Analyse Convexe

Cours M1 (4M057)

2017-2018, Période 2

Sorbonne Université

Préface

Objectifs de l'UE : L'analyse convexe est un des piliers des mathématiques appliquées. Elle intervient dans la modélisation et la résolution numérique de problèmes dans pratiquement tous les secteurs où la modélisation mathématique est pertinente : en ingénierie, en statistiques, en physique, en économie, en finance, dans les sciences de l'information, pour la simulation numérique et des données. L'objectif de ce cours de fournir les fondements de l'analyse convexe, ses implications en méthodes algorithmiques et ses applications vers le traitement des megadonnés (big data).

Le cours "Analyse convexe" constitue une UE d'orientation de M1 ; il est particulièrement approprié pour des étudiants visant les parcours *Mathématiques fondamentales, Mathématiques de la modélisation* ou *Ingénierie mathématique.* Dans certains cas il peut également être validé en M2 (à verifier avec son directeur d'études).

Prérequis : Algèbre linéaire, topologie élémentaire. Le cours L3 *Optimisation linéaire et convexité* (LM239) n'est pas requis ; pour les étudiants ayant cette expérience le présent cours servira en tant qu'un complément théorique et approfondi.

Ces notes sont principalement destinées aux étudiants qui suivent le cours à distance ; pour les autres elles seront un complément au cours oral, où parfois les idées essentielles seront présentées. Réciproquement, certains thèmes de base ont été inclus ici à des fins de référence ; ils seront traités seulement très succinctement dans le cours oral. Sauf mention explicite à l'effet contraire, les notes sont complètes et autonomes ; néanmoins, quelques livres sont indiqués dans la bibliographie pour des lecteurs désirants d'élargir ses connaissances.

Ce polycopié a été développé à partir d'un texte généreusement fourni par Sylvie Delabrière et Yves Raynaud, à qui je veux adresser mes remerciements. Cependant, la responsabilité pour tous les défauts et les erreurs est entièrement mienne.

Stanislaw Szarek

Table des matières

Chapitre premier
Convexes, fonctions convexes

1.1 Ensembles convexes dans un espace vectoriel, exemples et premières propriétés

Soit E est un espace vectoriel réel.

Définition 1.1.1 (Convexes)**.** *Une partie C de E est dite convexe si et seulement si*

$$\forall x, y \in C, \quad \forall \theta \in [0,1] : \quad \theta x + (1-\theta) y \in C$$

Géométriquement, $z(\theta) = \theta x + (1-\theta)y$ est le barycentre des points x et y affectés des coefficients (ou poids) respectifs θ et $1-\theta$. Lorsque θ décrit l'intervalle $[0,1]$, $z(\theta)$ décrit le segment fermé d'extrémités x et y que l'on notera $[x,y]$. Dire que C est convexe, c'est donc dire que C contient un segment dès qu'elle contient ses extrémités.

En fait un convexe contient toute combinaison convexe de ses points :

Définition 1.1.2 (Combinaison convexe)**.** *Pour toute famille finie $x_1, \ldots, x_n$ de points de E et tout système $\lambda_1, \ldots, \lambda_n$ de réels positifs ou nuls avec $\sum_{i=1}^n \lambda_i = 1$, le point $\sum_{i=1}^n \lambda_i x_i$ s'appelle une combinaison convexe des points $x_1, \ldots, x_n$.*

Proposition 1.1.3. *Soit C un convexe de E. Pour toute famille finie $x_1, \ldots, x_n$ de points de C et tout système $\lambda_1, \ldots, \lambda_n$ de réels positifs ou nuls avec $\sum_{i=1}^n \lambda_i = 1$, le point $\sum_{i=1}^n \lambda_i x_i$ appartient à C.*

Démonstration : On procède par récurrence sur n. C'est évident si $n = 1$. Supposons que c'est vrai jusqu'à l'ordre $n-1$ et considérons le point $\sum_{i=1}^n \lambda_i x_i$, où les x_i sont dans C. Si $\lambda_1 = 1$, les autres coefficients sont nuls et il n'y a rien à montrer. Si $\lambda_1 < 1$ alors $\sum_{i=2}^n \lambda_i = 1 - \lambda_1 > 0$. Posons $\mu_i = \dfrac{\lambda_i}{1-\lambda_1}$, on a $\sum_{i=2}^n \mu_i = 1$, d'où $y := \sum_{i=2}^n \mu_i x_i$ est dans C par l'hypothèse de récurrence et donc $x = \lambda_1 x_1 + (1-\lambda_1) y \in C$. $\square$

La démonstration de la propriété suivante est évidente :

Proposition 1.1.4. *Si $(C_i)_{i\in I}$ est une famille quelconque de convexes, leur intersection $\bigcap_{i\in I} C_i$ est convexe.*

Définition 1.1.5 (Enveloppe convexe)**.** *Soit A une partie de E. L'enveloppe convexe de A, notée* $\mathrm{conv}(A)$*, est l'intersection de tous les convexes contenant A.*

L'enveloppe convexe $\mathrm{conv}(A)$ est évidemment un convexe contenant A, et c'est le plus petit : si C est un convexe contenant A, il contient $\mathrm{conv}(A)$.

Proposition 1.1.6. *L'enveloppe convexe de A est l'ensemble de toutes les combinaisons convexes finies d'éléments de A :*

$$\mathrm{conv}(A) = \Big\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \Big\}$$

Démonstration : Comme $\mathrm{conv}(A)$ est convexe, cet ensemble contient toutes les combinaisons convexes finies de ses éléments (Proposition 1.1.3), donc en particulier celles des éléments de A.

Inversement il est facile de voir que l'ensemble $C = \{\sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1\}$ est convexe ; il contient A, donc aussi $\mathrm{conv}(A)$. $\square$

Proposition 1.1.7. *1) Si A est fini, soit $A = \{a_1, \ldots, a_p\}$, alors*

$$\mathrm{conv}(A) = \Big\{ \sum_{i=1}^{p} \lambda_i a_i \ /\lambda_i \geq 0, \sum_{i=1}^{p} \lambda_i = 1 \Big\}$$

2) Plus généralement, si $A_1, \ldots, A_p$ sont des convexes de E, alors

$$\mathrm{conv}(A_1 \cup A_2 \cup \ldots \cup A_p) = \Big\{ \sum_{i=1}^{p} \lambda_i x_i \ /x_i \in A_i, \lambda_i \geq 0, \sum_{i=1}^{p} \lambda_i = 1 \Big\}$$

Démonstration : *1)* est évident.

Montrons *2)* : l'inclusion

$$\Big\{ \sum_{i=1}^{p} \lambda_i x_i \ /x_i \in A_i, \lambda_i \geq 0, \sum_{i=1}^{p} \lambda_i = 1 \Big\} \subset \mathrm{conv}(A_1 \cup A_2 \cup \ldots \cup A_p)$$

est évidente. Montrons l'autre inclusion dans le cas où $p = 2$ pour simplifier les notations :

Si $x \in \mathrm{conv}(A_1 \cup A_2)$, x s'écrit $x = \sum_{i=1}^n \lambda_i x_i$ avec $\lambda_i \geq 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n \lambda_i = 1$. Les x_i sont dans $A_1 \cup A_2$, soit $I = \{i \in \{1, \ldots, n\}\ ,\ x_i \in A_1\}$, $J = \{i \in \{1, \ldots, n\}\ ,\ x_i \in A_2\}$ et $\lambda = \sum_{i \in I} \lambda_i$. Alors, $0 \leq \lambda \leq 1$ et $1 - \lambda = \sum_{i \in J} \lambda_i$. Si $\lambda = 0$ ou 1, x est alors dans $\mathrm{conv}(A_1)$ ou dans $\mathrm{conv}(A_2)$ et par suite dans $\mathrm{conv}(A_1 \cup A_2)$. Sinon, on écrit :

$$x = \lambda \sum_{i \in I} \frac{\lambda_i}{\lambda} x_i + (1 - \lambda) \sum_{i \in J} \frac{\lambda_i}{(1 - \lambda)} x_i = \lambda y_1 + (1 - \lambda) y_2\ ,\ y_1 \in A_1\ ,\ y_2 \in A_2$$

On en déduit bien que $x \in \mathrm{conv}(A_1 \cup A_2)$, d'où l'inclusion

$$\mathrm{conv}(A_1 \cup A_2) \subset \{\sum_{i=1}^2 \lambda_i x_i\ / x_i \in A_i, \lambda_i \geq 0, \sum_{i=1}^2 \lambda_i = 1\} \qquad \square$$

1.2 Cônes convexes

Dans un espace vectoriel réel E, on va s'intéresser à certains convexes particuliers, les cônes convexes, qui joueront un rôle important dans la structure des convexes :

Définition 1.2.1 (Cône). *1) Une partie Γ de E est un cône de sommet x_0 si elle est stable par les homothéties de centre x_0, de rapport $\lambda > 0$. Si on ne précise pas le sommet x_0, ce sera 0.*
2) Γ est un cône convexe de sommet x_0 si de plus Γ est un convexe.
3) Un cône est pointé s'il contient son sommet, épointé sinon.

Autrement dit, Γ est un cône de sommet x_0 si et seulement si quelque soit $x \in \Gamma$ et $\lambda > 0$ alors, $x_0 + \lambda(x - x_0) \in \Gamma$.
Dans le cas où $x_0 = 0$: Γ est un cône si et seulement si $\lambda\Gamma \subset \Gamma$, pour tout $\lambda > 0$.
On va maintenant donner une caractérisation des cônes convexes (de sommet 0) :

Notation 1.2.2. *Si A et B sont deux parties de E et $\lambda \in \mathbb{R}$, on note :*

$$A + B = \{a + b\ /\ a \in A, b \in B\}, \quad \lambda A = \{\lambda a\ /\ a \in A\}.$$

Proposition 1.2.3. *Une partie Γ de E est un cône convexe (de sommet 0) si et seulement si elle est stable par addition et par multiplication par les réels strictement positifs, c'est à dire, avec la notation précédente, si et seulement si $\Gamma + \Gamma \subset \Gamma$ et $\lambda\Gamma \subset \Gamma$ pour tout $\lambda > 0$.*

Démonstration : Si Γ est un cône convexe de sommet 0 et $x, y \in \Gamma$, on a $\dfrac{x+y}{2} \in \Gamma$. D'où $x + y = 2\,\dfrac{x+y}{2} \in \Gamma$. Donc $\Gamma + \Gamma \subset \Gamma$. La propriété $\lambda\Gamma \subset \Gamma$ est évidente.

La réciproque est immédiate. $\square$

Proposition 1.2.4. *1) Soit Γ un cône convexe non vide de sommet 0. Le sous-espace vectoriel engendré, en bref SEV, par Γ est $\Gamma - \Gamma$.*
2) Si $0 \in \Gamma$, le plus grand sous-espace vectoriel inclus dans Γ est $\Gamma \cap (-\Gamma)$.

Démonstration : *1)* Le sous-espace vectoriel engendré par Γ contient nécessairement Γ et $-\Gamma$, donc aussi $\Gamma - \Gamma$. Inversement, comme par hypothèse, Γ est un cône convexe, on vérifie aisément que $\Gamma - \Gamma$ est un SEV qui contient Γ, donc ausssi le SEV engendré par Γ. D'où l'égalité voulue.

2) Notons d'abord que l'hypothèse $0 \in \Gamma$ est nécessaire dans cette propriété car tout sous espace vectoriel contient 0. Comme Γ est un cône convexe contenant 0, on vérifie aisément que $\Gamma \cap (-\Gamma)$ est un SEV inclus dans Γ. Inversement, tout SEV inclus dans Γ est aussi inclus dans $-\Gamma$ et donc dans $\Gamma \cap (-\Gamma)$. D'où l'égalité voulue. $\square$

Notation 1.2.5. *Soit Γ un cône convexe non vide (de sommet 0). On notera*

$$N(\Gamma) = \{0\} \cup \big[\Gamma \cap (-\Gamma)\big]$$

Définition 1.2.6 (Cône convexe engendré)**.** *Soit $A \subset E$. Le plus petit cône convexe pointé de sommet 0 contenant A s'appelle le cône convexe engendré par A. On le note* $\mathrm{cone}(A)$.

Proposition 1.2.7. *1) L'intersection d'une famille quelconque de cônes convexes de même sommet est un cône convexe.*
2) Si $A \subset E$, le cône convexe engendré par A est l'ensemble des combinaisons linéaires finies à coefficients positifs ou nuls d'éléments de A, c'est à dire :

$$\mathrm{cone}(A) = \Big\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \geq 0 \Big\}$$

3) Si $C \subset E$ est convexe, le cône convexe engendré par C est $\bigcup_{\lambda \geq 0} \lambda C$.

Démonstration : *1)* est immédiat.

2) L'inclusion $\big\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \geq 0 \big\} \subset \mathrm{cone}(A)$ est claire. Inversement, l'ensemble $\big\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \geq 0 \big\}$ est un cône convexe contenant A donc contenant aussi $\mathrm{cone}(A)$, d'où l'égalité.

3) L'inclusion $\bigcup\limits_{\lambda \geq 0} \lambda C \subset \mathrm{cone}(C)$ est claire. Inversement, grâce à la convexité de C, on vérifie aisément que $\bigcup\limits_{\lambda \geq 0} \lambda C$ est un cône convexe. Comme il contient C, il contient aussi $\mathrm{cone}(C)$. □

1.3 Sous-espaces affines

Dans un espace vectoriel réel E, on va s'intéresser à une deuxième catégorie de convexes particuliers, les sous-espaces affines, qui joueront aussi un rôle important dans la structure des convexes :

Définition 1.3.1 (Sous-espaces affines)**.** *Un sous-ensemble $M \subset E$ est un sous-espace affine, en bref SEA de E si :*

$$\forall x, y \in M, \quad \forall \theta \in \mathbb{R} : \quad \theta x + (1-\theta) y \in M$$

Géométriquement cela signifie que M contient une droite D dès qu'il contient deux points distincts de D. Un SEA de E est clairement un convexe.

Définition 1.3.2. *Si A est une partie de E et $x_0 \in E$, le translaté de A par x_0 est*

$$x_0 + A = \{x_0 + a \mid a \in A\}$$

Si $y_0 \in E$, on a évidement $(A + x_0) + y_0 = A + (x_0 + y_0)$.

Proposition 1.3.3. *1) Si M est un SEA de E, tout translaté $x_0 + M$ de M est un SEA de E.*
2) Un SEA est un sous-espace vectoriel de E, en bref SEV, si et seulement s'il contient 0.
3) Si M est un SEA de E, et $x_0 \in M$, alors $M - x_0$ est un SEV de E. Inversement, s'il existe $x_0 \in E$ tel que $M - x_0$ soit un SEV de E, alors M est un SEA de E.

Démonstration : *1)* est une conséquence immédiate de la Définition 1.3.2.

2) Un SEV est évidemment un SEA contenant 0. Réciproquement, si M est un SEA contenant 0, alors :

$$\forall x \in M \ , \ \forall \lambda \in \mathbb{R} \ , \ \lambda x = \lambda x + (1-\lambda) 0 \in M$$

et si $x, y \in M$, $\dfrac{x+y}{2} \in M$, d'où :

$$x + y = 2 \, \frac{x+y}{2} \in M$$

M est donc bien un SEV.

3) Si M est un SEA de E, et $x_0 \in M$, alors $M - x_0$ est un SEA de E contenant 0 donc est un SEV. Inversement, s'il existe $x_0 \in E$ tel que $M - x_0$ soit un SEV de E, alors $0 \in M - x_0$ d'où $x_0 = 0 + x_0 \in (M - x_0) + x_0 \in M$. Donc M est un SEA. □

Définition 1.3.4 (Direction d'un sous-espace affine)**.** *Soit M un SEA et $x_0 \in M$. Le SEV : $V = M - x_0$ s'appelle la direction de M.*

Remarque 1.3.5. On a aussi $V = \{x - y \ , \ x, y \in M\}$, donc V ne dépend pas de $x_0 \in M$.

Définition 1.3.6 (Sous-espace affine engendré)**.** *Si A est une partie de E, le sous-espace affine engendré par A est le plus petit SEA contenant A c'est à dire l'intersection des SEA contenant A. On le notera* $\mathrm{Aff}(A)$.

Proposition 1.3.7. *1) On a*

$$\mathrm{Aff}(A) = \Big\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ n \geq 1, x_i \in A, \lambda_i \in \mathbb{R}, \sum_{i=1}^{n} \lambda_i = 1 \Big\}.$$

2) Si $x_0 \in A$, alors $\mathrm{Aff}(A) = x_0 + V$ où V est l'espace vectoriel engendré par $A - x_0$.
3) Si C est convexe, $\mathrm{Aff}(C) = \big\{\theta x + (1-\theta) y \ / \ \theta \in \mathbb{R}, x, y \in C\big\}$.

Démonstration : Le *1)* se démontre comme pour l'enveloppe convexe, voir 1.1.6.

2) est évident par construction.

Pour montrer le *3)*, on prend deux points $\theta_1 x_1 + \theta_2 x_2$ et $\theta_3 x_3 + \theta_4 x_4$ avec $x_i \in C$ et $\theta_i \in \mathbb{R}, i = 1, \dots, 4$, $\theta_1 + \theta_2 = 1$, $\theta_3 + \theta_4 = 1$ et $\theta \in \mathbb{R}$. Alors, le point $\theta(\theta_1 x_1 + \theta_2 x_2) + (1-\theta)(\theta_3 x_3 + \theta_4 x_4)$ s'écrit $\sum_{i=1,\dots,4} \alpha_i x_i$ où $\sum_{i=1,\dots,4} \alpha_i = 1$. Grâce à la convexité de C, en regroupant les x_i tels que $\alpha_i \geq 0$ et ceux tels que $\alpha_i < 0$, on montre facilement que le vecteur $\sum_{i=1,\dots,4} \alpha_i x_i$ est bien de la forme $\lambda x + (1-\lambda) y$, avec $\lambda = \displaystyle\sum_{\alpha_i \geq 0} \alpha_i \in [0,1]$ tandis que $x = \dfrac{1}{\lambda} \displaystyle\sum_{\alpha_i \geq 0} \alpha_i x_i$, $y = \dfrac{1}{(1-\lambda)} \displaystyle\sum_{\alpha_i < 0} \alpha_i x_i$ sont deux points de C.
On en déduit que l'ensemble $\big\{\theta x + (1-\theta) y \ / \ \theta \in \mathbb{R}, x, y \in C\big\}$ est un SEA qui contient C donc aussi $\mathrm{Aff}(C)$.
L'inclusion inverse est évidente par définition d'un SEA. □

Définition 1.3.8 (Indépendance affine)**.** *Les points $x_0, \dots, x_k$ de E sont dits affinement indépendants si aucun d'eux n'appartient au sous-espace affine engendré par les k autres.*

Proposition 1.3.9. *1) Les points $x_0, \dots, x_k$ sont affinement indépendants si et seulement si les vecteurs $x_1 - x_0, \dots, x_k - x_0$ sont linéairement indépendants.*
2) Une autre condition nécessaire et suffisante pour que les points $x_0, \dots, x_k$ soient affinement indépendants est qu'il n'existe pas $\alpha_0, \dots, \alpha_k$ non tous nuls avec $\sum_{i=0}^k \alpha_i x_i = 0$ et $\sum_{i=0}^k \alpha_i = 0$.
3) Si E est de dimension finie p, tout système de $p+1$ points affinement indépendants $x_0, \dots, x_p$ engendre affinement E, c'est à dire $E = \mathrm{Aff}(x_0, \dots, x_p)$.

Démonstration : *1)* et *3)* résultent facilement de *2)* de la Proposition 1.3.7. *2)* est une conséquence du fait que si $\sum_{i=0}^k \alpha_i = 0$, alors $\sum_{i=0}^k \alpha_i x_i = \sum_{i=1}^k \alpha_i (x_i - x_0)$. □

Définition 1.3.10 (Repère affine)**.** *1) Si $x_0, \dots, x_p$ sont $p+1$ points affinement indépendants et si $E = \mathrm{Aff}(x_0, \dots, x_p)$, on dit que $(x_0, \dots, x_p)$ est un repère affine de E. Il revient au même de dire que $(x_1 - x_0, \dots, x_p - x_0)$ est une base de E.*
2) On dira qu'un SEA est de dimension k s'il est engendré par $k+1$ points affinement indépendants.
3) Plus généralement, on dira qu'une partie A d'un espace vectoriel E est de dimension k si $\mathrm{Aff}(A)$ est de dimension k.

Une catégorie importante de sous-espaces affines est celle des hyperplans affines.

Définition 1.3.11 (Hyperplan affine)**.** *Un hyperplan affine est un sous-espace affine propre maximal.*

Autrement dit M est un hyperplan affine si et seulement si
i) $M \neq E$
ii) E est le seul SEA contenant strictement M.

Remarque 1.3.12. *ii)* signifie que pour tout $y \notin M$, $\mathrm{Aff}(M \cup \{y\}) = E$.

Proposition 1.3.13. *Le SEA M est un hyperplan affine si et seulement si sa direction V est un hyperplan vectoriel, c'est à dire que pour tout $u \in E \setminus V$, on a $E = V \oplus \mathbb{R}u$.*

Démonstration : Soit M un hyperplan affine de E et soit V sa direction. Alors $M = x_0 + V$ où $x_0 \in M \subset E$. Alors $V \neq E$ sinon M serait aussi égal à E. Supposons que V ne soit pas un hyperplan de E, c'est à dire qu'il existe un vecteur $u \in E$, tel que $u \notin V$ et $V \oplus \mathbb{R}u \neq E$. Alors $M' = x_0 + (V \oplus \mathbb{R}u)$ est un SEA contenant M et non égal à E, ce qui contredit la maximalité de M. Donc V est un hyperplan vectoriel de E.

Réciproquement, supposons maintenant que V est un hyperplan vectoriel de E et soit $u \notin V$ tel que $E = V \oplus \mathbb{R}u$. Alors $u \notin M$ car sinon $M = E$ et V ne serait pas la direction de M. De plus, si M' est un SEA contenant M et non égal à E, on peut écrire, $M' = x_0 + W$, où W est un SEV contenant V et non égal à E, ca qui contredit la maximalité de V. M est donc bien un hyperplan affine de E. □

Proposition 1.3.14. *Les hyperplans affines de E sont exactement les ensembles de la forme $\{x \in E \ / \ f(x) = a\}$, où f est une forme linéaire non nulle sur E et a un réel.*

Démonstration : Soit f une forme linéaire non nulle sur E, $a \in \mathbb{R}$, $H = \{x \in E \ / \ f(x) = a\}$ et $V = \{x \in E \ / \ f(x) = 0\}$. H est évidemment un SEA de E et V est sa direction. Soit $x_0 \in E$ tel que $f(x_0) = 1$. Si $y \in E$, $y - f(y)x_0 \in V$ d'où $y = (y - f(y)x_0) + f(y)x_0 \in V \oplus \mathbb{R}x_0$ et par suite, $E = V \oplus \mathbb{R}x_0$. Donc V est un hyperplan vectoriel de E et par suite, H un hyperplan affine de E.

Réciproquement, soit H est un hyperplan affine, V sa direction, x_0 un point de E n'appartenant pas à V, on pose $f(v + \lambda x_0) = \lambda$ pour tout $v \in V$ et $\lambda \in \mathbb{R}$. Alors f est une forme linéaire sur E, nulle sur V, donc constante sur H : en effet si $y_0 \in H$, on a, pour tout $y \in H$, $f(y) = f(y_0) + f(y - y_0) = f(y_0)$ car $y - y_0 \in V$. Si a est cette valeur, l'ensemble $\{x \in E \ / \ f(x) = a\}$ est un hyperplan affine contenant H, donc égal à H. □

Remarque 1.3.15. Si $\{x \in E \ / \ f(x) = a\}$ et $\{x \in E \ / \ g(x) = b\}$ sont deux équations d'un même hyperplan affine H, alors, il existe $\alpha \in \mathbb{R}$ tel que $g = \alpha f$.

Démonstration : En effet, si V est la direction de H, on a $V = \operatorname{Ker} f = \operatorname{Ker} g$.
Soit $x_0 \in E$, $x_0 \notin V$ tel que $f(x_0) = 1$ et posons $\alpha = g(x_0)$. Alors, $\alpha \neq 0$ car $g \neq 0$. Si $x = v + \lambda x_0$ est un point quelconque de E, on a : $f(x) = \lambda f(x_0) = \lambda$ et $g(x) = \lambda g(x_0) = \lambda\alpha$ donc $g(x) = f(x)$ pour tourt $x \in E$. □

Définition 1.3.16 (Demi-espaces)**.** *Soit H un hyperplan affine ; soient f une forme linéaire non nulle sur E et a un réel tels que $H = \{x \in E \ / \ f(x) = a\}$. On appelle demi-espaces associés à l'hyperplan H les ensembles $D_+ = \{x \in E \ / \ f(x) \leq a\}$ et $D_- = \{x \in E \ / \ f(x) \geq a\}$. Les demi-espaces stricts associés à l'hyperplan H sont les ensembles $D'_+ = \{x \in E \ / \ f(x) < a\}$ et $D'_- = \{x \in E \ / \ f(x) > a\}$.*

Ces ensembles sont convexes. Les paires $\{D_+, D_-\}$ et $\{D'_+, D'_-\}$ ne dépendent pas du couple (f, a) tel que $H = \{x \in E \ / \ f(x) = a\}$, à permutation près : ceci résulte de la Remarque 1.3.15.
On a évidemment $D'_+ = D_+ \setminus H$ et $D'_- = D_- \setminus H$.

Remarque 1.3.17. Une conséquence immédiate de la Proposition 1.1.4 est que toute intersection de demi-espaces affines, respectivement de demi-espaces affines stricts, est un convexe, éventuellement vide.
La question inverse est naturelle : quels sont les convexes représentables comme intersection de demi-espaces affines ? On y répondra dans les chapitres suivants.

1.4 Points extrémaux et faces des convexes dans un espace vectoriel

Définition 1.4.1 (Points extrémaux d'un convexe)**.** *Soit C un convexe d'un espace vectoriel réel E. On dit qu'un point x de C est un point extrémal de C si x n'appartient à aucun segment ouvert dont les extrémités sont dans C.*

Proposition 1.4.2. *Un point x est un point extrémal de C si et seulement si*

$$\textit{Il n'existe pas } \; y, z \in C \; \textit{ tels que } \; y \neq z \text{ et } x = \frac{y+z}{2}$$

ou encore si et seulement si :

$$y, z \in C, \; x = \frac{y+z}{2} \quad \Longrightarrow \quad x = y = z$$

DÉMONSTRATION : En effet, on remarque d'abord que la deuxième condition de la Proposition 1.4.2 est équivalente à la première en passant à la contraposée. D'autre part, si x est extrémal dans le convexe C, il est clair que la condition $y, z \in C, \; x = \dfrac{y+z}{2} \Rightarrow x = y = z$ est vérifiée.
Réciproquement, si cette condition est vérifiée et si x appartient à un segment ouvert dont les extrémités y, z sont dans C, alors le symétrique z' de l'une des deux extrémités, par exemple y, par rapport à x est sur ce segment, donc dans C ; il s'ensuit que les points, y, z', x, sont confondus d'où contradiction . Donc x est extrémal dans C. $\square$

Notation 1.4.3. *On désignera par* $\operatorname{Ext} C$ *l'ensemble des points extrémaux de C.*

Proposition 1.4.4. *1) Le point x du convexe C est extrémal si et seulement si $C \setminus \{x\}$ est encore convexe.*
2) Si A est une partie de E, on a $\operatorname{Ext}\big(\operatorname{conv}(A)\big) \subset A$.

DÉMONSTRATION : *1)* Supposons que x est extrémal dans C et soient $y, z \in C \setminus \{x\}$, $y \neq z$. Alors, $\dfrac{y+z}{2} \in C$ et $\dfrac{y+z}{2} \neq x$ par extémalité de x. Donc $\dfrac{y+z}{2} \in C \setminus \{x\}$ et $C \setminus \{x\}$ est bien convexe.

Réciproquement, si $C \setminus \{x\}$ est convexe, supposons $\dfrac{y+z}{2} = x$ avec $y, z \in C$. Alors, si $y \neq z$, on a nécessairement $y, z \in C \setminus \{x\}$ et par convexité, $x \in C \setminus \{x\}$, ce qui est une contradiction. Donc x est bien extrémal dans C.
2) Posons $C = \mathrm{conv} A$ et soit $x \in C \setminus A$. Soit $x = \sum_{i=1}^{k} t_i x_i$ une combinaison convexe avec $x_i \in A$ telle que la longueur k est minimal ; alors $t_i > 0$ pour tout i et, étant donné que $x \notin A$, on a forcement $k \geq 2$. On déduit que $x = t_1 x_1 + (1-t_1) y$, ou $y = (1-t_1)^{-1} \sum_{i=2}^{k} t_i x_i \in C$. En particulier, $x \in]x_1, y[$, un segment ouvert avec des extrémités dans C et donc $x \notin \mathrm{Ext}\, C$. □

Définition 1.4.5 (Face d'un convexe)**.** *1) Une face d'un convexe C est un convexe non vide $F \subset C$ tel que :*

$$x \in F, \quad y, z \in C, \quad x = \frac{y+z}{2} \Rightarrow y \text{ et } z \in F$$

2) Une face de dimension 1 est appelée une arête. Si $\dim F = \dim C - 1$, on dit que F est une facette de C.

Proposition 1.4.6. *F est une face d'un convexe C si et seulement si F est convexe et que si $x \in F$ appartient à un segment ouvert dont les extrémités sont dans C, ces extrémités sont en fait dans F.*

Démonstration : On passe d'un point quelconque d'un segment de F au milieu de ce segment en appliquant l'argument de la Proposition 1.4.2. □

Proposition 1.4.7. *Toute intersection non vide de faces de C est une face de C.*

Exemples 1.4.8. *1) Un exemple trivial de face de C est le convexe C lui même.*
2) Un point $x_0 \in C$ est extrémal pour C si et seulement si $\{x_0\}$ est une face de C.

Dans le cas des cônes, la notion de point extrémal n'a pas d'intérêt car seul le sommet du cône est un point extémal. Dans ce cas on peut introduire et étudier un concept d'une *génératrice extrémale*, (voir Définition 3.6.8).

Définition 1.4.9 (Hyperplan d'appui d'un convexe)**.** *On dit qu'un hyperplan H est un hyperplan d'appui du convexe C si C est inclus dans l'un des deux demi-espaces définis par H et rencontre H.*

Lemme 1.4.10. *Si H est un hyperplan d'appui du convexe C alors $F = H \cap C$ est une face de C.*

DÉMONSTRATION : En effet soit $f(x) = a$ une équation de H, f étant une forme linéaire et $a \in \mathbb{R}$. On a (par exemple) $f(x) \leq a$ pour tout $x \in C$. Si $x = \dfrac{y+z}{2}$, avec $x \in F$, $y, z \in C$, on a :

$$f(y) \leq a, \qquad f(z) \leq a \tag{$*$}$$

et

$$\frac{f(y) + f(z)}{2} = f(x) = a$$

d'où $0 = 2a - \big(f(y) + f(z)\big) = \big(a - f(y)\big) + \big(a - f(z)\big)$.
Or, une somme de nombres réels positifs n'est nulle que si chacun de ces deux réels est nul, donc les inégalités (*) sont des égalités et par suite $y, z \in F$. □

Il est clair que si $C_1 \subset C$ et $x \in C_1 \cap \operatorname{Ext} C$, alors $x \in \operatorname{Ext} C_1$. L'inverse n'est pas toujours vrai mais on a le résultat suivant, qui est une application immédiate des définitions :

Lemme 1.4.11. *Si F est une face de C alors $\operatorname{Ext} F \subset \operatorname{Ext} C$; et plus généralement toute face de F est une face de C.*

Il faut souligner qu'il existent des faces des convexes qui ne sont pas de la forme donnée par Lemme 1.4.10 (voir Exemple 1.4.13 ci-dessous), qui justifie la définition suivante.

Définition 1.4.12 (Points exposés at faces exposées)**.** *Une face de la forme donnée par Lemme 1.4.10 est appelée une face exposée de C. De la même façon, $x \in C$ tel que, pour un certain hyperplan d'appui H de C on a $H \cap C = \{x\}$ est appelé un point exposé de C.*

Exemple 1.4.13. *Soit $D = \big\{(x_1, x_2) \in \mathbb{R}^2 / \ x_1^2 + x_2^2 \leq 1\big\}$ le disque unité et soit $y = (2, 0) \in \mathbb{R}^2$, alors il existent deux points extrémaux du convexe $C = \operatorname{conv}(D \cup \{y\})$ qui ne sont pas exposés.*

1.5 Fonctions convexes sur un espace vectoriel

Soit E est un espace vectoriel réel.

Définition 1.5.1 (Fonction convexe)**.** *Une fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ est dite convexe si pour tous x, y dans E avec $f(x) < +\infty, f(y) < +\infty$ et tout $\theta \in [0, 1]$, on a :*

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

On remarque aisément que si f est convexe, tous les ensembles $A_f(a) = \{x \in E/\ f(x) \leq a\}$ et $B_f(a) = \{x \in E/\ f(x) < a\}$ sont convexes ; la réciproque est fausse : si $f :\ \mathbb{R} \to \mathbb{R}$ est seulement supposée monotone, les ensembles $A_f(a)$, $B_f(a)$ sont convexes car ce sont des intervalles.

Définition 1.5.2 (Domaine effectif d'une fonction)**.** *Le domaine effectif d'une fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ est l'ensemble*

$$\mathrm{dom}\ f = \big\{x \in E\ /\ f(x) < +\infty\big\}$$

Remarque 1.5.3. La fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ est convexe si et seulement si
1) son domaine effectif est convexe
2) sa restriction $f|_{\mathrm{dom}\ f}$ à son domaine effectif est convexe.

Définition 1.5.4 (Fonction strictement convexe)**.** *f est dite strictement convexe (sur son domaine effectif) si c'est une fonction convexe et si pour tous $x \neq y$ dans* dom f *et tout $\theta \in]0, 1[$, on a l'inégalité stricte :*

$$f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y)$$

Proposition 1.5.5. *Soit C un convexe d'un espace vectoriel et $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction strictement convexe dont le domaine effectif contient C. Si f atteint son maximum sur C en un point $x_0 \in C$, alors x_0 est un point extrémal de C.*

Démonstration : Soit M la borne supérieure de f sur C. Si $x_0 \in C$, avec $f(x_0) = M$, n'est pas extrémal, soient $y, z \in C$ distincts tels que $x_0 = (y+z)/2$. On a $M = f(x_0) < (f(y)+f(z))/2 \leq \max(f(y), f(z))$ donc f prend une valeur supérieure à M en l'un des points y ou z, ce qui constitue une contradiction. □

On verra des propriétés analogues sans hypothèse de stricte convexité dans les chapitres suivants.

Lorsqu'on la restreint à son domaine effectif, une fonction convexe devient une fonction convexe usuelle définie sur un convexe et à valeurs réelles. Inversement, on a :

Proposition 1.5.6 (Prolongement convexe)**.** *Si $f : C \to \mathbb{R}$ est une fonction convexe définie sur le convexe C, on peut l'étendre en une fonction convexe $\tilde{f} : E \to \mathbb{R} \cup \{+\infty\}$ en posant*

$$\tilde{f}(x) = \begin{cases} f(x) \ \text{si } x \in C \\ +\infty \ \text{sinon} \end{cases}$$

Cet artifice permet de manier simultanément des fonctions convexes $f, g, \ldots$ définies sur des convexes différents, et de considérer $f + g$, $\max(f, g)$,...

En particulier si A est une partie de E, on définit sa fonction caractéristique par :

$$\chi_A(x) = \begin{cases} 0 & \text{si } x \in A \\ +\infty & \text{sinon} \end{cases}$$

Alors l'ensemble A est convexe si et seulement si la fonction χ_A est convexe.

Définition 1.5.7 (Epigraphe d'une fonction)**.** *Si f est une fonction $E \to \mathbb{R} \cup \{+\infty\}$, son épigraphe est le sous-ensemble de $E \times \mathbb{R}$ défini par :*

$$\text{epi } (f) = \big\{(x, t) \in E \times \mathbb{R} \ / \ f(x) \leq t\big\}$$

l'épigraphe strict de f est

$$\text{epist } f = \big\{(x, t) \in E \times \mathbb{R} \ / \ f(x) < t\big\}$$

Remarque 1.5.8. *1)* L'épigraphe de f est un sous-ensemble de dom $f \times \mathbb{R}$.
2) La projection naturelle $E \times \mathbb{R} \to E$, $(x, t) \mapsto x$ envoie epi f exactement sur dom f.

Proposition 1.5.9. *La fonction $f : \ E \to \mathbb{R} \cup \{+\infty\}$ est convexe si et seulement si son épigraphe est convexe.*

Démonstration : Sif est convexe, alors pour tous (x, t), (y, u) dans epi (f), et tout $\theta \in [0, 1]$, on a $f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y) \leq \theta t + (1 - \theta)u$, donc $(\theta x + (1 - \theta)y, \theta t + (1 - \theta)u)$ appartient à epi f, qui est donc convexe.

Réciproquement, si epi f est convexe, alors pour tous $x, y \in$ dom f, et $\theta \in [0, 1]$, on a $(x, f(x)) \in$ epi f, $(y, f(y)) \in$ epi f, donc $(\theta x + (1 - \theta)y, \theta f(x) + (1 - \theta)f(y)) \in$ epi f, ce qui signifie que $f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$, et donc f est convexe. $\square$

Proposition 1.5.10. *1) Si $(f_i)_{i \in I}$ est une famille quelconque de fonctions convexes, l'enveloppe supérieure des $(f_i)_{i \in I}$, définie par $f = \sup_{i \in I} f_i$, est convexe.*
2) Si g est limite pour la convergence simple des fonctions d'une suite de fonctions convexes, g est convexe.

Dans certaines questions on peut avoir besoin de considérer des fonctions convexes à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. On prend alors la propriété de convexité de l'épigraphe comme définition :

Définition 1.5.11 (Fonction convexe à valeurs dans $\overline{\mathbb{R}}$)**.** *La fonction $f : E \to \overline{\mathbb{R}}$ est dite convexe si son épigraphe* epi $f = \big\{(x, t) \in E \times \mathbb{R} / \ f(x) \leq t\big\}$ *est convexe.*

Remarque 1.5.12. Il est équivalent de dire que l'épigraphe strict de f est convexe.

DÉMONSTRATION : En effet soit T_ε la translation $(x,t) \to (x, t+\varepsilon)$ dans $E \times \mathbb{R}$: elle transforme un convexe en un convexe. L'ensemble epist f s'écrit comme réunion de la famille croissante des ensembles $T_\varepsilon(\text{epi } f)$, $\varepsilon > 0$, tandis que epi f s'écrit comme intersection de la famille d' ensembles $T_\varepsilon(\text{epist } f)$, $\varepsilon < 0$. □

La définition 1.5.11 coïncide avec la définition usuelle lorsque f est à valeurs $> -\infty$. Lorsque ce n'est pas le cas, on a la caractérisation suivante :

Lemme 1.5.13. *La fonction $f : E \to \overline{\mathbb{R}}$ est convexe si et seulement si pour tous $x, y \in E$ tels que $f(x), f(y) < +\infty$ et pour tout $\lambda \in [0,1]$, on a l'inégalité de convexité*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

avec les conventions usuelles $-\infty + t = -\infty$, $\forall t < +\infty$, $a \cdot (-\infty) = -\infty$, $\forall a > 0$, et $0 \cdot (-\infty) = 0$.

Notons que dans cette caractérisation, les points où f prend la valeur $+\infty$ n'interviennent pas.

Remarque 1.5.14. La Proposition 1.5.10 reste vraie pour les fonctions à valeurs dans $\overline{\mathbb{R}}$.

Définition 1.5.15 (Fonction convexe propre)**.** *Une fonction convexe est dite propre si elle ne prend pas la valeur $-\infty$, et n'est pas la constante $+\infty$. Dans le cas contraire, la fonction est dite impropre.*

Chapitre 2
Espaces vectoriels normés, topologie des convexes et fonctions s.c.i.

2.1 Rappels de topologie des espaces vectoriels normés

Soit $(E, \|\cdot\|)$ un espace vectoriel normé sur $\mathbb{R}$ ou $\mathbb{C}$, en bref un EVN. E est donc un cas particulier d'espace métrique où la distance est donnée par

$$d(x, y) = \|x - y\|$$

Définition 2.1.1 (Normes équivalentes)**.** *Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur un EVN E sont dites équivalentes s'il existe deux constantes $a, b > 0$ telles que*

$$\forall x \in E \ , \ a \|x\|_1 \leq \|x\|_2 \leq b \|x\|_1$$

Pour $r > 0$, on notera $B(x_0, r) = \{x \in E \ / \ \|x - x_0\| < r\}$, la boule ouverte de centre x_0 et de rayon r et $\bar{B}(x_0, r) = \{x \in E \ / \ \|x - x_0\| \leq r\}$, la boule fermée de centre x_0 et de rayon r.

Exemples 2.1.2. *1) Soit $p \in \mathbb{N}$. L'espace vectoriel $\mathbb{R}^p$ peut être muni des normes $\|x\|_1 = \sum_{i=1}^{p} |x_i|$, $\|x\|_\infty = \sup_{i=1,\ldots,p} |x_i|$, $\|x\|_2 = \left(\sum_{i=1}^{p} {x_i}^2\right)^{1/2}$, où $x = (x_1, \ldots, , x_p)$.*
2) Si E_1 et E_2 sont deux EVN munis des normes respectives $\|\cdot\|_1$ et $\|\cdot\|_2$, l'espace produit $E_1 \times E_2$ est un EVN munis de l'une des normes (équivalentes) suivantes

$$\|x_1\|_1 + \|x_2\|_2 \ , \ \sup(\|x_1\|_1 , \|x_2\|_2) \ , \ \left(\|x_1\|_1^2 + \|x_2\|_2^2\right)^{1/2},$$

où $x = (x_1, x_2) \in E_1 \times E_2$.

Plus généralement, on peut définir des espaces analogues en dimension infinie :

Exemples 2.1.3. *ℓ_1, c_0, ℓ_∞, ℓ_2. Ce sont des espaces de suites de nombres réels définis par :*

$$\ell_1 = \left\{x = (x_n)_{n\in\mathbb{N}} \text{ telles que } \sum_{i=1}^{\infty} |x_i| < +\infty\right\}, \text{ muni de la norme : } \|x\|_1 = \sum_{i=1}^{\infty} |x_i|$$

$$\ell_2 = \Big\{ x = (x_n)_{n\in\mathbb{N}} \text{ telles que } \sum_{i=1}^{\infty} {x_i}^2 < +\infty \Big\}, \text{ muni de la norme : } \|x\|_2 = \Big(\sum_{i=1}^{\infty} {x_i}^2 \Big)^{1/2}$$

$$c_0 = \Big\{ x = (x_n)_{n\in\mathbb{N}} \text{ telles que } x_n \to_{n\to\infty} 0 \Big\}, \text{ muni de la norme : } \|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$$

$$\ell_\infty = \Big\{ x = (x_n)_{n\in\mathbb{N}} \text{ telles que } \sup_{i\in\mathbb{N}} |x_i| < +\infty \Big\}, \text{ muni de la norme : } \|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$$

Dans les exemples ci dessus, 2.1.2 et 2.1.3, en dimension finie ou non, il est très facile de démontrer qu'on a bien des normes, sauf dans le cas de $\|\cdot\|_2$. Dans le cas de cette dernière norme, seule l'inégalité triangulaire n'est pas évidente à démontrer : il faut utiliser le fait qu'elle provient du produit scalaire défini par $\langle x, y\rangle = \sum_{i=1}^{p} x_i y_i$ sur $\mathbb{R}^p$ et par $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i y_i$ dans le cas de ℓ_2.
Rappelons l'inégalité de **Cauchy-Schwarz** qui s'applique dans ce cadre :

$$|\langle x, y\rangle| \leq \|x\|_2 \, \|y\|_2$$

Cette inégalité implique alors facilement l'inégalité de **Minkowski** :

$$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$$

En effet, il suffit d'écrire :

$$\|x + y\|_2^2 = \|x\|_2^2 + 2\langle x, y\rangle + \|y\|_2^2 \leq \|x\|_2^2 + 2\,\|x\|_2\,\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$$

ce qui prouve le résultat.

Définition 2.1.4 (Espace euclidien et espace préhilbertien)**.** *Un espace vectoriel muni d'un produit scalaire est appelé un espace euclidien s'il est de dimension finie et un espace préhilbertien sinon.*

Définition 2.1.5 (Ensemble ouvert)**.** *Une partie $A \subset E$ d'un EVN est ouverte si ou bien $A = \emptyset$ ou bien*

$$\forall x_0 \in A \ , \ \exists r > 0 \text{ tel que } B(x_0, r) \subset A$$

On remarquera que l'espace E lui-même est ouvert.

Proposition 2.1.6. *Toute réunion d'ouverts et toute intersection finie d'ouverts est ouverte.*

Démonstration : Si $x \in \bigcup_{i\in I} A_i$, où les A_i sont ouverts, il existe un indice $i \in I$ tel que $x \in A_i$. Par suite il existe une boule $B(x,r)$, avec $r > 0$ incluse dans A_i et donc dans $\bigcup_{i\in I} A_i$, ce qui prouve bien que $\bigcup_{i\in I} A_i$ est ouvert.

