].

$\lhd$ On le montre par récurrence sur la dimension de l'espace affine $< F >$ engendré par F. La chose est claire en dimension 1 où les convexes compacts étant de la forme $[a,b]$, ils possèdent a et b pour points extrêmaux [2]. Supposons le résultat acquis lorsque dim $< F > \leq n$ et donnons-nous un convexe compact C engendrant un espace de dimension $n+1$, un point x de ∂C, ainsi qu'un hyperplan d'appui H à C en x, défini comme le lieu des points annulant une forme affine f telle que $C \subset \{f \leq 0\}$. D'après l'hypothèse de récurrence, $C \cap H$ possède un point extrêmal y ; ce point de C est aussi extrêmal dans C. En effet, si tel n'était pas le cas, y serait contenu dans l'intérieur d'un segment $[a,b]$ [3] inclus dans C ; $f(y)$ étant nul et $f(a) \leq 0$, $f(b) \leq 0$, $f(a)$ et $f(b)$ devraient être nuls, *i.e.* a et b appartiendraient à H ; cela interdirait à y d'être un point extrêmal de $C \cup H$. Ainsi,

$$\mathcal{E}(C \cap H) \subset \mathcal{E}(C) \cap H, \tag{1.0.1}$$

ce qui permet de conclure la récurrence. $\rhd$

Les choses sont plus compliquées en dimension infinie. La boule unité de $L^\infty(\mathbb{R})$ est un exemple de convexe fermé borné sans points extrêmaux : aucun des points de la boule ouverte ne peut être extrêmal, et pour ceux de la sphère unité, si f désigne l'un d'entre eux, le fait qu'il existe un borélien B_f de mesure (de Lebesgue) non nulle sur lequel f est presque sûrement non nulle permettant d'écrire f comme barycentre non trivial de points de la sphère unité :

$$f = \mu(B_f)\frac{f_{|B_f}}{\mu(B_f)} + \mu({B_f}^c)\frac{f_{|{B_f}^c}}{{B_f}^c},$$

clôt l'exemple [4].

L'inclusion 1.0.1 est l'outil principal de la démonstration du

[1] n est un entier non nul quelconque.

[2] Ces points ne sont pas supposés distincts.

[3] non trivial.

[4] Cette formule répond presque à la question ***(2)***...

Théorème 1. – ***Krein-Millman*** – *Soit C un convexe compact de $\mathbb{R}^n$. C est l'enveloppe convexe de ses points extrêmaux.*

$\triangleleft$ Là encore, on raisonne par récurrence sur la dimension de l'espace affine engendré par C. Le résultat est vrai en dimension 1 ; supposons-le vrai en dimension n et donnons-nous un convexe compact C engendrant un espace affine de dimension $n+1$. Etant donné que C est l'enveloppe convexe de sa frontière, il suffit de voir que celle-ci est dans l'enveloppe convexe $\mathcal{C}onv(\mathcal{E}(C))$ [5] de l'ensemble des points extrêmaux de C. Or si x est un point du bord de C et H un hyperplan d'appui à C en x, x est, d'après l'hypothèse de récurrence, dans $\mathcal{C}onv(\mathcal{E}(C \cap H))$ et partant dans $\mathcal{C}onv(C)$, puisque l'inclusion 1.0.1 a lieu. $\triangleright$

Un contre-exemple – Etant donné un espace topologique normal $\mathcal{N}$, regardons l'ensemble des points extrêmaux de la boule unité fermée B de l'ensemble des fonctions continues sur $\mathcal{N}$, muni de la norme uniforme. Les deux fonctions constantes égales à ± 1 sont clairement extrêmales. Si $|f|$ est < 1 en un point, le *lemme d'Uryshon* nous donne l'existence d'une fonction continue g telle que $||f \pm g|| < 1$, ce qui permet d'écrire f comme un barycentre de points de B : $f = \frac{1}{2}((f+g)+(f-g))$; f n'est donc pas extrêmal, et $\mathcal{E}(B) = \{\pm\mathbf{1}\}$; cependant, $\mathcal{C}(\{\pm\mathbf{1}\}) \subsetneq B$.

Questions – ***(1)*** L'ensemble des points extrêmaux d'un convexe compact de $\mathbb{R}^n$ est-il fermé, ouvert, de nature indéterminée ?

(2) En dimension finie, le *théorème de Carathéodory* nous dit entre autres que l'enveloppe convexe d'un compact est compacte, en est-il de même en dimension infinie ?

Commençons par répondre à la seconde question : Non ! Si $\{e_n\}_{n\geq 0}$ désigne la base canonique de l^2, $K := \{\frac{e_n}{n+1}\}_{n\geq 0} \cup \{0\}$ est compact ; tout point de $\mathcal{C}onv(K)$ aura ses coordonnées nulles à partir d'un certain rang ; on voit cependant que $\frac{6}{\pi^2}\sum_{n=0}^{+\infty}\frac{1}{(n+1)^2}\frac{e_n}{n+1}$ est dans $\overline{\mathcal{C}(K)}$: $\mathcal{C}(K)$ n'est même pas fermée.

Exercice : Montrez qu'on a tout de même le résultat suivant : "*Dans un espace de Banach, l'enveloppe convexe fermée d'un compact est compacte*" [6].

Apportons maintenant une réponse à la première question.

Si l'objet de la figure 1 permet de voir que l'ensemble des points extrêmaux d'un convexe compact de dimension finie, n'est pas toujours fermé ou ouvert dès que la dimension est ≥ 3, il est surprenant d'avoir le fait suivant.

Proposition 2. *En dimension 2, l'ensemble des points extrêmaux d'un convexe compact est toujours fermé, a fortiori compact.*

$\triangleleft$ Soit C un convexe compact du plan, d'intérieur non vide [7] ; on va voir que tout point x de C qui n'est pas extrêmal possède *dans* C un voisinage formé de points non extrêmaux.

L'affirmation étant claire lorsque $x \in \overset{\circ}{C}$, supposons x sur le bord de C. L'hypothèse sur x s'écrit : $x \in]a,b[$, où a et b sont deux points de ∂C [8]. Soit q un point $\overset{\circ}{C}$; l'intérieur du triangle de

[5] On notera toujours $\mathcal{C}onv(E)$ l'enveloppe convexe d'une partie E d'un $\mathbb{R}$-espace vectoriel - non nécessairement de dimension finie.

[6] Ouf !... C'est rassurant.

[7] Le cas dégénéré est immédiat.

[8] Que a et b soient dans ∂C et non dans $\overset{\circ}{C}$ résulte du fait que si $p \in \partial C$ et $q \in \overset{\circ}{C}$, $]p,q]$ est tout entier inclus dans $\overset{\circ}{C}$.

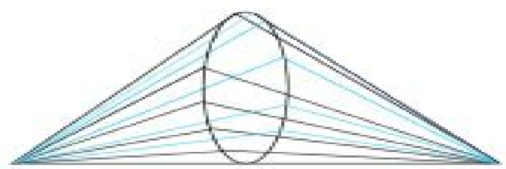

FIG. 1 – Un convexe compact particulier

sommets a, b, q, étant inclus dans $\overset{\circ}{C}$, est constitué des points non extrêmaux, de même, les points du segment ouvert $]a, b[$ ne sont pas extrêmaux ; la réunion de ceux-ci et des précédents forme un voisinage ouvert de x dans C formé de points non extrêmaux, affirmation qui tombe en défaut lorsque l'on est en dimension supérieure. $\triangleright$

Pour compléter la réponse à la seconde question, démontrons le

Théorème 2. – ***Carathéodory*** – *Soit X une partie de $\mathbb{R}^n$. Tout point de $\mathcal{C}onv(X)$ peut s'exprimer comme le barycentre d'exactement $n+1$ points de X :*

$$\mathcal{C}onv(X) = \bigcup \mathcal{C}onv(\{x_1, x_2, ..., x_{n+1}\}),$$

la réunion se faisant sur l'ensemble des $n+1$-uplets de points de X.

$\triangleleft$ Soit x un point de $\mathcal{C}onv(X)$, barycentre de $r > n+1$ points x_1,...,x_r, de X. On va voir que l'on peut également écrire x comme un barycentre de $r-1$ points de X.

Cela résulte directement du fait que la famille x_1,...,x_r étant affinement liée, on peut trouver des réels s_1,...,s_r, *de somme nulle*, tels que $\sum_{k=1}^{r} s_k x_k = 0$. On a donc pour tout réel α, $x = \sum_{k=1}^{r} (p_k + \alpha s_k) x_k$; il n'y a plus qu'à prendre un α pour lequel chacun des poids $p_k + \alpha s_k$ est ≥ 0 et qui annule l'un d'entre eux pour voir x écrit comme le barycentre de $r-1$ points de X. $\triangleright$

Le *théorème de Krein-Milman* demeure vrai sous des hypothèses plus générales, sans considérations sur les dimensions.

Théorème 3. – ***Krein-Millman*** – *Soit E un espace vectoriel*[9] *topologique dont le dual topologique E^* sépare les points. Tout convexe compact non vide $C \subset E$ est l'adhérence de l'enveloppe convexe de ses points extrêmaux : $C = \overline{\mathcal{C}onv(\mathcal{E}(C))}$; en particulier, l'ensemble de ses points extrêmaux est non vide.*

$\triangleleft$ On va voir que *dans tout sous-ensemble de C, extrêmal et compact, se trouve un point extrêmal* ; on va pour cela chercher "le plus petit des sous-ensembles de ce compact extrêmal qui ait encore cette propriété". A cette fin, on notera $\mathcal{P}$ l'ensemble non vide des parties compactes, extrêmales, de C – il contient C lui-même – et, on associera à tout élément $S \in \mathcal{P}$ et toute forme linéaire Λ, continue sur E, l'ensemble S_Λ des points de S où la partie réelle de Λ atteint son maximum ; c'est encore un élément de $\mathcal{P}$.

[9] a priori complexe.

