

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 L'espace euclidien <math>\mathbb{R}^d</math></b>	<b>3</b>
1.1 La structure vectorielle de $\mathbb{R}^d$ . . . . .	3
1.2 La structure affine de $\mathbb{R}^d$ . . . . .	7
1.3 Produit scalaire, norme et métrique . . . . .	17
1.4 Topologie . . . . .	18
<b>2 Ensembles convexes</b>	<b>22</b>
2.1 Définition et propriétés de base . . . . .	22
2.2 Enveloppe convexe . . . . .	25
2.3 L'intérieur relatif d'un convexe . . . . .	30
2.4 Le théorème de Radon et le théorème de Helly . . . . .	32
2.5 Séparation . . . . .	34
2.6 Appui et faces . . . . .	40
2.7 Convexes non bornés . . . . .	46
<b>3 Polytopes et ensembles polyédriques</b>	<b>51</b>
3.1 Définitions et théorème fondamental . . . . .	51
3.2 Familles d'exemples . . . . .	58
3.3 Résultats généraux élémentaires . . . . .	61
3.4 Le treillis des faces . . . . .	62
3.5 Polytopes Cycliques . . . . .	67
<b>4 Systèmes d'inégalités linéaires</b>	<b>72</b>
4.1 Le théorème de Fredholm . . . . .	72
4.2 Le lemme de Farkas . . . . .	73
4.3 La méthode de Fourier-Motzkin . . . . .	74
4.4 Programmation linéaire . . . . .	76

# Introduction

Le but de ce cours est de constituer une introduction à l'étude des ensembles convexes de  $\mathbb{R}^d$ , et en particulier des polytopes convexes (qui sont une généralisation des polygones convexes de  $\mathbb{R}^2$ , et des polyèdres convexes de  $\mathbb{R}^3$ .)

Ces notes sont basées sur le cours de mathématiques combinatoires donné par les professeurs J-P. Doignon et S. Fiorini en BA3 mathématiques à l'ULB, durant l'année académique 2007-2008. Des ajouts basés sur les références de Brøndsted [1], Matousek[6], Webster [7] et Ziegler[8] ont été faits. Mentionnons aussi l'ouvrage de Grünbaum [5], qui est un classique de la littérature mathématique.

# Chapitre 1

## L'espace euclidien $\mathbb{R}^d$

Ce chapitre est destiné à des rappels de géométrie dans  $\mathbb{R}^d$ . Nous rappelons la structure vectorielle, affine et euclidienne de  $\mathbb{R}^d$ . La dernière section est consacrée à des rappels sur les bases de la topologie usuelle de  $\mathbb{R}^d$  d'une part, et à une première introduction à l'intérieur relatif d'un ensemble d'autre part.

### 1.1 La structure vectorielle de $\mathbb{R}^d$

Cette section est constituée d'un enchaînement de définitions et de propriétés qui sont supposées connues. Le lecteur ayant du mal à se convaincre de la véracité des affirmations ira consulter un ouvrage d'algèbre linéaire.

L'ensemble  $\mathbb{R}^d$  forme un espace vectoriel sur le corps  $\mathbb{R}$  des réels. Lorsque cette structure est utilisée, les éléments de  $\mathbb{R}^d$  sont donc appelés *vecteurs* (lorsque la structure affine de  $\mathbb{R}^d$  est utilisée, les éléments sont appelés *points*). Le vecteur nul est comme d'habitude noté 0. Un sous-ensemble  $V$  de  $\mathbb{R}^d$  est un *sous-espace vectoriel* (ou un *sous-espace linéaire*) lorsqu'il forme un espace vectoriel pour l'addition et la multiplication vectorielle induite de celle de  $\mathbb{R}^d$ , c'est-à-dire lorsque  $u, v \in V \Rightarrow u + v \in V$  et  $v \in V, \lambda \in \mathbb{R} \Rightarrow \lambda v \in V$ .

Si  $V_i, i \in I$  est une collection de sous-espace linéaires de  $\mathbb{R}^d$ , alors il en va de même pour leur intersection  $\bigcap_{i \in I} V_i$ . Si  $X$  est un sous-ensemble de  $\mathbb{R}^d$ , l'intersection de tous les sous-espaces de  $\mathbb{R}^d$  contenant  $X$  est appelée le *sous-espace engendré* par  $X$ . On parle aussi de l'*enveloppe vectorielle* ou de l'*enveloppe linéaire* de  $X$ . Il s'agit du plus petit sous-espace de  $\mathbb{R}^d$  contenant  $X$ . On le notera  $Vect(X)$ .

Si  $x_1, \dots, x_n$  sont des vecteurs de  $X \subseteq \mathbb{R}^d$  et si  $\lambda_1, \dots, \lambda_n$  sont des réels quelconques, alors le vecteur

$$\sum_{i=1}^n \lambda_i x_i$$

est une *combinaison linéaire* des  $x_i$ . L'enveloppe linéaire de  $X$  est exactement

égale à l'ensemble des combinaisons linéaires des éléments de  $X$  :

$$\text{Vect}(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_i \in \mathbb{R}, x_i \in X, n \in \mathbb{N} \right\}.$$

Nous utilisons la convention qu'une somme vide (c'est-à-dire pour  $n = 0$ ) est égale au vecteur nul  $0$  de sorte que pour l'ensemble vide  $\emptyset$ , nous avons bien  $\text{Vect}(\emptyset) = \{0\}$ .

Un ensemble  $X$  de vecteurs de  $\mathbb{R}^d$  est dit *linéairement indépendant* si aucun vecteur de  $X$  n'est dans l'enveloppe linéaire des autres, c'est-à-dire si aucun vecteur de  $X$  n'est combinaison linéaire des autres. Sinon,  $X$  est dit *linéairement dépendant*. La partie  $X$  est linéairement indépendante si et seulement si  $\sum_{i=1}^n \lambda_i x_i = 0$  pour certains  $x_i \in X$  et  $\lambda_i \in \mathbb{R}$  implique  $\lambda_i = 0$  pour tout  $i$ . Une partie linéairement indépendante de vecteurs de  $X$  ne contient jamais le vecteur nul.

L'ensemble  $X$  est dit *générateur* de  $\mathbb{R}^d$  (ou plus généralement de tout sous-espace  $V$  de  $\mathbb{R}^d$ ) si  $\text{Vect}(X) = \mathbb{R}^d$  (respectivement  $\text{Vect}(X) = V$ ). Les parties génératrices minimales (pour l'inclusion) d'un sous-espace  $V$  de  $\mathbb{R}^d$  correspondent aux parties linéairement indépendantes maximales de  $V$ . Ces parties sont appelées *bases*. Tout espace vectoriel admet une base, et toutes les bases d'un même espace ont le même cardinal, à savoir  $d$  dans le cas où  $V = \mathbb{R}^d$ . La dimension d'un sous-espace vectoriel  $V$  de  $\mathbb{R}^d$  est la cardinalité d'une base, on la note  $\dim(V)$ . Un sous-espace de  $\mathbb{R}^d$  de dimension 1 est appelé une *droite vectorielle*, un sous-espace de dimension 2 un *plan vectoriel*, un sous-espace de dimension  $d - 1$  un *hyperplan vectoriel* (les hyperplans vectoriels correspondent donc aux sous-espaces propres maximaux de  $\mathbb{R}^d$ ).

Si  $B$  est une base de  $\mathbb{R}^d$ , alors un vecteur  $x$  s'exprime toujours d'une et une seule façon comme combinaison linéaire des vecteurs de  $B$ . La base constituée des  $d$  vecteurs

$$(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 0, 1)$$

est la *base canonique*.

Comme une partie libre de  $\mathbb{R}^d$  comporte au plus  $d$  éléments, une combinaison linéaire peut toujours s'exprimer avec au plus  $d$  éléments et nous avons

$$\text{Vect}(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_i \in \mathbb{R}, x_i \in X, n \in \mathbb{N}, 0 \leq n \leq d \right\}.$$

Les morphismes  $f : V \rightarrow \mathbb{R}^e$  (où  $V$  est un sous-espace vectoriel de  $\mathbb{R}^d$ ) de la structure d'espace vectoriel sont appelés *applications linéaires*. Il s'agit donc d'applications telles que  $f(u + v) = f(u) + f(v)$  et  $f(\lambda u) = \lambda f(u)$  pour tous  $u, v \in V$  et tout  $\lambda \in \mathbb{R}$ . Il revient bien sûr au même de dire que

$$f\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i f(v_i)$$

pour toute combinaison linéaire de vecteurs  $v_1, \dots, v_n \in V$ . Si  $e = d$  et  $f(V) \subseteq V$ , on parle d'*opérateur linéaire*. Un opérateur linéaire bijectif est une *permutation linéaire*. Si  $B_1 = \{e_1, \dots, e_d\}$  et  $B_2 = \{f_1, \dots, f_d\}$  sont deux bases de  $\mathbb{R}^d$ , alors il existe une et une seule permutation linéaire  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  envoyant  $B_1$  sur  $B_2$  c'est-à-dire telle que

$$f(e_i) = f_i$$

pour  $i = 1, \dots, d$ . Si  $e = 1$ , on parle de *forme linéaire sur  $V$* . Les formes linéaires  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  sont exactement les applications s'écrivant

$$f(x_1, x_2, \dots, x_d) = \sum_{i=1}^d a_i x_i$$

pour  $a_1, \dots, a_d$  des nombres réels. Une bijection linéaire  $f : V_1 \rightarrow V_2$  avec  $V_1$  un sous-espace de  $\mathbb{R}^d$  et  $V_2$  un sous-espace de  $\mathbb{R}^e$  est un *isomorphisme linéaire*, les deux sous-espaces étant alors dit *isomorphes*. Deux sous-espaces sont isomorphes si et seulement s'ils ont même dimension.

Le *noyau* d'une application linéaire quelconque  $f$  est par définition  $\text{Ker}(f) = \{x \in \mathbb{R}^d \mid f(x) = 0\}$ , l'*image* de  $f$  est  $\text{Im}(f) = \{f(x) \mid x \in \mathbb{R}^d\}$ . Le noyau et l'image d'une application linéaire forment toujours des sous-espaces de  $\mathbb{R}^d$  et  $\mathbb{R}^e$  respectivement. Une application linéaire telle que  $\text{Ker}(f) = \mathbb{R}^d$ , c'est-à-dire telle que  $f(x) = 0$  pour tout  $x \in \mathbb{R}^d$  est dite *dégénérée*. On démontre alors qu'un hyperplan vectoriel  $H$  est toujours égal au noyau d'une forme linéaire non dégénérée  $f$ , c'est-à-dire

$$H = \{x \in \mathbb{R}^d \mid \sum_{i=1}^d a_i x_i = 0\}$$

pour certains  $a_i$  non tous nuls (les  $a_i$  étant tous nuls si et seulement si la forme linéaire est dégénérée).

Pour deux sous-ensembles  $A, B \subseteq \mathbb{R}^d$ , la *somme de Minkowski* de  $A$  et  $B$  est

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

Si  $A = \{a\}$  est un singleton, on notera souvent  $a + B$  au lieu de  $A + B$ . Si  $A$  et  $B$  sont des sous-espaces, alors il en va de même pour  $A + B$ . De plus la dimension de  $A + B$  vérifie alors

$$\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B).$$

Démontrons maintenant le joli théorème suivant.

**Théorème 1.1** *Soit  $V$  un sous-espace vectoriel de  $\mathbb{R}^d$ . Alors  $V$  s'écrit comme intersection d'un nombre fini d'hyperplans*

$$V = \bigcap_{i=1}^n H_i.$$

De plus, le plus petit  $n$  qui convient est

$$n = d - \dim(V).$$

Si  $V = \mathbb{R}^d$ , on a  $n = 0$  et  $V = \cap \emptyset$ .

**Démonstration** : Le théorème est clair si  $V = \mathbb{R}^d$ . Sinon, soit  $B_1 = \{e_1, \dots, e_p\}$  une base de  $V$  avec  $p = \dim(V)$  et complétons  $B_1$  pour obtenir  $B_2 = \{e_1, \dots, e_p, \dots, e_d\}$  une base de  $\mathbb{R}^d$  (si  $V = \emptyset$ , alors  $p = 0$ .) Pour  $1 \leq i \leq d$ , définissons les hyperplans

$$H_i = \text{Vect}(\{e_j \mid 1 \leq j \leq d, j \neq i\})$$

et montrons que

$$V = \bigcap_{j=p+1}^d H_j.$$

Notons  $D$  le membre de droite de l'égalité. Comme  $e_1, \dots, e_p \in D$  par construction, on a

$$V = \text{Vect}(\{e_j \mid 1 \leq j \leq p\}) \subseteq D$$

vu que  $D$  est un sous-espace vectoriel. De plus, si  $u \in D$ , alors  $u$  s'écrit d'une et une seule façon comme combinaison linéaire d'éléments de  $B_2$ . Mais cette combinaison linéaire ne peut contenir aucun vecteur  $e_j$  avec  $j > p$  sinon  $u$  ne serait pas de  $H_j$ . Donc  $u \in V$  et l'égalité est démontrée.

L'intersection d'hyperplans ci-dessus comporte  $d-p = d - \dim(V)$  hyperplans. Il n'est pas difficile de démontrer (par récurrence sur le naturel non nul  $l$ ) que pour des hyperplans  $H_1, \dots, H_l$  on a toujours

$$\dim(H_1 \cap \dots \cap H_l) \geq d - l$$

de sorte que  $d - \dim(V)$  est bien le plus petit  $n$  possible.  $\square$

Voici un corollaire utile.

**Corollaire 1.2** Soit  $V$  un sous-espace vectoriel de  $\mathbb{R}^d$  de dimension  $n \geq 1$ . Soit  $G$  un hyperplan de  $V$ . Alors il existe un hyperplan  $H$  de  $\mathbb{R}^d$  tel que

$$G = H \cap V.$$

**Démonstration** : Soit

$$B_1 = \{e_1, \dots, e_{n-1}\}$$

une base de  $G$  que l'on complète en une base

$$B_2 = \{e_1, \dots, e_n\}$$

de  $V$  que l'on complète à son tour en une base

$$B_3 = \{e_1, \dots, e_d\}$$

de  $\mathbb{R}^d$ . Considérons les hyperplans  $H_i$  de la preuve du théorème 1.1. Alors cette même preuve montre que

$$V = \bigcap_{j=n+1}^d H_j$$

et

$$G = \bigcap_{j=n}^d H_j$$

de sorte qu'on peut prendre  $H = H_n$ .  $\square$

Terminons cette section en remarquant que l'application  $Vect$  définie sur l'ensemble des parties  $\mathcal{P}(\mathbb{R}^d)$  de  $\mathbb{R}^d$  par

$$Vect : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d) : X \rightarrow Vect(X)$$

est un *opérateur de fermeture* sur l'ensemble des parties ordonné par la relation d'inclusion à savoir

1.  $X \subseteq Vect(X)$ ,
2.  $Vect(Vect(X)) = Vect(X)$ ,
3.  $X \subseteq Y \Rightarrow Vect(X) \subseteq Vect(Y)$ ,

pour tous  $X, Y \in \mathcal{P}(\mathbb{R}^d)$ .

Le lecteur désirant en savoir plus sur la théorie générale des opérateurs de fermeture consultera un ouvrage sur les ensembles ordonnés, par exemple Davey et Priestley [2]. Nous aurons l'occasion de rencontrer d'autres opérateurs de fermeture dans la suite.

## 1.2 La structure affine de $\mathbb{R}^d$

La structure affine de  $\mathbb{R}^d$  est déduite de la structure vectorielle. Les éléments de  $\mathbb{R}^d$  sont maintenant appelés *points*. Les *droites* sont les classes latérales des droites vectorielles de  $\mathbb{R}^d$  c'est-à-dire les ensembles de la forme

$$a + D$$

avec  $D$  une droite vectorielle de  $\mathbb{R}^d$ . On parle parfois de *droite affine* pour les distinguer des droites vectorielles, une droite affine étant vectorielle si et seulement si elle passe par 0. On définit de même les *plans affins*, les *hyperplans affins*.

Plus généralement, les *sous-espaces affins* sont soit un translaté d'un sous-espace vectoriel  $V$  soit la partie vide  $\emptyset$ . Deux sous-espaces affins sont dits *parallèles* s'ils sont les translatsés d'un même sous-espace vectoriel de  $\mathbb{R}^d$ . La *dimension*  $\dim(A)$  d'un sous-espace affine  $A$  de  $\mathbb{R}^d$  est la dimension du sous-espace vectoriel correspondant si  $A \neq \emptyset$ , sinon on pose  $\dim(\emptyset) = -1$ .

Le *rapport de section* de trois points  $a, b, c$  alignés avec  $a \neq b$  est le réel  $k$  tel que  $(c - a) = k(b - a)$ . La structure affine de  $\mathbb{R}^d$  est constituées des droites, des droites parallèles et du rapport de section.

Si  $p, q$  sont des points de  $\mathbb{R}^d$ , il existe une et une seule droite passant par  $p$  et  $q$ , on la note  $\langle p, q \rangle$  et on a

$$\langle p, q \rangle = \{p + \lambda(q - p) \mid \lambda \in \mathbb{R}\}.$$



On définit de même le *segment*  $[p, q]$  par

$$[p, q] = \{p + \lambda(q - p) \mid 0 \leq \lambda \leq 1\} = \{\mu p + \lambda q \mid \mu + \lambda = 1, \mu \geq 0, \lambda \geq 0\}.$$



Le *segment ouvert*  $]p, q[$  par

$$]p, q[ = \{p + \lambda(q - p) \mid 0 < \lambda < 1\} = \{\mu p + \lambda q \mid \mu + \lambda = 1, \mu > 0, \lambda > 0\}.$$



La *demi-droite*  $[p, q >$  est

$$[p, q > = \{p + \lambda(q - p) \mid \lambda \geq 0\}.$$



Et enfin la *demi-droite ouverte*  $]p, q >$

$$]p, q > = \{p + \lambda(q - p) \mid \lambda > 0\}.$$



La définition des *segments semi-ouverts*  $]p, q]$  et  $[p, q[$  est similaire. On démontre aisément :



**Proposition 1.3** *Un sous-ensemble de points  $S \subseteq \mathbb{R}^d$  est un sous-espace affín si et seulement si  $\langle p, q \rangle \subseteq S$  pour tous  $p, q \in S$ .*

Et on en déduit :

**Proposition 1.4** *Toute intersection de sous-espaces affíns de  $\mathbb{R}^d$  est encore un sous-espace affín.*

On dispose d'un théorème analogue au théorème 1.1 :

**Théorème 1.5** *Tout sous-espace affín non vide  $A$  de  $\mathbb{R}^d$  est une intersection d'un nombre fini d'hyperplans affíns  $H_1, \dots, H_n$  (même si  $A = \mathbb{R}^d$ , auquel cas  $n = 0$  et  $\cap \emptyset = \mathbb{R}^d$ ). De plus le plus petit  $n$  qui convient est  $n = d - \dim(A)$ .*

**Démonstration** : Il suffit de constater que pour  $A = p + V$  avec  $V$  un sous-espace vectoriel de  $V$  tel que

$$V = \bigcap_i H'_i$$

avec  $H'_i$  des hyperplans vectoriels, on a

$$p + \bigcap_i H'_i = \bigcap_i (p + H'_i).$$

Si l'égalité de l'énoncé à lieu pour un certain

$$n < d - \dim(A) = d - \dim(V)$$

alors

$$V = A - p = \bigcap_{i=1}^n (H_i - p)$$

et les  $H_i - p$  sont des hyperplans vectoriels, d'où la contradiction avec le théorème 1.1. La preuve est donc terminée.  $\square$

Et on démontre de même le corollaire :

**Corollaire 1.6** *Soit  $A$  un sous-espace affín de  $\mathbb{R}^d$  et  $G$  un hyperplan affín de  $A$ . Alors il existe un hyperplan affín  $H$  de  $\mathbb{R}^d$  tel que*

$$G = H \cap A.$$

Voici la réciproque du corollaire précédent (voyez par exemple [4]) :

**Proposition 1.7** *Soit  $A$  un sous-espace affín de  $\mathbb{R}^d$  et  $H$  un hyperplan affín de  $\mathbb{R}^d$  tel que  $H \cap A$  est non vide. Alors  $H \cap A$  est un hyperplan de  $A$ .*

L'intersection de tous les sous-espaces affins de  $\mathbb{R}^d$  contenant un certain sous-ensemble  $X \subseteq \mathbb{R}^d$  est appelée l'*enveloppe affine* de  $X$ . C'est le plus petit sous-espace affine contenant  $X$ . On la note  $Aff(X)$ . Il n'est pas difficile de démontrer que :

**Proposition 1.8** *Pour tout  $X \subseteq \mathbb{R}^d$  et tout  $p \in X$  il vient*

$$Aff(X) = p + Vect(X \setminus \{p\}).$$

Et il en découle :

**Corollaire 1.9** *Si  $0 \in Aff(X)$  alors*

$$\dim(Aff(X)) = \dim(Vect(X)).$$

*Sinon*

$$\dim(Aff(X)) = \dim(Vect(X)) - 1.$$

**Démonstration** : Si  $0 \in Aff(X)$  alors  $Aff(X) = Vect(X)$  donc il n'y a rien à démontrer. Sinon,  $Aff(X) = p + Vect(X \setminus \{p\})$  pour un certain  $p \in X$  non nul (si  $X = \emptyset$  l'égalité est évidente). Mais alors  $p \notin Vect(X \setminus \{p\})$  sinon  $Aff(X) = Vect(X \setminus \{p\})$  donc  $Aff(X)$  contiendrait 0. Ceci implique donc, si  $\dim(Aff(X)) = n$  et si  $\{u_1, \dots, u_n\}$  est une base de  $Vect(X \setminus \{p\})$ , que  $\{p, u_1, \dots, u_n\}$  est une partie linéairement indépendante et engendre  $p \cup Vect(X \setminus \{p\})$  c'est-à-dire  $Vect(X)$ . Il s'agit donc d'une base de  $Vect(X)$ , et la propriété est démontrée.  $\square$

De même que pour l'application  $Vect$ , l'application  $Aff : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$   $X \rightarrow Aff(X)$  est un opérateur de fermeture sur l'ensemble ordonné  $(\mathcal{P}(\mathbb{R}^d), \subseteq)$ .

Si  $x_1, \dots, x_n \in \mathbb{R}^d$  et si  $\lambda_1, \dots, \lambda_n$  sont des réels tels que  $\sum_{i=1}^n \lambda_i = 1$  alors

$$x = \sum_{i=1}^n \lambda_i x_i$$

est appelée *combinaison affine* des  $x_i$ .

On a alors (ce résultat est à mettre en parallèle avec celui concernant l'enveloppe linéaire) :

**Proposition 1.10** *L'enveloppe affine de  $X \subseteq \mathbb{R}^d$  correspond à l'ensemble des combinaisons affines des éléments de  $X$ , donc*

$$Aff(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid x_i \in X, n \in \mathbb{N}_0, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

**Démonstration** : Soit  $D$  le membre de droite de l'égalité. Montrons que  $Aff(X) \subseteq D$ . Clairement  $X \subseteq D$ , et il suffit donc de montrer que  $D$  est un sous-espace affiné de  $\mathbb{R}^d$ . Soient donc

$$x = \sum_{i=1}^n \lambda_i x_i$$

et

$$y = \sum_{j=1}^m \mu_j y_j$$

deux éléments de  $D$  et montrons que la droite passant par ces deux points est toujours dans  $D$ . Pour  $\alpha, \beta \in \mathbb{R}$  tels que  $\alpha + \beta = 1$  nous avons donc

$$\alpha \sum_{i=1}^n \lambda_i x_i + \beta \sum_{j=1}^m \mu_j y_j = \sum_{i=1}^n (\alpha \lambda_i) x_i + \sum_{j=1}^m (\beta \mu_j) y_j$$

et ceci est un élément de  $D$  car  $m + n \in \mathbb{N}_0$ ,  $x_i, y_j \in X$  et  $\sum_{i=1}^n \alpha \lambda_i + \sum_{j=1}^m \beta \mu_j = \alpha.1 + \beta.1 = \alpha + \beta = 1$ .