Pour montrer qu'une intersection finie d'ouverts est ouverte, il suffit de remarquer qu'une intersection finie de boules de même centre et de rayons non nuls est encore une boule de même centre et de rayon non nul. □

Définition 2.1.7 (Ensemble fermé)**.** *Une partie $A \subset E$ d'un EVN est fermée si son complémentaire est ouvert.*

Proposition 2.1.8. *Toute intersection et toute réunion finie de fermés est fermée.*

Démonstration : On passe aux complémentaires dans la Proposition 2.1.6. □

Définition 2.1.9. *Une suite de vecteurs $(x_n)_{n\in\mathbb{N}}$ d'un EVN converge vers un vecteur $x \in E$ si et seulement si $\|x_n - x\| \to 0$ quand $n \to \infty$.*

Remarque 2.1.10. Soient E_1 et E_2 sont deux EVN. Dans l'espace vectoriel produit $E = E_1 \times E_2$ muni de l'une des normes définies ci-dessus, 2.1.2, on remarque qu'une suite $(x_n)_{n\in\mathbb{N}} = ((y_n, z_n))_{n\in\mathbb{N}}$ converge vers un point $x = (y, z)$ si et seulement si $y_n \to y$ et $z_n \to z$ quand $n \to \infty$.

Exemples 2.1.11. *On peut considérer $\mathbb{R}^p$ comme le produit de p copies de $\mathbb{R}$. La convergence d'une suite de vecteurs pour l'une des normes définies dans l'exemple 2.1.2 est donc la convergence coordonnées par coordonnées.*

On a alors une autre caratérisation des parties fermées en termes de convergence de suites :

Proposition 2.1.12. *Dans un EVN E, une partie A est fermée si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}} \subset A$, convergeant vers $x \in E$, alors $x \in A$.*

Démonstration : Supposons que A est fermée soit soit $(x_n)_{n\in\mathbb{N}}$ une suite de A convergeant vers $x \in E$. Si $x \notin A$, alors, comme A^c est ouvert, il existe une boule $B(x,r)$, avec $r > 0$ incluse dans A^c. Cette boule ne peut donc contenir aucun point de la suite $(x_n)_{n\in\mathbb{N}}$, ce qui contredit la convergence de cette suite vers x. Donc $x \in A$.

Réciproquement, supposons que pour toute suite $(x_n)_{n\in\mathbb{N}} \subset A$, convergeant vers $x \in E$, alors $x \in A$ et que A^c n'est pas ouvert. Il existe donc $x \in A^c$

tel que pour tout $n \in \mathbb{N}$, $B(x, \frac{1}{n})$ rencontre A. Pour chaque $n \in \mathbb{N}$, on choisit un point $x_n \in B(x, \frac{1}{n})$. La suite ainsi construite converge vers $x \notin A$, ce qui constitue une contradiction donc A^c est bien ouvert. □

Définition 2.1.13 (Adhérence d'une partie)**.** *1) Un point x est adhérent à un sous-ensemble A d'un EVN E s'il est limite d'une suite de points de A.*
2) L'adhérence ou fermeture $\overline{A}$ de A est l'ensemble des points adhérents à A.

Proposition 2.1.14. *L'ensemble A est fermé si et seulement s'il contient tous ses points adhérents, c'est à dire $A = \overline{A}$.*

Démonstration : On remarque d'abord que, par définition, on a toujours $A \subset \overline{A}$ et que par la Proposition 2.1.12, $\overline{A}$ est fermé. Supposons que A est fermé. Alors, d'après la Proposition 2.1.12, A contient tous ses points adhérents et donc $\overline{A} \subset A$. D'où l'égalité.

Réciproquement, si $A = \overline{A}$, A est évidemment fermé. □

Définition 2.1.15. *Soit A une partie d'un EVN E et $x \in E$. La distance de x à A est*

$$d(x, A) = \inf \left\{ \|x - y\| \ , \ y \in A \right\}$$

Proposition 2.1.16. *1)* $\overline{A} = \{x \in E \ , \ d(x, A) = 0\}$.
2) A est fermé si et seulement si $\forall x \notin A$, $d(x, A) > 0$.

Démonstration : *1)* La fonction $d(\cdot, A)$ est évidemment continue et donc $\{x \in E \ , \ d(x, A) = 0\}$ est fermé. Comme cet ensemble contient A, il contient aussi $\overline{A}$.
Soit $x \in E$ tel que $d(x, A) = 0$. Il existe donc une suite $(x_n)_{n \in \mathbb{N}} \subset A$ qui converge vers x. Par définition, $x \in \overline{A}$. D'où l'autre inclusion $\{x \in E \ , \ d(x, A) = 0\} \subset \overline{A}$.

2) A est fermé si et seulement si $A = \overline{A} = \{x \in E \ , \ d(x, A) = 0\}$. Donc A est fermé si et seulement si $\forall x \notin A$, $d(x, A) \neq 0$. □

Définition 2.1.17 (Intérieur d'une partie)**.** *1) Un point x est intérieur à une partie A d'un EVN E s'il existe une boule $B(x, r)$ de centre x, de rayon $r > 0$ incluse dans A.*
2) L'intérieur $\overset{o}{A}$ de A est l'ensemble des ses points intérieurs.

Remarque 2.1.18. Il revient au même de dire qu'un point x est intérieur à une partie A et que x n'est pas adhérent au complémentaire A^c de A.

Démonstration : En effet si x est intérieur à une partie A, il existe une boule de centre x et de rayon $r > 0$ incluse dans A donc dont l'intersection avec A^c est vide. Par suite aucune suite de A^c ne peut appartenir à cette boule, donc ne peut converger vers x.
Réciproquement si une suite de A^c converge vers x, toute boule centrée en x contient des points de cette suite et ne peut donc pas être incluse dans A. Donc x ne peut être intérieur à A. □

Proposition 2.1.19. *L'ensemble A est ouvert si et seulement s'il est égal à son intérieur.*

Démonstration : Il suffit de remarquer que $\overline{A}^c = \overset{o}{(A^c)}$ et de passer aux complémentaires dans la Proposition 2.1.14. □

Proposition 2.1.20. *Soit $A \in E$;*
1) L'adhérence de A est le plus petit fermé contenant A.
2) L'intérieur de A est le plus grand ouvert contenu dans A.

Démonstration : *1)* $\overline{A}$ est un fermé qui contient A. Si F est un fermé contenant A, F contient les limites de suites convergentes de points de A, il contient donc $\overline{A}$. $\overline{A}$ est donc bien le plus petit fermé contenant A.
2) On passe aux complémentaires dans *1)*. □

Définition 2.1.21 (Compact)**.** *Une partie K d'un EVN est compacte si de toute suite de points de K on peut extraire une sous-suite convergeant dans K.*

Exemple 2.1.22. *$E = \mathbb{R}$: les intervalles fermés bornés sont compacts.*

Démonstration : Rappelons que dans $\mathbb{R}$ toute ensemble majoré a une borne supérieure. Si $(x_n)_{n\in\mathbb{N}}$ est une suite dans l'intervalle $[a, b]$ posons pour tout entier p :

$$M_p = \sup \{x_n \ / \ n \geq p\}$$

On a $a \leq M_p \leq b$ et la suite M_p est décroissante, minorée par a, donc admet une limite M. Pour tout entier $k \geq 1$ on peut trouver un entier $p(k) \geq k$ tel que $M_{p(k)} < M + \dfrac{1}{k}$ puis un entier $n(k) \geq p(k)$ tel que $x_{n(k)} > M_{p(k)} - \dfrac{1}{k}$. Alors $n(k) \geq k$ et $M - \dfrac{1}{k} < x_n < M + \dfrac{1}{k}$ donc $n(k) \to \infty$ et $x_{n(k)} \to M$ quand $k \to \infty$.
On peut trouver $k_1 < k_2 < \ldots < k_i < k_{i+1} \ldots$ tels que $n(k_1) < n(k_2) < \ldots < n(k_i) < n(k_{i+1}) \ldots$. La suite $(x_{n(k_i)})_{i\geq 1}$ est une suite extraite de (x_n) et convergente vers $M \in [a, b]$. □

Proposition 2.1.23. *1) Toute partie fermé d'un compact est compacte.*
2) L'image d'un compact par une application continue est compacte.
3) Si K est un compact et $f : K \to \mathbb{R}$ une application continue alors f est bornée et atteint ses bornes.
4) Tout compact d'un EVN est fermé et borné.

DÉMONSTRATION : *1)* Si K est compact et $F \subset K$ est fermé, de toute suite $(x_n)_{n\in\mathbb{N}}$ de points de F on peut extraire une sous suite convergeant vers un point x de K. Alors $x \in F$ puisque F est fermé.

2) Soit K un compact de l'EVN E_1 et $f : K \to E_2$ une application continue. Soit (y_n) une suite de points de $f(K)$: pour tout n il existe donc $x_n \in K$ tel que $y_n = f(x_n)$.
Soit $(x_{n_k})_{k\geq 1}$ une suite extraite convergeant vers $x \in K$.
Par continuité on a $y_{n_k} = f(x_{n_k}) \to f(x) \in f(K)$.

3) Soit $M = \sup_{x\in K} f(x)$ (éventuellement $M = +\infty$). Il existe une suite $(x_n)_{n\in\mathbb{N}} \in K$ telle que $f(x_n) \to M$ quand $n \to \infty$.
On extrait une sous suite $(x_{n_k})_{k\in\mathbb{N}}$ convergeant vers $x \in K$. Alors par continuité $(f(x_{n_k}))_{k\in\mathbb{N}}$ converge vers $f(x)$ et donc $M = f(x)$. En particulier $M < +\infty$ et f atteint son maximum M sur K au point x.
Le raisonnement est le même pour le minimum.

4) En appliquant le point *3)* à la fonction continue $f(x) = \|x\|$ on voit que K est borné.
Si une suite $(x_n)_{n\in\mathbb{N}}$ de points de K converge vers un point x de E il en est de même de toutes ses sous-suites. Or l'une d'entre elles converge vers un point de K donc $x \in K$ et par suite K est fermé. □

Proposition 2.1.24. *Si E_1 et E_2 sont deux EVN et si K_1 et K_2 sont des compacts de, respectivement, E_1 et E_2 alors $K_1 \times K_2$ est un compact de $E_1 \times E_2$.*

DÉMONSTRATION : Soit $(x_n)_{n\in\mathbb{N}}$ une suite de points de $K_1 \times K_2$. On a $x_n = (y_n, z_n)$ avec $y_n \in K_1$ et $z_n \in K_2$.
Soit $(y_{n(k)})_{k\in\mathbb{N}}$ une suite extraite de $(y_n)_{n\in\mathbb{N}}$ qui converge vers $y \in K_1$. Posons $\eta_k = y_{n_k}$ et $\zeta_k = z_{n(k)}$.
Soit $(\zeta_{k(m)})_{m\in\mathbb{N}}$ une suite extraite convergeant vers $z \in K_2$. La suite $(\eta_{k(m)})_{m\in\mathbb{N}}$ étant extraite de la suite convergente $(\eta_k)_{k\in\mathbb{N}}$ converge vers le même point y.
Donc $(\eta_{k(m)}, \zeta_{k(m)})_{m\in\mathbb{N}} \to (y, z) \in K_1 \times K_2$ lorsque $m \to \infty$. Or c'est une suite extraite de $(x_n)_{n\in\mathbb{N}}$ et $K_1 \times K_2$ est bien compact. □

Lemme 2.1.25. *Munissons l'espace vectoriel $\mathbb{R}^p$ de la norme $\|(x_1, \ldots, x_p)\|_1 = \sum_{i=1}^{p} |x_i|$. Alors les fermés bornés de $\mathbb{R}^p$ muni de cette norme sont compacts.*

DÉMONSTRATION : D'après la proposition précédente, pour tout $M > 0$ l'ensemble $[-M, M]^n = [-M, M] \times [-M, M] \times \ldots \times [-M, M]$ est compact dans $\mathbb{R}^p$ muni de la norme $\| \|_1$. Si K est un fermé borné de $\mathbb{R}^p$, soit $M = \sup\limits_{x \in K} \|x\|_1$. On a $K \subset [-M, M]^p$; K est donc fermé dans un compact, donc compact. $\square$

Remarque 2.1.26. Lorsque deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur un espace vectoriel E sont équivalentes, toutes les notions topologiques (convergence, fermés, ouverts, compacts, continuité. . . .) sont les mêmes relativement à $\|\cdot\|_1$ ou à $\|\cdot\|_2$.

Définition 2.1.27 (EVN isomorphes ou isométriques). *Plus généralement, si E_1 et E_2 sont deux EVN, munis respectivement des normes $\|\cdot\|_1$ et $\|\cdot\|_2$ et si $T : E_1 \to E_2$ est une application linéaire bijective telle qu'il existe $a, b > 0$ vérifiant*

$$\forall x \in E_1 \ , \ a \|x\|_1 \leq \|Tx\|_2 \leq b \|x\|_1$$

on dira que E_1 et E_2 sont isomorphes et ils seront dits isométriques si $a = b = 1$.

On peut alors généraliser la remarque ci-dessus :

Remarque 2.1.28. Lorsque deux EVN E_1 et E_2 sont isomorphes, toutes les notions topologiques (convergence, fermés, ouverts, compacts, continuité. . . .) sont conservées par T et T^{-1}.

Théorème 2.1.29. *Sur un EVN de dimension finie toutes les normes sont équivalentes.*

DÉMONSTRATION : Soit E un EVN de dimension finie p, $(e_1, \ldots, e_p)$ une base de E et $\|\cdot\|$ une norme sur E. La formule

$$\left\| \sum_{i=1}^{p} \lambda_i e_i \right\|_1 = \sum_{i=1}^{p} |\lambda_i|$$

définit une autre norme $\|\cdot\|_1$ sur E.
Posons $b = \max\limits_{i=1,\ldots,p} \|e_i\|$. On a

$$\left\| \sum_{i=1}^{p} \lambda_i e_i \right\| \leq \sum_{i=1}^{p} |\lambda_i| \, \|e_i\| \leq b \sum_{i=1}^{p} |\lambda_i|$$

autrement dit $\|x\| \leq b \|x\|_1$ pour tout $x \in E$.
En particulier $\|\cdot\|$ est continue par rapport à $\|\cdot\|_1$ puisque

$$\forall x, y \in E, \quad \big| \|x\| - \|y\| \big| \leq \|x - y\| \leq b \|x - y\|_1$$

L'ensemble $S = \{x \in E \ / \ \|x\|_1 = 1\}$ est compact : en effet c'est exactement l'image du fermé borné de $\mathbb{R}^p$, muni de la norme $\|\cdot\|_1$ du Lemme 2.1.25, $\Sigma = \{(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p \ / \ \sum_{i=1}^{p} |\lambda_i| = 1\}$ par l'application continue $(\lambda_1, \ldots, \lambda_p) \mapsto \sum_{i=1}^p \lambda_i e_i$.

La fonction continue $\|\cdot\|$ atteint donc son minimum sur S ; autrement dit il existe $x_0 \in E$ avec $\|x_0\|_1 = 1$ et $\forall x \in S, \|x\| \geq \|x_0\|$. Comme $x_0 \neq 0$ on a $\|x_0\| > 0$. Posons $a = \|x_0\|$; pour tout $x \in E$ non nul on a $\dfrac{x}{\|x\|_1} \in S$ donc :

$$\|x\| = \|x\|_1 \left\| \frac{x}{\|x\|_1} \right\| \geq a \|x\|_1$$

Si $x = 0$ on a $\|x\| = 0 = \|x\|_1$.

Finalement on a montré que toute norme $\|\cdot\|$ sur E est équivalente à $\|\cdot\|_1$. Si $\|\cdot\|, \|\cdot\|'$ sont deux normes sur E elles sont équivalentes à $\|\cdot\|_1$, donc équivalentes entre elles. □

Corollaire 2.1.30. *Tout EVN de dimension p est isomorphe à $\mathbb{R}^p$.*

DÉMONSTRATION : Comme dans la démonstration précédente, on fixe une base $(e_1, \ldots, e_p)$ et on définit une norme sur $\mathbb{R}^p$ par :

$$\|(\lambda_1, \ldots, \lambda_p)\| = \|\lambda_1 e_1 + \ldots + \lambda_p e_p\|$$

Pour cette norme, l'opérateur linéaire T tel que $T(\lambda_1 e_1 + \ldots + \lambda_p e_p) = (\lambda_1, \ldots, \lambda_p)$ est une isométrie. □

Corollaire 2.1.31. *Dans un EVN de dimension finie E, tout ensemble fermé borné est compact.*

DÉMONSTRATION : En effet, ce résultat est vrai dans $\mathbb{R}^p$ pour la norme $\|\cdot\|_1$ par le Lemme 2.1.25. Par le Théorème 2.1.29 et la Remarque 2.1.28, elle est vraie sur $\mathbb{R}^p$ pour toute norme et sur tout EVN de dimension finie par le Corollaire 2.1.30 et la Remarque 2.1.26. □

Remarque 2.1.32. Le résultat devient faux en dimension infinie.

Par exemple considérons l'espace ℓ_1 des suites $(x_n)_{n\in\mathbb{N}}$ de réels telles que $\sum_{n=1}^{\infty} |x_n| < \infty$, muni de la norme $\|(x_n)\| = \sum_{n=1}^{\infty} |x_n|$, voir l'Exemple 2.1.3 . La boule unité fermée de ℓ_1 est un fermé borné non compact. En effet notons e_n la suite de réels $(0, 0, \ldots, 0, 1, 0 \ldots)$ (le 1 en $n$ième place). On a $\|e_n\| = 1$ et $\|e_n - e_m\| = 2$ pour tous n, m entiers distincts. Donc la suite $(e_n)_{n\in\mathbb{N}}$ n'est pas de Cauchy dans ℓ_1, de même que ses sous-suites, qui ne convergent donc pas.

Remarque 2.1.33. Ce phénomène est général : le théorème de F. Riesz dit dans un EVN de dimension infinie la boule unité fermée n'est jamais compacte.

Proposition 2.1.34. *Si K est un compact de E, alors pour tout $x \in E$, il existe $x_0 \in K$ tel que*
$$d(x, K) = \|x - x_0\|$$

DÉMONSTRATION : La fonction de $K \to \mathbb{R}_+$ qui à $y \in K$ associe $\|x - y\|$ est continue. Il suffit donc d'appliquer la Proposition 2.1.23, *3)*. □

Définition 2.1.35 (Dual d'un EVN)**.** *Le dual (topologique) E' d'un espace vectoriel normé E est l'ensemble des applications linéaires continues de E dans $\mathbb{R}$, appellées formes linéaires continues.*

Proposition 2.1.36. *Soit f une forme linéaire sur un EVN E. Les propriétés suivantes sont équivalentes :*
1) f est continue.
2) f est continue à l'origine.
3) f est bornée sur la boule unité $B(0,1)$ de E.
4) $\exists M \geq 0$ tel que $\forall x \in E$, $\|f(x)\| \leq M\|x\|$.

DÉMONSTRATION : Les implications *1)* $\Rightarrow$ *2)* $\Rightarrow$ *3)* sont évidentes.
Montrons *3)* $\Rightarrow$ *4)*. Soit $M = \sup\{\|f(x)\| \ , \ \|x\| \leq 1\}$. Si $x \neq 0$, on peut écrire : $\left\| \dfrac{f(x)}{\|x\|} \right\| \leq M$ et par linéarité, ceci implique bien $\|f(x)\| \leq M\|x\|$. L'inégalité est évidente si $x = 0$.
Montrons *4)* $\Rightarrow$ *1)* : Soit M la constante donnée par *4)* et soient $x, y \in E$, alors : $\|f(x) - f(y)\| = \|f(x - y)\| \leq M\|x - y\|$. Ceci prouve la continuité de f sur E. □

Définition 2.1.37 (Norme d'une forme linéaire)**.** *La norme d'une forme linéaire continue sur un EVN est définie par*
$$\|f\| = \sup_{x \in E, \|x\| \leq 1} |f(x)|$$

Proposition 2.1.38. *Le dual E' d'un EVN E est un EVN, la norme de E' est donnée par :*
$$\forall f \in E' \ , \ \|f\| = \sup_{x \in E, \|x\| \leq 1} |f(x)|$$

Notations 2.1.39. *Les élément du dual E' de E seront notés de deux façons différentes selon le contexte :*
i) ou bien la notation fonctionnelle $f : E \to \mathbb{R}$
ii) ou bien la notation vectorielle par analogie avec le cas préhilbertien (Chapitre 5) : $y \in E'$ agit sur les éléments de E par le crochet de dualité $\langle y, x \rangle$, pour $x \in E$.

Propriétés 2.1.40. *1) Dans un espace vectoriel de dimension finie, toutes les formes linéaires sont continues.*
2) Le dual d'un EVN de dimension finie est un EVN de même dimension.
3) Tout sous-espace affine d'un EVN de dimension finie est fermé.

Démonstration : Soit E un EVN de dimension p.
1) On fixe une base $(e_1, \dots, e_p)$ de E et on utilise la norme $\|\cdot\|_1$ définie par

$$\left\|\sum_{i=1}^{p} \lambda_i e_i\right\|_1 = \sum_{i=1}^{p} |\lambda_i| \,.$$

Si f est une forme linéaire sur E, on écrit :

$$\begin{aligned} \left\| f\Big(\sum_{i=1}^{p} \lambda_i e_i\Big)\right\|_1 &= \left\|\sum_{i=1}^{p} \lambda_i f(e_i)\right\|_1 \\ &\leq \big[\sup_{i=1,\dots,p} \|f(e_i)\|\big] \sum_{i=1}^{p} |\lambda_i| = \big[\sup_{i=1,\dots,p} \|f(e_i)\|\big] \left\|\sum_{i=1}^{p} \lambda_i e_i\right\|_1 \end{aligned}$$

Donc, par la Proposition 2.1.36, f est continue de norme inférieure ou égale à $\sup\limits_{i=1,\dots,p} \|f(e_i)\|$. Par Théorème 2.1.29, f est continue par rapport à toute norme.

2) On fixe une base $(e_1, \dots, e_p)$ de E et on introduit la base duale $(e_1^*, \dots, e_p^*)$, définie pour $i, j = 1, \dots, p$ par $e_i^*(e_j) = \delta_i j$, où $\delta_i j$ est le symbol de Kronecker qui vaut 1 si $i = j$ et 0 sinon. $(e_1^*, \dots, e_p^*)$ est une partie libre de E' car si $\sum\limits_{i=1}^{p} \lambda_i e_i^* = 0$, on applique cette forme linéaire à e_j pour chaque $j = 1, \dots, p$, ce qui implique $\lambda_j = 0$. C'est aussi une partie génératrice de E' car si $f \in E'$, on note $\alpha_j = f(e_j)$ pour $j = 1, \dots, p$ et on peut écrire, pour $x = \sum\limits_{i=1}^{p} \lambda_i e_i$, $f(x) = f\big(\sum\limits_{i=1}^{p} \lambda_i e_i\big) = \sum\limits_{i=1}^{p} \lambda_i f(e_i) = \sum\limits_{i=1}^{p} \lambda_i \alpha_i$. Or, par définition de la base duale, $\lambda_j = e_j^*(x)$ pour tout $j = 1, \dots, p$ et par suite $f(x) = \sum\limits_{i=1}^{p} e_i^*(x) \alpha_i = \big(\sum\limits_{i=1}^{p} \alpha_i e_i^*\big)(x)$. D'où l'égalité $f = \sum\limits_{i=1}^{p} \alpha_i e_i^*$.
La base duale $(e_1^*, \dots, e_p^*)$ est donc bien une base de E', ce qui implique en particulier que E' est de dimension p.

3) Il suit de la partie *1)* que tout hyperplan $H \subset E$ est fermé. Soit M un sous-espace vectoriel de E de dimension k, soit $(e_1, \dots, e_p)$ une base de E telle que $(e_1, \dots, e_k)$ est une base de M, et soit $(e_1^*, \dots, e_p^*)$ la base duale. Alors $M = \bigcap_{i=k+1}^{p} H_i$, où $H_i = \{x \in E \ / \ e_i^*(x) = 0\}$, et il suit que M est fermé (une intersection de hyperplans fermés). Si $Y \subset E$ est un sous-espace affine, M la direction de Y et $x_0 \in Y$, il suit que $Y = x_0 + M$ (voir Proposition 1.3.3) et alors Y est fermé. □

Remarque 2.1.41. Plus généralement, tout SEA de dimension finie (d'un EVN de dimension quelconque) est fermé. Effectivement, soit $Y \subset E$ avec $\dim Y < \infty$ et soit (y_n) une suite dans Y qui converge (vers la limite a qui a priori peut être dans $E \setminus V$). Alors (y_n) est bornée, disons $\|y_n\| \leq R$ pour tout $n \in \mathbb{N}$. Donc (y_n) est une suite dans un ensemble compact $K = \{y \in Y \ / \ \|y\| \leq R\}$ et sa limite appartient forcément à $K \subset Y$.

Dans un espace qui n'est pas de dimension finie, il existe des formes linéaires qui ne sont pas continues.

Exemple 2.1.42. *Soit $E = c_{00}$ l'espace des suites de réels nulles à partir d'un certain rang, muni de la norme $\|\cdot\|_\infty$. La forme linéaire f sur c_{00}, définie par $f\big((x_n)_{n\in\mathbb{N}}\big) = \sum_{i=1}^{\infty} x_i$ n'est pas continue.*

En effet, f est bien définie et linéaire sur E mais n'est pas bornée sur la boule unité car en considérant la suite finie de réels comportant n termes non nuls : $s_n = (1, 1, 1, \dots, 1, 0, 0, \dots)$, on a $\|s_n\|_\infty = 1$ et $f(s_n) = n \to +\infty$.

Définition 2.1.43 (Espace EVN separable)**.** *On dit qu'un EVN E est séparable s'il existe une suite d'espaces vectoriels de dimension finie dont la réunion est dense dans E.*

On remarquera qu'il est équivalent de dire qu'il existe une suite $(x_n)_{n\in\mathbb{N}}$, dense dans E (la définition habituelle de la séparabilité).

On peut montrer que ℓ_1, ℓ_2, c_0 sont séparables mais que ℓ_∞ ne l'est pas.

Dans les EVN, on a des exemples naturels de convexes :

Exemples 2.1.44. *(Boules de normes)*
Soit $\|\cdot\|$ une norme sur E. Les boules de E, ouvertes ou fermées, $B(x_0, r)$, $\bar{B}(x_0, r)$, pour $r > 0$ sont des convexes.

Les boules centrées à l'origine sont des convexes particuliers, en ce qu'ils sont symétriques, c'est à dire que si $x \in B(0, r)$ alors $-x \in B(0, r)$ et en particulier $0 \in B(0, r)$.

Plus généralement, considérons une application $p : E \to \overline{\mathbb{R}_+}$, appelée jauge, homogène, sous-additive, nulle en 0, c'est à dire qui vérifie :
i) $p(\lambda x) = \lambda p(x)$ $(\forall x \in E, \forall \lambda > 0)$
ii) $p(x + y) \leq p(x) + p(y)$ $(\forall x, y \in E)$.
iii) $p(0) = 0$.
Alors les ensembles $\{x \in E \ / \ p(x) < 1\}$ et $\{x \in E \ / \ p(x) \leq 1\}$ sont des convexes contenant 0.

Inversement, à tout convexe contenant 0, on peut associer une fonction jauge :

Définition 2.1.45 (Jauge d'un convexe)**.** *Si C est un convexe contenant* 0*, sa jauge est la fonction $j_C : \ E \to \overline{\mathbb{R}}_+$ définie par :*

$$j_C(x) = \inf\{\lambda > 0 \ / \ \frac{x}{\lambda} \in C\}$$

(avec la convention $\inf \emptyset = +\infty$). Alors j_C vérifie les propriétés i) à iii) ci-dessus, et de plus :

$$\big\{x \in E \ / \ j_C(x) < 1\big\} \subset C \subset \big\{x \in E \ / \ j_C(x) \leq 1\big\}.$$

On remarquera :
1) La jauge j_C d'un convexe C est une jauge au sens général défini ci-dessus.
2) j_C est à valeurs finies si et seulement si l'intersection de C avec toute demi-droite issue de 0 n'est pas réduite à $\{0\}$ (C est dit absorbant).
3) La fonction j_C est une norme si et seulement si de plus C est symétrique et ne contient aucune demi-droite issue de 0.

2.2 Topologie des ensembles convexes dans les EVN

Soit $(E, \|\cdot\|)$ est un espace vectoriel normé.

Proposition 2.2.1. *L'adhérence d'un convexe (resp., d'un SEA, d'un cône convexe, SEV) est un convexe (resp., un SEA, un cône convexe, un SEV).*

DÉMONSTRATION : Si C est un convexe et x et y sont adhérents à C, il existe deux suites $(x_n)_n$ et $(y_n)_n$ dans C convergeant respectivement vers x et y. Pour tout $\theta \in [0, 1]$, la suite $(\theta x_n + (1 - \theta)y_n)_n$ est dans C et converge vers $\theta x + (1 - \theta)y$, qui est donc adhérent à C. Donc $\overline{C}$ est convexe. Même preuve pour un SEA, un cône convexe ou SEV. □

Remarque 2.2.2. Le fait que $F \subset E$ soit fermé n'entraîne pas que $\mathrm{conv}(F)$ soit fermé.

En effet soit $E = \mathbb{R}^2$, F_1 la branche d'hyperbole $\{(x, y) \in \mathbb{R}^2 \ / \ x > 0, xy = 1\}$, et F_2 son symétrique par rapport au premier axe. Alors F_1, F_2 sont fermés, donc aussi $F = F_1 \cup F_2$; mais $\mathrm{conv}(F)$ est le demi-plan ouvert $\{(x, y) \ / \ x > 0, y \in \mathbb{R}\}$, qui n'est pas fermé.

Définition 2.2.3 (Enveloppe convexe fermée)**.** *L'enveloppe convexe fermée d'une partie A de E est l'adhérence de son enveloppe convexe. On la note* $\overline{\mathrm{conv}}(A)$.

Exercice 2.1. $\overline{\mathrm{conv}}(A)$ *est le plus petit convexe fermé de E contenant A.*

Proposition 2.2.4. *1) Si A est un ensemble fini,* $\mathrm{conv}(A)$ *est compacte.*
2) Plus généralement si $K_1, \ldots, K_p$ sont des convexes compacts, $\mathrm{conv}(K_1 \cup \ldots \cup K_p)$ *est compacte.*

Démonstration : Posons $T_p = \{(\lambda_1, \ldots, \lambda_p) \in [0, 1]^p \ / \ \sum_{i=1}^{p} \lambda_i = 1\}$: T_p est un compact car fermé dans $[0, 1]^p$.
1) Si $A = \{a_1, \ldots, a_p\}$, on considère l'application continue $f : T_p \to E$, $(\lambda_1, \ldots, \lambda_p) \mapsto \sum_{i=1}^{p} \lambda_i a_i$; on a $\mathrm{conv}(A) = f(T_p)$, donc est compact comme image continue d'un compact.

2) La preuve est analogue en considérant $g : T_p \times K_1 \times \cdots \times K_p \to E$, $(\lambda_1, \ldots, \lambda_p, x_1, \ldots, x_p) \mapsto \sum_{i=1}^{p} \lambda_i x_i$. □

Remarque 2.2.5. On verra au chapitre 3 que si E est de dimension finie, l'enveloppe convexe de tout compact de E est encore compacte. Toutefois ceci n'est plus vrai en dimension infinie.

Prenons par exemple pour E l'espace ℓ_∞ des suites réelles bornées, muni de la norme $\|(x_n)_n\| = \sup\limits_{n \in \mathbb{N}} |x_n|$, voir l'Exemple 2.1.3. Pour tout k soit $e_k = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$ où le 1 est placé à la k-ième place. Soit $K = \{\dfrac{e_k}{k} \ / \ k \geq 1\} \cup \{0\}$. Alors K est compact : en effet, toute suite de points de K soit contient une suite extraite constante, soit converge vers 0.

L'adhérence de $\mathrm{conv}(K)$ contient tous les points de la forme $x = \displaystyle\sum_{k=1}^{\infty} \lambda_k \frac{e_k}{k}$ tels que $\lambda_k > 0$ et $\sum\limits_{k=1}^{\infty} \lambda_k \leq 1$. En effet, un tel point est limite dans ℓ_∞ de la suite $x_n = \displaystyle\sum_{k=1}^{n} \lambda_k \frac{e_k}{k}$, qui est clairement dans $\mathrm{conv}(K)$.
Mais un tel point n'appartient pas à l'ensemble $\mathrm{conv}(K)$ lui-même, car si $y = (y_n)_{n \in \mathbb{N}}$ est dans $\mathrm{conv}(K)$, il y a seulement un nombre fini de coordonnées y_n non nulles. □

On va maintenant étudier l'intérieur d'un convexe. Commençons par une remarque :

Remarque 2.2.6. Soit C un convexe d'un EVN E , alors aucun point de l'intérieur $\overset{o}{C}$ de C n'est extrémal.

Lemme 2.2.7. *Soit C un convexe ayant un point intérieur x_0. Pour tout point x adhérent à C, tout point y du segment semi-ouvert $[x_0, x[$ est intérieur à C.*

DÉMONSTRATION : i) Supposons d'abord que $x \in C$.
Soit $B(x_0, r)$ une boule de centre x_0, de rayon $r > 0$ incluse dans C. Soit f l'homothétie de centre x, de rapport λ, amenant x_0 sur y.
On a :

$$\begin{aligned} \forall t \in E, \quad f(t) &= x + \lambda(t - x) \\ &= (1 - \lambda)x + \lambda t \end{aligned}$$

où λ est tel que $y = f(x_0) = (1 - \lambda)x + \lambda x_0$, donc $0 < \lambda \leq 1$. On a $f(C) \subset C$ par convexité de C, et $f\big(B(x_0, r)\big) = B(y, \lambda r)$ car $f(t) - y = f(t) - f(x_0) = \lambda(x - x_0)$, donc $B(y, \lambda r) \subset C$ et y est intérieur à C.

ii) Soit maintenant $x \in \overline{C}$ et $y \in]x_0, x[$.
Soit g l'homothétie de centre y amenant x_0 sur x.
On a $g(t) = y + \lambda(t - y)$; la condition $g(x_0) = x$ s'écrit $x - y = \lambda(x_0 - y)$, d'où $\lambda < 0$. Alors $g(B(x_0, r)) = B(x, |\lambda|\, r)$. Puisque x est adhérent à C, il existe $z \in C \cap B(x, |\lambda|\, r)$. Soit $u = g^{-1}(z)$. Alors $u \in B(x_0, r)$ est lui aussi intérieur à C, et $z - y = \lambda(u - y)$, $\lambda < 0$ montre que y est situé sur le segment $]u, z[$. D'après le i) ci-dessus, on voit que y est intérieur à C. □

Proposition 2.2.8. *Si C est un convexe d'intérieur non vide, alors son intérieur $\overset{o}{C}$ est convexe ; on a de plus : $\overline{\overset{o}{C}} = \overline{C}$ et $\overset{o}{\overline{C}} = \overset{o}{C}$.*

DÉMONSTRATION : i) Si $x, y \in \overset{o}{C}$, tout le segment $[x, y]$ est inclus dans $\overset{o}{C}$ d'après le Lemme précédent. Donc $\overset{o}{C}$ est convexe.

ii) L'inclusion $\overline{\overset{o}{A}} \subset \overline{A}$ est vraie pour toute partie A de E ; inversement si $x \in \overline{C}$, alors x est extrémité d'un segment ouvert inclus dans $\overset{o}{C}$, d'après le Lemme précédent, donc $x \in \overline{\overset{o}{C}}$.
On notera que l'adhérence d'un segment ouvert non vide $]x_0, x[$ est le segment fermé $[x_0, x]$.

iii) L'inclusion $\overset{o}{\overline{A}} \supset \overset{o}{A}$ est vraie pour toute partie A de E. Inversement soit x intérieur à $\overline{C}$; soit $r > 0$ tel que $B(x, r) \subset \overline{C}$. Comme $\overline{C} = \overline{\overset{o}{C}}$, d'après ii), il existe $y \in \overset{o}{C} \cap B(x,r)$. Soit $z \in E$ tel que x soit milieu du segment $[z,y]$; on a encore $z \in B(x,r)$, donc $z \in \overline{C}$; d'après le Lemme précédent, le segment $]z,y]$ est inclus dans $\overset{o}{C}$; en particulier x est intérieur à C. □

Remarque 2.2.9. Si E est de dimension infinie, l'égalité $\overline{\overset{o}{C}} = \overset{o}{C}$ peut être fausse lorsque $\overset{o}{C}$ est vide.

Par exemple, soit $E = c_0$, l'espace des suites de réels convergeant vers 0, muni de la norme $\|(x_n)_{n\in\mathbb{N}}\| = \sup_{n\in\mathbb{N}} |x_n|$, voir l'Exemple 2.1.3 et $C = c_{00}$, le sous-espace des suites nulles à partir d'un certain rang. Alors $\overline{C} = E$, mais $\overset{o}{C} = \emptyset$ car toute boule de E contient des vecteurs dont toutes les coordonnées sont non nulles.

Toutefois ce phénomène ne se produit pas en dimension finie, comme on le verra au chapitre 3.

Dans le cas des hyperplans et des demi-espaces affines en dimension infinie, on peut dire un peu plus :

Soit H un hyperplan affine, D_+ et D_- les deux demi-espaces affines définis par H, et $D'_+ = D_+ \setminus H$, $D'_- = D_- \setminus H$ les demi-espaces stricts associés, voir Définitions 1.3.11 et 1.3.16.

Proposition 2.2.10. *Si $H = \{x \in E/\ f(x) = a\}$ où f est une forme linéaire non nulle sur E et $a \in \mathbb{R}$, alors les assertions suivantes sont équivalentes :*
1) f est continue, c'est à dire $f \in E'$.
2) D'_+ est ouvert.
3) D_+ est fermé.
4) H est fermé.
5) D_+ est d'intérieur non vide.
Les propriétés analogues à 2), 3) et 5) avec D_- et D'_- sont également équivalentes aux précédentes.

Démonstration : *1) ⇒ 2)* : Si f est continue, D'_+ est évidemment ouvert.

2) ⇒ 3) : Si D'_+ est ouvert, son complémentaire D_- est fermé et D_+ également car c'est limage de D_- par n'importe quelle symétrie par rapport à H, centrée en un point de H.

3) ⇒ 4) : Si D_+ est fermé, par l'argument précédent, D_+ l'est aussi et comme $H = D_+ \cap D_-$, il est aussi fermé.

4) ⇒ 5) Supposons que H soit fermé et soit $x_0 \notin H$. Alors $f(x_0) \neq a$, par exemple $f(x_0) < a$. Comme le complémentaire H^c de H est ouvert, il existe

$r > 0$ tel que $B(x_0, r) \subset H^c$. Il est alors clair que $f\big(B(x_0, r)\big) < a$ car si $x_1 \in B(x_0, r)$ vérifie $f(x_1) \geq a$, la fonction $g(t) = f\big((1-t)x_0 + tx_1\big)$ prend les valeurs $f(x_0) < a$ en $t = 0$ et $f(x_1) \geq a$ en $t = 1$. Comme g est une fonction affine sur $\mathbb{R}$, elle prend la valeur a en un $t \in [0, 1]$. Le point $u = (1-t)x_0 + tx_1$ appartient alors à H, ce qui est contradictoire. Donc $B(x_0, r) \in D_+$ et D_+ est ouvert.

5) ⇒ 1) Soit $x \in \overset{o}{D}_+$ et $B(x, r) \subset D_+$.

On a donc : $\forall z \in B(0,1)\ ,\ f(x+rz) \geq a$, soit $-f(z) \leq \dfrac{f(x)-a}{r}$. Par linéarité, ceci implique aussi en changeant z en $-z$, $\forall z \in B(0,1)\ ,\ f(z) \leq \dfrac{f(x)-a}{r}$. Par suite, nécessairement $\dfrac{f(x)-a}{r} \geq 0$ et $\forall z \in B(0,1)\ ,\ |f(z)| \leq \dfrac{f(x)-a}{r}$. Donc f est continue de norme inférieure à $\dfrac{f(x)-a}{r}$, ce qui implique bien *1)*. □

Corollaire 2.2.11. *Si H est fermé, alors $D'_+ = \overset{o}{D}_+$ et $D'_- = \overset{o}{D}_-$.*

Démonstration : Soit $x \in \overset{o}{D}_+$. Il est clair que $x \in D'_+$ car toute boule centrée en $u \in H$ rencontre à la fois D_+ et D_- et u ne peut pas être intérieur à D_+.

Réciproquement si H est fermé, D'_+ est ouvert et inclus dans D_+ il est donc inclus dans $\overset{o}{D}_+$. □

Définition 2.2.12. *On dit qu'une partie A d'un EVN est dense dans E si son adhérence est égale à E.*

Proposition 2.2.13. *Une partie A d'un EVN est dense dans E si et seulement si son complémentaire A^c ne contient pas de boule de rayon strictement positif.*

Démonstration : En effet, si A^c contient une boule, $\overline{A}$ ne contient pas le centre de cette boule donc A n'est pas dense dans E.

Réciproquement, si A n'est pas dense dans E, il existe un point x de E qui n'est pas dans l'adhérence de A, donc qui est dans l'intérieur de A^c, d'où l'existence d'une boule centrée en x, de rayon strictement positif, qui ne rencontre pas A. □

Corollaire 2.2.14. *Un hyperplan affine est soit fermé, soit dense dans E.*

Démonstration : Si l'hyperplan H n'est pas dense dans E, le complémentaire de $\overline{H}$ contient une boule $B(x, r)$ de rayon non nul, donc $D'_+ \cup D'_-$ aussi. Cette boule ne peut intersecter à la fois D'_+ et D'_- car sinon, elle intersecterait alors H, donc est incluse par exemple dans D'_+. Donc D_+ est d'intérieur non vide, et H est fermé d'après la Proposition 2.2.10. □

Exemple 2.2.15. *On reprend l'Exemple 2.1.42 : soit $E = c_{00}$ l'espace des suites de réels nulles à partir d'un certain rang, muni de la norme $\|\cdot\|_\infty$. La forme linéaire f sur c_{00}, définie par $f\big((x_n)_{n\in\mathbb{N}}\big) = \sum_{i=1}^{\infty} x_i$ n'est pas continue. L'hyperplan $H = \{x \in E \ , \ f(x) = 0\}$ n'est donc pas fermé en appliquant l'équivalence 1) $\iff$ 4) de la Proposition 2.2.10 avec $a = 0$.*

Explicitons une suite de H qui ne converge pas dans H : on part d'un vecteur de c_{00} qui n'est pas dans H, soit $x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ avec $\alpha = \sum_{i=1}^{n} x_i \neq 0$ et on définit la suite $(\xi_k)_{k\in\mathbb{N}^*}$ par :

$$\xi_k = \Big(x_1, x_2, \ldots, x_n, -\frac{\alpha}{k}, -\frac{\alpha}{k}, \ldots, -\frac{\alpha}{k}, 0, 0, \ldots\Big)$$

où on a mis k termes en $-\dfrac{\alpha}{k}$. Alors, pour tout $k > 0$, $\xi_k \in H$ et de plus :

$$\|x - \xi_k\| = \frac{|\alpha_k|}{k} \to 0 \text{ quand } k \to +\infty$$

Donc la suite $(\xi_k)_{k\in\mathbb{N}^*}$ est dans H et converge vers $x \notin H$.

Bien entendu, en dimension finie, tous les hyperplans sont fermés.

Corollaire 2.2.16. *Si C est un convexe d'intérieur non vide, tout hyperplan d'appui (voir 1.4.9) de C est fermé. Si de plus C est fermé, l'intersection d'un hyperplan d'appui H avec C, $F = H \cap C$, est une face fermée de C.*

Démonstration : En effet, l'un des demi espaces associé à l'hyperplan d'appui contient C et est donc d'intérieur non vide, donc H est fermé par la Proposition 2.2.10. F est une face par le lemme 1.4.10. □

Dans la pratique les hyperplans d'appui d'un convexe fermé d'intérieur non vide C seront obtenus au moyen d'une forme linéaire f qui atteint son maximum et son minimum sur C. C'est le cas en particulier si f est continue et C compact. Dans ce cas, soient M et m, le maximum et le minimum de f sur C. Alors $H_M = \{x \in E \ / \ f(x) = M\}$ et $H_m = \{x \in E \ / \ f(x) = m\}$ sont des hyperplans d'appui de C.

2.3 Fonctions s.c.i. sur un EVN

Soit $(E, \|\cdot\|)$ un EVN. On sait qu'une caractérisation classique de la continuité d'une fonction numérique définie sur E, est la suivante :

Une fonction $f : E \to \mathbb{R}$ est continue si et seulement si l'image réciproque de tout intervalle ouvert de $\mathbb{R}$ est un ouvert de E.

On passe à la notion plus faible de semi-continuité inférieure si l'on se borne à demander que les images réciproques des intervalles ouverts du type $]a, +\infty[$ soient ouverts.

Si f est autorisée à prendre les valeurs $\pm\infty$, ce sont les intervalles $]a, +\infty]$ que l'on considère. On a donc :

Définition 2.3.1 (Fonction s.c.i.)**.** *On dit que la fonction numérique $f : E \to \overline{\mathbb{R}}$ est semi-continue inférieurement (s.c.i.) si pour tout $a \in \overline{\mathbb{R}}$, l'ensemble $\{x \in E \ / \ f(x) > a\}$ est un ouvert de E. Il revient au même de dire que pour tout $a \in \overline{\mathbb{R}}$, l'ensemble $\{x \in E \ / \ f(x) \leq a\}$ est un fermé de E.*

En effet les ensembles $\{x \in E \ / \ f(x) > a\}$ et $\{x \in E \ / \ f(x) \leq a\}$ sont complémentaires ; le premier est ouvert si et seulement si le second est fermé.

Exemple 2.3.2. *1) Un sous-ensemble A de E est fermé si et seulement si sa fonction caractéristique χ_A est s.c.i.*
2) Toute fonction continue est s.c.i.
3) Une fonction $f : E \to \overline{\mathbb{R}}$ est continue si et seulement si f et $-f$ sont s.c.i.

