

# CONVEXITÉ ET OPTIMISATION

*Guy Cohen*

École Nationale des Ponts et Chaussées  
et  
INRIA

Édition 2000  
correction 2006



# Table des Matières

<b>Notations</b>	<b>v</b>
<b>Préface</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Objet de l'optimisation . . . . .	1
1.1.1 "Le mieux possible" . . . . .	1
1.1.2 Contradictions et compromis . . . . .	2
1.1.3 Optimisation, théorie et pratique . . . . .	2
1.1.4 L'optimisation là où on ne l'attend pas . . . . .	3
1.2 Nomenclature en optimisation . . . . .	3
1.2.1 Optimisation continue et optimisation discrète (ou combinatoire) . . . . .	4
1.2.2 Optimisation déterministe et optimisation dans l'incertain . . . . .	4
1.2.3 Optimisation statique et dynamique . . . . .	4
1.2.4 Un critère/plusieurs critères, un décideur/plusieurs décideurs . . . . .	5
1.2.5 Optimisation linéaire et optimisation non linéaire . . . . .	6
1.3 Aperçu du cours . . . . .	6
1.4 Résumons nous . . . . .	7
<b>2 Éléments d'analyse fonctionnelle</b>	<b>9</b>
2.1 Espaces métriques, espaces complets . . . . .	9
2.2 Espaces normés, espaces de Banach . . . . .	10
2.3 Exemples . . . . .	11
2.4 Fonctions linéaires continues, dual topologique . . . . .	12
2.5 Topologie faible . . . . .	13
2.6 Opérateur adjoint . . . . .	13
2.7 Espaces de Hilbert . . . . .	14
2.7.1 Produit scalaire et dualité . . . . .	14
2.7.2 Opérateur adjoint dans les espaces de Hilbert . . . . .	16
2.8 Résumons nous . . . . .	16
2.9 Corrigé des exercices . . . . .	17
Corrigé de l'Exercice 2.5 . . . . .	17
Corrigé de l'Exercice 2.10 . . . . .	17
<b>3 Ensembles convexes</b>	<b>19</b>
3.1 Sous-espaces affines et enveloppe affine . . . . .	19
3.2 Premières définitions sur les sous-ensembles convexes . . . . .	19
3.3 Opérations préservant la convexité . . . . .	21
3.4 Enveloppe convexe . . . . .	22

3.5	Cônes convexes . . . . .	23
3.5.1	Cône convexe et relation d'ordre . . . . .	23
3.5.2	Cône normal (ou orthogonal) à un sous-ensemble convexe . . . . .	25
3.5.3	Ordre dual dans les espaces de Hilbert . . . . .	26
3.5.4	Enveloppe cônica . . . . .	27
3.5.5	Cône tangent à un sous-ensemble . . . . .	27
3.6	Propriétés topologiques des convexes . . . . .	28
3.6.1	Ouverture et fermeture des convexes . . . . .	28
3.6.2	Intérieur relatif . . . . .	29
3.6.3	Enveloppe convexe fermée . . . . .	30
3.6.4	Enveloppe cônica fermée . . . . .	31
3.6.5	Fermeture des convexes dans la topologie faible . . . . .	31
3.7	Orthogonalité du cône tangent et du cône normal . . . . .	31
3.8	Projection sur les convexes dans les espaces de Hilbert et séparation des convexes . . . . .	33
3.8.1	Projection sur un convexe fermé . . . . .	33
3.8.1.1	Existence et unicité . . . . .	33
3.8.1.2	Propriétés de la projection . . . . .	35
3.8.1.3	Projection sur un cône . . . . .	35
3.8.2	Séparation . . . . .	36
3.8.3	Hyperplans d'appui et description externe des convexes . . . . .	38
3.8.4	Lemme de Farkas . . . . .	40
3.9	Résumons nous . . . . .	42
3.10	Corrigé des exercices . . . . .	43
	Corrigé de l'Exercice 3.9 . . . . .	43
	Corrigé de l'Exercice 3.12 . . . . .	43
	Corrigé de l'Exercice 3.17 . . . . .	43
	Corrigé de l'Exercice 3.18 . . . . .	44
	Corrigé de l'Exercice 3.22 . . . . .	44
	Corrigé de l'Exercice 3.25 . . . . .	44
	Corrigé de l'Exercice 3.27 . . . . .	44
<b>4</b>	<b>Fonctions convexes</b> . . . . .	<b>45</b>
4.1	Épigraphe et convexité . . . . .	45
4.2	Autres notions de convexité . . . . .	47
4.3	Autres associations de sous-ensembles et de fonctions convexes . . . . .	47
4.3.1	Ensembles de niveau . . . . .	47
4.3.2	Fonction indicatrice . . . . .	48
4.3.3	Fonction support . . . . .	48
4.4	Opérations préservant la convexité des fonctions . . . . .	50
4.5	Transformée de Fenchel . . . . .	53
4.6	Enveloppe convexe . . . . .	54
4.7	Propriétés topologiques des fonctions convexes . . . . .	55
4.7.1	Continuité . . . . .	55
4.7.2	Fermeture de l'épigraphe et semi-continuité inférieure . . . . .	57
4.7.3	Enveloppe convexe s.c.i. . . . .	59
4.8	Différentiabilité et sous-différentiabilité des fonctions convexes . . . . .	62
4.8.1	Diverses notions de différentiabilité . . . . .	62
4.8.2	Dérivées directionnelles des fonctions convexes . . . . .	63
4.8.3	Sous-différentiabilité des fonctions convexes . . . . .	65

4.8.3.1	Sous-gradients et interprétation géométrique . . . . .	65
4.8.3.2	Existence de sous-gradients . . . . .	66
4.8.3.3	Une autre définition géométrique du sous-différentiel, . . . . .	67
4.8.3.4	Cas des fonctions convexes différentiables . . . . .	68
4.8.3.5	Cas des fonctions strictement et fortement convexes . . . . .	69
4.8.3.6	Le sous-différentiel comme opérateur multivoque . . . . .	69
4.8.3.7	Calcul sous-différentiel . . . . .	71
4.8.3.8	Sous-différentiel et transformée de Fenchel . . . . .	73
4.9	Résumons nous . . . . .	74
4.10	Corrigé des exercices . . . . .	75
	Corrigé de l'Exercice 4.17 . . . . .	75
	Corrigé de l'Exercice 4.18 . . . . .	75
	Corrigé de l'Exercice 4.60 . . . . .	75
	Corrigé de l'Exercice 4.69 . . . . .	75
<b>5</b>	<b>Optimisation sous contraintes : conditions d'optimalité locales</b>	<b>77</b>
5.1	Optimisation sur un ensemble admissible et inéquations variationnelles . . . . .	77
5.1.1	Formulation du problème et existence d'une solution . . . . .	77
5.1.2	Caractérisation des solutions . . . . .	79
5.1.3	Interprétation géométrique . . . . .	81
5.1.4	Autre caractérisation des solutions optimales . . . . .	82
5.1.5	Pénalisation exacte . . . . .	83
5.2	Optimisation sous contraintes explicites . . . . .	86
5.2.1	Description des contraintes . . . . .	86
5.2.2	Description du cône normal aux contraintes . . . . .	87
5.2.2.1	Une inclusion entre deux cônes . . . . .	87
5.2.2.2	Les écarts complémentaires . . . . .	88
5.2.2.3	L'égalité des deux cônes n'a pas toujours lieu . . . . .	88
5.2.3	Qualification des contraintes et conditions de Lagrange-Karush-Kuhn-Tucker . . . . .	88
5.2.4	Interprétation marginaliste des multiplicateurs : explication géométrique heuristique . . . . .	92
5.2.5	Retour sur la pénalisation exacte . . . . .	93
5.3	Résumons nous . . . . .	95
<b>6</b>	<b>Conditions d'optimalité globales et techniques Lagrangiennes</b>	<b>97</b>
6.1	Formulation compacte des contraintes, Lagrangien, point selle et condition suffisante . . . . .	97
6.1.1	Contraintes égalité et inégalité en dimension finie ou infinie . . . . .	97
6.1.2	Lagrangien . . . . .	98
6.1.3	Point selle . . . . .	99
6.1.3.1	Définition et caractérisations . . . . .	99
6.1.3.2	Cas convexe-concave . . . . .	101
6.1.4	Condition suffisante d'optimalité . . . . .	101
6.2	Le cas convexe, condition nécessaire et suffisante . . . . .	102
6.2.1	Hypothèses de convexité . . . . .	102
6.2.2	Point selle du Lagrangien et multiplicateurs de Kuhn et Tucker . . . . .	103
6.2.3	Application d'un théorème général d'existence de point selle . . . . .	104
6.3	Un point de vue géométrique sur la dualité . . . . .	105
6.3.1	La fonction "perturbation" . . . . .	105
6.3.2	Multiplicateurs et sous-différentiel de la fonction perturbation . . . . .	107
6.3.3	Fonction perturbation, transformée de Fenchel et Lagrangien . . . . .	109

6.3.4	Explication géométrique des résultats . . . . .	111
6.3.4.1	Le cas convexe . . . . .	111
6.3.4.2	Cas non convexes : remarques diverses sur le saut de dualité . . . . .	112
6.4	Résumons nous . . . . .	113
6.5	Corrigé des exercices . . . . .	114
	Corrigé de l'Exercice 6.6 . . . . .	114
	Corrigé de l'Exercice 6.7 . . . . .	114
<b>7</b>	<b>Lagrangien augmenté</b>	<b>115</b>
7.1	Justification géométrique . . . . .	115
7.1.1	Cas non convexes . . . . .	115
7.1.2	Cas convexe . . . . .	117
7.2	Lagrangien augmenté et régularisation duale . . . . .	118
7.2.1	Diverses expressions et propriétés du Lagrangien augmenté . . . . .	119
7.2.2	Quelques propriétés de la régularisation de Yosida-Moreau . . . . .	121
7.2.3	Régularisation du Lagrangien par rapport $\lambda$ et conséquences . . . . .	123
7.3	Conclusion . . . . .	126
7.4	Résumons nous . . . . .	127
	<b>Bibliographie</b>	<b>129</b>

# Notations

$\stackrel{\text{def}}{=}$  définition (le membre de gauche est défini par l'expression de droite)

$A \setminus B$  différence ensembliste (ensemble des éléments de  $A$  qui n'appartiennent pas à  $B$ )

$2^{\mathcal{X}}$  famille de tous les sous-ensembles de  $\mathcal{X}$

$\mathbb{N}$  ensemble des nombres entiers

$\mathbb{R}, \mathbb{R}^n$  ensemble des nombres réels, respectivement des vecteurs de dimension  $n$  à coordonnées réelles (éventuellement muni de la structure d'espace vectoriel euclidien habituelle)

$\mathbb{R}_+, \mathbb{R}_+^n$  ensemble des nombres réels non négatifs, respectivement des vecteurs de dimension  $n$  à coordonnées réelles non négatives

$\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$

$\mathcal{X}, \mathcal{Y}, \dots$  en général, espaces vectoriels

$\mathcal{L}(\mathcal{X}, \mathcal{Y})$  applications linéaires continues de  $\mathcal{X}$  dans  $\mathcal{Y}$  (voir §2.4)

$\mathcal{X}^*$  dual topologique de  $\mathcal{X}$  (voir §2.5)

$\text{im } A$  image d'un opérateur linéaire  $A$

$\ker A$  noyau d'un opérateur linéaire  $A$  (en Anglais, "kernel")

$A^*$  opérateur adjoint d'un opérateur linéaire  $A$  (voir §2.6)

$\langle \cdot, \cdot \rangle$  produit scalaire

$\|\cdot\|$  norme (un indice peut préciser la nature de cette norme)

$\Sigma_n$  simplexe de dimension  $n$  (voir (3.3))

$\overline{A}$  fermeture ou adhérence d'un sous ensemble  $A$  d'un espace topologique

$\overset{\circ}{A}$  intérieur d'un sous-ensemble  $A$  d'un espace topologique

$\overset{\circ}{\circ} C$  intérieur relatif d'un sous-ensemble convexe  $C$  d'un espace vectoriel topologique (voir Définition 3.37) ; on trouve dans la littérature la notation  $\text{ri } C$  (pour "relative interior")

$\aleph(A)$  enveloppe affine d'un sous-ensemble  $A$  (voir Définition 3.3) ; on trouve parfois dans la littérature la notation  $\text{aff } A$

$\odot(A)$  enveloppe convexe d'un sous-ensemble  $A$  (voir Définition 3.13) ; on trouve dans la littérature la notation  $\text{co } A$  ; on utilise la même notation pour l'enveloppe convexe de fonctions (voir §4.6)

- $\overline{\odot}(A)$  enveloppe convexe fermée d'un sous-ensemble  $A$  (voir Définition 3.39) ; on trouve dans la littérature la notation  $\overline{\text{co}} A$  ; on utilise la même notation pour l'enveloppe convexe s.c.i. de fonctions (voir §4.7.3)
- $\angle(A)$  enveloppe cônica d'un sous-ensemble  $A$  (voir Définition 3.29)
- $\overline{\angle}(A)$  enveloppe cônica fermée d'un sous-ensemble  $A$  (voir Définition 3.42)
- $A_a^\perp$  cône orthogonal à un sous-ensemble convexe  $A$  au point  $a \in A$  (voir Définition 3.20) ; on trouve dans la littérature la notation  $N_A(a)$
- $A^\perp$  notation abrégée pour  $A_0^\perp$  (définie seulement si  $0 \in A$ )
- $A_a^\top$  cône tangent à un sous-ensemble  $A$  au point  $a \in A$  (voir Définition 3.32) ; on trouve dans la littérature la notation  $T_A(a)$
- $A^\top$  notation abrégée pour  $A_0^\top$  (définie seulement si  $0 \in A$ )
- $A^\circ$  cône polaire d'un sous-ensemble  $A$  (voir Remarque 3.26)
- $C^*$  cône dual d'un cône convexe  $C$  ( $C^* = -C^\perp$  — voir §3.5.3)
- $\Pi_C(x)$  projection de  $x$  sur  $C$  (voir §3.8.1)
- $\delta_C(x)$  distance de  $x$  à  $C$  (voir Définition 5.12)
- $\mathfrak{U}(f), \mathfrak{U}_s(f)$  épigraphe (respectivement, épigraphe strict) d'une fonction  $f$  (voir Définition 4.1) ; on trouve dans la littérature la notation  $\text{epi } f$
- $\mathfrak{T}(f)$  domaine de  $f$  (voir Définition 4.5) ; on trouve dans la littérature la notation  $\text{dom } f$
- $\overline{\nabla}_\beta(f)$  ensemble de niveau de la fonction  $f$  au niveau  $\beta$  (voir Définition 4.10)
- $I_A$  fonction indicatrice d'un sous-ensemble  $A$  (voir Définition 4.11)
- $\sigma_A$  fonction support d'un sous-ensemble  $A$  (voir Définition 4.12)
- $f \square g$  inf-convolution de  $f$  et  $g$  (voir (4.6))
- $f^F$  transformée de Fenchel de  $f$  (voir Définition 4.20) ; on trouve dans la littérature la notation  $f^*$
- $\partial f(x)$  sous-différentiel de  $f$  en  $x$  (voir Définition 4.49)



# Préface

Ce cours (électif) a été donné pendant les années scolaires 1999–2000 et 2000–2001, au premier trimestre, sur une trentaine d’heures à l’École Nationale des Ponts et Chaussées. Il a été suivi par des étudiants de deuxième et troisième année de cette école. Il a bénéficié des remarques et observations de ce public.

*Guy Cohen  
Champs sur Marne,  
le 6 décembre 2000*

*L’art de l’ingénieur est de ne pas s’arrêter là où la théorie n’apporte plus de réponses. Le devoir de l’ingénieur est de ne pas ignorer les réponses que peut apporter la théorie.*



# Chapitre 1

## Introduction

### 1.1 Objet de l'optimisation

#### 1.1.1 “Le mieux possible”

Faire “le mieux possible” est somme toute une attitude naturelle dans la vie courante. Mais ceci n’a qu’un sens très relatif par rapport aux contraintes subies ou volontairement acceptées, contraintes qui peuvent résulter elles-mêmes de choix passés. Pour un ingénieur également, “faire le mieux possible” devrait être un objectif permanent lorsqu’il a par exemple en charge la conception d’un équipement ou le dimensionnement d’une installation. Mais là aussi, l’expression doit être relativisée en fonction de contraintes de budget, de sécurité, ou autre, contraintes dont le niveau a lui même fait l’objet de décisions préalables et souvent extérieures.

Les mots “optimiser, optimisation, etc.” sont supposés refléter cette idée du “mieux possible”. Dans la vie courante, les choix possibles se limitent souvent à deux (on appelle alors cela une “alternative”) ou quelques unités, de sorte que l’“algorithme” de prise de décision se réduit à envisager, explicitement ou pas, toutes les possibilités, à considérer et à évaluer leurs conséquences probables, et à retenir celle qui paraît “la meilleure” (mais cette notion de “meilleur” n’a pas de sens tant que le critère de choix n’a pas été explicité). Dans les problèmes techniques, les choix possibles représentent souvent un continuum (quelles dimensions donner à une poutre ?) et une telle énumération exhaustive des possibilités de choix est impensable. Il faut alors un “algorithme” plus performant, c’est-à-dire une méthode pour cheminer vers “la” solution en explorant le moins de “moins bonnes” solutions possible. Et il faut d’abord savoir “reconnaître”, c’est-à-dire, caractériser “la” solution, puisque le fait de pouvoir la comparer à toutes les autres a été exclu.

Les mots “optimiser” ou “optimisation” sont bien souvent galvaudés dans le langage courant (dans la communication publicitaire par exemple) dans la mesure où ils sont utilisés dans le sens d’une simple “amélioration” par rapport à un état antérieur. On ne peut strictement parler d’optimisation que si l’on est sûr qu’*aucun autre* choix ou aucune autre décision n’est possible — dans le cadre de contraintes explicitement spécifiées — qui fera mieux que celle réputée “optimale”, et “mieux” n’a de sens qu’une fois précisé l’échelle de mesure de performance (ou “critère”<sup>1</sup>). Un autre euphémisme consiste à parler de “solution *sous*-optimale” lorsqu’on est prêt à admettre qu’il se pourrait qu’on puisse encore faire mieux. Mais, sans précision sur le “combien mieux”, ou “combien moins bien”, la notion de “sous-optimalité” reste très floue, et une façon plus claire de s’exprimer consisterait à parler tout simplement de “*non* optimalité”.

---

<sup>1</sup>Nous emploierons aussi indifféremment le vocable “fonction coût”

### 1.1.2 Contradictions et compromis

Si tout le monde s'accorde à admettre qu'"il vaut mieux être riche et en bonne santé que pauvre et malade", on se trouve rarement dans la pratique en face de choix aussi clairs. Le plus souvent, on n'a pas *un* mais *plusieurs* critères en tête, et ils varient rarement tous dans le même sens. C'est pourquoi beaucoup de choix sont difficiles. C'est pourquoi aussi l'optimisation est souvent décriée, parce que l'illusion de faire mieux d'un certain point de vue se traduit par une détérioration d'un autre point de vue. Donc, bien souvent les choix ne sont que des compromis ("compromis" sonne déjà de façon moins noble que "choix optimal").

On verra que l'optimisation est une technique pour "quantifier" les compromis entre des critères parfois non commensurables, c'est-à-dire peu susceptibles de s'évaluer dans les mêmes unités. En effet, si tout se ramenait à des Francs<sup>2</sup>, il suffirait d'additionner tous les effets d'une décision pour la comparer à une autre. Mais que vaut la vie humaine ? Penser à une politique de sécurité routière qui coûte des investissements en Francs et qui rapporte des économies de . . . morts.<sup>3</sup> La théorie de l'optimisation fournit des outils permettant de répondre à des questions du type "si on accepte de perdre  $x$  sur un critère, combien pourra-t-on gagner au mieux sur un autre critère ?" au moins de façon marginale (c'est-à-dire lorsque  $x$  est petit). Voilà d'où vient le "prix de la vie humaine" : dans un contexte (et avec des contraintes) bien précis(es),  $x$  Francs de plus de tel investissement en matière de sécurité permettront (statistiquement) d'économiser  $y$  morts, d'où le prix marginal du mort. Mais ce prix marginal pourra être très différent dans un autre cadre décisionnel, avec d'autres niveaux de contraintes par exemple . . . ce qui doit donner à réfléchir. L'optimisation ne se substitue pas à la responsabilité du décideur, elle rend les conséquences de certains choix a priori (fixation d'un budget, d'un niveau de sécurité) plus explicites en ayant fait le mieux possible à l'intérieur du contexte créé par ces choix a priori.

### 1.1.3 Optimisation, théorie et pratique

Au vu de la discussion qui précède, on doit comprendre que l'optimisation, c'est au moins trois choses.

**L'art de formuler des problèmes de décision** grâce à des concepts et outils mathématiques précis.

Comme tout art, son acquisition demande du temps et de la pratique même lorsqu'on en a compris les fondements (ce qui reste de toute façon indispensable). Que mettre dans la "fonction coût" ? Que mettre dans les contraintes ? Comment formuler ces contraintes ? Comment revenir sur cette formulation lorsque la solution a été obtenue et ne paraît pas adéquate ou intuitivement satisfaisante ? Autant de questions auxquelles il est peu probable que le lecteur sache facilement répondre dès la fin de la lecture de ce document. Comme disait mon bon professeur de Recherche Opérationnelle, Michel Sakarovitch, lorsque j'étais moi-même étudiant dans cette École (oh, c'était il y a bien longtemps . . . ) : "je n'enseigne pas la Recherche Opérationnelle, j'enseigne les mathématiques de la Recherche Opérationnelle. Si vous voulez apprendre la Recherche Opérationnelle, venez la pratiquer."

**Une théorie mathématique.** Comme on l'a dit, il est facile de définir une solution optimale (c'est celle qui donne à une fonction coût la valeur la plus basse — en principe on cherche à minimiser les coûts), il est plus difficile de la reconnaître (des fois que vous tomberiez dessus par hasard !) puisqu'il est exclu de vérifier alors qu'elle obéit bien à la définition. Donc, il faut savoir *caractériser* une solution par des conditions nécessaires, suffisantes, ou les deux, en somme connaître certaines de ses propriétés pour les reconnaître. C'est un objectif important qui sera abordé dans ce cours.

<sup>2</sup>ou des Euros, . . . soyons modernes !

<sup>3</sup>Ne parlons pas des blessés qui occasionnent, eux, des dépenses de soins.

**Une “cuisine” algorithmique.** Une fois que le mathématicien a su caractériser une solution, et d’abord se prononcer sur son existence voire son unicité, l’ingénieur voudrait bien pouvoir *calculer* cette solution. Il est hors de question de passer en revue tout ce qui est imaginable ou autorisé ; il s’agit d’aller au plus vite vers la solution en améliorant par touches successives une ébauche de celle-ci. Autant dire que la plupart des algorithmes sont *itératifs*. Itératifs, mais rarement optimaux eux-mêmes, c’est l’une des contradictions de ce domaine. C’est pourquoi nous préférons parler de “cuisine algorithmique” (on rappelle que la cuisine aussi peut être un art . . . lorsqu’elle est bien faite). Cela n’exclut pas les fondements mathématiques, les idées, les démonstrations de convergence . . . mais il restera toujours un je ne sais quoi de “tour de main”.

L’algorithmique de l’optimisation peut faire l’objet d’un cours (d’informatique ?) à elle toute seule. De plus, à cause du “tour de main” qui ne s’acquiert pas du premier coup, il est vivement recommandé de ne jamais programmer un algorithme par soi-même, même si on a le sentiment de l’avoir bien compris. Il vaut mieux faire confiance aux nombreux programmes de bibliothèques scientifiques aujourd’hui disponibles.

Le volume horaire de ce cours ne permettra pas malheureusement d’aborder ne serait-ce que les grandes idées qui président à la conception et à l’analyse d’algorithmes ou de décrire les principales catégories d’algorithmes.

### 1.1.4 L’optimisation là où on ne l’attend pas

On l’a vu, l’optimisation se situe principalement sur le terrain de l’“aide à la décision”. Mais la “Nature” ou le “Hasard” aussi prend des décisions, et bien souvent des décisions optimales. Prenez une boîte d’allumettes à moitié pleine, renversez les allumettes sur la table, mélangez les, et remettez les en vrac dans la boîte. Puis agitez la boîte dans un mouvement de va et vient saccadé, et observez que les allumettes tendent à se ranger bien parallèlement les unes aux autres. Comme l’expérience est reproductible, c’est qu’il y a une loi derrière ce phénomène. Cette loi peut s’énoncer comme le fait que la configuration stable des allumettes minimise un certain critère ou “potentiel” (la Nature aurait horreur du désordre). Il y a bien d’autres lois de la Physique qui peuvent se formuler comme la solution d’un problème d’optimisation (cheminement d’un rayon lumineux dans un milieu d’indice de réfraction variable, répartition des débits ou des intensités dans un réseau hydraulique ou électrique maillé, nombreux problèmes d’équilibre en Mécanique). Ce point de vue *variationnel* sur les problèmes d’équilibre ou de résolution d’équations est souvent extrêmement fructueux, tant sur le plan de l’étude de l’existence et de l’unicité des solutions, que sur ceux de l’étude de leurs propriétés et de leur calcul numérique (problème des algorithmes de résolution).

Il y a d’autres problèmes d’équilibre qui ne peuvent stricto sensu s’interpréter comme la solution de problèmes d’optimisation au sens de la minimisation ou la maximisation d’un seul critère, mais qui relèvent toutefois de la famille des “problèmes variationnels” (*inéquations variationnelles* que l’on rencontre par exemple dans les problèmes d’équilibre de réseaux de transport, parfois interprétables comme la solution d’un “jeu” à plusieurs joueurs). Le sujet des inéquations variationnelles est intimement lié à celui de l’optimisation et nous le rencontrerons forcément dans ce cours, même si nous ne le traiterons pas toujours dans son contexte le plus général.

## 1.2 Nomenclature en optimisation

L’idée d’optimisation étant très générale, il faut entrer un peu plus dans la description de sous-catégories qui correspondent à des outils mathématiques et des niveaux de difficultés assez variés ou variables.

### 1.2.1 Optimisation continue et optimisation discrète (ou combinatoire)

La première distinction à faire concerne la nature des “espaces” dans lesquels les variables de décision prennent leurs valeurs : c’est la dichotomie “discret-continu” bien marquée en Mathématiques et qui conditionne évidemment beaucoup les possibilités de recourir à certains outils. Nous ne dirons pas grand’chose de l’optimisation en variables *discrètes*, que l’on appelle aussi *optimisation combinatoire* car le plus souvent, la combinatoire du nombre des valeurs prises par l’ensemble des variables est “explosive”, ce qui exclut l’énumération exhaustive comme méthode de recherche de la solution. L’optimisation combinatoire est un domaine en soi et elle ne sera pas du tout abordée dans ce cours.

On se situera donc dans le cadre de l’optimisation en variables continues, ce qui permet de considérer des variations infinitésimales et l’usage de la notion de dérivée des fonctions impliquées (les fonctions pourront ne pas être parfaitement “lisses” et seulement “sous-différentiables” en un sens à préciser ; toujours est-il qu’on disposera d’une possibilité de parler de petites variations).

### 1.2.2 Optimisation déterministe et optimisation dans l’incertain

Bien souvent, le décideur doit faire face à un certain nombre de facteurs inconnus, ou plus précisément seulement connus soit statistiquement, soit par des plages de valeurs (maxima, minima) que ces facteurs sont susceptibles de prendre. Il existe plusieurs approches de la “décision dans l’incertain” (que veut-dire “optimal” dans ce cas ?), et on peut au moins distinguer les cas où

- un cadre probabiliste est utilisé pour décrire mathématiquement l’incertain (on parle alors d’*optimisation stochastique* et on cherche en général à minimiser l’espérance mathématique de la fonction coût, ce qui suppose implicitement qu’on accepte un comportement “optimal en moyenne”), ou bien
- des plages de valeurs possibles (mais non probabilisées) des aléas sont supposées et on cherche à se prémunir contre la pire des occurrences (du point de vue de la fonction coût), ce qui correspond à un comportement pessimiste ou très prudent (on parle alors d’*optimisation min-max*, dans le cas le plus défavorable, ou *robuste*).

À nouveau, ces situations d’optimisation dans l’incertain ne seront pas abordées dans ce cours. On se limitera donc au cas de l’optimisation *déterministe* où tous les facteurs intervenant dans le problème, en dehors des variables de décisions à notre disposition, sont supposés connus avec précision. C’est bien sûr une idéalisation par rapport à la plupart des situations concrètes.

### 1.2.3 Optimisation statique et dynamique

Le mot “dynamique” est le plus souvent utilisé lorsque le temps intervient, c’est-à-dire qu’un système évolue, qu’il faut prendre des décisions à diverses étapes de cette évolution, et que ces décisions auront une influence sur tout le futur (mais pas sur le passé en raison du principe de causalité). L’optimisation dans ce contexte dynamique est appelée *commande optimale*<sup>4</sup> qui peut être à nouveau déterministe ou stochastique. Une fois de plus, ces domaines peuvent faire l’objet de cours spécifiques et ne seront pas abordés ici explicitement (sauf éventuellement à l’occasion d’un exercice).

Il existe en fait un sens plus subtil et plus profond au mot “dynamique” dans le contexte de problèmes de décision. C’est ce que nous allons tenter d’expliquer maintenant.

- Si on a affaire à une formulation déterministe, le futur de l’évolution est parfaitement calculable à partir du modèle du système et des décisions planifiées à chaque instant. Les objets à manipuler (“commandes”, c’est-à-dire variables de décision, états du système, etc.) sont des objets indexés

<sup>4</sup>On rencontre souvent le mot “contrôle” au lieu de “commande” dans ce contexte, mais il s’agit là de Français.

par le temps, ou “trajectoires” : ces objets sont de dimension finie si le temps est un temps discret (comme sur un ordinateur) et si l’horizon d’optimisation est borné, et ils sont de dimension infinie en particulier lorsque le temps est considéré comme une variable continue. L’implication du mot “dynamique” dans ce contexte déterministe est donc plutôt ici que l’on manipule des “trajectoires”, ce qui peut justifier de travailler dans des espaces fonctionnels de dimension infinie. Dans ce cours, nous avons choisi de développer la théorie dans le cadre des espaces de Hilbert de dimension finie ou infinie, et donc on couvre implicitement ces situations dynamiques, même si la notion de récursivité incluse dans l’évolution d’un système dynamique justifie(ra)it de consacrer un développement spécial à la commande optimale.

- Supposons maintenant que des facteurs incertains influencent l’évolution du système dynamique en plus de nos propres décisions. Alors, à un instant  $t$ , il n’est plus possible de prédire exactement l’évolution future même si l’on est capable de dire quelles décisions seront prises jusqu’à la fin de l’horizon. En fait, on a intérêt à ne pas prendre toutes les décisions futures tout de suite, mais à se limiter à celles qui seront immédiatement appliquées (celles de l’instant courant). En effet, la plupart du temps dans ces situations dynamiques et en présence d’incertain, on dispose d’observations à chaque instant qui nous renseignent de façon plus ou moins complète sur ce qu’ont été les valeurs passées des “bruits”, et sur l’état actuel du système, nous permettant de “corriger le tir” éventuellement. On a donc intérêt à prendre les décisions à chaque instant au vu des observations disponibles jusqu’à cet instant. On entre là dans un autre monde, celui de la commande en *boucle fermée* ou en *feedback*<sup>5</sup> qui, du point de vue mathématique, requiert un arsenal et des techniques entièrement différents. Ces véritables situations dynamiques, caractérisées par l’alternance de décisions et d’observations<sup>6</sup>, et qui relèvent le plus souvent d’une approche par la *programmation dynamique*, ne seront pas l’objet de ce cours. On se limitera donc à l’optimisation “statique” (par opposition à “dynamique”) même si la théorie sera applicable dans une certaine mesure aux problèmes de commande optimale déterministes pour les raisons expliquées plus haut.

#### 1.2.4 Un critère/plusieurs critères, un décideur/plusieurs décideurs

Comme cela a été évoqué plus haut, un même décideur est généralement confronté à plusieurs critères de choix qui, sans être nécessairement antagonistes, ne varient pas forcément tous dans le même sens en fonction des décisions envisageables. Une théorie de l’*optimisation multi-critère* ou *vectorielle* (c’est le critère qui est à valeurs vectorielles) existe. Nous ne l’aborderons pas explicitement dans ce cours. Nous avons indiqué qu’en choisissant lesquelles des préoccupations seront agrégées dans une fonction coût unique (quand elles sont commensurables) et lesquelles apparaîtront sous forme de contraintes, on peut parvenir à formuler de tels problèmes de compromis, quitte à itérer sur les choix a priori (par exemple sur les niveaux assignés aux contraintes). Comme on y a déjà fait allusion, la théorie de la dualité fournit une aide dans cette optique.

Nous resterons donc dans le cadre de l’optimisation “mono-critère”. Nous resterons aussi dans le cadre de l’optimisation “mono-décideur”. À l’opposé, il existe des situations où l’on est en présence de plusieurs décideurs ayant chacun leur propre fonction coût et leurs propres contraintes. Mais bien sûr, la décision de chacun agit sur le critère de tous (il n’est pas difficile d’imaginer des exemples en Économie). Nous entrons là sur le terrain de la *théorie des jeux* dans laquelle plusieurs concepts de “solution” ou d’“équilibre” (Pareto, Nash, Stackelberg) ont été définis.

<sup>5</sup>Eh oui, ça nous arrive aussi de parler Français.

<sup>6</sup>ce qui, encore une fois, n’a de véritable importance qu’en présence de facteurs incertains

### 1.2.5 Optimisation linéaire et optimisation non linéaire

La dichotomie “linéaire/non linéaire” est assez classique en Mathématiques même si elle représente une partition faussement symétrique (dans le monde du “non linéaire” il y a beaucoup de sous-catégories). En optimisation, on parle, du côté linéaire, de *programmation linéaire*.<sup>7</sup> La programmation linéaire, qui pourrait n’être considérée que comme un cas particulier de la *programmation convexe*, elle-même sous-catégorie de la *programmation non linéaire*, est bien plus que cela en fait. Historiquement, elle a occupé le devant de la scène en optimisation dans les années soixante sous l’impulsion de George B. Dantzig, et on peut dire que dans beaucoup de cercles, “optimisation” a été pour un temps synonyme de “programmation linéaire”. Techniquement, bien que cas particulier de la programmation convexe, le contexte de la programmation linéaire permet des développements qui lui sont spécifiques. Ce fut en particulier le cas sur le plan algorithmique avec le fameux *algorithme du simplexe* qui exploite un certain aspect combinatoire de la programmation linéaire même en variables continues. Les développements algorithmiques plus récents par les méthodes de points intérieurs ont sérieusement réduit l’écart entre programmation linéaire et non linéaire.

Toujours est-il que la programmation linéaire, par ses spécificités, constitue à elle seule un objet de cours. Elle ne sera pas abordée explicitement ici, même si les techniques de dualité qui y seront développées s’appliquent aussi bien à son cas (mais avec quelques difficultés techniques particulières au niveau de la justification mathématique).

Ce cours se situe donc sur le terrain de l’optimisation non linéaire, et plus exactement de l’optimisation convexe, comme le suggère son titre.

## 1.3 Aperçu du cours

Comme on vient donc de le voir, l’objet principal de ce cours est celui de l’optimisation

- à un seul critère (sous contraintes),
- déterministe,
- en variables continues,
- statique (sans pour autant sacrifier la dimension infinie),
- et dans un cadre convexe.

Ce dernier point signifie que les fonctions du problème (coût et contraintes) seront supposées *convexes* (et même affines pour les contraintes égalité). On parle alors d’optimisation ou de programmation convexe. Ce sujet de recherche a connu un fort développement après le succès initial de la programmation linéaire dans les années soixante et il s’est constitué en une théorie cohérente représentant la première incursion structurée dans le domaine de l’optimisation non linéaire.

Bien que la convexité ait une interprétation économique claire (phénomène des coûts marginaux croissants), beaucoup de problèmes rencontrés dans la réalité ne sont pas convexes. Cependant, puisque l’on s’intéresse à l’optimisation de fonctions — disons à la minimisation pour fixer les idées — au voisinage immédiat de leur minimum, les fonctions ont nécessairement une forme convexe. L’hypothèse de convexité globale peut donc être considérée comme une idéalisation (nécessaire dans toute théorie mathématique) où l’on pourra étudier sur tout l’espace (globalement) des phénomènes qui risqueraient de n’apparaître en réalité que localement dans la région qui nous intéresse. Tant que l’on n’utilise que des concepts locaux (variations infinitésimales autour d’un point), il y a peu de chances que l’aspect qualitatif

<sup>7</sup>Le mot “programmation”, comme dans les expressions “programmation dynamique”, “convexe”, etc., n’a pas ici de connotation informatique, mais il fait sans doute référence à l’idée de “planification des décisions”.



des phénomènes changent radicalement du fait de cette extension à tout l'espace de propriétés seulement locales. On s'est simplement placé dans un cadre mathématique plus agréable. Cependant, dès que l'on commence à tirer des conclusions globales d'hypothèses de convexité (et on rencontrera plusieurs circonstances de cette nature), alors il y a fort à parier que ces conclusions soient qualitativement différentes si on renonce aux hypothèses de convexité.

Avant d'aborder le sujet de l'optimisation proprement dit, il nous faudra donc explorer le monde de la convexité. C'est un monde où l'intuition géométrique des concepts et résultats principaux est très éclairant pour leur compréhension. Nous userons et abuserons de "petits dessins" pour appréhender ces résultats, sans sacrifier cependant à la rigueur mathématique. Celle-ci est d'autant plus nécessaire que l'on s'aventure sur le terrain de la dimension infinie où l'intuition est plus facilement prise en défaut. Nous le ferons dans le cadre le plus simple et à nouveau le plus géométrique possible, celui des *espaces de Hilbert*. C'est pourquoi le chapitre suivant passe en revue quelques faits essentiels d'*analyse fonctionnelle* sans prétendre ici à l'exhaustivité ni à l'approfondissement. Le but est seulement de poser quelques prérequis qui devront éventuellement être étudiés dans les ouvrages de référence du domaine.

## 1.4 Résumons nous

L'optimisation est une branche des "sciences de la décision" : son but est de spécifier la décision souhaitable parmi toutes les possibilités d'un espace de décisions, par le choix d'un "critère" ou "fonction coût" et par le choix de contraintes. Ces choix préalables sont "arbitraires", mais une fois ceux-ci arrêtés, une décision<sup>8</sup> devient "meilleure" que toutes les autres. Il s'agit alors de

- caractériser cette décision pour la reconnaître (conditions d'optimalité) car la définition initiale de l'optimalité n'est pas utilisable de façon opérationnelle ;
- puis de la calculer (algorithmes).

La théorie mathématique de l'optimisation traite de ces deux questions mais seule la première est abordée dans ce cours.

L'arbitraire du choix du critère et des contraintes ne fait donc pas partie de la théorie mathématique stricto sensu. Cet arbitraire est la marge de manœuvre qui permet à l'utilisateur d'exprimer ses désirs plus ou moins précis et souvent contradictoires. Mais la compréhension de la théorie mathématique qui vient ensuite permet de guider ces choix, en évitant par exemple des formulations difficiles à résoudre et en fournissant un certain "feedback" : la théorie de la dualité et l'interprétation des "multiplicateurs" qu'elle introduit permet de quantifier dans une certaine mesure les choix a priori les uns par rapport aux autres.

Les conditions d'optimalité qu'elle produit placent la théorie de l'optimisation dans une famille plus large dite des "problèmes variationnels" qui contient notamment tous les problèmes d'équilibre rencontrés dans de nombreuses branches de la Physique, des problèmes de transport, de théorie des jeux, etc. Inversement, certains états d'équilibre de la Nature peuvent se réinterpréter comme les solutions de problèmes d'optimisation, ce qui donne souvent des moyens efficaces de calculer ces équilibres ou d'étudier leurs propriétés.

L'optimisation en variables continues, avec une fonction coût unique, par ailleurs convexe, et des contraintes également convexes, telle qu'elle est traitée dans ce cours, avec de plus un point de vue "statique" et "déterministe", est également un cas particulier de théories envisageant la prise en considération de plusieurs critères, de plusieurs décideurs aussi (théorie des jeux), de phénomènes évolutifs (commande ou "contrôle" optimal(e)) et/ou aléatoires (optimisation stochastique). Les outils mathématiques diffèrent aussi sensiblement de ceux utilisés dans le cas de variables de décision à valeurs discrètes.

---

<sup>8</sup>ou un sous-ensemble de décisions en cas de non-unicité



## Chapitre 2

# Éléments d'analyse fonctionnelle

Dans ce chapitre, on propose des rappels minimaux d'analyse fonctionnelle que l'on peut approfondir dans de nombreux ouvrages (voir par exemple Aubin [1] ou Kolmogorov et Fomine [10]).

Les espaces de Hilbert sont un bon cadre pour la formulation de problèmes d'optimisation rencontrés par l'ingénieur. L'espace  $\mathbb{R}^n$  muni de son produit scalaire classique est l'exemple le plus connu. Le fait de ne pas se limiter à la dimension finie permet par exemple d'aborder les problèmes de commande optimale en temps continu où les grandeurs pertinentes sont des fonctions du temps que l'on peut munir d'une norme  $L^2$  ou  $H^1$  (nous y reviendrons en temps utile).

Le prix à payer pour travailler en dimension infinie, par rapport au fait de se limiter à la dimension finie, est de prendre un certain nombre de précautions et de s'imprégner de quelques différences somme toute pas très difficiles à appréhender. Parmi elles, il faut mentionner en premier lieu la distinction à opérer entre *topologie forte* et *topologie faible* et le fait que, dans la première, les sous-ensembles *bornés* fermés ne sont pas nécessairement *compacts* (comme c'est le cas en dimension finie, et aussi en dimension infinie mais pour la topologie faible — s'ils sont bien sûr faiblement fermés).