Montrons maintenant que  $D \subseteq Aff(X)$  par récurrence sur  $m \in \mathbb{N}_0$ . C'est clair si  $m = 1$ . Soit  $m > 1$ , alors pour  $x = \sum_{i=1}^n \lambda_i x_i \in D$  il existe un  $i$  tels que  $\lambda_i \neq 0$ , disons  $i = k$ . Dans ce cas nous pouvons écrire  $x$  ainsi

$$x = (1 - \lambda_k) \sum_{i=1, i \neq k}^n \frac{\lambda_i}{1 - \lambda_k} x_i + \lambda_k x_k.$$

Mais  $y_k = \sum_{i=1, i \neq k}^n \frac{\lambda_i}{1 - \lambda_k} x_i$  est dans  $Aff(X)$  par l'hypothèse de récurrence car

$$\sum_{i=1, i \neq k}^n \frac{\lambda_i}{1 - \lambda_k} = \frac{1 - \lambda_k}{1 - \lambda_k} = 1$$

et comme  $x_k \in Aff(X)$ , on a  $x \in Aff(X)$  car  $x = (1 - \lambda_k)y_k + \lambda_k x_k$  et  $Aff(X)$  est un sous-espace affiné.  $\square$

De même que pour les combinaisons linéaires de  $\mathbb{R}^d$  on peut se limiter à  $d$  éléments, pour les combinaisons affines on se limite à  $d + 1$  éléments :

**Proposition 1.11** *Pour tout  $X \subseteq \mathbb{R}^d$  on a*

$$Aff(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid x_i \in X, n \in \mathbb{N}_0, 1 \leq n \leq d + 1, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

**Démonstration** : Il suffit de prouver que toute combinaison affine de  $\sum_{i=1}^n \lambda_i x_i$  avec  $n > d + 1$  et  $\lambda_i \neq 0$  pour tout  $i$  peut s'écrire comme combinaison affine de  $n - 1$  éléments parmi les  $x_i$ . Nous savons donc que les éléments  $x_2 - x_1, x_3 - x_1,$

$\dots, x_n - x_1$  sont nécessairement linéairement dépendants (car il y en a  $n - 1 > d$ ). Il existe donc des réels  $\mu_2, \dots, \mu_n$  non tous nuls tels que

$$\sum_{j=2}^n \mu_j (x_j - x_1) = 0.$$

Mais alors en choisissant dans le membre de droite  $\mu_k \neq 0, k \neq 1$  on a

$$\sum_{i=1}^n \lambda_i x_i = \sum_{i=1}^n \lambda_i x_i - \frac{\lambda_k}{\mu_k} \sum_{j=2}^n \mu_j (x_j - x_1).$$

Ainsi  $x_k$  disparaît et il nous reste une combinaison affine de  $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$  car la somme des coefficients vaut bien

$$\sum_{i=1}^n \lambda_i - \frac{\lambda_k}{\mu_k} \sum_{j=2}^n \mu_j + \frac{\lambda_k}{\mu_k} \sum_{j=2}^n \mu_j = 1. \quad \square$$

Un ensemble  $X$  de points de  $\mathbb{R}^d$  est dit *affinement indépendant* si aucun point n'est dans l'enveloppe affine des autres. Sinon  $X$  est dit *affinement dépendant*.

**Proposition 1.12** *Un ensemble  $X \subseteq \mathbb{R}^d$  est affinement dépendant si et seulement si il existe un  $n \in \mathbb{N}_0, x_1, \dots, x_n \in X, \lambda_1, \dots, \lambda_n \in \mathbb{R}$  non tous nuls tels que*

$$\sum_{i=1}^n \lambda_i x_i = 0$$

et

$$\sum_{i=1}^n \lambda_i = 0.$$

**Démonstration** : Si  $x \in \text{Aff}(x_1, \dots, x_{n-1})$  pour  $x_1, \dots, x_{n-1} \in X$  et  $n \geq 2$  alors il existe  $\beta_1, \dots, \beta_{n-1}$  tels que  $x = \sum_{i=1}^{n-1} \beta_i x_i$  et  $\sum_{i=1}^{n-1} \beta_i = 1$ . En posant  $x_n = x$  ainsi que  $\lambda_i = -\beta_i$  et  $\beta_n = 1$  on a bien

$$\sum_{i=1}^n \lambda_i x_i = 0$$

ainsi que

$$\sum_{i=1}^n \lambda_i = 0.$$

De même, si  $\sum_{i=1}^n \lambda_i x_i = 0$  avec  $\sum_{i=1}^n \lambda_i = 0$  et  $\lambda_k \neq 0$  pour un certain  $k$ , alors

$$x_k = - \sum_{i=1, i \neq k}^n \frac{\lambda_i}{\lambda_k} x_i = 0.$$

et

$$- \sum_{i=1, i \neq k}^n \frac{\lambda_i}{\lambda_k} = -\frac{0 - \lambda_k}{\lambda_k} = 1. \quad \square$$

Voici encore une caractérisation des ensembles affinement indépendants :

**Proposition 1.13** *Un ensemble  $X \subseteq \mathbb{R}^d$  est affinement indépendant si et seulement si pour tout  $a \in X$ , l'ensemble*

$$X \setminus \{a\} - a = \{x - a \mid x \in X, x \neq a\}$$

*est linéairement indépendant.*

**Démonstration** : Si  $X \setminus \{a\} - a = \{x - a \mid x \in X, x \neq a\}$  est linéairement dépendant, il existe des réels  $\beta_1, \dots, \beta_n$  non tous nuls tels que pour certains  $x_1 - a, \dots, x_n - a$  on ait

$$\sum_{i=1}^n \beta_i (x_i - a) = 0.$$

Sans perte de généralité, on peut supposer  $\beta_1 \neq 0$ . Du coup

$$x_1 - a = -\frac{1}{\beta_1} \sum_{i=2}^n \beta_i (x_i - a) = \left(1 + \sum_{i=2}^n \frac{\beta_i}{\beta_1}\right) a - \frac{1}{\beta_1} \sum_{i=2}^n \beta_i x_i$$

avec

$$1 + \sum_{i=2}^n \frac{\beta_i}{\beta_1} - \sum_{i=2}^n \frac{\beta_i}{\beta_1} = 1$$

donc  $X$  est bien affinement dépendant.

Supposons maintenant  $X$  affinement dépendant. Il existe donc des éléments  $x, x_1, \dots, x_n$  de  $X$  tels que  $x = \sum_{i=1}^n \lambda_i x_i$ ,  $x$  différent de tous les  $x_i$  et  $\sum_{i=1}^n \lambda_i = 1$ . Alors pour tout  $a \in X$  on a

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

avec au moins un  $x_i$  différent de  $a$  (sinon  $x = x_i = a$  pour tout  $i$ ). Ainsi  $X \setminus \{a\} - a$  est bien linéairement dépendant.  $\square$

Un *repère affín* d'un sous-espace affín  $A$  de  $\mathbb{R}^d$  est un ensemble de points de  $A$  affinement indépendants de cardinal maximal. Si  $R$  est un repère affín de  $A$ , alors bien sûr  $A = \text{Aff}(R)$ .

Les repères affíns de  $A$  sont exactement les parties affinement indépendantes de cardinal  $\dim(A) + 1$  comme le montre le théorème ci-dessous :

**Théorème 1.14** Soit  $A \subseteq \mathbb{R}^d$  un sous-espace affiné. Alors  $\dim(A) = n$  si et seulement si le cardinal maximum d'une partie affinement indépendante de  $A$  est  $n + 1$ . Dans ce cas, si  $R = \{x_1, \dots, x_{n+1}\}$  est un repère affiné de  $A$ , tout point de  $A$  s'écrit d'une et une seule façon comme combinaison affine de points de  $R$ .

**Démonstration** : Soit  $A = p + V$  avec  $V$  un sous-espace vectoriel. Alors la dimension de  $A$  vaut  $n$  si et seulement si la dimension de  $V$  vaut  $n$  si et seulement si il existe une partie libre de cardinal maximum  $n$  dans  $V$ , soit  $\{a_1 - p, \dots, a_n - p\}$ , si et seulement si  $\{a_1, \dots, a_n, p\}$  est une partie affinement indépendante dans  $A$  (par la proposition 1.13).

De plus on a bien  $A = \text{Aff}(R)$  pour un repère affiné  $R \subseteq A$  (sinon  $R$  ne serait pas de cardinal maximum) et si  $x \in A$  s'écrit de deux façons comme combinaison affine d'éléments de  $R$  soit

$$x = \sum_{i=1}^{n+1} \lambda_i x_i$$

et

$$x = \sum_{i=1}^{n+1} \beta_i x_i$$

(les  $\lambda_i$  et  $\beta_i$  non tous égaux) alors on a

$$\sum_{i=1}^{n+1} (\lambda_i - \beta_i) x_i = 0$$

et

$$\sum_{i=1}^{n+1} (\lambda_i - \beta_i) = 1 - 1 = 0$$

donc  $R$  n'est pas affinement indépendante par la proposition 1.12.  $\square$

En particulier, les repères affins de  $\mathbb{R}^d$  sont les parties affinement indépendantes de cardinal  $d + 1$ . Les repères affins sont donc l'analogie affine des bases d'espaces vectoriels.

Les morphismes  $f : A \rightarrow \mathbb{R}^e$  (où  $A$  est un sous-espace affiné de  $\mathbb{R}^d$ ) de la structure affine de  $\mathbb{R}^d$  sont appelées *applications affines* (ou bien *affinités*). Il s'agit des applications envoyant toute droite sur une droite, deux droites parallèles sur deux droites parallèles, trois points ayant un rapport de section  $k$  sur trois points ayant un rapport de section  $k$ . Toute *translation* de  $\mathbb{R}^d$  est une application affine, c'est-à-dire toute application de la forme

$$\alpha_t : \mathbb{R}^d \rightarrow \mathbb{R}^d : x \rightarrow x + t$$

pour  $t$  un élément fixé de  $\mathbb{R}^d$ . Il est possible de montrer que toute affinité peut s'écrire comme la composée d'une application linéaire et de translations :

**Théorème 1.15** Toute affinité  $f : A \rightarrow \mathbb{R}^e$  (avec  $A = u + V$  un sous-espace affiné de  $\mathbb{R}^d$ ) s'écrit  $f = \alpha_t \circ \beta \circ \alpha_{-u}$  avec  $\beta$  une application linéaire de  $V$  vers  $\mathbb{R}^e$ ,  $\alpha_t$  la translation de  $\mathbb{R}^e$  de vecteur  $t$  et  $\alpha_{-u}$  la translation de vecteur  $-u$  dans  $\mathbb{R}^e$ . En d'autres termes,  $f$  s'écrit

$$f(u + x) = t + \beta(x)$$

pour tout  $x + u \in A$ .

L'image d'un sous-espace affiné par une application affine est toujours un sous-espace affiné. Une application affine bijective  $f : A_1 \rightarrow A_2$  avec  $A_1$  un sous-espace affiné de  $\mathbb{R}^d$  et  $A_2$  un sous-espace affiné de  $\mathbb{R}^e$  est un *isomorphisme affiné*, et dans ce cas  $A_1$  et  $A_2$  sont *isomorphes*. De même que dans le cas vectoriel, deux sous-espaces affines sont isomorphes si et seulement si ils ont même dimension.

Si  $A_1 = \mathbb{R}^d$  et  $e = 1$ , on parle de *forme affine*. Toute forme affine  $f$  s'écrit donc

$$f(x) = \sum_{i=1}^d a_i x_i + t$$

pour certains  $t, a_i \in \mathbb{R}$ . De façon analogue au cas vectoriel, tout hyperplan affiné  $H$  est l'image réciproque d'un point par une forme affine non triviale (c'est-à-dire telle qu'un des  $a_i$  est non nul), à savoir (il n'est bien sûr pas restrictif de supposer que le point en question est 0, quitte à changer la valeur de  $t$ )

$$H = \left\{ x \in \mathbb{R}^d \mid \sum_{i=1}^d a_i x_i + t = 0 \right\}$$

pour certains  $a_i$  non tous nuls. L'équation ci-dessus est appelée l'*équation de l'hyperplan*. Le lecteur démontrera à propos des équations d'hyperplans que :

**Théorème 1.16** Si  $y_1, \dots, y_d$  sont  $d$  points affinement indépendants de  $\mathbb{R}^d$  et si  $y_{ij}$  désigne la  $j$ -ième coordonnée de  $y_i$  (avec  $1 \leq i, j \leq d$ ) alors l'hyperplan engendré par les  $y_i$  s'écrit

$$\det \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \\ 1 & y_{11} & y_{12} & \dots & y_{1d} \\ 1 & y_{21} & y_{22} & \dots & y_{2d} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & y_{d1} & y_{d2} & \dots & y_{dd} \end{bmatrix} = 0.$$

où *det* désigne le déterminant.

On démontre également :

**Proposition 1.17** Une application  $f : \mathbb{R}^d \rightarrow \mathbb{R}^e$  est une application affine si et seulement si pour toute combinaison affine  $\sum_{i=1}^n a_i x_i$  de  $\mathbb{R}^d$  on a  $f(\sum_{i=1}^n a_i x_i) = \sum_{i=1}^n a_i f(x_i)$ .

Ceci établit un liens avec la définition d'une application linéaire dans le cas vectoriel.

Le résultat suivant n'est pas étonnant :

**Proposition 1.18** *Soient  $\{f_1, \dots, f_{d+1}\}$  et  $\{g_1, \dots, g_{d+1}\}$  deux repère affins de  $\mathbb{R}^d$ . Alors il existe une et une seule permutation affine  $f$  telle que  $f(f_i) = g_i$ .*

**Démonstration** : Nous savons donc que  $\{f_1 - f_{d+1}, \dots, f_d - f_{d+1}\}$  et  $\{g_1 - g_{d+1}, \dots, g_d - g_{d+1}\}$  sont deux bases de  $\mathbb{R}^d$ , il existe donc une et une seule permutation linéaire  $g$  telle que

$$g(f_i - f_{d+1}) = g_i - g_{d+1}$$

pour  $i = 1, \dots, d$ . Mais alors en posant  $t = g_{d+1} - g(f_{d+1})$  et  $f = f_t \circ g$  ou  $f_t$  désigne la translation de vecteur  $t$  on a bien

$$f(f_i) = (g(f_{d+1}) + g_i - g_{d+1}) + g_{d+1} - g(f_{d+1}) = g_i$$

pour  $1 \leq i \leq d$  et  $f(f_{d+1}) = g_{d+1}$ . On vérifie aisément que l'unicité de  $f$  résulte de l'unicité de  $g$ , et que  $f$  est une permutation car  $g$  l'est.  $\square$

Introduisons maintenant la notion fondamentale de demi-espace. Soit  $H = g^{-1}(u)$  un hyperplan affín, avec  $g$  une forme affine non triviale et  $u \in \mathbb{R}$ . On définit alors les deux *demi-espaces strictes* (ou *demi-espaces ouverts*) par

$$H^< = \{x \in \mathbb{R}^d \mid g(x) < u\}$$

et

$$H^> = \{x \in \mathbb{R}^d \mid g(x) > u\}.$$

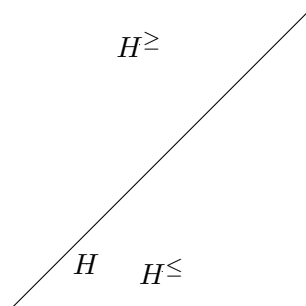
Les *demi-espaces larges* (ou *demi-espaces fermés*)  $H^{\leq}$  et  $H^{\geq}$  sont obtenus en remplaçant les inégalités strictes dans les définitions par des inégalités larges. On a bien sûr

$$H^{\leq} = H^< \cup H$$

ainsi que

$$H^{\geq} = H^> \cup H.$$

Dans le cas où  $d = 1$ , les demi-espaces sont les demi-droites précédemment définies. Dans le cas où  $d = 2$ , les demi-espaces sont les demi-plans.

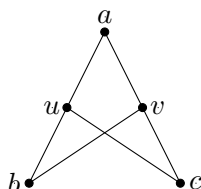




Signalons pour terminer deux propriétés des segments de droite.

**Proposition 1.19** Soient  $a, b, c$  trois points non alignés de  $\mathbb{R}^d$ . Soient  $u \in ]a, b[$  et  $v \in ]a, c[$ . Alors

$$]b, v[ \cap ]c, u[ \neq \emptyset.$$



**Proposition 1.20 (Propriété de Pasch-Peano)** Soient  $a, b, c$  trois points de  $\mathbb{R}^d$ . Alors il existe  $r$  tel que  $r \in ]a, c[$  et  $q \in ]b, r[$ .

### 1.3 Produit scalaire, norme et métrique

Jusqu'à présent nous n'avons pas fait appel à la structure euclidienne de  $\mathbb{R}^d$ , c'est-à-dire au produit scalaire. Rappelons quel le *produit scalaire usuel* sur  $\mathbb{R}^d$  est défini par

$$u.v = \sum_{i=1}^d u_i v_i \in \mathbb{R}$$

pour deux vecteurs  $u = (u_1, \dots, u_d)$  et  $v = (v_1, \dots, v_d)$  de  $\mathbb{R}^d$ , et vérifie

1.  $u.v = v.u$ ,
2.  $(u + v).w = u.w + v.w$ ,
3.  $(\lambda u).v = \lambda(u.v)$ ,
4.  $u.u \geq 0$ ,
5.  $u.u = 0$  si et seulement si  $u = 0$

pour tous  $u, v \in \mathbb{R}^d$  et  $\lambda \in \mathbb{R}$ . La *norme* associée au produit scalaire est l'application

$$\| \cdot \| : \mathbb{R}^d \rightarrow \mathbb{R} : x \rightarrow \sqrt{x.x}$$

qui elle vérifie

1.  $\| x \| = 0$  si et seulement si  $x = 0$ ,
2.  $\| \lambda x \| = |\lambda| \cdot \| x \|$ ,
3.  $\| x + y \| \leq \| x \| + \| y \|$

pour tous  $x, y \in \mathbb{R}^d$  et  $\lambda \in \mathbb{R}$ . Enfin, la *métrique* déduite de  $\| \cdot \|$  est

$$d : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R} : (x, y) \rightarrow \| x - y \|^2$$

et les propriétés habituelles d'une métrique sont vérifiées à savoir

1.  $d(x, y) = 0$  si et seulement si  $x = y$ ,
  2.  $d(x, y) = d(y, x)$ ,
  3.  $d(x, y) \leq d(x, z) + d(z, y)$
- pour tous  $x, y, z \in \mathbb{R}^d$ .

On vérifie que pour tous  $u, v \in \mathbb{R}^d$  on a

$$|\|u\| - \|v\|| \leq \|u - v\|.$$

Dans le membre de gauche figure la valeur absolue de la différence de deux normes.

Les vecteurs  $u \in \mathbb{R}^d$  de norme égal à 1 sont appelés vecteurs *unitaires*. Remarquons que l'équation d'un hyperplan affiné  $H$  de  $\mathbb{R}^d$  peut toujours s'écrire  $x \cdot u = t$  pour un certain vecteur  $u \in \mathbb{R}^d$  unitaire et un certain  $t \in \mathbb{R}$ .

Nous allons maintenant rappeler rapidement les définitions et propriétés de base de la topologie usuelle de  $\mathbb{R}^d$ , c'est-à-dire celle déduite de l'espace métrique  $(X, d)$ .

## 1.4 Topologie

La *boule ouverte* de centre  $x \in \mathbb{R}^d$  et de rayon  $r > 0$  sera notée  $B(x, r)$ . Ainsi

$$B(x, r) = \{y \in \mathbb{R}^d \mid \|x - y\| < r\}.$$

La boule fermée associée sera notée  $\overline{B}(x, r)$ . Lorsque  $d = 2$ , les boules correspondent aux segments de droite (ouverts ou fermés selon le cas).

Pour un ensemble  $A \subseteq \mathbb{R}^d$  quelconque, nous dirons qu'un point  $x \in \mathbb{R}^d$  est *intérieur* à  $A$  s'il existe un  $\epsilon > 0$  tel que  $B(x, \epsilon) \subseteq A$ . L'*intérieur* de  $A$ , noté  $\text{int}(A)$ , est l'ensemble constitué des points intérieurs à  $A$ . Nous dirons que  $x$  est *adhérent* à  $A$  si pour tout  $\epsilon > 0$ , l'intersection  $B(x, \epsilon) \cap A$  est non vide (ou de façon équivalente si tout ouvert contenant  $x$  contient au moins un point de  $A$ ). L'ensemble des points adhérents à  $A$ , noté  $\text{adh}(A)$  forment l'*adhérence* de  $A$  (qu'on appelle parfois la *fermeture* de  $A$ ). On démontre qu'un point  $a \in \mathbb{R}^d$  est dans l'adhérence de  $A$  si et seulement si il existe une suite d'éléments de  $A$  qui converge vers  $a$ .

Voici quelques propriétés élémentaires de l'intérieur et de l'adhérence d'un ensemble :

**Proposition 1.21** *Soient  $A, B$  des sous-ensembles de  $\mathbb{R}^d$ . Alors*

1. *Si  $A \subseteq B$  alors  $\text{int}(A) \subseteq \text{int}(B)$  et  $\text{adh}(A) \subseteq \text{adh}(B)$ ,*
2.  *$\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$ ,*
3.  *$\text{adh}(A \cup B) = \text{adh}(A) \cup \text{adh}(B)$ .*

Un ensemble  $O \subseteq \mathbb{R}^d$  est *ouvert* si  $\text{int}(O) = O$ . Il est *fermé* si  $\text{adh}(O) = O$ . On démontre qu'un ensemble est ouvert si et seulement si son complémentaire est fermé. On démontre aussi qu'une réunion quelconque d'ensembles ouverts est toujours un ensemble ouvert, une intersection finie d'ouverts est toujours un ouvert. Une intersection quelconque de fermés est toujours un fermé, une réunion finie de fermés est un fermé. De plus,  $\text{int}(O)$  est le plus grand ouvert inclus dans  $O$ . C'est la réunion de tous les ouverts inclus à  $O$ . De même  $\text{adh}(O)$  est l'intersection de tous les fermés contenant  $O$ . C'est le plus petit fermé contenant  $O$ . Sans surprise, on constate que l'application envoyant un ensemble sur sa fermeture est un opérateur de fermeture sur  $\mathcal{P}(\mathbb{R}^d, \subseteq)$ . Notons que ceci n'est pas vrai pour l'application envoyant un ensemble sur son intérieur (on a bien sûr  $\text{int}(O) \subseteq O$ , l'inclusion étant stricte en général).

Translater et étirer un ensemble ouvert redonne un ensemble ouvert, en effet

**Lemme 1.22** *Si  $U$  est un ouvert de  $\mathbb{R}^d$  alors il en va de même pour  $t + \lambda U$  avec  $\lambda$  un réel non nul et  $t \in \mathbb{R}^d$ .*

**Démonstration** : Si  $u_0 \in U$  alors il existe un  $\epsilon > 0$  tel que  $B(u_0, \epsilon) \subseteq U$ . Montrons alors que  $B(t + \lambda u_0, |\lambda|\epsilon) \subseteq t + \lambda U$ .

Soit  $s$  un élément de  $\mathbb{R}^d$  tel que  $\|t + \lambda u_0 - s\| \leq |\lambda|\epsilon$  c'est à dire tel que

$$|\lambda| \left\| u_0 - \frac{s-t}{\lambda} \right\| \leq |\lambda|\epsilon.$$

Ainsi  $\frac{s-t}{\lambda} \in U$  vu les hypothèses sur  $u_0$ , ce qui termine la preuve.  $\square$

Il est possible de démontrer que

**Lemme 1.23** *Tous les sous-espaces affins de  $\mathbb{R}^d$  sont fermés. En particulier, les droites, les plans, les hyperplans sont fermés.*

Il en résulte que pour tout  $A$  on a  $\text{adh}(A) \subseteq \text{Aff}(A)$ .

La *frontière* de  $A$  (ou le *bord* de  $A$ ), noté  $\partial A$ , est l'ensemble  $\text{adh}(A) \setminus \text{int}(A)$ .

Pour un sous-ensemble  $S \subseteq \mathbb{R}^d$  et un ouvert  $O$  de  $\mathbb{R}^d$ , la *trace* de  $O$  sur  $S$  est l'intersection  $S \cap O$ . On définit de même la trace d'un fermé  $F$  de  $\mathbb{R}^d$  sur  $S$ .

Une application  $f : A \rightarrow \mathbb{R}^e$  avec  $A \subseteq \mathbb{R}^d$  est *continue* si pour tout ouvert  $U \subseteq \mathbb{R}^e$ ,  $f^{-1}(U)$  est la trace d'un ouvert de  $\mathbb{R}^d$  sur  $A$ . En particulier si  $A = \mathbb{R}^d$ , une fonction est continue si et seulement si l'image réciproque d'un ouvert de  $\mathbb{R}^e$  est un ouvert de  $\mathbb{R}^d$ . Une bijection  $f : \mathbb{R}^d \rightarrow \mathbb{R}^e$  telle que  $f$  et  $f^{-1}$  sont continues est un *homéomorphisme*. On démontre qu'un tel homéomorphisme n'existe en fait que lorsque  $e = d$ . On a

**Lemme 1.24** *Toute application affine (donc en particulier toute translation) de  $\mathbb{R}^d$  dans  $\mathbb{R}^e$  est continue. Si  $f$  est bijective, alors  $e = d$  et  $f$  est un homéomorphisme.*

L'ensemble  $A \subseteq \mathbb{R}^d$  est dit *compact* si de tout recouvrement de  $A$  des ouverts  $O_i, i \in I$  on peut extraire un sous-recouvrement fini. Il est équivalente d'exiger que toute suite d'éléments de  $A$  admet une sous-suite qui converge vers un élément de  $A$ . Un ensemble de  $\mathbb{R}^d$  est compact si et seulement si il est fermé et borné (un sous-ensemble  $A \subseteq \mathbb{R}^d$  étant dit *borné* s'il existe une constante  $M \in \mathbb{R}$  telle que  $\|x\| \leq M$  pour tout  $x \in A$ ). Si  $A$  est compact et si  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  est continue, alors  $f(A)$  est toujours compact. Si  $A$  est un compact de  $\mathbb{R}^d$  et  $B$  est un compact de  $\mathbb{R}^e$  alors le produit cartésien  $A \times B$  est un compact de  $\mathbb{R}^{d+e}$ . Toute fonction continue définie sur un compact atteint un maximum et un minimum sur ce compact (ceci est un résultat classique de calcul différentiel et intégral).