Démonstration : Pour montrer *3)*, on commence par voir que si f est continue, f et $-f$ sont s.c.i.
Pour la réciproque, on écrit $]a, b[=]a, +\infty] \cap [-\infty, b[$ et par suite
$f^{-1}(]a, b[) = f^{-1}(]a, +\infty]) \cap f^{-1}([-\infty, b[) = f^{-1}(]a, +\infty]) \cap (-f)^{-1}(]-b, +\infty])$,
ce qui prouve bien que si f et $-f$ sont s.c.i., f est continue. □

Proposition 2.3.3. *La fonction numérique $f : E \to \overline{\mathbb{R}}$ est s.c.i. si et seulement si son épigraphe est fermé.*

Démonstration : *i)* Supposons que f est s.c.i. Soit $(x, t) \in E \times \mathbb{R}$ un point n'appartenant pas à epi f. On a donc $t < f(x)$. Choisissons $\tau \in \mathbb{R}$ avec $t < \tau < f(x)$. L'ensemble $U_\tau = \{y \in E \ / \ f(y) > \tau\}$ est un ouvert contenant x. L'ensemble $U_\tau \times] - \infty, \tau[$ est un ouvert de $E \times \mathbb{R}$ contenant (x, t) et disjoint de epi f. Donc le complémentaire de epi f est ouvert et epi f est bien fermé.

ii) Réciproquement si epi f est fermé, alors pour tout $a \in \mathbb{R}$, l'ensemble $S_a := \text{epi } f \cap (E \times \{a\})$ est fermé comme intersection de deux fermés de $E \times \mathbb{R}$; cet ensemble s'écrit aussi $S_a = F_a \times \{a\}$, où l'on a posé $F_a = \{x \in E \ / \ f(x) \leq a\}$. Donc F_a est fermé dans E comme image réciproque de S_a par l'application continue $x \mapsto (x, a)$. Dans le cas spécial où $a = -\infty$, on note que $F_{-\infty} = \bigcap_{t \in \mathbb{R}} F_t$ est fermé comme intersection d'une famille de fermés. □

On a une autre caractérisation des fonctions s.c.i. en termes de limites inférieures :

Définition 2.3.4. *1) Si $f : E \to \overline{\mathbb{R}}$, $x_0 \in E$, et $Y \subset E$, on pose :*

$$\liminf_{\substack{x\to x_0\\ x\in Y}} f(x) = \sup_{r>0} \inf_{\substack{d(x,x_0)<r\\ x\in Y}} f(x)$$

Lorsque $Y = E$, on écrit simplement $\liminf\limits_{x\to x_0} f(x) = \sup\limits_{r>0} \inf\limits_{d(x,x_0)<r} f(x)$

2) Si $(x_n)_{n\in\mathbb{N}}$ est une suite de E, on pose :

$$\liminf_{n\to+\infty} x_n = \lim_{p\to+\infty} \inf\{x_p \ , \ p \geq n\}$$

Remarque 2.3.5. Si $x_0 \in Y$, on a $\liminf\limits_{\substack{x\to x_0\\ x\in Y}} f(x) \leq f(x_0)$.

Lemme 2.3.6. *Soit $Y \subset E$ et $x \in \overline{Y}$.*
1) $\liminf\limits_{\substack{y\to x\\ y\in Y}} f(y) = \lim\limits_{r\to 0} \inf\limits_{\substack{d(y,x)<r\\ y\in Y}} f(y)$.
2) Si $x_n \to x$ avec $x_n \in Y$ pour tout $n \in \mathbb{N}$, on a : $\liminf\limits_{n\to+\infty} f(x_n) \geq \liminf\limits_{\substack{y\to x\\ y\in Y}} f(y)$.
3) Il existe $(x_n)_{n\in\mathbb{N}} \subset Y$ telle que $f(x_n) \to \liminf\limits_{\substack{y\to x\\ y\in Y}} f(y)$.

Démonstration : Dans tout ce qui suit, on utilise le fait général suivant :

$$A \subset B \Rightarrow \inf A \geq \inf B$$

1) est vrai car la fonction $r \to \inf_{\substack{d(y,x)<r\\ y\in Y}} f(y)$ est croissante.

2) Pour tout $n_0 \in \mathbb{N}$, on a : $\inf_{n\geq n_0} f(x_n) \geq \inf\{f(y) \ , \ y \in Y \cap B(x, r_{n_0})\}$ où $r_{n_0} = \sup\{d(x, x_n) \ , \ n \geq n_0\}$, ce qui prouve le résultat.

3) On choisit $x_n \in B(x, \dfrac{1}{n}) \cap Y$ tel que $f(x_n) \leq \inf\{f(y) \ , \ y \in B(x, \dfrac{1}{n}) \cap Y\} + \dfrac{1}{n}$. □

Proposition 2.3.7. *La fonction numérique $f : E \to \overline{\mathbb{R}}$ est s.c.i. si et seulement si pour tout $x_0 \in X$ on a* $\liminf\limits_{x\to x_0} f(x) = f(x_0)$.

Démonstration : Si f est s.c.i., et $f(x_0) > -\infty$, alors pour tout $t < f(x_0)$ l'ensemble $\{x \in E \ / \ f(x) > t\}$ est un ouvert contenant x_0, donc contenant aussi une boule ouverte $B(x_0, r)$, avec $r > 0$. Donc $\liminf_{x\to x_0} f(x) \geq t$ et en faisant $t \to f(x_0)$ on obtient $\liminf_{x\to x_0} f(x) \geq f(x_0)$. Ceci reste évidemment vrai si

$f(x_0) = -\infty$. L'inégalité inverse est toujours vraie, donc $\liminf_{x\to x_0} f(x) = f(x_0)$.

Réciproquement, supposons cette égalité est vraie en tout point $x_0 \in E$. Soit $t \in [-\infty, +\infty[$ et supposons que $f(x_0) > t$. Il existe alors $r > 0$ tel que $\inf_{x\in B(x_0,r)} f(x) > t$. Le point x_0 est donc intérieur à l'ensemble $U_t = \{x \in X \ / \ f(x) > t\}$. Comme ceci est vrai pour tout $x_0 \in U_t$, cet ensemble est donc ouvert. Le cas $t = +\infty$ est trivial car $U_{+\infty} = \emptyset$. □

On traduit familièrement cette propriété en disant que les fonctions s.c.i. ne peuvent sauter que vers le bas :

Remarque 2.3.8. On dira que f est s.c.i. en $x_0 \in E$ si $\liminf\limits_{x\to x_0} f(x) = f(x_0)$.

Proposition 2.3.9. *1) Si $(f_i)_{i\in I}$ est une famille quelconque de fonctions s.c.i. de $E \to \overline{\mathbb{R}}$, l'enveloppe supérieure des $(f_i)_{i\in I}$, $f = \sup_i f_i$, est également s.c.i.*
2) Si f et g sont des fonctions s.c.i. de X dans $\mathbb{R} \cup \{+\infty\}$, et α, β des réels positifs, la fonction $\alpha f + \beta g$ est s.c.i.

DÉMONSTRATION : *1)* On a epi $f = \bigcap_i$ epi f_i. C'est donc un fermé, comme intersection de fermés.

2) Pour tout sous-ensemble A de E, on a :

$$\inf_{x\in A}(\alpha f(x) + \beta g(x)) \geq \alpha \inf_{x\in A} f(x) + \beta \inf_{x\in A} g(x)$$

Il en résulte que pour tout $x_0 \in X$, on a :

$$\liminf_{x\to x_0}(\alpha f(x) + \beta g(x)) \geq \alpha \liminf_{x\to x_0} f(x) + \beta \liminf_{x\to x_0} g(x)$$

Si f et g sont s.c.i., on en déduit $\liminf_{x\to x_0}(\alpha f(x)+\beta g(x)) \geq \alpha f(x_0)+\beta g(x_0)$. L'inégalité inverse est triviale. □

Remarque 2.3.10. Proposition 2.3.9 *1)* est fausse si on remplace partout "s.c.i." par "continue" (Exercice !)

Considérons maintenant un sous-ensemble Y de E et une fonction $f : Y \to \overline{\mathbb{R}}$ définie sur Y.

Définition 2.3.11 (Fonction s.c.i. définie sur une partie de E)**.** *On dit que la fonction numérique $f : Y \to \overline{\mathbb{R}}$ est semi-continue inférieurement (s.c.i.) si pour tout $a \in \overline{\mathbb{R}}$, l'ensemble $\{x \in Y \ / \ f(x) > a\}$ est un ouvert de Y, c'est à dire l'intersection d'un ouvert de E avec Y.*

Proposition 2.3.12 (Prolongement s.c.i.). *Si $f : Y \to \overline{\mathbb{R}}$ est une fonction s.c.i., la fonction $\tilde{f} : E \to \overline{\mathbb{R}}$ définie par :*

$$\tilde{f}(x) = \begin{cases} \liminf\limits_{y \to x\ ,\ y \in Y} f(x) \text{ si } x \in \overline{Y} \\ +\infty \qquad\qquad \text{sinon} \end{cases}$$

est un prolongement s.c.i. de f, et on a $\text{epi } \tilde{f} = \overline{(\text{epi } f)}$.

Démonstration : Le fait que $\tilde{f}(x) = f(x)$ pour $x \in Y$ résulte de la semi-continuité inférieure de f. Montrons que $\text{epi } \tilde{f} = \overline{(\text{epi } f)}$, il en résultera que $\tilde{f}$ est s.c.i.

a) Si $(x,t) \in \overline{\text{epi } f}$, il existe une suite $(x_n, t_n) \in \text{epi } f$ convergeant vers (x,t). En particulier $x_n \to x$ et donc $x \in \overline{Y}$. Pour tout $\varepsilon > 0$, on a $t_n \leq t + \varepsilon$, pour $n > n_0(\varepsilon)$, d'où $\tilde{f}(x) \leq \liminf f(x_n) \leq t + \varepsilon$; en faisant $\varepsilon \to 0$, on voit que $\tilde{f}(x) \leq t$, c'est à dire $(x,t) \in \text{epi}\tilde{f}$.

b) Réciproquement, si $(x,t) \in \text{epi } \tilde{f}$, on a $\tilde{f}(x) < +\infty$, d'où $x \in \overline{Y}$ et $\tilde{f}(x) = \liminf_{y \to x} f(y)$. Il existe alors $x_n \to x$, $x_n \in Y$ telle que $f(x_n) \to \tilde{f}(x)$. Soit $\varepsilon > 0$; comme $t + \varepsilon > \tilde{f}(x)$, on a $f(x_n) \leq t + \varepsilon$ pour $n \geq n_0(\varepsilon)$, c'est-à-dire $(x_n, t+\varepsilon) \in \text{epi } f$, et par conséquent $(x, t+\varepsilon) \in \overline{\text{epi } f}$. En faisant $\varepsilon \to 0$, on obtient $(x,t) \in \overline{\text{epi } f}$. □

En particulier si Y est un ensemble fermé, toute fonction $f : Y \to \mathbb{R}$ continue se prolonge en une fonction s.c.i. sur E en posant

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in Y \\ +\infty & \text{sinon} \end{cases}$$

2.4 Fonctions convexes s.c.i. sur un EVN

Soit $(E, \|\cdot\|)$ un EVN. En réunissant les résultats sur les fonctions convexes et les fonctions s.c.i. des deux sections précédentes, on obtient immédiatement les résultats suivants :

Proposition 2.4.1. *Une fonction $E \to \overline{\mathbb{R}}$ est convexe s.c.i. si et seulement si son épigraphe est convexe fermé.*

Exemple 2.4.2. *Si $A \subset E$ alors A est convexe fermé si et seulement si χ_A est convexe s.c.i.*

Proposition 2.4.3. *1) L'enveloppe supérieure d'une famille quelconque de fonctions convexes s.c.i. est convexe s.c.i.*
2) Si f et g sont des fonctions convexes s.c.i. de E dans $\mathbb{R} \cup \{+\infty\}$, et α, β des réels positifs, la fonction $\alpha f + \beta g$ est convexe s.c.i.

Soit C un convexe de E.

Proposition 2.4.4 (Prolongement convexe s.c.i.). *Si $f : C \to \overline{\mathbb{R}}$ est une fonction convexe s.c.i., la fonction $\tilde{f} : E \to \overline{\mathbb{R}}$ définie par :*

$$\tilde{f}(x) = \begin{cases} \liminf_{\substack{y \to x \\ y \in C}} f(x) & \textit{si } x \in \overline{C} \\ +\infty & \textit{sinon} \end{cases}$$

est un prolongement convexe s.c.i. de f à E, et on a epi $\tilde{f} = \overline{(\text{epi } f)}$.

Démonstration : En effet, d'après 2.3.12, $\tilde{f}$ est s.c.i., epi $\tilde{f} = \overline{(\text{epi } f)}$ et cet ensemble est un convexe fermé d'après 2.2.1. □
En particulier si C est un convexe fermé, toute fonction $f : C \to \mathbb{R}$ convexe continue se prolonge en une fonction s.c.i. en posant

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in C \\ +\infty & \text{sinon} \end{cases}$$

Remarque 2.4.5. Le prolongement convexe s.c.i. $\tilde{\chi_C}$ de la fonction caractéristique du convexe C est égal à $\chi_{\overline{C}}$.

Lemme 2.4.6. *Si la fonction convexe $f : E \to \overline{\mathbb{R}}$ prend la valeur $-\infty$ en un point x_0 et si $f(x_1) > -\infty$, alors $f(x) = +\infty$ pour tout x tel que $x_1 \in [x_0, x[$.*

Démonstration : Si $f(x) < +\infty$, l'inégalité de convexité sur le segment $[x_0, x[$ entraîne $f(x_1) = -\infty$, ce qui est une contradiction. □

Corollaire 2.4.7. *Si la fonction convexe $f : E \to \overline{\mathbb{R}}$ prend la valeur $-\infty$ en un point x_0, elle prend la valeur $-\infty$ sur tout l'intérieur de* dom f.

Démonstration : Soit $x_1 \in \overset{o}{\text{dom}} f$. Alors, il existe $z \in \overset{o}{\text{dom}} f$ tel que $x_1 \in [x_0, z]$. Si $f(x_1) > -\infty$, on obtient une contradiction avec le lemme. □

En particulier, en prenant la contraposée de cette proposition, on a :

Corollaire 2.4.8. *S'il existe $x_0 \in \overset{o}{\text{dom}} f$ tel que $f(x_0) > -\infty$, f ne prend nulle part la valeur $-\infty$.*

Corollaire 2.4.9. *Si* dom$f$ *est d'intérieur non vide et si f est finie et s.c.i. en un point $x_0 \in$ dom$f$, f ne prend nulle part la valeur $-\infty$.*

Démonstration : En effet, comme $f(x_0) > -\infty$ et f est s.c.i. en x_0, il existe une boule $B(x_0, r)$ telle que $f(x) > -\infty$ pour tout $x \in B(x_0, r)$. Si $f(x_1) = -\infty$, il existe $z \in B(x_0, r)$ tel que $z \in [x_1, x_0]$. D'après le Lemme 2.4.6, puisque $f(x_1) = -\infty$ et $f(z) > -\infty$, $f(x_0) = +\infty$, ce qui constitue une contradiction. $\square$

Chapitre 3
Structure des convexes de dimension finie

Dans ce chapitre, E est un espace vectoriel de dimension finie. On a vu au chapitre précédent que sur les espaces vectoriels de dimension finie, toutes les normes sont équivalentes. Sans perte de généralité, on peut donc munir E d'un produit scalaire $\langle \cdot, \cdot \rangle$: il suffit d'identifier E à $\mathbb{R}^p$ à l'aide d'une base et de prendre le produit scalaire naturel de $\mathbb{R}^p$, défini par $\langle x, y \rangle = \sum_{i=1}^p x_i y_i$ où $x = (x_1, \ldots, x_p)$ et $y = (y_1, \ldots, y_p)$. La norme euclidienne associée est alors $\|x\| = \langle x, x \rangle^{1/2} = \left(\sum_{i=1}^p x_i^2 \right)^{1/2}$.

On a également vu au chapitre 2, que dans ce cas, toutes les formes linéaires sont continues et que le dual de E s'identifie à E, c'est à dire que toute forme linéaire (continue) $f \in E'$ sur E est représentée par un élément $y \in E$, avec :

$$\forall x \in E \ , \ f(x) = \langle y, x \rangle \text{ et } \|f\| = \|y\| = \langle y, y \rangle^{1/2}$$

3.1 Intérieur d'un convexe en dimension finie

Proposition 3.1.1. *Soit C un convexe dans un espace vectoriel de dimension finie E. Les assertions suivantes sont équivalentes :*
i) $\overset{o}{C} \neq \emptyset$
ii) $\mathrm{Aff}(C) = E$
iii) C contient un repère affine.

Démonstration : *i) ⇒ iii)* : Toute boule $B(x_0, r)$ contient un repère affine : $(x_0, x_0 + \rho e_1, x_0 + \rho e_2, \ldots, x_0 + \rho e_p$, où $(e_1, \ldots, e_p)$ est une base de E et $\rho < \inf\{r/\, \|e_i\|\}$.

iii)⇒ i) : Si $(x_0, \ldots, x_p)$ est un repère affine de E, alors $\mathrm{conv}(x_0, \ldots, x_p)$ est inclus dans C et d'intérieur non vide : en effet soit $u_i = x_i - x_0$; $(u_1, \ldots, u_p)$ est une base de E. Soit $(u_1^*, \ldots, u_p^*)$ la base duale. Alors $x = \sum_{i=0}^p \lambda_i x_i$, équivaut lorsque $\sum_{i=0}^p \lambda_i = 1$ à $x - x_0 = \sum_{i=1}^p \lambda_i (x_i - x_0)$, et dans ce cas $\lambda_i = u_i^*(x - x_0)$

pour tout $i = 1, 2, \ldots, p$. Par conséquent :

$$\operatorname{conv}(x_0, \ldots, x_p) = \Big\{x \in E \;/\; u_i^*(x - x_0) \geq 0, i = 1, \ldots, p \text{ et } \sum_{i=1}^{p} u_i^*(x - x_0) \leq 1\Big\}$$

qui contient l'ouvert

$$\Big\{x \in E \;/\; u_i^*(x - x_0) > 0, i = 1, \ldots, p \text{ et } \sum_{i=1}^{p} u_i^*(x - x_0) < 1\Big\}.$$

Cet ouvert est non vide car il contient le point $x = \displaystyle\sum_{i=0}^{p} \frac{x_i}{p+1}$ pour lequel $u_i^*(x - x_0) = \dfrac{1}{p+1}$ pour tout $i = 1, 2, \ldots, p$.

iii) $\Rightarrow$ *ii)* : clair.

ii) $\Rightarrow$ *iii)* : Si *iii)* est faux, soit k le nombre maximum de points affinement indépendants de C, alors $k < n + 1$. Soit $(x_1, \ldots, x_k)$ k points affinement indépendants de C. Tout $x \in C$ peut s'écrire $x = \sum_{i=1}^{k} \lambda_i x_i$, avec $\sum_{i=1}^{k} \lambda_i = 1$ sinon $(x, x_1, \ldots, x_k)$ serait affinement libre et donc k ne serait pas maximal. Il en est de même de tout $x \in \operatorname{Aff}(C)$. On a donc $\operatorname{Aff}(C) = \operatorname{Aff}(x_1, \ldots, x_k)$. Or la direction de $\operatorname{Aff}(x_1, \ldots, x_k)$ est de dimension $k - 1 < n$, donc $\operatorname{Aff}(C) \neq E$, ce qui contredit *ii)*. □

Définition 3.1.2 (Dimension d'un convexe)**.** *On appelle dimension d'un convexe C la dimension de $\operatorname{Aff}(C)$. C'est aussi le nombre p maximal tel que C contienne $p + 1$ points affinement indépendants.*

Remarque 3.1.3. La définition ci-dessus est un cas particulier de la Définition 1.3.10.

Corollaire 3.1.4. *Si $\overset{o}{C} = \emptyset$, alors $\overset{o}{\overline{C}} = \emptyset$.*

Démonstration : En effet, par la Proposition 3.1.1 si $\overset{o}{C} = \emptyset$, $\operatorname{Aff}(C) \neq E$. Comme $\overline{C} \subset \overline{\operatorname{Aff}(C)} = \operatorname{Aff}(C)$, alors, $\overset{o}{\overline{C}} = \emptyset$, en utilisant à nouveau la Proposition 3.1.1. □

On va introduire une notion, le cône engendré par un convexe, qui permet de caractériser d'une façon géométrique l'intérieur d'un convexe en dimension finie :

Définition 3.1.5 (Cône tangent)**.** *Si C est un convexe d'un espace vectoriel E, et x un point de C, le cône tangent en x au convexe C est le cône de sommet x engendré par le convexe C ; on le notera $T_C(x)$. On a donc :*

$$T_C(x) = \{x + \lambda y \ / \ y \in C \ \textit{et}\ \lambda \geq 0\} = x + \mathrm{cone}(C - x).$$

La Proposition suivante suit presque immédiatement dès définitions.

Proposition 3.1.6. *Soit C est un convexe d'un espace vectoriel E, et x un point de C. Les assertions suivantes sont équivalentes :*
i) $T_C(x) = E$
ii) $\mathrm{cone}(C - x) = E$
iii) pour tout $u \in E$ il existe $\delta > 0$ tel que $x + \delta u \in C$
iv) Pour toute droite Δ contenant x l'intersection $\Delta \cap C$ contient un segment ouvert contenant x.

Démonstration : L'équivalence i) $\iff$ ii) suit de la définition du cône tangent. L'équivalence avec iii) suit du fait que, pour un convexe C, $\mathrm{cone}(C) = \bigcup_{\lambda>0} \lambda C$ (Proposition 1.2.7). Finalement, l'équivalence iii) $\iff$ iv) suit de la convexité de C : si $\Delta = \{x + tu \ / \ t \in \mathbb{R}\}$ est une droite passant par x, l'intersection $\Delta \cap C$ est soit $\{x\}$, soit un segment qui contient x dans son intérieur si et seulement si il contient $x + \delta u$ et $x + \delta'(-u)$ pour certains $\delta, \delta' > 0$.

Définition 3.1.7 (Coeur d'un convexe)**.** *Si les conditions de la Proposition 3.1.6 sont réalisées on dit que x appartient au coeur du convexe C.*

Proposition 3.1.8. *Soit C un convexe d'un EVN. Si l'intérieur de C n'est pas vide il coïncide avec le coeur de C.*

Démonstration : Si x est un point intérieur de C soit $B(x, r)$ une boule de centre x, de rayon $r > 0$ incluse dans C. Pour tout $u \in E$ le point $x + \dfrac{r}{2}\dfrac{u}{\|u\|}$ et dans C, donc x est dans le coeur de C.

Réciproquement, supposons que C admet un point intérieur x_0 et soit x est un point de C non intérieurà C ; alors x n'appartient à aucun segment ouvert $]x_0, y[$ avec $y \in C$ (d'après le Lemme 2.2.7). Donc x n'est pas dans le coeur de C. □

Proposition 3.1.9. *Si C est un convexe d'un espace vectoriel de dimension finie, son intérieur coïncide avec son coeur.*

DÉMONSTRATION : Si l'intérieur de C n'est pas vide il coïncide avec son coeur (Proposition 3.1.8). Si l'intérieur de C est vide alors $\mathrm{Aff}(C) \neq E$ (Proposition 3.1.1). Il existe donc un vecteur u n'appartenant pas à la direction du SEA $\mathrm{Aff}(C)$. Pour tout $x \in C$, la droite Δ passant par x de vecteur directeur u vérifie $\Delta \cap \mathrm{Aff}(C) = \{x\}$. A fortiori $\Delta \cap C = \{x\}$ donc x n'est pas dans le coeur de C : celui-ci est donc vide. □

Corollaire 3.1.10. *(Caractérisation des points intérieurs d'un convexe en dimension finie) Soit C un convexe d'un EVN de dimension finie E et $x \in C$. Les propriétés suivantes sont équivalentes :*
i) $x \in \overset{\circ}{C}$
ii) Pour toute droite affine Δ contenant x, l'intersection $\Delta \cap C$ contient un segment ouvert contenant x.
iii) $\forall u \in E$, $\exists \delta > 0$ *tel que* $x + \delta u$ *et* $x - \delta u \in C$.

Exercice 3.1. *Les résultats de cette section sont en général faux en dimension infinie. Considérons $E = c_{00}$ (voir Exemple 2.1.42) muni de la norme $\| \cdot \|_\infty$.*
1) Soit $C = \{(x_n) \in c_{00} \ / \ \forall n \in \mathbb{N} \ |x_n| \leq \frac{1}{n}\}$. Montrer que $\overset{\circ}{C} = \emptyset$ tandis que $\mathrm{Aff}(C) = \mathrm{cone}(C) = E$ *et alors le coeur de C contient* 0.
2) Soi H un hyperplan dense dans E (donné par Example 2.2.15). Montrer que l'affirmation du Corollaire 3.1.4 n'est pas vraie pour $C = H$.

3.2 Théorèmes de séparation en dimension finie

Ces théorèmes reposent sur le théorème fondamental des espaces euclidiens de dimension finie, qui sera généralisé en dimension quelconque dans le chapitre 6 :

Théorème 3.2.1 (Projection métrique dans les espaces euclidiens)**.** *Soit C un convexe fermé non vide d'un espace euclidien de dimension finie et $x_0 \notin C$. Il existe un point unique $y_0 \in C$, appelé projection (métrique) de x_0 sur C, qui réalise la distance de x_0 à C. La projection de x_0 sur C est caractérisée par*

$$\forall y \in C \ , \ \langle x_0 - y_0, y - y_0 \rangle \leq 0 \qquad (*)$$

DÉMONSTRATION : Soit $r = d(x_0, C) > 0$ et $B = C \cap \bar{B}(x_0, 2r)$. On écrit :

$$d(x_0, C) = \inf_{y \in C} \|y - x_0\| = \inf_{y \in B} \|y - x_0\| \, .$$

(La deuxième égalité suit de $d(x_0, C) = \min\{d(x_0, B), d(x_0, C \setminus B)\}$ combinée avec $d(x_0, C \setminus B) \geq 2r$.) Or B est un fermé borné dans un espace de dimension finie, il est donc compact. La fonction continue $y \to \|y - x_0\|$ atteint donc son minimum sur B. Un point y_0 où cette fonction atteint son minimum réalise la distance de x_0 à C.

Un tel point y_0 vérifie : $\forall y \in C\ ,\ \|x_0 - y\|^2 \geq \|x_0 - y_0\|^2$, d'où en écrivant $x_0 - y = (x_0 - y_0) + (y_0 - y)$ et en développant : $\langle x_0 - y_0, y - y_0\rangle \leq 1/2\,\|y - y_0\|^2$. Soit $0 < \theta < 1$. On applique cette inégalité à $y_\theta = \theta y + (1-\theta)y_0 \in C$ et on obtient :

$$\theta\langle x_0 - y_0, y - y_0\rangle \leq (1/2)\theta^2\,\|y - y_0\|^2$$

En divisant par θ et en faisant tendre θ vers 0, on obtient bien la condition $(*)$.

Réciproquement, si $y_0 \in C$ vérifie $(*)$, on a bien :

$$\begin{aligned} \|x_0 - y\|^2 &= \|x_0 - y_0\|^2 + 2\langle x_0 - y_0, y_0 - y\rangle + \|y_0 - y\|^2 \\ &\geq \|x_0 - y_0\|^2 + \|y_0 - y\|^2 \geq \|x_0 - y_0\|^2 \end{aligned}$$

et y_0 réalise la distance de x_0 à C. En plus, la dernière inégalité est stricte pour $y \neq y_0$, d'où l'unicité de la projection de x_0 sur C. □

Les deux théorèmes de séparation qui suivent, séparation d'un point et d'un convexe fermé et séparation d'un point et d'un convexe ouvert, sont vrais également en dimension quelconque mais la démonstration présentée ici est particulière à la dimension finie :

Théorème 3.2.2 (Séparation d'un point et d'un convexe fermé dans un espace de dimension finie)**.** *Si C est un convexe fermé non vide d'un espace E de dimension finie et x_0 un point de E n'appartenant pas à C il existe un hyperplan affine H séparant strictement x_0 de C ; c'est à dire que C et x_0 sont inclus chacun dans un demi-espace ouvert différent défini par H. En particulier C est inclus dans l'un des demi-espaces affines associés à H.*

Démonstration : Soient x_0 et C comme dans l'énoncé. La distance

$$d(x_0, C) = \inf\{\|y - x_0\|\ /\ y \in C\}$$

de x_0 à C est strictement positive, puisque $x_0 \notin C$ et C est fermé. Le théorème de projection sur un convexe fermé d'un espace euclidien, 3.2.1, dit qu'il existe un unique point y_0 de C réalisant cette distance, c'est à dire $\|y_0 - x_0\| = d(x_0, C)$, caractérisé par les inéquations :

$$\forall y \in C,\ \langle y - y_0, x_0 - y_0\rangle \leq 0. \qquad (\dagger)$$

Posons $f(y) = \langle y, x_0 - y_0\rangle$; la fonction f est une forme linéaire non nulle sur E. La condition $(\dagger)$ peut être reformulée :

$$\forall y \in C,\ f(y) \leq f(y_0).$$

Par ailleurs, on a $f(x_0) = f(y_0) + f(x_0 - y_0) = f(y_0) + \|x_0 - y_0)\|^2 > f(y_0)$
Soit H l'hyperplan défini par :

$$y \in H \iff f(y) = c,$$

où $c = \frac{f(x_0)+f(y_0)}{2} = f\left(\frac{x_0+y_0}{2}\right)$. (Notons que H est l'hyperplan médiateur du segment $[x_0, y_0]$.) On voit que

$$f(x_0) > c \quad \text{et} \quad \forall y \in C,\ f(y) < c.$$

Autrement dit, x_0 et C sont inclus dans les demi-espaces ouverts différents définis par H. □

Remarque 3.2.3. (i) Le même argument montre que il existe un hyperplan d'appui H de C tel que $C \subset D_+$ et $x_0 \in D'_-$. En effet, on peut prendre $H = f^{-1}(c')$ avec $c' = f(y_0)$, assurant que $y_0 \in H \cap C$ et alors $H \cap C$ est non-vide.
(ii) Sous les hypothèses du Théorème 3.2.2, il revient au même de dire que si x_0 n'appartient pas au convexe fermé C, il existe une forme linéaire f sur E, de la forme $x \mapsto \langle u, x\rangle$ avec $u \in E'$, telle que $\sup_{x\in C} f(x) < f(x_0)$. En effet, si l'hyperplan H, d'équation $f(x) = a$, sépare strictement x_0 et C, on a par exemple : $\forall x \in C\ ,\ f(x) < a$ et $f(x_0) > a$, d'où $f(x_0) > a \geq \sup_{x\in C} f(x)$. Inversement, s'il existe une forme linéaire f sur E, non nulle avec $f(x_0) > \sup_{x\in C} f(x)$, on pose $a = \dfrac{1}{2}\big[f(x_0) + \sup\limits_{x\in C} f(x)\big]$ et alors $f(x_0) > a$ et $\sup_{x\in C} f(x) < a$.

Théorème 3.2.4 (Séparation d'un point et d'un convexe ouvert en dimension finie)**.** *Soit E espace de dimension finie, C un convexe ouvert de E. Par tout point x_0 n'appartenant pas à C il passe un hyperplan disjoint de C.*

Démonstration : On peut supposer $x_0 = 0$. Soit $\Gamma = \bigcup\limits_{\lambda>0} \lambda C$, c'est un cône convexe ; il est ouvert comme réunion des convexes ouverts λC pour $\lambda > 0$ et ne contient pas 0 puisque $C \not\ni 0$; autrement dit $\Gamma = \text{cone}(C) \setminus \{0\}$ (voir Proposition 1.2.7). En particulier $\Gamma \neq E$, donc son adhérence $\overline{\Gamma}$ est aussi distincte de E, car $\overline{\Gamma}$ a même intérieur que Γ, c'est à dire Γ. Soit y_0 un point de E n'appartenant pas à $\overline{\Gamma}$, le théorème de séparation d'un point et d'un convexe fermé, Théorème 3.2.2 donne l'existence d'une forme linéaire f telle que $f(y_0) < \inf\limits_{z\in\overline{\Gamma}} f(z)$. Si $z \in \Gamma$, il en est de même de tz pour tout $t > 0$, donc $f(y_0) < f(tz) = tf(z)$ pour tous $t > 0, z \in \Gamma$. En divisant les deux membres par t et faisant $t \to +\infty$, on obtient $0 \leq f(z)$ pour tout $z \in \Gamma$. Donc Γ est inclus dans le demi-espace fermé $D = \{x\ /\ f(x) \geq 0\}$ défini par l'hyperplan $H = \operatorname{Ker} f$ qui passe bien par 0. Comme Γ est ouvert il est en fait inclus dans l'intérieur de D, c'est-à-dire le demi-espace ouvert $\{x\ /\ f(x) > 0\}$. Il en est a fortiori de même pour C. □

Corollaire 3.2.5. *Soit C un convexe d'intérieur non vide. Par tout point x_0 non intérieur à C il passe un hyperplan affine H disjoint de $\overset{o}{C}$. De plus l'un des demi-espaces fermés associés à H contient C.*

DÉMONSTRATION : On applique le théorème précédent à $\overset{o}{C}$. On trouve un hyperplan H contenant $\overset{o}{C}$, tel que l'un des demi-espaces ouverts associés à H contienne $\overset{o}{C}$. Le demi-espace fermé correspondant contient l'adhérence de $\overset{o}{C}$, qui coïncide avec celle de C par la Proposition 2.2.8, donc contient C. □

Remarque 3.2.6. Sous les hypothèses du Théorème 3.2.4, il revient au même de dire qu'il existe une forme linéaire non nulle f sur E, de la forme $x \mapsto \langle u, x\rangle$, avec $u \in E'$, telle que $\sup_{x\in C} f(x) \leq f(x_0)$.

De ces deux théorèmes de séparation, on déduit ce qu'on appelle habituellement les théorèmes de Hahn-Banach :

Corollaire 3.2.7 (Théorème de Hahn-Banach géométrique dans les espaces de dimension finie #1)**.** *Soient C_1 et C_2 deux convexes non vides et disjoints de l'espace vectoriel normé E. On suppose que C_1 est ouvert. Il existe alors un hyperplan affine H séparant strictement C_1 de C_2. Plus précisément il existe une forme linéaire f sur E et un réel a tels que $\forall x_1 \in C_1, \forall x_2 \in C_2$, $f(x_1) < a \leq f(x_2)$.*

DÉMONSTRATION : On applique le Théorème 3.2.4 au point 0 et au convexe non vide $C_1 - C_2 = \{y - z \ / \ y \in C_1, z \in C_2\}$. En effet $C_1 - C_2 = \bigcup_{z\in C_2}(C_1 - z)$ est ouvert comme réunion des ouverts $C_1 - z$, translatés de C_1. D'où il existe f forme linéaire telle que $\forall y \in C_1, z \in C_2, f(y - z) < f(0) = 0$; on en déduit, en notant $a = \inf\limits_{z\in C_2} f(z)$ que $\forall x_1 \in C_1, \forall x_2 \in C_2$, $f(x_1) \leq a \leq f(x_2)$. La première inéquation signifie que C_1 est inclus dans le demi-espace $f(x) \leq a$; comme C_1 est ouvert, il est en fait inclus dans l'intérieur du demi-espace, c'est à dire le demi-espace ouvert $f(x) < a$. □

Corollaire 3.2.8 (Théorème de Hahn-Banach géométrique en dimension finie #2)**.** *Soit C un convexe d'intérieur non vide de l'espace vectoriel normé E. Tout sous-espace affine M disjoint de C est inclus dans un hyperplan affine H disjoint de $\overset{o}{C}$ et tel que C soit inclus dans l'un des demi-espaces limités par H.*

DÉMONSTRATION : Le convexe $\overset{o}{C} - M = \{x - y \ / \ x \in C, y \in M\}$ est ouvert, non vide et ne contient pas le point 0. Il existe donc d'après le Théorème 3.2.4 une forme linéaire f telle que $f(x - y) > f(0) = 0$ pour tout $x \in \overset{o}{C}, y \in M$. Alors la forme linéaire f est constante sur M : en effet, supposons qu'elle prend des valeurs distinctes en deux points $y_1, y_2 \in M$ et soit $y(t) = y_1 + t(y_2 - y_1)$ une paramétrisation de cette droite. Alors $f\big(y(t)\big) = f(y_1)+t\big(f(y_2)-f(y_1)\big)$ et la restriction de f à la droite y_1y_2 prend toutes les valeurs réelles, donc ne peut être majorée par $f(x)$ pour un $x \in C$. La forme linéaire f est donc bien égale à une constante a sur M. On a donc $M \subset H = \{y \ / \ f(y) = a\}$ qui est l'hyperplan affine recherché : en effet, comme le demi espace fermé $\{y \ / \ f(y) \geq a\}$ contient $\overset{o}{C}$, il contient $\overline{\overset{o}{C}}$ qui coïncide avec $\overline{C}$ d'après la Proposition 2.2.8, donc ce demi espace contient aussi C. □

Corollaire 3.2.9. *Soit C un convexe fermé de l'espace vectoriel normé E et et $K \subset E$ un convexe compact vérifiant $K \cap C = \emptyset$. Alors, il existe un hyperplan H qui sépare strictement C et K.*

Démonstration : La fonction $x \to d(x, C)$ est continue et strictement positive sur K puisque C est fermé. Comme K est compact, elle atteint son minimum $m > 0$ en un point $x \in K$. L'ensemble $K' = \{x \in E / \ d(x, K) < m\}$ est un convexe ouvert contenant K et disjoint de C. On peut donc, par le Corollaire 3.2.7, séparer strictement K' et C, donc aussi K et C. □

Corollaire 3.2.10 (Théorème de Hahn-Banach dans les espaces de dimension finie #3)**.** *Soit E un EVN et $F \subset E$ un SEV. Soit $f \in F'$, le dual de F. Alors, il existe $\tilde{f} \in E'$ qui est une prolongement de f (c'est-à-dire, $\tilde{f}(x) = f(x)$ si $x \in F$) et telle que $\|\tilde{f}\|_{E'} = \|f\|_{F'}$.*

Démonstration : On utilise le lemme suivant (Exercice !).

Lemme 3.2.11. *Soit X un EVN, $B = B(0, 1)$ la boule unité ouverte de X, et soit $f \in X'$, $f \neq 0$. Alors $f(B) \ =] - c, c[$, où $c = \|f\|$.*

Pour le Corollaire, on peut supposer que $c = \|f\| > 0$. Posons $M = f^{-1}(c)$, alors M est un hyperplan (fermé) dans F. Par Lemme 3.2.11, $M \cap B_F(0, 1) = \emptyset$ et – étant donné que $M \subset F$, il suit que $M \cap B_E(0, 1) = \emptyset$. Par Corollaire 3.2.8, il existe un hyperplan H de E tel que $H \supset M$ et $H \cap B_E(0, 1) = \emptyset$. Comme dans la démonstration de la Proposition 1.3.14, il existe une forme lineaire $\tilde{f}$ telle que $H = \tilde{f}^{-1}(c)$. Cette forme vérifie $\tilde{f}(x) = f(x)$ si $x \in M$ et si $x = 0$, alors "egalement pour tout $x \in \mathrm{Aff}(M \cup \{0\}) = f$. Autrement dit, $\tilde{f}$ qui est une prolongement de f. Par ailleurs, $H \cap B_E(0, 1) = \emptyset$ montre que $\tilde{f}\big(B_E(0, 1)\big) \subset] - c, c[$ et alors il suit du Lemme 3.2.11 que $\|\tilde{f}\| \leq c = \|f\|$. L'inégalité inverse étant évidente, le Corollaire suit. □

3.3 Autres théorèmes de séparation en dimension finie

Dans cette partie, on va montrer des théorèmes de séparation qui n'ont pas d'analogue en dimension quelconque. On utilise des notions, le cône engendré par un convexe et l'intérieur relatif d'un convexe qui n'ont d'intérêt qu'en dimension finie et qui sont essentielles pour ces théorèmes de séparation.

Définition 3.3.1 (Intérieur relatif d'un convexe de dimension finie)**.** *Soit C un convexe non vide d'un espace vectoriel de dimension finie E. On appelle intérieur relatif de C, et on note* $\mathrm{ri}(C)$ *l'intérieur de C relatif au sous-espace affine* $\mathrm{Aff}(C)$.

Un point x de C est donc relativement intérieur à C s'il existe $r > 0$ tel que la boule $B(x, r)$ vérifie $\big[B(x, r) \cap \mathrm{Aff}(C)\big] \subset C$.

Remarque 3.3.2. Cette notion d'intérieur relatif ne dépend pas de la norme considérée sur l'espace de dimension finie E.

Grâce à cette notion d'intérieur relatif, on peut préciser la Proposition 2.2.8 dans les espaces de dimension finie :

Proposition 3.3.3. *Si C est un convexe non vide de dimension finie, son intérieur relatif* $\mathrm{ri}(C)$ *est aussi non vide. De plus,* $\overline{C} = \overline{\mathrm{ri}(C)}$ *et* $\mathrm{ri}(\overline{C}) = \mathrm{ri}(C)$.

Démonstration : Soit C un convexe non vide de E. Par translation on se ramène au cas où $0 \in \mathrm{Aff}(C)$. Alors on peut considérer $E_1 = \mathrm{Aff}(C)$ comme un espace vectoriel normé (par la restriction de la norme de E), et C comme un convexe de E_1 tel que $\mathrm{Aff}(C) = E_1$. L'intérieur de C relatif à E_1 est donc non vide (Proposition 3.1.1) et c'est exactement $\mathrm{ri}(C)$.

De plus, C et $\mathrm{ri}(C)$ ont même adhérence dans E_1, d'après la Proposition 2.2.8 ; mais l'adhérence d'une partie A de E_1 est la même dans E et E_1 car E_1 est fermé. De même, puisque E_1 est fermé, $\overline{C}$ est dans E_1 et l'intérieur de C et de $\overline{C}$ relativement à E_1 coïncident d'après la proposition 2.2.8, ce qui veut bien dire que ces deux convexes ont même intérieur relatif dans E. □

Corollaire 3.3.4. *(Caractérisation des points relativement intérieurs d'un convexe) Soit C un convexe de dimension finie et $x \in C$. Les propriétés suivantes sont équivalentes :*
i) $x \in \mathrm{ri}(C)$
ii) pour toute droite Δ incluse dans $\mathrm{Aff}(C)$ *et contenant x, l'intersection $\Delta \cap C$ contient un segment ouvert contenant x.*
iii) pour tout point y de D, avec $y \neq x$, il existe $z \in C$ tel que $x \in [y, z[$.
iv) Le cône tangent $T_C(x)$ est un sous-espace affine de E (on a alors $T_C(x) =$ $\mathrm{Aff}(C)$*).*

Démonstration : $i) \iff ii)$ résulte de la définition de l'intérieur relatif car (ii) signifie que x est dans le coeur du convexe C relatif à $\mathrm{Aff}(C)$.
$(ii) \Rightarrow (iii)$ est trivial.

$(iii) \Rightarrow (iv)$: on se ramène par translation au cas où $x = 0$. Il suffit de montrer que $T_C(0)$ est symétrique (un cône symétrique est un SEV). Si $u \in T_C(0)$, $u \neq 0$, il existe $y \in C$, $y \neq 0$ et $\lambda > 0$ tels que $u = \lambda y$. Soit $z \in C$ tel que $0 \in]y, z[$ (hypothèse (iii)). on a $z = -\rho y$ pour un certain $\rho > 0$. Alors $-u = \frac{\lambda}{\rho} z$, donc $-u \in T_C(0)$.
$(iv) \Rightarrow (ii)$: Si $T_C(x)$ est un SEA alors $T_C(x) \supset \mathrm{Aff}(C)$ puisque par définition $T_C(x)$ contient C et $\mathrm{Aff}(C)$ est le plus petit SEA contenant C.
Inversement $\mathrm{Aff}(C)$ est lui-même un cône de sommet x contenant C, donc contient $T_C(x)$. Donc $\mathrm{Aff}(C) = T_C(x)$. Toute demi-droite de $\mathrm{Aff}(C)$ contenant x est donc incluse dans $T_C(x)$: elle passe donc par un point $y \in C$, $y \neq x$. Donc toute droite de $\mathrm{Aff}(C)$ contenant x passe par deux points $y, z \in C$, distincts de x (un sur chacune des deux demi-droites correspondantes). Alors $x \in]y, z[$. □

Grâce à cette notion, on peut obtenir d'autres théorèmes de séparation, valables uniquement en dimension finie et qui sont susceptibles de repondre à des cas de figure particuliers non couverts par les théorèmes de séparation précédents.

Théorème 3.3.5. *Si C est un convexe non vide d'un espace E de dimension finie et x_0 un point de E n'appartenant pas à* $\mathrm{ri}(C)$*, il existe un hyperplan affine H de E, contenant x_0 et disjoint de* $\mathrm{ri}(C)$*. De plus C est inclus dans l'un des demi-espaces affines associés à H.*

DÉMONSTRATION : Si $x_0 \notin \mathrm{Aff}(C)$, La remarque 3.2.3(ii) appliqué à x_0 et au convexe fermé $\mathrm{Aff}(C)$ implique l'existence d'une forme linéaire f telle que $\sup_{x \in \mathrm{Aff}(C)} f(x) < f(x_0)$. Alors, $H = \{x \in E / \ f(x) = f(x_0)\}$ contient x_0 et est disjoint de $\mathrm{Aff}(C)$, a fortiori de $\mathrm{ri}(C)$.