Soit donc $S \in \mathcal{P}$. Si I désigne l'intersection non vide de tous les éléments d'une famille *maximale* décroissante de parties compactes de S [10] qui sont extrêmales dans C, I est lui-même compact et extrêmal dans C, et, en vertu de la maximalité de la famille considérée, aucune sous-partie compacte de I n'est extrêmale dans C. Ainsi, $I_\Lambda = I$, quelle que soit la forme linéaire continue Λ sur E ; l'hypothèse de séparation est là pour permettre la conclusion : I est un singleton ; c'est "le" point de $\mathcal{E}(C) \cap S$ que l'on cherchait.

Revenons à nos moutons... C étant fermé et *convexe*, on sait a priori que $\overline{\mathcal{C}onv(\mathcal{E}(C))} \subset C$. Si un point x_0 de C n'était pas dans $\overline{\mathcal{C}onv(\mathcal{E}(C))}$, le *théorème de Hahn-Banach*, pour lequel l'hypothèse de séparation des points est nécessaire, nous fournirait une forme linéaire continue Λ telle que $\Lambda(x_0) > \max_{\overline{\mathcal{C}onv(\mathcal{E}(C))}} \Lambda$. L'ensemble extrêmal C_Λ devrait alors contenir un point extrêmal de C et être disjoint de $\overline{\mathcal{C}onv(\mathcal{E}(C))}$, ce qui est impossible. $\rhd$

Ce que montrent les deux premiers paragraphes de la démonstration, c'est à quelques détails près la propriété fondamentale suivante des fonctions convexes.

Théorème 4. *Toute fonction convexe réelle, semi-continue supérieurement, définie sur un convexe compact, atteint son maximum en un point extrêmal.*

Corollaire 1. *Tout hyperplan d'appui à un convexe compact contient un point extrêmal.*

Comme on va le voir, tout l'interet d'un tel théorème réside dans le fait qu'il nous assure l'*existence* de certains objets... à nous d'en tirer partie. Cependant, ce théorème serait miteux si les espaces dont le dual sépare les points étaient des curiosités de la nature. Le théorème suivant nous donne une grande famille d'espaces topologiques dont le dual sépare les points.

Théorème 5. *Le dual topologique de tout espace vectoriel topologique localement convexe sépare les points.*

C'est le cas, par exemple, des espaces munis de la topologie faible ou de la topologie faible-*, si l'on a affaire à un espace de Banach qui est le dual d'un autre espace de Banach.

$\lhd$ Il s'agit pour l'essentiel d'une application du *théorème de Hahn-Banach*.

Soit $x \neq 0$ un point d'un espace vectoriel topologique E, localement convexe. On va trouver une forme linéaire continue séparant 0 et x. Soit V un voisinage ouvert de 0, convexe et équilibré, ne contenant pas x ; on lui associe sa jauge J_V [11], qui est < 1 sur V. Le *théorème de Hahn-Banach* nous fournit alors une forme linéaire ℓ qui vaut 1 en x, 0 au point 0 et qui est inférieure à J_V : quel que soit $y \in E$, $\ell(y) \leq J_V(y)$.

Il nous reste à montrer que cette forme linéaire est continue. On montre pour cela que les demi-espaces $\{y \in E; \ell(y) < c\}$, $c \in \mathbb{R}$, sont ouverts, ainsi que les demi-espaces $\{y \in E; \ell(y) > c\}$, $c \in \mathbb{R}$. Les deux démonstrations se faisant de la même manière, regardons $\{y \in E; \ell(y) < c\}$. Pour un tel y, et $v \in V$, $r > 0$, on a :

$$\ell(y + rv) \leq \ell(y) + rJ_V(v) < c + r,$$

puisque $J_V(v) < 1$. Si donc l'on prend $r = c - \ell(y)$, l'inégalité précédente nous dit que l'ouvert $y + rV$ est tout entier inclus dans $\{z \in E; \ell(z) < c\}$. $\rhd$

Question : où sert la locale convexité de l'espace ?

[10] Par Zorn et Toutatis !!!

[11] Les propriétés de V en font bien une fonction finie positivement homogène, sous-additive.

Une application : existence d'une probabilité invariante ergodique

Etant donné un espace *métrique compact* X et une application continue $f : X \to X$, notons $\mathcal{M}(f)$ le convexe non vide [12] formé des probabilités boréliennes f-invariantes. Supposons qu'une telle mesure μ ne soit pas ergodique, nous donnant ainsi l'existence d'un borélien f-invariant A de μ-mesure non triviale [13] ; les restrictions renormalisées de μ à A et A^c, $\mu_{A^{(c)}}(.) := \frac{\mu(. \cap A^{(c)})}{\mu(A^{(c)})}$, sont alors deux probabilités (boréliennes) f-invariantes dont μ est un barycentre non trivial : $\mu = \mu(A)\mu_{|A} + \mu(A^c)\mu_{|A^c}$. Les points extrêmaux de $\mathcal{M}(f)$, s'il y en a, sont donc des probabilités boréliennes ergodiques. Si l'on veut s'assurer leur existence en appliquant le *théorème de Krein-Milman*, il nous faut au préalable vérifier que le dual topologique de l'ensemble des mesures positives sur X sépare les points... ce qui se fait sans difficulté [14].

On peut aussi obtenir directement des points extrêmaux en utilisant une méthode analogue à celle de la démonstration du théorème et qui a l'avantage d'être visuelle : une façon de trouver des points extrêmaux d'un convexe compact du plan est de commencer par regarder les points du compact de plus grande abscisse, puis, parmis ceux-ci, ceux de plus grande ordonnée... On généralise comme suit cette idée.

On va pour ce faire tirer avantage de la séparabilité de $\mathcal{C}(X)$ [15] en associant à une suite dense $\{\varphi_n\}_{n\geq 0}$ de $\mathcal{C}(X)$ la suite décroissante $\{\mathcal{M}_n\}_{n\geq 0}$ de sous-ensembles de $\mathcal{M}(f) =: \mathcal{M}_0$ définie par récurrence par la formule :

$$\mathcal{M}_{n+1} := \{\mu \in \mathcal{M}_n ; \int \varphi_{n+1} d\mu = \sup_{\nu \in \mathcal{M}_n} \int \varphi_{n+1} d\nu\}^{16}.$$

Les formes linéaires $\mu \mapsto \int \varphi_k d\mu$ étant continues [17], les $\mathcal{M}_n$ sont des convexes compacts dont l'intersection $\mathcal{I}$ est non vide et convexe ; on va voir que tous ses points sont extrêmaux. Supposons qu'un point μ de $\mathcal{I}$ soit barycentre de points de $\mathcal{M}(f)$: $\mu = t\nu + (1-t)\tilde{\nu}$, $t \in [0,1]$; on voit par récurrence que ν et $\tilde{\nu}$ sont dans tous les $\mathcal{M}_n$, ce qui signifie qu'on a pour pour chaque élément φ_n de la suite dense :

$$\int \varphi_n d\mu = \int \varphi_n d\nu = \int \varphi_n d\tilde{\nu};$$

ces égalités sont donc vraies pour toute fonction continue, imposant à ν et $\tilde{\nu}$ d'être égales à μ, ce qui montre que μ est extrêmal.

Ainsi, toute mesure extrêmale est ergodique. Ces deux ensembles de mesures sont en fait égaux. Pour le voir, supposons donnée une mesure ergodique μ, barycentre non trivial de deux mesures distinctes ζ et ϑ : $\mu = t\zeta + (1-t)\vartheta$, $t \in]0,1[$. ζ étant absolument continue par rapport à μ a une densité f ; cette fonction non constante (μ-presque partout), puisque ϑ est non nulle, est invariante par f ; son existence contredit l'ergodicité de μ [18].

[12] On démontre ce fait en appendice.

[13] *i.e.* $\mu(A) \in]0,1[$.

[14] Exercice !

[15] Qui n'est plus assurée si le compact X n'est pas métrisable.

[16] On reconnait ici l'analogue des ensembles décrits dans le plan.

[17] Par définition de la topologie faible-*.

[18] Voir l'Appendice.

Un autre exemple : un théorème de Liapounov

Rappelons qu'une ***mesure positive***, μ, définie sur une tribu $\mathcal{A}$, est dite ***sans atome*** si tout ensemble mesurable E, de mesure > 0, admet un sous-ensemble mesurable, de mesure non nulle strictement inférieure à celle de E. On dira d'une ***mesure signée***, μ, qu'elle est ***sans atome*** si sa variation totale, $|\mu|$, est sans atome.

Théorème 6. – ***Liapounov*** – *Soient μ_1,...,μ_n, des mesures signées, sans atome, définies sur l'espace mesurable $(\Omega, \mathcal{A})$. La mesure $\mu := (\mu_1, ..., \mu_n)$, à valeurs dans $\mathbb{R}^n$, a une image convexe et compacte.*

$\lhd$ Associons à μ_1,...,μ_n, la mesure positive $\sigma := |\mu_1| + ... + |\mu_n|$. C'est une mesure *sans atome*, par rapport à laquelle chacune des mesures μ_i est absolument continue ; on note $h_i := \frac{d\mu_i}{d\sigma}$. Notons aussi Λ l'application linéaire $g \in L^\infty(\sigma) \mapsto (\mu_1(g), ..., \mu_n(g)) \in \mathbb{R}^n$ [19]. Λg s'écrit aussi $(\int gh_1 d\sigma, ..., \int gh_n d\sigma)$; sous cette forme, Λ apparait comme une application continue pour la topologie faible-* qu'il y a naturellement sur $L^\infty(\sigma) = (L^1(\sigma))^*$.

Si K désigne le convexe $\{g \in L^\infty(\sigma); 0 \leq g \leq 1\}$, le fait que g appartient à K si, et seulement si, on a pour toute fonction $f \in L^1(\sigma)$, $0 \leq \int fg d\sigma \leq \int f d\sigma$, nous montre que K est fermé (pour la topologie faible-*) ; comme c'est un sous-ensemble de la boule unité de $L^\infty(\sigma)$, qui est compacte pour la topologie faible-* (c'est exactement ce qu'affirme le *théorème de Banach-Alaoglu* [20]), K est (*-faiblement) compact. Son image $\Lambda(K)$ est donc un convexe compact de $\mathbb{R}^n$.