Contrairement à ce qui précède, la notion d'intérieur relatif sera peut-être nouvelle pour le lecteur. L'*intérieur relatif* d'un ensemble  $A \subseteq \mathbb{R}^d$ , noté  $intrel(A)$  (parfois  $relint(A)$ ) est défini comme l'intérieur de  $A$  calculé dans  $Aff(A)$  plutôt que dans  $\mathbb{R}^d$ . Rappelons que  $Aff(A)$  est isomorphe à  $\mathbb{R}^{dim(A)}$ . De plus cet isomorphisme un homéomorphisme par le lemme précédant, et préserve donc les ensembles ouverts. Il est en fait possible de montrer que les ouverts et les fermés (respectivement) de  $Aff(A)$  ne sont rien d'autre que les traces des ouverts et des fermés (respectivement) de  $\mathbb{R}^d$  sur  $Aff(A)$ . On a donc

$$intrel(A) = \{x \in A \mid \exists \epsilon > 0, B(x, \epsilon) \cap Aff(A) \subseteq A\}.$$

Si  $Aff(A) = \mathbb{R}^d$  alors bien sûr  $int(A) = intrel(A)$ . Sinon, les notions sont distinctes. A titre d'exemple, le lecteur remarquera que l'intérieur relatif d'un segment fermé  $[a, b] \subseteq \mathbb{R}^3$  est le segment ouvert  $]a, b[$  correspondant tandis que l'intérieur de  $A$  est vide. Remarquons en passant que les segments ouverts, les demi-droites ouverts et les demi-espaces ouverts ne sont ouverts que dans leur enveloppe affine. Une terminologie plus adaptée aurait donc été "segments relativement ouverts", "demi-droite relativement ouverte", etc. mais nous ne l'utiliserons pas en raison de sa relative lourdeur.

Illustrons encore une différence de l'intérieur relatif et de l'intérieur usuel. Nous avons déjà mentionné le fait que si  $A \subseteq B$ , alors  $int(A) \subseteq int(B)$ . Cette propriété est cependant fautive dans le cas de l'intérieur relatif (c'est une conséquence du fait que l'enveloppe affine de  $A$  peut être différente de l'enveloppe affine de  $B$ ). Considérons par exemple un segment de droite  $[a, b]$  pour  $a, b \in \mathbb{R}^d$ . Alors  $intrel([a, b]) = ]a, b[$  mais  $intrel(\{a\}) = \{a\}$ . En règle générale nous avons de plus

$$intrel(A \cap B) \neq intrel(A) \cap intrel(B)$$

pour deux sous-ensembles  $A$  et  $B$  de  $\mathbb{R}^d$ .

Notons enfin que l'"adhérence relative" de  $A \subseteq \mathbb{R}^d$  n'est rien d'autre que l'adhérence de  $A$ . En effet si  $\{T_i \mid i \in I\}$  désigne l'ensemble des fermés de  $\mathbb{R}^d$

contenants  $A$  on a (comme  $adh(A) \subseteq Aff(A)$ ) que

$$adh(A) = \bigcap_{i \in I} (T_i \cap Aff(A))$$

et de même pour l'adhérence relative (les fermés de  $Aff(A)$  étant exactement l'intersection des fermés de  $\mathbb{R}^d$  avec  $Aff(A)$ ).

La *frontière relative* de  $A$  (ou le *bord relatif* de  $A$ ) est l'ensemble  $adh(A) \setminus relint(A)$ .

Nous verrons que la notion d'intérieur relatif sera utile dans notre étude des ensembles convexes.

## Chapitre 2

# Ensembles convexes

Après un chapitre de rappels sur les propriétés de base de  $\mathbb{R}^d$ , nous commençons l'étude des ensembles convexes proprement dit. Bien que nous travaillerons uniquement dans  $\mathbb{R}^d$  (espace vectoriel de dimension finie sur le corps des réels), une partie des résultats et des définitions de ce chapitre s'étend aux espaces vectoriels de dimension infinie. Ces résultats sont généralement étudiés dans des cours d'analyse fonctionnelle plutôt que dans des cours de géométrie. A titre d'exemple, mentionnons que la séparation des ensembles convexes (voir la section appropriée) s'étudie aussi en dimension infinie, une version du célèbre *théorème de Hahn-Banach* n'est rien d'autre qu'une généralisation du théorème de séparation pour des convexes disjoints de  $\mathbb{R}^d$ .

### 2.1 Définition et propriétés de base

Un ensemble  $C \subseteq \mathbb{R}^d$  est dit *convexe* si pour tous  $a, b \in C$ , nous avons  $[a, b] \subseteq C$ . Cette définition est illustrée sur le dessin ci-dessous (tout segment passant par deux points de  $C$  doit être inclus dans  $C$ ).



La figure de gauche est convexe, contrairement à celle de droite. Pour un convexe  $C$ , la *dimension* de  $C$  est par définition la dimension de son enveloppe affine  $\dim(\text{Aff}(C))$ , on la note  $\dim(C)$ . Voici des exemples élémentaires d'ensembles convexes :

**Proposition 2.1** *Les ensembles suivants sont des convexes de  $\mathbb{R}^d$*

1. *Tout sous-espace affine (donc en particulier  $\emptyset$ , les singletons, les droites, les plans, les hyperplans et  $\mathbb{R}^d$  lui même),*

2. Tout segment de droite  $[a, b]$ ,
3. Tout demi-espace  $H^{\leq}$ ,
4. Toute boule ouverte  $B(x, r)$ .

**Démonstration** : La vérification des points 1) et 2) est triviale. Pour vérifier 3), soit  $H = g^{-1}(u)$  l'hyperplan associé au demi-espace. Soit  $p$  et  $q$  deux points de  $H^{\leq}$ . Alors pour tout  $\lambda$  avec  $0 \leq \lambda \leq 1$  on a

$$g(\lambda p + (1 - \lambda)q) = \lambda g(p) + (1 - \lambda)g(q) \leq \lambda u + (1 - \lambda)u = u$$

et donc  $[a, b] \subseteq H^{\leq}$ .

Soient enfin  $p$  et  $q$  deux points de la boule  $B(x, r)$ . Alors pour tout  $0 \leq \lambda \leq 1$  il vient  $x = \lambda p + (1 - \lambda)q$  de sorte que pour tout point  $s = \lambda p + (1 - \lambda)q$  du segment  $[p, q]$  il vient

$$\|x - s\| \leq \|\lambda(x - p)\| + \|(1 - \lambda)(x - q)\|$$

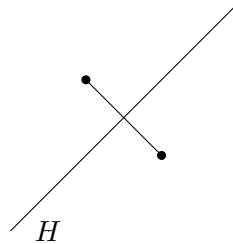
et le membre de droite vaut

$$\lambda \|x - p\| + (1 - \lambda) \|x - q\| \leq \lambda r + (1 - \lambda)r = r$$

donc  $s \in B(x, r)$  et la boule est bien convexe  $\square$ .

Un ensemble convexe de  $\mathbb{R}^d$  qui est à la fois compact et non vide est parfois appelé *corps convexe*.

L'intersection d'un hyperplan avec deux points se trouvant chacun dans un des deux demi-espaces est toujours un singleton :



**Proposition 2.2** Soit  $p \in H^<$  et  $q \in H^>$ . Alors  $[p, q] \cap H$  est un singleton.

**Démonstration** Soit  $H = g^{-1}(u)$  l'hyperplan. Soit  $s = \lambda p + (1 - \lambda)q$  un point du segment. Alors

$$g(s) = \lambda g(p) + (1 - \lambda)g(q) = g(q) + \lambda(g(p) - g(q)).$$

De plus  $g(p) - g(q) \neq 0$  vu que  $g(p) > u$  et  $g(q) < u$ . Pour

$$\lambda = \frac{(u - g(q))}{g(p) - g(q)}$$

on a bien  $0 < \lambda < 1$  et  $g(s) = u$ , donc  $s \in H \cap [p, q]$ . Un tel  $\lambda$  est clairement unique.  $\square$

Voici quelques propriétés élémentaires :

**Proposition 2.3** *Soient  $A$  et  $B$  deux ensembles convexes de  $\mathbb{R}^d$  et  $t \in \mathbb{R}$ . Alors les ensembles suivants sont aussi convexes*

1.  $A \cap B$ ,
2.  $A + B$ ,
3.  $tA = \{ta \mid a \in A\}$ ,
4.  $\text{int}(A)$  et  $\text{adh}(A)$ ,
5.  $f(A) = \{f(a) \mid a \in A\}$  où  $f$  est une application affine.

**Démonstration** : Le point 1) est évident. Soient  $a_1 + b_1 \in A + B$  et  $a_2 + b_2 \in A + B$ . Pour  $0 \leq \lambda \leq 1$  il vient

$$\lambda(a_1 + b_1) + (1 - \lambda)(a_2 + b_2) = (\lambda a_1 + (1 - \lambda)a_2) + (\lambda b_1 + (1 - \lambda)b_2)$$

et  $\lambda a_1 + (1 - \lambda)a_2$  est dans  $A$  car  $A$  est convexe, de même  $\lambda b_1 + (1 - \lambda)b_2$  est dans  $B$ . Ainsi  $A + B$  est bien convexe. Le point 3) se démontre de la même façon et le point 5) résulte de la proposition 1.17. Démontrons maintenant le point 4).

Soient  $x$  et  $y$  deux points intérieurs de  $A$ . Il existe donc  $\epsilon > 0$  tel quel  $B(x, \epsilon) \subseteq A$  et  $B(y, \epsilon) \subseteq A$ . Ainsi, pour tous  $u, v \in B(0, \epsilon)$  on a  $x + u \in A$  et  $y + v \in A$ . Comme  $A$  est convexe, on a pour tout  $0 \leq \lambda \leq 1$  que  $\lambda(x + u) + (1 - \lambda)(y + v) \in A$ . En particulier si  $v = u$ , on a que  $\lambda x + (1 - \lambda)y + u \in A$  quel que soit  $u \in B(0, \epsilon)$ . Ceci implique que pour tout  $z \in B(\lambda x + (1 - \lambda)y, \epsilon)$  on a

$$z = \lambda x + (1 - \lambda)y + (z - \lambda x + (1 - \lambda)y)$$

avec

$$(z - \lambda x + (1 - \lambda)y) \in B(0, \epsilon)$$

donc  $z \in A$ . Ainsi  $\lambda x + (1 - \lambda)y$  est bien un point intérieur.

Soient enfin  $w$  et  $t$  deux points dans l'adhérence de  $A$ . Il existe ainsi deux suites  $w_i$  et  $t_i$  d'éléments de  $A$  qui tendent vers  $w$  et  $t$  respectivement lorsque  $i \rightarrow \infty$ . Or la suite

$$s_i = \lambda w_i + (1 - \lambda)t_i$$

est une suite de  $A$  quel que soit  $0 \leq \lambda \leq 1$  car  $A$  est convexe, et comme cette suite tend vers  $\lambda w + (1 - \lambda)t$ , on a bien  $\lambda w + (1 - \lambda)t \in \text{adh}(A)$ .  $\square$

Le point 1) de la proposition précédente se généralise bien sûr à une collection quelconque de convexes de  $\mathbb{R}^d$ , toute intersection de convexes est encore convexe. Ceci n'est naturellement plus vrai pour une réunion, une réunion de deux convexes n'est généralement pas convexe.



## 2.2 Enveloppe convexe

Pour un ensemble quelconque  $X \subseteq \mathbb{R}^d$ , l'intersection de tous les convexes contenant  $X$  est appelée l'*enveloppe convexe* de  $X$ . C'est le plus petit ensemble convexe contenant  $X$ . On la note  $\text{Conv}(X)$ . De même que pour les enveloppes linéaires et affines, ainsi que la fermeture (au sens topologique), l'enveloppe convexe est un opérateur de fermeture sur l'ensemble des parties de  $\mathbb{R}^d$ . Bien sûr, si  $A$  est convexe on a  $\text{Conv}(A) = A$ . Dans les autres cas, le lecteur illustrera ce concept dans le cas  $d = 2$  ou  $d = 3$  au moyen de dessins.

Si  $x_1, \dots, x_n \in \mathbb{R}^d$  et si  $\lambda_1, \dots, \lambda_n$  sont des réels positifs tels que  $\sum_{i=1}^n \lambda_i = 1$  alors

$$x = \sum_{i=1}^n \lambda_i x_i$$

est appelée *combinaison convexe* des  $x_i$ . De même que dans le cas affiné et le cas vectoriel on a alors :

**Proposition 2.4** *L'enveloppe convexe de  $X \subseteq \mathbb{R}^d$  correspond à l'ensemble des combinaisons convexes des éléments de  $X$ , donc*

$$\text{Conv}(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid x_i \in X, n \in \mathbb{N}_0, \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0 \right\}.$$

**Démonstration** : Celle-ci est en tout point similaire à celle concernant l'enveloppe affine, voyez le chapitre 1.  $\square$

Le fait qu'on puisse se limiter à des combinaisons convexes ayant  $d + 1$  points uniquement est dû à un mathématicien du nom de Constantin Carathéodory (1873-1950) :

**Théorème 2.5 (Théorème de Carathéodory)** *Soit  $X \subseteq \mathbb{R}^d$ . Alors*

$$\text{Conv}(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid x_i \in X, n \in \mathbb{N}_0, 1 \leq n \leq d + 1, \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0 \right\}.$$

**Démonstration** : Par la proposition précédente, il suffit de démontrer que le membre de gauche est inclus au membre de droite. Soit donc  $q = \sum_{i=1}^n \lambda_i x_i$  avec  $n \geq d + 2$  (sinon il n'y a rien à prouver),  $\sum_{i=1}^n \lambda_i = 1$ ,  $\lambda_i > 0$  un élément de  $\text{conv}(X)$ . Ainsi les points  $x_1, \dots, x_n$  sont linéairement dépendants. Il existe ainsi  $\mu_1, \dots, \mu_n$  des réels non tous nuls vérifiant  $\sum_{i=1}^n \mu_i = 0$  et

$$\sum_{i=1}^n \mu_i x_i = 0.$$

Pour chaque  $\mu_i$  non nul on peut donc écrire

$$x_i = -\frac{1}{\mu_i} \sum_{j=1, j \neq i}^n \mu_j p_j.$$

En reportant cela dans l'expression de  $q$  on obtient

$$q = \sum_{j=1, j \neq i}^n (\lambda_j - \lambda_i \mu_i^{-1} \mu_j) p_j.$$

La somme des coefficients de cette combinaison vaut bien 1 car

$$\sum_{j=1, j \neq i}^n (\lambda_j - \lambda_i \mu_i^{-1} \mu_j) = \sum_{j=1, j \neq i}^n \lambda_j - \frac{\lambda_i}{\mu_i} \sum_{j=1, j \neq i}^n \mu_j$$

et  $\sum_{j=1, j \neq i}^n \mu_j = -\mu_i$ . Il nous reste à nous assurer que chaque coefficient est bien positif ou nul de façon à exprimer  $q$  comme une combinaison convexe de  $n - 1$  points et à terminer la preuve. Il nous faut donc prouver l'existence d'un  $i$  tel que pour tout  $j$  on ait  $\lambda_j - \lambda_i \mu_i^{-1} \mu_j \geq 0$  ce qui peut se réécrire

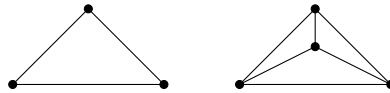
$$\lambda_j \mu_j^{-1} - \lambda_i \mu_i^{-1} \geq 0$$

si  $\mu_j > 0$  ou bien

$$\lambda_j \geq \lambda_i \mu_i^{-1} \mu_j$$

si  $\mu_i \leq 0$ . Parmi les  $\mu_i > 0$  il suffit alors de choisir l'indice  $i$  minimisant  $\lambda_i \mu_i^{-1}$  et les deux conditions seront bien satisfaites.  $\square$

Définissons maintenant un *simplexe* de  $\mathbb{R}^d$  comme l'enveloppe convexe d'un repère affín de  $\mathbb{R}^d$ . Le théorème de Carathéodory affirme donc que tout convexe peut s'écrire comme une réunion de simplexes. Les simplexes constituent une généralisation des triangles (pleins) dans le plan. Lorsque  $d = 1$ , il s'agit de segments de droites, lorsque  $d = 2$  de triangles pleins et lorsque  $d = 3$  de tétraèdres pleins.



De façon plus générale, nous définissons un *polytope* comme l'enveloppe convexe d'un nombre fini de points. Les polytopes généralisent les polygones du plan et les polyèdres de l'espace. Leur étude plus approfondie fera l'objet d'un chapitre ultérieur.

Voici un corollaire presque immédiat du théorème de Carathéodory :

**Corollaire 2.6** Soit  $A \subseteq \mathbb{R}^d$  avec  $\dim(\text{Aff}(A)) = m \leq d + 1$ . Alors

$$\text{Conv}(A) = \left\{ \sum_{i=1}^n \lambda_i a_i \mid a_i \in A, n \in \mathbb{N}_0, 1 \leq n \leq m + 1, \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0 \right\}$$

**Démonstration** : Si  $a = \sum_{i=1}^n \lambda_i a_i$  est un élément de  $\text{Conv}(A)$  avec  $n > m + 1$  alors par hypothèse, les  $a_i$  sont affinement dépendants et par le même argument que dans la preuve du théorème de Carathéodory, il est possible d'écrire  $a$  comme une combinaison convexe de  $n - 1$  éléments.  $\square$

Ceci nous permet de prouver que :

**Corollaire 2.7** Soit  $A \subseteq \mathbb{R}^d$  un compact de  $\mathbb{R}^d$ . Alors  $\text{Conv}(A)$  est aussi compact.

**Démonstration** : Soit  $n = \dim(\text{Aff}(A))$  et définissons l'ensemble

$$S = \left\{ (\lambda_1, \dots, \lambda_{n+1}) \mid \sum_{i=1}^{n+1} \lambda_i = 1, \lambda_i \geq 0 \right\} \subseteq \mathbb{R}^{n+1}.$$

On prouve aisément que  $S$  est compact (à savoir borné et fermé). Mais le produit cartésien  $A^{n+1}$  est aussi compact. Si on définit l'application  $f : A^{n+1} \times S \rightarrow \mathbb{R}^d$  par

$$f((a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1})) = \sum_{i=1}^{n+1} \lambda_i a_i$$

alors on a par ce qui a été démontré ci-dessus que  $\text{Conv}(A) = f(A^{n+1} \times S)$ . Comme  $f$  est continue (vu que  $f$  est une fonction polynomiale) et que  $A^{n+1} \times S$  est compact (produit cartésien de compacts) on a bien que  $\text{Conv}(A)$  est compact.  $\square$

Et en particulier, un ensemble fini de points étant toujours compact :

**Corollaire 2.8** Tout polytope de  $\mathbb{R}^d$  est compact.

Terminons cette section par quelques propriétés supplémentaires de l'enveloppe convexe. Nous laissons au lecteur le soin de prouver le lien suivant entre enveloppe affine et enveloppe convexe :

**Proposition 2.9** Si  $X$  est un sous-ensemble de  $\mathbb{R}^d$  alors

$$\text{Conv}(\text{Aff}(X)) = \text{Aff}(X) = \text{Aff}(\text{Conv}(X)).$$

**Proposition 2.10** Soient  $A, B \subseteq \mathbb{R}^d$ , alors

$$1. \text{Conv}(A \cup B) = \text{Conv}(\text{Conv}(A) \cup \text{Conv}(B)),$$

2.  $Conv(A + B) = Conv(A) + Conv(B)$ .

**Démonstration** : Soit  $D$  le membre de droite du point 1). Alors  $A \subseteq D$  et  $B \subseteq D$  c'est-à-dire  $A \cup B \subseteq D$ . Comme  $D$  est convexe, on a bien

$$Conv(A \cup B) \subseteq D.$$

De même comme  $A$  et  $B$  sont contenus dans  $Conv(A \cup B)$  on a

$$Conv(A) \cup Conv(B) \subseteq Conv(A \cup B)$$

et  $Conv(A \cup B)$  étant convexe on a bien  $D \subseteq Conv(A \cup B)$ .

Soit  $x = \sum_{i=1}^n \lambda_i(a_i + b_i)$  un élément de  $Conv(A + B)$ . Alors

$$x = \sum_{i=1}^n \lambda_i a_i + \sum_{i=1}^n \lambda_i b_i$$

donc  $x \in Conv(A) + Conv(B)$ . Soit maintenant

$$y = \sum_{i=1}^n \lambda_i a_i + \sum_{i=1}^m \gamma_i b_i$$

un élément de  $Conv(A) + Conv(B)$ . Il n'est pas restrictif de supposer  $m = n$  et comme

$$\sum_{i=1}^n \gamma_i = \sum_{i=1}^n \lambda_i = 1$$

on peut réécrire  $y$  ainsi

$$\sum_{i=1, j=1}^n \lambda_i \gamma_j (a_i + b_j)$$

ce qui donne bien un combinaison convexe d'éléments de  $A + B$ .  $\square$

En particulier nous avons :

**Corollaire 2.11** *Si  $P$  est un polytope de  $\mathbb{R}^d$ , alors tout translaté  $x + P$  est aussi un polytope. De plus si  $Q$  est un autre polytope de  $\mathbb{R}^d$  alors  $P + Q$  est toujours un polytope.*

**Démonstration** : La proposition précédente nous apprend en particulier que pour  $P = Conv(\{x_1, \dots, x_n\})$  et  $Q = Conv(\{y_1, \dots, y_m\})$  deux polytopes de  $\mathbb{R}^d$  et pour  $x \in \mathbb{R}^d$  on a

$$x + P = Conv(\{x_1 + x, \dots, x_n + x\})$$

ainsi que

$$P + Q = \text{Conv}(\{x_i + y_j \mid i \in \{1, \dots, n\}, j \in \{1, \dots, m\}\})$$

ce qui termine la preuve.  $\square$

Nous avons plus généralement :

**Proposition 2.12** *Si  $f : \mathbb{R}^d \rightarrow \mathbb{R}^e$  est une application affine et si  $P \subseteq \mathbb{R}^d$  est un polytope, alors  $f(P)$  est encore un polytope.*

Cette proposition étant une conséquence directe de :

**Proposition 2.13** *Si  $f : \mathbb{R}^d \rightarrow \mathbb{R}^e$  est une application affine et  $A \subseteq \mathbb{R}^d$  alors*

$$f(\text{Conv}(A)) = \text{Conv}(f(A)).$$

**Démonstration** : Ceci résulte de la proposition 1.17.  $\square$

Voici finalement une façon de calculer l'enveloppe convexe de la réunion de deux convexes, il suffit de prendre la réunion des segments de droite dont l'une extrémité est dans le premier ensemble, l'autre dans le second.