### 2.1 Espaces métriques, espaces complets

Un espace vectoriel  $\mathcal{X}$  (réel) est un espace métrique lorsqu'il est muni d'une *distance*  $d$ , c'est-à-dire d'une application de  $\mathcal{X} \times \mathcal{X}$  dans  $\mathbb{R}$  vérifiant :

- définie positivité :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, d(x, y) \geq 0 \quad \text{et} \quad (d(x, y) = 0 \Leftrightarrow x = y), \quad (2.1a)$$

- symétrie :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, d(x, y) = d(y, x), \quad (2.1b)$$

- inégalité triangulaire :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, z \in \mathcal{X}, d(x, z) \leq d(x, y) + d(y, z), \quad (2.1c)$$

ce qui en fait un espace topologique pour la topologie associée à cette distance : la convergence d'une suite  $\{x^k\}$  vers  $x$  dans cette topologie se traduit par la convergence vers zéro de  $d(x^k, x)$ .

Une suite convergente est une *suite de Cauchy*, c'est-à-dire que

$$\forall \varepsilon > 0, \exists k_0 : \forall k \geq k_0, k' \geq k_0, d(x^k, x^{k'}) \leq \varepsilon. \quad (2.2)$$

Inversement, lorsque dans un espace métrique, toute suite de Cauchy converge (la limite est alors unique), on dit que l'espace métrique est *complet*.

## 2.2 Espaces normés, espaces de Banach

Un espace normé est un espace métrique dont la distance est définie à partir d'une *norme*. Plus précisément, la distance  $d(x, y)$  est définie par

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

la norme  $\|\cdot\|$  étant une application de  $\mathcal{X}$  dans  $\mathbb{R}$  vérifiant :

- définie positivité :

$$\forall x \in \mathcal{X}, \|x\| \geq 0 \quad \text{et} \quad (\|x\| = 0 \iff x = 0) , \quad (2.3a)$$

- inégalité triangulaire :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, \|x + y\| \leq \|x\| + \|y\| , \quad (2.3b)$$

- homogénéité de degré 1 :

$$\forall x \in \mathcal{X}, \alpha \in \mathbb{R}, \|\alpha x\| = |\alpha| \|x\| . \quad (2.3c)$$

**Exercice 2.1.** Une norme vérifie les propriétés supplémentaires :

- symétrie :

$$\forall x \in \mathcal{X}, \|-x\| = \|x\| , \quad (2.3d)$$

- autre forme de l'inégalité triangulaire :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, \left| \|x\| - \|y\| \right| \leq \|x - y\| . \quad (2.3e)$$

**Exercice 2.2.** Une distance définie à partir d'une norme vérifie les propriétés supplémentaires :

- invariance par translation :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, z \in \mathcal{X}, d(x + z, y + z) = d(x, y) , \quad (2.4a)$$

- homogénéité de degré 1 :

$$\forall x \in \mathcal{X}, y \in \mathcal{X}, \alpha \in \mathbb{R}, d(\alpha x, \alpha y) = |\alpha| d(x, y) . \quad (2.4b)$$

Un *espace de Banach* est un espace *normé complet*.

### 2.3 Exemples

Dans  $\mathbb{R}^n$ , on peut considérer les normes

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n, p \in \mathbb{N} \setminus \{0\}, \quad \|x\|_p \stackrel{\text{def}}{=} \sqrt[p]{\sum_{i=1}^n |x_i|^p}, \quad \|x\|_\infty \stackrel{\text{def}}{=} \max_{i=1, \dots, n} |x_i|. \quad (2.5)$$

Comme dans tout espace de dimension finie, toutes ces normes sont équivalentes<sup>1</sup> et elles induisent la même topologie.

On appelle *boule (fermée) de centre  $x$  et de rayon  $\rho$* , l'ensemble

$$\overline{B}(x, \rho) \stackrel{\text{def}}{=} \{y \mid d(x, y) \leq \rho\}. \quad (2.6)$$

La Figure 2.1 représente, dans  $\mathbb{R}^2$ , les formes des boules associées à trois des normes introduites en (2.5).

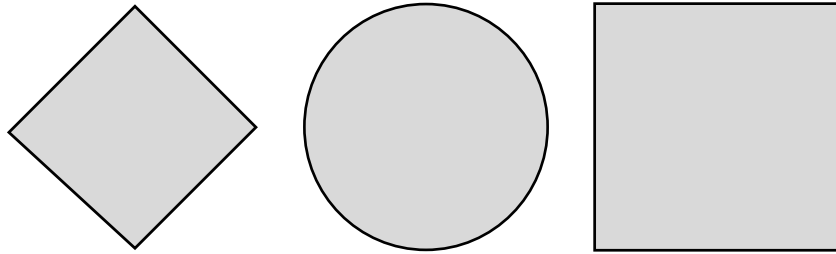


Figure 2.1: Formes des boules de  $\mathbb{R}^2$  associées aux normes  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  et  $\|\cdot\|_\infty$ , respectivement

Pour des suites infinies  $x = (x_0, x_1, \dots)$  dans  $\mathbb{R}$  (éléments de  $\mathbb{R}^{\mathbb{N}}$ ), on peut de même définir la “norme  $\ell^p$ ” ( $p \in \mathbb{N}$ ,  $p > 0$ ) de la façon suivante :

$$\|x\|_p \stackrel{\text{def}}{=} \sqrt[p]{\sum_{i \in \mathbb{N}} |x_i|^p}, \quad (2.7)$$

ce qui n'a de sens que pour les suites de “puissance  $p$  sommable” (ce qui implique au moins que  $x_i \rightarrow 0$  lorsque  $i \rightarrow +\infty$ ). On peut aussi définir la norme  $\|\cdot\|_\infty$  pour les suites bornées.

Pour les fonctions  $\varphi : [0, 1] \rightarrow \mathbb{R}$  continues sur un ouvert contenant  $[0, 1]$ , on peut définir, pour  $p \in \mathbb{N} \setminus \{0\}$ , la norme

$$\|\varphi\|_p \stackrel{\text{def}}{=} \sqrt[p]{\int_0^1 |\varphi(t)|^p dt}. \quad (2.8)$$

Pour des fonctions *non* continues, il faut d'abord se restreindre à celles pour lesquelles l'intégrale ci-dessus est définie. Mais alors, l'expression (2.8) ne définit plus une norme mais seulement une “semi-norme” : en effet, l'intégrale peut être nulle si la fonction est “presque partout nulle” au sens de la mesure de Lebesgue, sans être exactement nulle. On se tire classiquement de cette situation en considérant l'équivalence des fonctions  $\varphi$  et  $\psi$  telles que

$$\int_0^1 |\varphi(t) - \psi(t)|^p dt = 0.$$

<sup>1</sup>La norme  $\|\cdot\|_a$  est équivalente à la norme  $\|\cdot\|_b$ , s'il existe des constantes  $m$  et  $M$  positives telles que, pour tout  $x$ ,  $m \|x\|_a \leq \|x\|_b \leq M \|x\|_a$ .

Sur l'espace des fonctions de  $[0, 1]$  dans  $\mathbb{R}$  telles que l'intégrale (2.8) est définie, on considère alors l'espace quotient par cette relation d'équivalence, et, pour cet espace quotient, on obtient bien une norme. C'est la construction classique des espaces  $L^p([0, 1]; \mathbb{R})$ .

Revenant à des fonctions continues  $\varphi : [0, 1] \rightarrow \mathbb{R}$ , on peut, à l'instar de la norme  $\|\cdot\|_\infty$  introduite en (2.5) pour les espaces  $\mathbb{R}^n$ , introduire la "norme de la convergence uniforme" définie par

$$\|\varphi\|_\infty \stackrel{\text{def}}{=} \sup_{t \in [0, 1]} |\varphi(t)|. \quad (2.9)$$

## 2.4 Fonctions linéaires continues, dual topologique

Une application  $f : \mathcal{X} \rightarrow \mathcal{Y}$  entre deux espaces vectoriels normés est *linéaire* si

$$\forall x \in \mathcal{X}, x' \in \mathcal{X}, \alpha \in \mathbb{R}, \beta \in \mathbb{R}, \quad f(\alpha x + \beta x') = \alpha f(x) + \beta f(x'). \quad (2.10)$$

En particulier, on vérifie que  $f(0) = 0$ . Une application est *affine* si elle est égale à la somme d'une fonction linéaire et d'une fonction constante.

**Exercice 2.3.** Une fonction affine est caractérisée par (2.10) en restreignant  $\alpha$  et  $\beta$  par la condition supplémentaire  $\alpha + \beta = 1$ . (*indication* : montrer que  $x \mapsto f(x) - f(0)$  est linéaire).

On note  $\mathcal{L}(\mathcal{X}, \mathcal{Y})$  l'espace vectoriel des applications linéaires *continues* de  $\mathcal{X}$  dans  $\mathcal{Y}$  muni des opérations

$$\forall f \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), g \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), \alpha \in \mathbb{R}, \quad f + g : x \mapsto f(x) + g(x), \quad \alpha f : x \mapsto \alpha f(x).$$

Si  $\mathcal{X}$  est de dimension finie, toutes les applications linéaires de  $\mathcal{X}$  dans  $\mathcal{Y}$  sont continues. Dans le cas général, une application linéaire  $f : \mathcal{X} \rightarrow \mathcal{Y}$  est continue dès qu'elle est continue en  $x = 0 \in \mathcal{X}$  ; une autre condition nécessaire et suffisante est que la quantité

$$\sup_{x \in \mathcal{X} \setminus \{0\}} \|f(x)\|_{\mathcal{Y}} / \|x\|_{\mathcal{X}} \quad (2.11)$$

soit finie. Cette quantité est en fait une norme (notée  $\|f\|$ ) dans l'espace  $\mathcal{L}(\mathcal{X}, \mathcal{Y})$  qui est complet (donc "de Banach") si  $\mathcal{Y}$  est complet. Une troisième condition nécessaire et suffisante est que  $f$  soit Lipschitzienne, avec précisément  $\|f\|$  comme constante de Lipschitz, c'est-à-dire que

$$\exists \|f\| \in \mathbb{R} : \forall x \in \mathcal{X}, x' \in \mathcal{X}, \quad \|f(x) - f(x')\|_{\mathcal{Y}} \leq \|f\| \|x - x'\|_{\mathcal{X}}. \quad (2.12)$$

On dit que  $f$  est une *forme linéaire* sur  $\mathcal{X}$  lorsque c'est une application linéaire de  $\mathcal{X}$  dans  $\mathbb{R}$ . L'ensemble des formes linéaires *continues* est appelé *dual topologique* de  $\mathcal{X}$  et il est noté  $\mathcal{X}^*$ . Notons que puisque  $\mathbb{R}$  est complet,  $\mathcal{X}^*$  est un espace de Banach. Par ailleurs, comme on l'a déjà dit plus haut, si  $\mathcal{X}$  est de dimension finie, alors toutes les formes linéaires sont continues. Comme on va le voir dans l'exercice ci-après, ce n'est pas nécessairement le cas en dimension infinie.

**Exercice 2.4.** On considère l'espace  $\mathcal{X}$  des fonctions continues de  $[0, 1]$  dans  $\mathbb{R}$  muni de la norme  $\|\cdot\|_1$  telle que définie par (2.8) pour  $p = 1$ . Sur cet espace, on considère la forme linéaire qui, à une fonction continue  $\varphi(\cdot)$ , associe sa valeur  $\varphi(0)$  à  $t = 0$ . C'est évidemment une forme linéaire. Montrer qu'elle n'est pas continue pour la norme indiquée.<sup>2</sup> Pour cela, considérer la suite de fonctions continues  $\{\varphi_n\}$  telles que

$$\varphi_n(t) \stackrel{\text{def}}{=} \begin{cases} 1 - nt & \text{si } t \in [0, 1/n], \\ 0 & \text{si } t \in [1/n, 1], \end{cases}$$

et montrer que cette suite tend vers la fonction 0 dans la topologie considérée mais que la forme linéaire appliquée à la suite ne tend pas vers 0.


<sup>2</sup>Elle le serait par contre avec la norme de la convergence uniforme (2.9).

## 2.5 Topologie faible

La topologie associée à la norme dans un espace de Banach est désignée comme la “topologie forte”, et le dual topologique  $\mathcal{X}^*$  a été défini comme l'ensemble des formes linéaires (applications linéaires de  $\mathcal{X}$  dans  $\mathbb{R}$ ) continues pour cette topologie (et la topologie habituelle de  $\mathbb{R}$ ). La *topologie faible* sur  $\mathcal{X}$  peut être définie comme la topologie “minimale”, ou “la plus grossière” telle que les éléments de  $\mathcal{X}^*$  restent continus. Autrement dit, une suite  $\{x_n\}$  tend vers  $x$  *faiblement* (ce qui sera notée  $x_n \rightharpoonup x$ ) si, pour tout  $f \in \mathcal{X}^*$ ,  $f(x_n)$  tend vers  $f(x)$ .

Évidemment, de la définition même de la convergence faible, on déduit que si la suite  $x_n$  tend fortement vers  $x$ , elle converge aussi faiblement vers  $x$ . En dimension finie, il n'y a pas en fait de différence entre les topologies faible et forte. Ce n'est plus le cas en dimension infinie comme le montrera plus loin l'Exercice 2.10.

On montre par ailleurs que si  $x_n \rightharpoonup x$ , alors  $\|x_n\|$  est bornée et  $\|x\| \leq \liminf \|x_n\|$ .

**Exercice 2.5.**  Montrer que si  $x_n \rightharpoonup x$  et si  $f_n$  tend (fortement) vers  $f$  dans  $\mathcal{X}^*$ , alors  $f_n(x_n)$  tend vers  $f(x)$ .

## 2.6 Opérateur adjoint

On considère une application linéaire continue  $A$  (on dit aussi un “opérateur linéaire”) d'un espace de Banach  $\mathcal{X}$  dans un autre espace  $\mathcal{Y}$  et une forme linéaire continue  $f \in \mathcal{Y}^*$ . Il est clair que l'application composée  $g \stackrel{\text{def}}{=} f \circ A$  est une forme linéaire continue sur  $\mathcal{X}$ , c'est-à-dire un élément de  $\mathcal{X}^*$ . On vient donc d'associer, grâce à l'opérateur  $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ , un élément  $g \in \mathcal{X}^*$  à un élément  $f \in \mathcal{Y}^*$ . Cette application, notée  $A^* : f \mapsto g$ , est évidemment une fonction linéaire de son argument  $f$ . Par ailleurs, on montre facilement (résultat général sur les applications composées) que

$$\|g\|_{\mathcal{X}^*} \leq \|f\|_{\mathcal{Y}^*} \|A\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y})} ,$$

ce qui prouve que l'application  $A^*$  est continue avec  $\|A^*\|_{\mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)} \leq \|A\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y})}$ . En fait, on peut montrer l'égalité de ces deux normes d'opérateur. L'opérateur  $A^*$  est appelé *opérateur adjoint* de  $A$ .

On retiendra donc (voir Figure 2.2) que

$$\forall x \in \mathcal{X}, \forall f \in \mathcal{Y}^*, \forall A \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), \quad f(A(x)) = A^*(f)(x) . \tag{2.13}$$

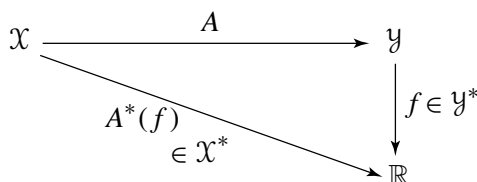


Figure 2.2: Diagramme définissant l'opérateur adjoint

Il deviendra plus clair dans le cadre des espaces de Hilbert ci-après que la notion d’“adjoint” généralise la notion plus familière de “transposée” d’une matrice (cas d’espaces de dimension finie). On poursuivra les développements sur l’opérateur adjoint dans ce cadre.

## 2.7 Espaces de Hilbert

### 2.7.1 Produit scalaire et dualité

Dans un espace vectoriel  $\mathcal{X}$ , un *produit scalaire* est une forme bilinéaire  $\omega$  symétrique et définie positive, c'est-à-dire une application de  $\mathcal{X} \times \mathcal{X}$  dans  $\mathbb{R}$  satisfaisant

- symétrie :  $\omega(x, y) = \omega(y, x)$  ;
- (bi-)linéarité :  $\omega$  est linéaire continue par rapport à son premier argument (et donc aussi par rapport à son deuxième argument par symétrie) ;
- définie positivité :  $\omega(x, x) \geq 0$  et  $\omega(x, x) = 0 \Rightarrow x = 0$ .

Un espace *euclidien* est un espace vectoriel normé dans lequel la norme est définie à partir d'un produit scalaire, c'est-à-dire que

$$\|x\| \stackrel{\text{def}}{=} \sqrt{\omega(x, x)} . \quad (2.14)$$

**Exercice 2.6.** Vérifier que (2.14) définit bien une norme.

**Exercice 2.7.** Démontrer l'inégalité de Cauchy-Schwarz :

$$|\omega(x, y)| \leq \|x\| \|y\| . \quad (2.15)$$

(*indication* : considérer  $\omega(x + \alpha y, x + \alpha y)$  pour  $\alpha \in \mathbb{R}$  quelconque).

Un *espace de Hilbert* est un espace euclidien complet (pour la topologie de la norme associée au produit scalaire). On parlera désormais d'espace de Hilbert même dans le cas où le caractère complet n'intervient pas.

L'intérêt des espaces de Hilbert est que l'on dispose d'une notion d'"angle" de deux vecteurs (c'est-à-dire de deux éléments de l'espace, ou plus précisément des vecteurs orientés qui joignent l'origine à ces deux éléments). En effet, grâce à l'inégalité (2.15), on voit que le rapport  $\omega(x, y) / \|x\| \|y\|$  est compris entre  $-1$  et  $1$ , et on peut donc l'interpréter comme le cosinus de l'angle des deux vecteurs : si le produit scalaire est nul, l'angle est de  $\pm\pi/2$  et les vecteurs sont "orthogonaux" ; si l'égalité à lieu dans (2.15), l'angle est  $0$  ou  $\pi$  et les vecteurs sont "colinéaires".

**Exercice 2.8.** Dans un espace de Hilbert, montrer l'"égalité du parallélogramme" :

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) . \quad (2.16)$$

En fait, un espace de Banach dans lequel la norme vérifie la propriété (2.16) est un espace de Hilbert. On montre en effet que l'expression

$$\omega(x, y) \stackrel{\text{def}}{=} \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \quad (2.17a)$$

définit un produit scalaire et que la norme est associée à ce produit scalaire. D'autres expressions possibles de ce produit scalaire sont (en utilisant (2.16)) :

$$\omega(x, y) = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2) \quad (2.17b)$$

$$= \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x - y\|^2) . \quad (2.17c)$$



Parmi toutes les normes (2.5) définies dans  $\mathbb{R}^n$ , la seule norme associée à un produit scalaire (donc vérifiant (2.16)) est la norme  $\|\cdot\|_2$ . Le produit scalaire associé est le produit scalaire usuel de  $\mathbb{R}^n$  :

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i . \quad (2.18a)$$

De la même façon, pour des fonctions de  $[0, 1]$  dans  $\mathbb{R}$  de carré sommable, la norme  $L^2$  (voir (2.8) avec  $p = 2$ ) est associée au produit scalaire

$$\langle \varphi(\cdot), \psi(\cdot) \rangle \stackrel{\text{def}}{=} \int_0^1 \varphi(t) \psi(t) dt . \quad (2.18b)$$

On utilisera désormais plutôt la notation  $\langle \cdot, \cdot \rangle$  que la notation  $\omega(\cdot, \cdot)$  pour désigner le produit scalaire.


Une propriété fondamentale des espaces de Hilbert est qu'ils sont identifiables à leur dual. Ceci est une conséquence du

**Théorème 2.9 (de représentation de Riesz).** *Dans un espace de Hilbert  $\mathcal{X}$ , tout élément  $x_f \in \mathcal{X}$  définit une forme linéaire continue par le procédé suivant :  $x \mapsto f(x) \stackrel{\text{def}}{=} \langle x_f, x \rangle$ . Réciproquement, à toute forme linéaire continue  $f \in \mathcal{X}^*$  peut être associée un unique élément  $x_f \in \mathcal{X}$  tel que*

$$\forall x \in \mathcal{X}, \quad f(x) = \langle x_f, x \rangle .$$

De plus,  $\|f\|_{\mathcal{X}^*} = \|x_f\|_{\mathcal{X}}$ . L'application  $\iota : x_f \mapsto f$  est donc une bijection isométrique entre  $\mathcal{X}$  et son dual.

Ce résultat est à la base de la relative simplicité de manipulation des espaces de Hilbert et de leur dual qui s'apparente assez à celle, plus familière, de l'espace  $\mathbb{R}^n$  et de son produit scalaire, pourvu qu'on omette la mention explicite de l'isomorphisme  $\iota$  et qu'on ne distingue pas  $\mathcal{X}$  et  $\mathcal{X}^*$ , ce que nous ferons la plupart du temps.<sup>3</sup> Une précaution à prendre, en dimension infinie, est de ne pas confondre la convergence forte (en norme) et la convergence faible définie au §2.5. L'exercice suivant montre un cas de convergence faible qui n'est pas forte.

**Exercice 2.10.**  On considère l'espace  $\ell^2$  des suites dans  $\mathbb{R}$  de carré sommable (voir (2.7)). C'est un espace de Hilbert et toutes les formes linéaires continues sont obtenues comme on vient de le voir par produit scalaire avec des éléments de même type. On considère dans cet espace la suite  $\{x^k\}$  telle que

$$x_i^k \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } i = k, \\ 0 & \text{sinon.} \end{cases}$$

Montrer que cette suite converge faiblement mais pas fortement vers 0.

Une autre précaution à prendre en dimension infinie est de ne pas assimiler les sous-ensembles bornés fermés (pour la norme) avec les sous-ensembles compacts (par rapport à la topologie forte).<sup>4</sup> En fait, les sous-ensembles bornés et faiblement fermés (donc aussi fortement fermés) sont *faiblement compacts* (c'est-à-dire, compacts dans la topologie faible) mais pas nécessairement *fortement compacts*.

<sup>3</sup>On rencontrera cependant dans une circonstance — lors de l'introduction d'une relation d'ordre par l'intermédiaire d'un cône convexe — la nécessité de bien distinguer l'espace primal de son dual.

<sup>4</sup>On rappelle que dans un sous-ensemble compact, toute suite infinie admet au moins un point d'accumulation.

### 2.7.2 Opérateur adjoint dans les espaces de Hilbert

Terminons cette section en revenant sur la notion d'opérateur adjoint qui a été défini par (2.13). Dans un espace de Hilbert, cette définition prend la forme

$$\forall x \in \mathcal{X}, \forall y \in \mathcal{Y}, \forall A \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), \quad \langle y, A(x) \rangle = \langle A^*(y), x \rangle. \quad (2.19)$$

Cette nouvelle relation fait mieux ressortir le lien avec la notion de transposition d'une matrice : lorsque  $\mathcal{X} = \mathbb{R}^n$  et  $\mathcal{Y} = \mathbb{R}^m$ , l'opérateur linéaire  $A$  est représentable par une matrice  $m \times n$  et  $A^*$  par la matrice  $n \times m$  transposée. D'une manière générale, par l'identification des espaces primal et dual, on peut considérer que  $A^* \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ .

Avec la relation (2.19), on voit que

$$x \in \ker A \stackrel{\text{def}}{=} \{x \mid A(x) = 0\} \Rightarrow \langle A^*(y), x \rangle = 0, \text{ c'est-à-dire que } x \in (\text{im } A^*)^\perp,$$

en désignant par  $\mathcal{V}^\perp$  le sous-espace orthogonal à un sous-espace vectoriel  $\mathcal{V}$ .<sup>5</sup> Autrement dit,

$$\ker A \subset (\text{im } A^*)^\perp. \quad (2.20)$$

De cela, on déduit que si  $A^*$  est *surjectif*, c'est-à-dire que  $\text{im } A^* = \mathcal{X}$  et donc  $(\text{im } A^*)^\perp = \{0\}$ , alors  $\ker A = \{0\}$ , c'est-à-dire que  $A$  est *injectif*. De la même façon, on montre que  $\ker A^* \subset (\text{im } A)^\perp$ , d'où l'on déduit que si  $A$  est *surjectif*, alors  $A^*$  est *injectif*. Les énoncés réciproques de ces énoncés sont vrais en dimension finie mais pas en dimension infinie, comme l'illustre l'exercice suivant.

**Exercice 2.11.** Dans l'espace  $\ell^2$  déjà utilisé à l'Exercice 2.10, on considère l'application de  $\ell^2$  dans lui-même définie par

$$A : x = (x_0, x_1, \dots, x_i, \dots) \mapsto y = (x_0, x_1/2, \dots, x_i/(i+1), \dots).$$

Montrer que  $A^* = A$ , que  $A$ , donc aussi  $A^*$ , sont injectifs mais pas surjectifs, et que l'égalité n'a donc pas lieu dans (2.20).

Comme vient de le montrer cet exercice, lorsque  $\mathcal{X} = \mathcal{Y}$ , un opérateur  $A$  peut être égal à son opérateur adjoint : on dit alors qu'il est *auto-adjoint*, ce qui généralise la notion de matrice symétrique en dimension finie. D'une manière générale, un opérateur  $A \in \mathcal{L}(\mathcal{X}, \mathcal{X})$  se décompose en la somme de sa partie auto-adjointe ou symétrique  $(A + A^*)/2$  et de sa partie antisymétrique  $(A - A^*)/2$ .

**Exercice 2.12.** Démontrer que les opérateurs antisymétriques, c'est-à-dire ceux qui sont tels que  $A^* = -A$ , “font tourner de  $\pm\pi/2$ ”, c'est-à-dire que  $\langle x, Ax \rangle = 0, \forall x$ . En déduire que pour construire une “forme quadratique” :  $x \mapsto \langle x, Ax \rangle$ , seule la partie symétrique de  $A$  est vraiment utile.

## 2.8 Résumons nous

Le but de ce court chapitre était de rappeler brièvement un certain nombre de notions et résultats supposés déjà connus avant d'aborder le cours. En allant du plus particulier vers le plus général, la structure euclidienne de  $\mathbb{R}^n$  résulte de l'existence d'un produit scalaire duquel est dérivée une norme qui définit elle-même une topologie. La même démarche peut être suivie dans des espaces “de Hilbert” de dimension infinie (comme par exemple, en commande optimale, l'ensemble des fonctions du temps — continu

<sup>5</sup>On donnera au chapitre suivant, avec une notation similaire, une définition du *cône orthogonal en un point à un sous-ensemble*. On verra que lorsque le sous-ensemble est un sous-espace vectoriel, les deux notions coïncident.

— de “carré sommable”) avec quelques précautions supplémentaires à prendre pour distinguer la topologie forte associée à la norme de la topologie faible qui est la topologie minimale dans laquelle toutes les formes linéaires continues pour la topologie forte restent encore continues.

Par le théorème de Riesz, l'ensemble des formes linéaires continues, dit “espace dual (topologique)” peut être mis en bijection avec l'espace “primal”, l'action d'une forme linéaire continue sur un élément de l'espace primal se “confondant” alors avec celle d'un produit scalaire entre deux éléments de l'espace primal. C'est, dans les espaces de Hilbert, l'équivalent du fait qu'il n'y a pas lieu de distinguer fondamentalement, dans  $\mathbb{R}^n$ , un vecteur colonne (primal) d'un vecteur ligne (dual). De même, la notion de “transposée d'une matrice” en dimension finie trouve son équivalent en dimension infinie sous la forme de la notion d’“adjoint” d'un opérateur linéaire (continu).

Il y a cependant, comme on l'a déjà dit, quelques précautions à prendre en dimension infinie vis-à-vis de la manipulation de deux topologies (qui se confondent en dimension finie) : dans la topologie forte, les bornés fermés ne sont pas nécessairement compacts.

## 2.9 Corrigé des exercices

**Corrigé de l'Exercice 2.5** On a

$$\begin{aligned} |f(x) - f_n(x_n)| &\leq |f(x) - f(x_n)| + |f(x_n) - f_n(x_n)| \\ &\leq |f(x) - f(x_n)| + \|f - f_n\| \|x_n\| . \end{aligned}$$

Le premier terme tend vers 0 du fait de la convergence faible de  $x_n$  vers  $x$ , et le second tend vers 0 du fait que  $\|x_n\|$  est bornée et que  $f_n$  tend fortement vers  $f$ .

**Corrigé de l'Exercice 2.10** On note d'abord que  $\|x^k\| = 1$  pour tout  $k$ , ce qui exclut la convergence forte de la suite vers 0. Montrons que la suite converge par contre faiblement vers 0. Comme on est dans un espace de Hilbert, les formes linéaires continues sont les produits scalaires par les éléments de  $\ell^2$  lui-même. Soit  $y = (y_0, y_1, \dots) \in \ell^2$ . Nécessairement,  $\lim_{i \rightarrow +\infty} y_i = 0$  puisque la suite est de carré sommable. Or

$$\langle y, x^k \rangle = y_k ,$$

ce qui montre que  $\lim_{k \rightarrow +\infty} \langle y, x^k \rangle = 0$  pour tout  $y \in \ell^2$  : c'est bien la convergence faible vers 0.



## Chapitre 3

# Ensembles convexes

Dans ce chapitre et celui qui suivra, on se limite à quelques faits fondamentaux de l'analyse convexe qui nous sont indispensables pour la suite du cours. De bonnes références pour aller plus loin sur ce sujet sont les ouvrages de Moulin et Fogelman-Soulié [12], Ekeland et Temam [8], [2], Hiriart-Urruty et Lemaréchal [9], ainsi que l'un des ouvrages "fondateurs", celui de Rockafellar [13].

### 3.1 Sous-espaces affines et enveloppe affine

**Définition 3.1.** Soit  $\mathcal{X}$  un espace vectoriel réel. Un sous-ensemble  $A$  de  $\mathcal{X}$  est un sous-espace affine si

$$\forall x \in A, y \in A, \forall \alpha \in \mathbb{R}, \quad \alpha x + (1 - \alpha)y \in A . \quad (3.1)$$

Autrement dit, un sous-espace affine contient toujours la "droite" passant par deux de ses points  $x$  et  $y$ .

**Exercice 3.2.** Montrer que si  $A$  est un sous-espace affine contenant 0, c'est un sous-espace vectoriel. En déduire que, pour tout  $a \in A$ , le translaté  $A - a$  est un sous-espace vectoriel.

Parce que l'intersection de sous-espaces affines est un sous-espace affine, et que l'ensemble des sous-espaces affines contenant  $A$  n'est pas vide (puisque  $\mathcal{X}$  en est un), la définition suivante a du sens.

**Définition 3.3.** L'enveloppe affine d'un sous-ensemble  $A$ , notée  $\mathfrak{N}(A)$ , est le plus petit sous-espace affine contenant  $A$ .

**Exercice 3.4.** Montrer que

$$\mathfrak{N}(A) = \left\{ x \mid n \in \mathbb{N} \setminus \{0\}, x = \sum_{i=1}^n \alpha_i x_i, x_i \in A, \sum_{i=1}^n \alpha_i = 1 \right\} .$$

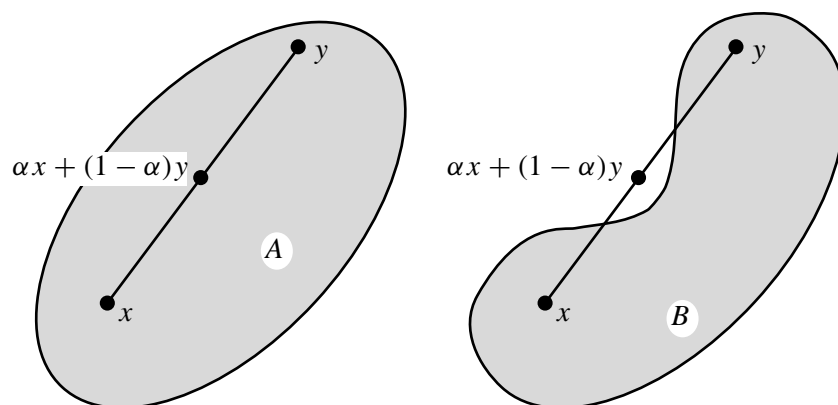
### 3.2 Premières définitions sur les sous-ensembles convexes

**Définition 3.5.** Soit  $\mathcal{X}$  un espace vectoriel réel. Un sous-ensemble  $A$  de  $\mathcal{X}$  est *convexe* si

$$\forall x \in A, y \in A, \forall \alpha \in [0, 1], \quad \alpha x + (1 - \alpha)y \in A . \quad (3.2)$$

Autrement dit, un sous-ensemble convexe contient toujours le segment  $[x, y]$  joignant deux de ses points  $x$  et  $y$ .<sup>1</sup> Une interprétation "optique" consiste à dire que dans une pièce convexe, deux personnes peuvent toujours s'apercevoir (voir Figure 3.1). Un sous-espace affine est évidemment convexe.

<sup>1</sup>Il est clair que ce que l'on appelle "segment"  $[x, y]$  est justement l'ensemble des points de la forme  $\alpha x + (1 - \alpha)y$  lorsque  $\alpha$  parcourt  $[0, 1]$ .

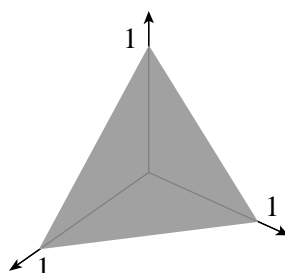
Figure 3.1:  $A$  est convexe,  $B$  n'est pas convexe

**Exercice 3.6.** Montrer que les seuls sous-ensembles convexes de  $\mathbb{R}$  sont les segments (au sens large, c'est-à-dire aussi les demi-droites), donc en fait les sous-ensembles *connexes* de  $\mathbb{R}$ .

**Définition 3.7.** On appelle *simplexe* de  $\mathbb{R}^n$  le sous-ensemble


$$\Sigma_n \stackrel{\text{def}}{=} \left\{ \alpha \in \mathbb{R}^n \mid \alpha_i \geq 0, \quad i = 1, \dots, n, \quad \sum_{i=1}^n \alpha_i = 1 \right\}. \quad (3.3)$$

Notons que nécessairement  $\alpha_i \leq 1$  pour  $i = 1, \dots, n$ . La Figure 3.2 représente le simplexe de  $\mathbb{R}^3$ .

Figure 3.2: Le simplexe de  $\mathbb{R}^3$ 

**Définition 3.8.** On appelle *combinaison convexe* de  $n$  points  $\{x_i\}_{i=1,\dots,n}$  tout point  $y$  obtenu par la formule

$$y = \sum_{i=1}^n \alpha_i x_i \quad \text{avec} \quad \alpha \in \Sigma_n. \quad (3.4)$$

**Exercice 3.9.**  Montrer qu'une combinaison convexe de  $n$  points peut être calculée récursivement par  $n - 1$  combinaisons convexes de 2 points.

Cet exercice montre que dans la Définition 3.5, on pouvait remplacer de façon équivalente la combinaison convexe de 2 points par celle de  $n$  points en acceptant des valeurs de  $n$  quelconques. En fait, dans un espace de dimension finie  $n$ , il est inutile de considérer des combinaisons convexes de plus de  $n + 1$  points, comme le montre le théorème de Caratheodory.<sup>2</sup>

<sup>2</sup>Nous énonçons ici un lemme que l'on peut aussi attribuer à Caratheodory. Le théorème de Carathéodory sera énoncé plus loin comme un corollaire immédiat de ce lemme.

**Lemme 3.10 (Caratheodory).** *Dans un espace vectoriel de dimension  $n$ , toute combinaison convexe de  $m$  points,  $m > n + 1$ , se ramène à une combinaison convexe de  $n + 1$  points au plus.*

*Démonstration.* Soit

$$y = \sum_{i=1}^m \alpha_i x_i, \quad \alpha \in \Sigma_m .$$

Il suffit de montrer que si  $m > n + 1$ , on peut faire décroître  $m$  de 1. Ce fait, utilisé de façon répétée, prouvera le résultat.

Si on considère les vecteurs  $z_i \stackrel{\text{def}}{=} x_i - x_1$  pour  $i = 2, \dots, m$ , en nombre au moins égal à  $n + 1$ , ces vecteurs ne peuvent être linéairement indépendants, et il existe donc une combinaison linéaire nulle, à coefficients  $\beta_i$  non tous nuls, de cet ensemble, donc

$$\sum_{i=2}^m \beta_i z_i = 0 \quad \Rightarrow \quad \sum_{i=1}^m \beta_i x_i = 0 \quad \text{avec} \quad \sum_{i=1}^m \beta_i = 0 ,$$

en ayant posé

$$\beta_1 \stackrel{\text{def}}{=} - \sum_{i=2}^m \beta_i .$$

On considère alors

$$\gamma \stackrel{\text{def}}{=} \max_{i=1, \dots, m} \left( \frac{\beta_i}{\alpha_i} \right) \quad \text{et} \quad \delta_i = \alpha_i - \frac{\beta_i}{\gamma}, \quad i = 1, \dots, m .$$

Noter que  $\gamma \neq 0$  car les  $\beta_i$  ne sont pas tous nuls. On vérifie maintenant que

$$y = \sum_{i=1}^m \alpha_i x_i = \sum_{i=1}^m \delta_i x_i ,$$

et que  $\delta \in \Sigma_m$  avec au moins un  $\delta_i$  égal à 0. On a donc bien exprimé  $y$  comme une combinaison convexe de moins de  $m$  points.  $\square$

L'ensemble des combinaisons convexes de  $n$  points isolés est un sous-ensemble convexe (appelé *polytope* ou *polyèdre*).

**Remarque 3.11.** Il ne faut pas conclure du lemme de Caratheodory qu'on peut, dans  $\mathbb{R}^2$  par exemple, engendrer tous les polyèdres avec au maximum 3 points (c'est-à-dire des triangles au plus) ! La Figure 3.3 prouve le contraire.

### 3.3 Opérations préservant la convexité

Les faits suivants sont énoncés sans démonstration car ils résultent presque immédiatement de la Définition 3.5 (ils peuvent donc être considérés par le lecteur comme des exercices faciles !). Partout,  $C$  désigne un sous-ensemble convexe de l'espace  $\mathcal{X}$ .

- L'intersection de sous-ensembles convexes est convexe.
- Pour  $a \in \mathcal{X}$ , le *translaté*  $a + C \stackrel{\text{def}}{=} \{a + x \mid x \in C\}$  est convexe.

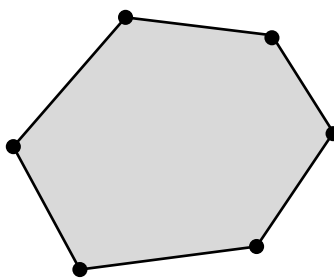


Figure 3.3: Un polyèdre à 6 sommets dans  $\mathbb{R}^2$ , c'est possible !

- Pour  $\alpha \in \mathbb{R}$ , l'*homothétie*  $\alpha C \stackrel{\text{def}}{=} \{\alpha x \mid x \in C\}$  est convexe.
- L'*image* de  $C$  par une *application affine* de  $\mathcal{X}$  dans un autre espace vectoriel  $\mathcal{Y}$  est convexe.
- L'*image réciproque* de  $C$  par une *application affine*  $f$  de  $\mathcal{Y}$  dans  $\mathcal{X}$ ,  $f^{-1}(C) \stackrel{\text{def}}{=} \{y \in \mathcal{Y} \mid f(y) \in C\}$ , est convexe.
- La *somme vectorielle* de  $C$  et d'un autre sous-ensemble convexe  $C' \subset \mathcal{X}$ , c'est-à-dire  $C + C' \stackrel{\text{def}}{=} \{x + x' \mid x \in C, x' \in C'\}$ , est convexe.
- Le *produit cartésien* de  $C \subset \mathcal{X}$  et  $C' \subset \mathcal{Y}$ , c'est-à-dire  $C \times C' \stackrel{\text{def}}{=} \{(x, y) \mid x \in C, y \in C'\}$ , est un sous-ensemble convexe de  $\mathcal{X} \times \mathcal{Y}$ .
- Inversement, la *projection* d'un sous-ensemble convexe d'un espace produit sur l'un de ses sous-espaces composants est convexe (voir Figure 3.4).
- L'union de sous-ensembles convexes n'est pas convexe en général, mais l'union *croissante* de convexes (famille emboîtée) est convexe.

Une conséquence de ce qui précède est que les solutions d'un ensemble d'égalités et d'inégalités affines constitue un sous-ensemble convexe.

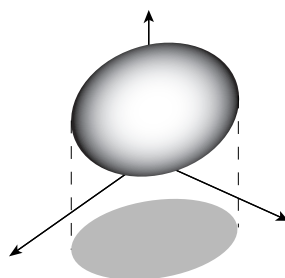



Figure 3.4: Projection d'un convexe

**Exercice 3.12.**  Montrer qu'étant donné un sous-ensemble convexe  $C$  et deux réels *positifs*  $\alpha$  et  $\beta$ , alors  $\alpha C + \beta C = (\alpha + \beta)C$ .

### 3.4 Enveloppe convexe

Étant donné un sous-ensemble  $A$  de  $\mathcal{X}$ , l'espace  $\mathcal{X}$  est un convexe contenant  $A$ . De plus, l'intersection de convexes contenant  $A$  étant convexe et contenant  $A$ , on peut poser la définition suivante.



**Définition 3.13.** L'enveloppe convexe d'un sous-ensemble  $A \subset \mathcal{X}$  quelconque est le plus petit convexe (au sens de l'inclusion) qui contient  $A$ . Elle est notée  $\odot(A)$ .