Si $x_0 \in \mathrm{Aff}(C)$, on peut par translation supposer que $\mathrm{Aff}(C)$ est un sous-espace vectoriel fermé V de E. Appliquant le Théorème 3.2.4 ci-dessus au convexe C considéré dans l'espace vectoriel V, on voit qu'il existe un hyperplan M de l'espace V contenant x_0 et disjoint de $\mathrm{ri}(C)$. La preuve du Théorème 3.2.4 montre en fait l'existence d'un vecteur $u \in V$ tel que $\mathrm{ri}(C)$ soit inclus dans le demi-espace ouvert (relatif à V) $\{y \in V \ / \ \langle u, y\rangle < \langle u, x_0\rangle\}$; donc aussi dans le demi-espace ouvert (relatif à E) $\{y \in E \ / \ \langle u, y\rangle < \langle u, x_0\rangle\}$. L'hyperplan affine de E défini par $H = \{y \in E \ / \ \langle u, y\rangle = \langle u, x_0\rangle\}$ contient donc x_0 et est disjoint de $\mathrm{ri}(C)$. Comme $C \subset \overline{\mathrm{ri}(C)} = \overline{C}$ (Proposition 3.3.3), il est inclus dans l'adhérence du demi-espace ouvert, c'est-à-dire le demi-espace fermé $\{y \in E \ / \ \langle u, y\rangle \leq \langle u, x_0\rangle\}$. □

3.4 Application aux hyperplans d'appui

Corollaire 3.4.1. *Par tout point non intérieur d'un convexe de E il passe un hyperplan d'appui.*

DÉMONSTRATION : Si le convexe C est d'intérieur non vide c'est une conséquence directe du Théorème 3.2.4 puisque $\overset{o}{C} = \mathrm{ri}(C)$. Si C est d'intérieur vide, $\mathrm{Aff}(C)$ est un sous-espace affine de E, différent de E. Il est donc inclus dans un hyperplan affine H ; celui-ci contient C, donc est bien un hyperplan d'appui de C. □

Corollaire 3.4.2. *Tout convexe fermé C de dimension finie distinct de l'espace entier est intersection des demi-espaces fermés qui le contiennent et qui sont associés à des hyperplans d'appui.*

DÉMONSTRATION : Soit C_1 l'intersection des demi-espaces fermés contenant C associés à des hyperplans d'appui. Alors, $C \subset C_1$. Il suffit donc de montrer l'autre inclusion, c'est-à-dire que $x_0 \notin C \implies x_0 \notin C_1$.

Soit $x_0 \notin C$. Par Remarque 3.2.3(i), il existe un demi-espace $D_+ \supset C$ associé à un hyperplan d'appui de C tel que $x_0 \in D'_-$. Il suit que $C_1 \subset D_+$ et – étant donné que $D'_- \cap D_+ = \emptyset$ – que $x_0 \notin C_1$. □

Proposition 3.4.3. *Soit K est un convexe compact d'un espace vectoriel de dimension finie E. Pour tout hyperplan vectoriel V de E il existe un hyperplan d'appui de K parallèle à V .*

DÉMONSTRATION : Tout hyperplan vectoriel V a pour équation $f(x) = 0$, où f est une linéaire non nulle. Cette forme linéaire est évidemment continue, elle atteint donc son maximum M et son minimum m sur K. Les hyperplans d'équation $f(x) = m$ et $f(x) = M$ rencontrent donc K. Comme K est inclus dans les demi-espaces $f(x) \geq m$ et $f(x) \leq M$, ces hyperplans sont bien des sont des hyperplans d'appui de K, de direction V. □

Remarque 3.4.4. En fait on voit que soit K est inclus dans un hyperplan parallèle à V, soit il est inclus dans une bande définie par deux hyperplans d'appui parallèles à V.

Dans le cas où K est un convexe compact, la démonstration du fait que K est l'intersection des demi-espaces le contenant déterminés par ses hyperplans d'appui est plus simple que celle du cas général : on remarque que si $x_0 \in E$ est un point n'appartenant pas à K, il existe d'après la Proposition 3.2.2 une forme linéaire continue f telle que $M := \sup_{x \in K} f(x) < f(x_0)$. Le demi-espace $f(x) \leq M$ contient K mais non x_0. Par suite, l'hyperplan $\{x \in E/\ f(x) = M\}$ est un hyperplan d'appui pour K.

3.5 Structure des convexes compacts de dimension finie

Le théorème qui suit est la version finie-dimensionnelle d'un théorème de Krein et Milman.

Théorème 3.5.1 (Krein-Milman en dimension finie)**.** *Tout convexe compact K de dimension finie est l'enveloppe convexe de ses points extrémaux.*

Démonstration : L'inclusion conv(Ext$K$) $\subset K$ est vraie pour tout convexe (même non-compact).

Pour l'autre inclusion, on raisonne par récurrence sur la dimension de l'espace vectoriel. Si $\dim E = 0$, c'est trivial. Soit $n > 0$, et supposons le théorème démontré pour $\dim E < n$. Soit K un convexe compact de E de dimension n. On peut supposer $\mathrm{Aff}(K) = E$ sinon on remplace E par $\mathrm{Aff}(K)$. Alors $\overset{o}{K} \neq \emptyset$ (Proposition 3.1.1). Soit $x \in K$, il s'agit de montrer que $x \in \mathrm{conv}(\mathrm{Ext}\, K)$:

• Si $x \in K \setminus \overset{o}{K}$, il existe un hyperplan d'appui H contenant x (Théorème 3.2.4), et on a $H \cap K \neq \emptyset$. Mais alors $F = H \cap K$ est une face de K, de sorte que $\mathrm{Ext}\, F \subset \mathrm{Ext}\, K$. D'autre part $\mathrm{Aff}(F) \subset H$ est un SEA de dimension strictement inférieure à n, par conséquent l'hypothèse de récurrence s'applique à F. On a donc $x \in \mathrm{conv}(\mathrm{Ext}\, F) \subset \mathrm{conv}(\mathrm{Ext}\, K)$.

• Si au contraire $x \in \overset{o}{K}$, soit D une droite affine quelconque contenant x. Alors $D \cap K$ est un convexe compact de D, c'est à dire un segment $[x_1, x_2]$. Les points x_1, x_2 ne sont pas dans $\overset{o}{K}$, et donc d'après ce qui précède, ils sont dans $\mathrm{conv}(\mathrm{Ext}\, K)$. Alors $x \in \mathrm{conv}(x_1, x_2) \subset \mathrm{conv}(\mathrm{Ext}\, K)$. □

Corollaire 3.5.2. *Soit K un convexe compact d'un espace vectoriel de dimension finie, et $f : K \to \mathbb{R} \cup \{\infty\}$ une fonction convexe. On a*

$$\sup_{x \in K} f(x) = \sup_{x \in \mathrm{Ext}\, K} f(x)$$

En particulier si la fonction f est continue, elle atteint son maximum en un point de $\mathrm{Ext}\, K$*.*

Démonstration : Soit $M = \sup_{x \in \mathrm{Ext}\, K} f(x) \in \overline{\mathbb{R}}$. Si $x \in \mathrm{conv}(\mathrm{Ext}\, K)$, cet élément est combinaison convexe d'une famille finie de points $x_i \in \mathrm{Ext}\, K$, $i = 1, \ldots, n$: soit $x = \sum_{i=1}^n \lambda_i x_i$, avec $\lambda_i > 0$ et $\sum_{i=1}^n \lambda_i = 1$. D'où

$$f(x) \leq \sum_{i=1}^n \lambda_i f(x_i) \leq M \sum_{i=1}^n \lambda_i = M$$

L'ensemble $\{x \in K \;/\; f(x) \leq M\}$ contient donc $\mathrm{conv}(\mathrm{Ext}\, K) = K$. Donc M est aussi le sup de f sur K.

Par ailleurs, si f est continue, elle atteint son maximum $M \in \mathbb{R}$ sur le compact K en un point y. Etant donné que y est une combinaison convexe d'une famille finie de points $x_i \in \mathrm{Ext}\, K$, le même argument montre que $\max_i f(x_i) \geq M$ et alors $\sup_{x \in \mathrm{Ext}\, K} f(x) \geq M$. Il suit que $\sup_{x \in \mathrm{Ext}\, K} f(x) = M$ et le sup est atteint. □

Remarque 3.5.3. *1)* Ce résultat est souvent utilisé sous la forme suivante : pour majorer la fonction f sur K, il suffit de la majorer sur $\operatorname{Ext} K$ ou $\overline{\operatorname{Ext} K}$.
2) Même si K est un convexe compact, $\operatorname{Ext} K$ n'est pas forcément fermé.

Pour voir *2)*, on peut considérer dans $\mathbb{R}^3$ l'enveloppe convexe K des points $A = (0,0,1)$, $B = (0,0,-1)$ et du cercle C d'équations : $z = 0, x^2+y^2-2x = 0$. Tous les points du cercle C sont extémaux dans K sauf le point $(0,0,0)$. Plus précisément, on a $\operatorname{Ext} K = \{A,B\} \cup \big(C \setminus \{0\}\big)$ qui n'est pas fermé.

Pour préciser le théorème de structure précédent, on utilise le théorème de Carathéodory :

Proposition 3.5.4 (Théorème de Carathéodory)**.** *Soit E un espace vectoriel de dimension finie p, et A un sous-ensemble de E.*
i) Tout point du cône convexe engendré par A est combinaison linéaire à coefficients positifs d'au plus p éléments de A.
ii) Tout point de l'enveloppe convexe de A est combinaison convexe d'au plus $p+1$ éléments de A.

Démonstration : *i)* Soit $x = \sum_{i=1}^n \lambda_i x_i$, avec $x_i \in A, \lambda_i \geq 0$. On veut montrer, par récurrence sur n, qu'il existe $I \subset \{1,\ldots,n\}$, avec $\operatorname{card}(I) \leq p$ et des nombres $\mu_i \geq 0$, tels que $\sum\limits_{i\in I} \mu_i x_i = x$. C'est clair si $n \leq p$. Supposons $n \geq p+1$. En plus, on peut supposer que $\lambda_i > 0$ pour tout i (sinon, on peut écarter le terme correspondant et appliquer l'hypothèse de récurrence). La famille $(x_i)_{1\leq i\leq n}$ est nécessairement liée, il existe donc des réels $(\alpha_1,\ldots,\alpha_n)$ non tous nuls tels que $\sum_{i=1}^n \alpha_i x_i = 0$. Pour tout $\tau \in \mathbb{R}$, on a alors :

$$x = \sum_{i=1}^n (\lambda_i - \tau\alpha_i) x_i$$

On peut toujours supposer que $\{i \;/\; \alpha_i > 0\} \neq \emptyset$, quitte à changer le signe de tous les α_i. Soit $\tau = \min\{\frac{\lambda_i}{\alpha_i} \;/\; \alpha_i > 0\}$. On a $\lambda_i \geq \tau\alpha_i$ si $\alpha_i > 0$, par définition de τ, avec l'égalité pour $i = i_0$, où $\tau = \frac{\lambda_{i_0}}{\alpha_{i_0}}$. Par ailleurs, on a $\tau > 0$ et alors $\lambda_i > 0 \geq \tau\alpha_i$ si $\alpha_i \leq 0$. L'élément x est donc une combinaison linéaire des $n-1$ éléments x_i, $i \neq i_0$, à coefficients $\lambda_i - \tau\alpha_i$ (positifs ou nuls d'après ce qui précède), et on applique l'hypothèse de récurrence.

ii) Dans l'espace vectoriel augmenté $\hat{E} = E \oplus \mathbb{R}$, on considère l'ensemble $\hat{A} = A \times \{1\}$. Il est clair que $\operatorname{conv}(\hat{A}) = \operatorname{conv}(A) \times \{1\}$. Si $(x,1) \in \operatorname{conv}(\hat{A})$, on a fortiori $(x,1) \in \operatorname{cone}(\hat{A})$, par conséquent il existe, d'après le (i) ci-dessus, $p+1$ points $(a_i,1)$ dans $\hat{A}$ dont $(x,1)$ soit combinaison linéaire à coefficients positifs

(α_i). Par conséquent :

$$\begin{cases} x = \sum_{i=1}^{p+1} \alpha_i a_i \\ 1 = \sum_{i=1}^{p+1} \alpha_i \end{cases}$$

c'est à dire que x est en fait combinaison convexe des $(p+1)$ points a_i de A. □

Corollaire 3.5.5. *Dans un espace de dimension finie p, tout point d'un convexe compact est combinaison convexe d'au plus $p+1$ points extrémaux du convexe.*

Une autre conséquence du théorème de Carathéodory est relative à l'enveloppe convexe d'un ensemble compact, ce résultat a été annoncé au chapitre 1, Remarque 2.2.5 :

Proposition 3.5.6. *Dans un espace vectoriel de dimension finie, l'enveloppe convexe de tout compact est encore compacte.*

Démonstration : Soit p la dimension de E, K un compact de E ; posons $T_p = \{(\lambda_1, \ldots, \lambda_p) \in [0,1]^p \ / \ \sum_{i=1}^{p} \lambda_i = 1\}$: T_p est un compact (car fermé dans $[0,1]^p$). Soit h l'application continue

$$T_{p+1} \times K^{p+1} \to E, \ (\lambda_1, \ldots, \lambda_{p+1}, x_1, \ldots, x_{p+1}) \mapsto \sum_{i=1}^{p+1} \lambda_i x_i.$$

On a $\mathrm{conv}(K) = h(T_{p+1} \times K^{p+1})$ d'après le théorème de Carathéodory, donc $\mathrm{conv}(K)$ est compacte en tant qu'image d'un compact par une application continue. □

Rappelons que l'on a vu au chapitre 2 que ce résulat est faux en dimension quelconque.

3.6 Structure des cônes convexes fermés de dimension finie

Définition 3.6.1 (Cône convexe saillant)**.** *Un cône convexe Γ est dit saillant si $\Gamma \cap (-\Gamma) \subset \{0\}$.*

Remarque 3.6.2. (i) Il revient au même de dire qu'un cône convexe Γ est saillant ou que $\Gamma \setminus \{0\}$ est encore un cône convexe, ou encore que Γ ne contient aucune droite vectorielle.
(ii) Rappelons (voir 1.2.4) la définition $N(\Gamma) = \{0\} \cup (\Gamma \cap (-\Gamma))$. Alors Γ est saillant si et seulement si $N(\Gamma) = \{0\}$.

Proposition 3.6.3 (Décomposition des cônes convexes fermés de dimension finie)**.** *Tout cône convexe* Γ *contenant* 0 *d'un espace de dimension finie admet une décomposition de la forme :*

$$\Gamma = N(\Gamma) + \Gamma_0$$

où Γ_0 *est un cône convexe saillant inclus dans un supplémentaire de* $N(\Gamma)$*. De plus, si* Γ *est fermé,* Γ_0 *l'est aussi.*

DÉMONSTRATION : Etant donné que $0 \in \Gamma$, $N(\Gamma) = \Gamma \cap (-\Gamma)$ et $N(\Gamma)$ est le plus grand sous-espace vectoriel inclus dans Γ. Soit W un supplémentaire de $N(\Gamma)$. On pose $\Gamma_0 = \Gamma \cap W$. Il est clair que $N(\Gamma) + \Gamma_0 \subset \Gamma$.

Réciproquement tout élément x de Γ se décompose selon $x = v + y$, avec $v \in N(\Gamma)$ et $y \in W$; alors $y = x - v \in \Gamma$ car $-v \in \Gamma$, et donc $y \in \Gamma_0$. Donc $\Gamma = N(\Gamma) + \Gamma_0$. Le cône Γ_0 ne peut contenir de droite vectorielle sinon elle serait incluse dans $N(\Gamma)$; il est donc saillant.

Comme $\Gamma_0 = W \cap \Gamma$ et que W est fermé, il est clair que si Γ est fermé, Γ_0 l'est aussi. □

Remarque 3.6.4. *1)* Par construction, Γ_0 est la projection de Γ sur W, parallèlement à $N(\Gamma)$.
2) Réciproquement si $\Gamma = V + \Gamma_0$ où V est un sous-espace vectoriel de E, et Γ_0 un cône convexe saillant inclus dans un supplémentaire W de V, alors $V = N(\Gamma)$.

DÉMONSTRATION : *2)* En effet la projection de $N(\Gamma)$ sur W parallèlement à V est un sous-espace vectoriel inclus dans Γ_0, donc est réduit à $\{0\}$; d'où $N(\Gamma) \subset V$, et l'inclusion inverse est claire. □

Corollaire 3.6.5. *Le cône convexe* $\mathrm{cone}(u_1, \ldots, u_s)$ *engendré par un ensemble fini de points est fermé.*

DÉMONSTRATION : Soit $\Gamma = \mathrm{cone}(u_1, \ldots, u_s)$, et considérons une décomposition $\Gamma = N(\Gamma) + \Gamma_0$ donnée par la Proposition 3.6.3 Comme Γ_0 est la projection de Γ sur un supplémentaire W de $N(\Gamma)$, on a $\Gamma_0 = \mathrm{cone}(y_1, \ldots, y_s)$, où les y_i sont les projections des u_i.

Montrons que Γ_0 est fermé. On peut supposer les y_i non nuls pour $i \leq k$, et nuls pour $k < i \leq s$. Alors $\Gamma_0 = \mathrm{cone}(y_1, \ldots, y_k)$. Le convexe $K = \mathrm{conv}(y_1, \ldots, y_k)$ est compact, et ne contient pas 0 : en effet, $\{y_1, \ldots, y_k\} \subset \Gamma_0 \setminus \{0\}$, qui est encore convexe car Γ_0 est saillant. La distance $\delta = d(0, K)$ est donc strictement positive. Soit (z_n) une suite dans Γ_0 convergeant vers le point $z \in E$. On peut écrire $z_n = \lambda_n x_n$, avec $\lambda_n \geq 0$ et $x_n \in K$. Alors la suite (λ_n) est bornée car $\lambda_n = \dfrac{\|z_n\|}{\|x_n\|} \leq \dfrac{\sup_n \|z_n\|}{\delta}$ (notons que (z_n) est convergente et donc bornée).

Quitte à passer à une sous-suite, on peut suppososer $\lambda_n \to \lambda \in \mathbb{R}_+$ et $x_n \to x \in K$. Alors $z_n \to z \in \Gamma_0$, qui est donc bien fermé.
Comme $N(\Gamma)$ est fermé et Γ_0 est inclus dans le supplémentaire W de $N(\Gamma)$, la somme $\Gamma = N(\Gamma) + \Gamma_0$ est fermée : si une suite de Γ converge dans E, les suites projetées sur $N(\Gamma)$ et W convergent, et leurs limites sont nécessairement dans $N(\Gamma)$ et W respectivement. □

Remarque 3.6.6. En revanche, même en dimension finie, le fait que A soit compact n'entraîne pas que cone(A) soit fermé. Soit par exemple dans $E = \mathbb{R}^2$ un disque fermé D centré sur l'axe $0x$, et tangent à l'axe $0y$. Alors cone(D) est le demi-plan ouvert $\{(x, y) \ / \ x > 0, y \in \mathbb{R}\}$, auquel on rajoute le point 0.

Le seul point extrémal d'un cône convexe fermé saillant est son sommet. Un théorème de structure ne peut donc être obtenu à partir des points extrémaux, il faut faire intervenir des faces particulières, de dimension un, les génératrices extrémales.

Notation 3.6.7. *Si u est un vecteur non nul, notons γ_u la demi-droite $\{\lambda u \ / \ \lambda \in \mathbb{R}, \lambda \geq 0\}$ issue de 0, de vecteur directeur u.*

Définition 3.6.8 (Génératrices extrémales)**.** *Une direction du cône convexe pointé Γ est une demi-droite γ_u, avec $u \in \Gamma$, $u \neq 0$. Une direction γ du cône est dite extrémale si c'est une face de Γ. Une direction extrémale de Γ est aussi appelée génératrice extrémale de Γ. On notera* Gext(Γ) *l'ensemble des génératrices extrémales de Γ.*

On voit donc que $u \in \Gamma$, $u \neq 0$ définit une génératrice extrémale si et seulement si

$$x, y \in \Gamma, u = x + y \quad \Longrightarrow \quad \exists \lambda_1, \lambda_2 \geq 0 \text{ tels que } x = \lambda_1 u, y = \lambda_2 u.$$

Remarque 3.6.9. Posons pour tout $u \neq 0$, $\gamma_u^* = \gamma_u \setminus \{0\}$. La direction γ_u est une génératrice extrémale du cône convexe Γ si et seulement si $u \in \Gamma$, et pour tous $v, w \in \Gamma$ tels que $\gamma_u^* \subset \gamma_v^* + \gamma_w^*$ on a $\gamma_u = \gamma_v = \gamma_w$.

Proposition 3.6.10. *Un cône convexe non saillant n'admet aucune génératrice extrémale.*

Démonstration : En effet si Γ est non saillant, il existe $u \in \Gamma \cap (-\Gamma)$, $u \neq 0$, alors pour tout $x \in \Gamma$ et tout $\alpha > 0$, les vecteurs $x \pm \alpha u$ appartiennent à Γ. Si x définit une direction extrémale, ces vecteurs doivent appartenir à la demi-droite vectorielle $\mathbb{R}_+ x$, pour tout $\alpha > 0$. Il en résulte que $\pm u \in \mathbb{R}_+ x$, ce qui est impossible. □

On peut maintenant énoncer la version cône du théorème de Krein-Milman (en dimension finie) :

Théorème 3.6.11 (Krein-Milman pour les cônes en dimension finie)**.** *Dans un espace de dimension finie, tout cône convexe fermé saillant est l'enveloppe convexe de ses génératrices extrémales. Plus précisément tout point du cône est dans l'enveloppe convexe d'au plus* $p = \dim E$ *génératrices extrémales.*

Pour démontrer ce théorème, on a besoin de deux lemmes. Le premier lemme permet de ramener l'étude des cônes convexes fermés saillants à celle des convexes compacts :

Lemme 3.6.12. *Soit* Γ *un cône convexe fermé saillant d'un espace de dimension finie. Il existe un hyperplan affine* H*, ne contenant pas* 0*, dont l'intersection avec* Γ *soit un convexe compact* K*. De plus* Γ *est le cône convexe engendré par* K*.*

DÉMONSTRATION : Soit $S_E = \{x \in E/\ \|x\| = 1\}$ la sphère unité de E . Alors $A = S_E \cap \Gamma$ est un compact de E car c'est un fermé borné. Il résulte de la Proposition 3.5.6 que $L = \mathrm{conv}(A)$ est aussi compact. De plus on a $L \not\ni 0$: en effet A est inclus dans $\Gamma \setminus \{0\}$ qui est convexe puisque Γ est saillant, de sorte que l'on a aussi $L \subset (\Gamma \setminus \{0\})$.
On peut alors séparer strictement le point 0 du convexe fermé L, et par conséquent trouver une forme linéaire f sur E telle que :

$$\inf_{x \in L} f(x) > f(0) = 0$$

On peut supposer que $\inf_{x\in L} f(x) = 1$. (On posera $H = \{x \in E\ /\ f(x) = 1\}$). A fortiori $\inf_{x\in A} f(x) \geq 1$, ce qui entraîne par homogénéité que $\forall x \in \Gamma$, $\|x\| \leq f(x)$. Alors le fermé $K = \{x \in \Gamma/\ f(x) = 1\}$ est inclus dans la boule unité de E, donc est compact.
Si $x \in \Gamma$, $x \neq 0$ on a $f(x) \geq \|x\| > 0$, donc $y = \dfrac{x}{f(x)}$ est bien défini, et appartient clairement à $H \cap \Gamma = K$. Comme $x = f(x)y$ pour tout $x \in \Gamma$, $x \neq 0$, on a bien $\Gamma = \mathrm{cone}(K)$. □

Lemme 3.6.13. *Soit* C *un convexe d'un hyperplan ne contenant pas* 0*, et* $\Gamma = \mathrm{cone}(C)$*. La bijection naturelle* $\gamma \to \gamma \cap H$ *entre les directions de* Γ *et les points de* C *fait correspondre les génératrices extrémales de* Γ *avec les points extrémaux de* C*.*

DÉMONSTRATION : Soit $f(x) = 1$ une équation de H. Si γ_x est une génératrice extrémale, donc une face, de Γ, il est trivial que $x \in \mathrm{Ext}\, C$.

Réciproquement, si x est un point extrémal de C et $x = x_1 + x_2$, $x_1, x_2 \in \Gamma \setminus \{0\}$, on a $1 = f(x) = f(x_1) + f(x_2)$, de sorte que, en posant $y_1 = \dfrac{x_1}{f(x_1)}$ et $y_2 = \dfrac{x_2}{f(x_2)}$ on peut écrire $x = f(x_1)y_1 + \varphi(x_2)y_2$ et x est donc le barycentre à coefficients non nuls des points y_1 et y_2. Ceci entraîne $y_1 = x = y_2$, donc x_1, x_2 appartiennent à la direction γ_x, qui est donc extrémale. □

On peut passer à la démonstration du Théorème 3.6.11 :

Démonstration : (Théorème 3.6.11) Soit K comme dans le Lemme 3.6.12. On a $K = \operatorname{conv}(\operatorname{Ext} K)$ d'après le théorème de Krein-Milman (Théorème 3.5.1), d'où

$$\Gamma = \mathbb{R}_+ K = \operatorname{conv}\{\mathbb{R}_+ x /\ x \in \operatorname{Ext} K\} = \operatorname{conv}\Big(\bigcup_{\gamma \in \operatorname{Gext}(\Gamma)} \gamma\Big) \qquad \square$$

3.7 Convexes fermés de dimension finie : le cas général

Si C est un convexe de l'espace vectoriel (de dimension finie) E, on considère dans l'espace vectoriel augmenté $\widehat{E} = E \times \mathbb{R}$ le convexe $\widehat{C} = C \times \{1\}$ (obtenu par translation à partir du convexe $C = C \times \{0\}$), et le cône convexe $\Gamma(C) = \operatorname{cone}(\widehat{C}) = \bigcup_{\lambda \geq 0} \lambda \widehat{C}$. C'est évidemment un cône convexe saillant puisque $\Gamma(C) \setminus \{0\}$ est inclus dans le demi-espace ouvert $E \times]0, +\infty[$ et $\Gamma(C) \cap (E \times \{0\}) = 0$ de sorte que $\Gamma(C) \cap -\Gamma(C) = 0$. Toutefois $\Gamma(C)$ n'est pas nécessairement fermé, même si C l'est, comme le montre le lemme suivant :

Lemme 3.7.1. *Le cône $\Gamma(C)$ est fermé si et seulement si C est compact.*

Démonstration : a) Supposons que C est compact. Soit (u_n) une suite dans $\Gamma(C)$ qui converge vers un point u de $\widehat{E}$. On a $u_n = (\lambda_n x_n, \lambda_n)$, avec $\lambda_n \geq 0$ et $x_n \in C$. Alors (λ_n) converge vers un réel $\lambda \geq 0$ (la seconde composante de u) ; par compacité de C, on peut (quitte à extraire une sous-suite) supposer que (x_n) converge vers $x \in C$. Alors $u = (\lambda x, \lambda)$, donc appartient à $\Gamma(C)$, qui est donc fermé.

b) Réciproquement supposons $\Gamma(C)$ fermé. Alors $\widehat{C} = \Gamma(C) \cap (E \times \{1\})$ est fermé, et donc C également. Si C n'est pas compact, c'est qu'il n'est pas borné, c'est à dire qu'il existe une suite $(x_n) \subset C$, avec $\|x_n\| \to \infty$. La suite $(x_n / \|x_n\|, 1/\|x_n\|)$ est bornée et incluse dans $\Gamma(C)$. Quitte à extraire, on peut supposer qu'elle converge vers un point $(u, 0)$, appartenant lui aussi à $\Gamma(C)$ (puisque ce cône est fermé) et tel que $\|u\| = 1$. C'est impossible car $\Gamma(C) \cap (E \times \{0\}) = \{0\}$. □

Dans le cas où C n'est pas compact, on s'intéresse donc à l'adhérence $\overline{\Gamma(C)}$:

Définition 3.7.2 (Cône asymptotique)**.** *On appelle cône asymptotique du convexe C le cône convexe fermé C_{as} tel que $\overline{\Gamma(C)} \cap (E \times \{0\}) = C_{as} \times \{0\}$.*

Autrement dit, le plongement naturel de E dans $E\times\mathbb{R}$ défini par $x \to (x, 0)$ fait apparaître le cône asymptotique C_{as} comme la trace de $\overline{\Gamma(C)}$ sur l'hyperplan $E \times \{0\}$ identifié à E.

Proposition 3.7.3. *Si C est un cône convexe fermé, on a $\overline{\Gamma(C)} = \Gamma(C) \cup (C_{as} \times \{0\})$.*

Démonstration : Il s'agit de montrer que les points de $\overline{\Gamma(C)}$ qui ne sont pas dans $\Gamma(C)$ sont nécessairement dans $E \times \{0\}$ et par suite dans $C_{as} \times \{0\}$. Si donc (y, λ) est un point de $\overline{\Gamma(C)}$, avec $\lambda > 0$, soit $(\lambda_n x_n, \lambda_n)$ ($x_n \in C$, $\lambda_n \geq 0$) une suite de points de $\Gamma(C)$ convergeant vers (y, λ). Alors le point $x = \dfrac{y}{\lambda} = \lim\limits_n x_n$ appartient à C, d'où $(y, \lambda) = (\lambda x, \lambda)$ appartient à $\Gamma(C)$. □

Proposition 3.7.4. *Soit C un convexe fermé. Les conditions suivantes sont équivalentes :*
i) u appartient au cône asymptotique de C ;
ii) Il existe une suite (x_n) dans C et une suite (λ_n) de réels positifs avec $\lambda_n \to 0$ et $\lambda_n x_n \to u$;
iii) Pour tout point x de C et tout réel $\lambda > 0$, le point $x + \lambda u$ appartient à C.
iv) Il existe un point x_0 de C tel que pour tout réel $\lambda > 0$, le point $x_0 + \lambda u$ appartienne à C.

Démonstration : L'équivalence de *i)* et *ii)* est la définition même du cône asymptotique.
i) ⇒ iii) : Si $x \in C$ et $u \in C_{as}$, alors $(x, 1) \in \widehat{C}$ et $(\lambda u, 0) \in \overline{\Gamma(C)}$; donc $(x+\lambda u, 1)$ appartient aussi au cône $\overline{\Gamma(C)}$, et comme ce n'est pas un élément de $E\times\{0\}$, c'est d'après la Proposition 3.7.3 un élément de $\Gamma(C)$; donc $x+\lambda u \in C$.

iii) ⇒ iv) : trivial.

iv) ⇒ iii) : On a $(x_0 + \lambda u, 1) \in \widehat{C}$, d'où $(x_0/\lambda + u, 1/\lambda) \in \Gamma(C)$ pour tout $\lambda > 0$. En passant à la limite $\lambda \to +\infty$, on obtient $(u, 0) \in \overline{\Gamma(C)}$, c'est à dire $u \in C_{as}$. □

Cette proposition montre que le cône asymptotique d'un convexe est formé des directions des demi-droites affines incluses dans ce convexe.

Exemples 3.7.5. *1) Dans $\mathbb{R}^2$, le cône asymptotique du convexe fermé défini par une parabole est la demi-droite issue de l'origine, dirigée selon l'axe de la parabole.*
2) Dans $\mathbb{R}^2$, le cône asymptotique du convexe fermé défini par une branche d'hyperbole est le cône saillant issue de l'origine, limité par les directions des axes de l'hyperbole.

Corollaire 3.7.6. *Le cône $\overline{\Gamma(C)}$ est saillant si et seulement si C ne contient aucune droite affine.*

DÉMONSTRATION : Comme $\overline{\Gamma(C)}$ est inclus dans le demi-espace $E \times \mathbb{R}_+$, $N\big(\overline{\Gamma(C)}\big) = \overline{\Gamma(C)} \cap (-\overline{\Gamma(C)})$ est inclus dans l'hyperplan $E \times \{0\}$, et donc dans $C_{as} \times \{0\}$. Il en résulte que $N\big(\overline{\Gamma(C)}\big) = (C_{as} \cap (-C_{as})) \times \{0\}$; donc $\overline{\Gamma(C)}$ est non saillant si et seulement si il existe un vecteur u non nul tel que u et $-u$ soient dans le cône asymptotique. Etant donné $x_0 \in C$ ceci équivaut à dire que la droite passant par x_0 de vecteur directeur u est incluse dans C, d'après *iii)* et *iv)* de la Proposition 3.7.4. $\square$

Théorème 3.7.7 (Décomposition des convexes fermés ne contenant pas de droite en dimension finie)**.** *Tout convexe fermé de dimension finie C ne contenant aucune droite affine est somme de son cône asymptotique et de l'enveloppe convexe de ses points extrémaux :*

$$C = \operatorname{conv}(\operatorname{Ext} C) + C_{as}$$

Plus précisément tout point x de C est dans la somme de k génératrices extrémales de C_{as} et de l'enveloppe convexe de ℓ points extrémaux de C, avec $k + \ell \leq \dim E + 1$.

DÉMONSTRATION : Il est clair d'après la Proposition 3.7.4 *iii)* que $\operatorname{conv}(\operatorname{Ext}(C)) + C_{as} \subset C$. Pour montrer l'autre inclusion, considérons le cône $\overline{\Gamma(C)}$: d'après le corollaire 3.7.6, le cône $\overline{\Gamma(C)}$ est saillant. Et d'après le théorème de Krein-Milman pour les cônes convexes fermés saillants (Théorème 3.6.11) on a

$$\overline{\Gamma(C)} = \operatorname{conv}\big(\operatorname{Gext}\overline{\Gamma(C)}\big)$$

On note que $C_{as} \times \{0\}$ est une face de $\overline{\Gamma(C)}$ comme intersection de ce cône convexe avec un hyperplan d'appui et il en résulte $\operatorname{Gext}\big(\overline{\Gamma(C)}\big) \cap (C_{as} \times \{0\}) = \operatorname{Gext}(C_{as}) \times \{0\}$, Lemme 1.4.11.

Montrons maintenant que $\mathrm{Gext}\big(\overline{\Gamma(C)}\big) \cap \Gamma(C) = \mathrm{Gext}\big(\Gamma(C)\big)$. Il suffit de vérifier l'inclusion $\mathrm{Gext}\big(\Gamma(C)\big) \subset \mathrm{Gext}\big(\overline{\Gamma(C)}\big)$. Si $x \in C$, notons $\gamma_{\widehat{x}} = \mathbb{R}_+.(x,1)$ la direction de $\Gamma(C)$ associée. Supposons que $\gamma_{\widehat{x}} \in \mathrm{Gext}\big(\Gamma(C)\big)$. Si $(x,1) = (y,\lambda)+(z,\mu)$, avec $(y,\lambda),(z,\mu) \in \overline{\Gamma(C)}$, et si $\lambda,\mu > 0$, on a en fait $(y,\lambda),(z,\mu) \in \Gamma(C)$, et par conséquent (y,λ) et (z,μ) sont colinéaires à $(x,1)$ par extrémalité de $\gamma_{\widehat{x}}$. Si au contraire on a par exemple $\mu = 0$, alors $(x,1) = (y,\lambda) + (z,0)$ avec $z \in C_{as}$; alors nécessairement $\lambda = 1$, $y \in C$, et $x = y + z$. Le point $x+z$ est aussi dans C, puisque $z \in C_{as}$. Le point $(x,1)$ étant milieu des points $(x \pm z, 1) \in \Gamma(C)$, ceux-ci doivent appartenir à la direction $\mathbb{R}_+(x,1)$, ce qui n'est possible que si $x+z = x = x-z$ donc $z = 0$. Donc dans ce cas $(z,\mu) = (0,0)$ et $(y,\lambda) = (x,1)$.
On voit donc que, dans les deux cas, $\gamma_{\widehat{x}}$ est bien une génératrice extremale du cône fermé $\overline{\Gamma(C)}$.
Comme $\Gamma(C) = \mathrm{cone}(\widehat{C})$ et $\widehat{C}$ est l'intersection de $\Gamma(C)$ avec l'hyperplan $E \times \{1\}$ qui ne contient pas $(0,0)$ et le Lemme 3.6.13 montre que $\gamma_{\widehat{x}} \in \mathrm{Gext}(\Gamma(C))$ si et seulement si $(x,1) \in \mathrm{Ext}\,(\widehat{C})$, c'est à dire si et seulement si $x \in \mathrm{Ext}\, C$.

Finalement soit $x \in C$. Il existe $m \leq p+1$ génératrices extrémales γ_i de $\overline{\Gamma}(C)$ telles que $(x,1) \in \sum_{i=1}^m \gamma_i$. D'après ce qui précède, il existe $x_1,\ldots,x_k \in \mathrm{Ext}\, C$, $\alpha_1,\ldots,\alpha_k \in \mathbb{R}_+$ et $u_{k+1},\ldots,u_m \in C_{as}$, tels que :

$$(x,1) = \sum_{i=1}^{k} \alpha_i (x_i,1) + \sum_{i=k+1}^{m} (u_i,0)$$

d'où $\sum_{i=1}^k \alpha_i = 1$ et $x = \sum_{i=1}^k \alpha_i x_i + \sum_{i=k+1}^m u_i$. □

Remarque 3.7.8. Si le convexe fermé C contient une droite affine, il n'a aucun point extrémal et son cône asymptotique n'a aucune génératrice extrémale.

Démonstration : En effet si $u \in N(C_{as})$, $u \neq 0$, et $x \in C$, on a $x \pm u \in C$, donc x n'est pas extrémal. Comme $N(C_{as}) \neq \{0\}$, C_{as} n'a pas de génératrice extrémale d'après la Remarque 3.6.10. □

Remarque 3.7.9. Si C est un convexe fermé, l'ensemble $\mathrm{conv}(\mathrm{Ext}\, C)$ n'est pas forcément fermé ni borné.

En effet, On a déja vu en 3.5.3 que $\mathrm{conv}(\mathrm{Ext}\, C)$ n'est pas nécessairement fermé et si C est le convexe fermé limité par une hyperbole dans $\mathbb{R}^2$, ses point extémaux sont exactement les points de l'hyperbole, qui est non borné.

On a pour les convexes fermés un théorème de décomposition généralisant la Proposition 3.6.3 :

Proposition 3.7.10 (Décomposition des convexes fermés en dimension finie). *Tout convexe fermé C d'un espace de dimension finie admet une décomposition de la forme :*

$$C = V + C_0$$

où V est un sous-espace vectoriel de E, et C_0 un convexe fermé ne contenant aucune droite affine et inclus dans un translaté d'un supplémentaire de V.

DÉMONSTRATION : Par translation, on peut supposer que $0 \in C$. Soit $N(C) := N(C_{as})$ le plus grand sous-espace vectoriel du cône asymptotique de C, W un supplémentaire de $N(C_{as})$ dans E, et $C_0 = C \cap W$; on conclut comme dans la preuve de la Proposition 3.6.3. $\square$

Chapitre 4
Structure des convexes polyédraux en dimension finie

On désigne par E un espace vectoriel de dimension finie.

4.1 Polaire d'une partie

Définition 4.1.1 (Polaire d'une partie)**.** *Soit A une partie d'un EVN E. Le polaire de A est le convexe fermé du dual E' défini par :*

$$A^\circ = \left\{x^* \in E' / \ \langle x^*, x\rangle \le 1, \forall x \in A\right\}$$

De même, si B est une partie de E', le polaire de B est le convexe fermé de E défini par :

$$B^\circ = \left\{x \in E / \ \langle x^*, x\rangle \le 1, \forall x^* \in B\right\}$$

Enonçons maintenant le Théorème du bipolaire :

Théorème 4.1.2 (Théorème du bipolaire)**.** *Pour toute partie A de E, on a :*

$$A^{\circ\circ} = \overline{\mathrm{conv}}(A \cup \{0\})$$

Démonstration : a) Il est clair que $A \subset A^{\circ\circ}$ et $0 \in A^{\circ\circ}$. Comme $A^{\circ\circ}$ est convexe fermé, il en résulte $A^{\circ\circ} \supset \overline{\mathrm{conv}}(A \cup \{0\})$.
b) Inversement, soit $x_0 \notin \overline{\mathrm{conv}}(A\cup\{0\})$. On peut donc, par le Théorème 3.2.2, séparer strictement le point x_0 du convexe fermé non vide $\overline{\mathrm{conv}}(A \cup \{0\})$; il existe donc $x^* \in E'$ telle que :

$$\sup_{x\in\overline{\mathrm{conv}}(A\cup\{0\})} \langle x^*, x\rangle < \langle x^*, x_0\rangle$$

En particulier $a := \langle x^*, x_0\rangle > b := \max(0, \sup_{x\in A}\langle x^*, x\rangle) \ge 0$. Par homogénéité, on peut supposer que $a > 1 \ge b$ (si $b = 0$, on multiplie x^* par $(2/a)$; si $b > 0$, on multiplie x^* par $(1/b)$). L'inégalité $1 \ge b$ signifie que $x^* \in A^\circ$. L'inégalité $a > 1$ entraîne ensuite que $x_0 \notin A^{\circ\circ}$. $\square$

Corollaire 4.1.3. *Pour tout convexe fermé contenant 0 on a :*

$$C^{\circ\circ} = C$$

Ces résultats sont vrais aussi en dimension quelconque : on pourra vérifier que la démonstration du Théorème 4.1.2 s'applique sans changement dans ce cadre en utlisisant les théorèmes de séparation du chapitre 6.

4.2 Définition et première représentation des convexes polyédraux

Définition 4.2.1 (Convexes polyédraux)**.** *On appelle convexe polyédral C de E une intersection d'un nombre fini de demi-espaces affines fermés, c'est à dire*

$$C = \left\{x \in E/\ g_i(x) \geq b_i;\ i = 1, \dots, N\right\}$$

où $g_1, \dots, g_N$ sont des formes linéaires sur E, et $b_1, \dots, b_N$ des réels.

On remarquera qu'un tel convexe polyédral C est toujours fermé. En plus, C est un cône – un cône polyédral – si tous les b_i sont nuls.

On a besoin d'un résulat sur les formes linéaires :

Lemme 4.2.2. *Soient $g_1, \dots, g_N$ sont des formes linéaires sur E. Alors, $g_1, \dots, g_N$ engendrent linéairement le dual E' de E si et seulement si* $\bigcap_{i=1}^{N} \operatorname{Ker} g_i = \{0\}$.

Démonstration : Si $x \in \bigcap_{i=1}^{N} \operatorname{Ker} g_i$ avec $x \neq 0$, il suit que pour tout $g \in \operatorname{Vect}\{g_1, g_2, \dots, g_n\}$ on a $g(x) = 0$. Donc $\operatorname{Vect}\{g_1, g_2, \dots, g_n\} \neq E'$.

Réciproquement, si $\operatorname{Vect}\{g_1, g_2, \dots, g_n\} \neq E'$, il existe un $g \in E'$ tel que $g \notin \operatorname{Vect}\{g_1, g_2, \dots, g_n\}$. Par le théorème de séparation d'un point et d'un convexe fermé (Théorème 3.2.2), il existe $x \in E = E''$ qui sépare g et $\operatorname{Vect}\{g_1, g_2, \dots, g_n\}$. Comme $\operatorname{Vect}\{g_1, g_2, \dots, g_n\}$ est un SEV, nécessairement, x annule tout élément de cet ensemble et donc est dans l'intersection des noyaux des g_i , $i = 1, 2, \dots, N$. Par contre, $g(x) \neq 0$. On en déduit qu'il existe un vecteur non nul dans $\bigcap_{i=1}^{N} \operatorname{Ker} g_i$. □

On peut maintenant étudier la structure des convexes polyédraux :

Proposition 4.2.3. *Si* $C = \{x \in E \;/\; g_i(x) \geq b_i;\; i = 1, \ldots, N\} \neq \emptyset$*, on a :*

$$C_{as} = \{x \in E/\; g_i(x) \geq 0;\; i = 1, \ldots, N\}$$

$$N(C) = N(C_{as}) = \{x \in E/\; g_i(x) = 0;\; i = 1, \ldots, N\} = \bigcap_{i=1}^{N} \operatorname{Ker} g_i$$

En particulier $N(C) = 0$ *si et seulement si* $g_1, \ldots, g_N$ *engendrent linéairement le dual* E' *de* E*.*

Démonstration : Montrons l'assertion relative à C_{as} ; les autres s'en déduisent facilement. Fixons $x_0 \in C$. Alors

$$\begin{aligned} u \in C_{as} &\iff \forall\lambda > 0,\; x_0 + \lambda u \in C \\ &\iff \forall\lambda > 0, \forall i = 1, \ldots, N,\; g_i(x_0 + \lambda u) \geq b_i \\ &\iff \forall\lambda > 0, \forall i = 1, \ldots, N,\; g_i(u) \geq \frac{b_i - g(x_0)}{\lambda} \end{aligned}$$

Faisant $\lambda \to +\infty$, on voit que cette dernière condition implique $g_i(u) \geq 0$, pour tous $i = 1, \ldots, N$. Inversement si ces inégalités sont vérifiées, on a bien :

$$g_i(x_0 + \lambda u) = g_i(x_0) + \lambda g(u) \geq g_i(x_0) \geq b_i$$

□

On cherche à caractériser les points extrémaux et les génératrices extrémales de convexes polyédraux. D'après 3.7.8 et 4.2.3, on supposera toujours que $\operatorname{Vect}(g_1, g_2, \ldots, g_N) = E'$. Commençons par les points extrémaux :

Proposition 4.2.4. *Soit* C *le convexe polyédral défini par les inéquations*

$$g_i(\cdot) \geq b_i \;,\; i = 1, \ldots, N$$

Pour tout $x \in C$*, soit* $I_C(x)$ *l'ensemble des indices* $i \in \{1, \ldots, N\}$ *pour lesquels* $g_i(x) = b_i$*. Alors* x *est extrémal si et seulement si*

$$\operatorname{Vect}\{g_i \;/\; i \in I_C(x)\} = E'$$

Démonstration : a) Supposons $\operatorname{Vect}\{g_i \;/\; i \in I_C(x)\} = E'$, montrons que x est extrémal. En effet si $x = (y + z)/2$, avec $y, z \in C$, on a nécessairement $g_i(y) = g_i(z) = b_i$ pour tout $i \in I_C(x)$. Comme les g_i, $i \in I_C(x)$ engendrent E' par hypothèse, on a donc $g(y) = g(z)$ pour toute $g \in E'$. Ceci entraîne bien que $y = z = x$, donc x est extrémal.

b) Supposons maintenant qu'au contraire $\mathrm{Vect}\{g_i \ / \ i \in I_C(x)\} \neq E'$. Il existe alors $u \in \bigcap_{i \in I_C(x)} \operatorname{Ker} g_i$, $u \neq 0$. Posons :

$$\varepsilon = \inf \frac{g_i(x) - b_i}{|g_i(u)|} \ , \ i \in I_C(x) \text{ et } g_i(u) \neq 0$$

Alors $\varepsilon > 0$ et $\varepsilon\,|g_i(u)| \leq g_i(x) - b_i$, pour tout $i \notin I_C(x)$. On a donc $g_i(x \pm \varepsilon u) \geq b_i$ pour $i \notin I_C(x)$.
D'autre part $g_i(x \pm \varepsilon u) = g_i(x) = b_i$ pour $i \in I_C(x)$ et si $g_i(u) = 0$.
D'où $x \pm \varepsilon u \in C$ et comme $\varepsilon u \neq 0$, x n'est pas extrémal. □

Corollaire 4.2.5. *Tout convexe polyédral a un nombre fini de points extrémaux.*

DÉMONSTRATION : Supposons comme ci-dessus le convexe polyédral C défini par les inéquations $g_i(\cdot) \geq b_i$, $i = 1, \ldots, N$. D'après la proposition précédente, si $x \in C$ est extrémal, il existe $I \subset \{1, \ldots, N\}$ tel que $\mathrm{Vect}(g_i)_{i \in I} = E'$ et $g_i(x) = b_i$ pour tout $i \in I$. On peut choisir I de cardinal $p = \dim E$. Inversement pour tout $I \subset \{1, \ldots, N\}$ de cardinal p tel que $\mathrm{Vect}(g_i)_{i \in I} = E'$, il existe exactement un point $z_I \in E$ tel que $g_i(z_I) = b_i$, pour tout $i \in I$. L'ensemble $\mathfrak{I}$ des parties I de $\{1, \ldots, N\}$ à p éléments telles que $\mathrm{Vect}(g_i)_{i \in I} = E'$ est de cardinal majoré par C_N^p, il en est donc de même de $\operatorname{Ext} C$ (qui est inclus dans $\{z_I / \ I \in \mathfrak{I}\}$). □

Remarque 4.2.6 (Sommets). Les points extrémaux d'un convexe polyédral sont aussi appelés ses sommets.