**Proposition 2.14** *Soit  $C$  et  $D$  deux convexes de  $\mathbb{R}^d$ . Alors*

$$\text{Conv}(C \cup D) = \bigcup_{c \in C, d \in D} [c, d].$$

**Démonstration** : Soit  $E$  le membre de droite de l'égalité. Alors  $E$  est contenu dans  $\text{Conv}(C \cup D)$  car ce dernier ensemble est convexe. Comme  $E$  contient  $C$  et  $D$ , il suffit de donc de démontrer que  $E$  est convexe pour terminer la preuve. Soit  $x = \alpha_1 c_1 + \beta_1 d_1$  et  $y = \alpha_2 c_2 + \beta_2 d_2$  deux éléments de  $E$ . On a donc  $\alpha_1, \beta_1 \geq 0$  ainsi que  $\alpha_2, \beta_2 \geq 0$  et de plus  $\alpha_1 + \beta_1 = 1$ ,  $\alpha_2 + \beta_2 = 1$ . Soit

$$q = \lambda_1(\alpha_1 c_1 + \beta_1 d_1) + \gamma_1(\alpha_2 c_2 + \beta_2 d_2)$$

un point du segment  $[x, y]$  et montrons que  $q \in E$ . Supposons dans un premier temps que  $\lambda_1 \alpha_1 + \gamma_1 \alpha_2$  et  $\lambda_1 \beta_1 + \gamma_1 \beta_2$  sont tous les deux non nuls. Dans ce cas  $q$  s'écrit

$$q = (\lambda_1 \alpha_1 + \gamma_1 \alpha_2) \left( \frac{\lambda_1 \alpha_1}{\lambda_1 \alpha_1 + \gamma_1 \alpha_2} c_1 + \frac{\gamma_1 \alpha_2}{\lambda_1 \alpha_1 + \gamma_1 \alpha_2} c_2 \right) + (\lambda_1 \beta_1 + \gamma_1 \beta_2) \left( \frac{\lambda_1 \beta_1}{\lambda_1 \beta_1 + \gamma_1 \beta_2} d_1 + \frac{\gamma_1 \beta_2}{\lambda_1 \beta_1 + \gamma_1 \beta_2} d_2 \right)$$

et comme  $C$  est convexe on a

$$\frac{\lambda_1 \alpha_1}{\lambda_1 \alpha_1 + \gamma_1 \alpha_2} c_1 + \frac{\gamma_1 \alpha_2}{\lambda_1 \alpha_1 + \gamma_1 \alpha_2} c_2 \in C$$

et de même

$$\frac{\lambda_1\beta_1}{\lambda_1\beta_1 + \gamma_1\beta_2}d_1 + \frac{\gamma_1\beta_2}{\lambda_1\beta_1 + \gamma_1\beta_2}d_2 \in D.$$

Comme de plus

$$\lambda_1\alpha_1 + \gamma_1\alpha_2 + \lambda_1\beta_1 + \gamma_1\beta_2 = 1$$

on a bien  $q \in E$ . Il nous reste à traiter les cas où  $\lambda_1\alpha_1 + \gamma_1\alpha_2$  ou  $\lambda_1\beta_1 + \gamma_1\beta_2$  sont nuls avec  $\gamma_1 = 1 - \lambda_1$ . Si

$$\lambda_1\alpha_1 + (1 - \lambda_1)\alpha_2 = 0$$

comme tous les nombres sont positifs ou nuls il faut obligatoirement  $\lambda_1\alpha_1 = 0$  et  $(1 - \lambda_1)\alpha_2 = 0$ . Ainsi soit  $\lambda_1 = 0$  et  $\alpha_2 = 0$  soit  $\alpha_1 = 0$  et  $\lambda_1 = 1$  soit  $\alpha_1 = 0$  et  $\alpha_2 = 0$ . Dans le premier cas on a  $q = d_2$ , dans le second  $q = c_1$  et dans le troisième  $q = \lambda_1d_1 + \gamma_1d_2 \in D$ . Ainsi  $q$  est trivialement dans  $E$  dans les trois cas. Le cas où  $\lambda_1\beta_1 + \gamma_1\beta_2$  est nul se traite de la même façon.  $\square$

### 2.3 L'intérieur relatif d'un convexe

L'intérieur d'un ensemble non vide (même convexe) peut être vide. En revanche, l'intérieur relatif d'un convexe non vide de  $\mathbb{R}^d$  est toujours non vide (ce qui n'est pas toujours le cas pour un ensemble non convexe de  $\mathbb{R}^d$ , voyez par exemple l'intérieur relatif de deux points distincts dans  $\mathbb{R}^d$  avec  $d \geq 1$ .)

**Théorème 2.15** *Soit  $C \subseteq \mathbb{R}^d$  un ensemble convexe avec  $C \neq \emptyset$ . Alors  $\text{intrel}(C) \neq \emptyset$ .*

La preuve est basée sur le lemme suivant :

**Lemme 2.16** *Soit  $A$  un ensemble de points affinement indépendants de  $\mathbb{R}^d$ . Alors  $\text{intrel}(\text{Conv}(A)) \neq \emptyset$ .*

**Démonstration** : Soit  $A = \{x_1, \dots, x_n\}$  avec  $n \leq d + 1$ . Alors  $A$  forme un repère affine de  $\text{Aff}(A)$  et chaque élément de  $\text{Aff}(A)$  s'écrit donc d'une et une seule façon comme combinaison affine d'éléments de  $A$ . Définissons maintenant la fonction  $f$  par

$$f : \text{Aff}(A) \rightarrow \mathbb{R}^n : \sum_{i=1}^n \lambda_i x_i \rightarrow (\lambda_1, \dots, \lambda_n).$$

Il est alors facile de vérifier que  $f$  est une application affine. Elle est donc continue.

Définissons maintenant pour  $i \in \{1, \dots, n\}$  les sous-ensembles suivants

$$S_i = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i > 0\}.$$

Ce sont des demi-espaces ouverts dans  $\mathbb{R}^n$  et donc,  $f$  étant continue, les ensembles  $f^{-1}(S_i)$  sont aussi ouverts dans  $Aff(A)$ . Il en va donc de même pour l'ensemble

$$\bigcap_{i=1}^n f^{-1}(S_i).$$

Nous avons bien sûr

$$\bigcap_{i=1}^n f^{-1}(S_i) = \left\{ \sum_{i=1}^n \lambda_i x_i \in Aff(A) \mid \lambda_i > 0 \forall i \right\}$$

et cet ensemble est non vide. De plus, comme toute combinaison affine  $\sum_{i=1}^n \lambda_i x_i$  avec  $\lambda_i$  positif pour tout  $i$  est une combinaison convexe, il en résulte que  $\bigcap_{i=1}^n f^{-1}(S_i)$  est aussi un sous-ensemble de  $Conv(A)$ . Cet ensemble est donc ouvert dans  $Aff(A)$ , non vide, et contenu dans  $Conv(A)$  ce qui termine la preuve.  $\square$

**Preuve du théorème 2.15 :** Soit  $n$  la dimension de  $C$  et  $A = \{x_1, \dots, x_{n+1}\}$  un ensemble de points affinement indépendants de cardinal maximum de  $C$ . Comme  $C$  est convexe,  $Conv(A)$  est contenu dans  $C$ . On a de plus  $Aff(A) = Aff(C)$  et  $Conv(A)$  est d'intérieur non vide dans  $Aff(A)$ . Ceci termine la preuve.  $\square$

La proposition suivante nous sera utile :

**Proposition 2.17** *Soit  $C$  un convexe de  $\mathbb{R}^d$ ,  $p \in relint(C)$  et  $q \in adh(C)$  avec  $p \neq q$ . Alors  $[p, q[ \subseteq relint(C)$ .*

**Démonstration :** Soit  $x = (1 - \lambda)p + \lambda q$  avec  $0 < \lambda < 1$  et montrons que  $x \in relint(C)$ . Comme  $p$  est dans l'intérieur relatif de  $C$ , il existe un ensemble ouvert  $U$  dans  $Aff(C)$  tel que  $x_o \in U \subseteq C$ . Considérons maintenant l'ensemble

$$V = \lambda^{-1}x - \lambda^{-1}(1 - \lambda)U.$$

Comme  $U$  est ouvert dans  $Aff(C)$ ,  $V$  aussi et est de plus contenu dans  $Aff(C)$  vu que  $\lambda^{-1} - \lambda^{-1}(1 - \lambda) = 1$  (tous les points de  $V$  sont combinaisons affines d'éléments de  $Aff(C)$ ), notons que  $x$  est bien dans  $Aff(C)$  vu que  $adh(C) \subseteq Aff(C)$ .

Comme  $q$  est dans  $adh(C)$  et que  $q \in V$  (étant donné  $q = \lambda^{-1}x - \lambda^{-1}(1 - \lambda)p$ ), il existe alors un point  $y \in V \cap C$ . Définissons

$$W = (1 - \lambda)U + \lambda y.$$

Cet ensemble est ouvert dans  $Aff(C)$  et est contenu dans  $C$  car  $y \in C$ ,  $U \subseteq C$  et  $C$  est convexe. Mais de plus comme  $y \in V$ , il existe  $z \in U$  tel que

$$y = \lambda^{-1}x - \lambda^{-1}(1 - \lambda)z$$

c'est-à-dire tel que

$$x = (1 - \lambda)z + \lambda y$$

donc  $x \in W$  par définition de  $W$ . On a donc bien trouvé un ouvert  $W$  de  $Aff(C)$  tel que  $x \in W \subseteq C$ .  $\square$

Comme corollaire nous obtenons :

**Corollaire 2.18** *Si  $C$  est un convexe de  $\mathbb{R}^d$ , alors  $relint(C)$  est aussi convexe.*

## 2.4 Le théorème de Radon et le théorème de Helly

Cette section est consacrée à la démonstration de deux théorèmes importants de convexité. L'un est dû à Johann Radon (1887 - 1956) l'autre à Eduard Helly (1884-1943).

**Théorème 2.19 (Théorème de Radon)** *Soit  $X \subseteq \mathbb{R}^d$  un ensemble contenant au moins  $d + 2$  éléments. Alors il existe une partition de  $X$  en deux ensembles  $X = X_1 \cup X_2$  tels que*

$$Conv(X_1) \cap Conv(X_2) \neq \emptyset.$$

**Démonstration** : Si l'ensemble  $X$  comporte plus de  $d + 1$  points, alors il est affinement dépendants. Il existe donc  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  ainsi que  $x_1, \dots, x_n \in X$  tels que  $\sum_{i=1}^n \lambda_i x_i = 0$  et  $\sum_{i=1}^n \lambda_i = 0$ . Sans perte de généralité, on peut supposer que  $\lambda_1, \dots, \lambda_{n_1} > 0$  et  $\lambda_{n_1+1}, \dots, \lambda_n < 0$  pour un certain  $n_1$  avec  $1 \leq n_1 \leq n$ . Ainsi

$$\sum_{i=1}^{n_1} \lambda_i x_i = \sum_{j=n_1+1}^n (-\lambda_j) x_j$$

et

$$\sum_{i=1}^{n_1} \lambda_i = \sum_{j=n_1+1}^n (-\lambda_j) > 0.$$

Ceci implique alors

$$\sum_{i=1}^{n_1} \gamma_i x_i = \sum_{j=n_1+1}^n \beta_j x_j$$

avec

$$\gamma_i = \frac{\lambda_i}{\sum_{k=1}^{n_1} \lambda_k}$$

et

$$\beta_j = \frac{-\lambda_j}{\sum_{k=n_1+1}^n \lambda_k}.$$

La somme des  $\gamma_i$  et des  $\beta_j$  égale 1 de sorte qu'il s'agit de combinaisons convexes. Ainsi, en prenant  $X_1 = \{x_1, \dots, x_{n_1}\}$  et  $X_2 = X \setminus X_1$  on a bien un point commun entre les deux enveloppes convexes.  $\square$

Nous allons maintenant utiliser ce théorème pour prouver :



**Théorème 2.20 (Théorème de Helly)** Soient  $C_1, \dots, C_n \subseteq \mathbb{R}^d$  des convexes avec  $n \geq d + 1$ . Si  $d + 1$  ensembles quelconques choisis parmi les  $C_i$  ont toujours une intersection non vide, alors il en va de même pour l'ensemble de tous les  $C_i$ .

**Démonstration** : Pour un certain  $d$  fixé, on procède par induction sur  $n$ . Le théorème est clair pour  $n = d + 1$ , supposons le vrai pour tout  $k$  avec  $d + 1 \leq k < n$ . Mais alors pour tout  $i \in \{1, \dots, n\}$  nous savons que l'intersection

$$\bigcap_{1 \leq j \leq n, j \neq i} C_j$$

est non vide. Soit  $a_i$  un élément de cette intersection. Par le théorème de Radon, il existe donc une partition de  $\{1, \dots, n\}$  soit  $I_1, I_2$  telle que

$$\text{Conv}(\{a_i \mid i \in I_1\}) \cap \text{Conv}(\{a_i \mid i \in I_2\}) \neq \emptyset.$$

Soit  $x$  un élément de cette intersection, et montrons que  $x$  est contenu dans tous les  $C_i$ . En effet, soit  $i \notin I_1$  et dans ce cas  $a_j \in C_i$  pour tout  $j \in I_1$  vu que  $j \neq i$  et  $a_j \in C_k$  dès que  $j \neq k$ . Dans ce cas on a

$$x \in \text{Conv}(\{a_j \mid j \in I_1\}) \subseteq C_i.$$

De même si  $i \notin I_2$  on a  $a_j \in C_i$  pour tout  $j \in I_2$  et

$$x \in \text{Conv}(\{a_j \mid j \in I_2\}) \subseteq C_i. \quad \square$$

Remarquons que ce théorème n'est bien sûr plus valable dans le cas où on considère une famille infinie  $\{C_l \mid l \in L\}$  de convexes, voici deux contre-exemples dans  $\mathbb{R}$  :

1.  $\{]0, \frac{1}{n}[ \mid n \in \mathbb{N}_0\}$ ,
2.  $\{[n, +\infty[ \mid n \in \mathbb{N}_0\}$ .

Il est cependant toujours possible de montrer que ce théorème est toujours valable pour une famille infinie de convexes compacts.

Signalons sans démonstration la généralisation suivante du théorème de Radon, dû au mathématicien Helge Arnulf Tverberg (1935 -).

**Théorème 2.21 (Théorème de Tverberg)** Soit  $A$  un sous-ensemble de  $(r - 1)(d + 1) + 1$  points de  $\mathbb{R}^d$  avec  $r \geq 2$  un entier. Alors il existe une partition de  $A$  en  $r$  sous-ensembles, soient  $A_1, \dots, A_r$  tels que

$$\bigcap_{i=1}^r \text{Conv}(A_i) \neq \emptyset.$$

Le cas où  $r = 2$  est précisément le théorème de Radon.

## 2.5 Séparation

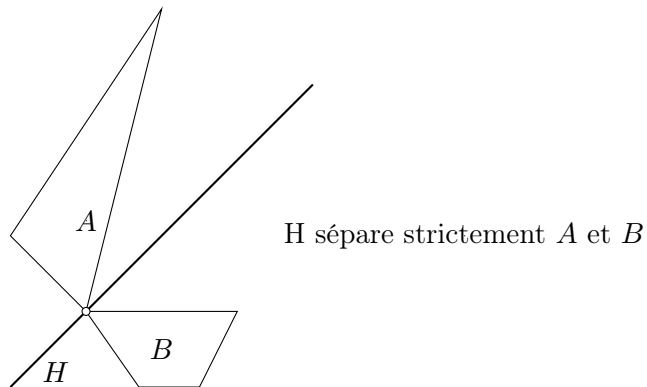
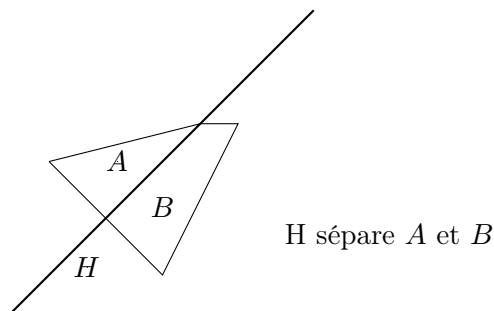
Soient  $A$  et  $B$  deux sous-ensembles de  $\mathbb{R}^d$ . Nous dirons qu'un hyperplan  $H \subseteq \mathbb{R}^d$  *sépare*  $A$  et  $B$  lorsque  $A \subseteq H^{\leq}$  et  $B \subseteq H^{\geq}$  (ou bien  $A \subseteq H^{\geq}$  et  $B \subseteq H^{\leq}$ ). Les ensembles  $A$  et  $B$  sont alors dit *séparables*. Il est bien sûr équivalent d'exiger l'existence d'une forme affine non dégénérée  $g$  et d'un réel  $k$  tels que  $g(a) \leq k$  pour tout  $a \in A$  et  $g(b) \geq k$  pour tout  $b \in B$ . Si on a de plus

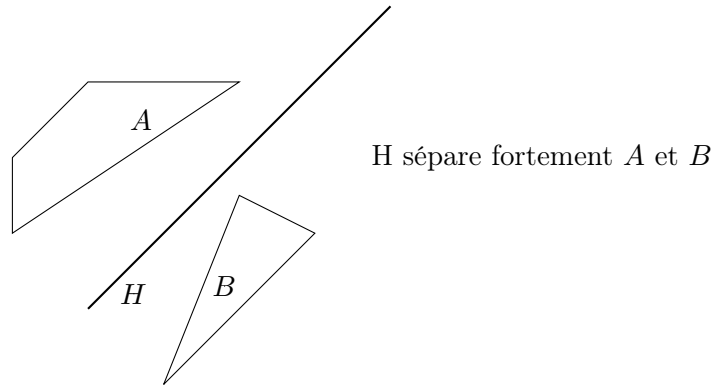
$$g(a) < k < g(b)$$

pour tous  $a \in A, b \in B$  alors  $A$  et  $B$  sont *fortement séparés*. Enfin, s'il existe un autre réel  $l$  tel que

$$g(c) \leq k < l \leq g(b)$$

les ensembles sont *strictement séparés*. Les dessins ci-dessous illustrent ces trois notions.





Remarquons que

**Proposition 2.22** *Un hyperplan  $H$  sépare  $A$  et  $B$  si et seulement si il sépare  $\text{Conv}(A)$  et  $\text{Conv}(B)$ .*

**Démonstration** : Si  $H$  sépare  $A$  et  $B$  alors on peut supposer  $A \subseteq H^{\leq}$  et  $B \subseteq H^{\geq}$  sans perte de généralité. Mais  $H^{\leq}$  et  $H^{\geq}$  sont des ensembles convexes contenant  $A$  et  $B$  et donc également  $\text{Conv}(A)$  et  $\text{Conv}(B)$  donc  $H$  sépare  $\text{Conv}(A)$  et  $\text{Conv}(B)$ . Si  $H$  sépare  $\text{Conv}(A)$  et  $\text{Conv}(B)$  alors  $H$  sépare clairement  $A$  et  $B$ .  $\square$

Nous pouvons donc nous restreindre à étudier la séparation des ensembles convexes. Notons que si  $\text{Aff}(A \cup B) \neq \mathbb{R}^d$  alors il existe un hyperplan affine contenant  $\text{Aff}(A \cup B)$  et ce dernier sépare trivialement  $A$  et  $B$ .

A présent, notre but est de prouver le théorème général suivant :

**Théorème 2.23 (Théorème de séparation)** *Soient  $C$  et  $D$  deux convexes non vides de  $\mathbb{R}^d$  tels que  $\text{Aff}(C \cup D) = \mathbb{R}^d$ . Alors  $C$  et  $D$  sont séparables si et seulement si*

$$\text{relint}(C) \cap \text{relint}(D) = \emptyset.$$

Nous aurons pour se faire besoin d'un certain nombre de résultats intermédiaires. Commençons par prouver un résultat plus faible

**Proposition 2.24** *Soient  $C$  et  $D$  deux convexes non vides, disjoints de  $\mathbb{R}^d$ . Alors il existe une partition de  $\mathbb{R}^d$  en deux convexes  $E$  et  $F$  tels que  $C \subseteq E$  et  $D \subseteq F$ .*

Une telle partition de  $\mathbb{R}^d$  en deux convexes est appelée une *bipartition convexe*. Pour démontrer cette proposition, nous travaillerons dans l'ensemble ordonné  $(\mathcal{P}(\mathbb{R}^d), \subseteq)$ . Rappelons donc brièvement les rudiments de la théorie des ensembles ordonnés. Soit  $(X, \leq)$  un *ensemble ordonné* (à savoir,  $\leq$  est une relation réflexive, antisymétrique et transitive sur  $X$ ). L'élément  $m \in X$  est le *maximum* de  $X$  si  $m \geq x$  pour tout  $x \in X$ . Un élément *maximal* de  $X$  est un élément  $m$  tel que

$m \leq x$  implique  $m = x$  pour tout  $x \in X$ . Un ensemble ordonné admet au plus un élément maximum, tandis que plusieurs éléments maximaux sont possibles.

Pour  $Y \subseteq X$  et  $x \in X$ , on dit que  $Y$  est *majoré* par  $x$  si  $x \geq y$  pour tout  $y \in Y$ . On dit alors que  $a$  est un *majorant* de  $Y$  ou que  $a$  *majoré*  $Y$ .

On définit de façon analogue le *minimum*, les éléments *minimaux* et les *minorants*.

Un sous-ensemble  $Y$  de  $X$  est une *chaîne* si la restriction de  $\leq$  à  $Y$  est un *ordre total* sur  $Y$  c'est-à-dire si deux éléments de  $T$  sont toujours en relation. De même, une *antichaîne* est un ensemble d'éléments tels que deux d'entre eux ne sont jamais en relation.

**Lemme 2.25 (Lemme de Zorn)** *Si toute chaîne de  $X$  est majorée, alors  $X$  possède un élément maximal.*

Ce lemme est trivial dans le cas où  $X$  est un ensemble fini, il est équivalent à l'*axiome du choix* sinon. Le lecteur se reportera à un livre de théorie des ensembles pour plus de détails.

**Démonstration de la proposition 2.24 :** Définissons

$$X = \{(P, Q) \mid C \subseteq P \subseteq \mathbb{R}^d, D \subseteq Q \subseteq \mathbb{R}^d, P \cap Q = \emptyset, P \text{ et } Q \text{ convexes}\}.$$

La relation d'ordre  $\leq$  est définie sur  $X$  par

$$(P_1, Q_1) \leq (P_2, Q_2)$$

si et seulement si  $P_1 \subseteq P_2$  et  $Q_1 \subseteq Q_2$ . Considérons une chaîne  $T$  de  $X$  et montrons que  $T$  est majorée par  $(R, S)$  avec

$$R = \bigcup \{P \mid \exists Q : (P, Q) \in T\}$$

et

$$S = \bigcup \{Q \mid \exists P : (P, Q) \in T\}.$$

Vérifions que  $(R, S)$  est bien dans  $X$ .

1.  $C \subseteq P \subseteq \mathbb{R}^d$  pour tout  $(P, Q) \in T$  donc  $C \subseteq R$  et bien sûr  $R \subseteq \mathbb{R}^d$ ,
2. De même  $D \subseteq S \subseteq \mathbb{R}^d$ ,
3.  $R$  est convexe car si  $a_1, a_2 \in R$  il existe  $(P_1, Q_1)$  ainsi que  $(P_2, Q_2)$  dans  $T$  tels que  $a_1 \in P_1$  et  $a_2 \in P_2$ . Comme  $T$  est une chaîne il faut  $P_1 \subseteq P_2$  ou  $P_2 \subseteq P_1$ . Ainsi  $a_1, a_2 \in P_1$  ou  $a_1, a_2 \in P_2$ . Comme  $P_1$  et  $P_2$  sont convexes, le segment  $[a_1, a_2]$  est soit dans  $P_1$  soit dans  $P_2$  qui sont tous deux inclus dans  $R$ . Donc  $R$  est bien convexe,
4. De même que pour  $R$ ,  $S$  est convexe,

5. Montrons que  $R \cap S = \emptyset$ . Si  $p \in R \cap S$  alors il existe  $(P_1, Q_1) \in T$  et  $(P_2, Q_2) \in T$  tels que  $p \in P_1$  et  $p \in Q_2$ . Si  $P_1 \subseteq P_2$  on a  $p \in P_2 \cap Q_2$ , une contradiction. De même si  $Q_2 \subseteq Q_1$  on a  $p \in P_1 \cap Q_1$  ce qui est contradictoire.

Il est maintenant clair que  $(R, S)$  majore  $T$ , soit donc  $(F, G)$  un élément maximal de  $(X, \leq)$  (par le lemme de Zorn). On a bien  $C \subseteq F$  ainsi que  $D \subseteq G$ . Les ensembles  $F$  et  $G$  sont non vides et d'intersection vide vu que  $(F, G) \in X$ . Montrons par l'absurde que

$$F \cup G = \mathbb{R}^d.$$

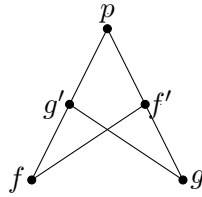
Supposons que  $p \in \mathbb{R}^d$  et que  $p$  n'est pas dans  $F \cup G$ . Ceci implique que

$$(Conv(F \cup \{p\}), G)$$

n'est pas dans  $X$  (cela serait en contradiction avec la maximalité de  $(F, G)$ ). Ceci implique à son tour que  $Conv(F \cup \{p\}) \cap G$  est non vide. Par la proposition 2.14 il existe donc  $f \in F$  et  $g' \in G$  tels que  $g' \in [p, f]$ . De même

$$(F, Conv(G \cup \{p\}))$$

n'est pas dans  $X$  et il existe  $g \in G$ ,  $f' \in F$  tels que  $f' \in [p, g]$ .



Par la proposition 1.19 du chapitre 1, il existe donc un élément  $z$  dans l'intersection

$$[f, f'] \cap [g, g']$$

ce qui contredit le fait que  $F$  et  $G$  sont convexes et d'intersection non vide.  $\square$

Pour une bipartition convexe  $F, G$  de  $\mathbb{R}^d$  définissons l'ensemble  $F | G = \{p \in \mathbb{R}^d \mid \text{il existe une droite } D \text{ par } p \text{ telle qu'une demi droite stricte de } D \text{ est dans } F, \text{ l'autre dans } G\}$ .

**Proposition 2.26** *Soit  $F, G$  une bipartition convexe de  $\mathbb{R}^d$ . Alors  $F | G$  est un hyperplan affiné de  $\mathbb{R}^d$ .*

**Démonstration** : Commençons par montrer que  $F | G$  est un sous-espace affiné de  $\mathbb{R}^d$ . Soient  $a$  et  $b$  deux points distincts de  $F | G$  et montrons que  $\langle a, b \rangle \subseteq F | G$ . Il existe donc des points  $a', b'$  avec  $a'$  différent de  $a$  et  $b'$  différent de  $b$  tels que

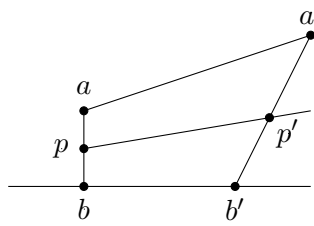
$$]a, a' > \subseteq F \text{ et } \langle a, a' \rangle \setminus ]a, a' > \subseteq G$$

ainsi que

$$]b, b' > \subseteq F \text{ et } < b, b' > \setminus ]b, b' > \subseteq G.$$

Pour  $p \in < a, b >$  nous avons soit  $p \in ]a, b[$  soit  $b \in ]a, p[$  soit  $a \in ]p, b[$ . Dans les trois cas, nous prouverons que  $p$  est dans  $F \mid G$ .