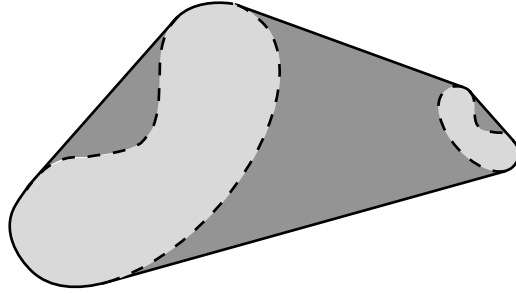


Figure 3.5: Enveloppe convexe (en sombre) d'un sous-ensemble (en clair)

La construction de l'enveloppe convexe suggérée ci-dessus (intersection des convexes contenant  $A$ ) est une construction "externe". Le théorème suivant propose une construction "interne" de  $\odot(A)$ .

**Théorème 3.14.** L'enveloppe convexe de  $A$  est égale à l'ensemble des combinaisons convexes d'éléments de  $A$ .

*Démonstration.* Soit  $B$  l'ensemble des combinaisons convexes d'éléments de  $A$ . Il est clair que  $B \subset \odot(A)$  puisque  $\odot(A)$ , en tant que sous-ensemble convexe, doit contenir les combinaisons convexes de ses points, et en particulier ceux de  $A$ . Par ailleurs, si on montre que  $B$  est convexe, alors comme évidemment  $B \supset A$ , on aura que  $B \supset \odot(A)$  d'après la définition de ce dernier sous-ensemble. Si on prend deux combinaisons convexes d'éléments de  $A$ , on doit montrer qu'une combinaison convexe de ces deux combinaisons convexes est encore une combinaison convexe d'éléments de  $A$ . Tout tient dans la remarque que, pour  $n$  nombres  $\alpha_i$ ,  $m$  nombres  $\beta_j$  et un nombre  $\gamma$ , tous dans  $[0, 1]$ , on a

$$\sum_{i=1}^n \alpha_i = 1 \text{ et } \sum_{j=1}^m \beta_j = 1 \quad \Rightarrow \quad \sum_{i=1}^n \gamma \alpha_i + \sum_{j=1}^m (1 - \gamma) \beta_j = 1 .$$

□

On est maintenant en mesure d'énoncer le théorème de Caratheodory.

**Théorème 3.15 (Caratheodory).** Dans un espace vectoriel de dimension  $n$ , l'enveloppe convexe d'un sous-ensemble  $A$  est égale à l'ensemble des combinaisons convexes de  $n + 1$  points de  $A$ .

Ce théorème est un simple corollaire du Lemme 3.10 et du Théorème 3.14.

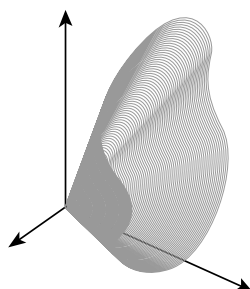
## 3.5 Cônes convexes

### 3.5.1 Cône convexe et relation d'ordre

Dans la formulation des contraintes inégalités, les cônes joueront ultérieurement un rôle essentiel.

**Définition 3.16.** Un sous-ensemble  $C$  est un cône si

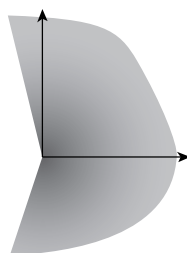
$$\forall x \in C, \forall \alpha \geq 0, \alpha x \in C . \quad (3.5)$$

Figure 3.6: Un cône non convexe de  $\mathbb{R}^3$ 

Un cône est donc une union de demi-droites fermées issues de l'origine.<sup>3</sup> La Figure 3.6 représente un cône de  $\mathbb{R}^3$  (d'ailleurs non convexe). Les espaces vectoriels sont évidemment des cônes. Un cône est dit *saillant* si

$$C \cap (-C) = \{0\}. \quad (3.6)$$

Un espace vectoriel n'est pas un cône saillant comme ne l'est pas non plus le cône représenté sur la Figure 3.7. Par contre, un espace vectoriel est un cône convexe.

Figure 3.7: Un cône non saillant dans  $\mathbb{R}^2$ 

**Exercice 3.17.** ☺ Montrer qu'un cône  $C$  est convexe si et seulement si il est stable par addition, c'est-à-dire que

$$\forall x \in C, \forall y \in C, x + y \in C. \quad (3.7)$$

Les cônes convexes permettent de définir une *relation de préordre* ou une *relation d'ordre* dans un espace vectoriel.

**Exercice 3.18.** ☺ On rappelle qu'une relation de préordre est une relation *réflexive* et *transitive* et qu'une relation d'ordre est de plus *antisymétrique*. Montrer que si  $C$  est un cône convexe d'un espace vectoriel  $\mathcal{X}$ , la relation

$$x \succeq y \Leftrightarrow x - y \in C \quad (3.8)$$

est une relation de préordre, et que c'est une relation d'ordre si de plus  $C$  est saillant.

On notera que la relation d'ordre ainsi définie est compatible avec la structure d'espace vectoriel, c'est-à-dire qu'à l'évidence, d'après (3.5), si  $x$  est "positif", sa multiplication par un élément positif

<sup>3</sup>Implicitement donc, le sommet d'un cône est toujours situé à l'origine de l'espace vectoriel. On peut si nécessaire déplacer le sommet en  $a \in \mathcal{X}$  en translatant le cône, c'est-à-dire en considérant le sous-ensemble  $a + C$ .

redonne un élément positif, et de plus, d'après (3.7), la somme de deux éléments positifs est un élément positif. Enfin, dans un espace de Banach, pour que la limite d'une suite convergente (pour la topologie forte) d'éléments positifs soit un élément positif, on prendra un "cône positif" *fermé*. On verra plus loin que, du fait que les ensembles convexes fermés sont aussi *faiblement* fermés, l'ordre est aussi continu pour la topologie faible.

**Remarque 3.19 (Ordre de  $\mathbb{R}^n$ ).** L'utilisation la plus classique de cette relation d'ordre est obtenue en choisissant, dans  $\mathbb{R}^n$ , le cône constitué par le premier "orthant", c'est-à-dire l'ensemble des  $x \in \mathbb{R}^n$  tels que  $x_i \geq 0, i = 1, \dots, n$ . La "non négativité" de  $x$  revient alors à celle de toutes ses composantes. Mais en choisissant un autre cône saillant comme "cône positif" de  $\mathbb{R}^n$ , on peut obtenir une relation d'ordre contenant une plus ou moins grande quantité d'éléments positifs.

### 3.5.2 Cône normal (ou orthogonal) à un sous-ensemble convexe

**Définition 3.20.** Dans un espace de Hilbert  $\mathcal{X}$ , le *cône normal* ou *orthogonal* à un sous-ensemble convexe  $A$  au point  $a \in A$  est défini par

$$A_a^\perp = \{z \in \mathcal{X} \mid \langle z, x - a \rangle \leq 0, \forall x \in A\}.$$

En premier lieu, il est facile de voir que ceci définit bien un cône, et même un cône convexe fermé comme intersection d'une collection de demi-espaces fermés. On notera  $A^\perp$  pour  $A_0^\perp$ , lorsque bien sûr  $0 \in A$ . Évidemment,  $(A - a)^\perp = A_a^\perp$ . La Figure 3.8 présente des exemples de cônes normaux à un sous-ensemble en divers points de la frontière

**Remarque 3.21.** Bien que ces cônes soient figurés avec leur sommet au point de la frontière où ils sont définis, la représentation correcte serait celle où leur sommet est ramené à 0 par translation. Autrement, dit, il faut à chaque fois traduire l'ensemble  $A$  pour que le point en question soit à l'origine, comme l'indique la définition.

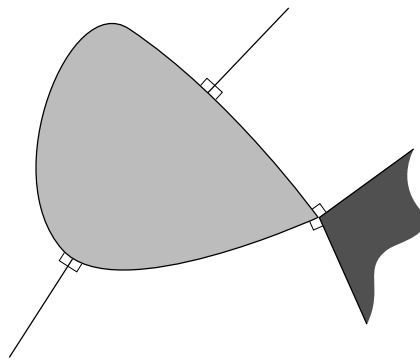



Figure 3.8: Cônes normaux en divers points de  $A$

L'exercice suivant montre que la notion de cône normal n'a d'intérêt qu'aux points appartenant à la frontière de  $A$  (c'est-à-dire, n'appartenant pas à l'intérieur de  $A$ ).

**Exercice 3.22.**  Montrer que si  $a \in \overset{\circ}{A}$  (intérieur de  $A$ , supposé donc non vide et dont on verra qu'il est convexe au §3.6), alors  $A_a^\perp = \{0\}$ .

**Remarque 3.23.** On aurait pu définir la notion de cône normal en un point pour un sous-ensemble quelconque et pas seulement pour les sous-ensembles convexes. Cela aurait conduit à des cônes éventuellement réduits à  $\{0\}$  y compris pour certains points de la frontière.

L'exercice suivant fait le lien entre cette notion de cône normal et celle de sous-espace orthogonal lorsque  $A$  est un sous-espace vectoriel ou affine.

**Exercice 3.24.** Si  $A$  est sous-espace vectoriel ou affine, montrer que  $A_a^\perp$  ne dépend pas de  $a$  et que c'est le sous-espace vectoriel des  $z$  tels que  $\langle z, x - a \rangle = 0, \forall x \in A, \forall a \in A$  (on le note donc  $A^\perp$ ).

Enfin, l'exercice suivant montre que si  $A$  est un cône convexe, le cône normal maximal est obtenu en 0.

**Exercice 3.25.** © Soit  $C$  un cône convexe. Montrer que pour  $a \in C, C_a^\perp = C^\perp \cap \{a\}^\perp$ , où  $\{a\}$  désigne le sous-espace vectoriel de dimension 1 engendré par le vecteur  $a$ . En déduire que  $\{a\}^\perp \cap C^\perp = \{0\}$  si  $a \in \overset{\circ}{C}$ .

**Remarque 3.26.** Dans la littérature, on trouve la notion de *cône polaire* d'un sous-ensemble  $A$ , généralement notée  $A^\circ$ , et définie comme

$$A^\circ \stackrel{\text{def}}{=} \{z \in \mathcal{X} \mid \langle z, x \rangle \leq 0, \forall x \in A\} .$$

On voit que cette définition coïncide avec ce que nous notons  $A_0^\perp$  ou plus simplement  $A^\perp$ , à cette différence importante près que nous ne pouvons pas définir  $A^\perp$  si 0 n'appartient pas à  $A$ . Avec la notion d'enveloppe cônica  $\angle(A)$  introduite plus loin (voir Définition 3.29 ci-après), qui, elle, contient toujours 0, on déduit facilement de la définition de  $A^\circ$  que (voir Figure 3.10)

$$A^\circ = (\angle(A))^\circ = (\angle(A))^\perp .$$

### 3.5.3 Ordre dual dans les espaces de Hilbert

Plaçons nous à nouveau dans un espace de Hilbert  $\mathcal{X}$  dont on a vu qu'il est identifiable à son dual. Cependant, il est important ici de distinguer, pour un élément donné  $\widehat{x}$ , le fait qu'il soit considéré comme un élément de l'espace primal  $\mathcal{X}$  ou comme un élément de l'espace dual, auquel cas il faudrait en toute rigueur parler de  $\iota(\widehat{x}) : x \mapsto \langle \widehat{x}, x \rangle$  (la bijection  $\iota$  a été introduite au Théorème 2.9). Nous ferons l'économie de la notation  $\iota$  mais nous préciserons  $\widehat{x} \in \mathcal{X}$  ou  $\widehat{x} \in \mathcal{X}^*$  pour indiquer si nous parlons d'un élément ou de sa forme linéaire associée.

Étant donné un cône positif (convexe, fermé, saillant) dans  $\mathcal{X}$ , on pose

$$C^* \stackrel{\text{def}}{=} \{\widehat{x} \in \mathcal{X}^* \mid \langle \widehat{x}, x \rangle \geq 0, \forall x \in C \subset \mathcal{X}\} . \quad (3.9)$$

Il est clair que modulo l'identification de  $\mathcal{X}$  à  $\mathcal{X}^*$ ,  $C^* = -C^\perp$ , et que  $C^*$  est un cône convexe fermé de  $\mathcal{X}^*$ . Ce cône n'est pas nécessairement saillant (considérer le cas où  $C$  est réduit à une demi-droite et alors  $C^*$  est un demi-espace), mais il peut servir à définir au moins un préordre sur l'espace  $\mathcal{X}^*$ , voire un ordre si le cône est saillant. Il est important de noter qu'en général, même si on identifie  $\mathcal{X}$  à son dual, il n'y a pas de raison a priori pour que  $C = C^*$ . Ceci se produit pourtant lorsqu'on choisit par exemple dans  $\mathbb{R}^n$  le cône positif  $\mathbb{R}_+^n \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$  (voir Remarque 3.19). Cependant, la Figure 3.9 montre, pour  $n = 2$ , deux cas où  $C$  est plus grand ou plus petit que  $C^*$ . Autrement dit, "être positif" ne veut pas dire la même chose dans le primal et dans le dual. C'est pourquoi, on préférera la notation explicite  $x \in C$  ou  $x \in C^*$  à la notation ambiguë  $x \geq 0$ .

**Exercice 3.27.** © Montrer qu'étant donnés deux cônes  $C$  et  $D$ ,  $(C \cup D)^* = (C + D)^* = C^* \cap D^*$ , et que par ailleurs  $(C \cap D)^* \supset C^* \cup D^*$  en donnant un exemple qui prouve que l'égalité n'a pas lieu en général.

**Remarque 3.28.** On peut s'intéresser à itérer l'opération de dualité en échangeant les rôles de l'espace primal et de l'espace dual pour calculer  $C^{**}$ . Il est immédiat de constater que  $C \subset C^{**}$ . En fait, il résulte du fait que  $C$  est un cône convexe fermé et du théorème sur la séparation des convexes, ou du lemme de Farkas qui en est une conséquence, que  $C^{**} = C$ . Nous y reviendrons plus loin (voir Lemme 3.61 ci-après).

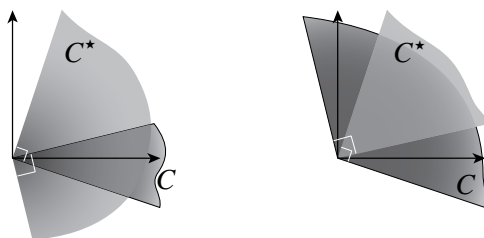


Figure 3.9: Cône et cône dual

### 3.5.4 Enveloppe cônique

À l'instar de l'enveloppe convexe d'un sous-ensemble  $A$ , on peut définir l'enveloppe cônique de  $A$ . L'espace  $\mathcal{X}$  est un cône convexe contenant  $A$  et l'intersection de cônes convexes contenant  $A$  contient aussi  $A$  et est un cône convexe. On adopte donc la définition suivante.

**Définition 3.29.** L'enveloppe cônique d'un sous-ensemble  $A \subset \mathcal{X}$  quelconque est le plus petit cône convexe (au sens de l'inclusion) qui contient  $A$ . Elle est notée  $\angle(A)$ .

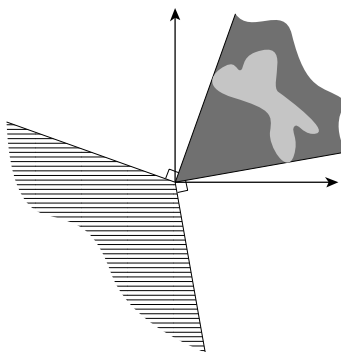


Figure 3.10: Enveloppe cônique (en sombre) d'un sous-ensemble (en clair) et cône polaire associé (rayé)

On montre facilement (exercice) le théorème suivant.

**Théorème 3.30.** L'enveloppe cônique de  $A$  est égale à l'enveloppe cônique de  $\odot(A)$  ; elle est également obtenue en effectuant toutes les combinaisons linéaires à coefficients non négatifs d'éléments de  $A$ .

Par un raisonnement très similaire à celui qui a conduit au Théorème de Caratheodory, on montre le théorème suivant.

**Théorème 3.31.** Dans un espace vectoriel de dimension  $n$ , l'enveloppe cônique d'un sous-ensemble  $A$  est égale à l'ensemble des combinaisons linéaires à coefficients non négatifs de  $n + 1$  points de  $A$ .

### 3.5.5 Cône tangent à un sous-ensemble

**Définition 3.32.** Soit  $A$  un sous-ensemble d'un espace vectoriel normé  $\mathcal{X}$ . On dit que  $x$  est *tangent* à  $A$  au point  $a \in A$  si il existe une suite  $\{\alpha^k\}_{k \in \mathbb{N}}$  de réels positifs et une suite  $\{a^k\}_{k \in \mathbb{N}}$  d'éléments de  $A$  telles que

$$\lim_{k \rightarrow +\infty} \alpha^k = a \quad \text{et} \quad \lim_{k \rightarrow +\infty} \alpha^k (a^k - a) = x .$$

L'ensemble des  $x$  tangents à  $A$  en  $a$  est clairement un cône qui est non vide (il contient au moins 0) ; il est appelé *cône tangent* à  $A$  en  $a$  et il est noté  $A_a^\top$ .

Nous noterons de façon abrégée  $A^\top$  pour  $A_0^\top$ , ce qui suppose bien sûr que  $0 \in A$ . La notion de cône tangent est une notion *locale* puisqu'elle ne fait référence qu'à des limites lorsqu'on s'approche du point  $A$  (autrement dit, on n'a pas besoin de savoir à quoi ressemble  $A$  en dehors d'un voisinage du point  $a$ ). On verra cependant au §3.7 que si  $A$  est *convexe*, il existe une définition *globale* de  $A_a^\top$ .

**Exercice 3.33.** Dessiner le cône tangent, au point d'intersection, à l'ensemble constitué dans  $\mathbb{R}^2$  par deux courbes sécantes.

La Figure 3.11 montre quelques exemples de cônes tangents en divers points d'un convexe (à nouveau, ce qui est en fait représenté, ce sont les cônes tangents *translatés* aux points où ils ont été calculés : un cône a toujours 0 comme sommet selon notre définition des cônes).

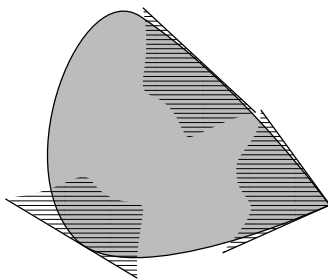


Figure 3.11: Cônes tangents à un sous-ensemble convexe

Comme pour les cônes normaux, la notion de cône tangent n'a d'intérêt qu'en des points  $a$  appartenant à la frontière de  $A$  comme le montre l'exercice suivant.

**Exercice 3.34.** Montrer que si  $a \in \overset{\circ}{A} \subset X$ , alors  $A_a^\top = X$ .

Nous poursuivrons au §3.7 l'étude des cônes tangents et de leur relation avec les cônes normaux une fois introduite, dans la section suivante, la notion d'enveloppe cônica fermée.

## 3.6 Propriétés topologiques des convexes

### 3.6.1 Ouverture et fermeture des convexes

On considère un convexe  $C$  dans un espace vectoriel normé.

**Théorème 3.35.** Si  $C$  est convexe, alors l'intérieur  $\overset{\circ}{C}$  de  $C$  et son adhérence  $\overline{C}$  sont aussi convexes.

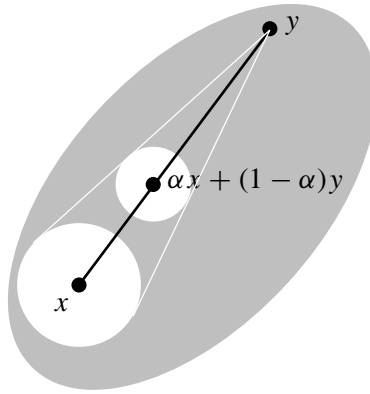
*Démonstration.* Si  $x$  et  $y$  appartiennent à  $\overline{C}$ , il existe des suites  $\{x_n\}$  et  $\{y_n\}$  d'éléments de  $C$  convergeant respectivement vers  $x$  et  $y$ , et pour tout  $\alpha \in [0, 1]$ ,  $\alpha x_n + (1 - \alpha)y_n$  appartient à  $C$  et converge vers  $\alpha x + (1 - \alpha)y$  qui appartient donc à  $\overline{C}$ , donc  $\overline{C}$  est convexe.

Si  $x$  et  $y$  appartiennent à  $\overset{\circ}{C}$ , il existe une boule  $B(x, \rho)$  incluse dans  $C$ . Pour tout  $\alpha \in [0, 1]$ , on montre que la boule  $B(\alpha x + (1 - \alpha)y, \alpha\rho)$  est incluse dans  $C$ , ce qui prouve que  $\alpha x + (1 - \alpha)y$  appartient à  $\overset{\circ}{C}$  qui est donc convexe. La Figure 3.12 explique la situation mieux qu'un long discours. Nous laissons le lecteur mettre en forme le fait qu'à tout  $z \in B(\alpha x + (1 - \alpha)y, \alpha\rho)$ , on peut faire correspondre un  $z' \in B(x, \rho)$  tel que  $z = \alpha z' + (1 - \alpha)y$  et que donc  $z$  appartient bien à  $C$ .  $\square$

**Exercice 3.36.** Montrer que si  $C$  est convexe et si  $\overset{\circ}{C}$  est non vide, alors

$$\overline{\overset{\circ}{C}} = \overset{\circ}{\overline{C}} \quad \text{et} \quad \overline{\overset{\circ}{C}} = \overline{\overline{C}}.$$

La condition  $\overset{\circ}{C} \neq \emptyset$  est évidemment essentielle (penser à un segment de droite dans  $\mathbb{R}^2$  par exemple). En fait, il existe une notion qui permet de s'affranchir de ce problème d'intérieur vide des convexes, au moins en *dimension finie*, c'est celle d'*intérieur relatif*.

Figure 3.12: Une boule centrée sur  $x$  et son homothétique de centre  $y$ 

### 3.6.2 Intérieur relatif

**Définition 3.37.** Pour un sous-ensemble convexe  $C$  non vide d'un espace vectoriel normé de dimension finie, on appelle *intérieur relatif* de  $C$ , noté  $\overset{\circ}{C}$ , l'intérieur de  $C$  pour la topologie induite sur  $\aleph(C)$ . On appelle *dimension* de  $C$  la dimension du sous-espace vectoriel parallèle à  $\aleph(C)$  (c'est-à-dire  $\aleph(C) - a$ , pour tout  $a \in \aleph(C)$ ).

**Théorème 3.38.** Pour tout sous-ensemble convexe  $C$  non vide d'un espace vectoriel normé de dimension finie,  $\overset{\circ}{C}$  est non vide.

*Démonstration.* Supposons que la dimension de  $\aleph(C)$  soit  $n$ . Soit  $x_0 \in C$ , il existe  $n$  vecteurs  $x_i$  de  $C$  tels que  $\{x_i - x_0\}_{i=1, \dots, n}$  constituent une base de l'espace vectoriel  $\aleph(C) - x_0$ . Autrement dit, tout élément de  $y \in \aleph(C)$  s'écrit, avec  $\alpha_i \in \mathbb{R}$ ,  $i = 1, \dots, n$ ,

$$y = \sum_{i=1}^n \alpha_i (x_i - x_0) + x_0 = \sum_{i=0}^n \alpha_i x_i ,$$

en ayant posé

$$\alpha_0 \stackrel{\text{def}}{=} 1 - \sum_{i=1}^n \alpha_i .$$

L'application  $f$  qui à  $\alpha \in \mathbb{R}^n$  associe le  $y \in \aleph(C)$  défini par la formule ci-dessus est un isomorphisme continu entre  $\mathbb{R}^n$  et  $\aleph(C)$ . Considérons alors l'ouvert suivant de  $\mathbb{R}^n$  :

$$\Omega \stackrel{\text{def}}{=} \left\{ \alpha \in \mathbb{R}^n \mid \alpha_i > 0, i = 1, \dots, n, \sum_{i=1}^n \alpha_i < 1 \right\} .$$

Son image  $f(\Omega)$  est donc un ouvert de  $\aleph(C)$ . Nous allons montrer que  $f(\Omega) \subset C$ , ce qui montrera que  $C$  contient un ouvert de  $\aleph(C)$  et donc l'intérieur de  $C$  dans  $\aleph(C)$  (que nous avons défini comme étant  $\overset{\circ}{C}$ ) est non vide. À cet effet, observons que le  $\alpha_0$  associé au  $\alpha \in \Omega$  (tel que  $\sum_{i=0}^n \alpha_i = 1$ ) est (strictement) positif, et que donc tout  $y$  dans l'ouvert  $f(\Omega)$  est une combinaison convexe d'éléments de  $C$  : il appartient donc bien à  $C$ .  $\square$

### 3.6.3 Enveloppe convexe fermée

Alors que l'enveloppe convexe d'un sous-ensemble ouvert est ouverte (exercice), l'enveloppe convexe d'un sous-ensemble fermé n'est pas nécessairement fermée.<sup>4</sup> Par exemple, si on considère dans  $\mathbb{R}^2$  le sous-ensemble fermé  $\{(0, 0)\} \cup \{(x, y) \mid x \geq 0, xy = 1\}$  (voir Figure 3.13, partie gauche), son enveloppe convexe est le sous-ensemble  $\{(0, 0)\} \cup \{(x, y) \mid x > 0, y > 0\}$  (voir Figure 3.13, partie droite) qui n'est pas fermé. Pour cette raison, il est nécessaire de recourir à la notion d'*enveloppe convexe fermée*

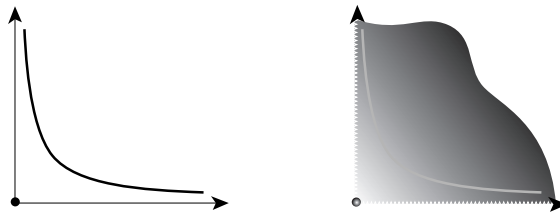


Figure 3.13: Un sous-ensemble et son enveloppe convexe

**Définition 3.39.** L'*enveloppe convexe fermée* d'un sous-ensemble  $A \subset \mathcal{X}$  quelconque est le plus petit convexe fermé (au sens de l'inclusion) qui contient  $A$ . Elle est notée  $\overline{\text{co}}(A)$ .

À nouveau, cette définition a un sens puisque l'intersection de convexes fermés est un convexe fermé. Le théorème ci-dessous indique une autre façon de définir  $\overline{\text{co}}(A)$ .

**Théorème 3.40.** L'*enveloppe convexe fermée* d'un sous-ensemble  $A$  est égale à la fermeture (ou l'adhérence) de son enveloppe convexe, c'est-à-dire  $\overline{\text{co}}(A)$ .

*Démonstration.* D'une part, il est clair que la fermeture de l'enveloppe convexe  $\overline{\text{co}}(A)$  est un convexe fermé qui contient le sous-ensemble  $A$ , et donc aussi son enveloppe convexe fermée  $\overline{\text{co}}(A)$  qui est le plus petit convexe fermé ayant cette propriété. D'autre part, puisque  $\text{co}(A)$  est contenue dans tout convexe (fermé ou pas) contenant  $A$ , son adhérence  $\overline{\text{co}}(A)$  est contenue dans tout convexe fermé contenant  $A$  et en particulier  $\overline{\text{co}}(A)$ .  $\square$

Il est facile de voir que l'enveloppe convexe fermée de  $A$  est aussi l'enveloppe convexe fermée de la fermeture  $\overline{A}$  de  $A$ . Mais, comme l'a montré l'exemple de la Figure 3.13,  $\overline{\text{co}}(\overline{A})$  n'est pas nécessairement fermée, donc n'est pas nécessairement égale à  $\overline{\text{co}}(A)$ . En résumé,

$$\overline{\text{co}}(A) = \overline{\text{co}(A)} = \overline{\text{co}(\overline{A})} = \overline{\text{co}}(\overline{A}) \supset \text{co}(\overline{A}) . \quad (3.10)$$

Il faut donc retenir que les opérations d'adhérence (ou de fermeture) et d'enveloppe convexe ne commutent pas et que la fermeture doit agir en dernier pour obtenir le plus grand sous-ensemble.

Le théorème suivant montre que, pour la dernière inclusion ci-dessus, l'égalité a lieu si l'espace est de dimension finie, et si le sous-ensemble  $A$  est borné.

**Théorème 3.41.** Dans un espace vectoriel normé  $\mathcal{X}$  de dimension finie, si  $A \subset \mathcal{X}$  est borné, respectivement compact, alors  $\overline{\text{co}}(A)$  est bornée, respectivement compacte.

<sup>4</sup>On verra cependant plus loin qu'en dimension finie, si l'ensemble est borné, en plus d'être fermé, donc s'il est compact, l'enveloppe convexe est aussi compacte, donc fermée.



*Démonstration.* Si  $n$  est la dimension de l'espace, par le Théorème 3.15, si  $x \in \odot(A)$ , alors  $x$  s'écrit comme une combinaison convexe de  $n + 1$  points d'éléments de  $A$  dont la norme est bornée par un nombre, et il est immédiat de borner la norme de  $x$  par le même nombre, ce qui montre que  $\odot(A)$  est bornée.

Si  $A$  est compact, pour montrer que  $\odot(A)$  est compacte, il faut montrer que toute suite infinie dans  $\odot(A)$  admet un point d'accumulation dans  $\odot(A)$ . Tout élément  $x^k$  de la suite s'écrit comme une combinaison convexe de  $n + 1$  d'éléments  $x_i^k$ ,  $i = 1, \dots, n + 1$ , avec un vecteur  $\alpha^k \in \Sigma_{n+1}$ , le simplexe de dimension  $n + 1$  (voir (3.3)). Ce dernier est aussi compact. On peut donc extraire de la suite  $\{\alpha^k\}$  une première sous-suite convergente. Pour cette sous-suite d'indices  $k$ , on peut extraire de la sous-suite  $\{x_1^k\}$  correspondante une sous-suite convergente. On continue ainsi à extraire des sous-suites successives convergentes des  $\{x_i^k\}$  jusqu'à  $i = n + 1$ . Pour la sous-suite d'indices finalement obtenue, on en déduit qu'il existe un point d'accumulation de la suite  $\{x^k\}$  qui s'exprime comme une combinaison convexe d'éléments de  $A$ , et donc il appartient à  $\odot(A)$ .  $\square$

### 3.6.4 Enveloppe cônica fermée

L'exemple de la Figure 3.13 peut aussi servir à montrer que l'enveloppe cônica d'un fermé n'est pas nécessairement fermée. On pose donc la définition suivante que l'on justifie par les arguments habituels.

**Définition 3.42.** L'enveloppe cônica fermée d'un sous-ensemble  $A \subset \mathcal{X}$  quelconque est le plus petit cône convexe fermé (au sens de l'inclusion) qui contient  $A$ . Elle est notée  $\overline{\angle}(A)$ .

On démontre un théorème analogue au Théorème 3.40. En résumé, à l'instar de (3.10),

$$\overline{\angle}(A) = \overline{\angle(A)} = \overline{\angle(\overline{A})} = \overline{\angle(\overline{A})} \supset \angle(\overline{A}) . \quad (3.11)$$

Par contre, l'analogue du Théorème 3.41 pour les enveloppes côniques, plus précisément le fait qu'en dimension finie, l'enveloppe cônica d'un compact  $A$  est fermée, est faux (considérer par exemple dans  $\mathbb{R}^2$  la boule fermée de centre  $(1, 0)$  et de rayon 1), mais cela devient vrai avec l'hypothèse que  $0$  n'appartient pas à  $\odot(A)$  (voir [9, Chap. III, Proposition 1.4.7]).

### 3.6.5 Fermeture des convexes dans la topologie faible

Le résultat exprimé par le Corollaire 3.56 énoncé plus loin aurait une place naturelle ici. Mais sa démonstration dépend de résultats sur la séparation des convexes qui seront abordés en fin de chapitre, ce qui nous oblige à reporter à plus tard cet énoncé. On fait d'ores et déjà remarquer à ce propos qu'une propriété — la convexité — ayant au départ une définition purement géométrique dans les espaces vectoriels a finalement une conséquence de nature topologique dans les espaces de Hilbert.

## 3.7 Orthogonalité du cône tangent et du cône normal

On commence par montrer que le cône tangent à un convexe  $A$  en  $a$  est un cône convexe fermé et qu'il ne dépend en fait que de l'enveloppe cônica de  $A - a$ .<sup>5</sup>

**Théorème 3.43.** Soit  $A$  un sous-ensemble convexe d'un espace vectoriel normé  $\mathcal{X}$ , et soit  $a \in A$ . Alors  $A_a^\top$  est un cône convexe fermé et c'est en fait l'enveloppe cônica fermée de  $A - a$  (ou encore de  $\overline{A} - a$  d'après (3.11)).

<sup>5</sup>Cette observation est également vraie pour le cône normal comme il résultera du lien entre ces deux cônes.

*Démonstration.* Montrons d'abord la convexité de  $A_a^\top$ . D'après l'Exercice 3.17, il suffit de montrer que ce cône est stable par addition. Pour cela, on considère deux éléments  $x_1$  et  $x_2$  dans  $A_a^\top$  et on montre que  $x_1 + x_2$  appartient aussi à  $A_a^\top$ . Pour  $i = 1, 2$ , il existe des suites de scalaires positifs  $\{\alpha_i^k\}_{k \in \mathbb{N}}$  et des suites d'éléments dans  $A$ ,  $\{a_i^k\}_{k \in \mathbb{N}}$ , telles que  $\alpha_i^k(a_i^k - a)$  converge vers  $x_i$ . Par conséquent,

$$\begin{aligned} x_1 + x_2 &= \lim_{k \rightarrow +\infty} (\alpha_1^k(a_1^k - a) + \alpha_2^k(a_2^k - a)) \\ &= \lim_{k \rightarrow +\infty} (\alpha_1^k + \alpha_2^k) \left( \frac{\alpha_1^k}{\alpha_1^k + \alpha_2^k} (a_1^k - a) + \frac{\alpha_2^k}{\alpha_1^k + \alpha_2^k} (a_2^k - a) \right) \\ &= \lim_{k \rightarrow +\infty} (\alpha_1^k + \alpha_2^k) \left( \frac{\alpha_1^k}{\alpha_1^k + \alpha_2^k} a_1^k + \frac{\alpha_2^k}{\alpha_1^k + \alpha_2^k} a_2^k - a \right). \end{aligned}$$

Comme  $A$  est convexe, le dernier terme entre parenthèses est de la forme  $(a^k - a)$  où  $a^k \in A$ . On a donc trouvé deux suites  $\{(\alpha_1^k + \alpha_2^k)\}_{k \in \mathbb{N}}$  et  $\{a^k\}_{k \in \mathbb{N}}$  qui montrent que  $x_1 + x_2 \in A_a^\top$ .

Montrons maintenant que  $A_a^\top$  est fermé. Pour cela, on considère une suite  $\{x_n\}_{n \in \mathbb{N}} \subset A_a^\top$  convergeant vers  $x$  et on montre que  $x \in A_a^\top$ . À chaque élément  $x_n$  de  $A_a^\top$  sont associées deux suites  $\{\alpha_n^k\}_{k \in \mathbb{N}}$  et  $\{a_n^k\}_{k \in \mathbb{N}}$  : il existe alors un indice  $k_n$  tel que

$$\|\alpha_n^{k_n} (a_n^{k_n} - a) - x_n\| \leq 1/n.$$

Les suites  $\{\alpha_n^{k_n}\}_{n \in \mathbb{N}}$  et  $\{a_n^{k_n}\}_{n \in \mathbb{N}}$  sont telles que  $\alpha_n^{k_n} (a_n^{k_n} - a)$  converge vers  $x$ , ce qui montre le résultat recherché.

Montrons l'inclusion  $\overline{Z}(A - a) \subset A_a^\top$ . Soit  $x \in A$ . Les suites  $\{\alpha^k\}_{k \in \mathbb{N}}$ , avec  $\alpha^k = k$ , et  $\{a^k\}_{k \in \mathbb{N}}$ , avec  $a^k = (1 - 1/k)a + x/k$ , sont telles que  $\lim \alpha^k(a^k - a) = x - a$ , et de plus  $a^k \in A$  parce que  $A$  est convexe. On a donc montré que  $x - a \in A_a^\top$ . Donc  $A - a \subset A_a^\top$ . Mais  $A_a^\top$  étant un cône convexe fermé qui contient  $A - a$ , il contient aussi  $\overline{Z}(A - a)$  qui est le plus petit cône convexe fermé ayant cette propriété.

Il nous reste enfin à montrer l'inclusion inverse. Tout  $x \in A_a^\top$  s'écrit comme la limite d'une suite  $\alpha^k(a^k - a)$  avec  $a^k \in A$ . Chacun de ces termes appartient à  $\overline{Z}(A - a)$  par définition de cet ensemble. Celui-ci étant fermé, la limite appartient aussi à cet ensemble.  $\square$

Un corollaire immédiat du théorème précédent montre que, lorsque  $A$  est convexe, le cône tangent et le cône normal au même point sont en étroite relation.

**Corollaire 3.44.** *Si  $A$  est convexe, et  $a \in A$ , le cône tangent à  $A$  en  $a$  est l'orthogonal du cône normal à  $A$  en  $a$ , c'est-à-dire que*

$$(A_a^\top)^\perp = A_a^\perp. \quad (3.12)$$

*Démonstration.* Pour  $A$  convexe et  $a \in A$ , grâce au théorème précédent, nous bénéficions d'une description simplifiée de  $A_a^\top$  :

$$A_a^\top = \overline{Z}(A - a) = \overline{\{y \mid y = \alpha(x - a), \alpha \in \mathbb{R}_+, x \in A\}}.$$

On en déduit immédiatement que

$$(A_a^\top)^\perp = \{z \mid \langle z, \alpha(x - a) \rangle \leq 0, \alpha \in \mathbb{R}_+, x \in A\}$$

(qui est automatiquement fermé). Il est évident que la présence du  $\alpha \geq 0$  n'apporte rien de plus à la définition de l'ensemble ci-dessus, et on reconnaît alors précisément la définition de  $A_a^\perp$ .  $\square$

Le rapprochement des Figures 3.8 et 3.11 illustre ce corollaire.

**Exercice 3.45.** En prolongement de l'Exercice 3.27, montrer que pour deux convexes  $A$  et  $B$  tel que  $a \in A \cap B \neq \emptyset$ , on a

$$(A \cap B)_a^\top \subset A_a^\top \cap B_a^\top \quad \text{et} \quad (A \cap B)_a^\perp \supset A_a^\perp + B_a^\perp ,$$

mais que l'égalité n'a pas lieu par exemple si, dans  $\mathbb{R}^2$ ,  $A$  et  $B$  sont les disques fermés de rayon 1 centrés en  $(-1, 0)$  et  $(1, 0)$ , respectivement, et  $a = (0, 0)$ .

## 3.8 Projection sur les convexes dans les espaces de Hilbert et séparation des convexes

### 3.8.1 Projection sur un convexe fermé

#### 3.8.1.1 Existence et unicité

On considère un point  $x$  dans un espace de Hilbert  $\mathcal{X}$  et un sous-ensemble  $A \subset \mathcal{X}$ . Il arrive souvent dans les applications que l'on doive se contenter d'une "approximation" de  $x$  par un élément de  $A$  parce que, pour une raison ou pour une autre, les éléments de  $A$  sont plus faciles à décrire sur un ordinateur. Au sens de la distance de l'espace de Hilbert, la meilleure approximation de  $x$  est alors l'élément  $y \in A$  qui minimise la distance  $\|y - x\|$  (cette quantité mesurant l'erreur faite dans l'approximation de  $x$  par  $y$ ). Lorsque  $A$  est convexe et fermé, on va voir que ce problème d'approximation est bien posé en ce sens qu'il y a existence et unicité de  $y$ .

**Théorème 3.46.** *Soit  $x$  un élément d'un espace de Hilbert  $\mathcal{X}$  et  $A$  un sous-ensemble convexe fermé de  $\mathcal{X}$ . Il existe un unique point  $y \in A$  tel que*

$$\|y - x\| = \min_{z \in A} \|z - x\| . \quad (3.13)$$

Cet élément  $y$  est caractérisé par l'inéquation variationnelle

$$\forall z \in A, \quad \langle x - y, z - y \rangle \leq 0 . \quad (3.14)$$

L'élément  $y$  est appelé projection de  $x$  sur  $A$  et sera noté  $\Pi_A(x)$ .

*Démonstration.* Pour commencer, on va montrer que (3.13) et (3.14) sont équivalents. En supposant d'abord (3.13) et en considérant un quelconque  $z \in A$  fixé, les éléments de la forme  $z_\alpha = (1 - \alpha)y + \alpha z$ ,  $\alpha \in ]0, 1[$ , appartiennent aussi à  $A$ . Donc

$$0 \geq (\|y - x\|^2 - \|z_\alpha - x\|^2) / \alpha = -\alpha \|z - y\|^2 + 2 \langle x - y, z - y \rangle ,$$

en ayant développé le carré  $\|z_\alpha - x\|^2$  par rapport à  $\alpha$ . En faisant tendre  $\alpha$  vers 0, on obtient (3.14).

Réciproquement, si (3.14) est vrai, alors, pour tout  $z \in A$ ,

$$0 \geq \langle y - x, y - z \rangle = \|y - x\|^2 + \langle y - x, x - z \rangle ,$$

d'où

$$\|y - x\|^2 \leq \langle y - x, z - x \rangle \leq \|y - x\| \|z - x\|$$

par l'inégalité de Schwartz, et finalement,

$$\|y - x\| \leq \|z - x\|$$

en simplifiant par  $\|y - x\|$  (à noter que cette inégalité est trivialement vérifiée si  $\|y - x\| = 0$ ) : on obtient donc (3.13).