Corollaire 4.2.7. *Tout cône convexe polyédral a un nombre fini de génératrices extrémales.*

DÉMONSTRATION : Si $N(\Gamma) \neq (0)$, Γ n'a aucune génératrice extrémale. Si $N(\Gamma) = (0)$, les génératrices extrémales de Γ sont en bijection avec les points extrémaux d'un convexe compact $K = \Gamma \cap H$, où H est un hyperplan bien choisi (Lemmes 3.6.12 et 3.6.13). Un hyperplan est trivialement polyédral, donc K est polyédral et par suite a un nombre fini de points extrémaux. □

On s'intéresse ensuite aux génératrices extrémales d'un cône convexe polyédral :

Proposition 4.2.8. *Soit* Γ *un cône polyédral défini par les inéquations*

$$g_i(\cdot) \geq 0 \ , \ i = 1, \ldots, N$$

On suppose que $\mathrm{Vect}(g_i)_{i=1}^N = E'$. *Alors* $\gamma = \mathbb{R}_+ u$ *est une génératrice extrémale de* Γ *si et seulement si* $\mathrm{Vect}(g_i)_{i \in I_\Gamma(u)}$ *est un hyperplan de* E'.

DÉMONSTRATION : Remarquons que la condition $\mathrm{Vect}(g_i)_{i=1}^N = E'$ équivaut à

$$\bigcap_i \operatorname{Ker} g_i = 0,$$

donc à $N(\Gamma) = 0$ (Lemme 4.2.3), et donc au fait que $\mathrm{Gext}(\Gamma) \neq \emptyset$ (Théorème 3.6.11 et Remarque 3.6.10).
Soit H un hyperplan affine tel que $K = H \cap \Gamma$ soit compact et $\Gamma = \mathrm{cone}(K)$ (Lemme 3.6.12). Alors K est un convexe polyédral, défini par les inéquations

$$g_i(\cdot) \geq 0 \ , \ i = 1, \ldots, N$$

et l'équation de H, soit

$$g_0(\cdot) = 1$$

Si $x \in K$, on a :

$$I_K(x) = \{i/ \ g_i(x) = 0\} \cup \{0\} = I_\Gamma(x) \cup \{0\}$$

On a donc :

$$\mathrm{Vect}(g_i)_{i \in I_K} = \mathrm{Vect}[(g_i)_{i \in I_\Gamma} \cup \{g_0\}]$$

Mais pour tout $x \in K$, on a $g_0 \notin \mathrm{Vect}(g_i)_{i \in I_\Gamma(x)}$, puisque $g_0(x) \neq 0$. On en déduit :

$$\begin{aligned} x \in \operatorname{Ext} K &\iff \operatorname{rang}(g_i)_{i \in I_K} = \dim E \\ &\iff \operatorname{rang}(g_i)_{i \in I_\Gamma} = \dim E - 1 \end{aligned}$$

□

On déduit de cette étude une représentation des convexes polyédraux :

Corollaire 4.2.9 (Décomposition des convexes polyédraux de dimension finie)**.** *Tout convexe polyédral peut s'écrire sous la forme :*

$$C = \mathrm{conv}(x_1, \ldots, x_q) + \mathrm{cone}(u_1, \ldots, u_s)$$

Si C ne contient pas de droite affine, on peut supposer que $x_1, \ldots, x_q$ sont des points extrémaux de C et $u_1, \ldots, u_s$ définissent des directions extrémales de C_{as}.

DÉMONSTRATION : Soit $C = N(C) + C_0$ une décomposition du convexe fermé C en la somme d'un sous-espace vectoriel et d'un convexe fermé ne contenant pas de droite (Proposition 3.7.10). Alors C_0 est encore un convexe polyédral (intersection d'un convexe polyédral avec un SEA). Soient $x_1, \ldots, x_q$ les points extrémaux de C_0 (en nombre fini, d'après Corollaire 4.2.5) et $\gamma_1, \ldots, \gamma_m$ les génératrices extrémales de son cône asymptotique (en nombre fini, d'après Corollaire 4.2.7) ; $u_1, \ldots, u_m$ des vecteurs directeurs de ces génératrices extrémales. Soit $(v_1, \ldots, v_\ell)$ une base de $N(C)$. On pose $(v_1, -v_1, \ldots, v_\ell, -v_\ell) = (u_{m+1}, \ldots, u_s)$. On a donc $\mathrm{cone}(u_{m+1}, \ldots, u_s) = N(C)$ et $\mathrm{conv}(x_1, \ldots, x_q) + \mathrm{cone}(u_1, \ldots, u_m) = C_0$. □

4.3 Convexes $*$-polyédraux et deuxième représentation des convexes polyédraux

On cherche à trouver une réciproque au Corollaire 4.2.9. Pour cela, on introduit les convexes $*$-polyédraux :

Définition 4.3.1 (Convexe $*$-polyédral)**.** *Appelons convexe $*$-polyédral un convexe de la forme*

$$C = \mathrm{conv}(x_1, \dots, x_q) + \mathrm{cone}(u_1, \dots, u_s)$$

Il résulte de la Proposition 4.2.9 que tout convexe polyédral (non vide) est $*$-polyédral.

Définition 4.3.2 (Polytope)**.** *Un convexe $*$-polyédral de la forme $C = \mathrm{conv}(x_1, \dots, x_q)$ est appelé un polytope.*

Proposition 4.3.3. *Tout convexe $*$-polyédral est fermé.*

DÉMONSTRATION : On a $C = K + \Gamma$, où $K = \mathrm{conv}(x_1, \dots, x_q)$ est compact (Proposition 2.2.4) et $\Gamma = \mathrm{cone}(u_1, \dots, u_s)$ est fermé (Corollaire 3.6.5). Il s'ensuit que C est fermé. En effet, si (x_n) est une suite de C convergeant vers $x \in E$, on a $x_n = y_n + z_n$, avec (y_n) dans K et (z_n) dans Γ. On peut supposer que (y_n) converge vers un point y de K. Alors (z_n) converge vers $z = x - y$, et $z \in \Gamma$ puisque Γ est fermé. D'où $x = y + z$ est dans $K + \Gamma = C$. $\square$

On va utiliser la notion de polaire introduite au chapitre 2, Définition 4.1.1 :

Proposition 4.3.4. *Le polaire d'un convexe $*$-polyédral est un convexe polyédral. Précisément, si*

$$C = \mathrm{conv}(x_1, \dots, x_q) + \mathrm{cone}(u_1, \dots, u_s)$$

alors

$$C^\circ = \{x^* \in E' / \ \langle x^*, u_j \rangle \leq 0 \ \ , \ j = 1, \dots, s \ ; \ \langle x^*, x_i \rangle \leq 1 \ \ , \ i = 1, \dots, q\}.$$

DÉMONSTRATION : Soit $C = \mathrm{conv}(x_1, \dots, x_q) + \mathrm{cone}(u_1, \dots, u_s)$ un convexe $*$-polyédral. Si $x^* \in C^\circ$, on a en particulier $\langle x^*, x_i + \lambda u_j \rangle \leq 1$ pour tous $1 \leq i \leq q$, $1 \leq j \leq s$ et $\lambda \geq 0$. En divisant par λ et faisant $\lambda \to +\infty$, on en déduit $\langle x^*, u_j \rangle \leq 0$ pour tout $1 \leq j \leq s$, tandis qu'en faisant $\lambda = 0$ il vient $\langle x^*, x_i \rangle \leq 1$ pour tout $1 \leq i \leq q$. Réciproquement si ces conditions sont vérifiées, on a $\langle x^*, x \rangle \leq 1$ pour tout $x = \sum_{i=1}^q \lambda_i x_i + \sum_{j=1}^s \rho_j u_j \in C$ ($\lambda_i \geq 0, i = 1, \dots, q$; $\sum_{i=1}^q \lambda_i = 1$; $\rho_j \geq 0, j = 1, \dots, s$). Donc :

$$C^\circ = \left\{x^* \in E' / \ \langle x^*, u_j \rangle \leq 0 \ j = 1, \dots, s \ ; \ \langle x^*, x_i \rangle \leq 1 \ i = 1, \dots, q\right\}$$

est bien un convexe polyédral de E'. $\square$

Théorème 4.3.5. *La classe des convexes $*$-polyédraux coïncide avec celle des convexes polyédraux.*

Démonstration : On sait déja que tout convexe polyédral est $*$-polyédral (Proposition 4.2.9).

Réciproquement soit C un convexe $*$-polyédral. Par translation, on peut supposer que $0 \in C$. D'après la Proposition 4.2.9, son polaire C° est polyédral, donc $*$-polyédral. Donc, à nouveau d'après la Proposition 4.2.9, son bipolaire $C^{\circ\circ}$ est encore polyédral. Comme C est fermé (Proposition 4.3.3) et contient 0, on a $C^{\circ\circ} = C$ par le théorème du bipolaire, 4.1.2 : donc C est polyédral. □

Corollaire 4.3.6. *L'image d'un convexe polyédral de E par une application linéaire $T : E \to F$ est un convexe polyédral de F.*

Démonstration : Si $$C = \mathrm{conv}(x_1, \ldots, x_q) + \mathrm{cone}(u_1, \ldots, u_s)$$
on a
$$T(C) = \mathrm{conv}(Tx_1, \ldots, Tx_q) + \mathrm{cone}(Tu_1, \ldots, Tu_s)$$
□

Corollaire 4.3.7. *La somme d'un nombre fini de convexes polyédraux est un convexe polyédral.*

Démonstration : Si $C_1, \ldots, C_n$ sont des convexes polyédraux de E, $C_1 \times C_2 \times \cdots \times C_n$ est un convexe polyédral de E^n. Soit $S :\ E^n \to E$ l'application linéaire somme : $(x_1, \ldots, x_n) \to \sum_{i=1}^n x_i$. Alors $S(C_1 \times C_2 \times \cdots \times C_n) = C_1 + C_2 + \ldots + C_n$ est polyédral. □

On a vu dans la preuve de la Proposition 4.3.4 que le polaire du convexe polyédral (ou, le même, $*$-polyédral) $C = \mathrm{conv}(x_1, \ldots, x_q) + \mathrm{cone}(u_1, \ldots, u_s)$ est le convexe polyédral d'inéquations $\langle x^*, x_i \rangle \leq 1$, $i = 1, \ldots, q$ et $\langle x^*, u_j \rangle \leq 0$, $j = 1, \ldots, s$. Réciproquement :

Remarque 4.3.8. Soient b_i, $i = 1, \ldots, q$ des réels strictement positifs. Le polaire du convexe polyédral C d'équations
$$C = \big\{x \in E/\ g_i(x) \leq b_i\ ,\ i = 1, \ldots, q\ ;\ h_j(x) \leq 0\ ,\ j = 1, \ldots, s\big\}$$
est le convexe
$$C^\circ = \mathrm{conv}(0, g_1/b_1, \ldots, g_q/b_q) + \mathrm{cone}(h_1, \ldots, h_s)$$

Démonstration : En remplaçant g_i par g_i/b_i, on peut supposer $b_i = 1$, $i = 1, \ldots, q$. Le polaire de $D = \mathrm{conv}(0, g_1, \ldots, g_q) + \mathrm{cone}(h_1, \ldots, h_s)$ est alors égal à C d'après ce qui précède. Alors $C^\circ = D^{\circ\circ} = D$ d'après le théorème du bipolaire, 4.1.2, (puisque D est fermé et contient 0). □

4.4 Application : le lemme de Farkas

Proposition 4.4.1 (Lemme de Farkas). *Soient $f_1, \dots, f_N$ et f des formes linéaires sur E. Une condition nécessaire et suffisante pour que le système d'inéquations*

$$f_j(x) \leq 0, \quad j = 1, \dots, N$$

entraîne l'inéquation $f(x) \leq 0$ est que f soit combinaison linéaire à coefficients positifs ou nuls des f_j, $j = 1, \dots, N$.

DÉMONSTRATION : Soit $\Gamma = \text{cone}(f_1, \dots, f_N)$. Son polaire n'est autre que le cône convexe polyédral de E défini par les équations $f_j(x) \leq 0$, $j = 1, \dots, N$. Dire que le système d'inéquations $f_j(x) \leq 0, \quad j = 1, \dots, N$ entraîne l'inéquation $f(x) \leq 0$ revient donc à dire que f appartient au bipolaire de Γ, qui n'est autre que Γ d'après le théorème du bipolaire 4.1.2 et le fait que Γ est fermé (Corollaire 3.6.5) et contient 0. $\square$

Chapitre 5
Application à la programmation linéaire

5.1 Programmation linéaire. Une méthode théorique de résolution

Dans tout ce chapitre, E est un EVN de dimension finie.

Définition 5.1.1 (Programmation linéaire)**.** *Un problème de programmation linéaire est un problème d'optimisation du type :*

$$(\mathcal{P})\begin{cases} \textit{déterminer} & \inf f(x) \\ \textit{sous les contraintes} & f_i(x) \leq b_i,\ i = 1, \ldots, m \end{cases}$$

où f et les f_i sont des formes linéaires sur E.

Il revient au même de dire que l'on cherche à minimiser f sur le convexe polyédral C déterminé par les inéquations $\{f_i(x) \leq b_i\ ,\ i = 1, \ldots, m\}$.

Définition 5.1.2 (Programme admissible)**.** *Le convexe C étant supposé non vide, ses éléments sont appelés les programmes admissibles de* $(\mathcal{P})$*.*

Soit $\operatorname{Ext} C$ l'ensemble des sommets (ou points extrémaux) de C, et $\operatorname{Gext}(C_{as})$ l'ensemble des génératrices extrémales du cône asymptotique C_{as}. On pose $\inf \mathcal{P} = \inf_{x \in C} f(x)$.

Définition 5.1.3 (Programme optimal)**.** *Un point $x_0 \in C$ tel que $f(x_0) = \inf \mathcal{P}$ (s'il en existe) est appelé programme optimal pour* $(\mathcal{P})$*.*

Proposition 5.1.4. *On suppose que C ne contient pas de droite. S'il existe une direction asymptotique $\gamma_u \in \operatorname{Gext}(C_{as})$ telle que $f(u) < 0$, alors $\inf \mathcal{P} = -\infty$. Dans le cas contraire $\inf \mathcal{P} > -\infty$, et est atteint en un point de $\operatorname{Ext} C$.*

Démonstration : Soit x_0 un point de C. Si $u \in C_{as}$ est tel que $f(u) < 0$, on a $f(x_0 + \lambda u) = f(x_0) + \lambda f(u) \to -\infty$ lorsque $\lambda \to +\infty$; comme $x_0 + \lambda u \in C$ pour tout $\lambda \geq 0$, on voit que $\inf \mathcal{P} = -\infty$.

Si au contraire $f(u) \geq 0$ pour tout u tel que $\gamma_u \in \mathrm{Gext}(C_{as})$, alors $f(u) \geq 0$ pour tout $u \in C_{as}$, puisque $C_{as} = \mathrm{cone}(\bigcup \mathrm{Gext}(C_{as}))$. On sait que $C = K + C_{as}$, avec $K = \mathrm{conv}(\mathrm{Ext}\, C)$. On a donc $\inf \mathcal{P} = \inf\{f(x) + f(u) \ / \ x \in K, u \in C_{as}\} = \inf\{f(x) \ / \ x \in K\}$. Soit $\mathrm{Ext}\, C = (s_1, \ldots, s_N)$ une énumération de l'ensemble (fini) des points extrémaux. Pour toute combinaison convexe $x = \sum_{i=1}^N \alpha_i s_i$, on a $f(x) = \sum_{i=1}^N \alpha_i f(s_i) \geq \min_{i=1,\ldots,N} f(s_i)$. Donc $\inf \mathcal{P} = \min_{i=1,\ldots,N} f(s_i) > -\infty$. $\square$

Remarque 5.1.5. (i) La première affirmation de la Proposition 5.1.4 est vraie pour un convexe quelconque. La deuxième affirmation n'est pas vraie en général, même pour des convexes fermés : Soit $E = \mathbb{R}^2$, $C = \{(x, y) \ / \ x > 0, xy \geq 1\}$, et $f(x, y) = y$.
(ii) Si C contient une droite, on a $C = N(C_{as}) + C_0$, où $C_0 = C \cap N(C_{as})^\perp$ est un convexe polyédral qui ne contient pas d'une droite. Si f n'annule pas tout element de $N(C_{as})$, il suit comme dans la démonstration de la Proposition 5.1.4 que $\inf \mathcal{P} = -\infty$. Dans le cas contraire, $\inf \mathcal{P}$ est équivalent au problème de minimisation de f sur C_0 et alors il suffit de solutionner le dernier.

En vue de la Proposition 5.1.4, une manière de résoudre le problème ($\mathcal{P}$) (sous l'hypothèse que C ne contient pas de droite) consiste donc à :

* Déterminer l'ensemble fini $\mathrm{Gext}(C_{as})$.
* Tester si il existe $\gamma_u \in \mathrm{Gext}(C_{as})$ tel que $f(u) < 0$: dans ce cas $\inf \mathcal{P} = -\infty$.
* Dans le cas contraire déterminer l'ensemble fini $\mathrm{Ext}\, C$.
* Calculer $f(s)$ pour tous les $s \in \mathrm{Ext}\, C$, et leur minimum.

Dans la pratique, ce schéma d'algorithme achoppe sur le nombre de points extrémaux ou de génératrices extrémales de C, qui peut être considérablement supérieur à la fois au nombre m de contraintes et à la dimension n de l'espace vectoriel des programmes E.

Par exemple si $E = \mathbb{R}^n$ et C est le cube de dimension n :

$$C = \left\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \ / \ 0 \leq x_i \leq 1, i = 1, \ldots, n\right\}$$

alors $m = 2n$. Les sommets de C sont les points $(x_1, \ldots, x_n)$ avec $x_i = 0$ ou 1, qui sont en nombre 2^n.

On est donc conduit à envisager un algorithme consistant à parcourir un chemin particulier dans l'ensemble $\mathrm{Ext}\, C$.

Définition 5.1.6 (Arêtes d'un convexe)**.** *Une arrête d'un convexe C est une face 1-dimensionnelle. On appelle $\mathcal{A}$ l'ensemble des arrêtes de C.*

Comme C est supposé ne pas contenir de droite, une arête est soit un segment $[s_1, s_2]$, et nécessairement dans ce cas $s_1, s_2 \in \operatorname{Ext} C$; soit une demi-droite $s + \gamma_u$, et alors nécessairement $s \in \operatorname{Ext} C$, et $\gamma_u \in \operatorname{Gext}(C_{as})$.

Lemme 5.1.7. *Un point non extrémal $x_0 \in C$ appartient à une arête de C si et seulement si, en posant $I_C(x_0) = \{i = 1, \ldots, m \ / \ f_i(x_0) = b_i\}$, la famille $(f_i)_{i \in I_C(x_0)}$ est de rang $(n-1)$ dans E'. Un vecteur non nul u est alors directeur pour cette arête si et seulement si $f_i(u) = 0$, $\forall i \in I_C(x_0)$. Pour tout x appartenant à la même arête, on a donc $I_C(x_0) \subset I_C(x)$ avec égalité quand x_0 n'est pas un sommet.*

Démonstration : i) Si $(f_i)_{i \in I_C(x_0)}$ est de rang $(n-1)$, l'intersection des $(n-1)$ hyperplans affines $H_i = \{x / \ f_i(x) = b_i\}$ est une droite affine Δ contenant x_0. Comme les hyperplans H_i sont d'appui pour C, chaque $H_i \cap C$ est une face de C, et il en est de même de $A = \Delta \cap C$. Comme x_0 n'est pas extrémal, la face A n'est pas réduite à $\{x_0\}$. C'est donc une arête de C, contenant x_0. Un vecteur directeur u de A vérifie évidemment $f_i(u) = 0$ pour tout $\forall i \in I_C(x_0)$; tout autre point x de a vérifie $f_i(x) = b_i$, pour tout $i \in I_C(x_0)$, c'est-à-dire $I_C(x_0) \subset I_C(x)$.

ii) Si $(f_i)_{i \in I_C(x_0)}$ est de rang n, x_0 est extrémal, il y a une contradiction.

iii) Si $(f_i)_{i \in I_C(x_0)}$ est de rang $r < (n-1)$, on peut trouver $n-r$ vecteurs indépendants u_ℓ, $\ell = 1, \ldots, n-r$, tels que $f_i(u_\ell) = 0$ pour tous $i \in I_C(x_0)$ et $\ell = 1, \ldots, n-r$.
Alors $x_0 \pm \delta u_\ell \in C$ pour tout $\ell = 1, \ldots, n-r$ et δ petit puisque $f_i(x_0 \pm \delta u_\ell) = b_i$ si $i \in I_C(x_0)$, tandis que $f_i(x_0 \pm \delta u_\ell) = f_i(x_0) \pm \delta f_i(u_\ell) < b_i$ pour $i \notin I_C(x_0)$ et δ suffisamment petit.
Il s'ensuit que toute face de C contenant x_0 contient aussi les $x_0 + \varepsilon u_\ell$, $\ell = 1, \ldots, n-r$, donc est de dimension $\geq n-r \geq 2$: donc x_0 n'appartient à aucune arête de C. □

Lemme 5.1.8. *Pour tout sommet s de C, le cône convexe de sommet s engendré par C coïncide avec le cône convexe de sommet s engendré par les arêtes issues de s.*

Démonstration : Par translation, on peut supposer que $s = 0$. Soit $\Gamma = \operatorname{cone}(C)$, le cône convexe engendré par C. Ce cône est encore polyédral, car si $C = \operatorname{conv}(0, s_2, \ldots, s_N) + \operatorname{cone}(u_1, \ldots, u_p)$, on a $\Gamma = \operatorname{cone}(s_2, \ldots, s_N, u_1, \ldots, u_p)$. Les génératrices extrémales de Γ sont des demi-droites engendrées par les arêtes de C issues de $s = 0$. En effet si u engendre une génératrice extrémale de Γ, soit $A = \gamma_u \cap C$; si $x \in A$, vérifie $x = \dfrac{v+w}{2}$, avec $v, w \in C$, on a a fortiori $v, w \in \Gamma$; si $x \neq 0$, on a $\gamma_x = \gamma_u$, d'où $v, w \in \gamma_x = \gamma_u$ par extrémalité de γ_u, donc $v, w \in A$; si $x = 0$, on a $v = w = 0$ par extrémalité de $s = 0$. Donc A est une arête de C.

Réciproquement, si A est une arête de C d'extrémité 0, soit $\gamma = \mathbb{R}_+ A$ la demi-droite issue de 0 engendrée par A. Si $u \in \gamma$ vérifie $u = \dfrac{v+w}{2}$ avec $v, w \in \Gamma$, soit $\lambda > 0$ tel que $\lambda u, \lambda v, \lambda w$ appartiennent à C. Alors $\lambda v, \lambda w$ appartiennent à A et donc v, w appartiennent à γ, qui est donc extrémale.
Le même raisonnement montre que 0 est un sommet de Γ, qui est donc saillant. Comme le cône polyédral saillant Γ est engendré par ses génératrices extrémales, il est donc inclus dans le cône convexe de sommet s engendré par les arêtes issues de s. L'inclusion inverse est triviale. □

Corollaire 5.1.9. *Si toute arête issue du sommet s de C est incluse dans le demi-espace $D = \{x \in E \ / \ f(x) \geq f(s)\}$, alors le convexe C tout entier est inclus dans ce demi-espace, c'est-à-dire que s est un programme optimal pour $(\mathcal{P})$.*

Définition 5.1.10 (Sommets voisins)**.** *On dira que deux sommets $s_1, s_2 \in \operatorname{Ext} C$ sont voisins si le segment $[s_1, s_2]$ est une arête de C.*

On peut alors modifier comme suit le schéma d'algorithme de recherche d'un programme optimal :

* initialisation : déterminer un sommet s_0 de C.
A la k-ième étape, on a déterminé un sommet s_k de C.
* s'il existe une arête infinie issue de s_k, dirigée par un vecteur u tel que $f(u) < 0$, alors $\inf \mathcal{P} = -\infty$, et l'algorithme s'arrête.
* sinon, deux cas se présentent :

** si pour tout sommet s voisin de s_k, on a $f(s) \geq f(s_k)$, alors s_k est un programme optimal (Corollaire 5.1.9) et l'algorithme s'arrête.

** sinon, il existe un sommet s voisin de s_k, tel que $f(s) < f(s_k)$; on pose $s_{k+1} = s$.

Puisque l'ensemble $\operatorname{Ext} C$ des sommets de C est fini, et que les sommets s_k ainsi construits sont tous distincts, l'algorithme s'arrête après un nombre fini d'étapes.

Cet algorithme consiste donc à joindre tout sommet initial de C à un sommet optimal s'il en existe par une chaîne de sommets voisins.

En fait il est possible de joindre ainsi deux sommets quelconques de C :

Remarque 5.1.11. Pour tous $s_0, s \in \operatorname{Ext} C$ il existe une chaîne $s_0, s_1, \ldots, s_N = s$ de sommets de C, telle que pour tout $i < N$, s_{i+1} soit voisin de s_i.

DÉMONSTRATION : Soit $\mathcal{V}(s_0)$ l'ensemble des sommets s de C qu'on peut joindre à s_0 par une telle chaîne. Soit $K_0 = \mathrm{conv}(\mathcal{V}(s_0))$. Alors, K_0 est inclus dans $\mathrm{convExt}(C)$.

Si $\mathrm{Ext}\, C$ n'est pas inclus dans le convexe fermé K_0, on peut trouver $s \in \mathrm{Ext}\, C$ et une forme linéaire f séparant strictement s de K_0, c'est-à-dire :

$$\inf_{x \in K_0} f(x) > f(s)$$

D'après la Proposition 5.1.4, on peut trouver $s_1 \in \mathcal{V}(s_0)$ tel que $f(s_1) = \inf_{x \in K_0} f(x)$.
Le Corollaire 5.1.9 appliqué à s_1 et $K = \mathrm{conv}(\mathrm{Ext}\, C)$ montre qu'il existe une arête $[s_1, s_2]$ de K telle que $f(s_2) < f(s_1)$. Alors $s_2 \in \mathrm{Ext}\, C$ est aussi dans $\mathcal{V}(s_0)$: en effet, une chaîne $t_0 = s_0, \ldots, t_N = s_1$ joignant s_0 à s_1 se prolonge en la chaîne $(t_0, \ldots, t_N, s_2)$ joignant s_0 à s_2 ; ce qui contredit $f(s_2) < \inf_{x \in K_0} f(x)$. □

5.2 Forme canonique et forme standard d'un problème linéaire

Dans la formulation canonique, on suppose que $E = \mathbb{R}^n$, et on désigne par $(e_j)_{j=1,\ldots,n}$ la base naturelle de $\mathbb{R}^n$.
La forme linéaire à optimiser est donnée par $f(x) = c.x$, où $c \in \mathbb{R}^n$ est identifiée à une matrice-ligne tandis que x est identifiée à une matrice colonne.

Définition 5.2.1 (Fonction objectif)**.** *f est appelée la fonction objectif.*

Les contraintes $f_i(x) \leq b_i$, $1 \leq i \leq m$ s'écrivent $A_i.x \leq b_i$, où les $A_i \in \mathbb{R}^n$ sont des matrices ligne. En notant $b \in \mathbb{R}^m$ la matrice colonne définie par les b_i, et A la matrice à m lignes et n colonnes dont les lignes sont les A_i, ces contraintes s'écrivent $Ax \leq b$, où $\leq$ désigne l'ordre partiel naturel sur $\mathbb{R}^m$, c'est à dire que l'inéquation $a \leq b$ signifie que leurs coordonnées a_i, b_i vérifient les inéquations $a_i \leq b_i$, $i = 1, \ldots, m$. On convient aussi de n'examiner que les programmes à coordonnées positives ou nulle.

Un problème linéaire sous forme canonique a donc la forme

$$(\mathcal{P}_c) \begin{cases} \text{minimiser } f(x) = c.x \\ \text{sous les contraintes } Ax \leq b \text{ et } x \geq 0 \end{cases}$$

Il est clair que le convexe polyédral défini par les contraintes ne contient aucune droite, puiqu'il est inclus dans le cône saillant $\mathbb{R}^n_+$.

Tout problème linéaire du type considéré au paragraphe 4.1 :

$$(\mathcal{P}) \begin{cases} \text{minimiser } f(x) = c.x, x \in \mathbb{R}^n \\ \text{sous les contraintes } \; Ax \leq b \end{cases}$$

se ramène à la forme canonique en dédoublant les variables :

$$(\mathcal{P}) \iff \begin{cases} \text{minimiser } g(x', x'') = c.(x' - x''), \; (x', x'') \in \mathbb{R}^{2n} \\ \text{sous les contraintes } A(x' - x'') \leq b \text{ et } x' \geq 0, x'' \geq 0 \end{cases}$$

Dans la formulation standard, les contraintes inégalités associées à la matrice A sont remplacées par des contraintes égalité :

$$(\mathcal{P}_s) \begin{cases} \text{minimiser } f(x) = c.x \\ \text{sous les contraintes } Ax = b \text{ et } x \geq 0 \end{cases}$$

On peut toujours passer de la formulation canonique à la formulation standard en introduisant une variable d'écart, "slack variable" en Anglais, $x' \in \mathbb{R}^m$:

$$(\mathcal{P}_c) \iff \begin{cases} \text{minimiser } g(x, x') = c.x + 0.x', \; (x, x') \in \mathbb{R}^{n+m} \\ \text{sous les contraintes } Ax + x' = b \text{ et } x \geq 0, x' \geq 0 \end{cases}$$

ce qui correspond à équation matricielle $\tilde{A} \begin{pmatrix} x \\ x' \end{pmatrix} = b$, avec la nouvelle matrice $\tilde{A} = (A \; / \; I_m)$.

5.3 Paramétrisation des programmes de base pour un problème standard

Dans toute la suite, le problème d'optimisation considéré est donné sous forme standard :

$$(\mathcal{P}_s) \begin{cases} \text{minimiser } f(x) = c.x, \; x \in \mathbb{R}^n \\ \text{sous les contraintes } Ax = b \text{ et } x \geq 0 \end{cases}$$

où A est une matrice à n colonnes, m lignes que l'on suppose de rang m.

Pour tout programme admissible x, notons $N(x)$ l'ensemble des indices $j = 1, \ldots, n$ pour lesquels les contraintes $x_j \geq 0$ sont actives, c'est à dire qu'on a $x_j = 0$:

$$N(x) = \{j \in \{1, \ldots, n\} \; / \; x_j = 0\}$$

Définition 5.3.1 (Programmes de base). *Les points extrémaux du convexe polyédral C défini par les contraintes $Ax = b$ et $x \geq 0$ sont appelés programmes de base.*

Lemme 5.3.2. *Un programme admissible x est un programme de base si et seulement si les vecteurs colonnes A^j de la matrice A correspondant aux indices $j \notin N(x)$ sont indépendants.*

Démonstration : D'après la Proposition 4.2.4, dire que x est un sommet de C équivaut à dire que les formes linéaires coordonnées $e_j^* : y \mapsto y_j$, pour $j \in N(x)$, et les formes linéaires associées aux contraintes égalité $f_i : y \mapsto A_i.y$, $i = 1, \dots, m$ engendrent E' ; il revient au même de dire que le seul vecteur de E annulant toutes ces formes linéaires est le vecteur nul :

$$\left.\begin{aligned} u_j = 0,\ \forall j \in N(x) \\ A_i.u = 0,\ i = 1 \dots m \end{aligned}\right\} \quad \Longrightarrow \quad u = 0$$

En notant A^j les vecteurs colonne de la matrice A, il revient encore au même de dire que :

$$\sum_{j \notin N(x)} u_j A_i^j = 0,\ i = 1, \dots, m \quad \Longrightarrow \quad u_j = 0, \forall j \notin N(x)$$

ou encore que les vecteurs colonne $(A^j)_{j \notin N(x)}$ sont indépendants. □

Remarque 5.3.3. Il résulte immédiatement de ce lemme que pour tout programme de base x, l'ensemble $N(x)$ a au moins $n-m$ éléments. Le programme de base est dit non dégénéré si $N(x)$ a exactement $n-m$ éléments, dégénéré s'il en a plus.

Si S est une partie non vide de l'ensemble $\{1, \dots, n\}$, soit $\mathbb{R}^S$ l'espace des suites de réels $(x_j)_{j \in S}$ indexées par S ; on a une application naturelle de restriction $r_S : \mathbb{R}^n \to \mathbb{R}^S$, $x = (x_j)_{j=1,\dots,n} \mapsto x_S = (x_j)_{j \in S}$.

Si (B, F) est une partition de $\{1, \dots, n\}$ en deux ensembles non vides, on écrira $x = \begin{pmatrix} x_B \\ x_F \end{pmatrix}$ pour signifier que x est obtenue par recollement des suites $(x_j)_{j \in B}$ et $(x_j)_{j \in F}$.

Notons que cette écriture est impropre et ambigüe lorsque les ensembles B et F ne sont pas consécutifs.

De même, la matrice A pourra s'écrire (improprement) $A = (A_B \ / \ A_F)$, où A_B est le bloc défini par les colonnes A^j, $j \in B$, et A_F celui défini par les colonnes A^j, $j \in F$. On a alors $Ax = A_B x_B + A_F x_F$.

Corollaire 5.3.4. *Si x est un programme de base, il existe une partition (B, F) de $\{1, \dots, n\}$ telle que* $\operatorname{card} B = m$*, que A_B soit une matrice régulière et que $x_B = A_B^{-1} b$, $x_F = 0$. Réciproquement pour toute sous-matrice A_B de A, carrée d'ordre m, inversible et telle que $A_B^{-1} b \geq 0$, le programme $s(B) := \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix}$ est un programme de base.*

DÉMONSTRATION : Si x est un programme de base, le système $(A^j)_{j\notin N(x)}$ est libre. Comme l'ensemble des vecteurs colonnes est de rang m par hypothèse, par le théorème de la base incomplète, on peut trouver une partie B à m éléments de $\{1,\ldots,n\}$, contenant le comlémentaire de $N(x)$, et tel que les colonnes A^j, $j\in B$, forment une base de $\mathbb{R}^m$. Alors la matrice A_B est inversible. Soit F le complémentaire de B, on a $F\subset N(x)$, donc $x_F=0$; alors $b=Ax=A_Bx_B$, d'où $x_B=A_B^{-1}b$.
Réciproquement si B est une partie à m éléments de $\{1,\ldots,n\}$, telle que A_B soit inversible et $A_B^{-1}b\geq 0$, alors $x=\begin{pmatrix}A_B^{-1}b\\0\end{pmatrix}$ est un programme admissible, et $N(x)$ contient le complémentaire F de B ; donc les colonnes A^j, $j\notin N(x)$ sont des colonnes de A_B, donc sont indépendantes, et x est un programme de base. □

Définition 5.3.5 (Paramétrisation des programmes de base)**.** *Soit $\mathcal{B}$ la famille des parties B à m éléments de $\{1,\ldots,n\}$ telles que A_B soit régulière et $A_B^{-1}b\geq 0$. Alors l'application $\mathcal{B}\to\mathbb{R}^n$, $B\to s(B)=\begin{pmatrix}A_B^{-1}b\\0\end{pmatrix}$ est une paramétrisation de l'ensemble* $\operatorname{Ext} C$ *des programmes de base. C'est en particulier une surjection de $\mathcal{B}$ sur* $\operatorname{Ext} C$.

Cette paramétrisation n'est pas nécessairement injective : il se peut que $s(B)=s(B')=x$, avec $B\neq B'$. Comme $N(s(B))\supset F=\overline{B}$, on voit que nécessairement $N(x)\supset F\cup F'$ a plus de $n-m$ éléments, et donc x est un programme de base dégénéré. A contrario, un programme de base non dégénéré, qui a donc exactement m coordonnées non nulles, est paramétrisé par un unique élément de $\mathcal{B}$.

Définition 5.3.6. *Si $B\in\mathcal{B}$, les coordonnées de x_B sont appelées variables de base, celles de x_F sont les variables libres.*

On s'intéresse ensuite aux programmes de base voisins :

Définition 5.3.7 (Programmes de base voisins)**.** *Deux éléments B,B' de $\mathcal{B}$ sont appelés voisins si $B\cap B'$ a $(m-1)$ éléments. Dans ce cas, $B\cup B'$ a $(m+1)$ éléments, et donc $F\cap F'$ en a $(n-m-1)$.*

Lemme 5.3.8. *Deux programmes de base différents sont voisins si et seulement si ils peuvent être paramétrisés par des éléments de $\mathcal{B}$ voisins.*

DÉMONSTRATION : Si $B,B'\in\mathcal{B}$ sont voisins, les programmes de bases $s(B),s(B')$ appartiennent tous deux à la variété affine Δ définie par :

$$\begin{cases} Ax=b \\ x_j=0,\ \forall j\in F\cap F' \end{cases}$$

Les m formes linéaires $f_i : y \to A_i.y$, $i = 1, \dots, m$ et les $n-m$ formes linéaires e_j^*, $i \in F$ forment un système de rang n. En effet, cela revient à l'indépendance des colonnes A^j, $j \in B$. Ces vecteurs forment donc une base de $\mathbb{R}^n$. Par ailleurs $F \cap F'$ a $n-m-1$ éléments.

Par conséquent le système $(f_i, i = 1, \dots, m; e_j^*, j \in F \cap F')$ est de rang $n-1$, et Δ est donc une droite affine.

Comme chaque ensemble $F_j = \{x \in C \ / \ x_j = 0\}$, lorsqu'il est non vide, est l'intersection de C avec un hyperplan d'appui, donc est une face de C, $a = \Delta \cap C$ est une face de C, contenant $s(B)$ et $s(B')$. C'est donc une arête de C, et elle est nécessairement égale à $[s(B), s(B')]$. Donc $s(B)$ et $s(B')$ sont voisins.

Réciproquement si s et s' sont deux programmes de base voisins, il résulte du Lemme 5.1.7 que le milieu $x = \dfrac{s+s'}{2}$ vérifie $N(x) \subset N(s) \cap N(s')$, et que le système de formes linéaires $f_i : y \mapsto A_i y$, $i = 1, \dots, m$, et $e_j^* : y \mapsto y_j$, $j \in N(x)$ est de rang $n-1$.

Par contre les systèmes $\{f_i, i = 1, \dots, m; e_j^*, j \in N(s)\}$ et $\{f_i, i = 1, \dots, m; e_j^*, j \in N(s')\}$ sont de rang n ; on peut donc trouver $j_0 \in N(s) \setminus N(x)$ et $j_0' \in N(s') \setminus N(x)$ tels que les systèmes $\{f_i, i = 1, \dots, m; e_j^*, j \in N(x) \cup \{j_0\}\}$ et $\{f_i, i = 1, \dots, m; e_j^*, j \in N(x) \cup \{j_0'\}\}$ soient de rang n.

En raisonnant comme dans la preuve du Lemme 5.3.2, on en déduit que le système de colonnes $(A^j)_{j \notin N(x)}$ n'est pas indépendant, mais qu'il le devient dès qu'on lui enlève la j_0-ième colonne (ou la j_0'-ième).

Complétons le système libre $(A^j)_{j \notin N(x), j \neq j_0}$ en une base $(A^j)_{j \in B}$ de $\mathbb{R}^m$. Alors $j_0 \notin B$, $j_0' \in B$ et si B' est déduit de B en remplaçant j_0' par j_0, le système $(A^j)_{j \in B'}$ est aussi générateur, donc une base de $\mathbb{R}^m$. Soient F et F' les complémentaires de B et B', on a $F \subset N(x) \cup \{j_0\} \subset N(s)$ et $F' \subset N(x) \cup \{j_0'\} \subset N(s')$.

Il en résulte que $A_B s_B = As = b$, $A_{B'} s'_{B'} = As' = b$, d'où $s_B = A_B^{-1} b$, $s'_{B'} = A_{B'}^{-1} b$, et finalement B, B' sont des éléments de $\mathcal{B}$ (voisins) et $s = s(B)$, $s' = s(B')$. $\square$

Dans le cas des arêtes infinies issues d'un programme de base, on a la paramétrisation suivante :

Lemme 5.3.9. *Si a est une arête infinie issue du sommet s, il existe $B \in \mathcal{B}$ paramétrisant s et un élément j_0 du complémentaire F de B tel que $u = \begin{pmatrix} -A_B^{-1} A^{j_0} \\ e_{j_0} \end{pmatrix}$ soit un vecteur directeur de a. Réciproquement pour tout $B \in \mathcal{B}$ et tout $j_0 \notin B$ tels que $A_B^{-1} A^{j_0} \leq 0$, la demi-droite issue de $s(B)$ de vecteur directeur $u = \begin{pmatrix} -A_B^{-1} A^{j_0} \\ e_{j_0} \end{pmatrix}$ est une arête de C.*

DÉMONSTRATION : Soit a une arête infinie issue de s, et $x = s$ un point de a, différent de s. Alors $N(x) \subset N(s)$. Comme dans la preuve du lemme précédent, le système de colonnes $(A^j)_{j \notin N(x)}$ est lié, mais on peut trouver $j_0 \in N(s) \backslash N(x)$ tel que le système $(A^j)_{j \notin N(x), \neq j_0}$ soit libre. On complète ce sytème en une base $(A^j)_{j \in B}$ de $\mathbb{R}^m$. Alors $j_0 \notin B$. Comme le complémentaire F de B est inclus dans $N(s)$, on a $s(B) = s$.
D'autre part $F \setminus \{j_0\} \subset N(x)$; donc $x_F = x_{j_0} e_{j_0}$. Par ailleurs $b = Ax = A_B x_B + A x_F = A_B x_B + x_{j_0} A^{j_0}$; comme $b = A_B s_B$, on en déduit par différence que $A_B(x_B - s_B) = -x_{j_0} A^{j_0}$ et donc $x_B - s_B = -x_{j_0} A_B^{-1} A^{j_0}$. Finalement $x - s = x_{j_0} \begin{pmatrix} -A_B^{-1} A^{j_0} \\ e_{j_0} \end{pmatrix}$.
Réciproquement, soient $B \in \mathcal{B}$ tel que $s(B) = s$, et $j_0 \notin B$ tels que $A_B^{-1} A^{j_0} \leq 0$. Le vecteur $u = \begin{pmatrix} -A_B^{-1} A^{j_0} \\ e_{j_0} \end{pmatrix}$ est ≥ 0, et vérifie $Au = 0$, o, en déduit que pour tout $\lambda \geq 0$, le point $x = s + \lambda u$ est dans C.
Par ailleurs, $u_j = 0$ pour tout $j \in F \backslash \{j_0\}$. Donc u annule $n-1$ formes linéaires indépendantes f_i,$i = 1, \ldots, m$ et E'_j, $j \in F \setminus \{j_0\}$, qui correspondent toutes à des contraintes actives en s (l'indépendance de ces formes découle du fait que f_i, $i = 1, \ldots, m$ et e_j^*, $j \in F$ forment une base de $\mathbb{R}^n$). D'après le Lemme 5.1.7, la demi-droite de sommet s et de vecteur directeur u est une arête de C. □

On déduit de cette étude du convexe C, le Test de la fonction objectif pour résoudre $(\mathcal{P}_s)$:

Définition 5.3.10 (Vecteur test)**.** *Soit (B, F) une partition de $\{1, \ldots, n\}$, avec $B \in \mathcal{B}$. On a pour tout programme admissible x :*

$$\begin{aligned} f(x) = c.x &= c_B x_B + c_F x_F \\ &= c_B(A_B^{-1}(b - A_F x_F)) + c_F x_F \\ &= c_B A_B^{-1} b + (c_F - c_B A_B^{-1} A_F) x_F \\ &= c_B A_B^{-1} b + tx = f(s(B)) + t(B)x \end{aligned}$$

où le vecteur ligne $t(B) = (0 \mid c_F - c_B A_B^{-1} A_F)$ est appelé vecteur test.