On va voir que $\mu(\mathcal{A})$ [21] et $\Lambda(K)$ coincident. Seule l'inclusion $\Lambda(K) \subset \mu(\mathcal{A})$ n'est pas évidente. Pour l'obtenir, associons à chaque $p \in \Lambda(K)$ le compact [22] $K_p := K \cap \Lambda^{-1}(p)$ et montrons qu'une fonction de la forme $\mathbf{1}_A$, $A \in \mathcal{A}$, appartient à K_p. On va la trouver sous la forme d'un point extrêmal de K_p (dont le *théorème de Krein-Milman* nous assure l'existence).

On aura la conclusion souhaitée une fois qu'on aura vu qu'une fonction g_0 de K_p qui n'est pas une fonction caractéristique n'est pas un point extrêmal de K_p. Or, on peut associer à une telle fonction g_0 un ensemble mesurable E, de σ mesure > 0, et un réel $r > 0$ tels qu'on a sur $E : r < g_0 < 1 - r$. Au vu de ces inégalités, il nous suffit maintenant de trouver une fonction $g \in L^\infty(\sigma)$, à support dans E, $-r < g < r$, appartenant au noyau de Λ, pour pouvoir écrire g_0 comme barycentre de deux points de $K_p : g_0 = \frac{g_0 - g}{2} + \frac{g_0 + g}{2}$; cela nous montrera que g_0 n'est pas extrêmal.

La condition $-r < g < r$ est sans importance puisqu'elle peut être obtenue en renormalisant g ; tout le travail consiste donc à trouver un élément du sous-espace Y de $L^\infty(\sigma)$ formé des fonctions à support dans E qui soit dans le noyau de Λ. Mais σ n'ayant pas d'atome et $\sigma(E) > 0$, Y est de dimension infinie, de sorte que le noyau de $\Lambda_{|Y}$ n'est pas restreint à $\{0\}$; un point bien choisi dans ce noyau fera l'affaire. $\rhd$

---

[19] On écrit $\mu_k(g)$ pour $\int g d\mu_k$.

[20] Rappelons-en un énoncé : *la boule unité fermée $\overline{B}$ du dual d'un espace de Banach X est compacte pour la topologie faible-*.*

Pour le démontrer, associons à chaque point u de $\overline{B}$ la famille le point $\{u(x)\}_{x \in X}$ du compact $\prod_{x \in X} [-|x|, |x|]$ (muni de la topologie produit). Cette application Ψ est une injection ; mieux encore... la topologie induite sur $\Psi(\overline{B})$ est la même que la topologie faible-* ; $\overline{B}$ et $\Psi(\overline{B})$ sont donc homéomorphes, de sorte qu'il nous suffit de montrer que $\Psi(\overline{B})$ est fermé (et partant compact) pour obtenir la conclusion souhaitée. Soit donc p un point adhérent à $\Psi(\overline{B})$; on définit une forme bornée en posant $u(x) := p_x$; pour voir qu'elle est linéaire, il suffit de remarquer qu'à x, y, $\in X$, et $r \in \mathbb{R}$, fixés, les relations

$$q_{x+y} = q_x + q_y, q_{rx} = rq_x$$

valables pour tout $q \in \Psi(\overline{B})$, ne concernant qu'un nombre fini de coordonnées, le passage à la limite est licite.

[21] qui est égal à $\Lambda(\{\mathbf{1}_A; A \in \mathcal{A}\})$

[22] compact pour la topologie faible-*, pour laquelle Λ est continue.

Voici une application amusante de ce théorème au *problème du partage du gâteau.* La situation est la suivante : deux amis se sont acheté un merveilleux gâteau : chantilly, fruits, napage, glaçage, chocolat... Peuvent-ils partager le gâteau en deux parts de façon que chacun estime avoir au moins la moitié du trésor, selon ses critères ? En d'autres termes, deux mesures μ_1 et μ_2 sur le gâteau étant données, est-il possible de partitionner le gâteau en deux sous-ensembles (mesurables [23]) G_1 et G_2 tels que $\mu_1(G_1) \geq \frac{1}{2}$ et $\mu_2(G_2) \geq \frac{1}{2}$?

Le procédé "l'un coupe, l'autre choisit" répond par l'affirmative à la question ; cependant, cette façon de faire peut ne pas être satisfaisante en pratique... Supposons par exemple que chacun connait la façon d'apprécier les parts de l'autre et qu'après découpe par le premier en deux parts qu'il estime égales, le second pense que l'une d'entre elles vaut les deux tiers du gâteau. Le premier saura que non seulement l'autre est satisfait mais qu'en plus il se réjouit sans vergogne de la découpe heureuse qu'il a sous les yeux ! Pour éviter ce déplaisir, il cherchera donc à avoir $\mu_2(G_2) = \frac{1}{2}$... et puisqu'il n'y a pas de raison que le second soit moins égoiste que le premier, celui-là exigera qu'on ait aussi $\mu_1(G_1) = \frac{1}{2}$.

Ce différent mettra-t-il fin à leur amitié ?!

Heureusement, NON ! et ceci grâce à Liapounov ! car pourvu qu'on puisse "couper en huit" d'éventuelles pépittes de chocolat, les mesures μ_1 et μ_2 seront sans atome ; les points $(0, 0)$ et $(1, 1)$ étant dans l'image de (μ_1, μ_2), le point $(\frac{1}{2}, \frac{1}{2})$ y est donc aussi ! Seul soucis, et de taille : on ne connait pas de manière constructive d'obtenir des ensembles...

Une troisième application : le théorème de Kakutani

Ce théorème donne une condition pour qu'une famille d'applications affines envoyant un convexe compact dans lui-même ait un point fixe commun.

En vue de sa démonstration, on va d'abord donner une proposition utile pour identifier le lieu des points extrêmaux d'un compact (le corollaire 2 ci-dessous).

Le lemme suivant servira dans cette proposition ; il est cependant digne d'intérêt en soit.

Lemme 1. *Dans un espace vectoriel topologique, l'enveloppe convexe d'un nombre fini de convexes, compacts, est compacte.*

$\triangleleft$ Si en effet K_1,..., K_n, désignent des compacts d'un espace vectoriel E, $\mathcal{S}$ le simplexe de $\mathbb{R}^n$, $\{(s_1, ..., s_n) \in \mathbb{R}^n; \forall i, s_i \geq 0, \sum_{i=1}^n s_i = 1\}$ et $K \subset E^n$ le produit des K_i, l'application continue $f : \mathcal{S} \times K \to E$,

$$(s, x) \mapsto s_1 x_1 + ... + s_n x_n,$$

a une image $\Im m f$, compacte, contenue dans $\mathcal{C}onv(K_1 \cup ... \cup K_n)$; cette image est même convexe, $\alpha f(s, x) + \beta f(t, y)$ s'écrivant aussi $f(\alpha s + \beta t, c)$, où $c_i = \frac{\alpha s_i x_i + \beta t_i y_i}{\alpha s_i + \beta t_i}$ appartient à K_i [24]. Il reste à noter que les inclusions $K_i \subset \Im m f$ et la convexité de $\Im m f$ nous assurent l'inclusion $\mathcal{C}onv(K_1 \cup ... \cup K_n) \subset \Im m f$ et partant l'égalité de ces ensembles, ce qui montre la compacité de $\mathcal{C}onv(K_1 \cup ... \cup K_n)$. $\triangleright$

Pour la clarté de la démonstration, mettons en avant une notion dont on ne se servira qu'ici :

[23] qui prétend pouvoir délimiter à l'aide d'un couteau une partie non borélienne de $\mathbb{R}^3$?!

[24] L'hypothèse de convexité des K_i est là pour ça ! ***On ne peut pas s'en passer.***

Définition 2. *Soit L une partie de E. On appelle* ***tranche de*** *L toute intersection (non vide) de L avec un demi-espace ouvert. Si L est convexe, un point $a \in L$ sera dit* ***fortement extrêmal*** *si, et seulement si, les tranches de L forment une base de voisinages de a dans L.*

Insistons sur le fait qu'il doit s'agir d'une base de voisinages ; par exemple, le point a est dans la figure suivante extrêmal sans être fortement extrêmal, alors que le point b est extrêmal et fortement extrêmal.

********figure********

Proposition 3. *Soit L un convexe. Tout point fortement extrêmal est extrêmal.*

$\lhd$ Soit a un point de L fortement extrêmal. Il suffit de noter que puisque les tranches de L forment une *base* de voisinages de a, $L\backslash\{a\} = \bigcup\{H \cap L; H$ demi-espace fermé ne contenant pas $a\}$ est convexe. $\rhd$

Proposition 4. *La réciproque est vraie si L est compact : tout point extrêmal est fortement extrêmal.*

$\lhd$ Cela résulte du fait que sur un compact la topologie naturelle et la topologie $\sigma(E, E')$ [25] coincident [26]. De fait, si a est un point extrêmal de L et V un voisinage de a dans L, V contient une intersection finie de tranches $H_i^c \cap L$, H_i^c demi-espace *fermé*, et donc $L\backslash V \subset \bigcup H_i \cap L$. Le convexe $A := \mathcal{C}onv(\bigcup H_i \cap L)$, compact d'après le lemme 1, ne contient pas a. Un hyperplan séparant a et A nous fournit un tranche de L incluse dans V, contenant a. $\rhd$

Corollaire 2. *Dans tout compact K de la forme $\overline{\mathcal{C}onv(F)}$, où F est un fermé, on a $\mathcal{E}(K) \subset F$.*

$\lhd$ Soit a un point extrêmal de K qui n'est pas dans F ; puisque a est fortement extrêmal, on peut trouver un demi-espace ouvert H séparant a et F qui fait de la tranche $H \cap K$ un voisinage de a dans K. Le convexe fermé $H^c \cap K$ contient $\overline{\mathcal{C}onv(F)} = K$: une contradiction. $\rhd$

Théorème 7. – ***Kakutani*** – *Soit E un espace vectoriel topologique dont le dual topologique sépare les points, K un convexe compact de E et G un groupe équicontinu d'applications affines envoyant K dans K. Alors les éléments de G ont un point fixe commun.*

$\lhd$ Puisqu'il s'agit de trouver un *point* de K invariant par tous les $g \in G$, cherchons parmis les convexes compacts non vide, inclus dans K, stables par tous les g, "le" plus petit d'entre eux. Cet ensemble de parties de K étant partiellement ordonné par l'inclusion, le *théorème de Zorn* nous en fournit un élément minimal $\boldsymbol{m}$; on va montrer, en raisonnant par l'absurde, qu'il s'agit d'un point ; appelons donc x et y deux points distincts de $\boldsymbol{m}$.