Supposons d'abord que  $p \in ]a, b[$ . Alors  $p = \lambda a + \mu b$  avec  $\lambda$  et  $\mu$  des réels positifs tels que  $\lambda + \mu = 1$ . Posons  $p' = \lambda a' + \mu b'$  et montrons que  $]p, p' > \subseteq F$ . Comme  $F, G$  partitionne  $\mathbb{R}^d$  on aura alors  $< p, p' > \setminus ]p, p' > \subseteq G$  et  $p$  sera dans  $F \mid G$  par définition.



Un point de  $]p, p' >$  s'écrit  $q = \alpha p + \beta p'$  avec  $\alpha + \beta = 1$  et  $\beta > 0$ . Mais alors

$$q = \alpha(\lambda a + \mu b) + \beta(\lambda a' + \mu b')$$

c'est-à-dire

$$q = \lambda(\alpha a + \beta a') + \mu(\alpha b + \beta b')$$

avec  $(\alpha a + \beta a')$  un point de  $]a, a' > \subseteq F$  et  $\alpha b + \beta b'$  un point de  $]b, b' > \subseteq F$ . Comme  $F$  est convexe, le point  $q$  est bien dans  $F$ .

Traitons maintenant le cas  $b \in ]a, p[$ , c'est-à-dire  $p = \lambda a + \mu b$  avec  $\lambda + \mu = 1$  et  $\mu > 1$ . Posons  $p' = \lambda(2a - a') + \mu b$  et montrons que  $]p, p' \subseteq F$ . Un point de ce segment s'écrit  $q = \alpha p + \beta p'$  ce qui s'écrit encore

$$\begin{aligned} q &= \alpha(\lambda a + \mu b) + \beta(\lambda(2a - a') + \mu b) \\ &= \lambda(\alpha a + \beta(2a - a')) + \mu(\alpha b + \beta b'). \end{aligned}$$

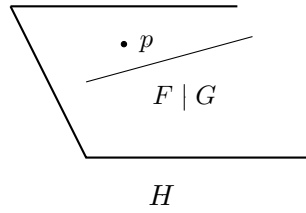
Le point  $2a - a'$  est dans  $< a, a' > \setminus ]a, a' >$  c'est-à-dire dans  $G$ . Ainsi  $\alpha a + \beta(2a - a')$  est dans  $]a, (2a - a') >$  et est donc aussi dans  $G$ . Le point  $\alpha b + \beta b'$  quant à lui est dans  $]b, b' >$  c'est-à-dire dans  $F$ . En utilisant les conditions sur  $\lambda$  et  $\mu$ , on voit alors aisément que si  $q \in G$  on aurait  $\alpha b + \beta b' \in G$  car  $G$  est convexe et

$$\alpha b + \beta b' \in [q, \alpha a + \beta(2a - a')]$$

ce qui est contradictoire. Donc  $q \in F$ .

Le cas  $a \in ]p, b[$  se traite de façon analogue. La preuve du fait que  $F \mid G$  est un sous-espace affiné est maintenant terminée.

Prouvons ensuite que  $F \mid G$  ne peut être strictement inclus dans un hyperplan affiné. Supposons au contraire que  $F \mid G \subseteq H$  avec  $H$  un hyperplan affiné de  $\mathbb{R}^d$ . Soit  $p \in H \setminus (F \mid G)$ .



Supposons de plus que  $p \in F$  (le cas où  $p \in G$  est semblable). Soit  $x \in \mathbb{R}^d \setminus H$  et  $D$  l'unique droite passant par  $p$  et  $x$  (de sorte que  $D \cap H = \{p\}$ ). Montrons que  $x$  est contenu dans  $F$ . Si ce n'est pas le cas alors  $x \in G$ , définissons l'ensemble

$$A_x = \{\lambda \in \mathbb{R} \mid \lambda x + (1 - \lambda)p \in F\}.$$

Alors  $0 \in A_x$  et  $1$  majore  $A_x$  (vu que  $1$  n'appartient pas à  $A_x$  par hypothèse, et tout réel supérieur à  $1$  ne peut appartenir à  $A_x$  par la convexité de  $F$ ). Donc l'ensemble  $A_x$  admet une borne supérieure, soit  $\mu = \sup A_x$ . Montrons que  $\mu x + (1 - \mu)p \in F \mid G$  (ce qui est une contradiction avec  $D \cap H = \{p\}$ ). Par définition de  $\mu$  nous avons

$$\mu' x + (1 - \mu')p \in G$$

pour tout  $\mu' > \mu$  ce qui montre qu'une demi-droite de  $D$  est dans  $G$ . Maintenant si  $\mu' < \mu$  alors par définition du suprémum il existe  $\mu'' \in A_x$  tel que  $\mu' < \mu'' < \mu$  de sorte que si  $\mu' x + (1 - \mu')p \in G$  nous aurons aussi  $\mu'' x + (1 - \mu'')p \in G$  par la convexité de  $G$ , ce qui est contradictoire.

Nous avons donc démontré que  $\mathbb{R}^d \setminus H \subseteq F$ . Mais comme  $F$  est convexe, il faut  $F = \mathbb{R}^d$  ce qui contredit que  $G \neq \emptyset$ .

Il reste à démontrer que  $F \mid G \neq \mathbb{R}^d$  pour terminer la preuve. Remarquons qu'on a  $Aff(F) = \mathbb{R}^d$  (dans le cas contraire  $Aff(F) \subseteq H$  pour un certain hyperplan  $H$  et le même raisonnement que ci-dessus montre que  $G = \mathbb{R}^d$ ). Mais alors  $relint(F) = int(F)$  qui est non vide (voyez la section consacré à l'intérieur relatif). Si  $q \in int(F)$  il n'est pas difficile de voir que  $q \notin F \mid G$  (il existe une boule de rayon  $\epsilon$  contenue dans  $F$  et tout segment de droite passant par  $q$  doit contenir un point de cette boule).  $\square$

La proposition suivante fournit une preuve du théorème 2.23 dans le cas où les deux convexes sont disjoints :

**Proposition 2.27** Soient  $C$  et  $D$  deux convexes non vides, disjoints de  $\mathbb{R}^d$  et  $F, G$  une bipartition convexe de  $\mathbb{R}^d$  telle que  $C \subseteq F$  et  $D \subseteq G$ . Alors l'hyperplan associé  $H = F \mid G$  sépare  $C$  et  $D$ .

**Démonstration** : Soit  $p \in H^<$ . Sans perte de généralité on peut supposer que  $p \in F$  (sinon  $p \in G$  et il suffit de renommer les deux éléments de la partition).

Il suffit de prouver que  $H^< \subseteq F$  et  $H^> \subseteq G$ , ceci impliquant (comme  $F$  et  $G$  partitionnent) que  $F \subseteq H^{\leq}$  et  $G \subseteq H^{\geq}$ . Soit  $f \in H^<$  un point différent de  $p$ . Si

$f \notin F$  alors  $f \in G$ . Or le segment  $[p, f]$  est inclus à  $H^<$  car  $H^<$  est convexe. En appliquant le même argument que dans la preuve de la proposition précédente, on trouve un point du segment qui est dans  $F \mid G$  c'est-à-dire dans  $H$ , d'où la contradiction.

Montrons enfin  $H^> \subseteq G$ . Soit  $r \in H^>$ , si  $r \in F$  alors  $[r, p] \subseteq F$ . En effet, un point de cette demi droite s'écrit

$$r + \lambda(p - r)$$

pour un certain  $\lambda \geq 0$ . On peut de plus supposer  $\lambda > 1$  (le segment  $[p, r]$  étant dans  $F$  par la convexité de  $F$ ) et dans ce cas, si  $H = g^{-1}(\alpha)$  on a

$$g(r + \lambda(p - r)) = (1 - \lambda)g(r) + \lambda g(p) < (1 - \lambda)\alpha + \lambda\alpha = \alpha$$

et comme  $H^< \subseteq F$  on a bien  $[r, p] \subseteq F$ .

Nous avons en fait  $\langle p, r \rangle \subseteq F$  vu que, si  $u \in G$  et  $u \in \langle p, r \rangle \setminus [r, p]$  alors par le même argument que dans la proposition précédente on aurait un point  $u' \in H$  situé sur la droite, sur le segment  $[u, r]$ , une contradiction car  $r \in H^>$  et  $p \in H^<$ . Mais donc

$$H^< \cup \langle p, r \rangle \subseteq F$$

et en utilisant la convexité de  $F$ , il n'est pas difficile de voir que forcément  $F = \mathbb{R}^d$ , et donc  $G = \emptyset$  (contradiction).  $\square$

Et enfin :

**Démonstration du théorème 2.23 :** Supposons  $C \subseteq H^{\leq}$  et  $D \subseteq H^{\leq}$  pour un certain hyperplan  $H$ . Par hypothèse, soit  $C$  soit  $D$  n'est pas inclus dans  $H$ , supposons sans perte de généralité que ce soit  $C$ . Alors  $\text{relint}(C) \subseteq H^<$ . En effet, on prouve aisément que sous ces conditions, tout ouvert de  $\text{Aff}(C)$  centré en  $c \in C \cap H$  rencontre  $H^<$  et  $H^>$  de sorte que  $c$  n'est pas dans l'intérieur relatif de  $C$ . Comme  $\text{relint}(D) \subseteq D \subseteq H^{\geq}$ , la première implication du théorème est prouvée.

Soient  $C$  et  $D$  deux convexes non vides tels que  $\text{relint}(C) \cap \text{relint}(D) = \emptyset$ . Mais  $\text{relint}(C)$  et  $\text{relint}(D)$  sont aussi des convexes non vides (voir la section sur l'intérieur relatif), et il existe donc un hyperplan  $H$  les séparant, soit  $\text{relint}(C) \subseteq H^{\leq}$  et  $\text{relint}(D) \subseteq H^{\geq}$ . Finalement, si  $c \in C$  et  $c \in H^>$  alors par la proposition 2.17 il faut

$$[p, c] \subseteq \text{relint}(C) \subseteq H^{\leq}$$

pour tout  $p \in \text{relint}(C)$ , ce qui n'est bien sûr pas possible. Ainsi  $C \subseteq H^{\leq}$  et de même  $D \subseteq H^{\geq}$ .  $\square$

## 2.6 Appui et faces

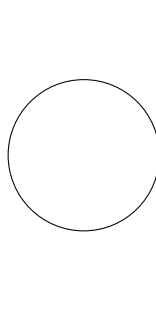
Soit  $C$  un convexe de  $\mathbb{R}^d$  ainsi que  $g$  une forme affine sur  $\mathbb{R}^d$  et  $k$  un réel. L'inégalité  $g(x) \geq k$  sera dite *valide* pour  $C$  si  $g(c) \geq k$  pour tout  $c \in C$  (c'est-



à-dire si  $C$  est contenu dans le demi-espace d'inégalité  $g(x) \geq k$ . Si de plus on a

$$k = \inf_{c \in C} g(c)$$

alors l'inégalité est dite d'*appui* pour  $C$ , de même que le demi-espace et l'hyperplan correspondant. La dessin ci-dessous montre une boule fermée de  $\mathbb{R}^2$  ainsi qu'une de ses droites d'appui. Les droites d'appui sont exactement les droites tangentes au cercle correspondant au bord de la boule.



On obtient immédiatement que :

**Proposition 2.28** *Si  $C$  est un convexe compact de  $\mathbb{R}^d$  (donc en particulier si  $C$  est un polytope), alors l'hyperplan  $g(x) = k$  est d'appui pour  $C$  si et seulement si il existe un point  $p$  de  $C$  tel que soit*

$$g(p) = k = \min_{c \in C} g(c)$$

soit

$$g(p) = k = \max_{c \in C} g(c).$$

**Démonstration** : L'hyperplan définit deux demi-espaces, à savoir  $I_1 : g(x) \leq k$  et  $I_2 : g(x) \geq k$ . D'autre part, une forme affine est une application continue, et toute application continue atteint un minimum et un maximum sur un compact, soit

$$u = \min\{g(x) \mid x \in C\}$$

et

$$v = \max\{g(x) \mid x \in C\}.$$

Il est maintenant clair que l'hyperplan sera d'appui si et seulement si une des deux inégalités est d'appui, c'est-à-dire si et seulement si  $u = k$  ou  $v = k$ .  $\square$

Une *face* de  $C$  est soit  $C$  lui même, soit le vide, soit l'intersection de  $C$  avec un hyperplan d'appui. Toute face d'un convexe est bien sûr toujours convexe (en tant qu'intersection de deux convexes). Les *facettes* de  $C$  sont les faces de dimension  $\dim(C) - 1$ . A titre d'exemple, remarquons qu'une boule fermée admet

une infinité de faces (tous les points de la sphère constituant le bord sont des faces), qu'un carré dans  $\mathbb{R}^2$  admet  $1 + 4 + 4 + 1 = 10$  faces, qu'un cube de  $\mathbb{R}^3$  admet  $1 + 8 + 12 + 6 + 1 = 28$  faces.

**Proposition 2.29** Soit  $F \subseteq C$ . Alors  $F$  est une face de  $C$  si et seulement si il existe une forme affine  $g$  telle que

1.  $g(c) \geq 0 \forall c \in C$ ,
2.  $F = \{p \in C \mid g(p) = 0\}$ .

**Démonstration** : Si  $F$  est une face de  $C$  il existe une forme affine  $g$  et un réel  $k$  tels que  $g(c) \geq k$  et  $F = \{p \in C \mid g(p) = k\}$ . On peut toujours supposer que  $k = 0$  quitte à considérer la forme affine  $g'(x) = g(x) - k$ . L'autre implication est évidente.  $\square$

On déduit de cette proposition :

**Proposition 2.30** Si  $F_1$  et  $F_2$  sont deux faces du convexes  $C \subseteq \mathbb{R}^d$  alors il en va de même pour  $F_1 \cap F_2$ .

**Démonstration** : Soient

$$F_1 = \{p \in C \mid g_1(p) \geq 0\}$$

et

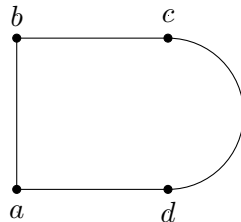
$$F_2 = \{p \in C \mid g_2(p) \geq 0\}.$$

La forme affine  $g_3 = g_1 + g_2$  est toujours valide pour  $C$  et en posant

$$F_3 = \{p \in C \mid g_3(p) = 0\}$$

on a bien sûr  $F_3 = F_1 \cap F_2$ .  $\square$

Un point  $p \in C$  est dit *exposé* si  $\{p\}$  est une face  $C$ . Il est dit *extrême* si pour tous  $q, r \in C$  tels que  $p \in [q, r]$  on a  $p = q$  ou  $p = r$ . On note  $exp(C)$  l'ensemble des points exposés de  $C$  et  $ext(C)$  l'ensemble des points extrêmes. Voici un exemple dans  $\mathbb{R}^2$ .



Les points exposés de la figure représentée ci-dessous sont  $a, b$  ainsi que les points du demi-cercle sauf  $c$  et  $d$ . Les points extrêmes sont les même avec  $c$  et  $d$  en plus. On remarque donc qu'un point extrême n'est pas forcément exposé. Nous allons en revanche démontrer qu'un point exposé est toujours extrême. Cela résulte du lemme suivant.

**Lemme 2.31** *Soit  $F$  une face du convexe  $C \subseteq \mathbb{R}^d$ . Alors*

$$\text{ext}(F) = F \cap \text{ext}(C).$$

**Démonstration** : Si  $p \in F \cap \text{ext}(C)$  alors clairement  $p \in \text{ext}(F)$  de sorte qu'il ne reste qu'une des deux inclusions à justifier. Soit  $p \in \text{ext}(F) \subseteq F$  et montrons que  $p \in \text{ext}(C)$ . Il existe donc une forme affine  $g$  telle que

$$g(c) \geq 0 \quad \forall c \in C$$

et

$$F = \{p \in C \mid g(p) = 0\}.$$

Si  $p \notin \text{ext}(C)$  il existe  $a, b \in C$  tels que  $p \in ]a, b[$  mais  $a \notin F$  ou bien  $b \notin F$  (vu que  $p \in \text{ext}(F)$ ). Ainsi soit  $g(a) > 0$  soit  $g(b) > 0$ . Mais il existe  $0 < \lambda < 1$  tel que

$$g(p) = g(\lambda a + (1 - \lambda)b) = \lambda g(a) + (1 - \lambda)g(b) > 0$$

ce qui est contradictoire vu que  $p \in F$ .  $\square$

**Proposition 2.32** *Pour tout convexe  $C$  de  $\mathbb{R}^d$ , on a*

$$\text{exp}(C) \subseteq \text{ext}(C).$$

**Démonstration** : Si  $t \in \text{exp}(C)$  alors  $T = \{t\}$  est une face de  $C$ . Par le lemme précédant

$$\text{ext}(T) = T \cap \text{ext}(C)$$

et bien sûr  $\text{ext}(T) = T$  de sorte que  $t \in \text{ext}(C)$ .  $\square$

Démontrons maintenant la caractérisation suivante des points extrêmes de  $C$  :

**Proposition 2.33** *Si  $C \subseteq \mathbb{R}^d$  est convexe et  $p \in C$ , alors  $p \in \text{ext}(C)$  si et seulement si  $C \setminus \{p\}$  est aussi convexe.*

**Démonstration** : Soit  $p$  un point extrême. Si  $C \setminus \{p\}$  n'est pas convexe alors il existe deux points  $q, r$  dans  $C \setminus \{p\}$  tels que  $p \in ]q, r[$ , contradiction avec le fait que  $p$  est extrême.

De même si  $C \setminus \{p\}$  est convexe pour un certain  $p \in C$ , alors  $p$  est clairement extrême.  $\square$

Le théorème suivant, dû à Hermann Minkowski (1864 - 1909), montre l'intérêt des points extrême.

**Théorème 2.34 (Théorème de Minkowski)** *Pour un convexe compact de  $\mathbb{R}^d$ , on a toujours*

$$C = \text{Conv}(\text{ext}(C)).$$

*De plus, si  $C = \text{Conv}(M)$  pour un certain  $M \subseteq C$ , alors  $\text{ext}(C) \subseteq M$ .*

**Démonstration** : Supposons tout d'abord que  $C = \text{Conv}(M)$ . Si  $x \in \text{ext}(C)$  et  $x \notin M$  alors  $M \subseteq C \setminus \{x\}$  mais  $C \setminus \{x\}$  est aussi convexe vu que  $x$  est extrême, donc  $\text{Conv}(M) \subseteq C \setminus \{x\}$ , d'où la contradiction.

Montrons maintenant que  $C \subseteq \text{Conv}(\text{ext}(C))$  (l'autre inclusion étant triviale) par induction sur la dimension  $\dim(C) = \dim(\text{Aff}(C))$ . L'énoncé est trivial si  $\dim(C) \leq 1$  car alors  $C$  est vide,  $C$  est un singleton ou  $C$  est un segment de droite. Supposons donc  $\dim(C) > 1$  et soit  $p \in C$ . Si  $p \in \text{ext}(C)$  alors c'est terminé. Sinon il existe deux points  $q, r \in C$  avec

$$p \in ]q, r[.$$

Cela implique que  $C \cap \langle p, q \rangle$  est un segment de droite fermé, soit  $[q', r']$ . Le point  $q'$  n'est clairement pas dans  $\text{relint}(C)$ . Par le théorème de séparation, il existe donc un hyperplan  $H$  séparant  $q'$  et  $H$ . On a alors que  $q' \in H$  de sorte que  $H \cap C$  est un convexe compact contenant  $q'$ . Ainsi  $F = H \cap C$  est une face de  $C$ . Le même raisonnement tient pour  $r'$ , soit donc  $G = H' \cap C$  une face de  $C$  contenant  $r'$ . Par l'hypothèse de récurrence on a

$$q' \in \text{Conv}(\text{ext}(H \cap C))$$

ainsi que

$$r' \in \text{Conv}(\text{ext}(H' \cap C)).$$

Mais alors  $p \in [q', r']$  et en utilisant les propositions 2.14 et 2.10 on a

$$\begin{aligned} [q', r'] &\subseteq \text{Conv}(\text{Conv}(\text{ext}(F)) \cup \text{Conv}(\text{ext}(G))) \\ &= \text{Conv}(\text{ext}(F) \cup \text{ext}(G)). \end{aligned}$$

Comme  $\text{ext}(F) \subseteq \text{ext}(C)$  et  $\text{ext}(G) \subseteq \text{ext}(C)$  par le lemme 2.31, la preuve est terminée.  $\square$

Notons que ce théorème n'est bien sûr plus nécessairement vrai dans le cas des convexes non compact, un tel convexe n'admettant pas forcément de points extrêmes (considérer par exemple une droite). En revanche, il existe une version de ce théorème dans le cas des espaces vectoriels de dimension infinie, il s'agit du *théorème de Krein-Milman* (voyez un cours d'analyse fonctionnelle pour son énoncé exact). Nous présenterons une généralisation de ce théorème pour les convexes fermés non bornés ne contenant pas de droite dans la section suivante.

Il existe un analogue du théorème de Minkowski pour les points exposés :

**Théorème 2.35 (Théorème de Straszewicz)** *Pour un convexe compact de  $\mathbb{R}^d$ , on a toujours*

$$C = \text{adh}(\text{Conv}(\text{exp}(C))).$$

**Démonstration** : Soit  $D$  le membre de droite. L'inclusion  $D \subseteq C$  est claire vu que,  $C$  étant convexe, on a  $\text{Conv}(\text{ext}(C)) \subseteq C$  et ainsi  $D \subseteq C$  (car  $C$  est fermé).

Supposons qu'il y ait un point  $p \in C \setminus D$ . Il existe alors une boule ouvert centrée en  $p$  et disjoint de  $\text{Conv}(\text{exp}(C))$  (par définition de l'adhérence). Soit  $H$  un hyperplan séparant  $\text{Conv}(\text{exp}(C))$  et la boule (par le théorème de séparation). Sans perte de généralité, nous avons donc  $p \in H^<$  et  $\text{exp}(C) \subseteq H^\geq$ . Nous allons en déduire une contradiction, en montrant qu'il existe un point exposé dans  $H^<$ .

Soit  $q$  le point situé à l'intersection de la droite passant par  $p$  et perpendiculaire à  $H$ . La fonction  $f$  définie par

$$f : C \cap H \rightarrow \mathbb{R} : x \rightarrow \|x - q\|$$

est continue sur le compact  $C \cap H$ , elle atteint donc un maximum  $\alpha$ . Soit  $s$  un point de  $H^>$  tel que  $q \in [p, s]$ . Par le théorème de Pythagore et la définition de  $\alpha$  nous avons

$$\|s - x\| = \sqrt{\|s - q\|^2 + \|q - x\|^2} \leq \sqrt{\|s - q\|^2 + \alpha^2}.$$

Choisissons maintenant  $s$  tel que

$$\|s - q\| > \frac{1}{2} \left( \frac{\alpha^2}{\|p - q\|} - \|p - q\| \right).$$

Alors

$$\alpha^2 < \|p - q\|(\|p - q\| + 2\|s - q\|)$$

c'est-à-dire (comme  $\|p - q\| = \|p - s\| - \|s - q\|$ )

$$\alpha^2 < (\|p - s\| - \|s - q\|)(\|p - s\| + \|s - q\|).$$

Mais donc

$$\|p - s\|^2 > \|s - q\|^2 + \alpha^2$$

et en utilisant la première inégalité ci-dessus, nous avons

$$\|x - s\| < \|s - p\|$$

pour tout  $x \in C \cap H$ .

Posons à présent

$$\beta = \max_{z \in C \cap H^\leq} \|z - s\|$$

(ce maximum existant par la compacité de  $C \cap H^\leq$ ). La boule fermée  $A$  de centre  $s$  et de rayon  $\beta$  contient donc  $C \cap H^\leq$ . Son bord (la sphère de centre  $s$  et de rayon

$\beta$ ) rencontre  $C \cap H^\leq$  en au moins un point, soit  $t$  (c'est le maximum en question). Mais  $t$  ne peut être dans  $C \cap H$  vu le choix de  $s$ . Donc  $t \in H^<$ . Remarquons pour finir que l'hyperplan  $G$  tangent à la sphère et passant par  $t$  est d'appui pour  $C \cap H^\leq$ . Cet hyperplan est aussi d'appui pour  $C$ . Il suffit en effet de vérifier que si  $C \cap H^\leq \subseteq G^\leq$  alors  $C \subseteq G^\leq$ . Ceci est évident car si  $u \in C \cap H^> \cap G^>$  alors  $]t, u] \subseteq G^>$ . Mais comme  $t \in H^<$ , le segment  $]t, u]$  doit contenir un point de  $C \cap H$ , contradiction.