Lemme 5.3.11. *i) Si $t(B) \geq 0$, le programme de base $s(B)$ est optimal.*
ii) S'il existe un indice ℓ tel que $t(B)_\ell < 0$ et que $A_B^{-1} A^\ell \leq 0$, on a $\inf \mathcal{P} = -\infty$.

DÉMONSTRATION : *i)* Le calcul précédent montre que pour tout programme admissible x, on a $f(x) = f(s(B)) + t(B)x$. Si $t(B) \geq 0$, on a $t(B)x \geq 0$, d'où $f(x) \geq f(s(B))$, et $s(B)$ est un programme de base optimal.

ii) Si $A_B^{-1}A^\ell \leq 0$, alors $u = \begin{pmatrix} -A_B^{-1}A^\ell \\ e_\ell \end{pmatrix}$ est vecteur du cône asymptotique. En fait u dirige une arête infinie issue de $s(B)$. Si $t_\ell > 0$, on a $f(u) = f(x + u) - f(x) = t(B)u = t(B)_\ell < 0$. D'où $\inf \mathcal{P} = -\infty$. $\square$

5.4 Application : l'algorithme du simplexe

Commençons par le passage d'un programme de base à un programme voisin :

A tout élément B de $\mathcal{B}$ on associe la matrice à $(m+1)$ lignes et $(n+1)$ colonnes obtenue en bordant la matrice $A_B^{-1}A$ (dont les colonnes sont indexées par les entiers $1, \dots, n$ et les lignes par les éléments de B) par une ligne et une colonne supplémentaire (indexées par l'entier $n + 1$) :

$$M(B) = \begin{pmatrix} A_B^{-1}A & A_B^{-1}b \\ t(B) & d(B) \end{pmatrix}$$

où l'on a posé $d(B) = -c_B A_B^{-1} b = -f(s(B))$.
D'après la définition du vecteur test $t(B)$, la valeur de la fonction objectif en tout programme admissible x est $f(x) = t(B)x - d(B)$.

Proposition 5.4.1. *Soient B et B' deux éléments de $\mathcal{B}$ voisins. Supposons qu'on passe de B à B' en enlevant k et ajoutant ℓ. On passe alors de $M(B)$ à $M(B')$ en pivotant autour de l'élément indexé par (k, ℓ), c'est à dire par les opérations élémentaires sur les lignes suivantes :*
Pour la ligne de $M(B')$ indexée par ℓ :

$$M(B')_\ell = \frac{1}{(A_B^{-1}A)_k^\ell} M(B)_k$$

Pour toutes les autres lignes, indexées par $j \in B \cap B'$ ou $j = n + 1$:

$$M(B')_j = M(B)_j - \frac{(A_B^{-1}A)_j^\ell}{(A_B^{-1}A)_k^\ell} M(B)_k$$

Remarque 5.4.2. On a $(A_B^{-1}A)_k^\ell \neq 0$. En effet la matrice $A_B^{-1}A_{B'}$ est régulière. Les colonnes de cette matrice sont identifiées aux vecteurs de $\mathbb{R}^B$ suivants : $(A_B^{-1}A_{B'})^j = e_j$ si $j \in B \cap B' = B \setminus \{k\}$, $(A_B^{-1}A_{B'})^k = \sum_{j \in B} (A_B^{-1}A_{B'})_j^k e_j$. Ce dernier vecteur n'est indépendant des précédents que si $(A_B^{-1}A)_k^\ell \neq 0$.

Démonstration : Considérons le système d'équations :

$$(I) \quad \begin{cases} Ax = b \\ c.x = -\xi \end{cases}$$

où $x \in \mathbb{R}^n$ et $\xi \in \mathbb{R}$. On sait que $c.x = t(B)x - d(B)$ dès que $Ax = b$. Donc le système (I) équivaut à

$$(II) \quad \begin{cases} A_B^{-1}Ax = A_B^{-1}b \\ \xi + t(B)x = d(B) \end{cases}$$

Quand on remplace B par B', on obtient un nouveau système équivalent :

$$(II') \quad \begin{cases} A_{B'}^{-1}Ax = A_{B'}^{-1}b \\ \xi + t(B')x = d(B') \end{cases}$$

Revenons au système (II) ; il s'écrit matriciellement :

$$\begin{pmatrix} 0 & I_B & A_B^{-1}A_F \\ 1 & 0 & t(B)_F \end{pmatrix} \begin{pmatrix} \xi \\ x_B \\ x_F \end{pmatrix} = \begin{pmatrix} A_B^{-1}b \\ d(B)) \end{pmatrix}$$

Quand on fait l'opération sur les lignes mentionnée dans l'énoncé, la colonne correspondant à l'indice ℓ ne comporte plus que des 0, sauf sur la ligne nouvellement indexée par ℓ, anciennement par k, où le coefficient vaut 1. Autrement dit on a obtenu un système équivalent de la forme :

$$\begin{pmatrix} 0 & I_{B'} & D \\ 1 & 0 & t \end{pmatrix} \begin{pmatrix} \xi \\ x_{B'} \\ x_{F'} \end{pmatrix} = \begin{pmatrix} z \\ d \end{pmatrix}$$

après réarrangement des colonnes pour regrouper celles indexées par B', où D est une matrice à m lignes, $(n-m)$ colonnes, t une matrice ligne (à $(n-m)$ colonnes), z une matrice colonne à m lignes et d un réel.
On a donc obtenu le système équivalent :

$$(III) \quad \begin{cases} x_{B'} + Dx_{F'} = z \\ \xi + tx_{F'} = d \end{cases}$$

Les systèmes II' et III ne peuvent être équivalents pour tous $x_{B'}, x_{F'}$ et ξ que si $D = A_{B'}^{-1}A_{F'}$, $t = t(B')_{F'}$, $z = A_{B'}^{-1}b$ et $d = d(B')$. $\square$

Progression dans les programmes de base. *Phase II.*

On part d'une solution de base $s(B)$, supposée trouvée dans une phase préliminaire (Phase I). On dispose donc de la matrice $\mathcal{A}(B) = (A_B)^{-1}A$, du programme de base $s(B) = A_B^{-1}b$, du vecteur test $t(B)$, de la valeur $d(B)$.

* Si le vecteur test est positif, le programme de base $s(B)$ est optimal et la valeur optimale de la fonction objectif est $-d(B)$.
* Sinon, on examine les colonnes de $\mathcal{A}(B)$ correspondant aux indices ℓ pour lesquelles la coordonnée $t(B)_\ell < 0$.

** Si pour l'une de ces colonnes $\mathcal{A}(B)^\ell \leq 0$, il n'y a pas de programme optimal, et la valeur optimale de la fonction objectif est $-\infty$.

** Sinon on choisit ℓ minimisant $t(B)_\ell$. On choisit k tel que $\mathcal{A}(B)_k^\ell > 0$, et minimisant $\dfrac{s(B)_k}{\mathcal{A}(B)_k^\ell}$, et on pivote autour du coefficient (k, ℓ).

Le nouveau programme obtenu s' a pour coordonnées : $s'_\ell = \dfrac{s(B)_k}{\mathcal{A}(B)_k^\ell}$, et pour $j \neq \ell$: $s'_j = s(B)_j - \dfrac{\mathcal{A}(B)_k^j}{\mathcal{A}(B)_k^\ell} s(B)_k$ Si $\mathcal{A}(B)_k^j \leq 0$, il est clair que $s(B')_j \geq 0$; sinon, d'après le choix de k, on a $s(B')_j = \mathcal{A}(B)_j^\ell [\dfrac{s(B)_j}{\mathcal{A}(B)_j^\ell} - \dfrac{s(B)_k}{\mathcal{A}(B)_k^\ell}] \geq 0$. Donc s' est bien un programme admissible, c'est à dire que $B' \in \mathcal{B}$ et $s' = s(B')$ est un nouveau programme de base.

La nouvelle valeur de la fonction objectif est :

$$f(s(B')) = -d(B') = -d(B) + \frac{s(B)_k}{\mathcal{A}(B)_k^\ell} t_\ell$$

donc $f(s(B')) - f(s(B)) = \dfrac{s(B)_k}{\mathcal{A}(B)_k^\ell} t_\ell \leq 0$
Si $s(B)_k \neq 0$ (en particulier si $s(B)$ n'est pas dégénéré), on obtient bien $f(s(B')) < f(s(B))$. Si en revanche $s(B)_k = 0$ alors en fait $s(B') = s(B)$, c'est à dire qu'on a obtenu une nouvelle paramétrisation du même programme de base.

Néanmoins même dans ce cas on peut continuer l'algorithme : on parcourt ainsi une suite de paramétrisations $B', B'', B''' \ldots$ de $s = s(B)$.
Deux cas se présentent :

*** soit ces paramétrisations successives de s sont distinctes : alors, puisque s a un nombre fini de paramétrisations, on finit par obtenir soit un nouveau programme de base et donc diminuer strictement la fonction objectif, soit par obtenir une arête infinie (d'où $\inf \mathcal{P} = -\infty$).

*** soit on retombe sur une paramétrisation de s déjà obtenue : il y a cyclage, l'algorithme boucle.

En l'absence de cyclage l'algorithme s'arrête en un nombre fini d'étapes et donne un programme de base optimal (ou bien $\inf \mathcal{P} = -\infty$).

On peut éviter théoriquement le cyclage en modifiant l'algorithme. Dans la pratique ce phénomène n'intervient quasi jamais.

Détermination d'un programme de base initial. *Phase I.*

Lorsqu'il n'y a pas de programme de base évident pour le problème $(\mathcal{P}_s)$, on considère le problème linéaire auxiliaire :

$$(\mathcal{P}_{aux}) \begin{cases} \text{minimiser} & g(x,y) = \sum_{i=1}^{m} y_i, \; x \in \mathbb{R}^n, y \in \mathbb{R}^m \\ \text{sous les contraintes} & Ax + y = b \text{ et } x \geq 0, y \geq 0 \end{cases}$$

Remarquons que, quitte à modifier la matrice A en multipliant certaines de ses lignes par -1, on peut toujours écrire le problème $(\mathcal{P}_s)$ de manière que $b \geq 0$. Dans ce cas le problème auxiliaire a un programme de base évident $(0, b)$, correspondant à la base $B = (n+1, \dots, n+m)$, la sous-matrice correspondante étant l'identité I_B. On peut donc lui appliquer l'algorithme du simplexe et obtenir un programme de base optimal (s, z).
Si $z \neq 0$ alors l'équation $Ax = b$ n'a pas de solution, ce qui veut dire que le problème $(\mathcal{P}_s)$ n'admet aucun programme admissible.
Dans le cas contraire le programme de base optimal de $(\mathcal{P}_{aux})$ ainsi obtenu est de la forme $(s, 0)$. Alors s est un programme de base pour $(\mathcal{P})$, car les colonnes de A correspondant aux coordonnées non nulles de s sont les mêmes que celle de la matrice $(A \mid I_m)$ correspondant aux coordonnées non nulles de $(s, 0)$, donc sont indépendantes.

5.5 Application : l'acquisition comprimée

Dans cette section nous esquissons (sans démonstrations) quelques techniques d'acquisition comprimée, qui – sous certains conditions – permet de récupérer un signal à partir d'un très petit nombre d'observations.

On considère un vecteur inconnu $x \in E = \mathbb{R}^n$ que on veut reconstruire en effectuant des mesures linéaires : $f_1(x), \dots, f_m(x)$, où $f_i \in E'$. Une approche équivalente est de considèrer une matrice de mesure M à m lignes et n colonnes, l'observation étant $Mx \in \mathbb{R}^m$.

Evidemment, dans le cas général, on a besoin n mesures pour récupérer x exactement et on peut prendre comme (f_i) n'importe quelle base de E'. (Dans la pratique, il pourrait être nécessaire d'augmenter un peu le nombre d'observations pour compenser des erreurs ou le bruit.) Cependant, ce nombre peut être considérablement réduit si nous savons *a priori* que le vecteur x est parcimonieux (ou éparse), c'est-à-dire que la plupart de coordonnées de x sont nulles. Il est convenable de quantifier cette proprieté.

Définition 5.5.1 (Parcimonie ou éparsité d'un vecteur). *Pour $x \in \mathbb{R}^n$ (ou pour $x \in c_{00}$), notons par $\|x\|_0$ le nombre de coordonnées de x qui sont non nulles. Si $\|x\|_0 \leq s$, on dit que le vecteur x est s-parcimonieux ou s-éparse.*

Proposition 5.5.2. *Si $n \geq m \geq 2s$, alors il existe une matrice de mesure M à m lignes et n colonnes telle que pour tout $y \in \mathbb{R}^m$, le problème*

$$(\mathcal{P})_s \qquad x \in \mathbb{R}^n,\ Mx = y,\ \|x\|_0 \leq s$$

admet au plus une solution.

DÉMONSTRATION : Supposons que rang $M = m$, $Mx' = Mx'' = y$ avec $\|x'\|_0 \leq s$ et $\|x''\|_0 \leq s$, et posons $x = x' - x''$. Alors $x \in \operatorname{Ker} M$ et $\|x\|_0 \leq 2s$; on veut montrer qu'il suit que $x = 0$. Notons que l'ensemble de de vecteurs $2s$-parcimonieux dans $\mathbb{R}^n$ est contenu dans une réunion d'un nombre fini de sous-espaces vectoriels de $\mathbb{R}^n$ de dimension $2s \leq m$ chacun, tandis que $\operatorname{Ker} M$ est de dimension $n - m$. Alors, génériquement, l'intersection de $\operatorname{Ker} M$ avec tous ces espaces contient seulement le point 0, comme requis. □

Malheureusement, l'algorithme suggéré par Proposition 5.5.2 n'est pas pratique : pour un entier s donné, on a $\binom{n}{s}$ des s-uplets de coordonnées à examiner. Il se trouve qu'il y a un moyen simple d'éviter cette difficulté.

Théorème 5.5.3. *Il existe une constante finie $C > 0$ tell que, pour tout $s \leq n$, il existent $m \leq Cs\big(1 + \log(n/s)\big)$ et une matrice de mesure M à m lignes et n colonnes vérifiant la conclusion de la Proposition 5.5.2 et telle que, pour tout $y \in \mathbb{R}^m$, les deux problèmes de minimisation*

$$(\mathcal{P}\,0) \qquad \begin{cases} \textit{minimiser} \quad f(x) = \|x\|_0 \\ \textit{sous les contraintes } x \in \mathbb{R}^n, Mx = y \end{cases}$$

et

$$(\mathcal{P}\,1) \qquad \begin{cases} \textit{minimiser} \quad g(x) = \|x\|_1 \\ \textit{sous les contraintes } x \in \mathbb{R}^n, Mx = y \end{cases}$$

sont équivalents.

Par conséquent, si le problème $Mx = y$ admet une solution s-parcimonieuse, cette solution est unique et peut être obtenue en résolvant le problème $(\mathcal{P}\,1)$. Le problème $(\mathcal{P}\,1)$ appartient *a priori* dans le cadre d'optimisation convexe (voir chapitres 7 et 8), mais on peut le facilement reformuler dans le language de la programmation linéaire (Exercice !) :

$$(\mathcal{P}) \qquad \begin{cases} \text{minimiser} \quad h(u) = \sum_i u_i \\ \text{sous les contraintes } u, x \in \mathbb{R}^n,\ Mx = y,\ -u \leq x \leq u \end{cases}$$

Remarque 5.5.4. Dans la pratique, on rencontre plusieurs complications. Premièrement, on peut avoir des erreurs dès mesure ou les solutions peuvent être seulement "approximativement" parcimonieuse ; ces difficultés peuvent être surmonter en utilisant des versions "robustes" du Théorème 5.5.3 (voir Réfs [6],[7], où une démonstration du Théorème peut être également trouvée). Deuxièmement, un aspect important des applications est d'identifier des bases dans lesquelles les vecteurs intéressants sont parcimonieux ; cet aspect appartient au domaine d'application ou à la modélisation plutôt qu'à l'analyse convexe.

Chapitre 6
Structure des convexes en dimension quelconque

6.1 Convexes dans l'espace de Hilbert

Soit $\mathcal{H}$ un espace de Hilbert réel, c'est à dire un espace vectoriel muni d'un produit scalaire $(x, y) \mapsto \langle x, y\rangle$, complet pour la norme hilbertienne associée $\|x\| = \sqrt{\langle x, x\rangle}$. Comme dans le cas de la dimension finie, si $x, y \in \mathcal{H}$, on pose $d(x, y) = \|x - y\|$ et pour toute partie $A \subset \mathcal{H}$, $d(x, A) = \inf\{d(x, y) \ / \ y \in A\}$.

Théorème 6.1.1 (Existence d'une solution optimale)**.** *Soit C un convexe fermé de $\mathcal{H}$, et $f : \mathcal{H} \to \mathbb{R}$ une fonction convexe sci. On suppose que C est borné, ou bien que $f(x) \to +\infty$ quand $\|x\| \to +\infty$. Le problème de minimisation sous contrainte*

$$\text{Minimiser } f(x) \text{ sous la contrainte } x \in C$$

admet une solution optimale au moins. Si de plus la fonction f est strictement convexe, cette solution optimale est unique.

Définition 6.1.2 (Solution optimale)**.** *Une solution optimale (d'un problème de minimisation) est un point $x_0 \in C$ tel que*

$$f(x_0) = \inf\{f(x) \ / \ x \in C\}$$

Remarque 6.1.3. Une fonction f définie sur un EVN (ou sur une partie non-bornée d'un EVN) est dite *coercive* si $\lim_{\|x\|\to\infty} f(x) = +\infty$. Le Théorème 6.1.1 dit alors qu'une fonction convexe sci coercive atteint son minimum sur toute partie convexe fermé d'un espace de Hilbert.

La preuve s'appuie sur le théorème de l'intersection suivant :

Théorème 6.1.4. *Dans un espace de Hilbert toute suite décroissante de convexes fermés bornés non vides a une intersection non vide.*

Démonstration : Soit C_n une suite décroissante de convexes fermés bornés de l'espace de Hilbert $\mathcal{H}$. La suite numérique $d(0, C_n)$ est croissante, majorée par le rayon R d'une boule de centre 0 contenant C_1, donc convergente. Soit d sa limite. Pour tout n, soit $x_n \in C_n$ tel que $\|x_n\| \leq d(0, C_n) + 1/n$. On utilise l'égalité du parallélogramme :

$$\forall u, v \in \mathcal{H}, \quad \|u - v\|^2 + \|u + v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

que l'on applique à $u = x_n$ et $v = x_m$. On a pour tous n, m avec $m \leq n$:

$$\|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|y_m\|^2) - \|x_n + x_m\|^2 \leq 2(\|x_n\|^2 + \|x_m\|^2) - 4d(0, C_m)^2$$

car $\left\| \dfrac{x_n + x_m}{2} \right\| \geq d(0, C_m)$, puisque C_m contient x_m et x_n, donc aussi leur milieu.

Comme les réels $\|x_n\|$, $\|x_m\|$ et $d(0, C_m)$ ont même limite finie d lorsque $n \geq m \to \infty$, on voit que $2(\|x_n\|^2 + \|x_m\|^2) - 4d(0, C_m)^2 \to 0$ lorsque $n \geq m \to \infty$, et donc aussi $\|x_n - x_m\|^2 \to 0$. Donc la suite (x_m) est de Cauchy. Soit x sa limite. Comme $x_n \in C_m$ pour tout $n \geq m$, et que C_m est fermé, on a donc $x \in C_m$, et ceci pour tout m. $\square$

Démonstration : (Preuve du Théorème 6.1.1) L'hypothèse entraîne que pour tout $a \in \mathbb{R}$, l'ensemble $\{x \in \mathcal{H} \ / \ f(x) \leq a\}$ est borné. Soit $m = \inf\{f(x) \ / \ x \in C\}$, qui a priori peut valoir $-\infty$. Soit (m_n) une suite de réels $m_n > m$, tendant vers m quand $n \to \infty$. Pour tout n, posons $C_n = \{x \in C \ / \ f(x) \leq m_n\}$: alors chaque C_n est un convexe fermé non vide, et cette suite est décroissante. De plus ces convexes sont bornés. Il existe donc un point x_0 dans leur intersection. On doit avoir $f(x_0) \leq m_n$ pour tout n, et donc $f(x_0) \leq m$ et par conséquent $f(x_0) = m$. Comme f est à valeur réelles, ceci entraîne $m > -\infty$.

L'ensemble des solutions $\{x \in C / f(x) \leq m\}$ est un convexe fermé non vide. Si f est strictement convexe cet ensemble est réduit à un point en raison de l'inéquation stricte :

$$x_1 \neq x_2 \quad \Longrightarrow \quad f\Big(\frac{x_1 + x_2}{2}\Big) < \frac{f(x_1) + f(x_2)}{2} \quad \square$$

Remarque 6.1.5. On seulement utilisé la quasi convexité de f, c'est à dire que pour tout $a \in \mathbb{R}$, l'ensemble $A_{f,a} = \{x \in \mathcal{H} \ / \ f(x) \leq a\}$ est convexe. On peut aussi autoriser f à prendre les valeurs $\pm\infty$.

Voici maintenant le théorème fondamental de projection sur les convexes fermés, que l'on a déja vu en dimension finie au chapitre 2, Théorème 3.2.1 :

Théorème 6.1.6. *Soit C un convexe fermé de $\mathcal{H}$, et x un point de $\mathcal{H}$. Il existe un unique point $y_0 \in C$ réalisant la distance de C à x, c'est à dire tel que $d(x, y_0) = d(x, C)$). Ce point, appelé projection de x sur C, est caractérisé par par la propriété*

$$\langle x - y_0, y - y_0 \rangle \leq 0 \text{ pour tout } y \in C$$

Démonstration : Existence et unicité de la projection :
La fonction $y \to d(x, y) = \|x - y\|$ est convexe, par convexité de la norme, et positive. Il s'ensuit que son carré est convexe. En appliquant l'égalité du parallélogramme à $u = x - x_1$ et $v = x - x_2$, on montre aisément que la fonction $y \to d(x, y)^2$ est strictement convexe. On peut appliquer le Théorème 6.1.1 au convexe C et à la fonction $y \to d(x, y)^2$.

La caractérisation de la projection se démontre comme en dimension finie, voir 3.2.1. □

On en déduit la représentation des formes linéaires sur un espace de Hilbert :

Corollaire 6.1.7. *Toute forme linéaire continue sur un espace de Hilbert $\mathcal{H}$ est de la forme $f_{x_0} : x \to \langle x_0, x \rangle$, où x_0 est élément de $\mathcal{H}$.*

Démonstration : Soit $f : \mathcal{H} \to \mathbb{R}$ une forme linéaire continue non nulle, et H l'hyperplan affine d'équation $f(x) = 1$. Soit x_0 la projection de 0 sur H. Alors $\langle x_0, x - x_0 \rangle = 0$ pour tout $x \in H$; c'est à dire que $f(x) = 1$ entraîne $\langle x_0, x \rangle = \langle x_0, x \rangle = \|x_0\|^2$. Par linéarité de f et f_{x_0} on en déduit $f_{x_0} = \|x_0\|^2 f$, donc $f = af_{x_0} = f_{ax_0}$ avec $a = \|x_0\|^{-2}$. □

Remarque 6.1.8. Il résulte du corollaire précédent que l'on peut identifier le dual $\mathcal{H}'$ de $\mathcal{H}$ à l'espace $\mathcal{H}$ lui même.

De manière plus précise on a une application linéaire naturelle $\mathcal{H} \to \mathcal{H}'$, $x \mapsto f_x$, qui est surjective d'après le corollaire précédent. On a $|f_x(y)| \leq \|x\| \|y\|$, d'après l'inégalité de Cauchy-Schwartz, donc $\|f_x\| \leq \|x\|$; l'égalité $f_x(x) = \|x\|^2$ montre qu'en fait $\|f_x\| = \|x\|$; donc l'application $x \to f_x$ est une isométrie linéaire et bijective.

Grâce à ces résultats, les théorèmes que nous avons vus dans le cadre des EVN de dimension finie au chapître 2 ont un analogue dans le cadre hilbertien :

Théorème 6.1.9 (Séparation dans un espace de Hilbert #1)**.** *Si C est un convexe fermé non vide d'un espace de Hilbert $\mathcal{H}$ et x_0 un point de $\mathcal{H}$ n'appartenant pas à C il existe un hyperplan affine H séparant strictement x_0 de*

C ; c'est à dire que C et x_0 sont inclus chacun dans un demi-espace ouvert différent défini par H. En particulier C est inclus dans l'un des demi-espaces affines associés à H.

Démonstration : C'est la même qu'en dimension finie, voir 3.2.2. □

Remarque 6.1.10. Il revient au même de dire que si x_0 n'appartient pas au convexe fermé C, il existe une forme linéaire f sur $\mathcal{H}$ telle que $\sup_{x\in C} f(x) < f(x_0)$ ou encore qu'il existe $u \in \mathcal{H}$ tel que $\sup_{x\in C}\langle u, x\rangle < \langle u, x_0\rangle$.

Le deuxième théorème de séparation pour des convexes quelconques en dimension finie (3.2.4) admet l'analogue suivant pour des convexes ouverts dans les espaces de Hilbert :

Théorème 6.1.11 (Séparation dans un espace de Hilbert #2)**.** *Si C est un convexe ouvert d'un espace de Hilbert $\mathcal{H}$ et x_0 un point de $\mathcal{H}$ n'appartenant pas à C, il existe un hyperplan affine H de $\mathcal{H}$, contenant x_0 et disjoint de C. De plus C est inclus dans l'un des demi-espaces affines associés à H.*

Démonstration : a) Cas où $\mathcal{H}$ est séparable.
Dans ce cas il existe une suite croissante de sous-espaces vectoriels de dimension finie E_n, contenant x_0 dont la réunion est dense dans $\mathcal{H}$. Pour tout n posons : $C_n = C \cap E_n$: c'est un convexe ouvert de E_n, ne contenant pas x_0.
On a $C_n \neq \emptyset$ pour n assez grand, par densité de $\bigcup\limits_n E_n$, on peut donc supposer $C_1 \neq \emptyset$.
On construit alors une suite de sous-espaces affines H_n, tels que $H_0 = \{x_0\}$ et que pour tout $n \geq 1$:
i) H_n est un hyperplan affine de E_n contenant H_{n-1} ;
ii) H_n ne rencontre pas C.

Pour trouver H_1, on applique le Théorème 3.2.4 en dimension finie à x_0 et C_1. Supposons avoir construit H_n. En appliquant le Théorème 3.2.8 en dimension finie à H_n et C_{n+1}, considérés dans l'espace de dimension finie E_{n+1}, on obtient un hyperplan affine H_{n+1} de E_{n+1}, contenant H_n et disjoint de C_{n+1}, donc de C.
Posons $L = \bigcup\limits_n H_n$ et $H = \overline{L}$. C'est un sous-espace affine de $\mathcal{H}$, qui contient x_0. H est disjoint de C car le complémentaire de C est fermé et contient L, donc son adhérence H.

Montrons que H est un hyperplan affine de $\mathcal{H}$, ce qui démontrera le théorème : Par translation, on peut supposer que $x_0 = 0$, de sorte que H est un sous-espace vectoriel de $\mathcal{H}$. Fixons $y_0 \in C_1$ et considérons $\hat{H} = H \oplus \mathbb{R}y_0$. Il s'agit de montrer que $\hat{H} = \mathcal{H}$.

On a $\hat{H} \supset H_n \oplus \mathbb{R}y_0 = E_n$, puisque H_n est un hyperplan de E_n ne contenant pas y_0, donc $\hat{H}$, contenant la réunion des E_n, est dense dans E. Il suffit donc de montrer que $\hat{H}$ est fermé.
Or si $x \in \mathcal{H}$ est limite d'une suite $(x_k + \rho_k y_0)_k$ de points de $\hat{H}$, on a :

$$d(x_k + \rho_k y_0, H) = d(\rho_k y_0, H) = |\rho_k|\, d(y_0, H)$$

et comme le membre de gauche tend vers $d(x, H)$, et que $d(y_0, H) \neq 0$, on a $|\rho_k| \to d(x, H)/d(y_0, H)$. Quitte à extraire une sous-suite, on peut supposer les ρ_k de signe constant, donc convergeant vers $\rho = \pm d(x, H)/d(y_0, H)$. Alors la suite x_k converge vers $y = x - \rho y_0$, qui doit appartenir à H (car H est fermé). Donc $x = y + \rho y_0 \in H \oplus \mathbb{R}x_0 = \hat{H}$.

b) Cas général.
Le cas général peut se démontrer de manière analogue au cas séparable ci-dessus par un argument de récurrence transfinie sur le caractère de densité de $\mathcal{H}$, c'est à dire le plus petit cardinal d'une partie dense dans $\mathcal{H}$. Pour un espace séparable le caractère de densité est $\aleph_0$, le cardinal de $\mathbb{N}$, qui est aussi le premier cardinal infini. □

Corollaire 6.1.12 (Théorème de Hahn-Banach géométrique dans les espaces de Hilbert)**.** *Soient C_1 et C_2 deux convexes non vides et disjoints de l'espace vectoriel de Hilbert $\mathcal{H}$. On suppose que C_1 est ouvert. Il existe alors un hyperplan affine H séparant strictement C_1 de C_2. Plus précisément il existe une forme linéaire f sur $\mathcal{H}$ et un réel a tels que $\forall x_1 \in C_1, \forall x_2 \in C_2$, $f(x_1) < a \leq f(x_2)$.*

Démonstration : On remarque d'abord que le convexe $C_1 - C_2$ est ouvert, comme réunion des convexes ouverts $C_1 - z$, $z \in C_2$. Comme en dimension finie, on applique alors le Théorème 6.1.11 au point 0 et au convexe ouvert non vide $C_1 - C_2 = \{y - z \,/\, y \in C_1, z \in C_2\}$. D'où il existe f forme linéaire telle que $\forall y \in C_1, z \in C_2, f(y - z) < f(0) = 0$; on en déduit, en notant $a = \inf_{z \in C_2} f(z)$ que $\forall x_1 \in C_1, \forall x_2 \in C_2,\ f(x_1) \leq a \leq f(x_2)$. La première inéquation signifie que C_1 est inclus dans le demi-espace $f(x) \leq a$; comme C_1 est ouvert, il est en fait inclus dans l'intérieur du demi-espace, c'est à dire le demi-espace ouvert $f(x) < a$. □

Corollaire 6.1.13. *Par tout point non intérieur d'un convexe d'intérieur non vide de $\mathcal{H}$ il passe un hyperplan d'appui.*

Démonstration : Comme le convexe C est d'intérieur non vide c'est une conséquence directe du Théorème 6.1.11 appliqué à $\overset{o}{C}$. □

Corollaire 6.1.14. *Tout convexe fermé C d'intérieur non vide distinct de l'espace entier est intersection des demi-espaces fermés qui le contiennent et qui sont associés à des hyperplans d'appui.*

Démonstration : Soit C_1 l'intersection des demi-espaces fermés contenant C associés à des hyperplans d'appui. Alors, $C \subset C_1$. Il suffit donc de montrer l'autre inclusion.

Soit $x_0 \in C_1$ n'appartenant pas à C, et y_0 un point intérieur à C. Alors l'intersection du segment $[y_0, x_0]$ avec C est un segment $[y_0, z_0]$; le point z_0 n'est pas intérieur à C puisqu'on peut l'approcher par des points de $]z_0, x_0]$. Soit H un hyperplan d'appui de C passant par z_0 et disjoint de $\overset{o}{C}$ (Corollaire 6.1.13), et D un demi-espace associé contenant C. Le demi-espace D ne peut contenir x_0 : en effet y_0 est intérieur à C, donc à D ; si x_0 est dans D, alors le point $z_0 \in]y_0, x_0[$ doit aussi être intérieur à D ; ce qui contredit $z_0 \in H$. Donc $x_0 \notin C_1$ ce qui est une contradiction. □

En dimension finie, les ensembles compacts sont exactement les ensembles fermés bornés (voir Chapître 1).
En dimension quelconque, cela n'est plus vrai, même dans les espaces de Hilbert : en effet la boule unité fermée d'un EVN qui n'est pas de dimension finie est fermée bornée mais non compacte à cause du théorème de F. Riecz qui affirme qu'un EVN est de dimension finie si et seulement si sa boule unité fermée est compacte, 8.6.8.
Cependant, dans les espaces de Hilbert, on peut démontrer des analogues des théorèmes concernant les convexes compacts en dimension finie pour les convexes fermés bornés.
Ceci ne sera plus vrai pour des EVN quelconques où il faudra l'hypothèse plus forte de compacité.

Proposition 6.1.15. *Soit C un convexe fermé borné d'un espace de Hilbert. Pour tout hyperplan vectoriel fermé V de $\mathcal{H}$ il existe un hyperplan d'appui fermé de C parallèle à V.*

Démonstration : La preuve est analogue à celle de la Proposition 3.4.3 et du corollaire 3.4.2. On invoque le Théorème 6.1.1 au lieu de la compacité pour assurer qu'une forme linéaire continue f atteint son minimum sur C. En fait, comme $-f$ est aussi linéaire et $\inf_C(-f) = -\sup_C f$, cette forme linéaire atteint aussi son maximum sur C. Le reste de la preuve est inchangé. □

Proposition 6.1.16. *Dans un espace de Hilbert $\mathcal{H}$, tout convexe fermé borné C possède au moins un point extrémal.*

DÉMONSTRATION : Donnons la preuve dans le cas où $\mathcal{H}$ est séparable, c'est à dire où il contient une suite partout dense $(u_n)_{n\geq 1}$. On construit par récurrence une suite décroissante de convexes fermés $(F_n)_{n\geq 0}$ de la manière suivante : $F_0 = C$, et

$$F_{n+1} = \left\{x \in F_n / \ \langle u_n, x\rangle = \inf_{y\in F_n} \langle u_n, y\rangle\right\}$$

Alors F_{n+1} est non vide et l'hyperplan $H_n = \left\{x \in \mathcal{H} \ / \ \langle u_n, x\rangle = \inf_{y\in F_n}\langle u_n, y\rangle\right\}$ est un hyperplan d'appui de F_n.
Donc F_{n+1} est une face fermée de F_n. D'après le Lemme 1.4.11, on voit par récurrence que F_n est une face de C lui-même. D'après le Théorème 6.1.4, l'intersection des F_n est non vide ; c'est une face F de C.
Il reste à remarquer que cette face n'a qu'un seul élément, qui est donc un point extrémal de C. Or supposons que $x_1, x_2 \in F$. En particulier $x_1, x_2 \in F_{n+1}$ pour tout n, et par conséquent $\langle u_n, x_1\rangle = \langle u_n, x_2\rangle$ $(= \inf_{y\in F_n}\langle u_n, y\rangle)$. Par conséquent $x_1 - x_2$ est orthogonal à tous les u_n, donc, par densité, à $\mathcal{H}$ tout entier. Ceci entraîne $x_1 - x_2 = 0$. □

Remarque 6.1.17. Cas non séparable.
Dans le cas non séparable on commence par étendre le Théorème 6.1.1 au cas d' une famille, non nécessairement dénombrable, totalement ordonnée de convexe fermés bornés non vides ; on considère ensuite la famille $\mathcal{F}$ des faces fermées de C. Pour l'inclusion, cet ensemble est inductif, c'est à dire que pour toute famille totalement ordonnée (F_α) d'éléments de $\mathcal{F}$ il existe un élément de $\mathcal{F}$ inclus dans tous les F_α, qui n'est autre que leur intersection.
D'après le théorème de Zorn, $\mathcal{F}$ possède des éléments minimaux. Soit F un tel élément minimal, il suffit de montrer que cette face de C n'a qu'un seul élément.
Or si x_1, x_2 sont deux éléments distincts de F, la forme linéaire $f(x) = \langle x_1 - x_2, x\rangle$ prend au moins deux valeurs distinctes sur F (puisque $f(x_1) - f(x_2) = \|x_1 - x_2\|^2 \neq 0$). Soient m et M le minimum et le maximum de f sur F. Les hyperplans d'appui de F correspondants, d'équations $\{x \in E \ , \ f(x) = m\}$ et $\{x \in E \ , \ f(x) = M\}$, définissent deux faces fermées F_1, F_2 de F disjointes, qui sont aussi des faces de C d'après 1.4.11.
Donc F n'est pas un élément minimal de $\mathcal{F}$, il y a contradiction. □

Théorème 6.1.18 (Krein-Milman, cas Hilbertien)**.** *Dans un espace de Hilbert $\mathcal{H}$, tout convexe fermé borné est l'enveloppe convexe fermée de ses points extrémaux.*

DÉMONSTRATION : Soit C un convexe fermé borné de l'espace de Hilbert $\mathcal{H}$, et $L = \overline{\text{convExt}}\, C$ l'enveloppe convexe fermée de l'ensemble de ses points extrémaux. Il est clair que $L \subset C$.

Supposons que $L \neq C$, il existe donc $x \in C \setminus L$. Soit H un hyperplan séparant strictement x et L ; Soit H_1 celui des hyperplans d'appui de C parallèles à H et situé du même côté que x ; alors $F = H_1 \cap C$ est une face de C disjointe de L. F est aussi un convexe fermé borné de $\mathcal{H}$, donc possède un point extrémal qui est aussi un point extrémal de C d'après 1.4.11 , et qui doit donc appartenir à L. Il y a donc contradiction. □

Corollaire 6.1.19. *Soit C un convexe fermé borné d'un espace de Hilbert, et $f : C \to \mathbb{R}$ une fonction convexe semi-continue inférieurement. On a*

$$\sup_{x \in C} f(x) = \sup_{x \in \operatorname{Ext} C} f(x)$$

En particulier si C est compact et f continue, f atteint son maximum en un point de $\overline{\operatorname{Ext} C}$.

Démonstration : Soit $M = \sup_{x \in \operatorname{Ext} C} f(x)$ (Ce nombre peut être $+\infty$). Si $x \in \operatorname{conv}\operatorname{Ext} C$, cet élément est combinaison convexe d'une famille finie de points $x_i \in \operatorname{Ext} C$: soit $x = \sum_i \lambda_i x_i$. D'où

$$f(x) \leq \sum_i \lambda_i f(x_i) \leq M \sum_i \lambda_i = M$$

L'ensemble $\{x \in C \,/\, f(x) \leq M\}$ contient donc $\operatorname{conv}\operatorname{Ext} C$; comme il est fermé, il contient en fait $\overline{\operatorname{conv}\operatorname{Ext} C} = C$. D'où l'égalité.
Par ailleurs si C est compact, il en est de même de $\overline{\operatorname{Ext} C}$ et si f est continue elle atteint son maximum sur ce compact. □

Dans les espaces de Hilbert qui ne sont pas de dimension finie, la notion d'intérieur relatif n'a pas d'intérêt car la Proposition 3.1.1, qui est essentielle dans l'utilisation de cette notion n'a pas d'analogue : en effet, si $(e_n)_{n \in \mathbb{N}}$ est une base orthonormée d'un espace de Hilbert séparable, l'enveloppe convexe fermée $\overline{\operatorname{conv}}\{\frac{e_N}{n}\ ,\ n \in \mathbb{N}\}$ est compacte donc ne contient pas de boule mais par contre, elle engendre l'espace tout entier.
La décomposition des convexes fermés non vides de dimension finie n'a pas non plus d'analogue bien que certains résultats perdurent, en utilisant la convergence faible.

6.2 Convexes dans les espaces vectoriels normés

Soit E un EVN de dimension quelconque. Contrairement aux espaces de dimension finie ou aux espaces de Hilbert, on n'a pas de théorème de projection sur les convexes fermés ni même sur les sous-espaces vectoriels fermés.

On est donc amené à démontrer d'abord le théorème de séparation pour les convexes ouverts :

Théorème 6.2.1 (Séparation, cas EVN #1)**.** *Soit E un espace vectoriel normé, C un convexe ouvert de E et x_0 un point de E n'appartenant pas à C. Il existe un hyperplan affine fermé passant par x_0, ne rencontrant pas C. Il revient au même de dire qu'il existe une forme linéaire continue f sur E telle que $\forall x \in C,\ f(x) < f(x_0)$.*

Démonstration : La démonstration est identique à celle des espaces de Hilbert (voir Théorème 6.1.11). □

Les corollaires qui suivent se déduisent du théorème comme dens les cas précédents de la dimension finie et des espaces de Hilbert.

Corollaire 6.2.2 (Théorème de Hahn-Banach géométrique)**.** *Soient C_1 et C_2 deux convexes non vides de l'espace vectoriel normé E. On suppose C_1 ouvert et disjoint de C_2. Il existe alors un hyperplan affine fermé H séparant strictement C_1 de C_2. Plus précisément il existe une forme linéaire continue f sur E et un réel a tels que $\forall x_1 \in C_1, \forall x_2 \in C_2,\ f(x_1) < a \leq f(x_2)$.*

Corollaire 6.2.3. *Par tout point non intérieur d'un convexe C d'intérieur non vide de E il passe un hyperplan d'appui.*

Démonstration : Comme le convexe C est d'intérieur non vide c'est une conséquence directe du Théorème 6.2.1 appliqué à $\overset{o}{C}$. □

Corollaire 6.2.4. *Soit E un espace vectoriel normé. Pour tout point x de E il existe une forme linéaire continue f de norme 1 qui norme x, c'est à dire telle que $f(x) = \|x\|$.*

Démonstration : Supposons par exemple $\|x\| = 1$. Alors on peut séparer le point x de la boule ouverte $B(0,1)$. Il existe donc une forme linéaire continue g telle que $\forall y \in B(0,1),\ g(y) \leq g(x)$. En particulier puisque $0 \in B(0,1)$, on obtient $0 < g(x)$. Alors $f = g/g(x)$ vérifie $f(x) = 1$ et $\sup_{y \in B(0,1)} f(y) \leq f(x) = 1$; cette dernière condition signifie $\|f\| \leq 1$. L'inégalité inverse résulte de $1 = f(x) \leq \|f\|\,\|x\| = \|f\|$. □

Corollaire 6.2.5. *Si E est un espace vectoriel normé séparable, il existe une suite (f_n) de formes linéaires continues de norme 1 qui est normante, c'est-à-dire que pour tout $x \in E$ on a $\|x\| = \sup_n f_n(x)$.*

Démonstration : Soit x_n une suite dense dans E, et pout tout n soit f_n une forme linéaire continue de norme 1 telle que $f_n(x_n) = \|x_n\|$ (Corollaire

6.2.4). Pour tout $x \in E$, et tout $\varepsilon > 0$, il existe n tel que $\|x - x_n\| \leq \varepsilon$; alors $|\|x\| - \|x_n\|| \leq \varepsilon$ et $|f_n(x) - f_n(x_n)| \leq \|f_n\| \, \|x - x_n\| \leq \varepsilon$. D'où $f_n(x) \geq \|x\| - 2\varepsilon$, et ceci montre que $\|x\| \leq \sup_k f_k(x)$. L'inégalité inverse est toujours vraie. □

Remarque 6.2.6. La suite (f_n) est alors séparante, c'est à dire que si $x, y \in E$ vérifient $f_n(x) = f_n(y)$ pour tout n, on a $x = y$.

Le théorème de séparation d'un point et d'un convexe fermé se déduit alors du théorème 6.2.1 :

Théorème 6.2.7 (Séparation, cas EVN #2)**.** *Soit C un convexe fermé de l'espace vectoriel normé E. Pour tout point $x_0 \in E$ n'appartenant pas à C il existe un hyperplan affine fermé H séparant strictement x de C ; il revient au même de dire qu'il existe une forme linéaire continue f telle que* $\sup_{x \in C} f(x) < f(x_0)$.

Démonstration : La distance $r = d(x_0, C)$ est non nulle. Alors $C_r = \{y \in E \ / \ d(y, C) < r\}$ est un convexe ouvert ne contenant pas x. Soit f une forme linéaire continue telle que $\forall y \in C_r, f(y) < f(x_0)$. On a donc pour tout $x \in C$ et $h \in B(0, r)$, $f(x) + f(h) = f(x + h) < f(x_0)$; d'où $f(x_0) \geq \sup_{h \in B(0,r)}(f(x) + f(h)) = f(x) + r \, \|f\|$ et finalement $f(x_0) \geq \sup_{x \in C} f(x) + r \, \|f\|$. □

Corollaire 6.2.8. *Soit E espace vectoriel normé. Tout convexe fermé de E, différent de E, est l'intersection des demi-espaces fermés qui le contiennent.*

Démonstration : Comme le convexe C considéré est différent de E, le Théorème 6.2.7 appliqué à un point $x_0 \notin C$ montre qu'il existe effectivement des demi-espaces fermés contenant E. Soit D l'intersection des demi-espaces fermés contenant C. Il est clair que D contient C ; le Théorème 6.2.7 montre que D ne contient aucun point n'appartenant pas à C. □

Proposition 6.2.9. *Si C est un convexe fermé d'intérieur non vide, alors, C est l'intersection des demi-espaces le contenant déterminés par ses hyperplans d'appui.*

Démonstration : Il suffit de montrer que tout point x n'appartenant pas à C peut être séparé strictement de C par un hyperplan d'appui de C. Or soit x_0 un point intérieur à C. L'intersection du segment fermé $[x_0, x]$ avec C est un sous-segment fermé $[x_0, y]$, et y n'est pas intérieur à C, puisqu'on peut l'approcher par les points de $]y, x]$ qui ne sont pas dans C. Soit H un hyperplan d'appui de C en y. Comme y appartient à H et x_0 appartient au demi-espace ouvert défini par H contenant $\overset{o}{C}$, x appartient à l'autre demi-espace ouvert. □

Comme on l'a déja remarqué au paragraphe précédent, dans les EVN quelconques, on a besoin d'une hypothèse supplémentaire, la compacité du convexe, pour obtenir des résultats analogues à ceux des espaces de Hilbert (ou de dimension finie) pour les convexes fermés bornés.