Faisons d'abord deux remarques : ***(1)*** $\boldsymbol{m}$ étant minimal, on a pour tout $z \in \boldsymbol{m}$, $\overline{\mathcal{C}onv(G.z)} = \boldsymbol{m}$; c'est en particulier le cas pour $z_0 := \frac{x+y}{2}$.

(2) A quoi peut bien servir l'hypothèse d'équicontinuité ?... Puisque x et y sont "éloignés" l'un de l'autre, leurs images par les éléments de G sont *uniformément* loins les unes des autres : si W est un voisinage de 0 tel que $x - y \notin W$, il existe un voisinage V de 0 tel que quel que soit $g \in G$, $g.x - g.y \notin V$ [27] – si $g.(x - y)$ était dans V pour un certain $g \in G$, $g^{-1}(g.(x - y)) = x - y$ serait dans W ; que G soit un *groupe* importe donc !

[25] Sa trace sur L, qui est a priori plus grossière.

[26] Ce n'est rien d'autre que de dire que $x_i \to x$ dans K si et seulement si $\ell(x_i) \to \ell(x)$, quelle que soit $\ell \in E'$

[27] ce V est celui associé à W dans la définition de l'équicontinuité.

Soit p un point extrêmal de $\boldsymbol{m}$ [28]. Au vu de la première remarque et du corollaire 2, ce point extrêmal est dans $\overline{G.z_0}$; de ce fait, on peut trouver une suite $\{g_n\}_{n\geq 0}$ d'éléments de G telle que $g_n.z_0 \to p$, $\{g_n.x\}_{n\geq 0}$ et $\{g_n.y\}_{n\geq 0}$ convergent et ont pour limites respectives des points $\underline{x}$ et $\underline{y}$ de $\boldsymbol{m}$. L'égalité $g_n.z_0 = \frac{g_n.x+g_n.y}{2}$ implique alors l'identité $p = \frac{\underline{x}+\underline{y}}{2}$, forçant $\underline{x}$ et $\underline{y}$ à être égaux et partant $g_n.x - g_n.y$ à tendre vers 0, ce qui contredit la seconde remarque. $\triangleright$

Il découle de ce théorème le très beau résultat suivant.

Théorème 8. *Tout groupe compact admet une mesure de Haar.*

$\triangleleft$ Notons E l'espace vectoriel des formes linéaires sur $C(G,\mathbb{R})$, muni de la topologie faible-$*$, et $K \subset E$ l'ensemble des probabilités sur G [29], c'est un convexe compact pour la topologie faible-$*$. Associons aussi à chaque $g \in G$ l'application $\tau_g : f \in C(G,\mathbb{R}) \mapsto \{x \mapsto f(g.x)\} \in C(G,\mathbb{R})$. La mesure qu'on recherche n'est autre qu'un élément de K fixé par chacune des applications ${}^t\tau_g : \varphi \in E \mapsto \{f \mapsto \varphi(\tau_g f)\}$.

Il nous faut donc vérifier d'une part que E a un dual topologique qui sépare les points et d'autre part que le groupe formé par les ${}^t\tau_g$ est équicontinu. Que la première condition soit assurée provient de ce que la topologie faible-* fait de $\mathcal{C}(G,\mathbb{R})$ un espace localement convexe ; pour la seconde, c'est une conséquence de la compacité de G, deux fonctions f et $\tau_g f$ étant uniformément proches sur K (*compact*) pourvu que g soit petit. $\triangleright$

Remarque : Sur un groupe de Lie muni d'une structure riemannienne, la donnée d'un produit scalaire sur l'espace tangent à l'identité détermine, par transport, une métrique invariante par translation, à laquelle est naturellement associée une mesure invariante par translation. Si celle-ci est finie (c'est le cas si le groupe est compact), c'est un multiple de la mesure de Haar.

[28] le théorème de Krein-Millman nous en assure l'existence.

[29] Le *théorème de représentation de Riesz* permet d'identifier les formes linéaires sur $\mathcal{C}(G,\mathbb{R})$ et l'ensemble des mesures signées sur G, ce qu'on fait ici.

2 La théorie de Choquet, une théorie des représentations intégrales

Un speech sur les représentations intégrales.

En un sens intuitif, le barycentre des points $(m_i, \alpha_i)_{i=1..r}$ de $\mathcal{E}(K)$, $\alpha_i \geq 0$, $\sum_{i=1..r} \alpha_i = 1$, s'interpréte comme le barycentre de la probabilité $\sum_{i=1..r} \alpha_i \delta_{m_i}$. En ces termes, le *théorème de Krein-Milman* nous dit que tout point de K est limite de barycentres de telles probabilités; cependant, rien ne nous assure qu'un point quelconque de K soit lui-même le barycentre d'une probabilité à support dans $\mathcal{E}(K)$.

C'est l'objet de la présente partie que de déterminer dans quelle mesure c'est ou non le cas.

Il y a une manière simple d'obtenir une telle représentation lorsque $\mathcal{E}(K)$ est *compact* [30] :

On l'a vu lors de la démonstration du théorème de Krein-Milman, pour chaque forme linéaire ℓ, continue sur K, l'ensemble des points de K où ℓ réalise son maximum (resp. minimum) est extrêmal et contient (au moins) un point extrêmal. Un point x de K étant donné, on a donc

$$\min_{\mathcal{E}(K)} \ell \leq \ell(x) \leq \max_{\mathcal{E}(K)} \ell,$$

ce dont il découle que ℓ est déterminée sur K par sa restriction à $\mathcal{E}(K)$. On peut donc noter $\ell(x) = x(f)$, où f est la restriction de ℓ à $\mathcal{E}(K)$. Cela définit sur l'ensemble L des restrictions à $\mathcal{E}(K)$ de formes linéaires continues sur K une forme linéaire x qui vérifie l'encadrement

$$\min f \leq x(f) \leq \max f;$$

x est en particulier une forme positive.

Si l'on rajoute la constante $\mathbf{1}$ à L et que l'on pose $x(\mathbf{1}) = 1$, x se prolonge à tout l'espace des fonctions définies sur $\mathcal{E}(K)$ en une forme linéaire positive. Le *théorème de représentation de Riesz* [31] nous fournit alors une mesure positive telle que $\ell(x) = x(f) = \int_{\mathcal{E}(K)} l(e) dm(e) = \langle \ell, m \rangle$, quelle que soit $\ell \in E'$; c'est bien affirmer que m a x pour barycentre.

Malheureusement, $\mathcal{E}(K)$ n'ayant aucune raison d'être compact [32], cette méthode ne nous permet de représenter x que par une mesure à support dans $\overline{\mathcal{E}(K)}$... ensemble qui n'a aucune signification géométrique particulière autre que sa définition.

Puisqu'il nous faudra dire des choses comme "toute probabilité est limite de probabilités discrètes", ou "cette suite de probabilités sur K admet une sous-suite convergente", deux affirmations qui n'ont rien d'évident pour qui sort à peine du lit, rappelons ce qu'il faut connaître sur l'ensemble $\mathcal{M}(F)$ des mesures de Radon sur un espace topologique F, *localement compact.*

[30] Ne lisez pas tout de suite cette démonstration, ou allez dès à présent voir la définition du barycentre d'une mesure donnée en 1.2.

[31] Rappelons que ce théorème recquiert pour son application la locale compacité de $\mathcal{E}(K)$, hypotèse non vérifiée dans le cas où l'on ne suppose pas $\mathcal{E}(K)$ compact.

[32] Sa nature topologique n'est d'ailleurs pas claire du tout.

2.1 Topologie sur $\mathcal{M}(F)$

Notations : $\mathcal{M}(F)$, $\mathcal{M}^+(F)$, $\mathcal{K}(F)$, $\mathcal{K}^+(F)$, $\mathcal{P}(S)$, S sous partie de F.

Rappelons qu'une ***mesure de Radon*** est une forme linéaire définie sur $\mathcal{K}(F)$ à laquelle on impose la condition de continuité quivante : il existe pour tout compact K une constante M_K telle qu'on a $|\mu(f)| \leq M_K||f||_\infty$, pour toute fonction continue f à support dans K. Sur un *espace localement compact*, le *théorème de représentation de Riesz* identifie ces formes linéaires à l'ensemble des mesures boréliennes finies régulières ; on ne distinguera pas ces deux visions d'un même objet.

On met sur $\mathcal{M}(F)$ la topologie $\mathcal{K}(F)^*$, dont les voisinages élémentaires d'une mesure μ sont de la forme :

$$\{\nu \in \mathcal{M}(F); \forall i = 1..r, |\nu(f_i) - \mu(f_i)| < \varepsilon\},$$

où les $f_i \in \mathcal{K}(F)$, $\varepsilon > 0$ et $r \in \mathbb{N}^*$.

Cette topologie munit $\mathcal{M}(F)$ d'une propriété sympathique.

Proposition 5. *$\mathcal{M}(F)$ et $\mathcal{M}^+(F)$ sont complets.*

$\lhd$ Pour $\mathcal{M}(F)$, il suffit de traduire le fait que $\{\mu_n\}_{n\geq 0}$ est de Cauchy si et seulement si chaque $\{\mu_n(f)\}_{n\geq 0}$, $f \in \mathcal{K}(F)$, est de Cauchy. Quant à $\mathcal{M}^+(F)$, c'est un fermé de $\mathcal{M}(F)$ en tant qu'intersection des fermés $\{\mu \in \mathcal{M}(F); \mu(f) \geq 0\}$, $f \in \mathcal{K}^+(F)$. $\rhd$

Comme dans tout espace topologique, il est intéressant de déterminer quels sont ses compacts.