Considérons maintenant le problème de l'existence de  $y$ . Soit  $\beta = \inf_{z \in A} \|x - z\|$ . Soit  $\{z^k\}_{k \in \mathbb{N}}$  une "suite minimisante", c'est-à-dire une suite telle que  $\|x - z^k\|^2 < \beta^2 + 1/k$ . Pour deux valeurs  $l$  et  $m$  de  $k$ , on utilise l'"égalité du parallélogramme" (2.16) pour écrire

$$\begin{aligned} \|z^l - z^m\|^2 &= \|(z^l - x) - (z^m - x)\|^2 \\ &= 2\|z^l - x\|^2 + 2\|z^m - x\|^2 - \|(z^l - x) + (z^m - x)\|^2 \\ &= 2\|z^l - x\|^2 + 2\|z^m - x\|^2 - 4\|x - (z^l + z^m)/2\|^2. \end{aligned}$$

Comme  $A$  est convexe,  $(z^l + z^m)/2$  appartient à  $A$  et le dernier terme (avec son signe moins) est inférieur à  $-4\beta^2$  par définition de  $\beta$ . On en déduit donc que

$$\|z^l - z^m\|^2 \leq 2(\beta^2 + 1/l) + 2(\beta^2 + 1/m) - 4\beta^2 = 2(1/l + 1/m).$$

La suite  $\{z^k\}$  est donc une suite de Cauchy et par conséquent elle converge. Soit  $y$  sa limite qui appartient à  $A$  puisque  $A$  est fermé. Donc  $\|y - x\| \geq \beta$ , mais par passage à la limite sur la suite  $\{z^k\}$ , on a aussi  $\|y - x\| \leq \beta$ , et par conséquent  $y$  réalise la distance minimale  $\beta$  de  $x$  à  $A$ .

Il reste enfin à montrer l'unicité d'un tel  $y$ . Supposons qu'il existe deux points  $y$  et  $y'$  vérifiant (3.13), et donc aussi (3.14) dont on a montré que c'est une condition équivalente. On utilise cette inéquation variationnelle pour  $y$  en faisant  $z = y'$ , mais aussi pour  $y'$  en faisant  $z = y$ , et on additionne les deux inégalités ainsi obtenues. On obtient  $\|y - y'\|^2 \leq 0$ , ce qui montre l'égalité de  $y$  et  $y'$ .  $\square$

Il est facile de comprendre que le fait que  $A$  soit *fermé* est nécessaire pour l'*existence* de la projection de  $x$  sur  $A$  (c'est-à-dire pour qu'il existe un point de  $A$  qui réalise le minimum sur  $A$  de la fonction  $z \mapsto \|z - x\|$ ). En dimension finie, cela suffit même à l'existence puisque si  $\beta = \inf_{z \in A} \|z - x\|$ , il suffit de chercher le minimum de la fonction sur l'intersection de  $A$  et de la boule fermée de centre  $x$  et de rayon  $\beta' > \beta$  : en dimension finie, cette boule est compacte et donc son intersection avec le fermé  $A$  l'est également ; on est donc amené à chercher le minimum d'une fonction sur un compact, minimum qui est alors atteint. Il en irait différemment en dimension infinie (voir [12] pour une discussion plus approfondie à ce sujet).

L'importance de la *convexité* de  $A$  pour l'*unicité* de la projection est facile à comprendre. La Figure 3.14 montre un cas de non unicité (mais il suffirait de considérer que  $A$  est constitué de deux points isolés et de choisir  $x$  exactement au milieu du segment joignant ces deux points).

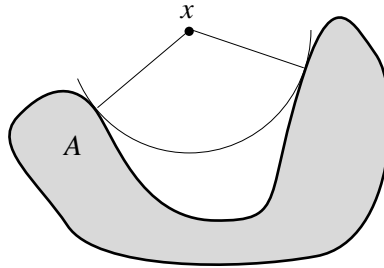


Figure 3.14: Le point  $x$  a deux projections sur  $A$

Le fait que l'espace soit un espace de Hilbert, c'est-à-dire que la distance soit liée à une norme elle-même liée à un produit scalaire joue un rôle également important pour l'unicité de la projection. Considérons les boules pour la norme  $\ell^1$  (norme qui ne provient pas d'un produit scalaire) dont la forme est représentée sur la Figure 2.1. La Figure 3.15 montre que l'origine n'a pas de projection unique au sens de cette distance sur la boule de centre  $(1, 1)$  et de rayon 1 dans  $\mathbb{R}^2$ .

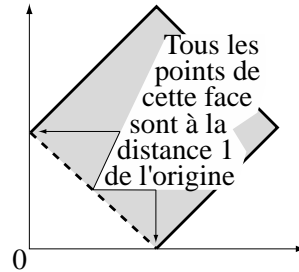


Figure 3.15: Non unicité de la projection avec la norme  $\ell^1$

### 3.8.1.2 Propriétés de la projection

**Théorème 3.47.** *La projection  $\Pi_A$  est un opérateur Lipschitzien de constante 1 et monotone, c'est-à-dire que*

$$\|\Pi_A(x') - \Pi_A(x)\| \leq \|x' - x\|, \quad (3.15)$$

$$\langle \Pi_A(x') - \Pi_A(x), x' - x \rangle \geq 0. \quad (3.16)$$

*Démonstration.* Utilisons la caractérisation (3.14) de la projection pour  $x, y = \Pi_A(x)$  et  $z = \Pi_A(x')$  :

$$\langle \Pi_A(x') - \Pi_A(x), x - \Pi_A(x) \rangle \leq 0,$$

puis additionnons cette inéquation avec celle analogue qui est obtenue en inversant le rôle de  $x$  et  $x'$ . On obtient, après quelques manipulations :

$$\langle \Pi_A(x') - \Pi_A(x), x' - x \rangle \geq \|\Pi_A(x') - \Pi_A(x)\|^2. \quad (3.17)$$

Cette inéquation montre d'une part que  $\Pi_A$  est bien monotone, et d'autre part, après application de l'inégalité de Schwartz et simplification par  $\|\Pi_A(x') - \Pi_A(x)\|$  (si cette quantité est nulle le résultat est trivial), que (3.15) est aussi vérifiée.  $\square$

Autrement dit, la projection est "non expansive", c'est-à-dire qu'elle ne dilate pas les distances.

### 3.8.1.3 Projection sur un cône

La projection sur des cônes convexes fermés jouera un rôle privilégié dans la dualité en optimisation. On a vu à l'Exercice 3.18 qu'un cône convexe fermé  $C$  dans un espace de Hilbert  $\mathcal{X}$  permet de définir une relation d'ordre, et au §3.5.3, on a introduit un ordre associé dans l'espace dual sous la forme d'un cône convexe fermé  $C^*$ . On va voir que tout élément  $x \in \mathcal{X}$  se décompose de façon unique en la somme de sa projection sur  $C$  et de sa projection sur  $-C^*$  et que cette décomposition est orthogonale (dans la suite nous aurons plutôt besoin de la décomposition sur  $-C$  et  $C^*$ ).

**Théorème 3.48.** Soit  $C$  un cône convexe fermé d'un espace de Hilbert  $\mathcal{X}$ . Pour tout  $x \in \mathcal{X}$ , la projection  $\Pi_C(x)$  est caractérisée par les conditions (en dehors bien sûr de la condition évidente  $\Pi_C(x) \in C$ )

$$\langle \Pi_C(x), x - \Pi_C(x) \rangle = 0, \quad (3.18a)$$

$$\forall z \in C, \quad \langle z, x - \Pi_C(x) \rangle \leq 0. \quad (3.18b)$$

La projection  $\Pi_C$  sur un cône est un opérateur positivement homogène de degré 1, c'est-à-dire que

$$\forall \alpha \geq 0, \quad \Pi_C(\alpha x) = \alpha \Pi_C(x). \quad (3.19)$$

De plus,  $x - \Pi_C(x)$  est égal à la projection  $\Pi_{-C^*}(x)$  et  $x$  se décompose de façon unique en

$$x = \Pi_C(x) + \Pi_{-C^*}(x) \text{ avec } \langle \Pi_C(x), \Pi_{-C^*}(x) \rangle = 0. \quad (3.20)$$

*Démonstration.* On considère la caractérisation (3.14) de la projection  $y = \Pi_C(x)$ , et, dans cette inéquation, on peut prendre  $z = 0$  et  $z = 2\Pi_C(x)$  parce que  $C$  est un cône. Des deux inégalités ainsi obtenues, on déduit (3.18a). Alors, de cette égalité et de (3.14), on déduit (3.18b). Réciproquement, la recombinaison des deux relations (3.18) redonne immédiatement (3.14) et ces deux relations sont donc bien une caractérisation de la projection sur le cône convexe fermé  $C$ .

L'inéquation (3.18b) montre immédiatement que  $\xi = x - \Pi_C(x)$  appartient à  $-C^*$ . Pour montrer que l'on a bien  $\Pi_{-C^*}(x) = \xi$ , il faut vérifier les conditions (3.18) pour cette nouvelle projection. La condition (3.18a) pour  $\Pi_C(x)$  se réécrit immédiatement  $\langle \xi, x - \xi \rangle = 0$ , ce qui constitue justement la même condition pour  $\Pi_{-C^*}(x)$ . Au passage, ceci montre l'orthogonalité de  $\xi$  et de  $x - \xi = \Pi_C(x)$ . Il faut maintenant vérifier la condition (3.18b) pour  $\Pi_{-C^*}(x)$ , c'est-à-dire  $\langle z, x - \xi \rangle \leq 0, \forall z \in -C^*$ , mais celle-ci découle immédiatement de la définition de  $-C^*$  et du fait que  $x - \xi = \Pi_C(x) \in C$ .  $\square$

La Figure 3.16 illustre cette décomposition orthogonale (on a noté  $x^+$  pour  $\Pi_C(x)$  et  $x^-$  pour  $\Pi_{-C^*}(x)$ ).

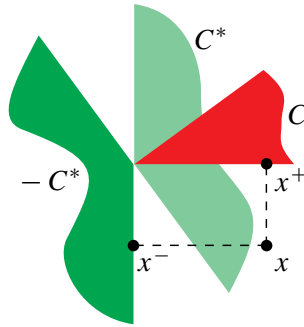


Figure 3.16: Décomposition orthogonale de  $x$

**Remarque 3.49.** On a donc  $\|x\|^2 = \|\Pi_C(x)\|^2 + \|x - \Pi_C(x)\|^2$ . De ce résultat, ou de (3.15) en remarquant que  $0 = \Pi_C(0)$  si  $C$  est un cône (convexe fermé), on déduit que  $\|\Pi_C(x)\| \leq \|x\|$ .

### 3.8.2 Séparation

Les théorèmes de séparation, dits “formes géométriques du théorème de Hahn-Banach” sont directement liés à l'existence d'une projection sur un convexe fermé. Il convient pour comprendre ces résultats d'interpréter géométriquement la caractérisation (3.14) de la projection  $y = \Pi_A(x)$ . Elle nous dit (voir

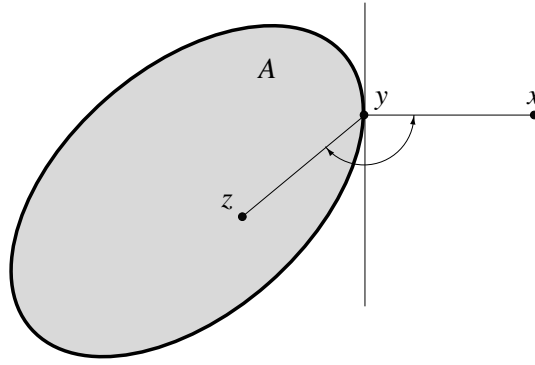
Figure 3.17: Séparation du point  $x$  et du convexe  $A$ 

Figure 3.17) que les vecteurs  $\vec{yx}$  et  $\vec{yz}$  forment un angle *obtus* lorsque  $z$  parcourt  $A$ , ou bien encore, que  $A$  est situé entièrement dans l'un des demi-espaces délimité par l'hyperplan passant par  $y$  et orthogonal au vecteur  $\vec{yx}$ , alors que  $x$  est situé dans l'autre demi-espace. Donc il existe un hyperplan qui "sépare"  $A$  de  $x$  dès que  $x$  n'appartient pas à  $A$ .

**Théorème 3.50.** Soit  $A$  un sous-ensemble convexe fermé d'un espace de Hilbert  $\mathcal{X}$  et  $x \notin A$ . Alors il existe  $r \in \mathcal{X}$  tel que

$$\sup_{z \in A} \langle r, z \rangle < \langle r, x \rangle . \quad (3.21)$$

*Démonstration.* Bien entendu, il suffit de prendre  $r = x - \Pi_A(x)$  qui est non nul si  $x \notin A$ , et de vérifier avec ce  $r$  l'inégalité (3.21). L'inéquation variationnelle (3.14) s'écrit encore (avec  $y = \Pi_A(x) = x - r$ ) :

$$\forall z \in A, \langle r, z - x + r \rangle \leq 0 \Rightarrow \langle r, z \rangle \leq \langle r, x \rangle - \|r\|^2 .$$

Le résultat annoncé en découle immédiatement étant donné que  $r \neq 0$ . □

**Définition 3.51.** Deux sous-ensembles convexes  $A_1$  et  $A_2$  d'un espace de Hilbert sont séparés strictement par la forme linéaire  $x \mapsto \langle r, x \rangle$  si

$$\sup_{x_1 \in A_1} \langle r, x_1 \rangle < \inf_{x_2 \in A_2} \langle r, x_2 \rangle , \quad (3.22a)$$

ou bien

$$\sup_{x_2 \in A_2} \langle r, x_2 \rangle < \inf_{x_1 \in A_1} \langle r, x_1 \rangle . \quad (3.22b)$$

Géométriquement, les deux convexes appartiennent chacun à l'un des demi-espaces délimité par l'hyperplan  $\langle r, x \rangle = \alpha$  pour tout  $\alpha$  compris strictement entre les deux nombres définis dans les deux membres de l'inégalité (3.22a) ou (3.22b) (selon celle qui est valide).

**Corollaire 3.52.** Dans un espace de Hilbert, on considère deux convexes fermés  $A_1$  et  $A_2$  et on suppose de plus que  $A_2$  est compact. Ces deux convexes peuvent être séparés strictement si et seulement si ils sont disjoints.

*Démonstration.* Il est clair que la condition d'être disjoints est nécessaire. Montrons qu'elle est suffisante. Si  $A_1$  et  $A_2$  sont disjoints, alors considérons  $A_1 - A_2 = \{x_1 - x_2 \mid x_1 \in A_1, x_2 \in A_2\}$  qui est convexe ; l'origine n'appartient pas à  $A_1 - A_2$ .

Montrons d'abord que  $A_1 - A_2$  est fermé. Soit une suite  $\{x^k = x_1^k - x_2^k, x_1^k \in A_1, x_2^k \in A_2\}$  convergeant vers  $\bar{x}$ . Il nous faut montrer que  $\bar{x} \in A_1 - A_2$ . Comme  $A_2$  est compact, la suite  $\{x_2^k\}$  admet des points d'accumulation : soit  $\bar{x}_2$  l'un d'entre eux et soit  $\{k_i\}$  une sous-suite d'indices telles  $x_2^{k_i}$  converge vers  $\bar{x}_2$  (évidemment,  $\bar{x}_2 \in A_2$ ). Puisque la suite  $\{x_1^{k_i} - x_2^{k_i}\}$  converge vers  $\bar{x}$ , la suite  $\{x_1^{k_i}\}$  converge vers  $\bar{x} + \bar{x}_2$  qui appartient à  $A_1$  puisque  $A_1$  est fermé. Donc  $\bar{x} \in A_1 - A_2$ .

Puisque  $A_1 - A_2$  est convexe fermé et que l'origine n'appartient pas à ce sous-ensemble, on peut utiliser le Théorème 3.50 pour ce sous-ensemble et ce point, ce qui donne :

$$\begin{aligned} \exists r : \sup_{z \in A_1 - A_2} \langle r, z \rangle &= \sup_{x_1 \in A_1, x_2 \in A_2} \langle r, x_1 - x_2 \rangle \\ &= \sup_{x_1 \in A_1} \langle r, x_1 \rangle + \sup_{x_2 \in A_2} \langle r, -x_2 \rangle \\ &= \sup_{x_1 \in A_1} \langle r, x_1 \rangle - \inf_{x_2 \in A_2} \langle r, x_2 \rangle \\ &< 0 . \end{aligned}$$

□

L'exercice qui suit permet de se convaincre que la compacité de l'un au moins des convexes est requise.

**Exercice 3.53.** Dans  $\mathbb{R}^2$ , on considère les deux convexes fermés suivants :  $A_1 = \mathbb{R}_+^2$  et  $A_2 = \{(x_1, x_2) \mid x_1 < 0, x_1 x_2 \leq -1\}$ . Dessiner la figure correspondante et montrer qu'on ne peut pas séparer strictement  $A_1$  et  $A_2$ .

### 3.8.3 Hyperplans d'appui et description externe des convexes

On peut faire encore une autre "lecture géométrique" de la caractérisation (3.14) de la projection de  $x$  sur un convexe fermé  $A$ . En effet, on voit, en rapprochant cette caractérisation de la Définition 3.20, que le vecteur  $x - \Pi_A(x)$  appartient au cône orthogonal à  $A$  au point  $\Pi_A(x)$ . Ce vecteur n'est pas nul si  $x \neq \Pi_A(x)$  donc si  $x \notin A$ . Réciproquement, en écho à la Remarque 3.23, et au moins en dimension finie, on peut montrer qu'en tout point  $a$  de la frontière de  $A$ , le cône normal  $A_a^\perp$  n'est pas réduit à 0.

**Lemme 3.54.** Soit  $A$  un convexe fermé d'un espace de Hilbert  $\mathcal{X}$  de dimension finie et  $a$  appartenant à la frontière de  $A$  ( $a \in A \setminus \overset{\circ}{A}$ ). Alors,  $A_a^\perp$  n'est pas réduit à 0.

*Démonstration.* Si  $a$  appartient à la frontière de  $A$ , il existe une suite  $\{x^k \notin A\}_{k \in \mathbb{N}}$  et tendant vers  $a$ . Pour tout  $x^k$ , on peut appliquer le Théorème 3.50, c'est-à-dire que

$$\exists r^k \in \mathcal{X} : \sup_{x \in A} \langle r^k, x \rangle < \langle r^k, x^k \rangle . \quad (3.23)$$

On peut de plus choisir les  $r^k$  de norme 1. La boule unité étant compacte (en dimension finie), il existe des points d'accumulation de la suite  $\{r^k\}_{k \in \mathbb{N}}$  : soit  $r$  l'un de ces points d'accumulation, et  $\{k_i\}$  une sous-suite d'indices telle que  $r^{k_i} \rightarrow r$ . Évidemment,  $\|r\| = 1$ , donc  $r \neq 0$ . Par passage à la limite dans (3.23) pour la sous-suite d'indices  $\{k_i\}$ ,

$$\forall x \in A, \quad \langle r, x \rangle \leq \langle r, a \rangle ,$$

ce qui montre que  $r \in A_a^\perp$ .

□



Ce résultat peut aussi être prouvé en dimension infinie avec l'hypothèse supplémentaire  $\check{A} \neq \emptyset$ , mais cela nécessite des théorèmes de séparation plus fins (séparation au sens large, c'est-à-dire séparation non stricte) que nous n'avons pas abordés ici.

Revenons en dimension quelconque. Parmi les hyperplans qui séparent un point  $x \notin A$  du convexe fermé  $A$ , on peut considérer l'hyperplan

$$\{z \mid \langle x - \Pi_A(x), z - \Pi_A(x) \rangle = 0\}$$

qui est orthogonal à la direction  $x - \Pi_A(x)$  et qui passe par  $\Pi_A(x)$ . Cet hyperplan a au moins le point  $\Pi_A(x)$  en commun avec  $A$ , mais on sait que les autres points de  $A$  sont tous situés dans l'un des demi-espaces fermés qu'il délimite. On appelle un tel hyperplan (ayant au moins un point commun avec  $A$  et incluant  $A$  dans l'un de ses demi-espaces fermés) un *hyperplan d'appui* de  $A$  (au point qu'il a en commun avec  $A$ ). La Figure 3.18 montre divers cas d'hyperplans d'appui.

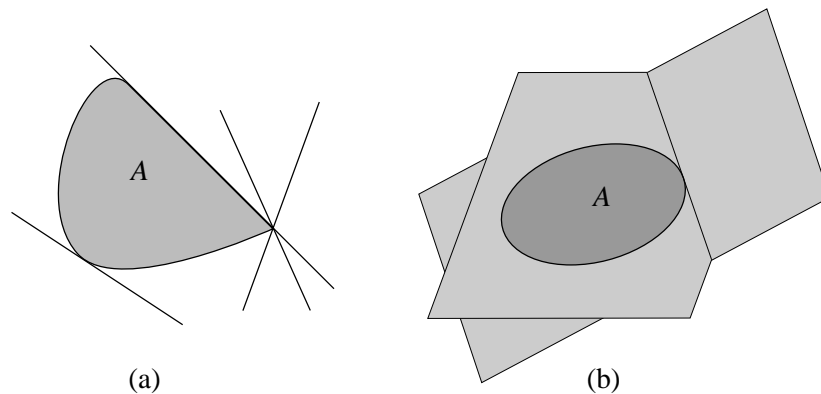


Figure 3.18: (a) Divers hyperplans d'appui à  $A \subset \mathbb{R}^2$  — (b) Deux hyperplans d'appui à  $A$  d'intérieur vide dans  $\mathbb{R}^3$  dont l'un est  $\mathfrak{N}(A)$

**Théorème 3.55.** *Tout convexe fermé  $A$  est égal à l'intersection de tous les demi-espaces fermés qui le contiennent et qui sont délimités par tous les hyperplans d'appui en tous les points de  $A$ .<sup>6</sup>*

*Démonstration.* Soit  $B$  l'intersection de tous les demi-espaces fermés délimités par tous les hyperplans d'appui de  $A$ . Évidemment,  $A \subset B$ . Réciproquement, si un point  $x$  appartenait à  $B$  mais pas à  $A$ , on pourrait le séparer de  $A$  par un hyperplan d'appui, ce qui est en contradiction avec le fait que  $x \in B$  étant donné la façon dont  $B$  a été défini (car tout  $x$  de  $B$  doit de trouver dans le même demi-espace que  $A$  pour tout hyperplan d'appui de  $A$  considéré). Finalement,  $B = A$ .  $\square$

**Corollaire 3.56.** *Dans un espace de Hilbert, un convexe fermé pour la topologie forte est aussi fermé pour la topologie faible.*

**Remarque 3.57.** Notons d'abord que ce résultat n'est pas trivial. En effet, la fermeture dans la topologie faible est *plus forte* que la fermeture dans la topologie forte. Si un sous-ensemble  $A$  quelconque est fermé pour la topologie faible, il est fermé pour la topologie forte : de fait, si une suite contenue dans  $A$  converge fortement vers une limite, elle converge aussi faiblement vers cette limite, et si  $A$  est faiblement fermé, on en conclut que la limite appartient à  $A$ , ce qui prouve alors la fermeture de  $A$  pour la convergence forte. Mais la réciproque est en général fautive. Le corollaire ci-dessus affirme que c'est vrai si  $A$  est *de plus* convexe.

<sup>6</sup>Bien sûr, aux points appartenant à  $\check{A}$ , il ne peut y avoir d'hyperplan d'appui, donc on peut se contenter des points de la frontière de  $A$ .

*Démonstration.* Un convexe fortement fermé est l'intersection de demi-espaces fermés d'après le Théorème 3.55. Par conséquent, il suffit de réaliser que les demi-espaces fermés pour la topologie forte le sont aussi pour la topologie faible. Un demi-espace fermé a été défini par la donnée d'une forme linéaire continue  $\langle r, \cdot \rangle$  et d'un scalaire  $\alpha$  de la façon suivante :

$$\{x \in \mathcal{X} \mid \langle r, x \rangle \leq \alpha\}.$$

C'est donc l'image réciproque d'une demi-droite fermée  $] - \infty, \alpha]$  de  $\mathbb{R}$  par une application linéaire continue (c'est pour cela qu'il est fermé). Mais, par définition même de la topologie faible, les formes linéaires continues sont aussi continues lorsque  $\mathcal{X}$  est muni de la topologie faible. C'est pourquoi le demi-espace est aussi faiblement fermé.  $\square$

Avec le point de vue exprimé par le Théorème 3.55, on obtient donc une “description externe” des convexes fermés illustrée par la Figure 3.19. Cette description est en quelque sorte “duale” de la “description interne” d'un convexe comme ensemble des combinaisons convexes de ses points. La description externe est bien souvent la plus utile dans la pratique.

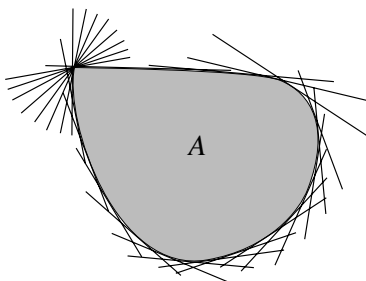


Figure 3.19: Description externe d'un convexe fermé

### 3.8.4 Lemme de Farkas

Une autre conséquence du théorème de séparation est le résultat suivant connu sous le nom de “Lemme de Farkas”.

**Lemme 3.58 (Farkas).** *Considérons une famille  $\{r_j\}_{j \in J}$  d'éléments d'un espace de Hilbert  $\mathcal{X}$  et une famille  $\{\alpha_j\}_{j \in J}$  de nombres réels associés. On suppose que le système d'inégalités*

$$\langle r_j, x \rangle \leq \alpha_j, \quad \forall j \in J,$$

*admet au moins une solution. Alors, étant donné  $s \in \mathcal{X}$  et  $\beta \in \mathbb{R}$ , les deux énoncés suivants sont équivalents :*

$$\forall x \in \mathcal{X}, [\langle r_j, x \rangle \leq \alpha_j, \forall j \in J] \Rightarrow [\langle s, x \rangle \leq \beta], \quad (3.24a)$$

$$(s, \beta) \in \overline{\mathcal{Z}(\{(r_j, \alpha_j)\}_{j \in J} \cup \{(0, 1)\})} \subset \mathcal{X} \times \mathbb{R}, \quad (3.24b)$$

où  $(0, 1) \in \mathcal{X} \times \mathbb{R}$ .

*Démonstration.* Supposons d'abord que  $(s, \beta)$  appartient à  $\mathcal{C} \stackrel{\text{def}}{=} \mathcal{Z}(\{(r_j, \alpha_j)\}_{j \in J} \cup \{(0, 1)\})$ , c'est-à-dire qu'il existe une sous-ensemble fini d'indices  $J_0 \subset J$  et des scalaires non négatifs  $\{\gamma_j\}_{j \in J_0}$  et  $\gamma_0$  tels que

$$s = \sum_{j \in J_0} \gamma_j r_j, \quad \beta = \sum_{j \in J_0} \gamma_j \alpha_j + \gamma_0.$$

Si  $x$  vérifie les inégalités de la première partie de (3.24a), par combinaison linéaire à coefficients non négatifs  $\{\gamma_j\}_{j \in J_0}$  de ces inégalités, on obtient :

$$\sum_{j \in J_0} \gamma_j \langle r_j, x \rangle \leq \sum_{j \in J_0} \gamma_j \alpha_j$$

soit

$$\langle s, x \rangle \leq \beta - \gamma_0 \leq \beta,$$

c'est-à-dire l'inégalité de la deuxième partie de (3.24a). Il est facile de voir que cette inégalité au sens large subsiste si l'on passe à la limite sur l'adhérence  $\overline{C}$ . On a donc montré que (3.24b) implique (3.24a).

Réciproquement, si (3.24a) est vrai mais si (3.24b) ne l'est pas, c'est-à-dire que  $(s, \beta) \notin \overline{C}$ , on peut séparer  $(s, \beta)$  du cône convexe fermé  $\overline{C}$ , ce qui signifie — voir Théorème 3.50 — qu'il existe  $(t, \tau) \in \mathcal{X} \times \mathbb{R}$  (cet espace étant muni du produit scalaire  $\langle (r, \alpha), (s, \beta) \rangle = \langle r, s \rangle_{\mathcal{X}} + \alpha\beta$ ) tel que

$$\sup_{(r, \alpha) \in \overline{C}} (\langle t, r \rangle + \tau\alpha) < \langle t, s \rangle + \tau\beta. \quad (3.25a)$$

En particulier, le sup dans le premier membre de (3.25a) est fini, et il est nécessairement non positif : en effet, s'il était strictement positif, il serait égal à  $+\infty$  puisque  $\overline{C}$  est un cône et on peut multiplier ses vecteurs par n'importe quel scalaire positif arbitrairement grand. Par conséquent,

$$\langle t, r_j \rangle + \tau\alpha_j \leq 0, \quad \forall j \in J, \quad (3.25b)$$

$$\tau \leq 0 \quad (3.25c)$$

(considérer le vecteur  $(0, 1) \in \overline{C}$ ), et

$$\langle t, s \rangle + \tau\beta > 0. \quad (3.25d)$$

À partir de là, il faut distinguer le cas  $\tau = 0$  du cas  $\tau < 0$ . Dans le second cas, on montre que l'élément  $x \stackrel{\text{def}}{=} t/(-\tau)$  viole (3.24a) (rapprocher (3.25b) de (3.25d), ces deux inégalités étant divisées par  $-\tau > 0$ ), ce qui aboutit à une contradiction. Dans le premier cas, partant d'un élément  $x$  qui vérifie les inégalités de la première partie de (3.24a), tous les éléments  $x + \lambda t$  pour  $\lambda \geq 0$  arbitraire les vérifient aussi (d'après (3.25b)), mais du fait de (3.25d), il va exister un  $\lambda$  assez grand pour lequel l'inégalité de la deuxième partie de (3.24a) sera violée. On obtient à nouveau une contradiction.  $\square$

**Corollaire 3.59.** *Si  $J$  est fini dans le lemme précédent, on peut remplacer (3.24b) par*

$$\forall j \in J, \exists \gamma_j \geq 0 : s = \sum_{j \in J} \gamma_j r_j, \quad \beta \geq \sum_{j \in J} \gamma_j \alpha_j.$$

*Démonstration.* Notons que la condition sur  $\beta$  est équivalente à

$$\exists \gamma_0 \geq 0 : \beta = \sum_{j \in J} \gamma_j \alpha_j + \gamma_0.$$

Ce corollaire repose entièrement sur le fait que l'enveloppe cônica d'un nombre fini d'éléments est (déjà) fermée. Ce dernier résultat est facile à montrer si les éléments constituent un ensemble de vecteurs linéairement indépendants (exercice), il nécessite une preuve un peu plus technique dans le cas contraire avec en particulier une construction similaire à celle utilisée dans la démonstration du Lemme de Carathéodory (voir par exemple [9, Lemma 4.3.3]).  $\square$

**Corollaire 3.60.** *Considérons une famille  $\{r_j\}_{j \in J}$  d'éléments d'un espace de Hilbert  $\mathcal{X}$ . On suppose que le système d'inégalités*

$$\langle r_j, x \rangle \leq 0, \quad \forall j \in J,$$

*admet au moins une solution. Alors, étant donné  $s \in \mathcal{X}$ , les deux énoncés suivants sont équivalents :*

$$\begin{aligned} \forall x \in \mathcal{X}, \quad [\langle r_j, x \rangle \leq 0, \forall j \in J] &\Rightarrow [\langle s, x \rangle \leq 0], \\ s \in \overline{\mathcal{Z}(\{r_j\}_{j \in J})} &\subset \mathcal{X} \times \mathbb{R}. \end{aligned}$$

*Démonstration.* Il suffit de constater que le  $\gamma_0$  de la démonstration de Lemme de Farkas peut toujours être pris nul (donc le vecteur  $(0, 1) \in \mathcal{X} \times \mathbb{R}$  introduit dans ce lemme est inutile).  $\square$

Pour terminer, revenons sur une question laissée en suspens à la Remarque 3.28 et qui va se résoudre ici grâce au Lemme de Farkas.

**Lemme 3.61.** *Si  $C$  est un cône convexe fermé, alors  $C^{**} = C$ .*

*Démonstration.* D'après (3.9),

$$C^{**} = \{y \mid \langle y, \hat{x} \rangle \geq 0, \forall \hat{x} : [\forall x \in C, \langle \hat{x}, x \rangle \geq 0]\}.$$

Autrement dit,

$$y \in C^{**} \iff [\forall x \in C : \langle x, \hat{x} \rangle \geq 0] \Rightarrow [\langle y, \hat{x} \rangle \geq 0],$$

ce qui est équivalent, d'après le lemme de Farkas à ce que  $y \in \overline{\mathcal{Z}(\{x\}_{x \in C})} = \overline{C} = C$  puisque  $C$  est fermé. Donc  $C^{**} = C$ .  $\square$

### 3.9 Résumons nous

Un convexe est d'abord défini comme un "espace" dans lequel deux observateurs évoluant à l'intérieur de cet espace ne se perdent jamais de vue : un espace sans "recoin" en quelque sorte ; on ne peut pas y jouer à "cache-cache" :-). C'est la définition "interne" des convexes. Mais on aboutit ensuite à une autre caractérisation des convexes, au moins dans le cas de convexes *fermés* : c'est aussi l'intersection — une opération qui préserve la convexité — de convexes fermés en quelque sorte "élémentaires", les demi-espaces fermés (ceux qui sont délimités par un "simple mur rectiligne"). C'est la définition "externe" des convexes fermés, et dans bien des circonstances, la plus opérationnelle.

Associées à ces caractérisations interne et externe des convexes, on peut définir de deux façons l'*enveloppe convexe* d'un sous-ensemble quelconque (par combinaisons convexes des points du sous-ensemble ou par intersection des convexes qui contiennent le sous-ensemble). En dimension finie, disons  $n$ , et pour la construction interne, le théorème de Caratheodory indique qu'il n'est pas utile d'utiliser plus de  $n + 1$  points à la fois.

La caractérisation externe, nous l'avons déduite d'un théorème fondamental de "séparation" des objets convexes lorsqu'ils ne se chevauchent pas. Ce résultat de séparation, nous l'avons lui-même déduit de la notion de projection d'un point sur un convexe fermé, projection dont on parvient à démontrer qu'elle est définie de façon unique pour les structures euclidiennes ou hilbertiennes. C'est une démarche possible, celle qui nous a paru la plus "self contained" : dans d'autres ouvrages, on démontre (ou on admet) d'abord le théorème de Hahn-Banach (dont la forme dite "géométrique" constitue le résultat sur

la séparation), puis on en déduit l'existence de la projection, qui est par ailleurs, notons le, le premier problème d'optimisation formulé et résolu dans ce cours.

Ce théorème de séparation (dont nous avons donné seulement les formes les plus simples — séparation au sens “strict”) doit être considéré comme le fait majeur de ce cours : tout le reste (notamment le calcul sous-différentiel pour les fonctions convexes, la théorie de la dualité — dans le cas convexe, etc.) n'est qu'une suite de conséquences directes de ce fait fondamental.

En dehors des demi-espaces, qui en sont d'ailleurs a posteriori un cas particulier, un autre type d'objet convexe joue un rôle éminent : les cônes convexes. Ceux-ci sont non seulement stables par combinaisons convexes de leurs points mais aussi par simple somme de leurs points. On a vu qu'un cône convexe *saillant* (et fermé) permet de définir une relation d'ordre compatible avec la structure d'espace vectoriel (topologique). Les notions de *cône normal* et de *cône tangent* sont d'un usage permanent en optimisation. On a aussi parlé de la notion d'enveloppe cônica (convexe) d'un sous-ensemble quelconque.

Un fait remarquable est que les convexes, définis initialement à partir de considérations purement géométriques, ont finalement aussi des propriétés topologiques singulières : tout convexe *fortement* fermé est aussi *faiblement* fermé. On a vu que ce résultat est lui aussi une conséquence de la définition externe des convexes (fortement) fermés (et du fait que, par définition même de la topologie faible, les demi-espaces fortement fermés sont faiblement fermés).

### 3.10 Corrigé des exercices

**Corrigé de l'Exercice 3.9** On démontre le résultat par récurrence sur le nombre  $n$ . Le résultat est évidemment vrai pour  $n = 2$ . Supposons qu'il soit vrai jusqu'à  $n - 1$ . Posons

$$y_1 = \left( \sum_{i=1}^{n-1} \alpha_i x_i \right) / \left( \sum_{i=1}^{n-1} \alpha_i \right).$$

Ce  $y_1$  est évidemment une combinaison convexe à  $n - 1$  points, et donc, d'après l'hypothèse de récurrence, il est calculable récursivement par des combinaisons convexes à 2 éléments. Il suffit alors de remarquer que  $y$ , défini par (3.4), est aussi donné par

$$y = (1 - \alpha_n)y_1 + \alpha_n x_n,$$

puisque  $\sum_{i=1}^{n-1} \alpha_i = 1 - \alpha_n$ .

**Corrigé de l'Exercice 3.12** L'inclusion  $(\alpha + \beta)C \subset \alpha C + \beta C$  est toujours vraie. Dans le cas où  $\alpha$  et  $\beta$  sont positifs, l'inclusion inverse provient du fait que

$$\forall x \in C, y \in C, \quad \alpha C + \beta C \ni \alpha x + \beta y = (\alpha + \beta) \underbrace{\left( \frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \right)}_z \in (\alpha + \beta)C,$$

du fait que  $z \in C$  puisque  $C$  est convexe.

**Corrigé de l'Exercice 3.17** Si  $C$  est un cône, étant donnés  $x$  et  $y$  dans  $\mathcal{X}$  et  $\alpha \in [0, 1]$ , alors  $x$  et  $y$  sont dans  $C$  si et seulement si  $x/\alpha$  et  $y/(1 - \alpha)$  sont dans  $C$ , et  $C$  est convexe si et seulement si

$$x + y = \alpha(x/\alpha) + (1 - \alpha)(y/(1 - \alpha)) \in C.$$

**Corrigé de l'Exercice 3.18** La relation (3.8) est réflexive parce que  $x - x = 0$  et  $0 \in C$ . La relation (3.8) est transitive parce que, si  $x - y \in C$  et  $y - z \in C$ , alors  $(x - y) + (y - z) = x - z \in C$  d'après l'Exercice 3.17. Enfin, supposons que  $x - y \in C$  et  $y - x \in C$ , donc  $x - y \in (-C)$ . Si  $C$  est saillant (voir (3.6)), le seul élément qui appartient à la fois à  $C$  et  $-C$  est  $0$ , donc  $x = y$ .

**Corrigé de l'Exercice 3.22** Si  $a \in \overset{\circ}{A}$ , il existe une boule de rayon assez petit, centrée sur  $a$ , et contenue dans  $A$ . Donc, il existe  $\alpha \in \mathbb{R}_+$  tel que, pour tout  $x \in \mathcal{X}$ ,  $a + \alpha x \in A$ . D'après la Définition 3.20,  $z \in A_a^\perp$  si  $\langle z, \alpha x \rangle \leq 0$ . Comme  $x$  parcourt tout  $\mathcal{X}$ , ceci n'est possible que si  $z = 0$ .

**Corrigé de l'Exercice 3.25** Si  $z \in C_a^\perp$ , alors  $\langle z, x - a \rangle \leq 0$ , pour tout  $x \in C$ . Si  $C$  est un cône, on peut prendre  $x = 0$  puis  $x = 2a$  dans cette inégalité pour déduire que  $\langle z, a \rangle = 0$  et alors  $\langle z, x \rangle \leq 0$  pour  $x \in C$ . Cette dernière inégalité définit les  $z \in C^\perp$ , et par ailleurs  $z \in \{a\}^\perp$ . Enfin, si  $a \in \overset{\circ}{C}$ , d'après l'Exercice 3.22,  $C_a^\perp = \{0\} = \{a\}^\perp \cap C^\perp$ .

**Corrigé de l'Exercice 3.27** On a

$$(C \cup D)^* = \{y \mid \langle y, x \rangle \geq 0, \forall x \in C \text{ et } \langle y, z \rangle \geq 0, \forall z \in D\} \iff \\ \left( \{y \mid \langle y, x \rangle \geq 0, \forall x \in C\} \text{ et } \{y \mid \langle y, z \rangle \geq 0, \forall z \in D\} \right) = C^* \cap D^* .$$

De plus, si  $\langle y, x \rangle \geq 0$  pour  $x \in C$  et  $\langle y, z \rangle \geq 0$  pour  $z \in D$ , alors  $\langle y, x + z \rangle \geq 0$  pour  $(x + z) \in (C + D)$ , ce qui montre que  $C^* \cap D^* \subset (C + D)^*$ . Réciproquement, si  $\langle y, x + z \rangle \geq 0$  pour tout  $(x + z) \in (C + D)$ , il suffit de considérer  $x = 0 \in C$  puis  $z = 0 \in D$  pour déduire l'inclusion inverse.

Si  $y \in C^* \cup D^*$ , alors soit pour tout  $x \in C$ ,  $\langle y, x \rangle \geq 0$ , soit pour tout  $z \in D$ ,  $\langle y, z \rangle \geq 0$ , et donc en particulier, pour tout  $x \in C \cap D$ , s'il en existe, on a nécessairement que  $\langle y, x \rangle \geq 0$ , ce qui prouve l'inclusion  $C^* \cup D^* \subset (C \cap D)^*$ . Mais il est concevable que par exemple  $C \cap D = \{0\}$ , auquel cas  $(C \cap D)^* = \mathcal{X}$  sans que  $C^* \cup D^*$  soit lui-même égal à tout  $\mathcal{X}$  (prendre par exemple, dans  $\mathcal{X} = \mathbb{R}^2$ , pour  $C$  et  $D$ , deux demi-droites formant un angle très faible, auquel cas  $C^* \cup D^*$  est à peine plus grand qu'un demi-espace).

# Chapitre 4

## Fonctions convexes

Dans ce chapitre, on aborde le sujet des fonctions convexes sur un espace de Hilbert  $\mathcal{X}$  à valeurs dans  $\mathbb{R}$  ou “ $\mathbb{R}$  étendu” (auquel on a adjoint la valeur  $+\infty$  ; on notera  $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ ). Dans une certaine mesure, ce sujet est étroitement lié à celui du chapitre précédent : on peut en effet associer à une fonction un sous-ensemble de  $\mathcal{X} \times \mathbb{R}$ , appelé “épigraphe”, qui est convexe si et seulement si la fonction est convexe ; inversement d’ailleurs, on peut associer à un sous-ensemble de  $\mathcal{X}$  des fonctions à valeurs réelles étendues qui sont convexes si (ou si et seulement si) le sous-ensemble est convexe. Il existe cependant des questions nouvelles (ou plus faciles à appréhender) lorsqu’on manipule les fonctions, par exemple les questions de continuité et de différentiabilité en divers sens.