Proposition 6.2.10. *Soit K est un convexe compact d'un espace vectoriel normé E. Pour tout hyperplan vectoriel fermé V de E il existe un hyperplan d'appui fermé de K parallèle à V .*

DÉMONSTRATION : La démonstration est analogue à celle du cas de la dimension finie, Proposition 3.4.3. □

Remarque 6.2.11. En fait, on voit que soit K est inclus dans un hyperplan parallèle à V, soit il est inclus dans une bande définie par deux hyperplans d'appui parallèles à V.

On va maintenant pouvoir démontrer le théorème de Krein-Milman, avec la proposition suivante :

Proposition 6.2.12. *Dans un espace vectoriel normé, tout convexe compact possède au moins un point extrémal.*

DÉMONSTRATION : Soit K un convexe compact de l'espace vectoriel E. Par définition de la compacité, pour tout N il existe un ensemble fini A_N de points de K tels que $K \subset \bigcup_{a \in A_N} B(a, 1/N)$. Il en résulte facilement que K contient une suite (x_n) dense dans lui-même. Alors K est inclus dans un sous-espace vectoriel normé F de E séparable (voir Définition 2.1.43), par exemple $\overline{\mathrm{Vect}\{x_n \;/\; n \geq 1\}}$. On peut donc supposer E lui même séparable.
D'après le Corollaire 6.2.5, on peut trouver une suite (f_n) de formes linéaires continues de normes 1, qui est normante, donc séparante. Posons pour tout $x \in E$:

$$F(x) = \sum_n 2^{-n} f_n(x)^2$$

Alors la fonction F est convexe, comme somme de fonctions convexes car le carré d'une forme linéaire est convexe. Cette fonction est en fait strictement convexe, car si $x \neq y$, il existe n_0 tel que $f_{n_0}(x) \neq f_{n_0}(y)$, d'où $f_{n_0}((x+y)/2)^2 < (f_{n_0}(x)^2 + f_{n_0}(y)^2)/2$ (puisque la fonction $t \mapsto t^2$ est strictement convexe). Pour $n \neq n_0$ on a l'inégalité large $f_n((x+y)/2)^2 \leq (f_n(x)^2 + f_n(y)^2)/2$, d'où l'inégalité stricte $F((x+y)/2) < (F(x) + F(y))/2$.
Par ailleurs la fonction F est continue sur E à cause de la convergence uniforme de la série sur toute boule $B(0, R)$; cette fonction atteint donc son maximum sur le compact K. D'après la Proposition 1.5.5, ce maximum est atteint en un point extrémal de K. □

Théorème 6.2.13 (Krein-Milman, cas EVN)**.** *Dans un espace vectoriel normé, tout convexe compact est l'enveloppe convexe fermée de ses points extrémaux.*

DÉMONSTRATION : C'est la même que dans les expaces de Hilbert, 6.1.18, en utilisant la Proposition 6.2.12. □

Corollaire 6.2.14. *Soit K un convexe compact d'un espace vectoriel normé, et $f : K \to \mathbb{R}$ une fonction convexe semi-continue inférieurement. On a*

$$\sup_{x \in K} f(x) = \sup_{x \in \overline{\operatorname{Ext} K}} f(x)$$

En particulier si la fonction f est continue, elle atteint son maximum en un point de $\overline{\operatorname{Ext} K}$.

Remarque 6.2.15. Ce résultat est souvent utilisé sous la forme suivante : pour majorer la fonction f sur K, il suffit de la majorer sur $\operatorname{Ext} K$.

DÉMONSTRATION : C'est la même que pour les espaces de Hilbert, (6.1.19). □

6.3 Une application : le théorème de Lyapounov

Définition 6.3.1 (Sous-espace facial d'un convexe)**.** *Soit C un convexe d'un espace vectoriel E, et x un point de C. Le sous-espace facial de x (relatif à C) est*

$$W_C(x) = \{u \in E / \ \exists \varepsilon > 0 \text{ tel que } x \pm \varepsilon u \in C\}$$

Autrement dit, $u \in W_C(x)$ si et seulement si $u = 0$ ou il existe un segment de direction u, de milieu x et inclus dans C. C'est un sous-espace vectoriel de E : $W_C(x)$ est clairement convexe et stable par les multiplications par les réels.

Remarque 6.3.2. *1)* Le point x est extrémal si et seulement si $W_C(x) = (0)$.
2) Soit $\Gamma_C(x)$ le cône engendré par $C - x$. Alors $W_C(x) = \Gamma_C(x) \cap (-\Gamma_C(x))$.

Exercice 6.1. *L'intersection $(x + W_C(x)) \cap C$ est une face de C : c'est la plus petite face de C contenant x.*

Lemme 6.3.3. *Soit K un convexe compact d'un EVN E (resp. un convexe fermé borné d'un espace de Hilbert). Soit $T : E \to \mathbb{R}^n$ une application affine continue. Alors pour tout y dans le convexe $C = T(K)$ il existe un point $x \in K$, tel que $Tx = y$, et $\dim W_K(x) \leq \dim W_C(y)$.*

DÉMONSTRATION : Une application affine est une composée $\tau_{x_0} \circ T_0$, où T_0 est linéaire et τ_{x_0} la translation de vecteur x_0. Il est alors facile de voir que $T_0(W_K(x)) \subset W_C(Tx))$. D'après le théorème de Krein-Milman, le convexe compact (resp. le convexe fermé) $T^{-1}(y) \cap K$ admet un point extrémal x. Dans ce cas, la restriction de T_0 à $W_C(x)$ est injective (et la conclusion sur les dimensions en résulte). En effet si $u \in W_C(x)$ vérifie $T_0 u = 0$, et si $x \pm \varepsilon u \in K$, alors $T(x \pm \varepsilon u) = T(x) = y$, i. e. $x \pm \varepsilon u \in T^{-1}(y) \cap K$. Comme x est extrémal dans $T^{-1}(y) \cap K$, il en résulte $u = 0$. $\square$

Etant donné un espace mesuré $(\Omega, \mathcal{A}, \mu)$, on considère l'espace de Banach $L_\infty(\Omega, \mathcal{A}, \mu)$ des (classes de) fonctions μ-mesurables bornées sur Ω, muni de la norme $\|f\|_\infty = \text{Ess}\sup_{\omega\in\Omega} |f(\omega)|$.

Rappelons qu'une mesure positive μ sur un espace $(\Omega, \mathcal{A})$ est dite non-atomique (ou diffuse) si tout ensemble $A \in \mathcal{A}$ de mesure $\mu(A) > 0$ contient un ensemble $B \in \mathcal{A}$ dont la mesure vérifie $0 < \mu(B) < \mu(A)$.

Lemme 6.3.4. *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Les points extrémaux de la boule-unité fermée $\overline{B}$ de l'espace $L_\infty(\Omega, \mathcal{A}, \mu)$ sont les éléments de valeur absolue $\mathbb{I}_\Omega$. Si la mesure est non-atomique, alors pour tout point non extrémal f de $\overline{B}$ l'espace $W_C(x)$ est de dimension infinie.*

Démonstration : a) Si $f \in \overline{B}$ vérifie $|f| = \mathbb{I}_\Omega$, on a $f = \mathbb{I}_A - \mathbb{I}_{A^c}$ (A^c désigne le complémentaire de A) et si $f = \dfrac{g+h}{2}$ avec $g, h \in \overline{B}$, alors $\mathbb{I}_A.g \leq \mathbb{I}_A$, $\mathbb{I}_A.h \leq \mathbb{I}_A$ et $\mathbb{I}_A.[(g+h)/2] = \mathbb{I}_A$ entraînent les égalités $\mathbb{I}_A.g = \mathbb{I}_A.h = \mathbb{I}_A$. De même $\mathbb{I}_{A^c}.g = \mathbb{I}_{A^c}.h = -\mathbb{I}_{A^c}$, et par conséquent $g = h = f$. Donc f est extrémal.

Inversement si $|f| \neq \mathbb{I}_\Omega$, posons $u = \mathbb{I}_\Omega - |f|$. On a $0 \leq u \leq \mathbb{I}_\Omega$, $u \neq 0$, et $|f \pm u| \leq |f| + u = \mathbb{I}_\Omega$, c'est à dire $\|f \pm u\| \leq 1$. Donc f n'est pas extrémale.

b) Supposons maintenant μ non-atomique. Soit $f \in \overline{B}$ non extrémale, et comme ci-dessus $u = \mathbb{I}_\Omega - |f|$. Alors $f \pm v \in \overline{B}$ ssi $|v| \leq u$. Il revient au même de dire que $v = h.u$ avec $h \in \overline{B}$. Il en résulte que $v \in \overline{B}(f)$ si et seulement si $v = h.u$, avec $h \in L_\infty(\Omega, \mathcal{A}, \mu)$. Soit S le support de u, c'est à dire $\mathbb{I}_S.u = u$ et $\mathbb{I}_{S^c}.u = 0$. Notons $L_\infty(S)$ le sous-espace de $L_\infty(\Omega, \mathcal{A}, \mu)$ formé par les éléments à support inclus dans S. Alors l'application $h \to h.u$ est un isomorphisme linéaire de $L_\infty(S)$ sur $W_{\overline{B}}(f)$. Mais comme μ est non-atomique, l'espace $L_\infty(S)$ est de dimension infinie, pour tout $S \in \mathcal{A}$ de μ-mesure non nulle. En effet on peut construire une suite infinie d'éléments $(S_k)_k$ de $\mathcal{A}$ inclus dans S et deux à deux disjoints, et la suite $(\mathbb{I}_{S_k})_k$ est alors linéairement indépendante. $\square$

Théorème 6.3.5 (Lyapounov)**.** *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, la mesure μ étant non-atomique σ-finie. Soient $f_1, \ldots, f_n$ des éléments de $L^1(\Omega, \mathcal{A}, \mu)$ fixés.*

Alors l'ensemble :

$$R(f_1,\ldots,f_n) := \{(\textstyle\int_A f_i\,d\mu)_{i=1}^n / \ A \in \mathcal{A}\}$$

est un convexe compact de $\mathbb{R}^n$.

Remarque 6.3.6. Quand $n = 1$ et $f_1 = \mathbb{I}_A$ le théorème dit que pour tout $t \in [0, \mu(A)]$ il existe $B \subset A$ avec $\mu(B) = t$; c'est une propriété bien connue des mesures diffuses.

Démonstration : Quitte à rempacer la mesure μ par $\mu_1 = \sup_{1\le i\le n} |f_i| .\mu$, et f_i par $f_i/\sup_i |f_i|$, on peut supposer que $|f_i| \le 1$, pour $i = 1,\ldots,n$, et que la mesure μ est une probabilité.
On considère le convexe $K = \{g \in L^2(\Omega,\mathcal{A},\mu)/ \ 0 \le g \le \mathbb{I}\}$, qui est fermé borné dans $L^2(\Omega,\mathcal{A},\mu)$ et l'application linéaire continue $T : L^2(\Omega,\mathcal{A},\mu) \to \mathbb{R}^n$ définie par $Tg = (\int_\Omega g.f_i\,d\mu)_{i=1}^n$.
Alors $C = T(K)$, image de K par l'application linéaire T, est un convexe borné qui contient l'ensemble $R(f_1,\ldots,f_n)$ puisque chaque caractéristique χ_A d'un ensemble $A \in \mathcal{A}$ est dans K.
Montrons d'abord que C est fermé, donc compact dans $\mathbb{R}^n$. Si $(a_1,\ldots,a_n) \in \overline{C}$, il existe une suite (g_ℓ) de K telle que $\int_\Omega g_\ell.f_i \to a_i$ quand $\ell \to \infty$, pour tout $i = 1,\ldots,n$. Soit $\Gamma_k = \{g \in K \ / \ \left|\int_\Omega g.f_i - a_i\right| \le 1/k, \ i = 1,\ldots,n\}$. Les Γ_k sont des convexes fermés bornés de L^2 car les applications $g \to \left|\int_\Omega g.f_i - a_i\right|$ étant convexes continues de L^2 dans $\mathbb{R}$, non vides car Γ_k contient g_ℓ pour ℓ assez grand, et la suite (Γ_k) est décroissante. D'après le Théorème 6.1.4, les Γ_k ont un point commun g, qui vérifie $\int_\Omega g.f_i = a_i$, $i = 1,\ldots,n$. Donc $(a_1,\ldots,a_n) \in C$ et C est fermé.
Montrons maintenant que $C = R(f_1,\ldots,f_n)$. Pour tout $y \in C$, il existe d'après le Lemme 6.3.3 un point x de K tel que $\dim W_K(x) \le \dim W_C(y)$, donc $W_K(x)$ est de dimension finie. Comme $\overline{B}$ est image de K par la bijection affine $z \to 2z - \mathbb{I}_\Omega$, on en déduit que pour $z = 2x - 1$ l'espace facial $W_{\overline{B}}(z)$ est de dimension finie. D'après le Lemme 6.3.4, z est un point extrémal de $\overline{B}$, c'est à dire de la forme $\mathbb{I}_A - \mathbb{I}_{A^c}$. Alors $x = (\mathbb{I}_\Omega + z)/2 = \mathbb{I}_A$. □

Chapitre 7
Régularité des fonctions Convexes

7.1 Convexité et continuité

Dans ce qui suit, E est un espace vectoriel normé quelconque.

Proposition 7.1.1 (Continuité des fonctions convexes)**.** *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, et x_0 un point intérieur à* dom*f. Si f est bornée supérieurement au voisinage de x_0, elle est continue, et même lipschitzienne, au voisinage de x_0.*

Démonstration : a) Montrons d'abord la continuité de f en x_0.
Soit $r_0 > 0$ tel que $f(x) \leq M$ pour tout x dans la boule fermée $\overline{B}(x_0, r_0)$. Pour tout $x \in \overline{B}(x_0, r_0)$ définissons y tel que $y - x_0 = \dfrac{r_0}{\|x - x_0\|}(x - x_0)$ et posons $\lambda = \dfrac{\|x - x_0\|}{r_0}$, alors $y \in \overline{B}(x_0, r_0)$, $\lambda \in [0, 1]$ et on a par convexité :

$$f(x) - f(x_0) = f(\lambda y + (1-\lambda)x_0) - f(x_0) \leq \lambda f(y) + (1-\lambda)f(x_0) - f(x_0) = \lambda\big(f(y) - f(x_0)\big)$$

$$\leq \lambda\big(M - f(x_0)\big) = \frac{\|x - x_0\|}{r_0}\big(M - f(x_0)\big)$$

Par ailleurs, posons $z = x_0 - (x - x_0)$. Alors, $z \in \overline{B}(x_0, r_0)$ et $x_0 = \dfrac{x + z}{2}$. On a donc :

$$f(x_0) = f\left(\frac{x + z}{2}\right) \leq \frac{1}{2}(f(x) + f(z))$$

soit encore :

$$f(z) - f(x_0) \geq -\big(f(x) - f(x_0)\big)$$

Il en résulte, en appliquant l'inégalité précédente à z :

$$\frac{\|x - x_0\|}{r_0}\big(M - f(x_0)\big) = \frac{\|z - x_0\|}{r_0}\big(M - f(x_0)\big) \geq f(z) - f(x_0) \geq -\big(f(x) - f(x_0)\big)$$

et finalement pour $x \in \overline{B}(0, r_0)$, on a

$$|f(x) - f(x_0)| \leq \frac{\|x - x_0\|}{r_0}\big(M - f(x_0)\big)$$

Ce qui prouve bien la continuité de f en x_0.

b) Montrons que f est lipschitzienne sur toute boule $B(x_0, r)$, $r < r_0$.

Si $x \in B(x_0, r)$, on a $\overline{B}(x, r_0 - r) \subset \overline{B}(x_0, r_0)$, donc f est bornée supérieurement par M sur $\overline{B}(x, r_0 - r)$. Il en résulte :

$$\forall y \in \overline{B}(x, r_0 - r), \qquad |f(y) - f(x)| \leq \frac{\|y - x\|}{r_0 - r}(M - f(x))$$

Mais comme $x \in B(x_0, r_0)$, on a $f(x) \geq f(x_0) - (M - f(x_0)) = 2f(x_0) - M$, d'où

$$\forall y \in \overline{B}(x, r_0 - r), \qquad |f(y) - f(x)| \leq 2\frac{\|y - x\|}{r_0 - r}(M - f(x_0))$$

Si maintenant $x, y \in B(x_0, r)$, on divise le segment $[x, y]$ en segments consécutifs $[x_i, x_{i+1}]$, $i = 0, \ldots, N$ de longueur inférieure à $r_0 - r$. Comme $[x, y] \subset B(x_0, r)$, on voit que :

$$|f(x) - f(y)| \leq \sum_i |f(x_{i+1}) - f(x_i)| \leq A \sum_i \|x_{i+1} - x_i\| = A\,\|x - y\|$$

où $A = 2\,\dfrac{M - f(x_0)}{r_0 - r}$. □

Définition 7.1.2 (Fonction localement lipschitzienne)**.** *On dira que f est localement lipschitzienne sur* $\overset{o}{\mathrm{dom}} f$ *si elle est lipschitzienne au voisinage de tout point de* $\overset{o}{\mathrm{dom}} f$.

Corollaire 7.1.3. *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Il est équivalent de dire :*
i) Il existe $x_0 \in \mathrm{dom} f$ tel que f soit majorée au voisinage de x_0 ;
ii) Il existe $x_0 \in \mathrm{dom} f$ tel que f soit continue en x_0 ;
iii) $\overset{o}{\mathrm{dom}} f \neq \emptyset$ *et f y est continue ;*
iv) $\overset{o}{\mathrm{dom}} f \neq \emptyset$ *et f y est localement lipschitzienne.*

DÉMONSTRATION : *iv)* $\Rightarrow$ *iii)* $\Rightarrow$ *ii)* $\Rightarrow$ *i)* est clair.
Montrons que *i)* $\Rightarrow$ *iv)* : Supposons f bornée sur la boule $\overline{B}(x_0, r_0)$. Alors nécessairement, $x \in \overset{o}{\mathrm{dom}} f$ et $\overset{o}{\mathrm{dom}} f \neq \emptyset$. Soit $y_0 \in \overset{o}{\mathrm{dom}} f$. Il existe z_0 un point intérieur à dom$f$ tel que y_0 appartienne au segment semi-ouvert $[x_0, z_0[$. Soit h l'homothétie de centre z_0 et de rapport λ amenant x_0 sur y_0. On a $h(x) = z_0 + \lambda(x - z_0) = \lambda x + (1-\lambda)z_0$, avec $0 < \lambda \le 1$ et $h(\overline{B}(x_0, r_0)) = \overline{B}(y_0, \lambda r_0)$ (voir la preuve du Lemme 2.2.7). Mais par convexité de f :

$$\forall x \in \overline{B}(x_0, r_0), \quad f(h(x)) = f(\lambda x + (1-\lambda)z_0) \le \lambda f(x) + (1-\lambda) f(z_0) \le \max(M, f(z_0))$$

ce qui entraîne que f est bornée supérieurement sur $\overline{B}(y_0, \lambda r_0)$, et est donc lipschitzienne sur $\overline{B}(y_0, \lambda r)$, pour tout $r < r_0$, d'après la proposition précédente et sa preuve. □

Corollaire 7.1.4. *Si E est de dimension finie, toute fonction convexe $f : E \to \mathbb{R} \cup \{+\infty\}$ est continue et même localement lipschitzienne sur l'intérieur relatif de son domaine effectif.*

La conclusion signifie que la restriction de f à ri(dom$f$) est continue.

DÉMONSTRATION : Si E est de dimension finie, quitte à remplacer E par Aff(dom$f$), on peut supposer que ri(dom$f$) $= \overset{o}{\mathrm{dom}} f$. Alors par la Proposition 3.1.1, dom$f$ contient un repère affine $(x_0, x_1, \dots, x_p)$. L'enveloppe convexe $S = \mathrm{conv}(x_0, x_1, \dots, x_p)$ est incluse dans dom$f$, et f est bornée supérieurement sur S, puisque pour tout système $(\lambda_0, \dots, \lambda_p)$ de coefficients positifs de somme 1,

$$f\Big(\sum_{i=0}^{p} \lambda_i x_i\Big) \le \sum_{i=0}^{p} \lambda_i f(x_i) \le \max_{i=1}^{p} f(x_i)$$

Puisque $\overset{o}{S}$ est non vide, inclus dans $\overset{o}{\mathrm{dom}} f$, il suit du corollaire précédent que f est continue, et même localement lipschitzienne, sur $\overset{o}{\mathrm{dom}} f$. □

Exercice 7.1. *Dans la situation du Corollaire 7.1.4, f est lipschitzienne sur tout compact inclus dans* ri(C).

Corollaire 7.1.5. *La fonction convexe $f : E \to \mathbb{R} \cup \{+\infty\}$ est continue en un point de son domaine effectif si et seulement si son épigraphe est d'intérieur non vide.*

DÉMONSTRATION : L'hypothèse de continuité entraîne que epi$f$ est d'intérieur non vide : en effet si $B(y_0, r)$ est une boule ouverte de centre y_0 sur laquelle f est majorée par M, l'épigraphe de f contient l'ouvert $B(y_0, r) \times]M, +\infty[$. La réciproque est une conséquence de 7.1.3. □

7.2 Convexité et différentiabilité

Commençons par examiner une notion très faible de différentiabilité :

Définition 7.2.1 (Dérivées directionnelles)**.** *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction numérique, x_0 un point de* dom$f$, *et h un vecteur non nul de E. On dit que f est dérivable à droite dans la direction h si le rapport $\dfrac{f(x_0+th)-f(x_0)}{t}$ a une limite, éventuellement infinie, lorsque t tend vers 0 par valeurs positives. Cette limite $f'_d(x_0,h)$ est appelée dérivée directionnelle de f dans la direction h.*

Proposition 7.2.2. *Une fonction convexe $f : E \to \mathbb{R} \cup \{+\infty\}$ est dérivable à droite dans toute direction en tout point x_0 où elle est finie (la dérivée directionnelle pouvant être infinie). De plus on a l'inégalité :*

$$\forall x \in E, \quad f(x) - f(x_0) \geq f'_d(x_0, x - x_0)$$

Démonstration : C'est une conséquence du lemme suivant :

Lemme 7.2.3. *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, $x_0 \in$* dom$f$ *et $h \in E$. L'application*

$$\Delta_{x_0,h} : \quad]0,+\infty[\to \mathbb{R}, \quad t \mapsto \frac{f(x_0+th)-f(x_0)}{t}$$

est croissante.

Démonstration : Posons $\varphi(t) = f(x_0+th) - f(x_0)$: c'est une fonction convexe de $[0,+\infty[$ dans $\mathbb{R} \cup \{+\infty\}$, avec $\varphi(0) = 0$ et par conséquent :

$$\forall \alpha, t \in [0,+\infty[, \varphi(\alpha t) = \varphi(\alpha t + (1-\alpha).0) \leq \alpha\varphi(t) + (1-\alpha)\varphi(0) = \alpha\varphi(t)$$

En posant $s = \alpha t$, on obtient $\varphi(s) \leq (s/t)\varphi(t)$ et donc si $s > 0$: $\dfrac{\varphi(s)}{s} \leq \dfrac{\varphi(t)}{t}$, c'est à dire $\Delta_{x_0,h}(s) \leq \Delta_{x_0,h}(t)$. Donc $\Delta_{x_0,h}$ est une fonction croissante de $]0,+\infty[$ dans $\mathbb{R} \cup \{+\infty\}$. □

Revenons à la preuve de la Proposition 7.2.2 : la fonction croissante $\Delta_{x_0,h}$ admet une limite dans $\overline{\mathbb{R}}$ quand $t \to 0$. Cette limite $f'_d(x_0,h)$ est nécessairement majorée par $\Delta_{x_0,h}(1)$. □

Corollaire 7.2.4 (Minimum d'une fonction convexe : Règle de Fermat #1)**.** *Une fonction convexe propre f admet un minimum au point x_0 si et seulement si toutes ses dérivées directionnelles à droite en x_0 sont positives ou nulles.*

Remarque 7.2.5. Posons $f'_d(x_0, 0) = 0$. Alors la fonction $f'_d(x_0, .) : E \to \overline{\mathbb{R}}$, $h \mapsto f'_d(x_0, h)$ est sous-linéaire, c'est à dire convexe et positivement homogène. En particulier $f'_d(x_0, h) \geq -f'_d(x_0, -h)$.

DÉMONSTRATION : La convexité résulte du fait que $f'_d(x_0, .)$ est limite simple des fonctions convexes $h \mapsto \Delta_{x_0,h}(t)$, lorsque $t \to 0^+$. L'homogénéité résulte de l'égalité $\Delta_{x_0,\lambda h}(t) = \lambda\Delta_{x_0,h}(\lambda t)$ pour tout $\lambda > 0$. Si les dérivées directionnelles selon $\pm h$ sont finies, (ou infinies de même signe) on a $f'_d(x_0, h) + f'_d(x_0, -h) \geq 2f'_d(x_0, 0) = 0$ par convexité, d'où l'inégalité $f'_d(x_0, h) \geq -f'_d(x_0, -h)$; celle-ci reste trivialement vraie dans les cas où ces dérivées directionnelles sont infinies de signe contraire. □

On remarquera que la fonction $f'_d(x_0, .)$ est un exemple naturel de fonction convexe pouvant prendre la valeur $-\infty$.

Une qualité de différentiabilité supérieure est atteinte lorsque la fonction $f'_d(x_0, .)$ est linéaire. En dimension infinie il est utile de requérir en sus la continuité :

Définition 7.2.6 (Gâteau différentiabilité)**.** *Soit $f : E \to \overline{\mathbb{R}}$ une fonction, et x_0 un point où f est finie. On dit que f est Gâteaux-différentiable en x_0 si pour tout $h \in E$, la fonction $\mathbb{R} \to \overline{\mathbb{R}}$, $s \to f(x_0 + sh)$ est dérivable au point* 0 *(on notera $f'(x_0, h)$ cette dérivée), et s'il existe une forme linéaire continue u^* telle que $f'(x_0, h) = \langle u^*, h\rangle$, pour tout $h \in E$. La forme linéaire continue u^* est alors unique, est appelée différentielle de Gâteaux (ou encore gradient) de f en x_0 et notée $f'(x_0)$ (ou parfois $\nabla f(x_0)$).*

Si la fonction f est Gâteaux-différentiable en x_0, on a en particulier $f'_d(x_0, h) = -f'_d(x_0, -h)$ pour tout $h \in E$. Le critère d'optimisation 7.2.4 devient alors

Corollaire 7.2.7 (Minimum d'une fonction convexe : Règle de Fermat #2)**.** *Une fonction f, convexe propre et Gâteaux-différentiable en tout point de son domaine effectif admet un minimum au point $x_0 \in \mathrm{dom} f$ si et seulement si $f'(x_0) = 0$.*

Il est relativement aisé pour une fonction convexe d'être Gâteaux-différentiable en un point donné. En effet :

Proposition 7.2.8. *Si f est une fonction convexe propre continue en $x_0 \in \mathrm{dom} f$ et si $f'_d(x_0, .)$ est impaire c'est à dire si $f'_d(x_0, h) = -f'_d(x_0, -h)$, pour tout $h \in E$, alors f est Gâteaux-différentiable en x_0.*

DÉMONSTRATION : Comme $f'_d(x_0, h) < +\infty$ pour tout $h \in E$, à cause de l'inégalité de la Proposition 7.2.2 et du fait que f est bornée au voisinage de x_0, l'égalité de l'énoncé montre qu'on a aussi $f'_d(x_0, h) > -\infty$. Les valeurs

$f'_d(x_0, h)$ et $-f'_d(x_0, -h)$ sont les dérivées à droite et à gauche de la fonction $t \to f(x_0 + th)$ au point 0. Leur égalité entraîne que cette fonction est dérivable en 0. La fonction $f'_d(x_0, .)$, $E \to \mathbb{R}$, sous-linéaire impaire est alors nécessairement linéaire.

Soit $r_0 > 0$ tel que $f(x_0 + h) - f(x_0) \leq 1$ pour $h \in B(0, r_0)$. Toujours à cause de l'inégalité de la Proposition 7.2.2, on a $f'_d(x_0, h) \leq 1$ pour $\|h\| \leq r_0$. Donc la forme linéaire $f'_d(x_0, .)$ est bornée, de norme $\leq 1/r_0$, c'est à dire continue. $\square$

On veut maintenant caractériser la convexité d'une fonction Gâteaux-différentiable $f : E \to \mathbb{R} \cup \{+\infty\}$. Le premier critère est maintenant trivial :

Proposition 7.2.9. *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$, Gâteaux-différentiable sur* dom$f$. *La fonction f est convexe si et seulement si pour tous $x \in$* dom$f$ *et $y \in E$ on a $f(y) - f(x) \geq \langle f'(x), y - x \rangle$.*

Démonstration : En effet si f est convexe, cette inégalité résulte de la Proposition 7.2.2.

Réciproquement supposons cette inégalité toujours vraie. Soient $x, y \in \mathrm{dom} f$, $\lambda \in [0, 1]$ et $z = \lambda x + (1 - \lambda)y$. On a

$$\begin{aligned} f(x) - f(z) &\geq \langle f'(z), x - z \rangle \\ f(y) - f(z) &\geq \langle f'(z), y - z \rangle \end{aligned}$$

d'où

$$\lambda f(x) + (1 - \lambda)f(y) - f(z) \geq \langle f'(z), \lambda(x - z) + (1 - \lambda)(y - z) \rangle = 0$$

donc f est convexe. $\square$

Pour obtenir un deuxième critère de convexité, on a besoin d'une définition :

Définition 7.2.10 (Application monotone)**.** *Une application T de $D \subset E$ dans E' est monotone si*

$$\forall x, y \in D \qquad \langle T(x) - T(y), x - y \rangle \geq 0$$

Proposition 7.2.11. *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$, Gâteaux-différentiable sur* dom$f$. *La fonction f est convexe si et seulement si f' est une application monotone de* dom$f$ *dans E'.*

Démonstration : a) Supposons f convexe.
Soient $x, y \in \mathrm{dom} f$. D'après la Proposition 7.2.2, on a :

$$\begin{aligned} f(y) - f(x) &\geq \langle f'(x), y - x \rangle \\ f(x) - f(y) &\geq \langle f'(y), x - y \rangle \end{aligned}$$

En sommant membre à membre ces deux inégalités, on trouve l'inégalité cherchée.
b) Réciproquement, supposons f' monotone.
Soient $x, y \in \mathrm{dom} f$; on considère l'application $\varphi : [0,1] \to \mathbb{R}$, $\lambda \mapsto \varphi(\lambda) = f(\lambda x + (1-\lambda)y) = f(y + \lambda(x-y))$. Elle est dérivable, de dérivée :

$$\varphi'(\lambda) = \langle f'(y + \lambda(x-y)), x-y \rangle$$

Si $\lambda_1 < \lambda_2$, on a :

$$\varphi'(\lambda_1) - \varphi'(\lambda_2) = \langle f'(y + \lambda_1(x-y)) - f'(y + \lambda_2(x-y)), x-y \rangle \leq 0$$

car la différence des arguments de f' vaut $(\lambda_1 - \lambda_2)(x-y)$. On en déduit que la fonction φ est à dérivée croissante, donc est convexe. En particulier $\varphi(\lambda) \leq \lambda\varphi(1) + (1-\lambda)\varphi(0) = \lambda f(x) + (1-\lambda)f(y)$, ce qui est l'inégalité de convexité pour f. □

Remarque 7.2.12. Supposons que l'application $\mathrm{dom} f \to E'$, $x \to f'(x)$ vérifie :

$$\forall h \in E, \langle \frac{f'(x+th) - f'(x)}{t}, h \rangle \to Q(x,h)$$

lorsque $t \to 0$, avec $Q(x,h) \geq 0$ pour tout h. Alors f est convexe. En particulier si $E = \mathbb{R}^n$ est de dimension finie, et $f : U \to \mathbb{R}$ est de classe C^2 sur l'ouvert convexe U, et si la matrice hessienne $H_x(f) = (\dfrac{\partial^2 f}{\partial x_i \partial x_j})$ des dérivées partielles secondes est positive c'est à dire ${}^t h H_x(f) h \geq 0$ pour tout $h \in \mathbb{R}^n$, alors f est convexe sur U.

Démonstration : En effet d'après le théorème des accroissements finis on a pour un certain $\theta \in]0,1[$:

$$\langle f'(x) - f'(y), x-y \rangle = \frac{d}{dt}\langle f'(y + t(x-y)), x-y \rangle_{t=\theta} = Q(\theta x + (1-\theta)y, x-y) \geq 0$$

Dans le cas C^2, on voit facilement que $Q(x;h) = {}^t h H_x(f) h$. □

Finalement, rappelons la notion la plus forte de différentiabilité, qui est appelée dans le cadre des EVN, la Fréchet-différentiabilité :

Définition 7.2.13 (Fréchet différentiabilité)**.** *Soit $f : E \to \overline{\mathbb{R}}$ une fonction, et x_0 un point où f est finie. On dit que f est Fréchet-différentiable en x_0 s'il existe une forme linéaire continue u^* telle que*

$$|f(x_0 + h) - f(x_0) - \langle u^*, h \rangle| = o(\|h\|)$$

Bien entendu, si f est Fréchet-différentiable, elle est Gâteaux-différentiable et $u^* = f'(x_0)$.

Par contre, même en dimension finie, en général, la Gâteaux-différentiabilité n'entraîne pas la Fréchet-différentiabilité. Néanmoins c'est vrai pour les fonctions convexes :

Proposition 7.2.14. *On suppose* E de dimension finie. *Si la fonction convexe* $f : E \to \mathbb{R}\cup\{+\infty\}$ *est Gâteaux-différentiable en* $x_0 \in \overset{o}{\mathrm{dom}} f$, *elle y est Fréchet-différentiable.*

Démonstration : Soit r_0 tel que $B(x_0, r_0) \subset \mathrm{dom} f$ et $r < r_0$. Posons pour $h \in S_E = \{h \in E/ \ \|h\| = 1\}$ et $0 \le t \le r$:

$$\Delta(h,t) = \begin{cases} \frac{f(x_0+th)-f(x_0)}{t} & \text{si } t > 0, \\ \langle f'(x_0), h\rangle & \text{si } t = 0. \end{cases}$$

La fonction $\Delta : \ S_E \times [0, r] \to \mathbb{R}$ est continue : c'est clair (d'après le Corollaire 7.1.4) aux points (h_0, t_0) avec $t_0 \neq 0$. Vérifions-le en un point $(h_0, 0)$; on a :

$$\begin{aligned} |\Delta(h,t) - \Delta(h_0,0)| &\le |\Delta(h,t) - \Delta(h_0,t)| + |\Delta(h_0,t) - \Delta(h_0,0)| \\ &= \frac{|f(x_0+th) - f(x_0+th_0)|}{t} \\ &+ \left|\frac{f(x_0+th_0) - f(x_0)}{t} - \langle f'(x_0), h\rangle\right| \end{aligned}$$

Le second terme tend vers 0 quand $t \to 0$ (Gâteaux-différentiabilité de f en x_0). Le premier terme se majore en utilisant la propriété de Lipschitz locale de f (Corollaire 7.1.3) :

$$\frac{|f(x_0+th) - f(x_0+th_0|}{t} \le M\,\|h - h_0\| \to 0, \text{ quand } h \to h_0$$

Comme $S_E \times [0, r]$ est un compact de $E \times \mathbb{R}$ (S_E est compacte car E est de dimension finie), la fonction continue Δ est uniformément continue, et par conséquent

$$\sup_{h\in S_E} |\Delta(h,t) - \Delta(h,0)| \to 0, \text{ quand } t \to 0$$

ce qui revient à dire que

$$\big|f(x_0+h) - f(x_0) - \langle f'(x_0), h\rangle\big| = o(\|h\|). \quad \square$$

7.3 Convexité et sous-différentiabilité

Définition 7.3.1 (Sous-gradient). *Soit* $f : E \to \overline{\mathbb{R}}$ *une fonction quelconque et* x_0 *un point donnant à* f *une valeur finie. On dit que* f *est* sous-différentiable *en $x_0 \in E$ s'il existe $u^* \in E'$ telle que :*

$$\forall x \in E, \quad f(x) - f(x_0) \geq \langle u^*, x - x_0 \rangle$$

On dit alors que u^* est un sous-gradient de f en x_0 ; on appelle sous-différentiel de f en x_0, et on note $\partial f(x_0)$, l'ensemble des sous-gradients de f en x_0.

Remarque 7.3.2. *1)* Par définition, f est sous-différentiable en $x_0 \in E$ si et seulement si $\partial f(x_0) \neq \emptyset$. *2)* Il est clair que si f est sous-différentiable en un point, elle ne prend nulle part la valeur $-\infty$.

Remarque 7.3.3 (Règle de Fermat #3). Soit $f : E \to \overline{\mathbb{R}}$ une fonction quelconque et $x_0 \in E$ tel que $f(x_0) \in \mathbb{R}$. Il suit directement de la définition que f admet un minimum au point x_0 si et seulement si $0 \in \partial f(x_0)$.

Exercice 7.2. *$\partial f(x_0)$ est un sous-ensemble convexe fermé de E' (éventuellement vide).*

La notion de sous-gradient s'interprète géométriquement dans $E \times \mathbb{R}$ (qu'on peut munir d'une structure d'EVN en posant par exemple $\|(x,t)\| = \|x\| + |t|$) :

Remarque 7.3.4. La forme linéaire continue u^* est un sous-gradient de f en x_0 si et seulement si l'hyperplan affine de $E \times \mathbb{R}$ d'équation $t = f(x_0) + \langle u^*, x - x_0 \rangle$ est un hyperplan d'appui de epi$f$.

La notion d'une fonction sci a été introduite en Définition 2.3.1. On peut énoncer :

Proposition 7.3.5. *Si la fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ a un domaine effectif convexe et y est sous-différentiable, elle est convexe, et coïncide sur son domaine effectif avec une fonction convexe sci.*

Démonstration : Pour tout $x \in \mathrm{dom} f$, choisissons $u_x^* \in \partial f(x)$. 0n pose :

$$g(y) = \sup_{x \in \mathrm{dom} f} (f(x) + \langle u_x^*, y - x \rangle)$$

La fonction g est convexe sci comme enveloppe supérieure de fonctions affines continues et $g(y) \leq f(y)$ pour tout $y \in E$.
De plus si $y \in \mathrm{dom} f$ on a $g(y) \geq f(y) + \langle u_x^*, y - y \rangle = f(y)$, d'où $g(y) = f(y)$et f coïncide bien avec une fonctions convexe continue sur son domaine effectif. Comme dom$f$ est convexe, ceci implique que f est convexe par la Remarque 1.5.3. □

Inversement les fonctions convexes sont sous-différentiables en tout point de continuité dans leur domaine effectif :

Théorème 7.3.6. *Si la fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ est convexe et continue en un point de son domaine effectif, elle est sous-différentiable en tout point de l'intérieur de son domaine effectif. De plus pour tout x_0 point intérieur de* $\mathrm{dom} f$ *et tout $h \in E$, il existe $u^* \in \partial f(x_0)$ tel que $f'_d(x_0, h) = \langle u^*, h \rangle$.*

Démonstration : Si f est continue en un point, par le Corollaire 7.1.5, $\mathrm{epi} f$ est d'intérieur non vide.

Soit x_0 un point intérieur à $\mathrm{dom} f$ et $h \in E$, $h \neq 0$. Soit $D_{x_0,h}$ la droite affine de $E \times \mathbb{R}$ passant par le point $(x_0, f(x_0))$ et de vecteur directeur $(h, f'_d(x_0, h)$. On a donc :

$$D_{x_0,h} = \big\{(x_0 + \tau h, f(x_0) + \tau f'_d(x_0, h) \ / \ \tau \in \mathbb{R}\big\}$$

Alors $\mathrm{epist} f \cap D_{x_0,h} = \emptyset$. En effet pour $\tau \geq 0$, on a (d'après la Proposition 7.2.2) :

$$f(x_0 + \tau h) \geq f(x_0) + f'_d(x_0, \tau h)) = f(x_0) + \tau f'_d(x_0, h)$$

tandis que pour $\tau \leq 0$, on a :

$$f(x_0 + \tau h) \geq f(x_0) + (-\tau) f'_d(x_0, -h) \geq f(x_0) + \tau f'_d(x_0, h)$$

(puisque $f'_d(x_0, -h) \geq -f'_d(x_0, h)$ et $-\tau \geq 0$).
Remarquons que $\mathrm{epist} f$ contient l'intérieur de $\mathrm{epi} f$ (les points de $\mathrm{epi} f \setminus \mathrm{epist} f$, c'est à dire du graphe de $f|_{\mathrm{dom} f}$ ne sont évidemment pas intérieurs à $\mathrm{epi} f$). Donc $D_{x_0,h}$ est disjointe de l'intérieur de $\mathrm{epi} f$. D'après le théorème de Hahn-Banach (Corollaire 6.2.2), il existe un hyperplan affine fermé de $E \times \mathbb{R}$ contenant $D_{x_0,h}$ et disjoint de l'intérieur de $\mathrm{epi} f$. L'un des demi-espaces fermés associés contient l'intérieur de $\mathrm{epi} f$, donc son adhérence $\overline{\mathrm{epi}} f$. Autrement dit il existe $\varphi \in (E \times \mathbb{R})'$ non nulle et $a \in \mathbb{R}$ tels que :

$$\inf_{(x,t) \in \mathrm{epi} f} \varphi(x, t) \geq a = \varphi(y, s)$$

pour tout $(y, s) \in D_{x_0,h}$. Mais $(E \times \mathbb{R})'$ s'identifie à $E' \times \mathbb{R}$, c'est à dire que : $\varphi(x, t) = \langle v^*, x \rangle + \alpha t$, pour un certain couple $(u^*, \alpha) \in E' \times \mathbb{R}$. On a alors en particulier, puisque $(x_0, f(x_0)) \in D_{x_0,h}$:

$$\forall (x, t) \in \mathrm{epi} f, \qquad \langle v^*, x \rangle + \alpha t \geq \langle v^*, x_0 \rangle + \alpha f(x_0) \qquad (*)$$

En faisant $t \to +\infty$, on voit que nécessairement $\alpha \geq 0$.

Si on avait $\alpha = 0$, alors on aurait

$$\forall x \in \mathrm{dom} f, \langle v^*, x - x_0 \rangle \geq 0$$

Comme $\mathrm{dom}f$ contient une boule $B(x_0, r)$ (puisque x_0 lui est intérieur), ceci entraînerait $\forall y \in B(0, r)$, $\langle v^*, y\rangle \geq 0$; ceci resterait vrai pour tout $y \in E$ par homogénéité, et en rempaçant y par $-y$, on voit qu'alors $\forall y \in E$ $\langle v^*, y\rangle = 0$, c'est à dire $v^* = 0$; mais alors $\varphi = 0$, ce qui est une contradiction.

On a donc $\alpha > 0$. En divisant par α, l'inégalité $(*)$ devient :

$$\forall (x,t) \in \mathrm{epi} f, \qquad \langle w^*, x\rangle + t \geq \langle w^*, x_0\rangle + f(x_0)$$

(où $w^* = (1/\alpha)v^*$). En appliquant ceci au cas $t = f(x)$, on obtient $f(x) - f(x_0) \geq -\langle w^*, x - x_0\rangle$, pour tout $x \in C$, c'est à dire $u^* = -w^*$ est un sous-gradient de f en x_0.

Par ailleurs utilisons le fait que φ est constante sur $D_{x_0,h}$:

$$\forall \tau \in \mathbb{R}, -\langle u^*, x_0 + \tau h\rangle + f(x_0) + \tau f'_d(x_0, h) = C^{ste}$$

ce qui entraîne $f'_d(x_0, h) = \langle u^*, h\rangle$. □

Corollaire 7.3.7. *Soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, et x_0 un point de son domaine effectif.*
i) Si la fonction f est Gâteaux-différentiable en x_0, elle est sous-différentiable en x_0 et $\partial f(x_0) = \{f'(x_0)\}$.
ii) Si la fonction f est continue en x_0, et si son sous-différentiel $\partial f(x_0)$ ne contient qu'un seul élément, elle est Gâteaux-différentiable en x_0.