Proposition 6. *Soit X une partie de $\mathcal{M}(F)$.* ***(i)*** *$\overline{X}$ est compacte si, et seulement si,* ***(ii)*** *X est vaguement bornée, i.e. l'image $\overline{X}(\varphi) := \{\mu(\varphi); \mu \in \overline{X}\}$ de chaque évalutation $\mu \mapsto \mu(\varphi)$, est bornée. Cette condition équivaut à la suivante :* ***(iii)*** *X est fortement bornée, i.e. on peut associer à chaque compact K de F une constante C_K telle que qu'on a pour toute mesure $\mu \in X$ et toute fonction continue f, à support dans K, $|\mu(f)| \leq C_K||f||_\infty$.*

$\lhd$ L'équivalence entre ***(ii)*** et ***(iii)*** résulte de la proposition précédente et du *théorème de Banach-Steinhaus* puisqu'à φ fixée, on ne regarde les mesures qu'à travers leur trace sur le support *compact* de φ.

Par ailleurs, il est clair que ***(i)*** implique ***(ii)***, puisque l'image de $\overline{X}$ par l'évaluation (continue) $\mu \mapsto \mu(f)$ est compacte et partant bornée.

Il nous reste donc à montrer, par exemple, ***(iii)*** implique ***(i)***. Or, si $\mathcal{F}$ un ultrafiltre de $\overline{X}$, $\overline{X}$ étant vaguement borné, l'image de $\mathcal{F}$ par toute évaluation est un ultrafiltre d'ensembles bornés de $\mathbb{R}$, qui converge donc vers un nombre $\mu_0(f)$. On voit sans mal que μ_0 est une forme linéaire sur $\mathcal{K}(E)$ et l'estimée ***(iii)*** nous dit que $\mu_0 \in \mathcal{M}(F)$; in other words, $\mathcal{F}$ converge vers μ_0. $\rhd$

Le *théorème de Banach-Alaoglu* nous fournit sans travail nouveau toute une classe de compacts.

Proposition 7. *Si $0 < a < +\infty$, l'ensemble $\{\mu \in \mathcal{M}(F); ||\mu|| \leq a\}$ est compact dans $\mathcal{M}(F)$. Si F est de plus supposé compact, alors $\{\mu \in \mathcal{M}^+(F); ||\mu|| = a\}$ est compact ; c'est en particulier le cas des probabilités sur F.*

$\lhd$ Pour la seconde affirmation, il suffit de remarquer qu'on regarde dans $\mathcal{M}^+(F)$ l'image inverse de a par l'application continue $\mu \to \mu(1)$, la fonction 1 n'étant dans $\mathcal{K}(F)$ que parce que F est compact. $\rhd$

Attention ! $\{\mu \in \mathcal{M}(F); ||\mu|| = a\}$, quoique relativement compact, n'est pas fermé en général ; dans un espace non compact, on peut perdre de la masse à l'infini.

On aura aussi besoin d'un ***théorème d'approximation*** ; l'expression "mesure discrète" y désigne une combinaison linéaire de masses de Dirac.

Théorème 9. ***(i)*** *L'espace vectoriel engendré par les* $\{\delta_x\}_{x\in F}$ *est dense dans* $\mathcal{M}(F)$.

(ii) *Le cone engendré par l'ensemble des* $\{\delta_x\}$, $x \in F$ [33] *est dense dans* $\mathcal{M}^+(F)$.

(iii) *Soit* K *un compact et* μ *une mesure positive, à support contenu dans* K. *Il existe une suite* $\{\mu_n\}_{n\geq 0}$ *de mesures discrètes à support dans* K, *de même masse que* μ, *ayant* μ *pour limite.*

$\lhd$ ***(i)*** Notons V l'espace vectoriel des mesures de Radon à support fini : $V = \langle \delta_x; x \in F\rangle$, et $V^\circ := \{f \in \mathcal{K}(F); \nu(f) = 0, \forall \nu \in V\}$ son annulateur. Il nous faut voir que $\mu \in (V^\circ)^\perp = \overline{V}$. Mais supposer $f(a) = \langle f, \delta_a\rangle = 0$ quel que soit $a \in \text{supp } \mu$ impliquant $\mu(f) = 0$, le tour est joué.

Pour démontrer ***(ii)*** et ***(iii)***, on va faire appel à un résultat géométrique dont la démonstration sera faite en annexe ; c'est une généralisation des considérations précédentes sur les ensembles "orthogonaux" Y°, $Z^\perp$.

Pour une partie $X \subset \mathcal{M}(F)$, on définit son ***ensemble polaire*** $B^\circ \subset \mathcal{K}(F)$ par $B^\circ := \{f \in \mathcal{K}(F); \mu(f) \geq -1, \text{ pour toute } \mu \in X\}$. De même, on définit l'ensemble polaire d'une partie D de $\mathcal{K}(F)$ par $D^\circ := \{\mu \in \mathcal{M}(F); \mu(f) \geq -1, \text{ pour toute } f \in D\}$. Alors *la bipolaire* $X^{\circ\circ}$ *de* X *est l'enveloppe convexe fermée de* $X \cup 0$.

(ii) : soit $X \in \mathcal{M}^+(F)$ le cone en question. D'un côté, le théorème sur les bipolaires nous dit que $X^{\circ\circ}$ et $\overline{X}$ sont égaux, d'un autre, on sait que $X^\circ = \{f \in \mathcal{K}(F); f \geq 0\}$ et $X^{\circ\circ} = \{\mu \in \mathcal{M}(F); \mu(f) \geq -1 \text{ pour toute } f \geq 0\ \} = \{\mu \in \mathcal{M}(F); \mu \geq 0\} = \mathcal{M}^+(F)$; la conclusion est là.

(iii) est immédiat une fois ***(ii)*** démontré. $\rhd$

Ces préliminaires sur la topologie des espaces de mesures étant terminés, on peut maintenant définir ce qu'est le barycentre d'une mesure.

2.2 Barycentre d'une mesure

Dans la suite, on se placera sur un espace localement convexe E, pour pouvoir disposer du *théorème de séparation de Hahn-Banach*. Toutes les mesures que l'on considèrera sont boréliennes, positives.

Notations : Pour un convexe compact K de E, on note $\mathcal{C}v(K)$ l'ensemble des fonctions convexes continues, définies sur K, à valeurs réelles, $-\mathcal{C}v(K)$ l'ensemble des fonctions concaves continues, et $\mathcal{A} = \mathcal{C}v(K) \cap -\mathcal{C}v(K)$, l'ensemble des fonctions affines continues sur K.

Le supremum de deux éléments de $\mathcal{C}v(K)$ est dans $\mathcal{C}v(K)$, $\mathcal{C}v(K)$ contient les constantes, et $\mathcal{C}v(K)$ sépare les points d'après le *théorème de Hahn-Banach* : $\mathcal{C}v(K) - \mathcal{C}v(K)$ est donc dense dans $\mathcal{C}(X)$, en vertu du *théorème de Stone-Weierstrass.*

S'il n'y a pas de danger de confusion, on notera $\mathcal{C}v$ plutôt que $\mathcal{C}v(K)$.

S'il est facile de définir le barycentre d'un nombre fini de points massiques, il est moins évident de définir celui d'une famille infinie de points, ou plus précisément d'une mesure signée (finie).

[33] *i.e.* l'ensemble des $\sum_{i=1}^n a_i\delta_{x_i}$, avec $a_i \geq 0$

Si l'on reformule la définition du barycentre r d'une famille finie de points $(m_i, \alpha_i)_{i=1..n}$, $\sum_{i=1}^n \alpha_i = 1$, à l'aide de formes linéaires :

$$x = r \Longleftrightarrow \forall \ell \in E', \ell(x) = \ell(\sum_{i=1}^{n} \alpha_i m_i) = \sum_{i=1}^{n} \alpha_i \ell(m_i)$$

et que l'on reconnaît dans $\ell(x)$, $\langle \ell, \delta_x \rangle$, et dans $\sum_{i=1}^n \alpha_i \ell(m_i)$, $\langle \ell, \sum_{i=1}^n \alpha_i \delta_{m_i} \rangle$, la définition suivante du barycentre d'une mesure ne peut paraître que naturelle.

Définition 3. *Soit K un compact de E et μ une mesure réelle sur K. Un point r de E est appelé* ***barycentre de*** *μ si, et seulement si, quelle que soit ℓ dans E', $\ell(r) = \langle \ell, \mu \rangle$. En d'autres termes, δ_r et μ coincident sur E'. S'il existe un unique barycentre, on le note $r(\mu)$.*

Remarques : ***(1)*** Le *théorème de Hahn-Banach* nous assure que s'il existe un barycentre celui-ci est unique.

(2) Une mesure *quelconque* sur un compact n'a pas forcément de barycentre...

Exercice : Trouver sur $\ell^2(\mathbb{N})$ une mesure à support compact [34] dont le "barycentre" ne peut qu'être dans $\ell^1(\mathbb{N}) \backslash \ell^2(\mathbb{N})$. Cette mesure est-elle finie ?