On vérifie aisément que  $G \cap C$  ne peut contenir d'autre point que  $t$ . Donc  $G$  est d'appui pour  $C$  et  $G \cap C = \{t\}$ , ce qui fait de  $t$  un point exposé de  $C$  contenu dans  $H^<$ .  $\square$

## 2.7 Convexes non bornés

Voici une jolie caractérisation des ensembles convexes non bornés de  $\mathbb{R}^d$  :

**Théorème 2.36** *Soit  $C \subseteq \mathbb{R}^d$  un convexe. Alors  $C$  est non borné si et seulement si il contient une demi-droite.*

**Démonstration** : Si la demi droite  $\{p + \lambda(q - p) \mid \lambda > 0\}$  est dans  $C$  alors

$$\| \|p\| - \lambda \|p - q\| \| \leq \|p + \lambda(q - p)\|$$

et le membre de gauche de cette inégalité tend vers l'infini quand  $\lambda$  tend vers l'infini de sorte que  $C$  ne peut être borné.

La deuxième implication est moins triviale. Nous pouvons supposer que  $0 \in \text{relint}(C)$  (sinon on translate  $C$  de façon à ce que ce soit le cas, si le translaté contient une demi droite alors il en va de même pour  $C$ ). Soit  $x_n$  une suite de  $C$  telle que  $\|x_n\| \rightarrow \infty$  (une telle suite existe car  $C$  est non borné). La suite

$$\frac{x_n}{\|x_n\|}$$

est dans la sphère unité et, cette dernière étant compact, tend vers un élément  $x_0$  qui est toujours dans la sphère unité, c'est à dire  $\|x_0\| = 1$ . Prouvons maintenant que la demi-droite  $[0, x_0[$  est dans  $C$ .

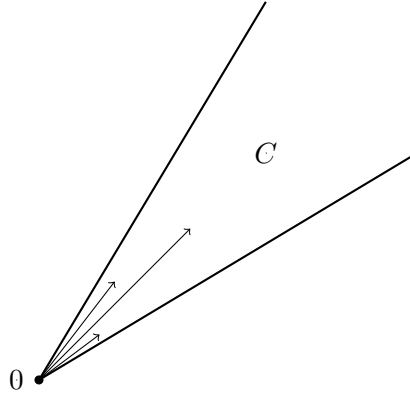
Soit  $r$  un réel positif. Il existe donc  $m \in \mathbb{N}_0$  tel que pour tout  $n > m$  on ait  $r < \|x_n\|$ . Alors pour un tel  $n$  on a  $0 < \frac{r}{\|x_n\|} < 1$  et

$$(1 - r\|x_n\|^{-1})0 + r\|x_n\|^{-1}x_n = r\|x_n\|^{-1}x_n \in C$$

par la convexité de  $C$ . Donc  $rx_0 \in \text{adh}(C)$  quel que soit  $r > 0$ . Ceci montre que la demi-droite est incluse dans  $\text{adh}(C)$ . Maintenant, pour un certain  $r$  on a  $rx_0 \in [0, sx_0[$  avec  $s > r$  quelconque et  $sx_0 \in \text{adh}(C)$ , donc  $rx_0 \in \text{relint}(C)$  par

la proposition 2.17.  $\square$

Un sous-ensemble  $C \subseteq \mathbb{R}^d$  est appelé un *cône convexe* si  $u, v \in C$  implique  $u + v \in C$  d'une part et si  $\lambda u \in C$  pour tout  $\lambda \in \mathbb{R}, \lambda \geq 0$  et tout  $u \in C$  d'autre part. En particulier, tout cône convexe non vide contient le vecteur nul. Ceci est illustré ci-dessous.



La démonstration de la propriétés ci-dessous est triviale.

**Proposition 2.37** *Si  $C$  est un cône convexe non vide de  $\mathbb{R}^d$  alors soit  $C = \{0\}$  soit  $C$  est un convexe non borné. De plus, toute intersection de cône convexe est encore un cône convexe.*

L'*enveloppe conique* d'un sous-ensemble  $X \subseteq \mathbb{R}^d$  est l'intersection de tous les cônes convexes de  $\mathbb{R}^d$  contenant  $X$ . On la note  $Cone(Y)$ . Le lecteur ne sera pas étonné d'apprendre que, de même que pour les enveloppes linéaires, affines et convexes, l'application  $Cone$  définit un opérateur de fermeture sur l'ensembles des parties de  $\mathbb{R}^d$ .

Si  $x_1, \dots, x_n$  sont des éléments de  $X$  et si  $\lambda_1, \dots, \lambda_n$  sont des réels positifs, alors

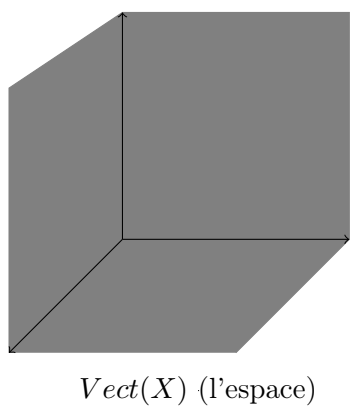
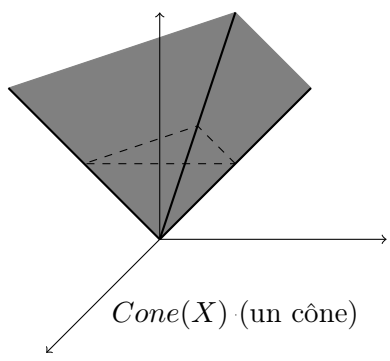
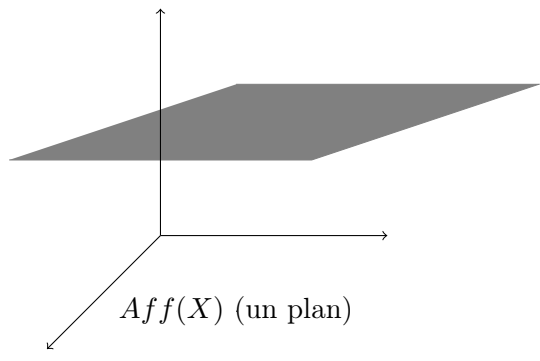
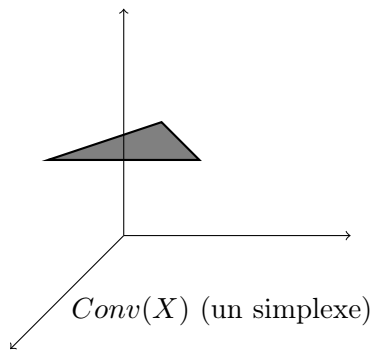
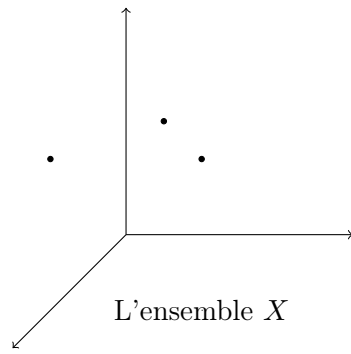
$$\sum_{i=1}^n \lambda_i x_i$$

est une *combinaison conique* des  $x_i$ . Il est alors facile de prouver que, comme on s'y attend :

**Proposition 2.38** *Soit  $X \subseteq \mathbb{R}^d$ . Alors*

$$Cone(X) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_i \geq 0, x_i \in X \right\}.$$

Sur le dessin de la page suivante figure une illustration (pour  $d = 3$ ) des enveloppes convexes, affines, coniques et linéaires d'un ensemble  $X$  de trois points.





Il est évident que :

**Proposition 2.39** *Si  $X$  est un sous-ensemble de  $\mathbb{R}^d$  alors*

$$\text{Conv}(X) \subseteq \text{Cone}(X) \subseteq \text{Vect}(X).$$

Soit  $C$  un convexe de  $\mathbb{R}^d$  et  $p \in C$ . Définissons le *cône caractéristique* de  $C$  en  $p$  comme

$$CC_p(C) = \{v \in \mathbb{R}^d \mid \forall \lambda \geq 0 : p + \lambda v \in C\}.$$

Il se fait que si  $C$  est fermé,  $CC_p(C) = CC_q(C)$  quels que soient  $p$  et  $q$  dans  $C$ . Pour démontrer cela nous utilisons le lemme :

**Lemme 2.40** *Soit  $C$  un convexe fermé de  $\mathbb{R}^d$  contenant la demi-droite  $L = \{\lambda z \mid \lambda \geq 0\}$ . Soient  $p, q \in C$ . Alors  $p + L \subseteq C$  si et seulement si  $q + L \subseteq C$ .*

**Démonstration** : Il suffit de démontrer que si  $p + L \subseteq C$  alors  $q + L \subseteq C$  (par symétrie de l'énoncé en  $p$  et  $q$ ). Soit donc  $q + \mu z$  un point de  $q + L$ . Par la convexité de  $C$  et le fait que  $C$  contient  $p + L$  nous savons que pour tout  $\lambda > \mu$

$$v_\lambda := \left(1 - \frac{\mu}{\lambda}\right)q + \frac{\mu}{\lambda}(p + \lambda z) \in C.$$

Mais comme  $C$  est fermé, nous avons

$$\lim_{\lambda \rightarrow \infty} v_\lambda \in C$$

pour autant que cette limite existe. Or elle existe bel et bien et vaut  $q + \mu z \in C$ .  
□

**Corollaire 2.41** *Si  $C$  est un convexe fermé de  $\mathbb{R}^d$  alors  $CC_p(C) = CC_q(C)$  quel que soient  $p, q \in C$ .*

Le *cône caractéristique* du convexe fermé  $C$  est l'unique ensemble  $CC_p(Q)$  pour  $p \in C$ . Il est noté  $CC(C)$ . Nous laissons au lecteur le soin de vérifier :

**Proposition 2.42** *Le cône caractéristique d'un convexe fermé  $C$  est un cône convexe non vide. De plus,  $C$  est borné si et seulement si  $CC(C) = \{0\}$ .*

Le théorème suivant généralise le théorème de Minkowski aux convexes non bornés ne contenant pas de droite :

**Théorème 2.43** *Si  $C$  est un convexe fermé de  $\mathbb{R}^d$  ne contenant aucune droite, alors*

$$C = CC(C) + \text{Conv}(\text{ext}(C)).$$

**Démonstration** : Si  $u \in \text{Conv}(\text{ext}(C))$  alors  $u \in C$ . Si  $p \in \text{CC}(C)$  alors  $p \in \text{CC}_u(C)$  donc  $p + u \in C$  ce qui prouve une des deux inclusions.

Pour l'autre inclusion on procède par induction sur  $\dim(C)$ . Le théorème est évident lorsque  $\dim(C) = 0$  (vu qu'alors  $C$  est réduit à un point). Pour  $\dim(C) > 0$ , nous pouvons supposer  $C$  non borné (sinon on retrouve le théorème de Minkowski). Soit donc  $v \in \text{CC}(C)$  non nul. Quel que soit  $p \in C$  nous avons donc  $\{p + \lambda v \mid \lambda \geq 0\} \subseteq C$ . Vu nos hypothèses, la demi-droite

$$\{p + \mu v \mid \mu \leq 0\}$$

n'est pas dans  $C$ . Il existe donc  $\mu \leq 0$  tel que  $q = p + \mu v$  soit dans la frontière de  $C$ . Alors  $q$  n'est pas dans  $\text{relint}(C)$ , soit donc  $H$  un hyperplan séparant  $q$  et  $\text{relint}(C)$ . On peut supposer que  $H$  ne contient pas  $\text{relint}(C)$ . Cet hyperplan détermine une face  $F$  de  $C$  contenant  $q$ . Par l'hypothèse de récurrence

$$q \in F = \text{CC}(F) + \text{Conv}(\text{ext}(F)).$$

Donc

$$p = q - \mu v \in q + \text{CC}(C) \subseteq \text{CC}(C) + \text{CC}(F) + \text{Conv}(\text{ext}(F))$$

mais  $\text{Conv}(\text{ext}(F)) \subseteq \text{Conv}(\text{ext}(C))$  d'une part et

$$\text{CC}(C) + \text{CC}(F) = \text{CC}(C)$$

d'autre part (vu que  $\text{CC}(F)$  est inclus dans  $\text{CC}(C)$  et que  $\text{CC}(C)$  est un cône convexe). Ceci termine la preuve.  $\square$

Ce théorème n'est plus vrai dans le cas des convexes non bornés contenant une droite (voyez par exemple une droite, qui n'admet aucun point extrême).

## Chapitre 3

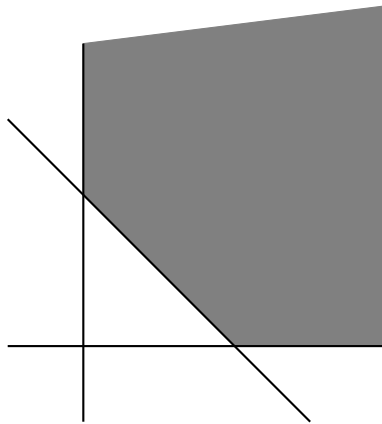
# Polytopes et ensembles polyédriques

Nous commençons ici l'étude plus approfondie des polytopes. Le théorème fondamental sur les polytopes affirme qu'un ensemble  $P$  de  $\mathbb{R}^d$  est un polytope (enveloppe convexe d'un ensemble fini de points) si et seulement si il est une intersection bornée d'un nombre fini de demi-espaces fermés. La première section de ce chapitre est consacrée à la démonstration de ce théorème, ainsi qu'à certaines de ses conséquences.

### 3.1 Définitions et théorème fondamental

Les polytopes ont été défini au chapitre 2 comme l'enveloppe convexe d'un nombre fini de points. Définissons maintenant un *ensemble polyédrique* comme l'intersection d'un nombre fini de demi-espaces fermés. Contrairement aux polytopes, les ensembles polyédriques ne sont pas toujours bornés. Un demi-plan est un exemple d'ensemble polyédrique. Voici ci-dessous représenté l'ensemble polyédrique de  $\mathbb{R}^2$  correspondant au système :

$$\begin{cases} x_1 & \geq 0, \\ x_2 & \geq 0, \\ x_2 & \geq 2 - x_1. \end{cases}$$



Nous avons alors :

**Théorème 3.1 (Théorème fondamental sur les polytopes)** *Un sous-ensemble de  $\mathbb{R}^d$  est un polytope*

$$P = \text{Conv}(X)$$

*avec  $X$  fini si et seulement si il est un ensemble polyédrique borné*

$$P = \bigcap_{k=1}^n H_k^{\leq}$$

*avec  $H_k$  des hyperplans affins.*

**Démonstration** : Soit

$$Q = \bigcap_{k=1}^n H_k^{\leq}$$

un ensemble polyédrique borné. Alors  $Q$  est un convexe compact et par le théorème de Minkowski (chapitre 2) nous savons que

$$Q = \text{Conv}(\text{ext}(Q))$$

de sorte qu'il suffit de prouver que  $\text{ext}(Q)$  est fini. Soit  $q \in \text{ext}(Q)$  et posons

$$E_q = \{k \in \{1, \dots, n\} \mid q \in H_k\}$$

ainsi que

$$S_q = \{k \in \{1, \dots, n\} \mid q \in H_k^{\lt}\}.$$

Ainsi

$$q \in \bigcap_{k \in E_q} H_k \cap \bigcap_{k \in S_q} H_k^{\lt}$$

et  $E_q \cup S_q = \{1, \dots, n\}$ . Remarquons que  $\bigcap_{k \in E_q} H_k$  est un sous-espace affiné et que  $\bigcap_{k \in S_q} H_k^{\lt}$  est un ouvert de  $\mathbb{R}^d$ . Nous affirmons que

$$\{q\} = \bigcap_{k \in E_q} H_k.$$

En effet si cette intersection contient un point distinct de  $q$ , disons  $r$ , alors elle contient la droite  $\langle q, r \rangle$ , et ainsi elle contient un segment ouvert  $]s, t[$  contenant  $q$ . Or,  $\bigcap_{k \in S_q} H_k^<$  étant ouvert, il contient une boule ouverte  $B(q, \epsilon)$  contenant  $q$ . Mais alors

$$B(q, \epsilon) \cap ]s, t[ \subseteq Q$$

contient un segment ouvert contenant  $q$ , ce qui contredit  $q \in \text{ext}(Q)$ . Nous avons donc prouvé que tout  $q \in \text{ext}(Q)$  peut s'obtenir comme une intersection des  $H_k$ . Ces hyperplans étant en nombre fini, il en résulte que  $\text{ext}(Q)$  est fini.

Soit maintenant  $P = \text{Conv}(X)$  un polytope (avec  $X$  fini). Nous avons déjà démontré que  $P$  est compact au chapitre 2, et en particulier,  $P$  est borné. On suppose  $X$  non vide (sinon il n'y a rien à prouver). Soient  $I_1, \dots, I_m$  des hyperplans tels que

$$\text{Aff}(X) = \bigcap_{i=1}^m I_i.$$

De tels hyperplans existent toujours (théorème 1.5), même si  $\text{Aff}(X) = \mathbb{R}^d$  auquel cas  $m = 0$ . Ainsi nous avons

$$\text{Aff}(X) = \bigcap_{i=1}^m (I_i^{\leq} \cap I_i^{\geq}).$$

Considérons à présent les demi-espaces  $G^{\leq}$  de  $\text{Aff}(X)$  tels que

1. Leur bord  $G$  est un hyperplan affinement engendré par un sous-ensemble  $S \subseteq X$ ,
2. Ils contiennent  $P$ .

Il n'existe qu'un nombre fini de tels demi-espaces vu que  $X$  est fini. De plus, par le corollaire 1.6, chacun des hyperplans  $G$  s'écrit  $G = H \cap \text{Aff}(X)$  avec  $H$  des hyperplans de  $\mathbb{R}^d$  (on peut supposer que  $\dim(\text{Aff}(X)) \geq 1$  sinon  $P$  est réduit à un point et l'énoncé devient trivial). Soit  $H_1, \dots, H_n$  ces hyperplans. Nous allons prouver que

$$P = \bigcap_{i=1}^m (I_i^{\leq} \cap I_i^{\geq}) \cap \bigcap_{k=1}^n H_k^{\leq}.$$

Le polytope  $P$  est clairement inclus dans cette intersection, il ne reste que l'autre inclusion à prouver.

Soit  $q \in \mathbb{R}^d \setminus P$  et montrons que  $q \notin D$  où  $D$  désigne le membre de droite de l'égalité. On peut supposer que  $q \in \text{Aff}(X)$  (sinon  $q \notin I_i$  pour un certain  $i$  et donc forcément  $q \notin D$ .) Donc  $q \in \text{Aff}(X) \setminus P$ . Considérons l'ensemble des hyperplans de  $\text{Aff}(X)$  engendrés par  $q$  et un sous ensemble  $S \subseteq X$  (ceux-ci sont donc en nombre fini,  $X$  étant fini). De plus, un nombre fini d'hyperplans ne pouvant jamais couvrir un ouvert, il existe un élément  $r \in \text{relint}(P)$  évitant tous ces hyperplans.

Considérons à présent le segment de  $[r, q]$ . Comme  $P$  est compact,  $P$  contient son bord relatif (rappelons qu'il s'agit de l'ensemble  $adh(P) \setminus relint(P)$ , voyez le chapitre 1) donc  $q$  n'est pas dans le bord relatif de  $P$ . Ainsi si on considère l'ensemble

$$A = \{\lambda \in [0, 1] \mid (1 - \lambda)r + \lambda q \in intrel(P)\}$$

nous avons  $A \neq \emptyset$  ainsi que

$$0 < sup(A) < 1.$$

Il est alors clair que toute boule de  $Aff(X)$  centrée en  $(1 - \mu)r + \mu q$  avec  $\mu = sup(A)$  rencontre  $P$  et son complémentaire, donc l'élément  $s = \mu r + (1 - \mu)q$  est dans le bord relatif de  $P$ . Soit  $J$  un hyperplan de  $Aff(X)$  séparant  $s$  de  $P$  (un tel hyperplan existe par le théorème de séparation vu que  $s$  n'est pas dans l'intérieur relatif de  $P$ ). Définissons maintenant l'ensemble  $Z$  comme un élément minimal de l'ensemble ordonné

$$(\{Y \subseteq X \mid s \in Aff(Y)\}, \subseteq).$$

L'ensemble  $Z$  est bien défini, vu que l'ensemble ordonné est fini et non vide. Nous prétendons que  $Aff(Z)$  est forcément un hyperplan. En effet, dans tous les cas  $Aff(Z)$  ne peut contenir  $q$  sinon

$$Aff(Z) = Aff(Z \cup q)$$

contiendrait aussi  $r$  (ce point étant sur la droite passant par  $s$  et  $q$ ) et il existerait un hyperplan engendré par  $q$  et un sous-ensemble de  $X$  qui contiennent  $r$ , une contradiction avec la définition de  $r$ . Ainsi si  $Aff(Z)$  n'est pas un hyperplan alors en définissant

$$W = Z \cup \{q\}$$

on a  $dim(Aff(W)) = dim(Aff(Z)) + 1$ , le sous-espace stricte  $Aff(W)$  contenant  $r$ , ce qui livre à nouveau une contradiction avec la définition de  $r$ .

Montrons finalement en utilisant le résultat qui vient d'être établi que

$$J = Aff(J \cap X)$$

ceci impliquant que  $J$  soit égal à un des  $H_k$  précédemment définis, on aura alors  $q \notin H_k^<$  (ce qui terminera la preuve). Nous avons bien sûr  $J \cap X \subseteq X$ , montrons que  $s \in Aff(J \cap X)$  de sorte que  $J \cap X$  doit contenir un ensemble minimal  $Z$  tel que  $s \in Aff(Z)$  ce qui impliquera que  $Aff(Z)$  est un hyperplan (par le raisonnement précédant) et ainsi  $J = Aff(Z)$ . Par définition de  $P$ , on sait que

$$s \in Conv(X) \subseteq Aff(X)$$

et donc  $s = \sum_{i=1}^l \lambda_i x_i$  avec  $\sum_{i=1}^l \lambda_i = 1$ . Si l'hyperplan  $J$  s'écrit

$$J = g^{-1}(\alpha)$$

pour un certaine forme affine  $g$  et un certain réel  $\alpha$ , il en résulte que tous les  $x_i$  sont dans  $J$  (sinon ils seraient dans  $J^<$  et un calcul évident montre qu'il en irait de même pour  $s$ ).

Ceci conclut la preuve du théorème fondamental.  $\square$

Voici deux autres versions de ce théorème, l'un pour les cônes convexes et l'autre pour les ensembles polyédriques. Les démonstrations peuvent être trouvées dans Ziegler [8].

**Théorème 3.2 (Théorème fondamental sur les cônes convexes)** *Un sous-ensemble  $C$  de  $\mathbb{R}^d$  est l'enveloppe conique d'un ensemble fini de points*

$$C = Cone(X)$$

*si et seulement si il est l'intersection d'un nombre fini de demi-espaces*

$$C = \bigcap_{k=1}^n H_k^{\leq}$$

*avec  $H_k$  des hyperplans vectoriels (passant par l'origine).*

**Théorème 3.3 (Théorème fondamental sur les ensembles polyédriques)** *Un sous-ensemble  $P$  de  $\mathbb{R}^d$  est la somme de Minkowski d'un ensemble fini de points  $X$  et de l'enveloppe conique d'un ensemble de points  $Y$*

$$P = Conv(X) + Cone(Y)$$

*si et seulement si il est l'intersection d'un nombre fini de demi-espaces*

$$C = \bigcap_{k=1}^n H_k^{\leq}$$

*avec  $H_k$  des hyperplans affins.*

Introduisons maintenant quelques définitions. Un *demi-espace définissant une facette* (DDF) du polytope  $P$  est un demi-espace  $H^{\leq}$  tel que  $P \subseteq H^{\leq}$ ,  $P \not\subseteq H$  et  $H \cap Aff(P)$  est engendré par des points extrêmes de  $P$ . Un tel hyperplan  $H$  est alors appelé *hyperplan définissant une facette* (HDF). Si de plus le demi-espace s'écrit

$$H^{\leq} = \{x \mid g(x) \leq \alpha\}$$

alors l'inégalité  $g(x) \leq \alpha$  est une *inégalité définissant une facette* de  $P$  (IDF). Notons qu'il ne s'agit pour le moment que de définitions, nous démontrerons plus loin que ces concepts définissent effectivement des facettes du polytope  $P$ . Notons aussi qu'un HDF de  $P$  est forcément d'appui pour  $P$ .

Le polytope  $P$  de  $\mathbb{R}^d$  est dit *plein* s'il est de dimension maximum, c'est-à-dire

$$\dim(P) := \dim(Aff(P)) = d.$$

Il est bien sûr équivalent d'exiger que  $P$  n'est contenu dans aucun hyperplan affini.

La démonstration du théorème fondamental sur les polytopes nous apprend déjà que tout polytope plein est l'intersection de ses DDF, et que ces derniers sont en nombre fini. Nous allons de plus démontrer que :

**Proposition 3.4** *Si  $P$  est un polytope plein de  $\mathbb{R}^d$ , toute intersection de demi-espaces dont l'intersection est  $P$  comprend tous les DDF.*

La démonstration est basée sur le lemme suivant.