### 4.1 Épigraphe et convexité

**Définition 4.1.** Soit  $\mathcal{X}$  un espace de Hilbert et  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ . On appelle *épigraphe* de  $f$ , noté  $\Psi(f)$ , le sous-ensemble

$$\Psi(f) \stackrel{\text{def}}{=} \{(x, \xi) \mid x \in \mathcal{X}, \xi \in \mathbb{R} : \xi \geq f(x)\} .$$

On aura parfois besoin de l’*épigraphe strict*, noté  $\Psi_s(f)$  et défini comme

$$\Psi_s(f) \stackrel{\text{def}}{=} \{(x, \xi) \mid x \in \mathcal{X}, \xi \in \mathbb{R} : \xi > f(x)\} .$$

La Figure 4.1 illustre cette notion. Il est clair qu’un sous-ensemble  $A$  de  $\mathcal{X} \times \mathbb{R}$  est un épigraphe si et

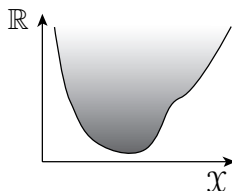


Figure 4.1: Épigraphe

seulement si il est “infini vers le haut”, c’est-à-dire que si  $(x, \xi) \in A$ , alors  $(x, \xi') \in A$  dès que  $\xi' \geq \xi$ . Il est non moins clair qu’alors

$$\Psi(f) = A \iff \forall x \in \mathcal{X}, f(x) = \inf_{(x, \xi) \in A} \xi , \tag{4.1}$$

ce qui donne un moyen constructif de retrouver  $f$  à partir de son épigraphe.

**Remarque 4.2.** Bien que nous nous autorisions à considérer des fonctions à valeurs dans  $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ , on définit l'épigraphe comme un sous-ensemble de  $\mathcal{X} \times \overline{\mathbb{R}}$  et non  $\mathcal{X} \times \mathbb{R}$ .

S'agissant de convexité, on peut d'abord énoncer le lemme suivant.

**Lemme 4.3.** Pour une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ ,  $\mathfrak{U}(f)$  est convexe si et seulement si  $\mathfrak{U}_s(f)$  est convexe.

La démonstration est laissée en exercice.

**Définition 4.4.** Soit  $\mathcal{X}$  un espace de Hilbert et  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ . La fonction  $f$  est *convexe* si son épigraphe est convexe.

D'après le Lemme 4.3, une définition équivalente peut être donnée en utilisant l'épigraphe strict. Si  $(x, f(x))$  et  $(y, f(y))$  appartiennent à  $\mathfrak{U}(f)$ , et si  $f$  est convexe, alors, pour tout  $\alpha \in [0, 1]$ ,

$$(\alpha x + (1 - \alpha)y, \alpha f(x) + (1 - \alpha)f(y))$$

appartient aussi à  $\mathfrak{U}(f)$ , donc

$$\forall x \in \mathcal{X}, y \in \mathcal{Y}, \forall \alpha \in [0, 1], f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y). \quad (4.2)$$

Réciproquement, si (4.2) est vrai, on en déduit facilement que  $\mathfrak{U}(f)$  est convexe, donc (4.2) peut aussi être prise (et c'est le plus souvent le cas dans la littérature) comme définition de la convexité d'une fonction  $f$ .

L'inéquation (4.2) a l'interprétation géométrique suivante : la valeur de la fonction  $f$  en tout point  $z$  du segment  $[x, y]$  est située sous la valeur de la "corde" (interpolation linéaire de  $f$  entre  $x$  et  $y$ ) au même point  $z$ . Ceci est illustré par la Figure 4.2. Pour une fonction affine, on aurait précisément l'égalité : une

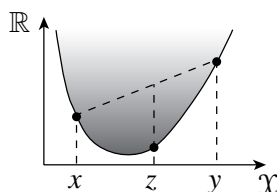


Figure 4.2: Interprétation de la convexité

fonction convexe est donc "sous-affine".

Une fonction  $f$  est *concave* si  $-f$  est convexe (si  $f$  prend des valeurs infinies, on suppose donc que ces valeurs sont à  $-\infty$ ). Il est clair que cela correspond à l'inégalité inverse dans (4.2). Si l'on veut relier cette notion à celle d'un sous-ensemble *convexe* de  $\mathcal{X} \times \overline{\mathbb{R}}$ , il faut considérer l'"hypographe", c'est-à-dire la partie de l'espace située *sous* le graphe. Une fonction est *affine* si et seulement si elle est à la fois convexe et concave.

**Définition 4.5.** On appelle *domaine* de  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ , noté  $\mathfrak{D}(f)$ , le sous-ensemble de  $\mathcal{X}$  où  $f$  prend des valeurs finies :

$$\mathfrak{D}(f) \stackrel{\text{def}}{=} \{x \in \mathcal{X} \mid f(x) < +\infty\} .$$

**Remarque 4.6.** En écho à la Remarque 4.2, on observe que  $\mathfrak{D}(f)$  est précisément la projection de  $\mathfrak{U}(f)$  sur  $\mathcal{X}$ .



## 4.2 Autres notions de convexité

On introduit maintenant les raffinements suivants.

**Définition 4.7.** Une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est *strictement convexe* si elle est convexe et si l'inégalité stricte a lieu dans (4.2) dès que  $x \neq y$  et  $\alpha \in ]0, 1[$ .

Autrement dit, le graphe de la fonction n'a pas de partie affine.

**Définition 4.8.** Une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est *fortement convexe* (de module  $b$ ) si

$$\exists b > 0 : \forall x \in \mathcal{X}, y \in \mathcal{Y}, \forall \alpha \in [0, 1], \\ f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - b\alpha(1 - \alpha) \|x - y\|^2 / 2. \quad (4.3)$$

Une fonction fortement convexe est strictement convexe mais la réciproque n'est pas vraie.

**Exercice 4.9.** Montrer que la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^4$  est strictement mais pas fortement convexe.

## 4.3 Autres associations de sous-ensembles et de fonctions convexes

### 4.3.1 Ensembles de niveau

Nous avons vu que l'épigraphe d'une fonction convexe est convexe. C'était même notre point de départ sur les fonctions convexes et une caractérisation possible de ces fonctions. On introduit maintenant les "ensembles de niveau" associés à une fonction à valeurs réelles.

**Définition 4.10.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  et  $\beta \in \overline{\mathbb{R}}$ . On appelle *ensemble de niveau de  $f$  au niveau  $\beta$* , noté  $\nabla_{\beta}(f)$ , le sous-ensemble (éventuellement vide)

$$\nabla_{\beta}(f) \stackrel{\text{def}}{=} \{x \in \mathcal{X} \mid f(x) \leq \beta\}.$$

La Figure 4.3 illustre cette notion. Il est immédiat de montrer (exercice) que les ensembles de niveau

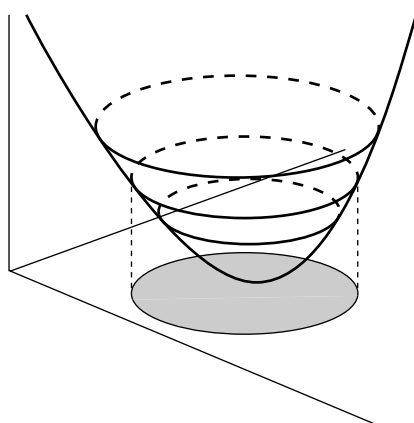


Figure 4.3: Ensembles de niveau

d'une fonction convexe sont convexes (en particulier, son domaine est convexe). Mais, contrairement à ce qui se passe pour l'épigraphe, cette propriété n'est pas caractéristique des fonctions convexes. Pour le voir, il suffit par exemple de considérer n'importe quelle fonction monotone, mais à part cela quelque, de  $\mathbb{R}$  dans  $\mathbb{R}$ , et de constater que les ensembles de niveau de cette fonction sont des demi-droites

(éventuellement vides), donc des convexes. Les fonctions dont les ensembles de niveau sont convexes constituent une classe intéressante et sont appelées “fonctions *quasi-convexes*”. Nous n’aurons malheureusement pas le loisir de les étudier dans ce cours. La Figure 4.4 représente une fonction quasi-convexe de  $\mathbb{R}^2$  dans  $\mathbb{R}$ .

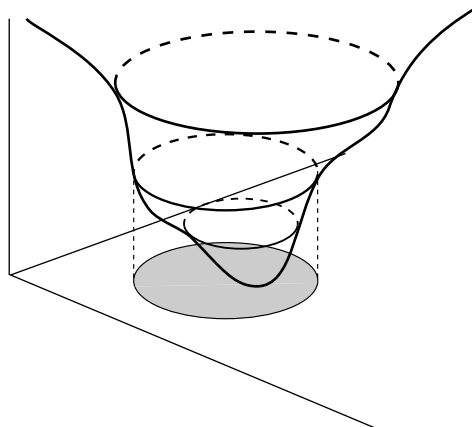


Figure 4.4: Une fonction quasi-convexe

### 4.3.2 Fonction indicatrice

Nous avons associé à des fonctions des sous-ensembles (épigraphe, ensembles de niveau). Nous allons inversement associer à des sous-ensembles des fonctions, et en premier lieu la “fonction indicatrice”.

**Définition 4.11.** Soit  $A$  un sous-ensemble d’un espace de Hilbert  $\mathcal{X}$ . On appelle *fonction indicatrice* de  $A$ , noté  $I_A$ , la fonction définie sur  $\mathcal{X}$  à valeurs dans  $\overline{\mathbb{R}}$  suivante :

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

On peut se demander pourquoi aller chercher la valeur infinie (et ne pas se contenter de 1 par exemple). Une raison apparaîtra lorsque nous aborderons le domaine de l’optimisation. Mais une autre raison est qu’une telle fonction ne serait pas convexe même lorsque  $A$  est convexe.

Il est immédiat de montrer que  $I_A$ , telle que définie, est convexe si et seulement si  $A$  est convexe. De plus,  $A = \overline{\bigcap_{\beta} (I_A \leq \beta)}$  pour tout  $\beta$  non négatif et fini, et  $A$  est aussi le domaine de  $I_A$ . La fonction  $I_A$  caractérise sans ambiguïté le sous-ensemble  $A$ .

### 4.3.3 Fonction support

Une autre fonction importante associée à  $A$  est la “fonction support”.

**Définition 4.12.** Soit  $A$  un sous-ensemble *non vide* d’un espace de Hilbert  $\mathcal{X}$ . On appelle *fonction support* de  $A$ , noté  $\sigma_A$ , la fonction définie sur  $\mathcal{X}$  à valeurs dans  $\overline{\mathbb{R}}$  suivante :

$$\sigma_A(p) = \sup_{x \in A} \langle p, x \rangle . \quad (4.4)$$

**Remarque 4.13.** La convention de prendre la valeur  $-\infty$  (respectivement,  $+\infty$ ) lorsque le sup (respectivement, l’inf) est pris sur un ensemble vide est classique en optimisation. On peut donc considérer que  $\sigma_{\emptyset} \equiv -\infty$  mais nous avons évité depuis le début de manipuler des fonctions convexes prenant la valeur  $-\infty$ .

On observe que la fonction support est convexe même lorsque  $A$  ne l'est pas : cela proviendra de résultats généraux énoncés plus loin et du fait que la fonction support est une enveloppe supérieure de fonctions linéaires. La fonction support est aussi positivement homogène de degré 1, c'est-à-dire que

$$\forall \alpha > 0, \forall p \in \mathcal{X}, \sigma_A(\alpha p) = \alpha \sigma_A(p) . \tag{4.5}$$

La Figure 4.5 illustre le calcul de la fonction support. Il n'est pas difficile de se rendre compte sur cette figure que  $A$  et  $\odot(A)$  (et même  $\overline{\odot}(A)$ ) fournissent les mêmes valeurs de la fonction support. Celle-

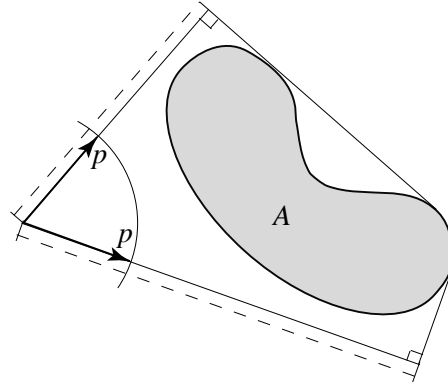


Figure 4.5: Valeurs de  $\sigma_A(p)$  — longueurs en pointillés — pour deux valeurs particulières de  $p$  avec  $\|p\| = 1$

ci ne caractérise donc pas  $A$  en général. En effet, comment retrouver  $A$  à partir de  $\sigma_A$  ? Il est clair que, par définition de la fonction support,

$$x \in A \Rightarrow \forall p \in \mathcal{X}, \langle p, x \rangle \leq \sigma_A(p) .$$

Le second membre de cette implication définit un sous-ensemble  $B$  de points  $x$  qui est égal à l'intersection d'une famille infinie (indexée par  $p$ ) de demi-espaces (fermés). C'est donc un sous-ensemble convexe fermé. Donc  $A$  est inclus dans ce convexe fermé qui contient donc aussi  $\overline{\odot}(A)$ . Le théorème suivant affirme qu'en fait  $B = \overline{\odot}(A)$  et que  $A$  et  $\overline{\odot}(A)$  sont indistingables à travers leur fonction support.

**Théorème 4.14.** *Soit  $A \subset \mathcal{X}$  et soit*

$$B = \{x \in \mathcal{X} \mid \forall p \in \mathcal{X}, \langle p, x \rangle \leq \sigma_A(p)\} .$$

*Alors,  $B = \overline{\odot}(A)$  et  $\sigma_{\overline{\odot}(A)} = \sigma_A$ .*

*Démonstration.* On vient de montrer l'inclusion  $B \supset \overline{\odot}(A)$ . Il faut montrer en fait la nécessité de l'égalité de ces deux sous-ensembles convexes fermés. Supposons qu'il existe  $a \in B$  et  $a \notin \overline{\odot}(A)$ . Alors, par le théorème de séparation, on pourrait trouver  $r \in \mathcal{X}$  tel que

$$\sup_{x \in \overline{\odot}(A)} \langle r, x \rangle < \langle r, a \rangle .$$

Mais, par ailleurs,

$$\sigma_A(r) \stackrel{\text{def}}{=} \sup_{x \in A} \langle r, x \rangle \leq \sup_{x \in \overline{\odot}(A)} \langle r, x \rangle .$$

En rapprochant ces deux inégalités, on déduit qu'il existe  $r$  tel que  $\sigma_A(r) < \langle r, a \rangle$ , ce qui contredit l'hypothèse que  $a \in B$ . Donc, finalement,  $B = \overline{\text{co}}(A)$ .

La définition même de  $B$  implique que  $\sigma_B \leq \sigma_A$  mais évidemment  $\sigma_A \leq \sigma_{\overline{\text{co}}(A)} = \sigma_B$ , et on obtient donc l'égalité des fonctions support.  $\square$

Finalement, ce que nous dit ce théorème, c'est que la fonction support de  $A$  ne caractérise que  $\overline{\text{co}}(A)$  et pas  $A$  lui-même.

Nous établirons ultérieurement un lien entre fonction support et fonction indicatrice d'un même sous-ensemble (voir Exercice 4.21).

**Exercice 4.15.** Pour deux sous-ensembles  $A_1$  et  $A_2$ , montrer les formules :

$$\sigma_{A_1 \cup A_2}(p) = \sup(\sigma_{A_1}(p), \sigma_{A_2}(p)), \quad \sigma_{A_1 + A_2}(p) = \sigma_{A_1}(p) + \sigma_{A_2}(p).$$

**Exercice 4.16.** Soit  $A$  un sous-ensemble d'un espace produit  $\mathcal{X} \times \mathcal{Y}$  et  $A_{\mathcal{X}}$  sa projection sur  $\mathcal{X}$ . On rappelle que l'on peut définir  $A_{\mathcal{X}}$  de la façon suivante :

$$A_{\mathcal{X}} \stackrel{\text{def}}{=} \{x \in \mathcal{X} \mid \exists y \in \mathcal{Y} : (x, y) \in A\}.$$

Montrer que  $\sigma_{A_{\mathcal{X}}}$  est donné par la formule :

$$\forall p \in \mathcal{X}, \quad \sigma_{A_{\mathcal{X}}}(p) = \sigma_A(p, 0).$$

La Figure 4.6 illustre ce dernier exercice.

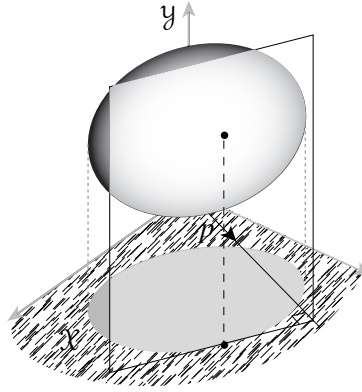


Figure 4.6: Calcul de la fonction support de la projection d'un sous-ensemble

## 4.4 Opérations préservant la convexité des fonctions

Les faits suivants sont énoncés sans démonstration : ils peuvent être considérés comme des exercices assez simples que le lecteur est cependant invité à faire.

- On a vu que l'intersection de sous-ensembles convexes est convexe. Que l'on se place du point de vue des épigraphes ou des ensembles de niveau associés à des fonctions convexes, l'intersection correspond à prendre l'*enveloppe supérieure* des fonctions correspondantes. En fonction de cette observation sur les épigraphes, on a donc que :

$$f : x \mapsto \sup_{j \in J} f_j(x) \text{ est convexe si } f_j \text{ est convexe, } \forall j \in J.$$

La Figure 4.7 illustre ce fait. C'est pour cette raison que la fonction support est convexe dès que  $A \neq \emptyset$  : en effet, d'après (4.4), c'est l'enveloppe supérieure d'une famille de fonctions affines.

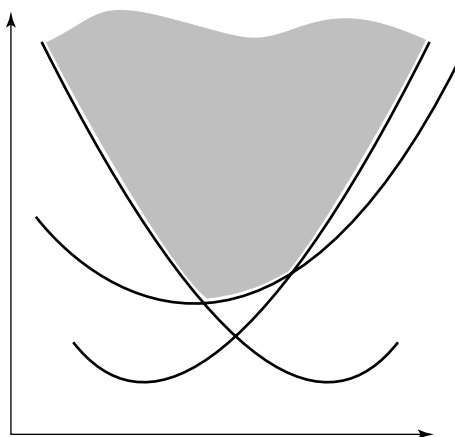


Figure 4.7: Enveloppe supérieure de fonctions convexes

- La multiplication par un scalaire *positif* d’une fonction convexe, la *somme* de fonctions convexes, et donc aussi la somme pondérée à coefficients positifs de fonctions convexes redonnent des fonctions convexes.
- Le changement d’échelle en  $x$  (homothétie dans  $\mathcal{X}$ ), la translation en  $x$ , et plus généralement une transformation affine sur  $\mathcal{X}$  préservent la convexité. Plus explicitement,  $f : x \mapsto f(Ax + b)$  est convexe si  $f$  est convexe, et  $A \in \mathcal{L}(\mathcal{X}, \mathcal{X}), b \in \mathcal{X}$ .
- L’homothétie d’un sous-ensemble convexe reste convexe. S’agissant d’épigraphe, il faut évidemment se limiter à des rapports d’homothétie positifs pour que l’épigraphe homothétique reste un épigraphe. On montre que

$$\forall \alpha > 0, \quad \mathfrak{U}(g) = \alpha \mathfrak{U}(f) \iff g(x) = \alpha f(x/\alpha) .$$

La fonction  $g$  est convexe lorsque  $f$  est convexe.

- La *composition* d’une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$  avec une fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$  est convexe si  $f$  est convexe et si  $g$  est convexe *et monotone non décroissante*.
- La *restriction* d’une application  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$  à l’un des espaces, c’est-à-dire par exemple l’application

$$g_y : \mathcal{X} \rightarrow \overline{\mathbb{R}}, x \mapsto f(x, y) \text{ (} y \text{ fixé)}$$

est convexe si  $f$  est convexe sur l’espace  $\mathcal{X} \times \mathcal{Y}$  (on dira qu’elle est *conjointement convexe* en  $(x, y)$ ). Par contre, il ne suffit pas que toutes les restrictions  $g_y$  pour tout  $y$  et les restrictions analogues à  $x$  fixé soient convexes pour pouvoir affirmer que  $f$  est (conjointement) convexe. La Figure 4.8 suggère le graphe d’une fonction “convexe-convexe” mais pas “conjointement convexe”.

On étudie sous forme d’exercices deux autres opérations importantes sur les fonctions convexes qui correspondent respectivement aux opérations de somme vectorielle et de projection sur les épigraphes ou les épigraphes stricts, opérations dont on a vu qu’elles préservent la convexité.

**Exercice 4.17.** ☺ On considère deux fonctions  $f$  et  $g$  sur un espace de Hilbert  $\mathcal{X}$  et on définit une nouvelle fonction  $h$ , appelée *inf-convolution* de  $f$  et  $g$ , opération notée  $f \square g$ , de la façon suivante :

$$h(x) = (f \square g)(x) \stackrel{\text{def}}{=} \inf_{y \in \mathcal{X}} (f(y) + g(y - x)) . \tag{4.6}$$

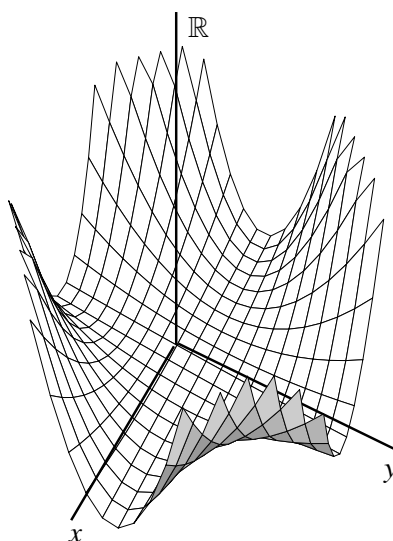



Figure 4.8: Graphe d'une fonction non convexe dont les sections parallèles aux axes sont des courbes convexes

Cette définition nécessite une hypothèse technique supplémentaire pour garantir que la valeur  $-\infty$  ne sera pas atteinte. On peut par exemple supposer que les deux fonctions  $f$  et  $g$  admettent une minoration par la même fonction affine. Montrer, sous cette condition, que  $h$  ne prend pas la valeur  $-\infty$  et que

$$\mathbb{U}_s(h) = \mathbb{U}_s(f) + \mathbb{U}_s(g) .$$

En déduire que si  $f$  et  $g$  sont convexes, alors leur inf-convolution est convexe.

**Exercice 4.18.**  On considère une fonction  $f$  sur le produit  $\mathcal{X} \times \mathcal{Y}$  de deux espaces de Hilbert. On suppose qu'elle est *conjointement* convexe en  $(x, y)$  et bornée inférieurement sur tout  $\mathcal{X} \times \mathcal{Y}$ . On lui associe la fonction “marginale”  $g$  définie sur  $\mathcal{X}$  de la façon suivante :

$$g(x) \stackrel{\text{def}}{=} \inf_{y \in \mathcal{Y}} f(x, y) . \quad (4.7)$$

Évidemment,  $g$  est également bornée inférieurement sur  $\mathcal{X}$ . Montrer que  $\mathbb{U}_s(g)$  est la projection de  $\mathbb{U}_s(f)$  sur la composante  $\mathcal{X}$  (voir Figure 3.4). En déduire que la fonction marginale  $g$  est convexe.

**Remarque 4.19.** On observe d'abord que (4.6) peut être considérée comme un cas particulier de (4.7) pour la fonction  $(x, y) \mapsto f(y) + g(y - x)$  qui est bien conjointement convexe en  $(x, y)$  (exercice). La convexité d'une inf-convolution de deux fonctions convexes aurait donc pu se déduire du résultat de l'Exercice 4.18. On note ensuite que la convexité de la fonction marginale n'est pas triviale si l'on réalise que cette fonction peut être vue comme l'enveloppe *inférieure* d'une famille de fonctions convexes (les fonctions  $x \mapsto f(x, y)$  indexées par  $y$ ). Or l'enveloppe inférieure, contrairement à l'enveloppe supérieure, ne préserve pas la convexité en général (elle fait l'union des épigraphes, au lieu de l'intersection pour l'enveloppe supérieure). Le résultat présent est dû à l'hypothèse de convexité *conjointe* de  $f$  en  $(x, y)$ .

## 4.5 Transformée de Fenchel

Ayant introduit la fonction support d'un ensemble, on peut chercher à calculer celle-ci pour l'épigraphe d'une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$ . Il faut alors travailler dans l'espace produit  $\mathcal{X} \times \mathbb{R}$  qui sera bien sûr muni du produit scalaire  $\langle \cdot, \cdot \rangle_{\mathcal{X}} + \cdot \times \cdot$ .

- Il est facile de réaliser que dès que l'on prend comme argument de la fonction support une valeur  $(p, \varpi) \in \mathcal{X} \times \mathbb{R}$  avec  $\varpi > 0$ , en raison de la forme même d'un épigraphe "infini vers le haut", on va trouver la valeur  $+\infty$ .
- Si l'on prend la valeur  $\varpi = 0$ , c'est-à-dire si l'on calcule  $\sigma_{\mathcal{U}(f)}(p, 0)$ , en fonction du résultat énoncé à l'Exercice 4.16, on est en train de calculer en fait, au point  $p$ , la fonction support de la projection de  $\mathcal{U}(f)$  sur  $\mathcal{X}$  qui n'est autre que  $\mathcal{T}(f)$  (voir Remarque 4.6).
- Finalement, considérant les valeurs négatives de  $\varpi$ , on peut se contenter de la valeur  $\varpi = -1$  en raison du caractère positivement homogène de la fonction support (voir (4.5)). On calcule alors

$$\begin{aligned} \sigma_{\mathcal{U}(f)}(p, -1) &= \sup_{(x, \xi) \in \mathcal{U}(f)} (\langle p, x \rangle - \xi) \\ &= \sup_{\xi \geq f(x)} (\langle p, x \rangle - \xi) \\ &= \sup_{x \in \mathcal{T}(f)} (\langle p, x \rangle - f(x)). \end{aligned}$$

Cette dernière formule définit ce qui est connu sous le nom de "transformée de Fenchel"<sup>1</sup> de  $f$ .

**Définition 4.20.** Pour une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$ , on appelle *transformée de Fenchel*, notée  $f^F$ , la fonction de  $\mathcal{X} \rightarrow \overline{\mathbb{R}}$  suivante :

$$f^F : \mathcal{X} \rightarrow \overline{\mathbb{R}}, \quad f^F(p) = \sup_{x \in \mathcal{T}(f)} (\langle p, x \rangle - f(x)).$$

On observe immédiatement que cette fonction est toujours convexe car elle s'exprime comme l'enveloppe supérieure d'une famille de fonctions affines en  $p$  (on a déjà fait cette observation pour les fonctions support ; or, comme on vient de le voir, la transformée de Fenchel est pratiquement la fonction support d'un épigraphe).

Par ailleurs, on a immédiatement l'inégalité :

$$\forall p, \forall x, \quad f^F(p) + f(x) \geq \langle p, x \rangle. \quad (4.8)$$

On verra que lorsque cette inégalité devient une égalité,  $p$  (respectivement,  $x$ ) joue un rôle particulier vis-à-vis de  $f$  (respectivement,  $f^F$ ).

**Exercice 4.21.** Montrer que la fonction support d'un sous-ensemble  $A$  est la transformée de Fenchel de la fonction indicatrice de ce même sous-ensemble.

**Exercice 4.22.** L'épigraphe de l'enveloppe inférieure  $h$  de deux fonctions  $f$  et  $g$  :

$$h(x) \stackrel{\text{def}}{=} \min(f(x), g(x))$$

est l'union des épigraphes de  $f$  et  $g$ . Grâce à l'Exercice 4.15, on en déduit que  $h^F = \sup(f^F, g^F)$ . Redémontrer ce résultat directement en utilisant l'expression de ces fonctions.

<sup>1</sup>On attribue aussi parfois les deux noms "Legendre-Fenchel" à cette transformation, tant il est vrai que la transformée de Legendre réfère à une transformation assez proche de celle-ci mais qui nécessite la différentiabilité de  $f$ .

De même, on a vu à l'Exercice 4.17 que l'épigraphe (strict) de  $h = f \square g$  est égal à la somme des épigraphes (stricts) de  $f$  et  $g$ . En recoupant cette observation avec celle de l'Exercice 4.15, on en déduit de la transformée de Fenchel de  $h$  est la somme des transformées de Fenchel de  $f$  et  $g$ . Redémontrer ce résultat directement en utilisant les expressions de ces fonctions.

**Exercice 4.23.** On a vu à l'Exercice 4.18 que l'épigraphe de la fonction marginale  $g$  (voir (4.7)) est la projection de l'épigraphe de la fonction initiale  $f$ . Par ailleurs, on a vu à l'Exercice 4.16 comment la fonction support d'une projection est reliée à la fonction support du sous-ensemble à projeter. On en déduit que  $g^F(p) = f^F(p, 0)$ . Redémontrer ce résultat directement en utilisant les expressions de ces fonctions.

Pour terminer cette section, résumons la discussion du début. On a vu que

$$\sigma_{\Psi(f)}(p, \varpi) = \begin{cases} +\infty & \text{si } \varpi > 0, \\ \sigma_{\neg(f)}(p) & \text{si } \varpi = 0, \\ (-\varpi) f^F(-p/\varpi) & \text{si } \varpi < 0. \end{cases} \quad (4.9)$$

## 4.6 Enveloppe convexe

On a discuté au §3.4 de la question de l'enveloppe convexe d'un sous-ensemble, à savoir le plus petit sous-ensemble convexe le contenant. Si on applique cette notion à l'épigraphe d'une fonction quelconque, on est confronté à deux difficultés potentielles.

- En premier lieu, si la fonction tend vers  $-\infty$  dans deux directions opposées de l'espace (par exemple, la fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ ,  $x \mapsto -x^2$ ), on risque d'obtenir une enveloppe convexe de l'épigraphe qui "passe à  $-\infty$ " pour des valeurs finies de l'abscisse ; on se prémunira contre ces situations en supposant que la fonction considérée admet une minorante affine.
- En second lieu, l'enveloppe convexe d'un épigraphe n'est pas "tout-à-fait" un épigraphe : il peut lui manquer un "bord inférieur" comme pour la fonction (voir Figure 4.9) :

$$x \mapsto \begin{cases} +\infty & \text{si } x \notin [0, 2[ \\ \lfloor x \rfloor \text{ (partie entière de } x) & \text{sinon.} \end{cases}$$

C'est pourquoi il faudra récupérer ces bords inférieurs par une opération d'infimum.<sup>2</sup>

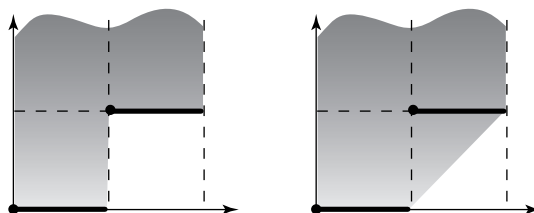


Figure 4.9: Une fonction, son épigraphe et l'enveloppe convexe de son épigraphe

<sup>2</sup>On pourrait sans doute se tirer de cette difficulté en s'exprimant en termes d'épigraphes *stricts* de la fonction et de son enveloppe convexe.



Du fait que l’inclusion des épigraphes correspond à l’ordre *inverse* des fonctions, c’est-à-dire que

$$\Psi(f) \subset \Psi(g) \iff f(x) \geq g(x), \forall x,$$

il est clair que pour une fonction  $f : \mathcal{X} \mapsto \mathbb{R}$ , l’enveloppe convexe de l’épigraphe de  $f$  — une fois “bordée” — correspondra à l’épigraphe de *la plus grande* fonction convexe qui est *plus petite* que  $f$  au sens de cet ordre. On notera cette enveloppe convexe  $\odot(f)$ , c’est-à-dire avec la même notation que pour les sous-ensembles, ce qui ne devrait pas causer de difficultés.

Par ailleurs, d’après le Théorème 3.14, l’enveloppe convexe de l’épigraphe est obtenue en effectuant toutes les combinaisons convexes des points de cet épigraphe. Comme on passe de l’épigraphe à la fonction en prenant le “bord inférieur” (voir (4.1)), on peut finalement énoncer le résultat/définition suivant(e) (voir [9, Proposition 2.5.1]).

**Théorème 4.24 (Enveloppe convexe d’une fonction).** *Pour une fonction  $f$  donnée, si elle est minorée par une fonction affine, il existe une plus grande fonction convexe inférieure ou égale à  $f$ . Elle est appelée enveloppe convexe de  $f$  notée  $\odot(f)$ . On a que*

$$\odot(f)(x) = \inf \{ \xi \mid (x, \xi) \in \odot(\Psi(f)) \} .$$

Par ailleurs,  $\odot(f)$  est aussi donnée par la formule :

$$\odot(f)(x) = \inf \left\{ \sum_{i=1}^n \alpha_i f(x_i) \mid n \in \mathbb{N}, \alpha \in \Sigma_n, \sum_{i=1}^n \alpha_i x_i = x \right\} . \tag{4.10}$$

Cette dernière construction de l’enveloppe convexe relève de la philosophie “construction interne” déjà appuie à la “construction externe” des convexes (fermés) par leurs hyperplans d’appui.

**Exercice 4.25.** On considère la fonction de  $\mathbb{R}$  dans  $\mathbb{R}$  suivante :

$$f(x) = \min (|x| + 1/2, |x - 1|) .$$

Dessiner le graphe de cette fonction et celui de son enveloppe convexe. Retrouver cette enveloppe convexe par le calcul en utilisant la formule (4.10).

## 4.7 Propriétés topologiques des fonctions convexes

### 4.7.1 Continuité

Il suffit de penser à la fonction indicatrice d’un sous-ensemble convexe qui passe brusquement de 0 à  $+\infty$  au bord du convexe pour se convaincre qu’une fonction convexe peut ne pas être continue. Cependant, les “ennuis” se concentrent en quelque sorte au bord du domaine de la fonction, c’est-à-dire là où elle n’est pas localement bornée.

**Lemme 4.26.** *Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe sur un espace de Hilbert  $\mathcal{X}$  et  $a \in \mathcal{X}$  un point au voisinage duquel la fonction est bornée. Alors,  $f$  est continue en  $a$ .*

*Démonstration.* Pour simplifier les écritures, on peut effectuer une translation dans  $\mathcal{X}$  qui amène  $a$  en 0 et une translation dans  $\overline{\mathbb{R}}$  qui amène  $f(a)$  en 0, ce qui ne modifie pas les hypothèses. On considère donc qu’il existe un voisinage  $V$  de l’origine où  $f$  est bornée par une constante  $M$ . On peut aussi supposer que ce voisinage est symétrique par rapport à l’origine en le restreignant éventuellement à  $V \cap (-V)$ .

Soit  $\alpha \in ]0, 1[$  et  $x$  quelconque dans  $\alpha V$ . Notons que  $(x/\alpha) \in V$  et  $(-x/\alpha) \in V$ , et donc  $f(x/\alpha) \leq M$  et  $f(-x/\alpha) \leq M$ . Alors,

$$\begin{aligned} x &= (1 - \alpha)0 + \alpha(x/\alpha) \Rightarrow f(x) \leq (1 - \alpha)f(0) + \alpha f(x/\alpha) \\ &\Rightarrow f(x) \leq \alpha M, \end{aligned}$$

et par ailleurs

$$\begin{aligned} 0 &= \frac{1}{1 + \alpha}x + \frac{\alpha}{1 + \alpha}(-x/\alpha) \Rightarrow f(0) \leq \frac{1}{1 + \alpha}f(x) + \frac{\alpha}{1 + \alpha}f(-x/\alpha) \\ &\Rightarrow f(x) \geq (1 + \alpha)f(0) - \alpha f(-x/\alpha) \\ &\Rightarrow f(x) \geq -\alpha M. \end{aligned}$$

Finalement,  $|f(x)| \leq \alpha M$  et lorsque  $\alpha \rightarrow 0$ ,  $x \in \alpha V$  tend aussi vers 0 et  $f(x) \rightarrow 0 = f(0)$  ce qui montre la continuité de  $f$  en 0.  $\square$

**Théorème 4.27.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe sur un espace de Hilbert  $\mathcal{X}$ . Alors,  $f$  est continue sur l'intérieur de son domaine qui est non vide si et seulement si il existe un ouvert sur lequel  $f$  est bornée.

*Démonstration.* La partie “seulement si” est immédiate. On se concentre sur le “si”. Supposons donc qu'il existe un ouvert  $V$  sur lequel  $f$  est bornée par une constante  $M$ . Par conséquent,  $V \subset \mathcal{I}(f)$ , par définition de  $\mathcal{I}(f)$  (voir Définition 4.5), et puisque  $\mathcal{I}(f)$  contient un ouvert, il est d'intérieur non vide. En fonction du Lemme 4.26, il suffit de montrer que pour tout point  $z$  de l'intérieur de  $\mathcal{I}(f)$ , on peut construire un voisinage où  $f$  est bornée. Pour construire un tel voisinage de  $z$ , on utilise le procédé déjà utilisé dans la démonstration du Théorème 3.35 et illustré par la Figure 3.12. Pour cela, on prend un point  $x \in V$  et on observe que pour  $\alpha$  assez petit (et en tout état de cause plus petit que 1),

$$y = (z - \alpha x)/(1 - \alpha)$$

appartient encore à l'intérieur de  $\mathcal{I}(f)$ . En fait,  $z$  est maintenant une combinaison convexe de  $x$  et de  $y$  (en effet,  $z = \alpha x + (1 - \alpha)y$ ). Il suffit de faire une homothétie de centre  $y$  et de rapport  $\alpha$  pour amener  $x$  sur  $z$  et, dans cette homothétie,  $V$  se transforme en un voisinage centré autour de  $z$ . Par l'inégalité de la convexité pour  $f$ , on montre alors facilement que pour tout point  $a$  de ce voisinage de  $z$ ,

$$f(a) \leq \alpha M + (1 - \alpha)f(y),$$

et donc  $f$  est bien bornée au voisinage de  $z$ .  $\square$

**Corollaire 4.28.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe sur un espace de Hilbert  $\mathcal{X}$  de dimension finie. Alors  $f$  restreinte à l'intérieur relatif de son domaine est continue.

*Démonstration.* On écarte bien sûr le cas peu intéressant où  $f$  vaut identiquement  $+\infty$ . Alors  $\mathcal{I}(f)$  est un convexe non vide, et par conséquent, par le Théorème 3.38, on sait que son intérieur relatif est non vide. On se ramène donc à raisonner sur l'enveloppe affine de cet intérieur relatif : désormais, lorsqu'on dira “ouvert”, ce sera relativement à la topologie induite sur cette variété affine. Soit  $n$  la dimension de cette variété affine. Il suffit maintenant d'appliquer le Théorème 4.27 à condition de vérifier que l'intérieur relatif de  $\mathcal{I}(f)$  contient un ouvert sur lequel  $f$  est bornée. Or l'intérieur relatif de  $\mathcal{I}(f)$  contient un polyèdre ouvert de  $(n + 1)$  sommets (la construction explicite de ce polyèdre suit le procédé utilisé dans la démonstration du Théorème 3.38). Par l'inégalité de la convexité de  $f$ , on borne facilement la valeur de  $f$  en tout point de ce polyèdre ouvert par le max des  $(n + 1)$  valeurs en ces sommets (qui sont évidemment finies puisque l'on est dans  $\mathcal{I}(f)$ ).  $\square$

Mentionnons pour terminer cette exploration des propriétés de continuité des fonctions convexes qu'on peut aller plus loin que la continuité sous les hypothèses précédentes, c'est-à-dire qu'on peut même montrer le caractère *localement Lipschitzien*<sup>3</sup> (à nouveau, en un point au voisinage duquel la fonction est bornée). Le caractère localement Lipschitzien exclut les "pentes verticales" comme pour la fonction  $\sqrt{|x|}$  de  $\mathbb{R}$  dans  $\mathbb{R}$  au point  $x = 0$ . Nous renvoyons pour ce type de résultat à [9, Tome I, p. 173–174] ou [8, Corollaire 2.4].

### 4.7.2 Fermeture de l'épigraphe et semi-continuité inférieure

On a vu au chapitre précédent que les convexes *fermés* peuvent être décrits de façon externe par l'intersection de demi-espaces fermés délimités par leurs hyperplans d'appui. On a dit que cette description externe est souvent plus importante que la description interne par des combinaisons convexes de points. Il n'est donc pas étonnant que la classe des fonctions convexes dont l'épigraphe est fermé soit d'un intérêt particulier. On va voir que l'épigraphe est fermé si et seulement si la fonction est "semi-continue inférieurement" (ce que nous abrègerons le plus souvent en s.c.i.). Rappelons cette définition.

**Définition 4.29.** Une fonction est *semi-continue inférieurement* en un point  $x$  si

$$f(x) \leq \liminf_{y \rightarrow x} f(y) .$$

La fonction peut donc être discontinue en  $x$  mais seules les discontinuités "vers le bas" sont admises. Sur la Figure 4.10, on voit (à gauche) une fonction s.c.i. (non convexe) et (à droite) une modification de cette fonction qui n'est plus s.c.i.. On observe que les ensembles de niveau (en gris sur l'espace de départ) sont fermés pour la version s.c.i. mais ne le sont plus pour la modification non s.c.i.. Pour cette dernière, on observe de plus que le minimum de la fonction n'est pas atteint : on soupçonne donc l'importance du caractère s.c.i. pour la théorie de l'optimisation. De même l'épigraphe est fermé pour la fonction s.c.i. et quelconque pour la version non s.c.i.. Ces observations sont précisées et complétées dans le théorème qui suit.