En dépit de ce contre-exemple, on a tout de même le *résultat* suivant, *fondamental pour toute la suite.*

Théorème 10. *Si μ est une probabilité sur le convexe compact K, son barycentre $r(\mu)$ existe ; c'est un point de K. De plus, l'application $\mu \in \mathcal{P}(K) \mapsto r(\mu) \in K$ est continue.*

$\lhd$ Approchons μ par une suite $\{\mu_n\}$ de probabilités discrètes, dont les barycentres sont dans K (qui est *convexe*). Puisque K est compact, on peut toujours [35] supposer que ces barycentres convergent vers un point r de K. On aura alors pour toute forme linéaire continue ℓ sur E

$$\ell(r) = \lim \ell(r(\mu_n)) = \lim \mu_n(\ell) = \mu(\ell).$$

La continuité de l'application r résulte de ce que K étant compact, sa topologie naturelle coincide avec la restriction à K de la topologie $\sigma(E, E')$, ce qui n'est rien d'autre que de dire que $a_i \to a$ dans K si et seulement si $\ell(a_i) \to \ell(a)$, quelle que soit $\ell \in E'$. De ce fait la convergence de μ_i vers μ entraîne celle de $r(\mu_i)$ vers $r(\mu)$. $\rhd$

(3) Revenons au théorème de Krein-Milman : "tout point x de K est la limite des barycentres $r(\mu_n) = \sum \lambda_i^{(n)} y_i^{(n)} \in Conv(\mathcal{E}(K))$ de probabilités $\mu_n = \sum \lambda_i^{(n)} \delta_{y_i^{(n)}}$ à support dans $\mathcal{E}(K)$". Si cela ne suffit pas pour obtenir de x qu'il soit lui aussi le barycentre d'une probabilité à support dans $\mathcal{E}(K)$, on peut toujours extraire de la suite $\{\mu_n\}$ une sous-suite convergente dont la limite μ vérifiera, d'après ce qu'on vient de voir, $x = r(\mu)$.

(4) On peut raffiner le théorème d'approximation 9.

Théorème 11. *Toute probabilité μ sur K est limite de probabilités discrètes ayant même barycentre que μ.*

[34] On rappelle que les compacts de $\ell^p(\mathbb{N})$, $p \geq 1$ sont les parties fermées, bornées, équisommables : étant donné $\varepsilon > 0$, il existe un entier n_ε tel qu'on a pour toute suite s du compact $\sum_{k \geq n_\varepsilon} |s_k|^p \leq \varepsilon$.

[35] Quitte à ne considérer qu'une sous-suite de $\{\mu_n\}$

$\triangleleft$ L'idée de la démonstration est simple et naturelle : on découpe μ en petits morceaux μ_k et on remplace chaque μ_k par une masse de Dirac située en $r(\mu_k)$.

Pratiquement, associons à un reconvrement fini $\mathcal{U} = \{U_k\}$ de K par des ouverts convexes une partition de l'unité$\{g_k\}$, et notons ν_k la trace de μ sur U_k : $\nu_k := \frac{g_k \mu}{\mu(g_k)}$, si $\mu(g_k) \neq 0$, 0 sinon ; notons aussi $x_k := r(\nu_k) \in U_k$ son barycentre, et

$$\mu_{\mathcal{U}} := \sum \mu(g_k)\delta_{x_k}.$$

On a bien $r(\mu_{\mathcal{U}}) = r(\mu)$:

$$\forall \ell \in E',\ \mu_{\mathcal{U}}(\ell) = \sum \mu(g_k)\ell(x_k) = \sum \mu(g_k)\nu_k(\ell) = \sum \mu(g_k \ell) = \mu(\ell).$$

La convergence de $\mu_{\mathcal{U}}$ vers μ lorsque les partitions s'affinent est une conséquence de l'uniforme continuité des fonctions continues sur K. $\triangleright$

Corollaire 3. *Un point x de K est extrêmal si, et seulement si, δ_x est la seule probabilité sur K ayant x pour barycentre.*

2.3 Diffusion d'une mesure

[36] Le problème est pour nous le suivant : étant donné un point x de K, on cherche une probabilité μ sur K portée par $\mathcal{E}(K)$, au sens où $\mu(\mathcal{E}(K)) = 1$, et de barycentre x.

Une méthode naturelle, suivie d'abord par Choquet, puis reprise sous forme plus élaborée par Loomis, est basée sur l'idée suivante : Soit Φ une application borélienne de K dans $\mathcal{P}(K)$ telle que pour tout $x \in K$, $\Phi(x) \in \mathcal{P}(K)$ soit une mesure discrète de barycentre x et portée par deux points x_1 et x_2, distincts de x si $x \neq \mathcal{E}(K)$. Pour toute $\mu \in \mathcal{P}(K)$ de barycentre x, la mesure $\Phi(\mu)$ a aussi x pour barycentre et on espère que $\Phi(\mu)$ est plus proche de $\mathcal{E}(K)$ que μ. Cette idée peut être précisée et fournit effectivement une solution au problème posé lorsque K est métrisable. Mais on devine qu'on aurait une preuve plus intuitive si l'on pouvait mesurer par un nombre ou une famille de nombre en quoi $\Phi(\mu)$ est meilleure que μ. C'est cette idée qui a été introduite par Bishop-de-Leeuw (1959), puis rendue plus commode par Choquet (1960) :

Définition 4. *Etant donné deux mesures sur K, μ et ν, on dit μ* ***plus diffuse que*** *ν* [37] *, ce qu'on écrit $\nu \prec \mu$, si, et seulement si, $\nu(f) \leq \mu(f)$ pour toute fonction continue f, convexe sur K.*

En quoi cette définition répond elle à notre attente ?

Si l'on se rappelle le théorème 1.6, l'utilisation de fonctions convexes comme fonctions tests est naturelle : idéalement, une fonction convexe atteint son maximum en des points extrêmaux ; dire que $\nu(f) \leq \mu(f)$, quelle que soit $f \in \mathcal{S}$, signifie donc que μ est plus concentrée que ν sur le voisinage des points extrêmaux de K. De ce fait, on s'attend à ce qu'une mesure complètement diffuse [38] soit à support dans $\mathcal{E}(K)$.

Puisque μ est plus grande que ν sur les fonctions *convexes*, et plus petite que ν sur les fonctions *concaves*, ces deux mesures coincident sur les fonctions affines, qui sont à la fois convexes et concaves. En prenant la fonction identiquement égale à 1 et la famille de toutes les formes linéaires, on en déduit que *deux mesures comparables ont même masse*

[36] Ce paragraphe est emprunté à G. Choquet ********

[37] ou ν moins diffuse que μ.

[38] Aucune mesure n'est plus diffuse qu'elle.

et même barycentre. Pour résumer en un slogan : *la diffusion d'une mesure ne change ni sa masse ni son barycentre.*

Si deux mesures quelconques n'ont aucune raison d'être comparables, on a toujours $\delta_{r(\mu)} \prec \mu$: si μ est discrète, l'inégalité $f(r(\mu)) \leq \mu(f)$ traduit la convexité même de f ; si μ n'est pas discrète, il suffit de l'approcher par de telles mesures, la continuité de r permettant un passage à la limite sans douleur.

Rappelons qu'on est la recherche d'une probabilité à support dans $\mathcal{E}(K)$ dont le barycentre soit un point x de K choisi d'avance. On l'a vu, ce qui pose problème est la condition supp $\mu \subset \mathcal{E}(K)$. Vu le sens de la définition de la relation $\prec$, et puisqu'on est assuré qu'une mesure plus diffuse que δ_x a x pour barycentre, on peut essayer de trouver μ en considérant une mesure maximale pour l'ordre $\prec$, parmis toutes les mesures plus diffuses que δ_x. Il nous faut pour cela vérifier que l'ordre $\prec$ est inductif sur $\mathcal{M}^+(K)$.

Ordre : seule l'antisymétrie n'est pas évidente, mais deux mesures telles que $\nu \prec \mu$ et $\mu \prec \nu$, coincidant sur $\mathcal{C}v$, coincident sur le treilli $\mathcal{C}v - \mathcal{C}v$ engendré par $\mathcal{C}v$, qui est dense dans $\mathcal{C}(K)$; elles sont donc égales.

Inductivité : soit $\{\mu_i\}_{i\in I}$ [39] une famille totalement ordonnée de mesures sur K ; oublié le fait que toutes ces mesures ont même barycentre x, une telle famille vérifie deux choses essentielles : ***(1)*** : toutes les μ_i ont même masse, et ***(2)*** : les $\mu_i(f)$ croissent, quelles que soit $f \in \mathcal{C}v$. En conséquence, ***(1)*** : la partie $\{\mu_i\}_{i\in I} \subset \mathcal{P}(K)$ est d'adhérence compacte, elle a donc une suite convergente [40], de limite une mesure μ, ce qui, d'après ***(2)*** entraîne $\mu_i(f) \to \mu(f)$. A nouveau, la densité du treilli $\mathcal{C}v - \mathcal{C}v$ dans $\mathcal{C}(K)$ suffit pour avoir $\mu_i(f) \to \mu(f)$, quelle que soit f dans $\mathcal{C}(K)$. Cette mesure μ vérifie $\mu_i \prec \mu$, quel que soit $i \in I$.

Conclusion : *Tout point de K est barycentre d'une mesure maximale.*

Puisque tout semble indiquer que c'est une mesure maximale qui sera solution de notre problème, on va chercher à comprendre à quoi ressemblent ces mesures.

2.4 Mesures maximales

Soit $f : K \to \mathbb{R}$ une fonction bornée. On pourrait appeler *majorante concave* et *minorante convexe* les fonctions $\widehat{f}$ et $\check{f}$ définies par les formules :

$$\widehat{f}(x) := \inf\{g(x); g \geq f, g \in -\mathcal{C}v\}$$

et

$$\check{f}(x) := \sup\{g(x); g \leq f, g \in \mathcal{C}v\};$$

c'est en quelques sortes ce qu'un opérateur ne mangeant que des fonctions concaves (resp. convexes) voit de f.

********Dessin********

$\widehat{f}$ jouit des propriétés suivantes [41] : elle est concave, semi-continue supérieurement, $f \leq \widehat{f}$; $f \mapsto \widehat{f}$ est continue, croissante, *sous-linéaire*, et f et $\widehat{f}$ ne sont égales que si f est déjà concave et semi-continue supérieurement [42] .

[39] I un ensemble ordonné.

[40] $\{\mu_{i_n}\}_{n\geq 0}$ avec $\{i_n\}$ croissant indéfiniment dans I : tout i est plus petit qu'un i_n.

[41] $\check{f}$ vérifie des propriétés analogues.

[42] On abrégera dorénavant "semi-continue supérieurement" en "*scs*" et "semi-continue inférieurement" en "*sci*"

$\widehat{f}$ et $\check{f}$ sont liés non seulement à f mais aussi à la géométrie de K.