**Lemme 3.5** *Si  $C$  est convexe, si  $p \in \text{relint}(C)$  et si  $H$  est un hyperplan d'appui de  $C$  passant par  $p$  alors  $H$  contient  $C$ .*

**Démonstration** : Supposons qu'il existe  $q \in C \setminus H$ . Comme  $p$  est dans l'intérieur relatif de  $C$ , il existe un point  $r \in ]p, q[$  tel que  $]r, 2p - r[ \subseteq C$ , le point  $2p - r$  n'étant rien d'autre que le symétrique de  $r$  par rapport à  $p$  situé sur la droite  $\langle p, q \rangle$ . Mais  $r$  et  $2p - r$  sont situés dans les cotés ouverts distincts de  $H$ , donc  $H$  n'est pas d'appui.  $\square$

**Démonstration de la proposition 3.4** : Soit  $H^\leq$  un DDF de  $P$  et supposons que

$$P = \bigcap_{k=1}^n I_k^\leq$$

pour certains demi-espaces  $I_k^\leq$ . Soit  $E$  un ensemble de points extrême de  $P$  engendrant  $H \cap \text{Aff}(P)$ . Prenons  $p \in \text{relint}(\text{Conv}(E))$  ainsi que  $q$  un point de  $\text{relint}(P)$ . Alors  $q \notin H$  et si  $p \in ]q, r[$  pour un certain point  $r$ , le point  $r$  est forcément hors de  $P$  (sinon  $p$  est dans l'intérieur relatif de  $P$  et on a une contradiction avec le lemme précédent et le fait que  $H$  ne peut contenir  $P$ ).

Ceci implique que  $p \in I_k$  pour au moins un  $k$ , sinon

$$p \in \bigcap_{k=1}^n I_k^\leq$$

et on en déduirait l'existence d'un point  $r$  contenu dans cette intersection et tel que  $p \in ]q, r[$ . Il n'est pas difficile de voir qu'alors, comme  $p$  est combinaison convexe de points de  $E$ , il faut  $E \subseteq I_k$  ce qui livre

$$H = H \cap \text{Aff}(P) = I_k$$

vu que  $E$  engendre  $H$ .  $\square$

Voici enfin une caractérisation des points extrêmes d'un polytope  $P$ .

**Proposition 3.6** *Soit  $p \in P$  un point du polytope plein  $P$ . Alors les affirmations suivantes sont équivalentes*

1.  $p$  est exposé,
2.  $p$  est extrême,
3.  $\{p\}$  est l'intersection d'hyperplans d'appui,



4.  $\{p\}$  est l'intersection de HDF.

**Démonstration** : Nous allons montrer  $1 \Rightarrow 2$ ,  $2 \Rightarrow 4$ ,  $4 \Rightarrow 3$ ,  $3 \Rightarrow 1$ .

Le point 1) implique le point 2) (démontré au chapitre 2 dans le cas d'un convexe quelconque  $C$ ). Le fait que 2) implique 4) a été démontré dans la preuve du théorème fondamental. Comme tout HDF est un hyperplan d'appui, 4) implique 3). Voyons maintenant que 3) implique 1). Soit

$$\{p\} = \bigcap_{k=1}^n H_k$$

avec  $H_k$  des hyperplans d'appui d'équations

$$g_k(x) = \alpha_k.$$

On a donc  $P \subseteq H_k^{\leq}$ . Considérons la forme affine définie par

$$g(x) = \sum_{k=1}^n g_k(x)$$

ainsi que le réel

$$\alpha = \sum_{k=1}^n \alpha_k.$$

L'inégalité  $g(x) \leq \alpha$  est bien sûr valide pour  $P$  et un point  $q \in P$  satisfait  $g(q) = \alpha$  si et seulement si  $g_k(q) = \alpha_k$  pour tout  $k$ , c'est-à-dire si et seulement si  $q$  est dans tous les hyperplans  $H_k$ , c'est-à-dire si et seulement si  $q = p$ . Ainsi  $\{p\}$  est exposé.  $\square$

Les points extrêmes du polytope  $P$  sont appelés les *sommets* de  $P$ . Leur ensemble est noté  $Vert(P)$  (pour *vertex*, en anglais).

Résumons à présent les résultats fondamentaux obtenus pour un polytope  $P = Conv(X)$  (avec  $X$  fini) de  $\mathbb{R}^d$  :

1. Parmi les ensembles  $X$  tels que  $P = Conv(X)$ , il en existe un contenu dans tous les autres et c'est  $Vert(P)$  (ceci résulte du théorème de Minkowski au chapitre 2). Cet ensemble est fini. En particulier,  $P$  est compact.
2. Si  $P$  est plein, la collection de tous les DDF de  $P$  (qui sont en nombre fini) est une collection minimale de demi-espaces dont l'intersection est  $P$ . De plus, tout sommet est intersection de certains de ces DDF.

Ceci implique en particulier que lorsque  $P$  est plein, il existe un seul système d'inéquations affines minimum (c'est-à-dire dont le nombre d'inéquations est minimum) dont l'intersection est  $P$ . Ceci n'est plus vrai lorsque le polytope n'est plus plein, essentiellement car les hyperplans dont il est question au théorème 1.5

ne sont pas uniques (rappelons que ce théorème est utilisé dans la démonstration du théorème fondamental).

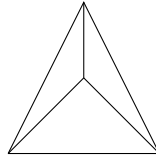
Un système minimum d'inéquations dont l'intersection est  $P$  comporte alors  $d - \dim(\text{Aff}(P))$  équations et autant d'inéquations que de DDF de  $P$  dans  $\text{Aff}(P)$ .

## 3.2 Familles d'exemples

Illustrons à présent quelques familles courantes de polytopes. Les démonstrations des descriptions et propriétés données pour les différents polytopes sont laissées comme exercice au lecteur.

### 1) Simplexes

Les simplexes de  $\mathbb{R}^d$  ont déjà été défini (au chapitre deux) comme l'enveloppe convexe d'un repère affine de  $\mathbb{R}^d$ . Ceux-ci constituent une généralisation des triangles ( $d = 2$ ) et des tétraèdres ( $d = 3$ ).



Si  $S_d = \text{Conv}(\{e_0, \dots, e_d\})$  est un simplexe, alors

$$\text{Vert}(S_d) = \{e_0, \dots, e_d\}$$

et les HDF de  $S_d$  sont les  $d + 1$  hyperplans définis par

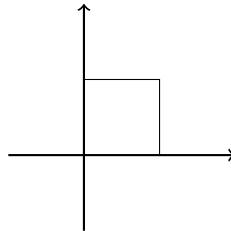
$$H_i = \text{Aff}(\{e_0, \dots, e_d\} \setminus \{e_i\})$$

pour  $i = 0, \dots, d$ .

### 2) Hypercubes

De même que les simplexes généralisent les triangles et les tétraèdres, les hypercubes généralisent les carrés de  $\mathbb{R}^2$  et les cubes de  $\mathbb{R}^3$ . On définit l'*hypercube*  $C_d$  de  $\mathbb{R}^d$  ainsi :

$$C_d = \text{Conv}(\{(x_1, \dots, x_d) \mid x_i = 0 \text{ ou } x_i = 1\}).$$



On a alors

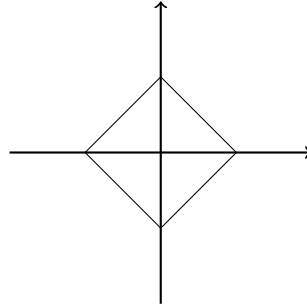
$$\text{Vert}(C_d) = \{(x_1, \dots, x_d) \mid x_i = 0 \text{ ou } x_i = 1\}$$

et les HDF de  $C_d$  sont exactement les hyperplans d'équations  $x_i = 0$  ou  $x_i = 1$  pour  $i = 1, \dots, d$ .

### 3) Hyperoctaèdres

Sans surprise, les *hyperoctaèdres*  $O_d$  généralisent les losanges de  $\mathbb{R}^2$  et les octaèdres de  $\mathbb{R}^3$ . Si  $e_1, \dots, e_d$  désignent les vecteurs de la base canonique de  $\mathbb{R}^d$  on définit :

$$O_d = \text{Conv}(\{\epsilon_i e_i \mid \epsilon_i \in \{-1, 1\}\}).$$



Et on obtient

$$\text{Vert}(O_d) = \{\epsilon_i e_i \mid \epsilon_i \in \{-1, 1\}\}$$

tandis que les HDF de  $O_d$  sont les hyperplans d'équations

$$\sum_{i=1}^d \epsilon_i x_i = 1$$

avec  $\epsilon_i \in \{-1, 1\}$ .

### 4) Polytopes 0/1

Un *polytope 0/1* est un polytope dont les coordonnées de chaque sommet sont 0 ou 1. Il s'agit donc d'un polytope dont l'ensemble des sommets est un sous-ensemble de celui des sommets de l'hypercube (qui est donc lui même un polytope 0/1). Le simplexe de  $\mathbb{R}^d$  engendré par les éléments de la base canonique et le vecteur nul est un autre exemple. Nous décrivons ci-dessous une famille importante de polytope 0/1.

### 5) Polytopes de vecteurs caractéristiques

Soit  $\mathcal{F}$  une collection de sous-ensembles d'un ensemble fini  $E = \{x_1, \dots, x_n\}$ . Chaque élément  $F$  de  $\mathcal{F}$  est encodé par son *vecteur caractéristique* noté  $\chi_F \in \mathbb{R}^n$  et défini par  $(\chi_F)_i = 1$  si  $x_i \in F$  et 0 sinon.

Le *polytope de  $\mathcal{F}$*  est

$$P_{\mathcal{F}} = \text{Conv}(\{\chi_F \mid F \in \mathcal{F}\}).$$

Ces polytopes apparaissent fréquemment dans des problèmes de programmation linéaire et de recherche opérationnelle (nous en reparlerons dans le chapitre suivant). Un cas particulier est le *polytope du voyageur de commerce*, associé au fameux *problème du voyageur de commerce*. Ce problème demande, étant donné  $n$  villes et les distances entre ces villes, de trouver un circuit passant par toutes les villes et dont la somme des distances parcourues est minimum. Si un circuit  $C$  est encodé par un ensemble de paires de villes

$$C = \{(x_1, x_2), (x_2, x_3), \dots, (x_{n-1}, x_n)\}$$

alors l'ensemble des circuits peuvent être encodés par des vecteurs caractéristiques de  $\mathbb{R}^{\frac{n(n-1)}{2}}$ . Il existe en effet

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

paires de villes. Le lecteur vérifiera qu'il existe

$$\frac{(n-1)!}{2}$$

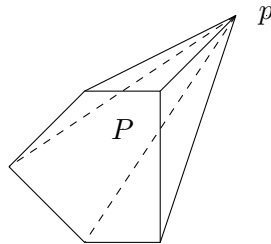
circuits possible, et donc autant de sommets du polytope du voyageur de commerce.

## 6) Pyramides et bipyramides

Considérons un polytope  $P$  de dimension  $d - 1$  dans  $\mathbb{R}^d$  ainsi qu'un point  $p$  hors de  $\text{Aff}(P)$ . Alors

$$\text{Conv}(P \cup \{p\})$$

est la *pyramide de sommet  $p$  et de base  $P$* , notée  $\text{Pyr}(P, p)$ .



On montre alors que :

**Proposition 3.7** Avec les définitions donnés ci-dessus nous avons

$$\text{Vert}(\text{Pyr}(p, P)) = \text{Vert}(P) \cup \{p\}.$$

Les HDF de  $\text{Pyr}(p, P)$  sont soit  $\text{Aff}(P)$ , soit un hyperplan  $H$  passant par  $p$  et tel que  $H \cap \text{Aff}(P)$  est un HDF de  $P$  dans  $\text{Aff}(P)$ .

Soit de plus  $q$  un point tel que  $[q, p] \cap \text{Aff}(P)$  soit un point de  $\text{relint}(P)$ . Alors

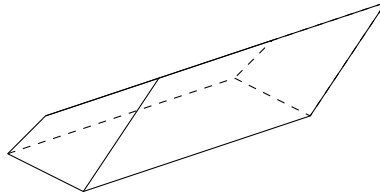
$$\text{Conv}(P \cup \{p, q\})$$

est la *bipyramide* engendrée par  $P$ ,  $p$  et  $q$ , notée  $\text{Bip}(P, p, q)$ .

## 7) Prismes

Soit  $P$  un polytope de dimension  $d - 1$  dans  $\mathbb{R}^d$  et  $v$  un vecteur non parallèle à  $\text{Aff}(P)$  (c'est-à-dire  $\text{Aff}(v)$  n'est parallèle à aucune droite affine de  $\text{Aff}(P)$ ). Alors le *prisme de bases*  $P$  et  $P + v$  est le polytope

$$P(P, v) = \text{Conv}(P \cup \{P + v\}).$$



## 3.3 Résultats généraux élémentaires

Nous savons déjà qu'un ensemble convexe est un polytope (ensemble convexe d'un ensemble fini de points) si et seulement si il est un ensemble polyédrique borné, qu'un polytope est toujours compact, que la somme de Minkowski et deux polytopes est toujours un polytope, et que l'image d'un polytope par une application affine (en particulier, par une translation) est toujours un polytope. Nous démontrons ici quelques autres propriétés similaires.

La démonstration de la propriété ci-dessous illustre bien l'utilité des deux représentations possibles d'un polytope.

**Proposition 3.8** Soient  $P$  et  $Q$  deux polytopes de  $\mathbb{R}^d$ . Alors les ensembles suivants sont aussi des polytopes :

1.  $P \times Q = \{(p, q) \in \mathbb{R}^{2d} \mid p \in P, q \in Q\}$ ,
2.  $P \cap Q$ ,
3.  $\text{Conv}(P \cup Q)$ .

**Démonstration** : Si

$$P = \text{Conv}(\{p_1, \dots, p_n\})$$

et

$$Q = \text{Conv}(\{q_1, \dots, q_m\})$$

alors il est facile de vérifier que

$$P \times Q = \text{Conv}(\{(p_i, q_j) \mid 1 \leq i \leq n, 1 \leq j \leq m\})$$

de sorte que  $P \times Q$  est bien un polytope.

De même si

$$P = \bigcap_{k=1}^n H_k^{\leq}$$

et

$$Q = \bigcap_{l=1}^m I_l^{\leq}$$

sont deux ensembles polyédriques bornés alors il en va de même pour

$$P \cap Q = \bigcap_{k=1}^n H_k^{\leq} \cap \bigcap_{l=1}^m I_l^{\leq}.$$

Enfin, par la proposition 2.10 nous savons que

$$\text{Conv}(P \cup Q) = \text{Conv}(\{p_1, \dots, p_n\} \cup \{q_1, \dots, q_m\})$$

et la preuve est terminée.  $\square$

Notons que le point 1 de la proposition est valable pour deux polytope  $P \subseteq \mathbb{R}^d$  et  $Q \subseteq \mathbb{R}^e$ , les naturels  $d$  et  $e$  non nécessairement égaux. Dans ce cas  $P \times Q \subseteq \mathbb{R}^{d+e}$ .

**Proposition 3.9** *Tout polytope  $P$  de  $\mathbb{R}^d$  est l'image d'un simplexe de  $\mathbb{R}^n$  par une application affine, avec  $n = |\text{Vert}(P)| - 1$ .*

**Démonstration** : Soit  $R$  un repère affiné de  $\mathbb{R}^n$ . On peut démontrer qu'il existe une application affine  $f$  envoyant  $R$  sur les sommets de  $P$ . Le résultat suit alors car

$$f(\text{Conv}(R)) = \text{Conv}(f(R)) = P.$$

$\square$

### 3.4 Le treillis des faces

L'ensemble des faces d'un polytope muni de la relation d'inclusion appartient à une famille importante d'ensembles ordonnés, les treillis. Rappelons brièvement leur définition. Soit  $(X, \leq)$  un ensemble ordonné. Deux éléments  $x, y \in X$  admettent l'élément  $s \in X$  comme *borne supérieure* si

1.  $s$  majore  $\{x, y\}$ ,
2. Pour tout majorant  $t$  de  $\{x, y\}$ , on a  $s \leq t$ .

La borne supérieure (on parle aussi de *suprémum*) de  $x$  et  $y$  est donc le minimum de l'ensemble des majorants de  $\{x, y\}$ , pour autant que ce minimum existe. De même, la *borne inférieure* (ou l'*infimum*) de  $x$  et  $y$  est le maximum de l'ensemble des minorants de  $\{x, y\}$ , pour autant que celui-ci existe. Si toute paire d'éléments admet un infimum et un suprémum, l'ensemble ordonné est appelé un *treillis*. Le suprémum de  $x$  et  $y$  est alors noté  $x \vee y$  tandis que l'infimum est noté  $x \wedge y$ . Le lecteur peut vérifier en guise d'exercice que les ensembles ordonnés suivants sont des treillis, les bornes supérieures et inférieures étant indiquées à côté :

1. Tout ensemble totalement ordonné, le suprémum de  $x$  et  $y$  étant le maximum de  $\{x, y\}$  et l'infimum le minimum,
2. L'ensemble  $\mathbb{N}$  muni de la relation de divisibilité, le suprémum étant le PPCM et l'infimum le PGCD,
3. L'ensemble  $(\mathcal{P}(E), \subseteq)$  des parties d'un ensemble  $E$ , le suprémum étant la réunion des deux ensembles et l'infimum leur intersection,
4. L'exemple précédant peut être facilement généralisé. Soit  $\mathcal{F}$  une collection de sous-ensemble de  $E$  contenant  $E$  et stable par intersection, à savoir

$$A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}.$$

Alors  $(\mathcal{F}, \subseteq)$  forme un treillis avec

$$A \wedge B = A \cap B$$

et

$$A \vee B = \bigcap_{C \in \mathcal{F}, A \cup B \subseteq C} C.$$

Notons que l'intersection ci-dessus est forcément non vide, vu que nous avons exigé  $E \in \mathcal{F}$ .

5. La famille de treillis introduite au point précédant contient de nombreux exemples qui se rencontrent fréquemment en algèbre et en géométrie : l'ensemble des sous-groupes d'un groupe, des sous-anneaux d'un anneau, des sous-espaces vectoriels d'un espace vectoriel, des sous-espaces affins d'un espace affins, des ensembles convexes de  $\mathbb{R}^d$ , ...

Se donner un ensemble de parties d'un ensemble  $E$  contenant  $E$  et stable par intersection est en fait équivalent à se donner un opérateur de fermeture sur  $E$ . Nous n'en dirons pas plus à ce sujet pour ne pas trop nous écarter du sujet qui nous préoccupe ici (les polytopes), mais le lecteur intéressé consultera l'ouvrage de Davey et Priestley [2].

Comme cas particulier de l'exemple 4 ci-dessus il vient :

**Proposition 3.10** *Soit  $P$  un polytope de  $\mathbb{R}^d$ . Alors l'ensemble  $\mathcal{F}(P)$  des faces de  $P$  muni de l'inclusion forme un treillis avec*

$$A \wedge B = A \cap B$$

et

$$A \vee B = \bigcup \{K \in \mathcal{F}(P) \mid A \cup B \subseteq K\}.$$

**Démonstration** : Nous avons déjà démontré que l'ensemble des faces d'un convexe est stable par intersection. Cet ensemble contient bien sûr  $P$  (par définition d'une face).  $\square$

La proposition précédente a été énoncée dans le cas des polytopes, mais la démonstration montre qu'elle reste vraie dans le cas d'un convexe quelconque. Cependant, alors qu'un convexe quelconque peut admettre une infinité de faces, pour un polytope nous avons :

**Proposition 3.11** *Soit  $F$  une face du polytope  $P \subseteq \mathbb{R}^d$ . Alors  $F$  est encore un polytope, et*

$$\text{vert}(F) = F \cap \text{vert}(P).$$

*De plus, l'ensemble des faces de  $P$  est fini.*

**Démonstration** : Comme  $F$  est l'intersection de  $P$  avec un hyperplan d'appui, il est évident que si  $P$  est une intersection bornée de demi-espaces alors  $F$  aussi. L'ensemble des sommets étant l'ensemble des points extrêmes (par définition), l'égalité ci-dessus n'est rien d'autre que le lemme 2.31 appliqué aux polytopes. Enfin, comme

$$F = \text{Conv}(\text{Vert}(F))$$

et qu'il n'existe qu'un nombre fini de possibilités pour  $\text{Vert}(F)$  (vu que  $\text{Vert}(P)$  est fini), le nombre de faces est clairement fini.  $\square$

L'ensemble des faces d'un polytope  $P$  est appelé le *treillis des faces* de  $P$ . Deux polytopes  $P$  et  $Q$  ont *même type combinatoire* lorsque leurs treillis des faces  $(\mathcal{F}(P), \subseteq)$  et  $(\mathcal{F}(Q), \subseteq)$  sont isomorphes. La définition des isomorphismes d'ensembles ordonnés est assez évidente : deux ensembles ordonnés  $(X, \leq_1)$  et  $(Y, \leq_2)$  (et donc en particulier deux treillis) sont isomorphes lorsqu'il existe une bijection

$$f : X \rightarrow Y$$

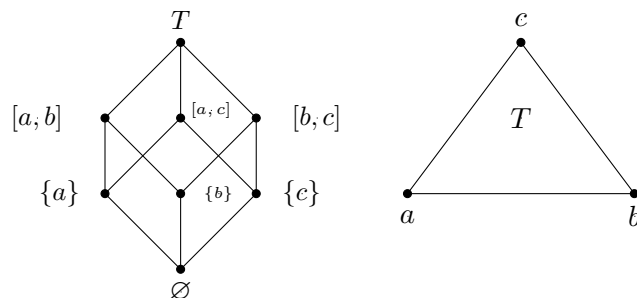
telle que

$$x \leq_1 y \Leftrightarrow f(x) \leq_2 f(y).$$

Par exemple, dans  $\mathbb{R}^2$ , les carrés, les losanges, et les rectangles ont tous même type combinatoire. Dessinons ci-dessous le treillis des faces d'un triangle de  $\mathbb{R}^2$ . Rappelons que les ensembles ordonnés finis sont représentés au moyen de leur *diagramme de Hasse* (du nom du mathématicien Helmut Hasse (1898 - 1979)), à savoir : chaque élément  $x$  de l'ensemble ordonné est représenté par un point  $p_x$  du plan, si  $x < y$  alors  $p_x$  est situé plus bas que  $p_y$  (la notation  $x < y$  signifie bien sûr  $x \leq y$  et  $x \neq y$ ), le point  $p_x$  et relié par un segment de droite à  $p_y$  si seulement



si  $x < y$  et il n'existe aucun  $z$  tel que  $x < z < y$  (un tel segment de droite ne contenant alors aucun autre  $p_z$  avec  $z \neq x$  et  $z \neq y$ ). Avec ces conventions,  $x < y$  si et seulement si il existe un chemin ascendant de  $x$  à  $y$  dans le diagramme. Voici donc le diagramme de Hasse de  $\mathcal{F}(\text{Triangle})$  :



Rappelons que les facettes de  $P$  sont les faces de dimension maximales (c'est-à-dire de dimension  $\dim(P)-1$ ) de  $P$ . La proposition suivante justifie la terminologie HDF, DDF, IDF.

**Proposition 3.12** Soit  $P$  un polytope de  $\mathbb{R}^d$ . Alors les facettes de  $P$  sont les intersections de  $P$  avec les HDF.

**Démonstration** : Soit  $H$  un HDF de  $P$ . Alors  $H$  est clairement d'appui pour  $P$ , donc  $H \cap P$  est une face de  $P$ . Comme  $H \cap \text{Aff}(P)$  est engendré par des points extrêmes de  $P$  on a

$$H \cap \text{Aff}(P) = \text{Aff}(H \cap P)$$

et  $H \cap \text{Aff}(P)$  est hyperplan de  $\text{Aff}(P)$  donc

$$\dim(\text{Aff}(H \cap P)) = \dim(P) - 1$$

ce qui montre bien que  $H \cap P$  est une facette de  $P$ .