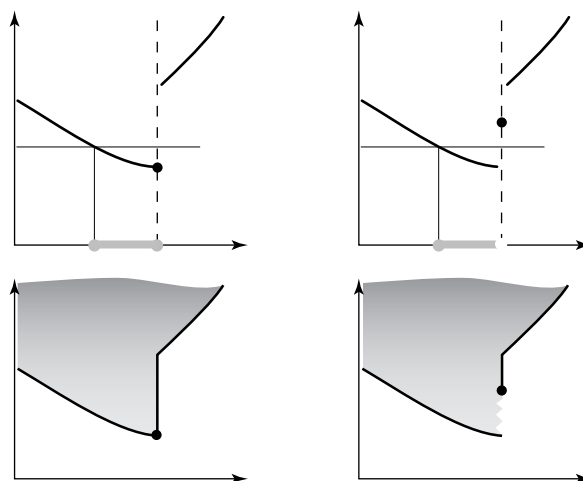


Figure 4.10: Une fonction s.c.i., une modification non s.c.i., les ensembles de niveau et les épigraphes associés

**Théorème 4.30.** Pour une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$  (non nécessairement convexe), les trois faits suivants sont équivalents :

<sup>3</sup>La propriété de Lipschitz est une propriété du type (2.12) (avec une constante  $L$  au second membre à la place de  $\|f\|$ ), et "localement" signifie qu'il existe une telle constante lorsqu'on se limite à un voisinage borné de tout point.

1.  $f$  est s.c.i. ;
2. son épigraphe est fermé ;
3. ses ensembles de niveau sont fermés.

*Démonstration.*

- 1  $\Rightarrow$  2 : Prenons une suite  $\{(x^k, \xi^k)\} \subset \mathbb{U}(f)$  convergeant vers  $(x, \xi)$ . Puisque  $f(x^k) \leq \xi^k$  et que  $\xi = \lim \xi^k \geq \liminf f(x^k) \geq f(x)$  du fait que  $f$  est s.c.i., on en déduit que  $(x, \xi) \in \mathbb{U}(f)$  et donc  $\mathbb{U}(f)$  est fermé.
- 2  $\Rightarrow$  3 : Dans  $\mathcal{X} \times \mathbb{R}$ , l'ensemble de niveau  $\overline{\mathbb{V}}_\beta(f)$  au niveau  $\beta$  de  $f$  est la section de  $\mathbb{U}(f)$  (supposé fermé) par l'hyperplan "horizontal" au niveau  $\beta$  (évidemment lui aussi fermé), cette section étant ensuite ramenée sur  $\mathcal{X}$  par une translation de  $(0, -\beta)$ . Cette section, et par conséquent  $\overline{\mathbb{V}}_\beta(f)$ , sont donc fermés.
- 3  $\Rightarrow$  1 : Si  $f$  n'est pas s.c.i. en un point  $x$ , il doit exister une suite  $\{x^k\}$  telle que  $x^k \rightarrow x$  et  $f(x^k) \rightarrow \beta < f(x)$ . Soit  $\beta' \in ]\beta, f(x)[$ . À partir d'un certain rang,  $f(x^k) \leq \beta'$  et donc, à partir de ce rang, la (sous-)suite  $\{x^k\}$  appartient à  $\overline{\mathbb{V}}_{\beta'}(f)$ . Comme  $x^k \rightarrow x$  mais que  $f(x) > \beta'$ ,  $x \notin \overline{\mathbb{V}}_{\beta'}(f)$ , et donc cet ensemble de niveau ne peut être fermé.

□

**Corollaire 4.31.** Une fonction convexe  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  qui est s.c.i. pour  $\mathcal{X}$  muni de la topologie forte est aussi s.c.i. pour  $\mathcal{X}$  muni de la topologie faible.

Ce corollaire résulte immédiatement de l'application du Corollaire 3.56 à  $\mathbb{U}(f)$ .

**Corollaire 4.32.** L'enveloppe supérieure d'une famille de fonctions s.c.i. est s.c.i.. En particulier, l'enveloppe supérieure d'une famille de fonctions affines continues ou convexes s.c.i. est convexe et s.c.i.. Donc la classe des fonctions convexes s.c.i. est stable par enveloppe supérieure.

*Démonstration.* L'enveloppe supérieure d'une famille de fonctions a pour épigraphe l'intersection des épigraphes des fonctions de cette famille et cette intersection est fermée si tous ces épigraphes le sont.

□

**Remarque 4.33.** Le domaine  $\mathbb{D}(f)$  d'une fonction convexe s.c.i. n'est pas nécessairement fermé. Considérer par exemple la fonction (voir Figure 4.11)

$$f : \mathbb{R} \rightarrow \overline{\mathbb{R}}, \quad f(x) = \begin{cases} |\operatorname{tang} x| & \text{pour } -\pi/2 < x < \pi/2 \\ +\infty & \text{sinon.} \end{cases} \quad (4.11)$$

Observer que cela montre au passage que la projection d'un convexe fermé n'est pas nécessairement fermée.

**Exercice 4.34.** Montrer que la fonction indicatrice d'un sous-ensemble  $A$  est convexe s.c.i. si et seulement si le sous-ensemble est convexe fermé.

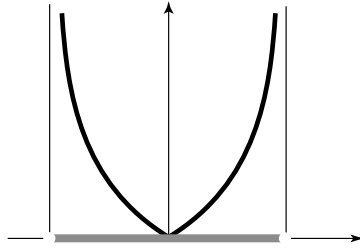


Figure 4.11: Graphe de la fonction (4.11) et son domaine

### 4.7.3 Enveloppe convexe s.c.i.

On peut reprendre toutes les considérations sur l’enveloppe convexe faites au §4.6 mais en y rajoutant maintenant la préoccupation d’approximer une fonction donnée par la plus grande fonction convexe *et s.c.i.* qui est en dessous de cette fonction. En effet, comme on vient de le dire dans la section précédente, les fonctions convexes s.c.i. sont appelées à jouer un rôle privilégié. Compte tenu des considérations faites ci-dessus sur les fonctions s.c.i. et les enveloppes convexes de fonctions d’une part, celles faites au chapitre précédent sur les enveloppes convexes fermées de sous-ensembles (voir §3.6.3) — considérations qui s’appliquent ici aux épigraphes — d’autre part, il est immédiat de démontrer le résultat suivant.

**Théorème 4.35 (Enveloppe convexe s.c.i. d’une fonction).** *Pour une fonction  $f$  donnée, si elle est minorée par une fonction affine, il existe une plus grande fonction convexe s.c.i. inférieure ou égale à  $f$ . Elle est appelée enveloppe convexe s.c.i. de  $f$  notée  $\overline{\odot}(f)$ . On a que*

$$\Psi(\overline{\odot}(f)) = \overline{\odot}(\Psi(f)) = \overline{\odot}(\overline{\Psi(f)}) .$$

Par ailleurs,

$$\overline{\odot}(f)(x) = \liminf_{y \rightarrow x} \odot(f)(y) .$$

Le dernier résultat indique que l’on peut d’abord calculer l’enveloppe convexe de  $f$  par la formule (4.10), puis modifier cette enveloppe convexe, partout où elle n’est pas continue, par une “régularisation s.c.i.” obtenue par la formule ci-dessus.

En fait, il existe une autre façon de calculer cette enveloppe convexe s.c.i., et cette autre façon est en rapport avec la remarque que les sous-ensembles convexes *fermés* font l’objet d’une description externe par l’intersection des demi-espaces fermés délimités par tous leurs hyperplans d’appui.

**Lemme 4.36.** *Considérons une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  convexe s.c.i.. Son épigraphe est égal à l’intersection des épigraphes de tous les fonctions affines continues qui minorent  $f$  ainsi que de tous les demi-espaces “verticaux” dans  $\mathcal{X} \times \mathbb{R}$  dont la projection (ou trace) sur  $\mathcal{X}$  est un hyperplan d’appui de  $\overline{\Upsilon}(f)$ .*

*Démonstration.* On sait que l’épigraphe est un convexe fermé et donc il est égal à l’intersection de tous les demi-espaces qui le contiennent (que ce soit des hyperplans d’appui ou des hyperplans parallèles à des hyperplans d’appui mais laissant le convexe tout entier dans un demi-espace qui le délimite — on les appellera “hyperplans séparateurs”). Ces hyperplans sont définis dans  $\mathcal{X} \times \mathbb{R}$  par

$$\{(x, \xi) \in \mathcal{X} \times \mathbb{R} \mid \langle r, x \rangle + \rho \xi = \beta\}$$

pour tout triplet  $(r, \rho, \beta)$  tels que  $r \in \mathcal{X}$ ,  $\rho \in \mathbb{R}$ ,  $\beta \in \mathbb{R}$  et tels que

$$\sup_{(x, \xi) \in \mathfrak{U}(f)} (\langle r, x \rangle + \rho \xi) \leq \beta . \quad (4.12)$$

On observe que nécessairement  $\rho \leq 0$ , sinon on aboutit à une contradiction dans l'inégalité (4.12) puisque  $\xi$  peut prendre des valeurs arbitrairement grandes dans un épigraphe.<sup>4</sup> Il faut alors distinguer les cas  $\rho < 0$  et  $\rho = 0$ . Observons aussi que si  $(x, \xi) \in \mathfrak{U}(f)$ , alors  $x \in \mathfrak{T}(f)$  (voir Remarque 4.6).

Si  $\rho < 0$ , il est clair que le sup en  $\xi$  à  $x$  fixé sera atteint pour la valeur  $f(x)$  (finie car  $x \in \mathfrak{T}(f)$ ) et on peut remplacer (4.12) par l'inéquation variationnelle équivalente

$$\forall x \in \mathfrak{T}(f), \quad \langle r, x \rangle + \rho f(x) \leq \beta .$$

En divisant par  $-\rho > 0$ , ce qui ne change pas le sens de l'inégalité, on obtient encore

$$\forall x \in \mathfrak{T}(f), \quad f(x) \geq \langle -r/\rho, x \rangle + \beta/\rho .$$

Autrement dit, les hyperplans séparateurs de l'épigraphe sont liés ici à des fonctions affines continues qui minorent la fonction  $f$  (et donc  $\mathfrak{U}(f)$  est inclus dans l'épigraphe de ces fonctions). Réciproquement, on peut remonter le calcul et montrer que toute fonction affine continue de ce type engendre un hyperplan séparateur de  $\mathfrak{U}(f)$ .

Si  $\rho = 0$ , on obtient que  $r$  est tel que

$$\forall x \in \mathfrak{T}(f), \quad \langle r, x \rangle \leq \beta ,$$

ce qui caractérise un hyperplan séparateur de  $\mathfrak{T}(f)$  (ou de  $\overline{\mathfrak{T}(f)}$ ) dans  $\mathcal{X}$ , ce que l'on peut voir aussi dans  $\mathcal{X} \times \mathbb{R}$  comme un hyperplan séparateur de  $\mathfrak{U}(f)$  “vertical” de “coefficients directeurs”  $(r, 0)$ . Là encore, la réciproque est immédiate.

□

**Corollaire 4.37.** *Toute fonction convexe s.c.i. est égale à l'enveloppe supérieure de toutes les fonctions affines continues qui minorent la fonction.*

*Démonstration.* La seule difficulté, en fonction de ce qui précède, est de montrer que l'on n'a pas besoin de considérer, dans l'opération d'enveloppe supérieure — c'est-à-dire en fait l'opération d'intersection de demi-espaces séparateurs de  $\mathfrak{U}(f)$ , les demi-espaces associés à des hyperplans verticaux s'appuyant dans  $\mathcal{X}$  sur les hyperplans séparateurs de  $\mathfrak{T}(f)$  (notons déjà que cette catégorie est vide si  $\mathfrak{T}(f) = \mathcal{X}$ ). Dans ce qui suit, on va considérer tout triplet  $(r, \rho, \beta) \in \mathcal{X} \times \mathbb{R} \times \mathbb{R}$  comme définissant un demi-espace fermé de  $\mathcal{X} \times \mathbb{R}$ , à savoir  $\{(x, \xi) \mid \langle r, x \rangle + \rho \xi \leq \beta\}$ , et on notera que l'intersection de ces demi-espaces pour une famille de triplets revient à considérer les inégalités pour l'union de ces triplets. Pour que les choses soient claires, on choisit de parler uniformément d'union de triplets sans perdre de vue qu'il s'agit de faire l'intersection de demi-espaces. Soit

- $\mathcal{F}_0$  la famille de triplets  $(r_0, 0, \beta_0)$  tels que les demi-espaces de  $\mathcal{X}$  associés,  $\langle r_0, x \rangle \leq \beta_0$ , contiennent tout  $\mathfrak{T}(f)$  ;
- $\mathcal{F}_1$  la famille de triplets  $(r_1, \rho_1, \beta_1)$  tels que  $\rho_1 < 0$  et tels que les demi-espaces de  $\mathcal{X} \times \mathbb{R}$  associés,  $\langle r_1, x \rangle + \rho_1 \xi \leq \beta_1$ , contiennent tout  $\mathfrak{U}(f)$  ;

<sup>4</sup>Au cas où le lecteur ne l'aurait pas déjà remarqué, on est en train de reprendre la discussion du §4.5 qui nous a finalement conduit à la formule (4.9)

- $\mathcal{F}_\alpha = \bigcup (r_1 + \alpha r_0, \rho_1, \beta_1 + \alpha \beta_0)$ , l'union étant prise pour tous les  $(r_i, \rho_i, \beta_i)$  parcourant leurs familles  $\mathcal{F}_i$  respectives ( $i = 0, 1$ , avec  $\rho_0 \equiv 0$ ) ;
- $\mathcal{F}_2 \stackrel{\text{def}}{=} \bigcup_{\alpha \geq 0} \mathcal{F}_\alpha$ .

On va montrer que l'intersection de demi-espaces associés à tous les éléments de  $\mathcal{F}_0 \cup \mathcal{F}_1$  engendre le même sous-ensemble de  $\mathcal{X} \times \mathbb{R}$  que celui engendré par  $\mathcal{F}_2$ . Or on note d'ores et déjà que pour tout élément de  $\mathcal{F}_2$ , le scalaire  $\rho$  est strictement négatif et on peut donc associer aux triplets de  $\mathcal{F}_2$  des minorantes affines continues.

D'une part, si  $(r_i, \rho_i, \beta_i) \in \mathcal{F}_i, i = 0, 1$ , alors pour tout  $\alpha \geq 0$ , on obtient, par combinaison linéaires à coefficients positifs d'éléments de  $\mathcal{F}_0$  et  $\mathcal{F}_1$  que

$$\langle r_1 + \alpha r_0, x \rangle + \rho_1 \xi \leq \beta_1 + \alpha \beta_0,$$

donc on a engendré un élément de  $\mathcal{F}_2$ . Réciproquement, si l'inégalité ci-dessus est vérifiée pour tout  $\alpha \geq 0$ , on peut en particulier prendre  $\alpha = 0$  (et on retrouve donc un élément de  $\mathcal{F}_1$ ), puis diviser par  $\alpha$  (ce qui ne change pas le sens de l'inégalité) et laisser tendre  $\alpha$  vers  $+\infty$  (et on retrouve alors un élément de  $\mathcal{F}_0$ ). On a donc montré l'équivalence recherchée.  $\square$

**Corollaire 4.38.** *L'enveloppe convexe s.c.i. d'une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est donnée par la formule*

$$\overline{\odot}(f)(x) = \sup_{(r, \alpha)} \{ \langle r, x \rangle + \alpha \mid \forall y \in \mathcal{X}, \langle r, y \rangle + \alpha \leq f(y) \}. \quad (4.13)$$

*Démonstration.* Notons

$$H(f) : x \mapsto H(f)(x) \stackrel{\text{def}}{=} \sup_{(r, \alpha)} \{ \langle r, x \rangle + \alpha \mid \forall y \in \mathcal{X}, \langle r, y \rangle + \alpha \leq f(y) \}$$

et observons que

- $H(f)(x) \leq f(x)$  pour tout  $x$  ;
- $H(f)$  est, pour toute fonction  $f$ , une fonction convexe s.c.i. comme enveloppe supérieure d'une famille de fonctions affines continues ;
- si  $f \geq g$ , alors il y a plus de couples  $(r, \alpha)$  qui vérifient les contraintes pour  $f$  que pour  $g$  et donc le sup qui définit  $H(f)(x)$  est une fonction croissante de la fonction  $f$ .

Par ailleurs, d'après le Corollaire 4.37,  $H(\overline{\odot}(f)) = \overline{\odot}(f)$  (comme pour toute fonction convexe s.c.i. qui peut être retrouvée par l'opération  $H$ ). Comme  $\overline{\odot}(f) \leq f$ , alors  $\overline{\odot}(f) = H(\overline{\odot}(f)) \leq H(f) \leq f$ , mais comme  $H(f)$  est convexe s.c.i. et que  $\overline{\odot}(f)$  est la plus grande minorante de  $f$  qui soit convexe s.c.i., on a aussi que  $H(f) \leq \overline{\odot}(f)$ , donc finalement l'égalité a lieu.  $\square$

On peut être plus précis en ce qui concerne la famille de fonctions affines continues à considérer dans les deux corollaires qui précèdent. En effet, dans la formule (4.13), pour  $r$  fixé, le plus grand  $\alpha$  possible tel que

$$\alpha \leq f(y) - \langle r, y \rangle, \quad \forall y \in \mathcal{X},$$

est égal à

$$\begin{aligned}\alpha &= \inf_{y \in \mathcal{X}} (f(y) - \langle r, y \rangle) \\ &= - \sup_{y \in \mathcal{X}} (\langle r, y \rangle - f(y)) \\ &= - \sup_{y \in \mathcal{T}(f)} (\langle r, y \rangle - f(y))\end{aligned}$$

(car si  $y \notin \mathcal{T}(f)$ , on obtient la valeur  $-\infty$  à éviter pour obtenir un sup)

$$= -f^F(r),$$

d'après la Définition 4.20. Par conséquent, (4.13) devient

$$\overline{\odot}(f)(x) = \sup_{r \in \mathcal{X}} (\langle r, x \rangle - f^F(r))$$

qui n'est rien d'autre que  $(f^F)^F(x)$  que l'on notera plus simplement  $f^{FF}(x)$ . On peut donc énoncer le corollaire suivant.

**Corollaire 4.39.** *L'enveloppe convexe s.c.i. d'une fonction  $f$  est égale à  $f^{FF}$ . La transformée de Fenchel est involutive<sup>5</sup> pour les fonctions convexes s.c.i..*

## 4.8 Différentiabilité et sous-différentiabilité des fonctions convexes

### 4.8.1 Diverses notions de différentiabilité

Sur un espace de Hilbert, il existe diverses notions de différentiabilité que nous examinons rapidement ici.

**Définition 4.40 (Dérivée directionnelle).** On appelle *dérivée directionnelle* de  $f : \mathcal{X} \rightarrow \mathcal{Y}$  ( $\mathcal{X}$  et  $\mathcal{Y}$  sont deux espaces de Hilbert) au point  $x \in \mathcal{X}$  et dans la direction  $d \in \mathcal{X}$ , notée  $Df(x; d)$ , la limite suivante, si elle existe dans  $\mathcal{Y}$  :

$$Df(x; d) = \lim_{\varepsilon \rightarrow 0_+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon}.$$

Considérons par exemple la fonction  $|x|$  de  $\mathbb{R}$  dans  $\mathbb{R}$ . L'application de la définition ci-dessus nous donne :

$$Df(0; 1) = 1 \quad \text{et} \quad Df(0; -1) = -1.$$

Sur cet exemple, il apparaît que la dérivée directionnelle en un point  $x$  peut ne pas être une fonction linéaire de la direction  $d$ .

**Définition 4.41 (Dérivée de Gâteaux).** Si  $f$  admet en  $x$  des dérivées directionnelles pour toutes les directions  $d$  et si  $Df(x; d)$  est une fonction linéaire continue de  $d$ , c'est-à-dire qu'il existe un opérateur  $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$  telle que  $Df(x; d) = A(d)$  pour tout  $d$ , alors la fonction  $f$  est dite *différentiable au sens de Gâteaux* au point  $x$  ; l'opérateur  $A$ , qui sera noté  $f'(x)$ , est appelée *dérivée de Gâteaux* de  $f$  au point  $x$ . Si une telle dérivée existe en tout point où  $f$  est définie, on dira que  $f$  est *différentiable au sens de Gâteaux*.

<sup>5</sup>“Involutive” signifie que si on applique deux fois de suite la transformation à un objet, on revient à l'objet d'origine.



**Remarque 4.42.** L'existence d'une dérivée de Gâteaux en un point n'implique pas du tout la continuité de la fonction en ce point. Considérons la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  suivante :

$$f(x, y) = \begin{cases} 1 & \text{si } y = x^2 \text{ et } x \neq 0, \\ 0 & \text{sinon,} \end{cases}$$

dont on a essayé de représenter le graphe sur la Figure 4.12. On peut vérifier que la dérivée de Gâteaux existe en  $(0, 0)$  mais la fonction n'est pas continue en ce point.

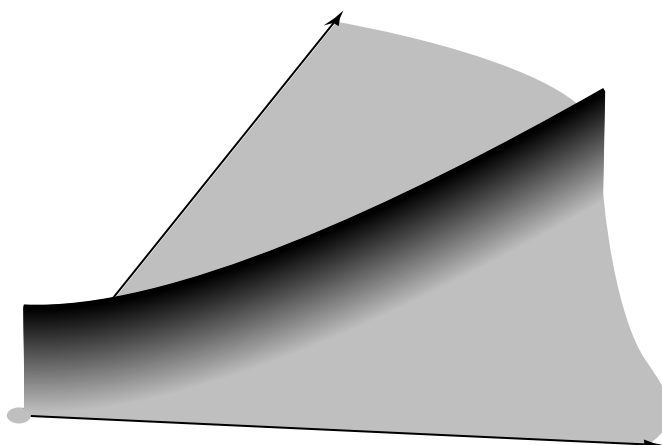


Figure 4.12: Graphe d'une fonction Gâteaux-différentiable mais non continue au point  $(0, 0)$

De tels "canulars" ne se produiront pas avec des fonctions convexes dont on a vu qu'elles sont essentiellement continues dans l'intérieur de leur domaine. Il existe une notion plus forte que la dérivée de Gâteaux qui interdit ce genre de paradoxe.

**Définition 4.43 (Dérivée de Fréchet).** La fonction  $f : \mathcal{X} \rightarrow \mathcal{Y}$  admet une dérivée de Fréchet au point  $x \in \mathcal{X}$  si il existe un opérateur linéaire continu, encore noté  $f'(x) \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ , tel que :

$$\lim_{\|d\| \rightarrow 0} \frac{f(x+d) - f(x) - f'(x)(d)}{\|d\|} = 0$$

et ce pour tout  $d \in \mathcal{X}$  tendant vers 0.

Évidemment, la différence fondamentale entre les notions de dérivées en  $x$  de Gâteaux et de Fréchet est que l'approche de  $x$  se fait uniquement le long de demi-droites dans le premier cas, et selon un chemin quelconque dans le second cas, ce qui est plus exigeant en terme d'existence de limite. Ainsi, la fonction de la Remarque 4.42 n'est pas Fréchet-différentiable en  $(0, 0)$ . La Fréchet-différentiabilité en  $x$  implique nécessairement la continuité en  $x$ .

Il se trouve qu'en Optimisation où l'on est surtout préoccupé de comparer des valeurs de fonctions coût, la notion de dérivée de Gâteaux suffit, et c'est ce que désignera désormais la notation  $f'(x)$ .

## 4.8.2 Dérivées directionnelles des fonctions convexes

Les fonctions convexes ne sont pas nécessairement dérivables sur leur domaine comme l'a déjà montré l'exemple de la fonction  $x \mapsto |x|$ . Cependant, on a le résultat suivant qui limite les comportements possibles d'une fonction convexe.

**Lemme 4.44.** Si  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est convexe, l'application  $\varepsilon \mapsto (f(x + \varepsilon d) - f(x))/\varepsilon$  est une fonction non décroissante de  $\varepsilon \in \mathbb{R}$  pour tout  $x \in \mathcal{X}$  et  $d \in \mathcal{X}$  (partout où elle est définie).

*Démonstration.* Soit  $\varepsilon$  et  $\varepsilon'$  tels que  $\varepsilon \leq \varepsilon'$ . Si  $\varepsilon \geq 0$ , on pose  $\alpha = \varepsilon/\varepsilon'$  et alors  $\alpha \in [0, 1]$  et

$$\begin{aligned} x + \varepsilon d &= \alpha(x + \varepsilon' d) + (1 - \alpha)x \\ \Rightarrow f(x + \varepsilon d) &\leq \alpha f(x + \varepsilon' d) + (1 - \alpha)f(x) \\ \Rightarrow f(x + \varepsilon d) - f(x) &\leq \alpha(f(x + \varepsilon' d) - f(x)) \\ \Rightarrow (f(x + \varepsilon d) - f(x))/\varepsilon &\leq (f(x + \varepsilon' d) - f(x))/\varepsilon'. \end{aligned}$$

Si  $\varepsilon \leq \varepsilon' \leq 0$ , on applique le résultat précédent en changeant  $d$  en  $-d$  et en changeant aussi les signes de  $\varepsilon$  et  $\varepsilon'$  (ce qui inverse l'ordre). Enfin, si  $\varepsilon \leq 0 \leq \varepsilon'$ , c'est  $x$  que l'on exprime comme la combinaison convexe  $\alpha(x + \varepsilon d) + (1 - \alpha)(x + \varepsilon' d)$  avec  $\alpha = \varepsilon'/(\varepsilon' - \varepsilon)$  qui appartient bien à  $[0, 1]$  et à partir de là, on exploite à nouveau la convexité de  $f$ .  $\square$

**Corollaire 4.45.** Une fonction convexe  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  admet toujours des dérivées directionnelles éventuellement égales à  $\pm\infty$ . De plus,

$$Df(x; d) = \inf_{\varepsilon > 0} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon}. \quad (4.14)$$

Enfin,  $-Df(x; -d) \leq Df(x; d)$ .

*Démonstration.* Le lemme précédent a montré que le ratio  $(f(x + \varepsilon d) - f(x))/\varepsilon$  est non croissant lorsque  $\varepsilon$  s'approche de 0 par valeurs positives. Il a donc une limite, si on admet aussi les valeurs  $\pm\infty$ , et cette limite est évidemment donnée par (4.14). Enfin, en utilisant à nouveau ce lemme pour  $\varepsilon \leq 0 \leq \varepsilon'$ , on montre facilement la deuxième affirmation.  $\square$

**Remarque 4.46.** La fonction

$$x \mapsto f(x) = \begin{cases} 1 - \sqrt{1 - x^2} & \text{si } -1 \leq x \leq 1, \\ +\infty & \text{sinon,} \end{cases}$$

(voir graphe sur la Figure 4.13) est convexe s.c.i. (et même différentiable) sur son domaine  $[-1, 1]$  avec  $Df(1; -1) = Df(-1; 1) = -\infty$  et donc (par linéarité)  $Df(1; 1) = Df(-1; -1) = +\infty$  (ceci pour montrer que les valeurs  $\pm\infty$  de la dérivée directionnelle sont possibles).

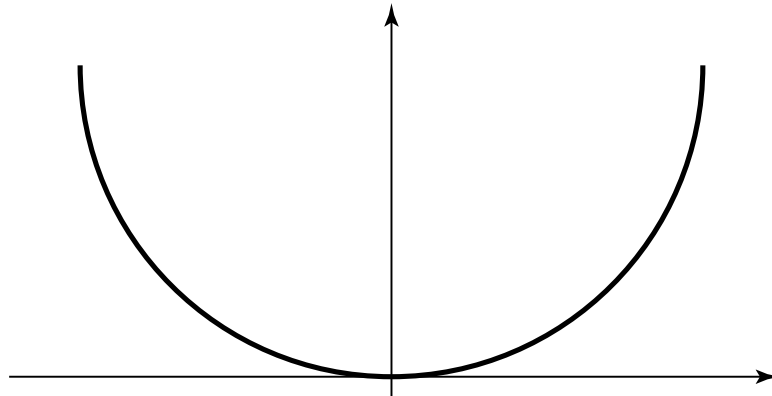


Figure 4.13: Graphe de la fonction  $x \mapsto 1 - \sqrt{1 - x^2}$  pour  $x \in [-1, 1]$

**Exercice 4.47.** Montrer que la fonction  $d \mapsto Df(x; d)$  est positivement homogène de degré 1 (voir (4.5)) et que de plus, elle est convexe si  $f$  est convexe.<sup>6</sup> En conséquence, elle est alors sous-additive, c'est-à-dire que

$$Df(x; d + d') \leq Df(x; d) + Df(x; d') .$$

**Remarque 4.48.** C'est un résultat général que les fonctions positivement homogènes de degré 1 sont convexes si et seulement si elles sont sous-additives. Si l'on observe que l'épigraphe d'une fonction positivement homogène est un cône — ce qui résulte immédiatement de la notion d'homogénéité — ce résultat est à rapprocher du résultat de l'Exercice 3.17.

Pour une fonction différentiable, on a bien sûr l'égalité  $-Df(x; -d) = Df(x; d)$  par linéarité par rapport à  $d$ . On peut donc considérer que  $Df(x; d) + Df(x; -d)$  (qui vaut 0 dans le cas différentiable et qui est non négative en général) mesure le “défaut de différentiabilité” de la fonction  $f$  au point  $x$  dans la direction (non orientée)  $d$ .

Par exemple, pour la fonction  $x \rightarrow |x|$ , on a vu que  $Df(0; 1) = 1$  et  $Df(0; -1) = 1$  et on vérifie que l'inégalité stricte  $-Df(0; -1) < Df(0; 1)$  a bien lieu. Cette inégalité exprime d'une manière générale que les “pentes” de la fonction convexe de  $\mathbb{R}$  dans  $\overline{\mathbb{R}}$  :

$$\varphi_{x,d}(\alpha) \stackrel{\text{def}}{=} f(x + \alpha d)$$

sont non décroissantes lorsque  $\alpha$  passe par 0 en croissant. On va voir que ceci est une propriété caractéristique des fonctions convexes qui se généralise dans  $\mathcal{X}$ .

### 4.8.3 Sous-différentiabilité des fonctions convexes

#### 4.8.3.1 Sous-gradients et interprétation géométrique

Une “minorante affine continue” d'une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est définie par un “coefficient directeur”  $r \in \mathcal{X}$  et une “ordonnée à l'origine”  $\alpha$  tels que

$$f(y) \geq \langle r, y \rangle + \alpha, \quad \forall y \in \mathcal{X},$$

et cette minorante est “exacte” en  $x$  si l'égalité a lieu en  $y = x$  dans l'égalité précédente. Alors,  $\alpha = f(x) - \langle r, x \rangle$  et l'inégalité précédente se réécrit

$$f(y) - f(x) \geq \langle r, y - x \rangle, \quad \forall y \in \mathcal{X}. \tag{4.15}$$

**Définition 4.49 (Sous-gradient, sous-différentiel).** Pour une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ , on appelle *sous-gradient* de  $f$  en  $x$  (avec  $f(x) < +\infty$ ) tout vecteur  $r \in \mathcal{X}$  qui vérifie (4.15). L'ensemble (vide ou non vide) de ces vecteurs est appelé *sous-différentiel* de  $f$  en  $x$  et est noté  $\partial f(x)$ .

Il est clair qu'au lieu de parler de minorante affine continue (et éventuellement exacte) d'une fonction, on aurait pu parler d'hyperplan séparateur (et éventuellement d'appui) de l'épigraphe associé. Si  $r \in \partial f(x)$ , il est équivalent de dire que l'hyperplan d'équation  $\langle r, y \rangle - \xi = \langle r, x \rangle - f(x)$  (les variables sont  $(y, \xi) \in \mathcal{X} \times \mathbb{R}$ ) est un hyperplan d'appui de  $\mathbb{U}(f)$  au point  $(x, f(x))$ . En particulier,  $(r, -1)$  est le “coefficient directeur” de cet hyperplan. On poursuit cette discussion géométrique du sous-différentiel sous la forme du lemme suivant (que l'on a tenté d'illustrer par la Figure 4.14).

**Lemme 4.50.** *Un vecteur  $r$  appartient à  $\partial f(x)$  pour une fonction convexe  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  si et seulement si  $(r, -1)$  appartient au cône normal à  $\mathbb{U}(f)$  au point  $(x, f(x))$  et si  $r$  appartient au cône normal à l'ensemble de niveau  $\overline{\nabla}_{f(x)}(f)$  au point  $x$ .*

<sup>6</sup>Pour rester consistant avec notre option de ne pas parler des fonctions convexes à valeurs  $-\infty$ , rajoutons l'hypothèse “et si  $Df(x; \cdot)$  ne prend pas la valeur  $-\infty$ ”.

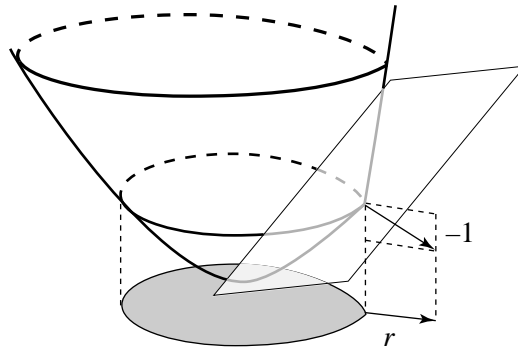


Figure 4.14: Interprétation géométrique du sous-gradient

**Remarque 4.51.** Attention, dans le deuxième énoncé sur le cône normal à l'ensemble de niveau, il ne s'agit bien que d'une condition nécessaire mais pas suffisante (on a en effet perdu trop d'information sur  $f$  en ne parlant que de son ensemble de niveau). Autrement dit, le cône orthogonal à un ensemble de niveau est stable par multiplication par un scalaire positif, ce qui n'est pas — en général — le cas d'un sous-différentiel — sauf exception examinée plus loin pour les fonctions indicatrices. Dans le cas du cône normal à l'épigraphe pour lequel la condition est nécessaire et suffisante, la “normalisation” adéquate provient du  $-1$  mis sur l’“axe vertical”.

*Démonstration.* Dire que  $r \in \partial f(x)$ , c'est dire que  $r$  vérifie (4.15) où l'on peut de façon équivalente remplacer  $f(y)$  par  $\xi \geq f(y)$  (l'équivalence est triviale), ce qui s'écrit encore :

$$\langle r, y - x \rangle + (-1)(\xi - f(x)) \leq 0, \quad \forall (y, \xi) \in \mathfrak{U}(f) .$$

On reconnaît bien là la définition du cône orthogonal à l'épigraphe au point  $(x, f(x))$  (voir Définition 3.20).

Par ailleurs, en considérant dans (4.15) des  $y \in \overline{\nabla}_{f(x)}(f)$ , c'est-à-dire tels que  $f(y) \leq f(x)$ , on obtient

$$\langle r, y - x \rangle \leq f(y) - f(x) \leq 0, \quad \forall y \in \overline{\nabla}_{f(x)}(f) ,$$

ce qui caractérise bien le cône normal à  $\overline{\nabla}_{f(x)}(f)$  en  $x$ . □

**Exercice 4.52.** Montrer que pour un convexe fermé  $A$ ,  $\partial I_A(x)$  est égal à  $A_x^\perp$ .

### 4.8.3.2 Existence de sous-gradients

En fonction du Lemme 4.50, l'existence d'un sous-gradient en  $x_0$ , donc aussi celle d'une minorante affine, suppose d'abord celle de la non réduction triviale à  $\{0\}$  du cône normal à l'épigraphe au point  $(x, f(x))$ .

Il faut d'abord que ce point appartienne à la frontière de l'épigraphe : on considèrera donc des points  $x \in \mathfrak{T}(f)$ .

Le Lemme 3.54 a traité, en dimension finie, de la non trivialité du cône normal à un convexe en un point de la frontière et, après la démonstration de ce lemme, on a indiqué que le même résultat en dimension infinie requiert la non vacuité de l'intérieur du convexe (ici l'épigraphe). Ce résultat en dimension infinie a été énoncé sans démonstration car il fait appel à un théorème de séparation au sens large que nous n'avons pas traité ici.

La question de la non vacuité de l'intérieur de l'épigraphe nécessite celle de la non vacuité de l'intérieur du domaine et elle lui est même équivalente s'il existe un ouvert  $O$  sur lequel  $f$  est bornée (par une constante  $\alpha$  : il suffit en effet de considérer l'ouvert  $O \times ]\alpha, +\infty[$  qui appartient alors à l'intérieur de  $\mathbb{U}(f)$ ). Le Théorème 4.27 a indiqué que l'existence d'un ouvert sur lequel  $f$  est bornée (qui implique déjà la non vacuité de l'intérieur du domaine) est équivalente à la continuité de  $f$  en tout point de l'intérieur du domaine. Tout ceci nous amène à énoncer, sans démonstration, le résultat suivant (pour une démonstration complète voir, par exemple [8, Proposition 5.2]).

**Théorème 4.53.** *Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe qui est finie et continue en au moins un point  $x$ . Alors, le sous-différentiel de  $f$  est non vide en tout point de l'intérieur de son domaine (qui est lui-même non vide) et en particulier en  $x$ .*

### 4.8.3.3 Une autre définition géométrique du sous-différentiel, propriétés et lien avec la dérivée directionnelle

On donne maintenant un autre point de vue sur le sous-différentiel d'une fonction convexe plus en rapport avec la notion de dérivée directionnelle.

**Théorème 4.54.** *Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe. En un point  $x \in \mathcal{X}$ , on a que :*

$$\partial f(x) = \{r \in \mathcal{X} \mid \forall d \in \mathcal{X}, \langle r, d \rangle \leq Df(x; d)\} . \quad (4.16)$$

*Démonstration.* Si  $r$  vérifie (4.16), d'après (4.14) (avec  $\varepsilon = 1$  et  $d = y - x$ ), on voit que (4.15) est vérifiée (pour tout  $y$ ) et donc  $r \in \partial f(x)$ . Réciproquement, en partant de l'inégalité (4.15) et en posant  $y = \varepsilon d$ , on parvient sans difficulté à montrer (4.16).  $\square$

**Corollaire 4.55.** *Le sous-différentiel  $\partial f(x)$  d'une fonction  $f$  convexe en un point  $x$  est un sous-ensemble convexe fermé (éventuellement vide).*

*Démonstration.* Ceci pouvait déjà se voir avec (4.15) mais c'est encore plus clair ici puisque  $\partial f(x)$  apparaît dans (4.16) comme l'intersection d'une famille de demi-espaces fermés.  $\square$

**Corollaire 4.56.** *Si  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  est une fonction convexe et si elle est finie et continue en  $x$ , alors la fonction  $d \mapsto Df(x; d)$  est la fonction support de  $\partial f(x)$ . Autrement dit,*

$$Df(x, d) = \sup_{r \in \partial f(x)} \langle r, d \rangle . \quad (4.17)$$

*Démonstration.* Si  $f$  est finie et continue en  $x$ , elle est bornée au voisinage de  $x$  et d'après le Théorème 4.53,  $\partial f(x)$  est non vide. Soit  $r \in \partial f(x)$ . Alors, d'après (4.16) et (4.14),

$$\langle r, d \rangle \leq Df(x; d) \leq f(x + d) - f(x) .$$

On peut se contenter d'utiliser cet encadrement pour  $\|d\|$  aussi petit que l'on veut (donc en restant au voisinage de  $x$ ) puisque  $Df(x; \cdot)$  est positivement homogène de degré 1 (voir Exercice 4.47). Par conséquent, on en déduit que  $Df(x; d)$  est bornée inférieurement et supérieurement pour toute direction  $d$  (dès que l'on borne  $\|d\|$  évidemment). S'agissant d'une fonction convexe de  $d$  (voir encore l'Exercice 4.47), on en déduit qu'elle est continue (voir Lemme 4.26). La conclusion que  $Df(x; \cdot)$  est la fonction support de  $\partial f(x)$  est alors une conséquence de (4.16) et du Théorème 4.57 ci-après. La formule (4.17) n'est rien d'autre que la même affirmation.  $\square$

**Théorème 4.57.** Soit  $h : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe s.c.i. et positivement homogène de degré 1. Alors  $h$  est la fonction support de l'ensemble  $S_h$  défini par

$$S_h \stackrel{\text{def}}{=} \{r \in \mathcal{X} \mid \langle r, d \rangle \leq h(d), \forall d \in \mathcal{X}\} .$$

*Démonstration.* Comme  $h$  est une fonction convexe s.c.i., elle est l'enveloppe supérieure de toutes les fonctions affines qui la minorent. Considérons l'une quelconque de ces minorantes affines :

$$\langle r, d \rangle + \alpha \leq h(d), \quad \forall d \in \mathcal{X} .$$

Puisque  $h$  est homogène, alors  $h(0) = 0$ . On en déduit que  $\alpha$  est négatif ou nul. Ensuite, toujours à cause du caractère homogène, pour tout  $\varepsilon > 0$ ,

$$\varepsilon \langle r, d \rangle + \alpha \leq \varepsilon h(d) ,$$

ce qui montre, avec  $\varepsilon$  infiniment grand, que la fonction *linéaire*  $d \mapsto \langle r, d \rangle$  est aussi une fonction qui minore  $h$ . Or cette fonction est une meilleure (plus grande) minorante de  $h$  que la fonction affine correspondante puisque  $\alpha \leq 0$ . Autrement dit,  $h$  est aussi égale à l'enveloppe supérieure des fonctions linéaires (uniquement) qui la minorent. Ceci s'écrit :

$$\forall d, \quad h(d) = \sup \{ \langle r, d \rangle \mid \forall r, \forall y, \langle r, y \rangle \leq h(y) \} ,$$

mais, d'après la définition de  $S_h$ , ceci se réécrit :

$$\forall d, \quad h(d) = \sup \{ \langle r, d \rangle \mid \forall r \in S_h \} = \sup_{r \in S_h} \langle r, d \rangle .$$

Ceci correspond exactement au calcul de la fonction support de  $S_h$ . □

#### 4.8.3.4 Cas des fonctions convexes différentiables

Si  $\partial f(x)$  est réduit à un singleton, la formule (4.17) montre que  $Df(x; d)$  est une fonction linéaire de  $d$ , autrement dit  $f$  est (Gâteaux-) différentiable au point  $x$ . La réciproque est également vraie.