Proposition 8. *Soit $f \in \mathcal{C}(K)$. On a pour tout $x \in K$*

$$\textbf{\textit{(i)}}\ \widehat{f}(x) = \sup\{\mu(f); \mu \in \mathcal{P}(K), r(\mu) = x\}$$

et

$$\textbf{\textit{(ii)}}\ \check{f}(x) = \inf\{\mu(f); \mu \in \mathcal{P}(K), r(\mu) = x\}$$

$\lhd$ La façon la plus simple de procéder est de s'apercevoir que l'énoncé exprime le fait que l'application continue $\varphi : \mu \in \mathcal{P}(K) \mapsto (r(\mu), \mu(f)) \in K \times \mathbb{R}$ a pour image l'intersection $\widehat{f}^- \cap \check{f}^+$ du sous-graphe (fermé) de $\widehat{f}$ (qui est scs) et du surgraphe fermé de $\check{f}$ (qui est sci).

Guidés par le dessin, on démontre cela en montrant qu'ils sont tous deux égaux à l'enveloppe convexe fermée du graphe Γ de f.

(1) $\Im m\varphi = \overline{\mathcal{C}onv(\Gamma)}$: Il n'y a qu'à constater que $\varphi(\sum \lambda_i \delta_{x_i}) = \sum \lambda_i (x_i, f(x_i)) \in \mathcal{C}onv(\Gamma)$ et laisser au théorème d'approximation 9 le soin de faire le reste.

(2) $\overline{\mathcal{C}onv(\Gamma)} = \widehat{f}^- \cap \check{f}^+$: L'inclusion $\overline{\mathcal{C}onv(\Gamma)} \subset \widehat{f}^- \cap \check{f}^+$ est évidente. S'il s'agissait d'une inclusion stricte, on pourrait séparer un point $(\zeta, \theta) \in \widehat{f}^- \cap \check{f}^+$ de $\overline{\mathcal{C}onv(\Gamma)}$ par une forme linéaire $L : (x, t) \mapsto g(x) + \alpha t$, $g \in E'$, continue sur $E \times \mathbb{R}$:

$$\exists m > 0, \forall (x,t) \in \overline{\mathcal{C}onv(\Gamma)}, L(x,t) < m < L(\zeta, \theta);$$

α ne peut être nul [43]

Si α était < 0, la majoration $g(x) + \alpha f(x) < m$, majoration sur K de f par la fonction concave $\frac{m-g(x)}{\alpha}$, forcerait $\widehat{f}$ à être $\leq \frac{m-g}{\alpha}$; on aurait en particulier $\widehat{f}(\zeta) \leq \frac{m-g(\zeta)}{\alpha} < \theta$, ce qui interdirait au point (ζ, θ) d'appartenir à $\widehat{f}^-$: cela contredit sa définition.

Dans l'hypothèse où $\alpha < 0$, on montre pareillement que (ζ, θ) ne peut appartenir à $\check{f}^+$. $\rhd$

Utilisée avec le corollaire 2, cette caractérisation de $\widehat{f}$ nous dit que

Corollaire 4. *f et $\widehat{f}$ coincident sur $\mathcal{E}(K)$.*

Remarque : Le théorème d'approximation 11 qui précède le corollaire 2 nous permet même de raffiner les caractérisations précédentes : *on peut se contenter de ne prendre pour μ que des mesures discrètes, de barycentre x.*

L'application $f \mapsto \widehat{f}$ induit sur $\mathcal{M}(K)$ une transformation naturelle $\mu \mapsto \widehat{\mu}$, où :

$$\widehat{\mu} : \mathcal{C}(K) \to \mathbb{R}, f \mapsto \mu(\widehat{f})^{44}$$

Mais ***attention*** ! L'opération $f \mapsto \widehat{f}$ n'étant *que sous-linéaire*, $\widehat{\mu}$ n'a aucune raison d'être mieux que sous-linéaire. La continuité de $f \mapsto \widehat{f}$ fait tout de même de $\widehat{\mu}$ une forme continue.

On est maintenant à même de donner une caractérisation des mesures maximales.

Théorème 12. *$\mu \in \mathcal{M}^+(K)$ est maximale si, et seulement si,* ***(i)*** *$\widehat{\mu}$ est linéaire, condition équivalente à la suivante :* ***(ii)*** *μ et $\widehat{\mu}$ coincident sur $\mathcal{C}(K)$.*

[43] Sans quoi on a $L(\zeta) < m < L(\zeta)$.

[44] $\widehat{f}$ étant bornée est partant mesurable.

Avant de lire la démonstration, nous vous conseillons d'aller lire en Appendice ce qu'on rappelle sur le prolongement des formes linéaires.

$\lhd$ [45] On sait que la forme $\widehat{\mu}$, positivement homogène, sous-linéaire, est linéaire si, et seulement si, il existe une unique forme linéaire qui lui est inférieure : $\nu \leq \widehat{\mu}$. Mais une telle mesure $\nu \leq \widehat{\mu}$ vérifie $\nu(f) \leq \mu(f) = \widehat{\mu}(f)$, pour toute fonction concave continue, elle est donc plus diffuse que μ. Ainsi, dire que μ est maximale nous donne l'existence d'une unique ν (égale à μ !) inférieure à $\widehat{\mu}$ [46] ; réciproquement, si $\widehat{\mu}$ est linéaire, aucune autre mesure que $\mu(= \widehat{\mu})$ n'est inférieure à $\widehat{\mu}$, *i.e.* n'est plus diffuse que μ. $\rhd$

(ii) nous dit exactement que quelle que soit f dans $\mathcal{C}(K)$, μ ne charge que $\{x \in K; \widehat{f}(x) = f(x)\}$. On donne à cet ensemble le nom d'***ensemble bordant de*** f, et on le note K_f. En ces termes, μ *est maximale si, et seulement si, son support est inclus dans tous les ensembles bordants.*

Remarque : $\widehat{\mu}$ et μ étant toutes deux continues, il suffit qu'elles coincident sur $\mathcal{S}$ [47] pour qu'elles soient égales sur $\mathcal{C}(K)$.

2.5 Représentation intégrale dans les convexes compacts

On a vu au corollaire 3 que pour toute fonction f, continue sur K, f et $\widehat{f}$ coincident sur $\mathcal{E}(K)$; on a donc l'inclusion $\mathcal{E}(K) \subset K_f$. La caractésrisation des mesures maximales obtenue au théorème précédent permet donc d'affirmer que

Théorème 13. *Toute mesure à support dans $\mathcal{E}(K)$ est maximale.*

Si l'inclusion $\mathcal{E}(K) \subset K_f$ est en général stricte, il est vrai cependant que

Proposition 9. $\mathcal{E}(K) = \bigcap_{f \in \mathcal{S}} K_f$

$\lhd$ Il nous suffit de voir qu'un point x non extrêmal, $x = \frac{x_1+x_2}{2}$, $x_1 \in K$, $x_2 \in K$, n'est pas dans un K_f particulier, $f \in \mathcal{S}$. On trouve ce f sous forme du carré d'une forme linéaire continue séparant x_1 et x_2 : $f(x) < \frac{f(x_1)+f(x_2)}{2} \leq \frac{\widehat{f}(x_1)+\widehat{f}(x_2)}{2} \leq \widehat{f}(x)$ [48]. $\rhd$

Au vu de ce résultat, on ne peut qu'être satisfait du théorème de représentation suivant, corollaire du théorème de caractérisation des mesures maximales.

Théorème 14. *Tout point de K est barycentre d'une mesure probabilité portée par tout les ensembles bordants K_f, $f \in \mathcal{S}$.*

Mais attention, aussi tentant que cela soit, on ne peut affirmer que μ est à support dans $\mathcal{E}(K)$; G. Choquet donne un contre-exemple du à Mokobodski dans *Lectures on analysis, volume II, Representation theory*, p153.

Le cadre adequat pour obtenir une telle conclusion est le cadre métrique.

Théorème 15. *Si K est métrisable, tout point de K est barycentre d'une probabilité à support dans $\mathcal{E}(K)$.*

[45] Vu ce que l'on a rappelé en Appendice sur le théorème de prolongement de hahn-Banach
[46] On a toujours $\mu \leq \widehat{\mu}$
[47] et donc sur la partie dense $\mathcal{S} - \mathcal{S}$.
[48] Ce qu'on fait ici, c'est trouver une fonction qui est strictement convexe sur le segment $[x_1, x_2]$.

$\triangleleft$ Cela vient de ce que *si K est métrisable, $\mathcal{E}(K)$ est l'ensemble bordant d'une certaine fonction convexe continue f.*

Comme dans la démonstration de l'identité $\mathcal{E}(K) = \bigcap_{f \in \mathcal{S}} K_f$, il nous suffit de trouver une fonction strictement convexe [49]. La métrisabilité de K nous aide en cela qu'elle implique la séparabilité de $\mathcal{C}(K)$, a fortiori celle de l'ensemble des fonctions affines continues. On peut donc trouver une suite $\{\ell_n\}$ d'applications affines de normes ≤ 1, *séparant les points de K*. L'application continue $f := \sum 2^{-n}\ell_n$ est strictement convexe : deux points distincts x et y étant séparés par une ℓ_p, $\ell_p(\frac{x+y}{2}) < \frac{\ell_p(x)+\ell_p(y)}{2}$, et $f(\frac{x+y}{2}) < \frac{f(x)+f(y)}{2}$. $\triangleright$

2.6 Unicité de la représentation

Que nous manque-t-il pour que tout point de K se représente de façon unique ?

Choisissons x dans K. Ce qu'il nous faut trouver, c'est un probabilité (de barycentre x) majorant toutes les autres [50] sur $\mathcal{S}$. C'est presque ce que nous fournit la proposition 7 :

$$\mu_x(f) := \widehat{f}(x) = \sup\{\mu(f); \mu \in \mathcal{P}(K), r(\mu) = x\}$$

à ceci près que μ_x n'est pas une mesure, l'additivité lui faisant défaut. Si on lui prète cette vertu [51], μ_x s'étendra à $\mathcal{S} - \mathcal{S}$, puis à $\mathcal{C}(K)$, en une probabilité sur K, de barycentre x, majorant sur $\mathcal{S}$ toutes les autres probabilités ayant x pour barycentre : μ_x est maximale, l'unique telle probabilité puisque la relation $\prec$ est un *ordre*.