Réciproquement, soit

$$P = \bigcap_{i=1}^m I_i \cap \bigcap_{k=1}^n H_k^{\leq}$$

où les  $I_k$  sont des hyperplans et les  $H_k$  sont les DDF. Soit  $F = H \cap P$  une facette de  $P$  (pour un certain hyperplan  $H$ ) ainsi que  $p \in F$ . Par le lemme 3.5 nous savons que  $p \notin \text{relint}P$  (sinon  $p$  ne serait pas sur  $F$ ). Donc il existe  $j$  tel que  $p \in H_j^<$ . Ceci implique en fait  $F \subseteq H_j$  (vu que  $F \subseteq H_j^{\leq}$  et  $p \in \text{relint}(F)$ ). Comme

$$\dim(H \cap P) = \dim(P) - 1 = \dim(H_j \cap P)$$

il faut

$$\dim(H \cap P) = \dim(H_j \cap P)$$

et donc  $H \cap P = H_j \cap P$ .  $\square$

**Proposition 3.13** *Toute face de  $P$  de dimension  $\dim(P) - 2$  est contenue dans exactement deux facettes.*

**Théorème 3.14** *Toute face de  $P$  est intersection de facettes.*

Notons en passant que :

**Proposition 3.15** *Soit  $P \subseteq \mathbb{R}^d$  un polytope et  $1 \leq k \leq \dim(P)$ . Si  $F$  est une face de dimension  $k$  de  $P$ , alors :*

$$|\text{Vert}(F)| \geq k + 1.$$

**Démonstration :** Si  $|\text{Vert}(F)| \leq k$  alors

$$\dim(\text{Aff}(F)) = \dim(\text{Aff}(\text{Conv}(\text{Vert}(F)))) = \dim(\text{Aff}(\text{Vert}(F))) \leq k - 1$$

d'où la contradiction.  $\square$

Les polytopes tels que toute face de dimension  $k$  (pour  $1 \leq k < \dim(P)$ ) possède exactement  $k + 1$  sommets sont appelés *polytopes simpliciaux*. Il est bien sûr équivalent de demander que toute face propre  $F$  du polytope soit un simplexe. En fait :

**Proposition 3.16** *Un polytope  $P \subseteq \mathbb{R}^d$  est simplicial si et seulement si toutes ses facettes sont des simplexes.*

**Démonstration :** Si  $P$  est simplicial alors bien sûr toute facette de  $P$  est un simplexe par définition. Si toute facette de  $P$  est un simplexe et si  $F$  est une face propre de  $P$  alors il existe une facette  $G$  telle que  $F \subseteq G$ . En particulier,  $\text{Vert}(F) \subseteq \text{Vert}(G)$ , donc les points de  $\text{Vert}(F)$  sont affinement indépendants. Comme  $F = \text{Conv}(\text{Vert}(F))$  et  $\text{Aff}(F) = \text{Aff}(\text{Vert}(F))$ , il suit que  $F$  est un simplexe.  $\square$

**Corollaire 3.17** *Les simplexes sont simpliciaux.*

Dans la section suivante nous étudierons une autre classe de polytopes simpliciaux, les polytopes cycliques.

Nous appellerons *arête* de  $P$  une face de dimension 1. Deux sommets seront dit *adjacents* s'ils sont les sommets d'une arête. Le *graphe* de  $P$  est le graphe

$$G(P) = (\text{Vert}(P), E)$$

où  $E$  est l'ensemble des couples  $(p, q)$  de sommets adjacents.

### 3.5 Polytopes Cycliques

Introduisons à présent une classe de polytopes ayant de jolies propriétés. Définissons pour ce faire la *courbe des moments*. Il s'agit de la courbe de  $\mathbb{R}^d$  paramétrée par

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^d : t \rightarrow (t, t^2, \dots, t^d).$$

Soient  $n \geq d + 1$  ainsi que  $t_1, \dots, t_n$  des réels distincts. Alors le *polytope cyclique* déterminé par  $t_1, \dots, t_n$  est par définition :

$$C_d(t_1, \dots, t_n) = \text{Conv}(\{\gamma(t_i) \mid i = 1, \dots, n\}).$$

Nous allons prouver que ce polytope est de dimension  $d$  et que son type combinatoire ne dépend que de  $d$  et  $n$  (ce qui fait que ce polytope est souvent abusivement noté  $C_d(n)$ ).

**Proposition 3.18** *Tout hyperplan de  $\mathbb{R}^d$  contient au plus  $d$  points de  $\gamma(\mathbb{R})$ .*

**Démonstration** : Si un hyperplan  $H$  a pour équation

$$\sum_{i=1}^d \alpha_i x_i = \alpha$$

alors il est clair qu'il existe au plus  $d$  valeurs de  $t$  telles que

$$\sum_{i=1}^d \alpha_i t^i = \alpha$$

ce qui termine la preuve.  $\square$

On déduit de la proposition précédente la résultat :

**Corollaire 3.19** *Soient  $t_1, \dots, t_n$  des réels distincts, avec  $n \leq d + 1$ . Alors les points*

$$\gamma(t_1), \dots, \gamma(t_n)$$

*sont affinement indépendants.*

**Démonstration** : Si ces points sont affinement dépendants, soient  $A$  un sous-espace affine de dimension au plus  $n - 2$  qui les contient. Si  $n = d + 1$  ceci est contradictoire avec la proposition précédente. Sinon, soient

$$t_{n+1}, \dots, t_{d+1}$$

des points tels que  $t_i \neq t_j$  pour  $1 \leq i, j \leq d + 1$  et  $i \neq j$ . Alors les points

$$\gamma(t_1), \dots, \gamma(t_{d+1})$$

sont distincts et contenus dans un espace affín  $B$  de dimension au plus

$$(n - 2) + (d + 1 - n) = d - 1$$

ce qui livre à nouveau une contradiction.  $\square$

Et par conséquent :

**Corollaire 3.20** *Quels que soient les réels distincts  $t_1, \dots, t_n$  nous avons*

$$\dim(C_d(t_1, \dots, t_n)) = d.$$

Déterminons maintenant les sommets du polytope.

**Théorème 3.21** *Les sommets de  $C_d(t_1, \dots, t_n)$  sont*

$$\text{Vert}(C_d(t_1, \dots, t_n)) = \{\gamma(t_1), \dots, \gamma(t_n)\}.$$

**Démonstration** : Les sommets du polytope sont certainement inclus dans le membre de droite par la définition de  $C_d(t_1, \dots, t_n)$  et le théorème de Minkowski. Nous allons montrer l'autre inclusion en montrant que  $\gamma(t_i)$  est exposé pour tout  $i = 1, \dots, n$ .

Définissons le polynôme  $p_i$  ainsi :

$$p_i(t) = -(t - t_i)^2 = -t^2 + 2tt_i - t^2.$$

Définissons de plus le vecteur  $y_i \in \mathbb{R}^d$  comme

$$y_i = (2t_i, -1, 0, \dots, 0).$$

Alors

$$p_i(t) = \gamma(t) \cdot y_i - t_i^2$$

ou  $\cdot$  désigne le produit scalaire usuel dans  $\mathbb{R}^d$ . Ceci implique d'une part que l'hyperplan  $H_i$  de  $\mathbb{R}^d$  d'équation

$$y_i \cdot x = t_i^2$$

est d'appui pour  $C$  vu que  $p_i(t) \leq 0$  pour tout  $t$ , et que

$$H_i \cap \{\gamma(t_1), \dots, \gamma(t_n)\} = \{\gamma(t_i)\}$$

vu que  $p_i(t) = 0$  si et seulement si  $t = t_i$ . Mais comme

$$C_d(t_1, \dots, t_n) = \text{Conv}(\{\gamma(t_1), \dots, \gamma(t_n)\})$$

on a clairement

$$H_i \cap C_d(t_1, \dots, t_n) = \{\gamma(t_i)\}.$$

Donc  $\gamma(t_i)$  est exposé.  $\square$

On en déduit que les polytopes cycliques sont simpliciaux, en effet :

**Théorème 3.22** *Tout polytope cyclique est simplicial.*

**Démonstration** : Il suffit de montrer que toute facette contient exactement  $d$  sommets. Comme elles en contiennent au moins  $d$ , que les sommets sont des points de la courbe des moments et qu'un hyperplan ne peut jamais contenir plus de  $d$  points de la courbe, le théorème est démontré.  $\square$

Le prochain résultat caractérise les facettes de  $C_d(t_1, \dots, t_n)$ . Il est nommé d'après David Gale (1921-2008) :

**Théorème 3.23 (Condition de parité de Gale)** *Soit  $C_d(t_1, \dots, t_n)$  un polytope cyclique de  $\mathbb{R}^d$  avec*

$$t_1 < \dots < t_n.$$

*Alors le sous-ensemble de  $d$  sommets*

$$S = \{\gamma(t_{l_1}), \dots, \gamma(t_{l_d})\}$$

*définit une facette du polytope si et seulement si pour tous*

$$i, j \in \{1, \dots, n\} \setminus \{l_1, \dots, l_d\}$$

*le nombre d'indices*

$$k \in \{l_1, \dots, l_d\}$$

*tels que  $i < k < j$  est pair.*

Avant de démontrer ce théorème, mentionnons le résultat classique suivant (la démonstration constitue un exercice d'algèbre linéaire) :

**Théorème 3.24 (Identité de Vandemonde)** *Si  $t_0, \dots, t_d$  sont  $d+1$  réels alors*

$$\det \begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^d \\ 1 & t_1 & t_1^2 & \dots & t_1^d \\ 1 & t_2 & t_2^2 & \dots & t_2^d \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & t_d & t_d^2 & \dots & t_d^d \end{bmatrix} = \prod_{0 \leq i < j \leq d} (t_j - t_i).$$

**Démonstration du théorème 3.23** : L'hyperplan engendré par les  $d$  points de  $S$  à donc pour équation

$$\det \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \\ 1 & t_{l_1} & t_{l_1}^2 & \dots & t_{l_1}^d \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & t_{l_d} & t_{l_d}^2 & \dots & t_{l_d}^d \end{bmatrix} = 0.$$

Si les sommets de  $S$  définissent une facette, cet hyperplan doit être un HDF, et en particulier le polytope doit se trouver dans un des demi-espaces larges définis par celui-ci. Utilisant l'identité de Vandermonde, nous voyons que le membre de gauche évalué en un point  $\gamma(t)$  de la courbe des moments est exactement

$$\prod_{j=1}^d (t_{l_j} - t) \cdot \prod_{1 \leq i < j \leq d} (t_{l_j} - t_{l_i})$$

et cette expression doit avoir le même signe en les points

$$\{\gamma(t_k) \mid k \in \{1, \dots, n\} \setminus \{l_1, \dots, l_d\}\}.$$

Vu que  $t_{l_j} > t_{l_i}$  pour  $j > i$  le signe de cette expression est égal au signe de

$$p(t) = \prod_{j=1}^d (t_{l_j} - t) = (t_{l_1} - t) \dots (t_{l_d} - t).$$

Mais  $p(t)$  est un polynôme de degré  $d$  en  $t$  admettant  $d$  racines, le signe de  $p(t)$  change chaque fois que  $t$  passe par un des  $t_{l_i}$ . La condition sera donc satisfaite exactement lorsque pour deux indices  $i, j \in \{1, \dots, n\} \setminus \{l_1, \dots, l_d\}$ , le nombre d'indices  $k \in \{l_1, \dots, l_d\}$  tels que  $i < k < j$  est pair.  $\square$ .

Le résultat annoncé en début de section suit :

**Corollaire 3.25** *Le type combinatoire de  $C_d(t_1, \dots, t_n)$  ne dépend que de  $d$  et  $n$ .*

**Démonstration** : Il suffit d'utiliser la condition de parité de Gale et le fait que toute face est intersection de facettes.  $\square$

Un polytope  $P \subseteq \mathbb{R}^d$  est dit  $k$ -serré pour un naturel  $0 \leq k \leq d$  lorsque tout ensemble formé d'au plus  $k$  de ses sommets est toujours l'ensemble des sommets d'une face de  $P$ . Pour  $k = 0$  ou  $k = 1$ , tout polytope est bien sûr  $k$ -serré. Il est en fait possible de démontrer que tout polytope  $k$ -serré avec  $k > \lfloor \frac{d}{2} \rfloor$  est un simplexe (où  $\lfloor x \rfloor$  désigne le plancher du réel  $x$ , voyez [1]), de sorte que la notion n'a pas beaucoup d'intérêt pour de tels  $k$ . En ce qui concerne les polytopes cycliques :

**Proposition 3.26** *Tout polytope cyclique est  $\lfloor \frac{d}{2} \rfloor$ -serré.*

**Démonstration** : Soient  $\gamma(t_{i_1}), \dots, \gamma(t_{i_k})$  des sommets de  $C_d(t_1, \dots, t_n)$  avec  $1 \leq k \leq \lfloor \frac{d}{2} \rfloor$ . Il suffit d'exhiber un hyperplan d'appui contenant exactement ces sommets. Choisissons un réel  $\epsilon > 0$  suffisamment petit pour que pour tout  $1 \leq i < n$  on ait

$$t_i < t_i + \epsilon < t_{i+1}.$$

Choisissons de plus un certain réel  $M > t_n + \epsilon$ . Posons  $s_{i_k} = t_{i_k} + \epsilon$ . Alors l'hyperplan  $H$  engendré par les  $d$  points

$$\gamma(t_{i_1}), \gamma(s_{i_1}), \gamma(t_{i_2}), \gamma(s_{i_2}), \dots, \gamma(t_{i_k}), \gamma(s_{i_k}), \gamma(M+1), \dots, \gamma(M+d-2k)$$

est un hyperplan approprié. En effet les sommets de  $C_d(t_1, \dots, t_n)$  contenus dans  $H$  sont exactement  $\gamma(t_{i_1}), \dots, \gamma(t_{i_k})$  (vu que  $H$  ne peut contenir plus de  $d$  points de la courbe des moments) et en utilisant un argument similaire à celui du théorème précédant, on montre que deux sommets de  $C_d(t_1, \dots, t_n)$  sont toujours du même côté de  $H$ .  $\square$

Mentionnons sans démonstration le théorème important suivant (le lecteur intéressé par la démonstration consultera [1] ou bien [8]) :

**Théorème 3.27 (Théorème de la borne supérieure)** *Si  $P \subseteq \mathbb{R}^d$  est un polytope plein avec  $n$  sommets, alors le nombre de faces de dimension  $k$  de  $P$  (pour  $0 \leq k \leq d$ ) est au plus égal au nombre de faces de dimension  $k$  de  $C_d(n)$ .*

Ce théorème a été prouvé par MacMullen en 1970.

## Chapitre 4

# Systemes d'inégalités linéaires

Le fait qu'un polytope puisse être vu comme intersection d'un nombre fini de demi-espaces (eux-mêmes décrits par des inégalités linéaires) nous incite à nous pencher sur les systèmes d'inégalités linéaires.

### 4.1 Le théorème de Fredholm

Rappelons en quelques lignes le résultat principal de la théorie des systèmes d'équations linéaires. Soit  $Ax = b$  un tel système, où  $A \in \mathbb{R}^{m \times n}$  est la matrice des coefficients,  $b \in \mathbb{R}^{m \times 1}$  est le vecteur colonne des termes indépendants et  $x \in \mathbb{R}^{n \times 1}$  est le vecteur colonne des inconnues. Si  $W$  désigne l'ensemble des solutions de  $Ax = 0$  alors  $W$  est un sous-espace de  $\mathbb{R}^n$  (vu qu'il s'agit d'une intersection d'hyperplans). Si  $S$  désigne l'ensemble des solutions de  $Ax = b$ , alors pour tout  $p \in S$  il vient

$$S = p + W$$

ce qui fait de  $S$  un sous-espace affine de  $\mathbb{R}^n$ . Résoudre le système  $Ax = b$  signifie donc soit prouver  $S = \emptyset$  soit trouver un  $p \in S$  et une base de  $W$ .

Le théorème suivant est dû au mathématicien Erik Ivar Fredholm (1866-1927). On note comme d'habitude  $u^t$  le vecteur transposé du vecteur  $u$ .

**Théorème 4.1 (Théorème de Fredholm)** *Le système  $Ax = b$  admet une solution si et seulement si pour tout  $u \in \mathbb{R}^{m \times 1}$  :*

$$u^t A = 0 \Rightarrow u^t b = 0.$$

**Démonstration** : Si  $x$  est une solution du système  $Ax = b$ , alors pour tout  $u$  tel que  $u^t A = 0$  il vient

$$u^t b = u^t (Ax) = (u^t A)x = 0$$

donc une des implications du théorème est prouvée.

Etablissons maintenant la contraposée de l'autre implication. Considérons l'application linéaire

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \rightarrow Ax.$$



Comme nous supposons que le système n'admet pas de solution, nous savons que  $b$  n'est pas dans l'image de  $f$ . L'image de  $f$  étant un sous-espace de  $\mathbb{R}^m$ , il existe un hyperplan vectoriel  $H$  tel que  $Im(f) \subseteq H$  et  $b \notin H$ . Soit  $c^t y = 0$  l'équation de cet hyperplan. Alors  $c^t A = 0$  (car chaque colonne de  $A$  est dans  $Im(f)$ ) et  $c^t b \neq 0$ . Ceci termine la preuve.  $\square$

Ce théorème est souvent appelé *théorème de l'alternative* car exactement un des deux cas suivants se produit :

- $\exists x \in \mathbb{R}^{n \times 1} : Ax = b$ ,
- $\exists u \in \mathbb{R}^{m \times 1} : u^t A = 0$  et  $u^t b \neq 0$ .

## 4.2 Le lemme de Farkas

Le théorème suivant est l'analogie du théorème de Fredholm dans le cas où les égalités sont remplacées par des inégalités. Il s'agit d'une des versions du lemme de Farkas, nommé d'après le mathématicien Farkas Gyula (1847 - 1930). Le lecteur désirent plus de détails sur les autres versions de ce lemme ira voir [8].

**Théorème 4.2 (Lemme de Farkas)** *L'ensemble polyédrique  $Ax \leq b$  est non vide si et seulement si pour tout  $u \in \mathbb{R}^{m \times 1}$  :*

$$u \geq 0 \text{ et } u^t A = 0 \Rightarrow u^t b \geq 0.$$

**Démonstration** : Soit  $p$  un point de l'ensemble polyédrique. Alors pour tout  $u \geq 0$  tel que  $u^t A = 0$  on a

$$u^t b \geq u^t A p = (u^t A) p = 0 p = 0$$

ce qui prouve une des implications.

Pour l'autre implication, nous considérons à nouveau l'application linéaire

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \rightarrow Ax$$

et nous travaillerons avec le dominant de l'image de  $f$  qui est défini comme

$$Im(f)^\downarrow = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : Ax \leq y\}$$

ou encore

$$Im(f)^\downarrow = \{y \in \mathbb{R}^m \mid \exists t \in Im(f) : t \leq y\}.$$

Il s'agit d'un ensemble convexe. En effet soient  $y, z \in Im(f)^\downarrow$ , avec  $x$  et  $x'$  tels que  $Ax \leq y$  et  $Ax' \leq z$ . Si  $\lambda y + \mu z$  est une combinaison convexe de  $y$  et  $z$  alors

$$A(\lambda x + \mu x') = \lambda Ax + \mu Ax' \leq \lambda y + \mu z$$

donc  $\lambda y + \mu z$  est bien dans  $Im(f)^\downarrow$ .

Supposons que l'ensemble polyédrique est vide, c'est-à-dire que  $b \notin \text{Im}(f)^\downarrow$ . Il n'est pas difficile de voir que  $\text{Im}(f)^\downarrow$  est fermé, de sorte que  $\mathbb{R}^d \setminus \text{Im}(f)^\downarrow$  est ouvert. Il existe ainsi une boule ouverte  $B$  centrée en  $b$  est disjointe de  $\text{Im}(f)^\downarrow$ . Soit maintenant  $H$  un hyperplan affín séparant  $B$  de  $\text{Im}(f)^\downarrow$ . On a bien sûr  $\text{Im}(f) \subseteq \text{Im}(f)^\downarrow$ . Comme  $\text{Im}(f)$  est un sous-espace de  $\mathbb{R}^d$ , il faut soit  $\text{Im}(f) \cap H = \emptyset$  soit  $\text{Im}(f) \subseteq H$ . Dans les deux cas, on déduit l'existence d'un hyperplan  $I$  de  $\mathbb{R}^d$  parallèle à  $H$  contenant  $\text{Im}(f)$  et tel que  $b \in I^<$  et  $\text{Im}(f)^\downarrow \subseteq I^\geq$ . Si l'équation de cet hyperplan est  $c^t \cdot y = 0$  alors  $c^t A = 0$  (chaque colonne de  $A$  étant dans  $\text{Im}(f)$ ),  $c^t b < 0$  et  $c \geq 0$  (POURQUOI??) ce qui contredit les hypothèses.  $\square$

### 4.3 La méthode de Fourier-Motzkin

Cette section est destinée à une très brève présentation de la méthode de Fourier-Motzkin dont le but est de déterminer si le système  $Ax \leq b$  admet une solution. Illustrons d'abord celle-ci sur un exemple dans  $\mathbb{R}^2$ . Nous aimerions voir si le système suivant admet une solution :

$$\begin{cases} -x + y & \leq 0 \\ y & \leq 1 \\ -x - 3y & \leq -3 \\ x + y & \leq 4 \\ x - 2y & \leq 0 \end{cases}$$

En vue d'éliminer  $y$ , nous réécrivons le système ainsi :

$$\begin{cases} y & \leq x \\ y & \leq 1 \\ 1 - \frac{x}{3} & \leq y \\ y & \leq 4 - x \\ \frac{x}{2} & \leq y \end{cases}$$

Une telle valeur de  $y$  existe si et seulement si le système suivant admet une solution :

$$\begin{cases} 1 - \frac{x}{3} \leq x \\ 1 - \frac{x}{3} \leq 1 \\ 1 - \frac{x}{3} \leq 4 - x \\ \frac{x}{2} \leq x \\ \frac{x}{2} \leq x \\ \frac{x}{2} \leq 4 - x \end{cases}$$

C'est-à-dire :

$$\begin{cases} \frac{3}{4} \leq x \\ 0 \leq x \\ x \leq \frac{9}{2} \\ x \leq 2 \\ x \leq \frac{8}{3} \end{cases}$$

Les  $x$  qui conviennent sont ceux de l'intervalle  $[\frac{3}{4}, 2]$ . Pour  $x = 1$  par exemple, on obtient pour  $y$  :

$$\begin{cases} y \leq 1 \\ \frac{2}{3} \leq y \\ y \leq 3 \\ \frac{1}{2} \leq y \end{cases}$$

Et les  $y$  qui conviennent sont ceux de l'intervalle  $[\frac{2}{3}, 1]$ .

De façon générale, une étape de la méthode consiste à éliminer la variable  $x_l$  en écrivant les inéquations suivantes :

- Celles ne contenant pas  $x_l$  (avec  $a_{il} = 0$ ),
- Puis celles avec  $a_{jl} > 0$  sous la forme

$$x_l \leq \frac{1}{a_{jl}} (b_j - \sum_{t=1, t \neq l}^n a_{jt} x_t) := B_j,$$

- Puis celles avec  $a_{kl} < 0$  sous la forme

$$x_j \geq \frac{1}{a_{kl}} (b_k - \sum_{t=1, t \neq l}^n a_{kt} x_t) := C_k.$$

Le système aura une solution si et seulement si le système obtenu avec les inégalités suivantes admet une solution :

- Celles ne contenant pas  $x_l$ ,
- $C_k \leq B_j$  pour tous  $k, j$ ,

ce système étant "plus simple" en ce sens que la variable  $x_l$  n'apparaît pas. Notons cependant que le nombre d'inégalités peut augmenter assez vite. Pour un système comportant  $m$  inégalités, on peut ainsi passer à un système comportant

$$\frac{m}{2} \cdot \frac{m}{2} = \frac{m^2}{4}$$

inégalités.

## 4.4 Programmation linéaire

Un *programme linéaire* est un problème où il est question d'optimiser une fonction linéaire soumise à des contraintes linéaires, c'est-à-dire :

$$\max\{c^t x \mid x \in \mathbb{R}^n, Ax \leq b\}$$

avec  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^{m \times 1}$  et  $c \in \mathbb{R}^{n \times 1}$ . Le *max* peut être remplacé par un *min*. Géométriquement, il s'agit d'optimiser la fonction linéaire

$$f : \mathbb{R}^n \rightarrow \mathbb{R} : x \rightarrow c^t x$$

sur l'ensemble polyédrique  $\{x \mid Ax \leq b\}$ .

Il est possible de démontrer que :

**Proposition 4.3** *Si un programme linéaire admet une solution, alors l'ensemble des ces solutions forment une face de l'ensemble polyédrique sous-jacent.*

La programmation linéaire constitue un domaine fort étudié, de nombreux problèmes de mathématiques appliquées pouvant se formuler comme des programmes linéaires. Elle est elle-même prolongée par la *programmation mathématique* (les contraintes et la fonction à optimiser ne sont plus nécessairement linéaires). Nous n'en dirons pas plus ici, le lecteur désirent étudier ce sujet ira consulter un cours ou un manuel approprié. Voici cependant une belle conclusion à ce cours :

**La géométrie convexe possède des liens importants non seulement avec la géométrie affine et l'algèbre linéaire (elle en constitue en quelques sorte le prolongement), mais aussi avec des domaines plus appliqués (la programmation linéaire est un exemple parmi d'autres).**

# Bibliographie

- [1] A.Brøndsted, *An introduction to convex polytopes*, Springer, 1983, New-York.
- [2] B.A Davey, H.A. Priestley, *Introduction to lattices and order*, Cambridge University Press, Second Edition, 2001, Cambridge.
- [3] J-P Doignon, S. Fiorini, *Mathématiques combinatoires*, Cours donné à l'ULB, 2008.
- [4] J. Doyen, F. Buekenhout, *Mathématiques linéaires*, Syllabus de cours, Troisième édition, 1994.
- [5] B. Grünbaum, *Convex polytopes*, 1967, Second edition, New-York.
- [6] J. Matousek, *Lectures on discrete geometry*, Springer, 2002, New-York.
- [7] R. Webster, *Convexity*, Oxford University Press, 1994, New-York.
- [8] G.M Ziegler, *Lectures on polytopes*, 7<sup>th</sup> printing, Springer, New-York.