**Théorème 4.58.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe, finie et continue en  $x \in \mathcal{X}$ . Elle admet une dérivée  $f'(x)$  si et seulement si  $\partial f(x)$  est réduit à un seul point, à savoir  $f'(x)$ , et donc

$$f(y) - f(x) \geq \langle f'(x), y - x \rangle , \quad \forall y \in \mathcal{X} . \tag{4.18}$$

*Démonstration.* Le “si” vient d'être justifié. Montrons le “seulement si”. Montrons d'abord (4.18), ce qui montrera déjà que  $f'(x) \in \partial f(x)$ . Si  $f$  est différentiable, alors par définition  $Df(x; y - x) = \langle f'(x), y - x \rangle$ . Mais, toujours avec (4.14),  $Df(x; y - x) \leq f(y) - f(x)$ , d'où (4.18).

Supposons maintenant qu'il existe un autre  $r \in \partial f(x)$ . Par définition de la différentiabilité,  $Df(x; d) = \langle f'(x), d \rangle$  pour tout  $d$ , mais par ailleurs, d'après (4.17),  $Df(x; d) \geq \langle r, d \rangle$ . Par conséquent,  $\langle f'(x), d \rangle \geq \langle r, d \rangle$  pour tout  $d$ . Utilisons cette inégalité pour  $d = r - f'(x)$ , ce qui donne  $-\|r - f'(x)\|^2 \geq 0$ , prouvant que  $r = f'(x)$ . □

L'inégalité (4.18) exprime que le graphe d'une fonction convexe différentiable se situe au dessus de tous ses hyperplans tangents, ce qui est à rapprocher du fait que ce graphe est situé sous toutes ses cordes (voir Figure 4.2). Ces deux propriétés caractérisent les fonctions convexes (différentiables).

**Théorème 4.59.** Une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$  différentiable est convexe si et seulement si elle vérifie (4.18) pour tous  $x$  et  $y$ .

*Démonstration.* Il reste à montrer le “si”. On suppose (4.18) pour tous  $x$  et  $y$  et on utilise cette inégalité pour  $y = u$  et  $x = u + \alpha(v - u)$ ,  $\alpha \in [0, 1]$ ,  $u \in \mathcal{X}$ ,  $v \in \mathcal{X}$ , ce qui donne

$$f(u) - f(u + \alpha(v - u)) \geq \alpha \langle f'(u + \alpha(v - u)), u - v \rangle .$$


Avec  $y = v$  et le même  $x$ , on obtient

$$f(v) - f(u + \alpha(v - u)) \geq (1 - \alpha) \langle f'(u + \alpha(v - u)), v - u \rangle .$$

La combinaison convexe de ces deux inégalités avec les “poids”  $(1 - \alpha)$ , respectivement  $\alpha$ , conduit finalement à l’inégalité de la convexité.  $\square$

#### 4.8.3.5 Cas des fonctions strictement et fortement convexes

Les fonctions sous-différentiables, voire différentiables, qui sont *strictement* (voir (4.7)) ou *fortement* (voir (4.8)) convexes peuvent être caractérisées par des inégalités du type (4.15), voire (4.18), “renforcées”.

**Exercice 4.60.**  Montrer qu’une fonction strictement convexe, respectivement fortement convexe de module  $b$ , et sous-différentiable satisfait

$$\forall y \in \mathcal{X}, \forall x \in \mathcal{X}, y \neq x, \forall r \in \partial f(x), \quad f(y) - f(x) > \langle r, y - x \rangle , \quad (4.19)$$

respectivement,

$$\forall y \in \mathcal{X}, \forall x \in \mathcal{X}, \forall r \in \partial f(x), \quad f(y) - f(x) \geq \langle r, y - x \rangle + b \|y - x\|^2 / 2 . \quad (4.20)$$

Réciproquement, montrer qu’une fonction sous-différentiable qui satisfait (4.19), respectivement (4.20), est strictement convexe, respectivement fortement convexe de module  $b$ .

Observons que (4.20) indique qu’en tout point, la fonction peut être minorée non plus seulement par une fonction affine, mais par une parabole.

#### 4.8.3.6 Le sous-différentiel comme opérateur multivoque

La correspondance  $x \mapsto \partial f(x)$  est une application *multivoque* ou *multi-application*, c’est-à-dire qu’elle associe à un point  $x \in \mathcal{X}$  le sous-ensemble  $\partial f(x)$  de  $\mathcal{X}$ , donc c’est une application  $x \rightarrow 2^{\mathcal{X}}$  (famille de tous les sous-ensembles de  $\mathcal{X}$ ). De plus, elle est ici à “valeurs” convexes fermées.

**Définition 4.61.** Une multi-application  $A : \mathcal{X} \rightarrow 2^{\mathcal{X}}$  est *monotone* si

$$\forall x \in \mathcal{X}, \forall y \in \mathcal{X}, \forall r \in A(x), \forall s \in A(y), \quad \langle r - s, x - y \rangle \geq 0 .$$

Elle est *strictement monotone* si l’inégalité stricte a lieu ci-dessus dès que  $y \neq x$ . Elle est *fortement monotone* (de module  $b$ ) si

$$\exists b > 0 : \forall x \in \mathcal{X}, \forall y \in \mathcal{X}, \forall r \in A(x), \forall s \in A(y), \quad \langle r - s, x - y \rangle \geq b \|x - y\|^2 . \quad (4.21)$$

On vérifie que pour une simple application de  $\mathbb{R}$  dans  $\mathbb{R}$ , cette propriété exprime le caractère monotone non décroissant (respectivement, croissant dans le cas “strictement”, ou croissant plus vite qu’une fonction linéaire de pente  $b$  dans le cas “fortement”). Dans le cas d’un opérateur linéaire continu  $A \in \mathcal{L}(\mathcal{X}, \mathcal{X})$ , cette propriété s’exprime plus simplement (exercice)

$$\forall x \in \mathcal{X}, \langle A(x), x \rangle \geq 0, \text{ resp. } > 0 \text{ si } x \neq 0, \text{ resp. } \geq b \|x\|^2 .$$

Dans le cas où  $\mathcal{X}$  est de dimension finie, pour ce qui est de la forme quadratique associée,  $A$  peut être représenté dans une base par une matrice symétrique (voir Exercice 2.12), et on reconnaît alors la définition des matrices semi-définies, respectivement définies, positives. Dans le cas de dimension finie, il n’y a pas de différence entre stricte et forte monotonie (dans les deux cas, la plus petite valeur propre est strictement positive). En dimension infinie, étant donné que le spectre d’un opérateur linéaire continu peut être continu, il peut s’approcher infiniment près de 0 dans le cas strictement monotone (sans l’atteindre), alors qu’il reste borné inférieurement par  $b$ , le module de forte monotonie dans ce cas.

**Lemme 4.62.** *Le sous-différentiel d’une fonction convexe (respectivement, strictement, fortement convexe de module  $b$ ) est une multi-application monotone (respectivement, strictement, fortement monotone de module  $b$ ). Réciproquement, pour une fonction sous-différentiable, si le sous-différentiel est monotone, (respectivement, strictement, fortement monotone de module  $b$ ), la fonction est convexe (respectivement, strictement, fortement monotone de module  $b$ ).*

*Démonstration.* On se contentera de la démonstration dans le cas fort. Le cas ordinaire correspond à un module  $b$  nul, le cas strict est une simple adaptation. Il est plus facile de démontrer l’équivalence de la forte monotonie du sous-différentiel avec la propriété de “minorante quadratique” exprimée par (4.20). L’Exercice 4.60 a déjà montré l’équivalence de cette dernière propriété avec la forte convexité de module  $b$  de  $f$ . Supposons donc d’abord que (4.20) est vraie : on réécrit une deuxième fois cette inégalité en inversant le rôle de  $y$  et  $x$ , donc à l’aide d’un  $s$  quelconque dans  $\partial f(y)$  ; par addition de ces deux versions de l’inégalité (4.20), on obtient (4.21). Pour la réciproque, considérons la fonction convexe  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$

$$\varphi(\alpha) \stackrel{\text{def}}{=} f(z_\alpha) \quad \text{avec} \quad z_\alpha \stackrel{\text{def}}{=} (1 - \alpha)x + \alpha y.$$

On a

$$\begin{aligned} D\varphi(\alpha; 1) &= \inf_{\varepsilon > 0} (\varphi(z_{\alpha+\varepsilon}) - \varphi(z_\alpha)) / \varepsilon \\ &\geq \langle r_\alpha, y - x \rangle \end{aligned} \tag{4.22}$$

pour  $r_\alpha \in \partial f(z_\alpha)$

$$\geq \langle r, y - x \rangle + \alpha b \|y - x\|^2 \tag{4.23}$$

pour  $r \in \partial f(x)$  et en utilisant (4.21). Utilisant à nouveau (4.22) (pour la valeur “infinitésimale”  $\varepsilon = d\alpha$ ) puis (4.23), on déduit que

$$\begin{aligned} f(y) - f(x) &= \varphi(1) - \varphi(0) \\ &\geq \int_0^1 D\varphi(\alpha; 1) \, d\alpha \\ &\geq \int_0^1 (\langle r, y - x \rangle + \alpha b \|y - x\|^2) \, d\alpha \\ &= \langle r, y - x \rangle + b \|y - x\|^2 / 2 . \end{aligned}$$

□



On retiendra donc l'équivalence des trois propriétés exprimées par (4.3), (4.20) et (4.21).

**Remarque 4.63.** En nous limitant, pour simplifier le discours, au cas d'une fonction convexe différentiable d'une variable scalaire  $x$ , interprétée comme la fonction coût associée à une production  $x$ , le fait que la dérivée soit une fonction croissante de  $x$  s'interprète comme une notion de "coût marginal croissant" : plus on doit produire, plus on produira à un coût marginal élevé, parce que, par exemple, on utilise d'abord les moyens de production les plus économiques avant d'avoir recours aux moyens les plus chers.

Comme on peut le comprendre intuitivement, si une fonction convexe est deux fois différentiable, puisque sa dérivée première est un opérateur monotone, sa dérivée seconde doit être "positive", c'est-à-dire plus précisément que la forme quadratique

$$y \mapsto \langle f''(x)(y), y \rangle$$

est semi-définie positive.

**Exercice 4.64.** Montrer qu'une fonction deux fois différentiable est convexe si et seulement si sa dérivée seconde en tout point  $x$  définit une forme quadratique semi-définie positive.

Indiquons pour finir que le sous-différentiel est non seulement un opérateur multivoque (à valeurs convexes fermées) monotone, mais qu'il est de plus *monotone maximal*, ce qui signifie que son graphe ne peut pas être strictement inclus dans le graphe d'un autre opérateur monotone. Ceci implique concrètement que le graphe n'a pas de "trou", comme illustré par la Figure 4.15. Cette propriété est

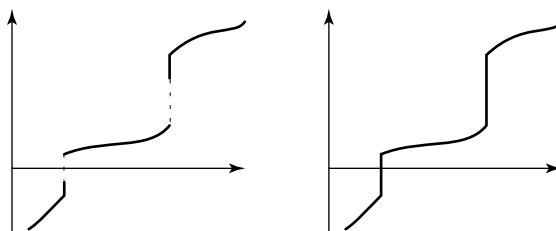


Figure 4.15: Graphe d'un opérateur monotone non maximal (à gauche) et version maximale monotone (à droite)

importante car elle signifie que la résolution de l'"équation"  $0 \in \partial f(x)$ , que l'on rencontrera bientôt en optimisation, aura une solution en  $x$  (avec toutefois des hypothèses supplémentaires qui garantissent que le graphe "va de part et d'autre de 0"), et on voit que ce n'est pas le cas sur la partie gauche de la figure (où le graphe "saute par dessus l'ordonnée 0"). On renvoie par exemple à [6] pour l'étude des opérateurs maximaux monotones.

#### 4.8.3.7 Calcul sous-différentiel

On considère ici un certain nombre d'opérations préservant la convexité des fonctions et on envisage le calcul du sous-différentiel des fonctions obtenues à partir de ceux des fonctions de départ. Très souvent, une formule d'inclusion est facile à démontrer et elle est proposée sous forme d'exercice. On indique ensuite l'hypothèse technique qui permet de démontrer l'inclusion inverse et donc l'égalité. La plupart du temps, la démonstration fait appel à un théorème de séparation au sens large que nous n'avons pas discuté en détail dans ce cours.

**Exercice 4.65.** Soit  $f$  et  $g$  deux fonctions convexes sous-différentiables de  $\mathcal{X}$  dans  $\overline{\mathbb{R}}$ . Montrer que

$$\forall x \in \mathcal{X}, \quad \partial(f + g)(x) \supset \partial f(x) + \partial g(x) . \tag{4.24}$$

L'égalité peut être démontrée si il existe un point dans  $\nabla(f) \cap \nabla(g)$  où au moins l'une des fonctions est continue (voir [8, Proposition 5.6]).

**Exercice 4.66.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  une fonction convexe et  $A : \mathcal{Y} \rightarrow \mathcal{X}$  un opérateur linéaire continu. Montrer que

$$\forall x \in \mathcal{X}, \quad \partial(f \circ A)(x) \supset A^* \circ \partial f(x). \quad (4.25)$$

L'égalité peut être démontrée si il existe un point dans  $\nabla(f) \cap \text{im } A$  où  $f$  est continue (voir [8, Proposition 5.7]).

**Exercice 4.67.** Soit  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$  définie comme l'enveloppe supérieure d'une famille de fonctions convexes  $\{f_j\}_{j \in J}$  :

$$f(x) \stackrel{\text{def}}{=} \sup_{j \in J} f_j(x).$$

A priori, rien n'est dit sur la nature de l'ensemble  $J$ . On pose

$$J(x) = \arg \max_{j \in J} f_j(x) \stackrel{\text{def}}{=} \{j \in J \mid f(x) = f_j(x)\}.$$

Montrer que

$$\partial f(x) \supset \overline{\bigcup_{j \in J(x)} \partial f_j(x)}. \quad (4.26)$$

Nous n'entrerons pas ici dans le détail des hypothèses techniques qui permettent de démontrer l'égalité. Il faut distinguer les cas extrêmes où  $J$  est fini d'une part, est un sous-ensemble d'un espace de Hilbert d'autre part. Dans le second cas, une hypothèse de semi-continuité supérieure<sup>7</sup> des fonctions  $j \mapsto f_j(x)$  pour tout  $x$  est souvent supposée.

La formule (4.26) montre en tout cas que lorsque  $J(x) = \arg \max_{j \in J} f_j(x)$  n'est pas réduit à un point (c'est-à-dire que, pour un point  $x$ , le sup en  $j$  est réalisé par plusieurs fonctions — voir par exemple la Figure 4.7), il y a peu de chance que  $f$  soit différentiable même si toutes les fonctions  $f_j$  le sont. Par exemple,  $x \mapsto |x|$  peut être vue comme l'enveloppe supérieure des deux fonctions linéaires  $x \mapsto x$  et  $x \mapsto -x$  et elle n'est pas différentiable en 0 parce que les deux fonctions linéaires valent simultanément 0. Notons de plus que les deux fonctions sont différentiables et ont pour dérivées 1 et  $-1$  respectivement, et que le sous-différentiel en 0 de  $|\cdot|$  est égal justement à l'enveloppe convexe de ces valeurs, c'est-à-dire à  $[-1, 1]$ . La formule (4.26) (avec égalité) est donc bien vérifiée dans ce cas.

Ceci dit, il ne faut pas en conclure que la réduction de l'arg max à un singleton est une condition *nécessaire* pour la différentiabilité de  $f$  (il se pourrait en effet que toutes les fonctions  $f_j$ ,  $j \in J(x)$  aient le même gradient en  $x$ , c'est-à-dire qu'elles soient "tangentes" : dessiner le graphe d'un tel exemple).

La formule suivante concerne la fonction marginale (4.7). On a vu à l'Exercice 4.18 qu'on peut ainsi obtenir une fonction convexe d'une variable à partir d'une opération d'enveloppe *inférieure*, à condition de partir d'une fonction *conjointement* convexe par rapport au couple de variables. On a également vu que cela correspond à une opération de projection sur l'un des espaces composants en terme d'épigraphe. On va d'abord examiner le cas de la projection d'un sous-différentiel.


**Exercice 4.68.** Soit  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$  une fonction convexe et sous-différentiable. Montrer que

$$\Pi_{\mathcal{X}}(\partial f(x, y)) = \partial_x f(x, y), \quad (4.27)$$

où  $\partial_x f(x, y)$  correspond au sous-différentiel de la fonction  $x \mapsto f(x, y)$  à  $y$  donné. En déduire que

$$\partial f(x, y) \subset \partial_x f(x, y) \times \partial_y f(x, y). \quad (4.28)$$

<sup>7</sup>Une fonction est "semi-continue supérieurement" (abréviation : s.c.s.) si son opposée est s.c.i..

**Exercice 4.69.**  Soit  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$  une fonction convexe et soit  $g$  définie par (4.7). On pose

$$Y(x) = \arg \min_{y \in \mathcal{Y}} f(x, y) \stackrel{\text{def}}{=} \{y \in \mathcal{Y} \mid g(x) = f(x, y)\} .$$

Montrer que

$$\partial g(x) \subset \bigcap_{y \in Y(x)} \partial_x f(x, y) . \tag{4.29}$$

Pour cet exercice, on a besoin en fait de savoir que  $y \in Y(x) \Leftrightarrow 0 \in \partial_y f(x, y)$ , ce que nous montrerons ultérieurement dans ce cours. En corollaire de la formule (4.29), on voit que si  $f$  est différentiable en  $x$ , alors soit  $g$  n'est pas sous-différentiable au point  $x$ , soit elle est obligatoirement différentiable : en effet, puisque  $\partial_x f(x, y) = \{f'_x(x, y)\}$  est un singleton pour tout  $y$ , alors ou bien tous les  $f'_x(x, y)$  sont égaux lorsque  $y$  parcourt  $Y(x)$  et  $g'(x)$  est égale à cette unique valeur commune, ou bien l'intersection de singletons est vide et c'est aussi le cas de  $\partial g(x)$ .

**Exercice 4.70.** On a défini par (4.6) l'inf-convolution  $h$  de deux fonctions convexes  $f$  et  $g$  et on a vu que c'est une fonction bien définie si  $f$  et  $g$  admettent une minorante affine commune. Montrer que

$$\partial h(x) = \partial f(\xi) \cap \partial g(\eta) \text{ en tous les points } \xi, \eta \text{ tels que } h(x) = f(\xi) + g(\eta) . \tag{4.30}$$

### 4.8.3.8 Sous-différentiel et transformée de Fenchel

Revenons pour terminer sur la notion de transformée de Fenchel introduite au §4.5 pour établir un lien avec la notion de sous-différentiel. On a vu que par définition de la transformée de Fenchel, l'inégalité (4.8) est toujours vraie. Le théorème qui suit discute de la situation du cas d'égalité.

**Théorème 4.71.** *Soit  $f$  une fonction convexe. L'inégalité (4.8) est une égalité si et seulement si  $p \in \partial f(x)$ . De plus, si  $f$  est s.c.i., alors l'égalité dans (4.8) est aussi équivalente au fait que  $x \in \partial f^F(p)$ .*

*Démonstration.* Si l'égalité a lieu dans (4.8) pour un certain  $x$  et un certain  $p$ , alors on peut retrancher cette égalité à l'inégalité écrite pour un autre  $y$  et pour le même  $p$ . Ce que l'on obtient ainsi signifie que  $p \in \partial f(x)$ . Réciproquement, si  $p \in \partial f(x)$ , on a

$$\forall y, \langle p, y \rangle - f(y) \leq \langle p, x \rangle - f(x) ,$$

ce qui montre que c'est  $x$  qui réalise le sup dans la Définition 4.20 de la transformée de Fenchel de  $f$  au point  $p$ . On obtient donc l'égalité dans (4.8).

Pour la seconde partie du théorème, on observe d'abord que la première partie qui vient d'être démontrée peut aussi s'appliquer au couple de fonctions  $(f^F, f^{FF})$  au lieu de  $(f, f^F)$ . Ceci se traduit notamment, en cas d'égalité pour un couple  $(p, x)$  par le fait que  $x \in \partial f^{FF}(p)$ . Mais dans le cas où  $f$  est s.c.i., elle est égale à son enveloppe convexe s.c.i. qui est égale à  $f^{FF}$  en général (Corollaire 4.39). Donc  $f^{FF} = f$ , ce qui achève la démonstration  $\square$

Introduisons la notion d'*inverse ensembliste* d'une application, voire même d'une multi-application  $A : \mathcal{X} \rightarrow 2^{\mathcal{Y}}$  : c'est la multi-application de  $\mathcal{Y}$  dans  $2^{\mathcal{X}}$  définie par

$$A^{-1}(y) = \{x \in \mathcal{X} \mid y \in A(x)\} .$$

Autrement dit,  $y \in A(x)$  est par définition équivalent à  $x \in A^{-1}(y)$ . Alors, du Théorème 4.71, et pour une fonction convexe s.c.i. et différentiable, on déduit que

$$x \in (\partial f)^{-1}(p) \Leftrightarrow p \in \partial f(x) \Leftrightarrow x \in \partial f^F(p) ,$$

et par conséquent

$$(\partial f)^{-1} = \partial f^F . \quad (4.31)$$

Ceci montre que, considérée plutôt au niveau des sous-différentiels qu'à celui des fonctions, la transformée de Fenchel agit comme une inversion !

**Exercice 4.72.** On a vu à l'Exercice 4.22 que la transformée de Fenchel  $h^F$  d'une inf-convolution  $h = f \square g$  est la somme des transformées de Fenchel  $f^F$  et  $g^F$ . En faisant l'hypothèse technique nécessaire à l'utilisation de la formule (4.24) avec égalité, et en s'appuyant sur (4.31), montrer que le sous-différentiel de l'inf-convolution est la *moyenne harmonique* (l'inverse de la somme des inverses) des sous-différentiels des fonctions de départ, c'est-à-dire

$$\partial h = \left( (\partial f)^{-1} + (\partial g)^{-1} \right)^{-1} .$$

Autrement dit, au niveau des sous-différentiels plutôt qu'au niveau des fonctions, l'inf-convolution agit comme la moyenne harmonique. Rapprocher cette formule de (4.30).

## 4.9 Résumons nous

Dans ce cours, on a choisi d'aborder l'étude des fonctions convexes après celle des ensembles convexes. On a vu qu'il est possible de passer de diverses façons des fonctions aux ensembles (en considérant les épigraphes ou les ensembles de niveau) et des ensembles aux fonctions (en considérant les fonctions indicatrices ou les fonctions support par exemple).

Comme pour les ensembles, la propriété géométrique de convexité des fonctions a des conséquences topologiques : en dimension infinie, continuité s.c.i. dans la topologie faible des fonctions convexes qui sont s.c.i. dans la topologie forte ; en dimension finie, les fonctions convexes ne peuvent pas avoir de "saut vertical fini", ce qui implique la continuité à l'intérieur relatif de leur domaine (les discontinuités, infinies, se limitent éventuellement au bord de ce domaine ; en dimension infinie, il faut supposer que la fonction est bornée au voisinage d'un point pour conclure à la continuité en ce point).

Le théorème de séparation pour les convexes fermés, que nous avons déduit au chapitre précédent de la notion de projection, avait eu pour conséquence la "définition externe" des convexes comme intersection des demi-espaces délimités par leurs hyperplans d'appui. En appliquant ces notions aux épigraphes de fonctions convexes s.c.i., on débouche sur la notion de sous-gradient et sur le calcul sous-différentiel que l'on a mis en relation avec la notion (plus faible) de dérivée directionnelle (qui apparaît alors comme la fonction support du sous-différentiel), et celle plus forte de dérivée de Gâteaux (correspondant à l'unicité du sous-gradient en tout point). Les approches interne et externe des convexes correspondent, pour l'épigraphe des fonctions convexes, au fait que le graphe est "sous les cordes" et "au-dessus des tangentes".

Une autre propriété caractéristique des fonctions convexes est le fait que le sous-différentiel est un opérateur *monotone*, ce qui, dans le cas le plus simple, correspond à une notion de "coût marginal croissant".

Une autre façon d'aborder la sous-différentiabilité est de s'attaquer au calcul de la fonction support de l'épigraphe. Ceci conduit à introduire naturellement la notion de *transformée de Fenchel* et, de façon non surprenante, il existe un lien très fort entre cette transformée de Fenchel et le calcul de sous-gradients. On a vu que la transformée de Fenchel agit finalement comme une simple inversion lorsque l'on considère son effet sur les sous-différentiels des fonctions transformées.

### 4.10 Corrigé des exercices

**Corrigé de l'Exercice 4.17** Si la fonction affine supposée minorer  $f$  et  $g$  est  $x \mapsto \langle r, x \rangle + b$ , alors  $f(y) \geq \langle r, y \rangle + b$  et  $g(x-y) \geq \langle r, x-y \rangle + b$  dans (4.6), ce qui montre que  $h(x) \geq \langle r, x \rangle + 2b > -\infty$ .

Si  $(x, \xi) \in \mathbb{U}_s(h)$ , donc  $\xi > \inf_{y \in \mathcal{X}} (f(y) + g(y-x))$ , alors il existe  $y_0$  tel que  $\xi > f(y_0) + g(y_0-x)$ , et donc  $\xi$  peut s'écrire  $\eta + \zeta$  avec  $\eta > f(y_0)$  et  $\zeta > g(x-y_0)$ . On a donc exprimé  $(x, \xi)$  comme la somme de  $(y_0, \eta) \in \mathbb{U}_s(f)$  et  $(x-y_0, \zeta) \in \mathbb{U}_s(g)$ . Par conséquent,  $\mathbb{U}_s(h) \subset \mathbb{U}_s(f) + \mathbb{U}_s(g)$ .

Réciproquement, si  $x = x_1 + x_2$  et  $\xi = \xi_1 + \xi_2$  avec  $\xi_1 > f(x_1)$  et  $\xi_2 > f(x_2)$ , alors

$$\xi > f(x_1) + g(x_2) > \inf_{y_1+y_2=x_1+x_2} (f(y_1) + g(y_2)) = h(x_1 + x_2) ,$$

donc  $(x_1 + x_2, \xi_1 + \xi_2) \in \mathbb{U}_s(h)$  et on a montré l'inclusion inverse.

Si  $f$  et  $g$  sont convexes, leur épigraphe strict est convexe et la somme vectorielle préserve la convexité, donc  $h$  est convexe.

**Corrigé de l'Exercice 4.18** Dire que  $(x, \beta) \in \mathbb{U}_s(g)$ , c'est dire que  $\beta > \inf_y f(x, y)$  ou bien qu'il existe  $y$  tel que  $\beta > f(x, y)$ , ou bien encore qu'il existe  $y$  tel que  $((x, y), \beta) \in \mathbb{U}_s(f)$ . Cette dernière affirmation est équivalente à dire que  $(x, \beta)$  appartient à la projection de  $\mathbb{U}_s(f)$  sur  $\mathcal{X}$ .

Comme on a vu que la projection d'un convexe est convexe (voir Figure 3.4), ceci montre que  $\mathbb{U}_s(g)$  est convexe, donc  $g$  est une fonction convexe.

**Corrigé de l'Exercice 4.60** On ne donne que les démonstrations du cas "fortement". Le cas "strictement" correspond à une adaptation immédiate. On suppose (4.3) que l'on réécrit

$$\begin{aligned} (1-\alpha)(f(y) - f(x)) &\geq f(\alpha x + (1-\alpha)y) - f(x) + b\alpha(1-\alpha) \|x-y\|^2 / 2 \\ &\geq (1-\alpha) \langle r, x \rangle + b\alpha(1-\alpha) \|x-y\|^2 / 2 \end{aligned}$$

pour tout  $r \in \partial f(x)$ . Il suffit de diviser cette inégalité par  $\alpha$  puis de prendre  $\alpha = 1$  pour obtenir (4.20). Réciproquement, on suit de près le schéma de la démonstration du Théorème 4.59, mais en partant de (4.20) au lieu de (4.18), ce qui conduit après calcul à (4.3).

**Corrigé de l'Exercice 4.69** Supposons la fonction marginale  $g$  sous-différentiable en  $x$ . Pour tout  $r \in \partial g(x)$  et tout  $x'$ ,

$$g(x') - g(x) \geq \langle r, x' - x \rangle .$$

Si  $y \in Y(x)$ , alors, par définition,  $g(x) = f(x, y)$  et par ailleurs (condition d'optimalité),  $0 \in \partial_y f(x, y)$ . Moyennant ces observations et celle que pour tout  $y'$ ,  $f(x', y') \geq g(x')$ , l'inégalité ci-dessus implique que :

$$f(x', y') - f(x, y) \geq \langle r, x' - x \rangle + \langle 0, y' - y \rangle .$$

Cette inégalité étant vraie pour tout  $(x', y')$ , elle montre que  $(r, 0) \in \partial f(x, y)$  et donc que  $r \in \Pi_{\mathcal{X}}(\partial f(x, y)) = \partial_x f(x, y)$ . Cette relation d'appartenance étant vraie pour tout  $y \in Y(x)$ , l'inclusion (4.29) se trouve justifiée.



## Chapitre 5

# Optimisation sous contraintes : conditions d'optimalité locales

Dans la première section de ce chapitre, on traite du problème général de la minimisation d'une fonction convexe sur un sous-ensemble "admissible" convexe. L'introduction d'un sous-ensemble admissible est une façon géométrique ou implicite de formuler des contraintes pesant sur le problème de décision optimale. Nous envisagerons dans les sections suivantes une autre façon plus explicite ou analytique d'introduire des contraintes sous la forme d'égalités ou d'inégalités à respecter.

### 5.1 Optimisation sur un ensemble admissible et inéquations variationnelles

Les conditions nécessaires et suffisantes obtenues ici se présentent sous forme d'inéquations variationnelles dont on examinera la signification géométrique. En général, les inéquations variationnelles ne s'interprètent pas forcément comme les conditions d'optimalité d'un problème d'optimisation (sous contraintes) ; elles peuvent apparaître dans les problèmes d'équilibre (en Mécanique, dans les réseaux de transport . . . ) ou en théorie des jeux.

#### 5.1.1 Formulation du problème et existence d'une solution

On va donc principalement dans ce chapitre traiter du problème générique suivant

$$\min_{u \in U^{\text{ad}}} J(u) , \quad (5.1)$$

ce qui signifie "chercher la valeur la plus petite possible de la fonction  $J$  et les décisions  $u$  qui la réalisent lorsque  $u$  est astreint à rester dans  $U^{\text{ad}}$ ". Ici  $U^{\text{ad}}$  est un sous-ensemble convexe fermé d'un espace de Hilbert  $\mathcal{U}$  (dit "ensemble admissible") et  $J$  est une fonction convexe s.c.i. de  $\mathcal{U}$  dans  $\overline{\mathbb{R}}$ . Un tel problème est alors qualifié de "problème d'optimisation convexe". On appellera "solution" tout  $u$  où  $J(u)$  atteint effectivement sa valeur infimale sur  $U^{\text{ad}}$ . L'ensemble des solutions est noté  $\arg \min_{u \in U^{\text{ad}}} J(u)$ .

Observons tout de suite que l'on pourrait d'emblée se limiter à des problèmes de minimisation où  $U^{\text{ad}}$  est toujours égal à l'espace  $\mathcal{U}$  tout entier dans la mesure où on s'est autorisé des fonctions  $J$  prenant éventuellement la valeur  $+\infty$ . En effet, le problème (5.1) est équivalent au problème

$$\min_{u \in \mathcal{U}} (J(u) + I_{U^{\text{ad}}}(u)) , \quad (5.2)$$

puisque la fonction indicatrice de l'ensemble  $U^{\text{ad}}$  vaut  $+\infty$  dès que  $u$  sort de  $U^{\text{ad}}$  — ce qui évidemment n'est pas le but recherché pour *minimiser* la fonction coût — et qu'elle est neutre (vaut uniformément 0)

lorsque  $u$  reste dans  $U^{\text{ad}}$  — ce qui laisse alors la fonction  $J$  seule juge du meilleur  $u$ . Il est clair que cette façon de voir les contraintes est plutôt un outil conceptuel que pratique : nous nous en servons dans les cas appropriés.

Une première conséquence importante de la convexité est que toute solution *locale* (on dit aussi “minimum local”) est une solution *globale* (“minimum global”). Par “solution locale” ou “minimum local”, on entend un point  $u^\circ \in U^{\text{ad}}$  tel que

$$J(u^\circ) \leq J(u), \quad \forall u \in U^{\text{ad}} \cap \overline{B}(u^\circ, \rho),$$

pour un certain  $\rho > 0$ . Autrement dit, l’ensemble admissible est réduit à un voisinage du prétendu minimum local. La Figure 5.1 illustre cette notion (dans un cas non convexe obligatoirement).

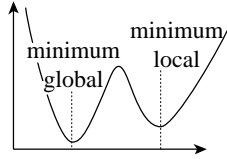


Figure 5.1: Minimum local et minimum global

**Lemme 5.1.** *Pour un problème convexe, toute solution locale est globale.*

*Démonstration.* Soit  $u \in U^{\text{ad}} \setminus \overline{B}(u^\circ, \rho)$ ,  $\alpha = \rho / \|u - u^\circ\| < 1$  et  $u_\alpha = (1 - \alpha)u^\circ + \alpha u$ . On a  $u_\alpha - u^\circ = \alpha(u - u^\circ)$  et ce vecteur est de norme égale à  $\rho$ , donc  $u_\alpha \in \overline{B}(u^\circ, \rho)$ , et par ailleurs, évidemment  $u_\alpha \in U^{\text{ad}}$ . Alors,

$$J(u^\circ) \leq J(u_\alpha) \leq (1 - \alpha)J(u^\circ) + \alpha J(u) \Rightarrow J(u^\circ) \leq J(u).$$

Ceci étant vrai pour  $u \in U^{\text{ad}} \setminus \overline{B}(u^\circ, \rho)$ , mais également pour  $u \in U^{\text{ad}} \cap \overline{B}(u^\circ, \rho)$  (par définition de  $u^\circ$ ) est donc vrai pour  $u \in U^{\text{ad}}$ . Ceci montre que  $u^\circ$  est bien une solution de (5.1).  $\square$

Le théorème suivant traite du problème de l’existence d’une solution du problème (5.1). Auparavant, on introduit un peu de terminologie.

**Définition 5.2.** Une fonction  $J$  est *coercive sur*  $U^{\text{ad}}$  si

$$\lim J(u) = +\infty \text{ lorsque } \|u\| \rightarrow +\infty \text{ et } u \in U^{\text{ad}}.$$

Autrement dit, dans toutes les directions de  $U^{\text{ad}}$  où  $u$  peut tendre vers l’infini,  $J(u)$  tend vers  $+\infty$ . Évidemment, si  $U^{\text{ad}}$  est borné, donc s’il n’existe aucune direction dans  $U^{\text{ad}}$  où  $u$  peut aller à l’infini, la propriété est réputée vraie. Observons que si  $J$  est *fortement convexe*, alors elle est coercive sur tout sous-ensemble  $U^{\text{ad}}$  d’après (4.20). Observons aussi qu’il est équivalent de dire que  $J$  est coercive sur  $U^{\text{ad}}$  ou que  $J + I_{U^{\text{ad}}}$  est coercive sur  $\mathcal{U}$ .

**Théorème 5.3.** *Si  $J$  est une fonction convexe s.c.i. et coercive sur  $U^{\text{ad}}$ , si  $U^{\text{ad}}$  est convexe et fermé, si*

$$\nabla(J) \cap U^{\text{ad}} \neq \emptyset, \tag{5.3}$$

*alors il existe au moins une solution au problème (5.1). L’ensemble des solutions est un sous-ensemble convexe fermé. Il est réduit à un point (la solution est donc unique) si  $J$  est strictement convexe.*



*Démonstration.* L'hypothèse (5.3) vise à écarter le cas trivial où la valeur de l'inf $_{u \in U^{\text{ad}}} J(u)$  serait égal à  $+\infty$  car, sous cette hypothèse, il existe au moins un  $u \in U^{\text{ad}}$  où  $J(u)$  est finie.

On considère alors une suite minimisante  $\{u^k\}_{k \in \mathbb{N}} \subset U^{\text{ad}}$ , c'est-à-dire telle que  $\lim_{k \rightarrow +\infty} J(u^k) = \inf_{u \in U^{\text{ad}}} J(u)$  (cette valeur peut a priori être égale à  $-\infty$ , ce qui ne gêne pas le raisonnement à venir, mais on va démontrer que ce n'est pas le cas finalement). On peut affirmer que cette suite est bornée car sinon, avec la coercivité de  $J$  sur  $U^{\text{ad}}$ , on obtiendrait une contradiction. Elle est donc compacte dans la topologie faible, et on peut extraire une sous-suite  $\{u^{k_i}\}$  faiblement convergente vers un certain  $u^\sharp$ , point d'accumulation de la suite, et, du fait que  $U^{\text{ad}}$  est un convexe fermé, donc faiblement fermé (voir Corollaire 3.56), alors  $u^\sharp \in U^{\text{ad}}$ . Du fait qu'une fonction convexe s.c.i. est aussi faiblement s.c.i. (voir Corollaire 4.31), on a

$$\inf_{u \in U^{\text{ad}}} J(u) = \lim_{k_i \rightarrow +\infty} J(u^{k_i}) \geq J(u^\sharp),$$

ce qui montre que l'égalité a nécessairement lieu et que  $u^\sharp$  est une solution de (5.1) (en particulier  $J(u^\sharp) > -\infty$  car nous avons implicitement exclu les fonctions prenant ces valeurs — a priori, on peut cependant s'approcher de  $-\infty$  en envoyant  $u$  vers l'infini, mais ici c'est exclu par l'hypothèse de coercivité).

Ayant établi la finitude de l'inf et l'existence d'au moins une solution  $u^\sharp$ , toutes les autres solutions peuvent être obtenues comme le sous-ensemble

$$U^\sharp \stackrel{\text{def}}{=} \overline{\nabla_{J(u^\sharp)}(J)} \cap U^{\text{ad}}.$$

On sait d'après le Théorème 4.30 que les ensembles de niveau de  $J$  sont convexes et fermés et comme c'est aussi le cas de  $U^{\text{ad}}$ , c'est aussi le cas de  $U^\sharp$ .

Enfin, si  $J$  est *strictement* convexe et si on suppose que l'on a deux solutions distinctes  $u_1^\sharp$  et  $u_2^\sharp$  (évidemment dans  $U^{\text{ad}}$ ), alors le point  $(u_1^\sharp + u_2^\sharp)/2$  (également dans  $U^{\text{ad}}$  par convexité de  $U^{\text{ad}}$ ) donne nécessairement à  $J$  une valeur strictement inférieure à  $(J(u_1^\sharp) + J(u_2^\sharp))/2 = J(u_1^\sharp) = J(u_2^\sharp)$  par convexité stricte (voir Définition 4.7), ce qui est une contradiction.  $\square$

### 5.1.2 Caractérisation des solutions

Une fois réglé le problème de l'existence de solutions, il faut pouvoir *caractériser* les solutions du problème (5.1). Le théorème qui suit propose un certain nombre de critères d'optimalité.

**Théorème 5.4.** *En plus des hypothèses de convexité, de fermeture de  $\mathbb{U}(J)$  et  $U^{\text{ad}}$ , on fait l'hypothèse que  $J$  est bornée dans un ouvert (alors son domaine est d'intérieur non vide, et  $J$  y est continue et sous-différentiable<sup>1</sup>) ainsi que l'hypothèse suivante :*

$$\overline{\mathbb{U}(J)} \cap U^{\text{ad}} \neq \emptyset. \quad (5.4)$$

Alors, chacune des conditions ci-dessous constitue une condition nécessaire et suffisante pour que  $u^\sharp$  soit une solution du problème (5.1), en plus de la condition évidemment nécessaire que  $u^\sharp \in U^{\text{ad}}$  :

$$\forall u \in U^{\text{ad}}, \quad DJ(u^\sharp; u - u^\sharp) \geq 0, \quad (5.5a)$$

$$0 \in \partial(J + I_{U^{\text{ad}}})(u^\sharp), \quad (5.5b)$$

$$\exists r^\sharp \in \partial J(u^\sharp) \cap (-(U^{\text{ad}})_{u^\sharp}^\perp), \quad (5.5c)$$

$$\exists r^\sharp \in \partial J(u^\sharp) : \forall u \in U^{\text{ad}}, \quad \langle r^\sharp, u - u^\sharp \rangle \geq 0. \quad (5.5d)$$

<sup>1</sup>Voir Théorèmes 4.27 et 4.53

*Démonstration.* Si  $u^\sharp$  est solution,  $J(u^\sharp) \leq J((1 - \varepsilon)u^\sharp + \varepsilon u)$  pour tout  $u \in U^{\text{ad}}$  et tout  $\varepsilon \in [0, 1]$ . Comme

$$DJ(u^\sharp; u - u^\sharp) = \lim_{\varepsilon \rightarrow 0} \left( J(u^\sharp + \varepsilon(u - u^\sharp)) - J(u^\sharp) \right) / \varepsilon ,$$

il est clair que cette limite est non négative. Réciproquement, comme on peut remplacer la limite ci-dessus par l'inf $_{\varepsilon > 0}$ , alors, en particulier, avec  $\varepsilon = 1$ , on a que  $DJ(u^\sharp; u - u^\sharp) \leq J(u) - J(u^\sharp)$ . Donc si  $DJ(u^\sharp; u - u^\sharp) \geq 0$ , c'est aussi le cas de  $J(u) - J(u^\sharp)$ , ce qui prouve que  $u^\sharp$  est solution.