Une façon de s'assurer de la linéarité de $f \in \mathcal{S} \mapsto \widehat{f}$ est de supposer chaque $\widehat{f}$ affine. En effet, $\widehat{f}$, scs et affine, étant limite décroissante de fonctions f_i affines continues, pour lesquelles on a $\widehat{f_i}(x) = \mu(\widehat{f_i}) = \mu(f_i)$, pour toute mesure *maximale* de barycentre x, un passage à la limite [52] nous donne $\widehat{f}(x) = \mu(f)$, qui est bien une expression linéaire en f.

Cependant, cette condition ne semble pas plus aisée à vérifier que la linéarité de $f \in \mathcal{S} \mapsto \widehat{f}$, d'autant plus que la seule chose devant décider de l'unicité de la représentation (comme de son existence) étant la géométrie de K, c'est en ses termes qu'une réponse satisfaisante devra être formulée.

Il nous faut, avant d'y arriver, préciser quelques notions géométriques.

2.6.1 Espaces vectoriels ordonnés

On appellera ***espace vectoriel ordonné*** un espace vectoriel E muni d'un ordre (partiel) compatible avec les opérations d'espace vectoriel : si $x \leq y$, on a pour tout $z \in E$ et tout $\alpha \geq 0$, $x + z \leq y + z$ et $\alpha x \leq \alpha y$.

De tels ordres sont totalement déterminés pas leur cone positif $E^+ := \{z \in E; z \geq 0\}$, x étant supérieur à y si, et seulement si, $x - y \in E^+$. En ce sens, le cone convexe E^+ donne une image géométrique de l'ordre $\leq$.

Exemple : On peut associer à tout compact de E l'ordre sur $E \times \mathbb{R}$ associé au cone de la figure suivante.

****

[49]Sur l'espace tout entier cette fois-ci.

[50]Ayant x pour barycentre.

[51] :-)

[52]Légitime puisque décroissant.

Si deux points quelconque de E ne sont pas forcément comparables, il se peut que l'ensemble des majorants communs aux deux points soit non vide et même qu'il possède un plus petit élément. S'il en va ainsi quels que soient x et y, on dit que $(E, \leq)$ est un ***treillis vectoriel*** et on note $x \smile y$ ce plus petit majorant de x et y, qu'on nomme "sup de x et de y". Dans un tel ensemble, deux points quelconque ont aussi un plus grand minorant, égal à $-(-x) \smile (-y)$, noté $x \frown y$, et appelé "inf de x et de y".

Géométriquement, $z = x \smile y$ si, et seulement si, (E^+ translaté en x) intersecte (E^+ translaté en y) selon (E^+ translaté en z).

Dans l'exemple ci-dessous, le cone de gauche n'est pas celui d'un treillis vectoriel, celui de droite en est un.

********FIGURE********

Exercice : En utilisant la caractérisation des mesures maximales à l'aide de leur support, montrez que l'ensemble $\mathcal{M}$ des mesures maximales est un cone convexe qui jouit de la propriété suivante : ********

A cette notion de treillis est naturellement associée celle d'***isomorphisme entre cones positifs*** : $\Phi : E^+ \to F^+$ est un isomorphisme si, et seulement si, Φ est une bijection positivement homogène et additive [53].

S'il n'y a qu'une proposition à retenir sur les treillis vectoriels, c'est la proposition suivante, appelée

Lemme 2. –***Lemme de décomposition de Riesz***– *Dans un treillis vectoriel* $(E, \leq)$, *si trois éléments positifs* x, y_1, y_2, *sont tels que* $x \leq y_1 + y_2$, *alors* x *peut s'écrire comme une somme* $x = z_1 + z_2$, *où* $z_1 \leq y_1$ *et* $z_2 \leq y_2$.

$\lhd$ Il n'y a qu'à prendre $z_1 = x \frown y_1$ et $z_2 = x - z_1 \geq 0$ et à constater que puisque $x \leq y_1 + y_2$, *i.e.* $y_1 \geq x - y_2$, on a $y_1 \frown x \geq (x - y_2) \frown x \geq x - y_2$, y_2 étant positif, ce qui montre que $z_2 \leq y_2$. $\rhd$

****

2.6.2 Simplexes

Pourquoi parler de treillis vectoriels dans l'histoire qui nous intéresse ?

En dimension finie, n, seuls les simplexes, enveloppe convexe de $n+1$ points affinement libres, jouissent de cette propriété d'unicité de la représentation. Cependant, leur définition ne s'étend pas telle quelle en dimension infinie. Voici une autre formulation qui a l'avantage de ne pas faire intervenir la dimension de l'espace ; $\widetilde{K}$ y désigne le cone construit sur K représenté à la figure 3.

Proposition 10. $K \subset \mathbb{R}^n$ *est un simplexe si, et seulement si, le cone* $\widetilde{K} \subset \mathbb{R}^n \times \mathbb{R}$ *fait de* $\mathbb{R}^n \times \mathbb{R}$ *un treillis vectoriel.*

On dit aussi que $\widetilde{K}$ est un treillis pour son ordre propre.

Si c'est un fait visuellement évident (*cf* Fig. 4), l'implication $\Leftarrow$ n'a rien d'évident ; elle résultera du théorème d'unicité à venir.

[53] En somme, c'est une bijection qui respecte la structure.

Oubliant la définition donnée en dimension finie, on dira que le convexe compact $K \subset E$ est un ***simplexe*** si, et seulement si, le cone $\widetilde{K}$ est un treillis pour son ordre propre.

Rappelons qu'on a noté $\mathcal{M}$ le cone des mesures maximales sur K et qu'on a vu qu'il s'agit d'un treillis pour son ordre propre.

Supposons maintenant que tout point de K est uniquement représenté par une probabilité maximale. L'application $\Phi : \mathcal{M} \to \widetilde{K}$, $\mu \mapsto (||\mu||r(\mu), ||\mu||)$, où $||\mu||$ est la masse de μ, est alors un isomorphisme : l'hypothèse d'unicité nous donne la bijection, l'homogène positivité est évidente, quant à l'additivité, elle résulte de ce que $r(\mu + \mu') = \frac{||\mu||r(\mu)+||\mu'||r(\mu')}{||\mu||+||\mu'||}$. $\mathcal{M}$ étant un treillis pour son ordre propre, cette propriété se transporte sur $\widetilde{K}$; in other words : *K est un simplexe.*

Nous en sommes donc rendus au point suivant :

" si chaque $\widehat{f}$, $f \in \mathcal{S}$, est affine sur K, alors il y a unicité de la représentation, ce qui entraîne que K est un simplexe ".

Si l'on arrive à montrer que l'hypothèse "K est un simplexe" implique que "chaque $\widehat{f}$, $f \in \mathcal{S}$, est affine sur K", on aura trouvé la caractérisation géométrique que l'on cherchait pour notre critère d'unicité.

Pour une fonction $g : K \to \mathbb{R}$, notons $\underline{g} : \widetilde{K} \to \mathbb{R}$, $(\lambda x, \lambda) \mapsto \lambda g(x)$ son extension par homogénéité à $\widetilde{K}$.

Supposons donc que K est un simplexe ; dans $E \times \mathbb{R}$, la conclusion recherchée se lit : "chaque $\underline{\widehat{f}}$, $f \in \mathcal{S}$, est *additive* sur $\widetilde{K}$".

Que vaut $\underline{\widehat{f}}(z)$? Comme $\widehat{f}(x) = \sup\{\sum \alpha_i f(x_i); \sum \alpha_i x_i = x, \sum \alpha_i = 1\}$, on a

$$\begin{aligned}\underline{\widehat{f}}(z) = \underline{\widehat{f}}(\lambda x, \lambda) = \lambda \widehat{f}(x) =& \sup\{\sum \lambda\alpha_i f(x_i); \sum \alpha_i x_i = x, \sum \alpha_i = 1\} \\ &= \sup\{\sum \lambda\alpha_i f(x_i); \sum \lambda\alpha_i x_i = \lambda x, \sum \lambda\alpha_i = \lambda\} \\ &= \sup\{\sum \lambda_i f(x_i); \sum (\lambda_i x_i, \lambda_i) = (\lambda x, \lambda)\} \\ &= \sup\{\sum \underline{f}(z_i); \sum z_i = z\}.\end{aligned} \tag{2.6.1}$$

Cette écriture rend claire l'inégalité : $\underline{\widehat{f}}(z_1 + z_2) \geq \underline{\widehat{f}}(z_1) + \underline{\widehat{f}}(z_2)$.

Pour obtenir l'autre inégalité, on se sert du *lemme de décomposition de Riesz, qu'il n'est licite d'appliquer que parce que K est un simplexe*, pour écrire :

$$\sup\{\sum \underline{f}(z_i); \sum z_i = z_1 + z_2\} = \sup\{\sum \underline{f}(z_i' + z_i''); \sum z_i' = z_1, \sum z_i'' = z_2\}$$

et puisque f est convexe,

$$\underline{f}(z + z') = \underline{f}((\alpha x, \alpha) + (\beta y, \beta)) = (\alpha + \beta) f(\frac{\alpha x + \beta y}{\alpha + \beta}) \leq \alpha f(x) + \beta f(y) = \underline{f}(z) + \underline{f}(z')$$

ce qui nous donne : $\underline{\widehat{f}}(z_1 + z_2) \leq \underline{\widehat{f}}(z_1) + \underline{\widehat{f}}(z_2)$, avec cela la linéarité de $\underline{\widehat{f}}$.

On a ainsi obtenu le crière d'unicité recherché :

Théorème 16. *Chaque point de K admet une unique représentation comme le barycentre d'une probabilité maximale si, et seulement si, K est un simplexe.*