Démonstration : *i)* Il résulte de la Proposition 7.2.2 et de la définition 7.2.6 que $f'(x_0)$ est un sous-gradient de f en x_0.
Inversement, soit $u^* \in \partial f(x_0)$. On a donc, pour tout $h \in E$:

$$\forall t \in]0,1], \quad f(x_0 + th) - f(x_0) \geq \langle u^*, th\rangle$$

On divise par t et on fait $t \to 0$, on obtient à la limite :

$$\langle f'(x_0), h\rangle \geq \langle u^*, h\rangle$$

En remplaçant h par $-h$, on voit que $\langle f'(x_0), h\rangle \leq \langle u^*, h\rangle$, et par conséquent $u^* = f'(x_0)$.
ii) Si $\partial f(x_0) = \{u^*\}$, il résulte du Théorème 7.3.6 que pour tout $h \in E$ on a :

$$f'_d(x_0, h) = \langle u^*, h\rangle$$

Il en résulte que f admet $f'(x_0)$ pour différentielle de Gâteaux en x_0. □

Pour étudier le sous-différentiel d'une fonction convexe, on a besoin d'un théorème de séparation fonctionnel :

Théorème 7.3.8 (Séparation fonctionnel)**.** *Soient $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, et $g : E \to \mathbb{R} \cup \{-\infty\}$ une fonction concave telles que $\forall x \in E, \quad f(x) \geq g(x)$. On suppose qu'il existe un point x_0 où f et g prennent des valeurs finies, et où l'une d'elle soit continue. Alors il existe une fonction affine continue h telle que $f(x) \geq h(x) \geq g(x)$.*

Démonstration : On note hypo$g$ l'hypographe de g, c'est à dire l'ensemble convexe

$$\text{hypo}g = \{(x,t) \in E \times \mathbb{R} \ / \ t \leq g(x)\}$$

Supposons que par exemple f est continue en x_0. Alors epi$f$ est d'intérieur non vide, inclus dans epist$f$. L'hypothèse $f \geq g$ entraîne que epist$f \cap$ hypo$g = \emptyset$. Comme ces deux ensembles sont convexes, on peut les séparer par un hyperplan affine fermé (Corollaire 6.2.2, théorème de Hahn-Banach géométrique) ; il existe donc une forme linéaire continue φ non nulle sur $E \times \mathbb{R}$ telle que

$$\inf_{(x,t) \in \text{epi}f} \varphi(x,t) \geq b \geq \sup_{(x,t) \in \text{hypo}g} \varphi(x,t)$$

il revient au même de dire qu'il existe $u^* \in E'$ et $a \in \mathbb{R}$, non tous deux nuls, tels que

$$\forall x \in \text{dom}f, \forall y \in \text{dom}(-g), \forall t \geq f(x), \forall s \leq g(y) : \qquad \langle u^*, x\rangle + at \geq b \geq \langle u^*, y\rangle + as$$

En faisant $t \to +\infty$, on voit que $a \geq 0$.
Si $a = 0$, en prenant $y = x_0$ on voit que :

$$\forall x \in \text{dom}f, \qquad \langle u^*, x - x_0\rangle \geq 0$$

et comme x_0 est intérieur à dom$f$ ceci entraîne $u^* = 0$, contradiction. Donc $a > 0$, et quitte à diviser φ par a, on peut supposer $a = 1$. On a donc :

$$\forall x \in \text{dom}f, \forall y \in \text{dom}(-g), \qquad \langle u^*, x\rangle + f(x) \geq b \geq \langle u^*, y\rangle + g(y)$$

c'est à dire :

$$\forall x \in E, f(x) \geq -\langle u^*, x\rangle + b \geq g(x) \qquad \square$$

Théorème 7.3.9. *Soient f, g deux fonctions convexes de E dans $\mathbb{R} \cup \{+\infty\}$. Soit x_0 un point de E.*
a) Si f et g sont sous-différentiables en x_0, il en est de même de $f + g$, et

$$\partial f(x_0) + \partial g(x_0) \subset \partial(f+g)(x_0)$$

b) Supposons que ces fonctions prennent des valeurs finies en x_0 et que l'une d'elle soit continue en x_0. Alors si $f+g$ est sous-différentiable en x_0, il en est de même de f et g, et

$$\partial(f+g)(x_0) = \partial f(x_0) + \partial g(x_0)$$

Démonstration : a) Si $u^* \in \partial f(x_0)$ et $v^* \in \partial g(x_0)$, on a

$$f(x) \geq f(x_0) + \langle u^*, x - x_0\rangle$$
$$g(x) \geq g(x_0) + \langle v^*, x - x_0\rangle$$

pour tout $x \in E^*$. En sommant membre à membre il vient

$$f(x) + g(x) \geq f(x_0) + g(x_0) + \langle u^* + v^*, x - x_0\rangle$$

donc $u^* + v^* \in \partial(f+g)(x_0)$.
b) Supposons que f est continue en x_0. Soit w^* un sous-gradient de $(f+g)$ en x_0. On a donc :

$$f(x) + g(x) \geq f(x_0) + g(x_0) + \langle w^*, x - x_0\rangle$$

pour tout $x \in E$; ce qu'on peut encore écrire :

$$f(x) - f(x_0) - \langle w^*, x - x_0\rangle \geq -(g(x) - g(x_0))$$

Soit $f_0(x) = f(x) - f(x_0) - \langle w^*, x - x_0\rangle$ et $k_0(x) = -(g(x) - g(x_0))$: alors f_0 est convexe, continue en x_0, k_0 est concave , et $f_0 \geq k_0$. Il existe donc (Théorème 7.3.8) une fonction affine continue h telle que $f_0 \geq h \geq k_0$. Comme $f_0(x_0) = 0 = g_0(x_0)$, on a $h(x_0) = 0$, c'est à dire que h est de la forme $h(x) = \langle -u^*, x - x_0\rangle$, avec $u^* \in E'$. Donc :

$$\forall x \in E, \quad \begin{cases} f(x) - f(x_0) \geq \langle w^* - u^*, x - x_0\rangle \\ g(x) - g(x_0) \geq \langle u^*, x - x_0\rangle \end{cases}$$

ce qui signifie que $u^* \in \partial g(x_0)$ et $w^* - u^* \in \partial f(x_0)$. □

Proposition 7.3.10. *Soient E, F deux espaces vectoriels normés ; $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, et $A : F \to E$ un opérateur linéaire continu. Soit x_0 un point de E.*
a) Si f est sous-différentiable au point Ax_0, alors $f \circ A$ est sous-différentiable en x_0, et

$$\partial(f \circ A)(x_0) \supset A^* \partial f(Ax_0)$$

(où $A^* : E' \to F'$ désigne l'opérateur transposé de A).
b) Si de plus f est continue en Ax_0, on a l'égalité

$$\partial(f \circ A)(x_0) = A^* \partial f(Ax_0).$$

Démonstration : a) Soit $u^* \in \partial f(Ax_0)$. On a :

$$\forall x \in F, (f \circ A)(x) - (f \circ A)(x_0) = f(Ax) - f(Ax_0) \geq \langle u^*, Ax - Ax_0 \rangle = \langle A^* u^*, x - x_0 \rangle\rangle$$

donc $A^* u^* \in \partial(f \circ A)(x_0)$.
b) Réciproquement, soit $v^* \in \partial(f \circ A)(x_0)$. On a donc :

$$\forall x \in F, f(Ax) - f(Ax_0) = (f \circ A)(x) - (f \circ A)(x_0) \geq \langle v^*, x - x_0 \rangle$$

Alors $\operatorname{Ker} A \subset \operatorname{Ker} v^*$: en effet, si $h \in \operatorname{Ker} A$, on a $0 = f(A(x_0 \pm h)) - f(Ax_0) \geq \pm \langle v^*, h \rangle$, donc $h \in \operatorname{Ker} v^*$. Autrement dit la valeur de v^* au point x ne dépend que de l'image $Ax = y$; on peut donc définir sans ambigüité $\ell : M = \operatorname{Im} A \to \mathbb{R}$ par $\ell(y) = \langle v^*, x \rangle$ pour tout $x \in A^{-1}(y)$. L'application ℓ est bien sûr linéaire. Posons

$$g(y) = \begin{cases} \ell(y) \text{ si } y \in \operatorname{Im} A \\ -\infty \text{ sinon} \end{cases}$$

Alors g est concave, $g(Ax_0) = \ell(Ax_0) \in \mathbb{R}$ et

$$\forall y \in E, f(y) - f(Ax_0) \geq g(y) - g(Ax_0)$$

D'après le Théorème 7.3.8, on en déduit qu'il existe une fonction affine continue h telle que

$$\forall y \in E, \qquad f(y) - f(Ax_0) \geq h(y) \geq g(y) - g(Ax_0)$$

Comme $h(Ax_0) = 0$, on a en fait $h(y) = \langle u^*, y - Ax_0 \rangle$, pour un certain $u^* \in E'$. On a donc $u^* \in \partial f(Ax_0)$. La restriction de u^* à $\operatorname{Im} A$ vérifie clairement :

$$\langle u^*, Au \rangle \geq \ell(Au) = \langle v^*, u \rangle$$

pour tout $u \in F$. En remplaçant u par $-u$, on obtient donc l'égalité : $\langle u^*, Au \rangle = \langle v^*, u \rangle$ pour tout $u \in F$. Donc $v^* = A^* u^*$, qui appartient à $A^* \partial f(Ax_0)$. $\square$

Chapitre 8
Fonctions convexes conjuguées

8.1 Enveloppes supérieures de fonctions affines continues

Théorème 8.1.1. *Une fonction $f : E \to \mathbb{R} \cup \{+\infty\}$ est convexe sci si et seulement si elle est enveloppe supérieure de fonctions affines continues.*

Démonstration : Une enveloppe supérieure de fonctions affines continues est convexe sci à valeurs dans $\mathbb{R} \cup \{+\infty\}$.

Réciproquement soit $f : E \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe sci. Si f est la fonction identique à $+\infty$, c'est trivialement une enveloppe supérieure d'une famille de fonctions affines continues, par exemple $f_n = n$, $n \geq 1$.

Sinon, $\mathrm{dom}f$ n'est pas vide. Considérons $x_0 \in \overline{\mathrm{dom}f}$. Pour tout t réel $t < f(x_0)$, l'ensemble $V_f(t) = \{x \in E \ / \ f(x) > t\}$ est un ouvert contenant x_0, donc contenant une boule ouverte $B(x_0, r)$, $r > 0$. Soit $g = t - \chi_{B(x_0,r)}$: alors g est une fonction concave, finie et continue sur $B(x_0, r)$ et majorée par f. Comme x_0 est adhérent à $\mathrm{dom}f$, on a $(\mathrm{dom}f) \cap B(x_0, r) \neq \emptyset$.
D'après le théorème de séparation fonctionnelle (Théorème 7.3.8) il existe une fonction affine continue h avec $g \leq h \leq f$. Donc h est une minorante affine continue de f telle que $h(x_0) \geq t$.

Ce qui précède montre d'une part que f admet au moins une minorante affine continue h_0, d'autre part qu'en tout point x adhérent à $\mathrm{dom}f$, il existe pour tout $t < f(x_0)$ une minorante affine continue h de f telle que $h(x_0) \geq t$.

Soit maintenant $x_0 \in E \setminus \overline{\mathrm{dom}f}$; on peut trouver d'après la Proposition 6.2.7 (séparation d'un point et d'un convexe fermé) une forme linéaire $u^* \in E'$ telle que

$$\alpha := \sup_{x \in \mathrm{dom}f} \langle u^*, x\rangle < \langle u^*, x_0\rangle$$

La fonction affine $h_{u^*,\alpha}$ définie par $h_{u^*,\alpha}(x) = \langle u^*, x\rangle - \alpha$ vérifie donc $h_{u^*,\alpha}(x) \leq 0$ pour tout $x \in \mathrm{dom}f$ et $h_{u^*,\alpha}(x_0) > 0$. Pour tout $t \in \mathbb{R}$ on peut donc trouver

$\rho \geq 0$ tel que $\rho h_{u^*,\alpha}(x_0) > t - h_0(x_0)$. Alors $h = \rho h_{u^*,\alpha} + h_0$ vérifie $h(x_0) > t$ et $h(x) \leq h_0(x) \leq f(x)$ pour tout $x \in \mathrm{dom} f$, donc $h \leq f$.

Désignons par $\mathcal{A}_f$ l'ensemble des minorantes affines continues de f ; il est clair que $f \geq \sup \mathcal{A}_f$, et ce qui précède montre que $f = \sup \mathcal{A}_f$. □

8.2 Régularisée convexe sci d'une fonction numérique

Définition 8.2.1 (La régularisée convexe sci d'une fonction)**.** *La régularisée convexe sci d'une fonction* f $E \to \overline{\mathbb{R}}$ *est la plus grande fonction convexe sci* $\hat{f} : E \to \overline{\mathbb{R}}$ *minorant* f.

Soit $\mathcal{S}_f$ l'ensemble des fonctions convexes sci minorant f (Noter que $\mathcal{S}_f$ contient en particulier la constante $-\infty$). Alors l'enveloppe supérieure de $\mathcal{S}_f$ est convexe sci, donc c'est $\hat{f}$.

Remarque 8.2.2. La fonction f et sa régularisée convexe sci ont les mêmes minorantes affines continues : $\mathcal{A}_f = \mathcal{A}_{\hat{f}}$. Il résulte donc du Théorème 8.1.1 que si $\mathcal{A}_f \neq \emptyset$, la régularisée convexe sci $\hat{f}$ est égale à l'enveloppe supérieure de $\mathcal{A}_f$.

Proposition 8.2.3. *On a* $\mathrm{epi}\hat{f} = \overline{\mathrm{conv}}(\mathrm{epi} f)$.

Démonstration : L'ensemble $\mathrm{epi}\hat{f}$ est convexe fermé et contient $\mathrm{epi} f$, puisque $\hat{f} \leq f$ et donc contient $\overline{\mathrm{conv}}(\mathrm{epi} f)$.

Inversement $\overline{\mathrm{conv}}(\mathrm{epi} f)$ est l'épigraphe d'une fonction convexe sci $g : E \to \overline{\mathbb{R}}$. En effet pour tout $s \geq 0$ soit T_s la tranlation de $E \times \mathbb{R}$ de vecteur $(0, s)$, on a $T_s(\mathrm{epi} f) \subset \mathrm{epi} f$, et il en résulte $T_s(\overline{\mathrm{conv}}(\mathrm{epi} f)) \subset \overline{\mathrm{co}}(\mathrm{epi} f)$. Autrement dit si $(x, t) \in \overline{\mathrm{conv}}(\mathrm{epi} f)$ alors pour tout $s \geq 0$, on a $(x, t + s) \in \overline{\mathrm{conv}}(\mathrm{epi} f)$. Posant $g(x) = \inf\{t/\ (x, t) \in \overline{\mathrm{conv}}(\mathrm{epi} f)\}$, avec la convention $\inf \emptyset = +\infty$, on a $(x, g(x)) \in \overline{\mathrm{conv}}(\mathrm{epi} f)$ lorsque $g(x) > -\infty$ car cet ensemble est fermé et finalement $\mathrm{epi} g = \overline{\mathrm{conv}}(\mathrm{epi} f)$.
La fonction g est alors convexe sci, et minore f, d'où $g \leq \hat{f}$ et donc $\overline{\mathrm{conv}}(\mathrm{epi} f) = \mathrm{epi} g \supset \mathrm{epi}\hat{f}$. □

8.3 Conjuguée de Fenchel-Moreau

Définition 8.3.1 (Fonction conjuguée)**.** *Soit* $f : E \to \overline{\mathbb{R}}$ *une fonction numérique. Sa fonction conjuguée (au sens de Fenchel-Moreau) est la fonction* $f^* : E' \to \overline{\mathbb{R}}$ *définie par*

$$f^*(x^*) = \sup_{x \in E} [\langle x^*, x \rangle - f(x)]$$

Remarque 8.3.2. La fonction conjuguée f^* est parfois aussi appelée polaire de la fonction f. Dans le cas $E = \mathbb{R}$, c'est la fonction conjuguée de Young de f.

Proposition 8.3.3. *La fonction conjuguée f^* est soit une fonction convexe sci de E' dans $\mathbb{R} \cup \{+\infty\}$, soit la fonction identiquement égale à $-\infty$.*

Démonstration : En effet si f est identiquement égale à $+\infty$, sa conjuguée est identiquement égale à $-\infty$. Sinon on peut se restreindre dans la définition de f^* aux éléments x de dom$f$, et f^* apparaît donc comme enveloppe supérieure de fontions affines continues, donc est convexe, sci et ne prend pas la valeur $-\infty$. □

Remarque 8.3.4. La valeur de la fonction conjuguée f^* au point x^* est la plus petite constante b telles que la fonction affine continue $h_{x^*,b} = \langle x^*, .\rangle - b$ minore la fonction f (avec la convention $\inf \emptyset = +\infty$). En particulier f^* est la constante $+\infty$ si et seulement si f n'a pas de minorante affine continue.

En effet $\langle x^*, .\rangle - b \leq f$ si et seulement si pour tout $x \in E$, on a $\langle x^*, x\rangle - f(x) \leq b$, ce qui équivaut à $f^*(x^*) \leq b$. □

Proposition 8.3.5. *1)* $f^*(0) = -\inf_{x\in E} f(x)$.
2) $f \leq g \Rightarrow f^* \geq g^*$.
3) Si $\lambda > 0$, *on a* $(\lambda f)^*(x^*) = \lambda f^*(\frac{x^*}{\lambda})$
4) Si $\lambda \neq 0$, *on a* $(D_\lambda f)^* = D_{1/\lambda} f^*$, *(où $D_\lambda f$ est la fonction dilatée $D_\lambda f(x) = f(\frac{x}{\lambda})$).*
5) Si $\alpha \in \mathbb{R}$, *on a* $(f + \alpha)^* = f^* - \alpha$
6) Si $a \in E$, *on a* $(\tau_a f)^* = f^* + \langle ., a\rangle$ *(où $\tau_a f$ est la fonction translatée : $\tau_a f(x) = f(x - a)$).*
7) $(\inf_{i\in I} f_i)^* = \sup_{i\in I} f_i^*$
8) $(\sup_{i\in I} f_i)^* \leq \inf_{i\in I} f_i^*$

Démonstration : Exercice !

Exemple 8.3.6. *Si A est une partie de E, χ_A sa fonction caractéristique, alors $(\chi_A)^*(x^*) = \sup_{x\in A}\langle x^*, x\rangle$: c'est la fonction d'appui de l'ensemble A. En particulier si B_E est la boule-unité de E, on a $\chi^*_{B_E}(x^*) = \|x^*\|$.*

Exemple 8.3.7. *Soit $u^* \in E'$: c'est une fonction sur E que nous notons $\langle u^*, .\rangle$ pour éviter les confusions. Alors sa conjuguée est très dégénérée : $\langle u^*, .\rangle^* = \chi_{\{u^*\}}$.*

Exemple 8.3.8. *Si $f(x) = \|x\|$, on a $f^* = \chi_{\bar{B}_{E'}}$.*

Exemple 8.3.9. *Soit* $E = \mathbb{R}$.
Si $1 < p < \infty$ *et* $f(t) = \dfrac{|t|^p}{p}$, *on a* $f^*(s) = \dfrac{|s|^q}{q}$, *où* q *est l'exposant conjugué de* p *c'est à dire* $1/p + 1/q = 1$.
Si $f(t) = |t|$, *on a* $f^*(s) = \chi_{[-1,1]}(s)$, *et inversement si* $f = \chi_{[-1,1]}$ *on a* $f^*(s) = |s|$.

Exemple 8.3.10. *Si* $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ *est paire et* $f : E \to E$ *est définie par* $f(x) = \varphi(\|x\|)$, *alors* $f^*(x^*) = \varphi^*(\|x^*\|)$.

En effet

$$\begin{aligned} f^*(x^*) &= \sup_{t\geq 0} \sup_{x\in E, \|x\|=t} [\langle x^*, x\rangle - \varphi(\|x\|)] \\ &= \sup_{t\geq 0} \sup_{x\in E, \|x\|=t} [\langle x^*, x\rangle - \varphi(t)] = \sup_{t\geq 0} [t\,\|x^*\| - \varphi(t)] \end{aligned}$$

et comme φ est paire :

$$= \sup_{t\in\mathbb{R}} [t\,\|x^*\| - \varphi(t)] = \varphi^*(\|x^*\|)$$

□

8.4 Biconjuguée

Si g est une fonction $E' \to \overline{\mathbb{R}}$, sa conjuguée g^* est stricto sensu une application $E'' \to \overline{\mathbb{R}}$. On ne considérera en général que la restriction $g^{(*)}$ de g^* à E, autrement dit la fonction $E \to \overline{\mathbb{R}}$ définie par :

$$g^{(*)}(x) = \sup_{x^*\in E'} [\langle x^*, x\rangle - g(x^*)].$$

Lorsque E est réflexif ($E'' = E$), par exemple si E est de dimension finie ou un espace de Hilbert, on a $g^{(*)} = g^*$. Par abus de notation, on note souvent g^* au lieu de $g^{(*)}$, même lorsque E n'est pas réflexif.

Remarque 8.4.1. La Proposition 8.3.5 est valable aussi pour cette notion de conjugaison.

Théorème 8.4.2 (Théorème de la biconjuguée)**.** *La biconjuguée* $f^{*(*)} : E \to \overline{\mathbb{R}}$ *d'une fonction* $f : E \to \overline{\mathbb{R}}$ *est égale à la régularisée convexe sci de* f *si* f *admet une minorante affine continue ; c'est la constante* $-\infty$ *sinon.*

DÉMONSTRATION : D'après la Remarque 8.3.4, la fonction affine continue $h_{x^*,b}$ définie par $h_{x^*,b}(x) = \langle x^*, x\rangle - b$ minore f si et seulement si $(x^*, b) \in \mathrm{epi} f^*$. Si une telle minorante affine existe, on a :

$$\hat{f}(x) = \sup_{(x^*,b)\in \mathrm{epi} f^*} h_{x^*,b}(x) = \sup_{x^*\in \mathrm{dom} f^*} [\langle x^*, x\rangle - f^*(x^*)] = f^{*(*)}(x)$$

□

Corollaire 8.4.3. *Si $f : E \to \overline{\mathbb{R}}$ est une fonction convexe sci ne prenant pas la valeur $-\infty$ (ou constamment égale à $-\infty$) on a $f^{*(*)} = f$.*

Remarque 8.4.4. De même si g est une fonction $E' \to \mathbb{R}$, sa biconjuguée $g^{(*)*}$ est l'enveloppe supérieure des minorantes affines de g de la forme : $h_{x,b}(x^*) = \langle x^*, x\rangle - b$ (minorantes affines w^*-continues). En particulier $g^{(*)*} \leq g$.

Corollaire 8.4.5. *Si f est une fonction $E \to \overline{\mathbb{R}}$, on a $f^{*(*)*} = f^*$.*

DÉMONSTRATION : De $f^{*(*)} \leq f$ on tire $f^{*(*)*} = (f^{*(*)})^* \geq f^*$; inversement on a $f^{*(*)*} = (f^*)^{(*)*} \leq f^*$. □

Notations 8.4.6. *Soit $\Gamma(E)$ la réunion de l'ensemble des fonctions $E \to \mathbb{R} \cup \{+\infty\}$ convexe sci et de la constante $-\infty$.*

Lorsque E est réflexif (par exemple si E est de dimension finie, ou un espace de Hilbert), toutes les fonctions affines continues sur E' sont de la forme $h_{x,b}(x^*) = \langle x^*, x\rangle - b$ ($x \in E$, $b \in \mathbb{R}$). Par conséquent on a :

Corollaire 8.4.7. *Supposons l'espace E réflexif (par exemple E de dimension finie, ou espace de Hilbert). La conjuguaison induit alors une bijection de $\Gamma(E)$ sur $\Gamma(E')$, qui échange les constantes $+\infty$ et $-\infty$.*

Exercice 8.1. *Si $(f_i)_{i\in I}$ est une famille de fonctions de $\Gamma(E)$ alors $(\sup_{i\in I} f_i)^* = (\inf_{i\in I} f_i^*)^{(*)*}$.*

8.5 Conjugaison et sous-différentiels

Remarque 8.5.1. Il résulte immédiatement de la définition de la conjugaison que pour toute fonction $f : E \to \overline{\mathbb{R}}$, et tous $x \in E, x^* \in E'$, on a :

$$\langle x^*, x\rangle \leq f(x) + f^*(x^*)$$

dès que le second membre a un sens. Cette inégalité s'appelle l'inégalité de Fenchel.

Le cas d'égalité est caractérisé par la proposition suivante :

Proposition 8.5.2. *Soit* $f : E \to \overline{\mathbb{R}}$ *et soit* f^* *sa conjuguée. Soient* $x \in E, x^* \in E^*$. *Alors* x^* *est un sous-gradient de* f *en* x *si et seulement si* $f(x)$ *est fini et*

$$\langle x^*, x\rangle = f(x) + f^*(x^*)$$

DÉMONSTRATION : La forme linéaire x^* est un sous-gradient de f en x si et seulement si la fonction affine continue $h(y) = \langle x^*, y - x\rangle + f(x)$ est une minorante de f. D'après la Remarque 8.3.4, c'est le cas si et seulement si $-h(0) = \langle x^*, x\rangle - f(x) \geq f^*(x^*)$. L'inégalité inverse étant toujours vraie, ceci équivaut à l'égalité. □

Remarque 8.5.3. De manière analogue si $g : E' \to \overline{\mathbb{R}}$, et $x \in E, x^* \in E^*$, alors x est un sous-gradient de g en x^* si et seulement si $g(x^*)$ est fini et

$$\langle x^*, x\rangle = g(x^*) + g^{(*)}(x)$$

Corollaire 8.5.4. *Si* $f : E \to \overline{\mathbb{R}}$ *et* $x \in E$ *vérifient* $f^{*(*)}(x) = f(x)$ *alors*

$$x^* \in \partial f(x) \iff x \in \partial f^*(x^*)$$

La condition $f^{*(*)}(x) = f(x)$ *est vérifiée en particulier dans les cas suivants :*
1) lorsque f *est sous-différentiable en* x *et dans ce cas* $\partial f(x) = \partial f^{*(*)}(x)$.
2) lorsque $f \in \Gamma(E)$.

DÉMONSTRATION : On a :

$$\begin{aligned} x^* \in \partial f(x) &\iff \langle x^*, x\rangle = f(x) + f^*(x^*) \\ x \in \partial f^*(x^*) &\iff \langle x^*, x\rangle = f^{*(*)}(x) + f^*(x^*) \end{aligned}$$

et ces conditions coïncident quand $f(x) = f^{*(*)}(x)$. Ceci est en particulier vrai pour tout x si $f = f^{*(*)}$ c'est à dire si $f \in \Gamma(E)$.

Si $\partial f(x) \neq \emptyset$, la fonction f admet une minorante affine continue qui lui est égale au point x, donc $f^{*(*)}(x) = f(x)$. Alors les assertions $x^* \in \partial f(x)$ et $x^* \in \partial f^{*(*)}(x)$ sont équivalentes à la même relation $\langle x^*, x\rangle = f(x) + f^*(x^*) = f^{*(*)}(x) + f^*(x^*)$. □

Corollaire 8.5.5. *Supposons que* $f : E \to \overline{\mathbb{R}}$ *est Gâteaux différentiable en* x, *et que* f^* *est Gâteaux différentiable en* $x^* = f'(x)$. *Alors* $x = f^{*\prime}(x^*)$.

8.6 Dualité dans les problèmes d'optimisation convexe

Dualité de Rockafellar.

Soient E, F des EVN et Φ une application convexe $E \times F \to \mathbb{R} \cup \{+\infty\}$.

L'espace vectoriel $E \times F$ est normé, par exemple par $\|(x,p)\| = \|x\| + \|p\|$. On identifie $(E \times F)'$ à $E' \times F'$, le crochet de dualité étant défini par :

$$\langle (x^*, p^*), (x,p) \rangle = \langle x^*, x \rangle + \langle p^*, p \rangle$$

On suppose que la fonction $\Phi(x, 0)$ n'est pas identiquement $+\infty$.

Définition 8.6.1 (Problème dual)**.** *On considère le problème :*

$$\mathcal{P} : \qquad \textit{Minimiser } \Phi(x,0), \ \textit{ pour } x \in E$$

Le problème dual (sous-entendu : relativement à la fonction Φ) est par définition :

$$\mathcal{P}^* : \qquad \textit{Maximiser } (-\Phi^*(0,p^*)), \ \textit{ pour } p^* \in F'$$

Le nombre $\inf \mathcal{P} = \inf_{x \in E} \Phi(x, 0)$*, (resp.* $\sup \mathcal{P}^* = \sup_{p^* \in F'} (-\Phi^*(0,p^*))$*), élément de* $\mathbb{R} \cup \{-\infty\}$*) est appelé valeur du problème $\mathcal{P}$ (resp. $\mathcal{P}^*$).*

Définition 8.6.2 (Solutions optimales d'un problème et de son dual)**.** *On dit que $\bar{x} \in E$ (resp. $\bar{p}^* \in F'$) est solution optimale de $\mathcal{P}$ (resp. de $\mathcal{P}^*$) si* $\inf \mathcal{P} = \Phi(\bar{x}, 0)$ *(resp.* $\sup \mathcal{P}^* = -\Phi^*(0, \bar{p}^*)$*). On notera dans ce cas* $\min \mathcal{P}$ *au lieu de* $\inf \mathcal{P}$ *(resp.* $\max \mathcal{P}^*$ *au lieu de* $\sup \mathcal{P}^*$*).*

Remarque 8.6.3. On a toujours $\sup \mathcal{P}^* \leq \inf \mathcal{P}$.

Démonstration : Ceci résulte trivialement de l'inégalité de Fenchel :

$$\forall x \in E, \forall p^* \in F' \qquad \Phi(x,0) + \Phi^*(0,p^*) \geq \langle 0, x \rangle + \langle p^*, 0 \rangle = 0 \quad \square$$

Posons $h(p) = \inf_{x \in E} \Phi(x,p)$. On a clairement $\inf \mathcal{P} = h(0)$. Notons aussi que h ne peut être la constante $+\infty$ (car $\Phi \not\equiv +\infty$).

Lemme 8.6.4. *On a $h^*(p^*) = \Phi^*(0,p^*)$, de sorte que* $\sup \mathcal{P}^* = h^{**}(0)$ *et, si h^{**} n'est pas la constante $-\infty$, l'ensemble des solutions optimales de $\mathcal{P}^*$ est égal à $\partial h^{**}(0)$.*

DÉMONSTRATION : On a :

$$h^*(p^*) = (\inf_{x\in E} \Phi(x,.))^*(p^*) = \sup_{x\in E} \Phi(x,.)^*(p^*)$$
$$= \sup_{x\in E}\sup_{p\in F}[\langle p^*, p\rangle - \Phi(x,p)] = \Phi^*(0,p^*)$$

Alors

$$h^{**}(0) = \sup_{p^*\in F'}[-h^*(p^*)] = \sup_{p^*\in F'}(-\Phi^*(0,p^*)) = \sup \mathcal{P}^*$$

si et seulement si $-h^*(p^*) = h^{**}(0)$, soit $h^*(p^*) + h^{**}(0) = \langle 0, p^*\rangle$
Notons que le membre de gauche existe puisque h^{**} diffère des constantes $\pm\infty$) ; d'après la Remarque 8.5.3, ceci équivaut à $p^* \in \partial h^{**}(0)$. □

Lemme 8.6.5. *La fonction h est convexe.*

DÉMONSTRATION : On a epist$h = \{(x,t) \in E\times\mathbb{R}/\ \exists p \in F : t > \Phi(x,p)\} = \{(x,t) \in E\times\mathbb{R}/\ \exists p \in F : (x,p,t) \in \text{epist}\Phi\}$: donc epist$h$ est l'image de epistΦ par la projection naturelle $E \times F \times \mathbb{R} \to E \times \mathbb{R}$. Donc la convexité de epistΦ entraîne celle de epist$h$. □

Théorème 8.6.6 (Égalité des solutions optimales)**.** *Supposons que la fonction Φ vérifie la condition :*

(H) Il existe $x_0 \in E$ telle que la fonction $\Phi(x_0,.)$ soit finie et continue au point $p = 0$

Alors le problème $\mathcal{P}^*$ a des solutions optimales et $\inf \mathcal{P} = \max \mathcal{P}^*$.

DÉMONSTRATION : La conclusion est trivialement vérifiée si $\inf \mathcal{P} = -\infty$. On suppose donc $h(0) = \inf \mathcal{P} > -\infty$. Comme la fonction $\Phi(x_0,.)$ est finie et continue en 0, elle est majorée au voisinage de 0. A fortiori, la fonction h est majorée au voisinage de 0, et comme $h(0) > -\infty$, h ne prend pas la valeur $-\infty$ (Lemme 1.5.13). Par conséquent, d'après la Proposition 7.1.1, h est continue en 0. D'après la Proposition 7.3.6 elle est sous-différentiable en 0. D'après le Corollaire 8.5.4, on a $h^{**}(0) = h(0)$ et $\partial h^{**}(0) = \partial h(0)$. D'après le Lemme 8.6.4, on a $\inf \mathcal{P} = \sup \mathcal{P}^*$, et $\mathcal{P}^*$ a des solutions optimales. □

Proposition 8.6.7. *Si les problèmes $\mathcal{P}$ et $\mathcal{P}^*$ admettent tous deux des solutions optimales, et si :*

$$\min \mathcal{P} = \max \mathcal{P}^* \in \mathbb{R}$$

Alors toute solution optimale $\bar{x}$ de $\mathcal{P}$ et toute solution optimale $\bar{p}^*$ de $\mathcal{P}^*$ sont liées par la relation d'extrémalité :

$$\Phi(\bar{x},0) + \Phi^*(0,\bar{p}^*) = 0$$

ce qui équivaut à : $(0,\bar{p}^*) \in \partial\Phi(\bar{x},0)$ (et si Φ est sci à $(\bar{x},0) \in \partial\Phi^*(0,\bar{p}^*)$).
Réciproquement, si $\bar{x} \in E$ et $\bar{p}^* \in F'$ vérifient la relation d'extrémalité, alors $\bar{x}$ est solution optimale de $\mathcal{P}$, $\bar{x}^*$ est solution optimale de $\mathcal{P}^*$, et $\min \mathcal{P} = \max \mathcal{P}^*$.

Démonstration : Comme $\inf \mathcal{P} = \Phi(\bar{x}, 0)$, $\sup \mathcal{P}^* = -\Phi^*(0, \bar{p}^*)$, ces deux quantités sont égales (et finies) si et seulement si $\Phi(\bar{x}, 0) + \Phi^*(0, \bar{p}^*) = 0$. En écrivant $0 = \langle(\bar{x}, 0), (0, \bar{x}^*)\rangle$, il résulte (égalité dans l'inégalité de Fenchel) que cette relation équivaut à : $(0, \bar{p}^*) \in \partial\Phi(\bar{x}, 0)$ (elle entraîne donc $(\bar{x}, 0) \in \partial\Phi^*(0, \bar{p}^*)$, et est équivalente à cette relation lorsque $\Phi^{**}(\bar{x}, 0) = \Phi(\bar{x}, 0)$, en particulier si Φ est sci).

Réciproquement, si cette relation d'extrémalité est vérifiée, on a :

$$\inf \mathcal{P} \leq \Phi(\bar{x}, 0) = -\Phi^*(0, \bar{p}^*) \leq \sup \mathcal{P}^*$$

mais comme on a toujours $\sup \mathcal{P}^* \leq \inf \mathcal{P}$ (Remarque 8.6.3), ces deux inégalités sont des égalités. □

Corollaire 8.6.8 (Théorème de Fenchel-Rockafellar). *Soient $f, g : E \to \mathbb{R} \cup \{+\infty\}$, deux fonctions convexes. On suppose qu'il existe $x_0 \in E$ tel que*

$$(H_1) \begin{cases} f(x_0) < +\infty \qquad g(x_0) < +\infty \\ f \text{ est continue en } x_0 \end{cases}$$

Alors

$$\inf_{x \in E} (f(x) + g(x)) = \max_{x^* \in E'} (-f^*(-x^*) - g^*(x^*))$$

Démonstration : On applique le Théorème 8.6.6 à la fonction $\Phi : E \times E \to \mathbb{R} \cup \{+\infty\}$ définie par :

$$\Phi(x, p) = f(x - p) + g(x)$$

L'hypothèse (H_1) entraîne bien que la fonction $\Phi(x_0, .)$ est finie et continue en 0. On vérifie aisément que :

$$\Phi^*(x^*, p^*) = f^*(-p^*) + g^*(x^* + p^*)$$

De sorte que le problème $\mathcal{P}$: minimiser $(f(x) + g(x))$
a pour problème dual $\mathcal{P}^*$: maximiser $(-f^*(-p^*) - g^*(p^*))$. □

Exemple 8.6.9. *Soit C un convexe fermé non vide de E. Si $x_0 \in E$, on a :*

$$d(x_0, C) = \inf_{x \in C} \|x_0 - x\| = \inf_{x \in E} [\|x_0 - x\| + \chi_C(x)]$$

Soit $f(x) = \|x - x_0\|$, $g(x) = \chi_C(x)$, on a :

$$f^*(x^*) = \chi_{\overline{B}_{E'}}(x^*) + \langle x^*, x_0\rangle$$

$$g^*(x^*) = \sup_{x \in C} \langle x^*, x\rangle = h_C(x^*)$$

On déduit donc du Corollaire 8.6.8 la relation :

$$d(x_0, C) = \max_{\|x^*\| \leq 1} [-h_C(x^*) + \langle x^*, x_0\rangle]$$

On considère maintenant un problème d'optimisation convexe ordinaire, c'est à dire du type :

$$(\mathcal{P})\begin{cases} \text{minimiser} & f(x) \\ \text{sous la contrainte} & g(x) \leq 0 \end{cases}$$

où f est une fonction convexe $E \to \mathbb{R} \cup \{+\infty\}$ et g est une fonction $E \to \mathbb{R}^m$ convexe, c'est à dire que $g(x) = (g_1(x), \ldots, g_m(x))$ avec g_i convexe pour tout $i = 1, \ldots, m$.

Définition 8.6.10 (Condition de Slater). *On dit que la contrainte $g(x) \leq 0$ vérifie la condition de qualification de Slater s'il existe x_0 tel que $f(x_0) < +\infty$ et $g_i(x_0) < 0$ pour tout $i = 1, \ldots, m$.*

Corollaire 8.6.11. *Sous la condition de qualification de Slater, il existe $\bar{p}^* \in \mathbb{R}^m$, $\bar{p}^* \leq 0$ tel que la valeur du problème $(\mathcal{P})$ coïncide avec celle du problème d'optimisation convexe sans contrainte :*

$$(\mathcal{P}_{\bar{p}^*}) : \qquad \textit{Minimiser } L(x, \bar{p}^*) = f(x) - \langle \bar{p}^*, g(x) \rangle, \quad x \in E$$

De plus la valeur de $(\mathcal{P})$ est en fait la plus grande des valeurs des problèmes $(\mathcal{P}_{p^*})$, $p^* \leq 0$.

Définition 8.6.12 (Multiplicateurs de Lagrange). *1) La fonction $L : E \times \mathbb{R}^m \to \mathbb{R} \cup \{-\infty\}$, définie par*

$$L(x, p^*) = \begin{cases} f(x) - \langle p^*, g(x) \rangle & \text{si } p^* \leq 0 \\ -\infty \text{ sinon} \end{cases}$$

est appelée Lagrangien du problème $(\mathcal{P})$.
2) Le m-uplet $\bar{p}^*$ est appelé un système de multiplicateurs de Lagrange du problème $(\mathcal{P})$.

Démonstration : Posons

$$\Phi(x, p) = f(x) + \chi_{C_p}(x)$$

où $C_p = \{x \in E \;/\; g(x) \leq p\}$. La fonction $\Phi : E \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ est convexe ; de plus $\Phi(x_0, p) = f(x_0)$ pour tout $p \geq g(x_0)$; l'ensemble $\{p \in \mathbb{R}^m \;/\; p > g(x_0)\}$ est un ouvert, contenant 0 d'après la condition de Slater ; autrement dit la fonction $p \to \Phi(x_0, p)$ est constante sur un voisinage de 0, et donc a fortiori continue au point 0. Les conditions du Théorème 8.6.6 sont donc réalisées.

Explicitons le problème dual $(\mathcal{P}^*)$, relativement à Φ. La fonction conjuguée est :

$$\begin{aligned} \Phi^*(x^*, p^*) &= \sup_{x \in E,\, p \in \mathbb{R}^m} [\langle x^*, x \rangle + \langle p^*, p \rangle - f(x) - \chi_{\{g(x) \leq p\}}] \\ &= \sup_{x, p : g(x) \leq p} [\langle x^*, x \rangle + \langle p^*, p \rangle - f(x)] \end{aligned}$$

d'où :

$$\Phi^*(0,p^*) = \sup_{x,p\,:\,g(x)\le p} [\langle p^*, p\rangle - f(x)]$$
$$= \begin{cases} \sup\limits_{x\in E}[\langle p^*, g(x)\rangle - f(x)] \text{ si } p^* \le 0 \\ +\infty \text{ sinon} \end{cases}$$

Le problème dual est donc :

$$(\mathcal{P}^*) \begin{cases} \text{Maximiser} & -\Phi(0,p^*) = \inf\limits_{x\in E}[f(x) - \langle p^*, g(x)\rangle] \\ \text{sous la contrainte} & p^* \le 0 \end{cases}$$

Le Théorème 8.6.6 montre que :

$$\inf_{g(x)\le 0} f(x) = \max_{p^*\le 0} \inf_{x\in E}[f(x) - \langle p^*, g(x)\rangle] \qquad \square$$

Proposition 8.6.13 (Théorème de Kuhn-Tucker)**.** *Sous les hypothèses du Corollaire 8.6.11, si le problème* $(\mathcal{P})$ *admet une solution optimale* $\bar{x}$*, celle-ci est également solution optimale d'un problème sans contrainte*

$$(\mathcal{P}_{\bar{p}^*}) : \qquad \textit{Minimiser } L(x,\bar{p}^*) = f(x) - \langle \bar{p}^*, g(x)\rangle, \quad x \in E$$

où $\bar{p}^* \le 0$ *vérifie la condition de Kuhn-Tucker*

$$\textit{(KT) :} \qquad \bar{p}^*.g(x) = 0$$

Réciproquement si $\bar{x}$ *est une solution optimale du problème sans contrainte* $(\mathcal{P}_{\bar{p}^*})$ *(avec* $\bar{p}^* \le 0$*) telle que* $g(\bar{x}) \le 0$*, et si le couple* $(\bar{x}, \bar{p}^*)$ *vérifie la condition de Kuhn-Tucker (KT), alors* $\bar{x}$ *est solution optimale du problème* $(\mathcal{P})$*, et* $\bar{p}^*$ *est un système de multiplicateurs de Lagrange pour* $(\mathcal{P})$*.*

Démonstration : i) Supposons que $\bar{x}$ est solution optimale du problème (P). Soit $\bar{p}^*$ un système de multiplicateurs de Lagrange pour $(\mathcal{P})$ (Corollaire 8.6.11). On a donc :

$$f(\bar{x}) = \inf_x [f(x) - \langle \bar{p}^*, g(x)\rangle] \le f(\bar{x}) - \langle \bar{p}^*, g(\bar{x})\rangle$$

Par conséquent $\langle \bar{p}^*, g(\bar{x})\rangle \le 0$ (noter que $f(\bar{x}) \le f(x_0) < +\infty$, donc $f(\bar{x})$ est finie). Mais $\bar{p}^*.g(\bar{x}) \ge 0$ puisque $\bar{p}^* \le 0$ et $g(\bar{x}) \le 0$. Les nombres $p_i^* g_i(\bar{x})$ sont positifs, de somme négative : ils sont donc tous nuls. D'où la condition (KT) est vérifiée et $\bar{x}$ est solution optimale de $\mathcal{P}_{p^*}$.

ii) Réciproquement, supposons que $\bar{x}$ est une solution optimale du problème sans contrainte $(\mathcal{P}_{\bar{p}^*})$, avec $\bar{p}^* \le 0$, $g(\bar{x}) \le 0$ et (KT). On a donc :

$$\forall x \in E, f(x) - \langle \bar{p}^*, g(x)\rangle \ge f(\bar{x}) - \langle \bar{p}^*, g(\bar{x})\rangle = f(\bar{x})$$

Si de plus $g(x) \leq 0$, on a $f(x) - \langle \bar{p}^*, g(x) \rangle \leq f(x)$, d'où $f(x) \geq f(\bar{x})$. □

Remarque 8.6.14. *1)* La condition de qualification de Slater n'est pas requise dans la démonstration de la réciproque de la Proposition 8.6.13.
2) Supposons que l'espace E est de dimension finie ou est un espace de Hilbert [ou plus généralement un espace réflexif]. Si la fonction f est sci de même que les fonctions g_i, $i = 1, \ldots, m$ et si par exemple $f(x) \to +\infty$ lorsque $\|x\| \to \infty$, alors $(\mathcal{P})$ admet des solutions optimales : en effet dans ce cas le convexe $C_0 = \{x \ / \ g(x) \leq 0\}$ est fermé, et on peut appliquer le Théorème 6.1.1. On a la même conclusion lorsque C_0 est borné, sans hypothèse sur le comportement de f à l'infini.
3) Dire que $\bar{x}$ est solution du problème sans contrainte $(\mathcal{P}_{p^*})$ revient à dire que 0 est dans le sous-différentiel de la fonction $x \to L(x, \bar{p}^*)$ au point $\bar{x}$. Supposons par exemple les fonctions g_i continues et la fonction f sci. Dans ce cas le Théorème 7.3.9 montre que cette condition équivaut à $0 \in \partial f(\bar{x}) - \sum_{i=1}^m p_i^* \partial g(\bar{x})$, ou encore :

$$\partial f(\bar{x}) \cap (\sum_{i=1}^{m} p_i^* \partial g(\bar{x})) \neq \emptyset$$

Bibiographie

Oeuvres générales pour d'élargir ses connaissances :

[1] L. D. Berkovitz, *Convexity and Optimization in* $\mathbb{R}^n$, Wiley-Interscience 2001

[2] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press 2004

[3] I. Ekeland, R. Temam, *Analyse convexe et problèmes variationnels*, Dunod 1974

[4] J.-B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*, Grundlehren Text Editions, Springer 2001

[5] R. Tyrell Rockafellar, *Convex Analysis*, Princeton University Press 1996

Références pour Théorème 5.5.3 :

[6] E. J. Candès, *The restricted isometry property and its implications for compressed sensing*, C. R. Acad. Sci. Paris, Ser. I 346 (2008), 589-592.

[7] R. Vershynin, *On the role of sparsity in Compressed Sensing and random matrix theory*, 2009 3rd IEEE International Workshop on CAMSAP, Aruba, Dutch Antilles, 2009, 189-192.

Index