Si  $u^\sharp$  est solution de (5.1), il est aussi solution de (5.2), et en fonction de ce qui vient d'être dit,  $D(J + I_{U^{\text{ad}}})(u^\sharp; u - u^\sharp) \geq 0$  pour tout  $u \in \mathcal{U}$ . Grâce à l'hypothèse (5.4), on sait qu'il existe un point dans l'intersection de l'intérieur du domaine de  $J$  et du domaine de  $I_{U^{\text{ad}}}$  où  $J$  est continue. Par conséquent  $G = J + I_{U^{\text{ad}}}$  est sous-différentiable avec  $\partial G(\cdot) = \partial J(\cdot) + \partial I_{U^{\text{ad}}}(\cdot)$  (voir les commentaires après (4.24)). Par ailleurs,  $DG(u^\sharp; \cdot)$  est la fonction support de  $\partial G(u^\sharp)$  (Corollaire 4.56). Par conséquent, on peut calculer la fonction indicatrice de  $\partial G(u^\sharp)$  en prenant sa transformée de Fenchel, puisque

- le sous-différentiel est convexe fermé (donc sa fonction indicatrice est convexe s.c.i.),
- que la fonction support est la transformée de Fenchel de la fonction indicatrice (Corollaire 4.21),
- et que la transformée de Fenchel est involutive pour les fonctions convexes s.c.i. (Corollaire 4.39).

On a donc

$$I_{\partial G(u^\sharp)}(u) = \sup_{p \in \mathcal{U}} (\langle u, p \rangle - DG(u^\sharp; p)) \Rightarrow I_{\partial G(u^\sharp)}(0) = \sup_{p \in \mathcal{U}} (-DG(u^\sharp; p)) .$$

Or, comme  $DG(u^\sharp; p) \geq 0$  pour tout  $p \in \mathcal{U}$ , on en déduit que  $I_{\partial G(u^\sharp)}(0) \leq 0$  et donc il est nul (une fonction indicatrice ne prend que les valeurs 0 et  $+\infty$ ). Finalement, on en déduit que  $0 \in \partial G(u^\sharp) = \partial(J + I_{U^{\text{ad}}})(u^\sharp)$ . Réciproquement, par l'inégalité de la convexité, il est immédiat que si  $0 \in \partial(J + I_{U^{\text{ad}}})(u^\sharp)$ , alors pour tout  $u \in \mathcal{U}$ ,

$$(J + I_{U^{\text{ad}}})(u) - (J + I_{U^{\text{ad}}})(u^\sharp) \geq \langle 0, u - u^\sharp \rangle = 0 ,$$

ce qui montre que  $u^\sharp$  est solution de (5.2) et donc aussi de (5.1).

Puisque  $\partial(J + I_{U^{\text{ad}}})(u^\sharp) = \partial J(u^\sharp) + \partial I_{U^{\text{ad}}}(u^\sharp)$ , dire que  $0 \in \partial(J + I_{U^{\text{ad}}})(u^\sharp)$  est équivalent à dire qu'il existe  $r^\sharp \in \partial J(u^\sharp)$  et  $s^\sharp \in \partial I_{U^{\text{ad}}}(u^\sharp)$  tels que  $r^\sharp + s^\sharp = 0$ , autrement dit,  $r^\sharp \in -\partial I_{U^{\text{ad}}}(u^\sharp)$ , mais  $I_{U^{\text{ad}}}(u^\sharp) = (U^{\text{ad}})_{u^\sharp}^\perp$  (voir Exercice 4.52), d'où l'on déduit (5.5c).

Enfin, dire que  $r^\sharp \in \partial J(u^\sharp)$  appartient aussi à  $-(U^{\text{ad}})_{u^\sharp}^\perp$ , c'est exactement énoncer la condition (5.5d) en fonction de la Définition 3.20 du cône orthogonal.  $\square$

**Remarque 5.5.** La forme (5.5d) est appelée “*inéquation variationnelle*”. Certaines inéquations variationnelles caractérisant des équilibres dans les problèmes de jeux, dans les réseaux de transport, en Physique (notamment en Mécanique), entre autres domaines, sont de la forme (5.5d), mais font intervenir, à la place de l'opérateur  $\partial J$ , un opérateur (éventuellement multivoque et monotone) qui ne s'interprète pas comme le sous-différentiel d'une fonction (on dit qu'il “ne dérive pas d'un potentiel”).

**Remarque 5.6.** Lorsque l'ensemble admissible  $U^{\text{ad}}$  est tout l'espace, ou si  $u^\sharp \in \overset{\vee}{U^{\text{ad}}}$ , alors  $(U^{\text{ad}})_{u^\sharp}^\perp = \{0\}$  et les conditions (5.5b)–(5.5d) se réduisent toutes les trois à

$$0 \in \partial J(u^\sharp) .$$

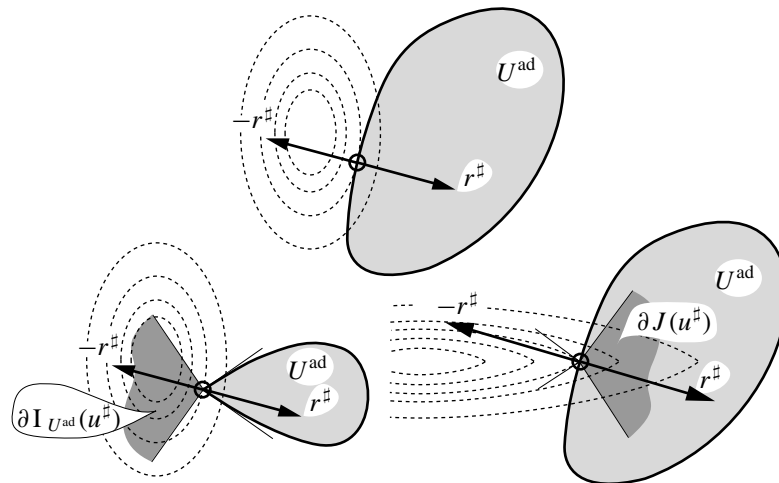


Figure 5.2: Interprétation géométrique des conditions d'optimalité

### 5.1.3 Interprétation géométrique

C'est la forme (5.5c) qui se prête sans doute le mieux à l'interprétation géométrique. La Figure 5.2 représente trois situations (nous n'avons pas osé montrer la quatrième !) sur lesquelles on peut essayer de comprendre ce que (5.5c) signifie. Quelques ensembles de niveau de la fonction  $J$  y sont délimités en pointillés tandis que le sous-ensemble  $U^{\text{ad}}$  est en gris clair. Évidemment, plus on va vers des ensembles de niveau proches du "centre", plus la fonction  $J$  diminue, ce qui va dans le sens souhaité par la minimisation. Mais on doit par ailleurs rester dans l'ensemble  $U^{\text{ad}}$ . Le point optimal est marqué par un petit cercle. Dans la partie supérieure de la figure, les frontières des ensembles de niveau sont "lisses" (ce qui correspond à une fonction  $J$  différentiable, donc  $r^\# = J'(u^\#)$  dans ce cas), ainsi que la frontière de  $U^{\text{ad}}$ , ce qui signifie que le cône normal en tout point de cette frontière est réduit à une demi-droite. L'ensemble au niveau  $J(u^\#)$  est dans ce cas tangent à l'ensemble  $U^{\text{ad}}$  au point  $u^\#$  et  $r^\#$  est orthogonal à la tangente commune à ces deux ensembles. L'orientation est telle que  $-r^\#$  sort de  $U^{\text{ad}}$ , ce qui est interdit aux déplacements admissibles. La tangente commune aux deux sous-ensembles ( $U^{\text{ad}}$  et  $\nabla J(u^\#)$  ( $J$ )) sépare le plan en deux demi-plans, l'un dans lequel la fonction  $J$  augmente et qui contient par ailleurs  $U^{\text{ad}}$ , donc les déplacements admissibles, et l'autre dans lequel la fonction  $J$  diminue mais les déplacements vers ce demi-espace sont interdits.

**Exercice 5.7.** Montrer que pour une fonction convexe différentiable, tout déplacement à partir d'un point faisant un angle aigu avec le gradient en ce point fait croître la fonction, et que tout déplacement *suffisamment petit* faisant un angle aigu avec l'opposé du gradient la fait décroître.

Dans ces explications, le mot "orthogonal" a son importance parce que sinon, il existerait une projection non nulle de  $-r^\#$  sur la tangente à  $U^{\text{ad}}$  (et donc par continuité sur une direction légèrement rentrante dans  $U^{\text{ad}}$ ), ce qui permettrait d'effectuer encore un déplacement restant dans  $U^{\text{ad}}$  mais améliorant  $J(u^\#)$  : alors  $u^\#$  ne serait pas optimal.

Le schéma en bas à gauche représente encore une situation où  $J$  est différentiable mais par contre la frontière de  $U^{\text{ad}}$  au point  $u^\#$  présente un point anguleux. Dans ce cas, le cône normal à  $U^{\text{ad}}$  au point  $u^\#$  (représenté en gris foncé sur la figure) n'est plus réduit à une demi-droite. Comme  $-r^\#$  appartient à ce cône, toutes les directions du cône tangent à  $U^{\text{ad}}$  au point  $u^\#$  (les seules correspondant à des déplacements admissibles) forment au moins un angle droit (donc en général un angle obtus) avec  $-r^\#$ , ce qui interdit à nouveau d'améliorer  $J$  par un déplacement admissible.

Enfin le schéma en bas à droite représente une situation où la frontière de  $U^{\text{ad}}$  est à nouveau lisse (le cône normal se réduit donc à une demi-droite) mais c'est la fonction  $J$  qui n'est pas différentiable en  $u^\#$

(donc le cône normal à l'ensemble de niveau auquel appartient nécessairement  $r^\sharp$  — voir Lemme 4.50 — n'est pas réduit à une seule direction ; il est représenté en gris foncé sur la figure). Comme dans le premier des trois cas considérés, les déplacements admissibles constituent un demi-espace, mais le cône des déplacements qui font diminuer  $J$  est situé entièrement dans l'autre demi-espace.

**Remarque 5.8.** Contrairement au cas différentiable, on observera dans ce cas que la direction  $-r^\sharp$  n'est pas nécessairement une direction de diminution de  $J$  même pour un déplacement infinitésimal, et même si c'est le cas qui est représenté sur la figure. En effet, les directions de diminution de  $J$  sont représentées par le cône tangent à l'ensemble de niveau  $J(u^\sharp)$  et il est possible (voir Figure 5.3) que ce cône tangent soit strictement contenu dans l'opposé du cône normal (cône de tous les vecteurs  $-r^\sharp$ ). La différence ensembliste  $\left(-(\nabla_{J(u^\sharp)}(J))_{u^\sharp}^\perp\right) \setminus (\nabla_{J(u^\sharp)}(J))_{u^\sharp}^\top$  contient tous les vecteurs de  $-\partial J(u^\sharp)$  qui ne sont pas des directions de descente de  $J$  au point  $u^\sharp$ . Par contre, on a la propriété suivante (mais qui n'est pas très utile pour *minimiser*  $J$ ).

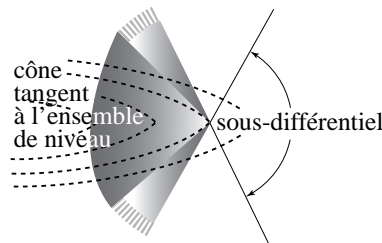


Figure 5.3: Les directions (striées) qui sont dans  $-(\nabla_{J(u)}(J))_u^\perp$  mais pas dans  $(\nabla_{J(u)}(J))_u^\top$  ne sont pas des directions de descente au point  $u$

**Exercice 5.9.** Montrer que pour une fonction convexe, un déplacement à partir d'un point d'amplitude quelconque dans la direction d'un sous-gradient quelconque en ce point fait augmenter la fonction.

**Remarque 5.10.** Il est intéressant de revenir sur la question de la projection sur un convexe fermé (voir §3.8.1) qui est un problème d'optimisation particulier (voir (3.13)) et de constater que (3.14) n'est rien d'autre que l'application de (5.5d) à cette situation. On retrouve dans l'interprétation géométrique le fait que le vecteur  $x - \Pi_A(x)$ , qui est l'opposé du gradient de la fonction  $\|\cdot - x\|^2/2$  à minimiser, gradient calculé à la solution  $\Pi_A(x)$ , appartient au cône normal au sous-ensemble  $A$  en ce point.

### 5.1.4 Autre caractérisation des solutions optimales

Revenons au Théorème 5.4. La caractérisation (5.5a) dit essentiellement que lorsque les fonctions sont convexes, une variation *locale* (ou “au premier ordre”) non négative est équivalente à une variation *globale* non négative. Il est possible, lorsque la fonction  $J$  est vue comme la somme de deux fonctions *convexes* (cette hypothèse est fondamentale !) de “mélanger” dans les conditions d'optimalité des variations locales et globales. Ceci résulte du lemme suivant.

**Lemme 5.11.** *Sous les hypothèses du Théorème 5.4 et si  $J$  est égale à la somme de deux fonctions convexes  $J_1$  et  $J_2$ , la condition (5.5a) est équivalente à la condition*

$$\forall u \in U^{\text{ad}}, \quad DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp) \geq 0. \quad (5.6)$$

*Démonstration.* On a que

$$\begin{aligned} D(J_1 + J_2)(u^\sharp; u - u^\sharp) &= \inf_{\varepsilon > 0} \left( (J_1 + J_2)(u^\sharp + \varepsilon(u - u^\sharp)) - (J_1 + J_2)(u^\sharp) \right) / \varepsilon \\ &\leq \inf_{\varepsilon > 0} \left( J_1(u^\sharp + \varepsilon(u - u^\sharp)) - J_1(u^\sharp) + (1 - \varepsilon)J_2(u^\sharp) + \varepsilon J_2(u) - J_2(u^\sharp) \right) / \varepsilon \end{aligned}$$

par convexité de  $J_2$

$$\begin{aligned} &= \inf_{\varepsilon > 0} \left( J_1(u^\sharp + \varepsilon(u - u^\sharp)) - J_1(u^\sharp) + \varepsilon(J_2(u) - J_2(u^\sharp)) \right) / \varepsilon \\ &= DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp), \end{aligned}$$

et donc la non négativité du premier membre entraîne celle du second membre.

Réciproquement, puisque  $DJ_1(u^\sharp; u - u^\sharp) \leq J_1(u) - J_1(u^\sharp)$ , alors  $DJ_1(u^\sharp; u - u^\sharp) + J_2(u) - J_2(u^\sharp) \leq (J_1 + J_2)(u) - (J_1 + J_2)(u^\sharp)$ , et la non négativité du premier membre entraîne celle du second. Mais avec le Théorème 5.4, cette dernière équivaut à (5.5a).  $\square$

La traduction de la condition (5.6) en la condition suivante (analogue de (5.5d))

$$\exists r_1^\sharp \in \partial J_1(u^\sharp) : \forall u \in U^{\text{ad}}, \left\langle r_1^\sharp, u - u^\sharp \right\rangle + J_2(u) - J_2(u^\sharp) \geq 0 \quad (5.7)$$

nécessite une hypothèse technique permettant d'affirmer que  $\partial(J_1 + J_2)(u^\sharp) = \partial J_1(u^\sharp) + \partial J_2(u^\sharp)$  (voir commentaire après (4.24)).

### 5.1.5 Pénalisation exacte

La considération de la fonction  $J + I_{U^{\text{ad}}}$  a été expliquée comme une "pénalisation" de la contrainte  $u \in U^{\text{ad}}$ , pénalisation qui vaut 0 si  $u$  reste "admissible" mais qui devient infinie sinon. En fait, une pénalisation finie mais "suffisante" peut jouer le même rôle dans la plupart des cas. On va introduire à cet effet une nouvelle notion.

**Définition 5.12.** On considère un sous-ensemble convexe fermé  $A$  d'un espace de Hilbert  $\mathcal{X}$ . On appelle *distance d'un point  $x$  à  $A$* , notée  $\delta_A(x)$ , l'expression

$$\delta_A(x) = \inf_{y \in A} \|y - x\|.$$

En fait, d'après (3.13),  $\delta_A(x)$  n'est rien d'autre que  $\|\Pi_A(x) - x\|$ . Évidemment, cette fonction vaut 0 sur  $A$  et elle est strictement positive à l'extérieur de  $A$ . Le lemme suivant énumère un certain nombre de propriétés utiles.

**Lemme 5.13.** *La fonction  $x \mapsto \delta_A(x)$  est convexe. Elle est différentiable en tout point  $x \notin A$  avec  $\delta'_A(x) = (x - \Pi_A(x)) / \|x - \Pi_A(x)\|$ , ainsi évidemment que dans  $\overset{\circ}{A}$  où elle vaut identiquement 0. Elle est sous-différentiable en tout point  $x$  de la frontière de  $A$  et alors,  $\partial\delta_A(x) = A_x^\perp \cap \overline{B}(0, 1)$ .*

*Démonstration.* La convexité résulte du fait que  $x \mapsto \delta_A(x)$  est la fonction marginale (voir (4.7)) d'une fonction (conjointement) convexe  $(x, y) \mapsto \|y - x\| + I_A(y)$ . Cette dernière fonction est convexe parce qu'elle est la somme de deux fonctions convexes,  $I_A$  d'une part, et  $(x, y) \mapsto \|y - x\|$  d'autre part. Cette dernière fonction est convexe comme la composition de la fonction linéaire  $(x, y) \mapsto y - x$  avec la fonction convexe  $\|\cdot\|$  (montrer que la fonction norme est convexe).

La fonction  $\delta_A$  est égale à  $\sqrt{\delta_A^2}$ . La fonction  $\delta_A^2$ , composée de  $\delta_A$  et de l'élevation au carré, qui est une fonction convexe *et croissante pour des arguments non négatifs*, est donc aussi convexe (grâce à

un résultat énoncé au §4.4). Elle est de plus différentiable, ce qui peut se démontrer directement par le calcul, ou bien en utilisant la formule (4.29) et les observations faites à la suite de l'Exercice 4.69. De plus, la même formule (ou encore une fois le calcul direct) montre que  $(\delta_A^2)'(x) = 2(x - \Pi_A(x))$ . Étant donné que la fonction  $\sqrt{\cdot}$  est différentiable en tout point strictement positif et que sa dérivée au point  $x$  est alors  $1/2\sqrt{x}$ , le théorème de la dérivée des fonctions composées nous dit que la fonction  $\delta_A = \sqrt{\delta_A^2}$  est différentiable dès que  $x \notin A$  (et donc  $x \neq \Pi_A(x)$  et  $\delta_A(x) \neq 0$ ) et que sa dérivée vaut

$$\frac{(\delta_A^2)'(x)}{2\sqrt{\delta_A^2(x)}} = \frac{x - \Pi_A(x)}{\|x - \Pi_A(x)\|}.$$

Il reste à considérer le cas où  $x$  appartient à la frontière de  $A$ . Soit  $r \in \partial\delta_A(x)$ , et dans ce cas  $\delta_A(x) = 0$ . Par définition,

$$\forall y \in \mathcal{X}, \quad \delta_A(y) \geq \langle r, y - x \rangle.$$

Avec  $y \in A$ , on en déduit que  $r \in A_x^\perp$  (voir Définition 3.20). Par ailleurs, avec  $y = x + r$ , on obtient que

$$\|r\|^2 \leq \delta_A(x + r) \leq \|x + r - x\| = \|r\|,$$

ce qui montre que  $\|r\| \leq 1$  et on a donc montré que  $\partial\delta_A(x) \subset A_x^\perp \cap \overline{B}(0, 1)$ . Réciproquement, soit  $r \in A_x^\perp \cap \overline{B}(0, 1)$ . On a que

$$\begin{aligned} \forall y \in \mathcal{X}, \quad \langle r, y - x \rangle &= \langle r, y - \Pi_A(y) \rangle + \langle r, \Pi_A(y) - x \rangle \\ &\leq \langle r, y - \Pi_A(y) \rangle \end{aligned}$$

parce que le second terme est non négatif puisque  $r \in A_x^\perp$ ,

$$\leq \|r\| \|y - \Pi_A(y)\|$$

par l'inégalité de Schwartz,

$$\leq \delta_A(y)$$

parce que  $\|r\| \leq 1$  et que  $\|y - \Pi_A(y)\| = \delta_A(y)$ . Ceci montre que  $r \in \partial\delta_A(x)$ . □

**Exercice 5.14.** Considérer le sous-ensemble convexe fermé  $A = [0, 1] \subset \mathbb{R}$ , tracer le graphe de  $\delta_A$  dans ce cas, et vérifier tous les résultats énoncés dans le Lemme 5.13.

Revenons maintenant au propos principal de cette section qui est la pénalisation exacte, et expliquons d'abord l'idée de cette technique. On a vu que les conditions d'optimalité indiquent qu'il existe un élément de  $\partial J(u^\sharp)$  qui appartient aussi à l'opposé du cône normal à  $U^{\text{ad}}$  au point  $u^\sharp$  (ce cône  $(U^{\text{ad}})_{u^\sharp}^\perp$  est aussi  $\partial I_{U^{\text{ad}}}(u^\sharp)$ ). Or ce cône contient des vecteurs de norme arbitrairement grande, et tous ne sont pas utiles : si  $M$  est une borne de  $\|r\|$  lorsque  $r$  parcourt  $\partial J(u^\sharp)$ , il suffit de se limiter à  $(U^{\text{ad}})_{u^\sharp}^\perp \cap \overline{B}(0, M)$ . Le raisonnement reste vrai pour tout scalaire positif  $\alpha$  remplaçant  $M$  et plus grand que  $M$ . Donc si  $M$  est fini, on va pouvoir se limiter à des  $\alpha$  finis. Or, grâce au Lemme 5.13, on reconnaît dans  $(U^{\text{ad}})_{u^\sharp}^\perp \cap \overline{B}(0, \alpha)$  le sous-différentiel de la fonction  $\alpha\delta_A$ . D'où l'idée de remplacer la fonction  $J + I_{U^{\text{ad}}}$  par la fonction  $J + \alpha\delta_{U^{\text{ad}}}$ , c'est-à-dire de considérer le problème

$$\min_{u \in \mathcal{U}} (J(u) + \alpha\delta_{U^{\text{ad}}}(u)), \tag{5.8a}$$

au lieu et place de (5.1) ou (5.2). Le théorème suivant donne un fondement à cette idée.

**Théorème 5.15.** *Si  $u^\sharp$  est une solution optimale de (5.8a) pour  $\alpha \geq 0$  et si  $u^\sharp \in U^{\text{ad}}$ , alors  $u^\sharp$  est une solution de (5.8a) pour tout  $\alpha' > \alpha$  et c'est aussi une solution de (5.1). Réciproquement, si  $u^\sharp$  est une solution de (5.1) et si  $J$  est bornée au voisinage de  $u^\sharp$ , il existe  $\alpha \geq 0$  tel que  $u^\sharp$  est solution de (5.8a) pour ce  $\alpha$  et tout autre  $\alpha' > \alpha$ .*

*Démonstration.* Si  $u^\sharp$  est solution de (5.8a) pour un certain  $\alpha$ , on a que

$$\forall u \in \mathcal{U}, \quad J(u^\sharp) + \alpha \delta_{U^{\text{ad}}}(u^\sharp) \leq J(u) + \alpha \delta_{U^{\text{ad}}}(u). \quad (5.8b)$$

Si de plus  $u^\sharp \in U^{\text{ad}}$ , alors  $\delta_{U^{\text{ad}}}(u^\sharp) = 0$ , donc le premier membre de (5.8b) est égal à  $J(u^\sharp)$  quel que soit  $\alpha$ , et le second membre est croissant avec  $\alpha$  car  $\delta_A$  est à valeurs positives. Ceci montre que  $u^\sharp$  est aussi solution pour tout  $\alpha' > \alpha$ . Par ailleurs, pour tout  $u \in U^{\text{ad}}$ , donc tel que  $\delta_{U^{\text{ad}}}(u) = 0$ , (5.8b) implique que  $J(u^\sharp) \leq J(u)$  et donc  $u^\sharp$  est bien une solution de (5.1).

Réciproquement, soit  $u^\sharp$  une solution de (5.1) (donc nécessairement  $u^\sharp \in U^{\text{ad}}$ ). L'hypothèse que  $J$  est bornée au voisinage de  $u^\sharp$  permet non seulement d'affirmer que  $J$  est continue en  $u^\sharp$  (Lemme 4.26) mais qu'elle est localement Lipschitzienne en ce point (voir commentaires à la fin de §4.7.1). Soit  $\alpha$  une constante de Lipschitz locale, c'est-à-dire valable dans une boule  $B(u^\sharp, \rho)$  pour  $\rho > 0$ . Soit  $u \in B(u^\sharp, \rho)$ , alors, puisque la projection n'augmente pas les distances (voir (3.15)),

$$\|\Pi_{U^{\text{ad}}}(u) - \Pi_{U^{\text{ad}}}(u^\sharp)\| = \|\Pi_{U^{\text{ad}}}(u) - u^\sharp\| \leq \|u - u^\sharp\| < \rho,$$

et donc  $\Pi_{U^{\text{ad}}}(u) \in B(u^\sharp, \rho)$ . Par conséquent, on a que

$$\begin{aligned} |J(\Pi_{U^{\text{ad}}}(u)) - J(u)| &\leq \alpha \|\Pi_{U^{\text{ad}}}(u) - u\| \\ \Rightarrow J(\Pi_{U^{\text{ad}}}(u)) &\leq J(u) + \alpha \|\Pi_{U^{\text{ad}}}(u) - u\| \\ &= J(u) + \alpha \delta_{U^{\text{ad}}}(u). \end{aligned}$$

Et comme  $\Pi_{U^{\text{ad}}}(u) \in U^{\text{ad}}$ , on a évidemment que  $J(u^\sharp) \leq J(\Pi_{U^{\text{ad}}}(u))$ , ce qui montre que

$$J(u^\sharp) = J(u^\sharp) + \alpha \delta_{U^{\text{ad}}}(u^\sharp) \leq J(u) + \alpha \delta_{U^{\text{ad}}}(u), \quad \forall u \in B(u^\sharp, \rho).$$

Donc  $u^\sharp$  est un minimum local de  $J + \alpha \delta_{U^{\text{ad}}}$  et c'est donc aussi un minimum global sur tout  $\mathcal{U}$  (Lemme 5.1).  $\square$

La condition que  $J$  soit bornée au voisinage de  $u^\sharp$  est essentielle pour assurer le caractère localement Lipschitzien. Considérer par exemple le cas de la fonction introduite à la Remarque 4.46 et supposer que l'on veuille la minimiser sur l'ensemble admissible  $[1, +\infty[$ . La solution est 1 mais on vérifiera que le Théorème 5.15 ne s'applique pas (il est vrai que dans cet exemple, même l'hypothèse (5.4) n'est pas vérifiée).

La technique de pénalisation exacte permet donc de remplacer un problème sous contrainte par un problème sans contrainte, et ce sans avoir recours à la fonction indicatrice de  $U^{\text{ad}}$  (qui a plutôt un intérêt conceptuel que pratique). Cependant, il faut être capable d'estimer la valeur de  $\alpha$  qui convient ; on peut cependant se tromper "vers le haut", c'est-à-dire que toute valeur supérieure à une valeur qui convient, convient également. De plus, le Théorème 5.15 a donné un test a posteriori de l'adéquation de ces valeurs  $\alpha$  : il suffit que la solution du problème pénalisé obtenue soit admissible pour être optimale.

Cependant, le désagrément de cette technique est que l'on doit manipuler une fonction coût *non différentiable*, ce qui, du point de vue des algorithmes numériques de résolution, représente un inconvénient certain. On peut se demander si ceci n'est pas évitable en considérant par exemple la fonction de pénalisation *différentiable*  $\delta_{U^{\text{ad}}}^2$  à la place de  $\delta_{U^{\text{ad}}}$ . Mais on se convainc facilement que cela ne donne pas une pénalisation "exacte" tant que  $\alpha$  reste fini : il faut que  $\alpha$  tende vers l'infini pour que la fonction de pénalisation force asymptotiquement la solution du problème pénalisé à entrer dans  $U^{\text{ad}}$  (condition sine

qua non pour être aussi une solution de (5.1)). La raison est que la dérivée d'une fonction de pénalisation, quelle qu'elle soit, est nulle vers l'intérieur de  $U^{\text{ad}}$  (où la fonction doit être identiquement nulle), et cette dérivée est donc aussi nulle vers l'extérieur au point  $u^\sharp$  (supposé appartenir ici à la frontière de  $U^{\text{ad}}$ ) si la fonction de pénalisation est différentiable. Dans ces conditions, elle reste nulle pour tout  $\alpha$  fini et elle ne modifie pas localement et au premier ordre la fonction  $J$  (c'est-à-dire en fait son sous-différentiel en  $u^\sharp$ ).

Ces explications intuitives sur le fait qu'une fonction de pénalisation, pour être exacte, doit nécessairement être non différentiable peuvent être justifiées rigoureusement [3].

## 5.2 Optimisation sous contraintes explicites

### 5.2.1 Description des contraintes

Dans cette section, on en vient à une description explicite des contraintes sous la forme de la satisfaction d'un ensemble d'égalités et d'inégalités. L'ensemble admissible  $U^{\text{ad}}$  est alors décrit par :  $u \in U^{\text{ad}} \Leftrightarrow$

$$\Theta_i(u) \leq 0, \quad i = 1, \dots, m, \quad (5.9a)$$

$$\Omega_j(u) = 0, \quad j = 1, \dots, p. \quad (5.9b)$$

Pour que cet ensemble de contraintes définisse un sous-ensemble  $U^{\text{ad}}$  convexe, il suffit que chaque contrainte individuellement définisse un sous-ensemble convexe ( $U^{\text{ad}}$  est alors l'intersection de ces sous-ensembles), et ceci est assuré si

- les fonctions  $\Theta_i : \mathcal{U} \rightarrow \mathbb{R}$  sont convexes, car la contrainte  $i$  de la forme (5.9a) définit l'ensemble de niveau  $\bar{\nabla}_0(\Theta_i)$  qui est alors convexe,
- les fonctions  $\Omega_j : \mathcal{U} \rightarrow \mathbb{R}$  sont affines ; en effet, la contrainte  $j$  de la forme (5.9b) peut être vue comme la réunion de deux contraintes

$$\Omega_j(u) \leq 0 \quad \text{et} \quad -\Omega_j(u) \leq 0 ;$$

si chacune doit correspondre à une fonction convexe, alors  $\Omega_j$  doit être affine.

Sous les hypothèses que la fonction coût  $J$  est convexe, que les fonctions  $\Theta_i$  sont convexes, et que les fonctions  $\Omega_j$  sont affines, on parle de “programmation convexe”.

Quant au caractère fermé de  $U^{\text{ad}}$ , il est assuré si les fonctions  $\Omega_j$  sont continues et si les fonctions  $\Theta_i$  sont au moins s.c.i..

Il faut noter qu'il n'y a pas unicité de la représentation d'un sous-ensemble admissible  $U^{\text{ad}}$  par un jeu de contraintes égalité et inégalité. Par exemple, on peut toujours ramener la description à une unique contrainte scalaire. En effet, posons

$$\Xi(u) \stackrel{\text{def}}{=} \sum_{i=1}^m \max(0, \Theta_i(u)) + \sum_{j=1}^p |\Omega_j(u)|, \quad (5.10a)$$

$$\Psi(u) \stackrel{\text{def}}{=} \max \left( \max_{i=1, \dots, m} \left( \max(0, \Theta_i(u)) \right), \max_{j=1, \dots, p} \left( |\Omega_j(u)| \right) \right), \quad (5.10b)$$

Alors, il est facile de vérifier que l'ensemble  $U^{\text{ad}}$  défini par (5.9) peut être aussi décrit par  $\Xi(u) \leq 0$  ou  $\Psi(u) \leq 0$ , ou même  $\Xi(u) = 0$  ou  $\Psi(u) = 0$  (on observe que sous les hypothèses précédentes,  $\Xi$  et  $\Psi$  sont bien convexes mais pas affines). La question de la “qualification des contraintes” qui sera abordée plus loin est une propriété qui est plutôt liée à une représentation particulière de  $U^{\text{ad}}$  par un jeu de contraintes qu'à une propriété intrinsèque de  $U^{\text{ad}}$ . Par exemple, la contrainte  $\Xi(u) \leq 0$  ou  $\Psi(u) \leq 0$  a peu de chance d'être “qualifiée” en général.



## 5.2.2 Description du cône normal aux contraintes

### 5.2.2.1 Une inclusion entre deux cônes

On a vu au Théorème 5.4 que dans l'écriture des conditions d'optimalité, le cône normal à  $U^{\text{ad}}$  au point  $u^\sharp$  joue un rôle primordial. La question est de savoir si on peut décrire ce cône normal explicitement à partir des fonctions  $\Theta_i$  et  $\Omega_j$  qui ont permis de décrire  $U^{\text{ad}}$  lui-même. Le lemme suivant montre que ce n'est qu'en partie vrai.

**Lemme 5.16.** *On suppose les fonctions  $\Theta_i$  intervenant dans (5.9a) sous-différentiables et on écrit les fonctions  $\Omega_j$  qui sont supposées affines et continues globalement sous la forme  $A(u) = b$  où  $b \in \mathbb{R}^p$  et  $A \in \mathcal{L}(\mathcal{U}, \mathbb{R}^p)$ .<sup>2</sup> Pour  $u \in U^{\text{ad}}$  (c'est-à-dire vérifiant (5.9)), on définit l'ensemble  $\Lambda(u) \subset \mathbb{R}_+^m$  de la façon suivante*

$$\Lambda(u) \stackrel{\text{def}}{=} \{ \lambda \in \mathbb{R}^m \mid \lambda_i \geq 0, \lambda_i \Theta_i(u) = 0, i = 1, \dots, m \} . \quad (5.11)$$

On définit le cône

$$C(u) \stackrel{\text{def}}{=} \sum_{i=1}^m \lambda_i \partial \Theta_i(u) + A^* \mu, \quad \lambda \in \Lambda(u), \quad \mu \in \mathbb{R}^p . \quad (5.12)$$

Alors,  $C(u) \subset (U^{\text{ad}})_u^\perp$ .

*Démonstration.* Soit  $r \in C(u)$ , il faut montrer que  $\langle r, v - u \rangle \leq 0$  pour tout  $v \in U^{\text{ad}}$ , donc satisfaisant (5.9). Le vecteur  $r$  s'écrit

$$r = \sum_{i=1}^m \lambda_i \theta_i + A^* \mu$$

pour  $\theta_i \in \partial \Theta_i(u)$ ,  $\mu \in \mathbb{R}^p$  et  $\lambda \in \Lambda(u)$  défini par (5.11). Alors,

$$\begin{aligned} \langle r, v - u \rangle &= \sum_{i=1}^m \lambda_i \langle \theta_i, v - u \rangle + \langle \mu, A(v - u) \rangle \\ &\leq \sum_{i=1}^m \lambda_i (\Theta_i(v) - \Theta_i(u)) + \langle \mu, A(v) - A(u) \rangle \end{aligned}$$

puisque  $\lambda_i \geq 0$  et  $\theta_i \in \partial \Theta_i(u)$ ,

$$= \sum_{i=1}^m \lambda_i \Theta_i(v)$$

du fait que  $\lambda \in \Lambda(u)$  et que  $A(u) = A(v) = b$  puisque  $u$  et  $v$  sont dans  $U^{\text{ad}}$ ,

$$\leq 0$$

puisqu'à nouveau  $v \in U^{\text{ad}}$ , donc  $\Theta_i(v) \leq 0$  et  $\lambda_i \geq 0$  pour tout  $i$ . □

<sup>2</sup>C'est-à-dire que  $A$  est l'opérateur linéaire constitué de la concaténation des "lignes"  $a_j$  telles que  $\Omega_j(u) = \langle a_j, u \rangle - b_j$ ,  $j = 1, \dots, p$ .

### 5.2.2.2 Les écarts complémentaires

La condition sur les  $\lambda_i$  (qui entrent dans la définition (5.11) de  $\Lambda(u)$ ), et qui peut-être réécrite, toujours en supposant que  $u \in U^{\text{ad}}$ , sous les formes équivalentes suivantes (exercice)

$$\lambda_i \geq 0, \quad \Theta_i(u) \leq 0, \quad \sum_{i=1}^m \lambda_i \Theta_i(u) = 0, \quad (5.13a)$$

$$\Theta_i(u) \leq 0, \quad I(u) \stackrel{\text{def}}{=} \{i \in [1, \dots, m] \mid \Theta_i(u) = 0\}, \quad \lambda_i \geq 0, \quad \lambda_i = 0 \text{ pour } i \notin I(u), \quad (5.13b)$$

est connue sous le nom de *condition d'écarts complémentaires*.<sup>3</sup> L'ensemble  $I(u)$  est l'ensemble des indices des contraintes dites *saturées* ou *actives* au point  $u$ . La condition d'écarts complémentaires indique que si  $\Theta_i(u) < 0$ , alors nécessairement  $\lambda_i = 0$ , et aussi, si  $\lambda_i > 0$ , alors nécessairement  $\Theta_i(u) = 0$ . Elle nous dit que seules les contraintes saturées en  $u$  sont à prendre en compte pour la construction du cône  $C(u)$  censé se substituer (on va se repencher sur ce point ci-après) au cône normal à  $U^{\text{ad}}$  au point  $u$ . Ceci est bien naturel puisque le bord de  $U^{\text{ad}}$  est localement défini en  $u$  par les contraintes saturées (et les contraintes égalité qui sont toujours saturées par définition). Nous verrons ultérieurement une interprétation économique de cette condition d'écarts complémentaires.

### 5.2.2.3 L'égalité des deux cônes n'a pas toujours lieu

Revenons maintenant à la question du lien entre  $C(u)$  et  $(U^{\text{ad}})_u^\perp$ . Le Lemme 5.16 a montré une inclusion ; il s'agit de savoir s'il y a toujours égalité. La réponse est négative comme le montre l'exemple suivant.

**Exemple 5.17.** Considérons dans  $\mathcal{U} = \mathbb{R}^2$  les deux contraintes suivantes :

$$x^2 + y^2 \leq 1 \quad \text{et} \quad -y \leq -1.$$

Le sous-ensemble  $U^{\text{ad}}$  ainsi défini se réduit au point  $u = (0, 1)$ . Par conséquent, le cône normal à  $U^{\text{ad}}$  en ce(t unique) point est tout  $\mathbb{R}^2$ . Cependant, le cône  $C(u)$  défini par (5.12) est constitué par la droite  $x = 0$ . Il n'y a donc pas égalité ici entre  $C(u)$  et  $(U^{\text{ad}})_u^\perp$  mais seulement inclusion.

Or, du fait que  $U^{\text{ad}}$  est réduit à un singleton, n'importe quelle fonction coût  $J$  (disons, différentiable, pour simplifier) conduira à la solution  $u^\sharp = (0, 1)$  du problème (5.1). Évidemment,  $J'(u^\sharp)$  pourra prendre une valeur quelconque, ce qui ne contredit pas la condition d'optimalité (5.5c) puisque le cône normal à  $U^{\text{ad}}$  est tout  $\mathbb{R}^2$ , mais ce qui montre qu'on ne pourra pas toujours affirmer que  $J'(u^\sharp) \in -C(u^\sharp)$ .

Là est donc le problème : dans les cas où  $C(u^\sharp) \neq (U^{\text{ad}})_{u^\sharp}^\perp$ , on ne pourra pas substituer l'un à l'autre cône dans l'écriture des conditions d'optimalité (or, rappelons que c'est  $C(u^\sharp)$  qui a une expression explicite par les fonctions définissant les contraintes).

## 5.2.3 Qualification des contraintes et conditions de Lagrange-Karush-Kuhn-Tucker

Afin de résoudre la difficulté que nous venons d'évoquer, il faut faire l'hypothèse que

$$C(u^\sharp) = (U^{\text{ad}})_{u^\sharp}^\perp, \quad (5.14)$$

que nous appellerons "hypothèse de *qualification des contraintes* dans sa forme minimale". On va montrer que cette hypothèse est toujours vérifiée si toutes les contraintes sont affines.

**Lemme 5.18.** *Si dans (5.9), toutes les fonctions sont affines continues et si elles définissent un sous-ensemble  $U^{\text{ad}} \neq \emptyset$ , alors (5.14) est satisfaite en tout point  $u \in U^{\text{ad}}$ .*

<sup>3</sup>"Complementary slackness" en Anglais.

*Démonstration.* Nous poserons

$$\begin{aligned}\Theta_i(u) &= \langle \theta_i, u \rangle - t_i, \quad \theta_i \in \mathcal{U}, t_i \in \mathbb{R}, \\ \Omega_i(u) &= \langle a_i, u \rangle - b_i, \quad a_i \in \mathcal{U}, b_i \in \mathbb{R}.\end{aligned}$$

Il faut montrer que si  $r \in (U^{\text{ad}})_u^\perp$ , alors  $r$  peut s'exprimer comme un élément de  $C(u)$  (voir (5.12)). Si  $r \in (U^{\text{ad}})_u^\perp$ , ceci signifie que

$$\forall v \in U^{\text{ad}}, \quad \langle r, v - u \rangle \leq 0,$$

ou encore

$$\left( \begin{array}{l} \langle \theta_i, v \rangle \leq t_i \\ \langle a_i, v \rangle \leq b_i \\ \langle -a_i, v \rangle \leq -b_i \end{array} \right) \Rightarrow \langle r, v - u \rangle \leq 0.$$

Cette dernière implication nous permet d'utiliser le Lemme de Farkas 3.58, ou mieux le Corollaire 3.59 et d'en déduire en particulier qu'il existe des scalaires non négatifs  $\lambda_i, i = 1, \dots, m$ , et  $\gamma_j, \gamma'_j, j = 1, \dots, p$ , tels que

$$r = \sum_{i=1}^m \lambda_i \theta_i + \sum_{j=1}^p (\gamma_j - \gamma'_j) a_j \quad \text{et} \quad \langle r, u \rangle \geq \sum_{i=1}^m \lambda_i t_i + \sum_{j=1}^p (\gamma_j - \gamma'_j) b_j.$$

En reportant l'expression de  $r$  dans l'inégalité et en tenant compte du fait que  $u \in U^{\text{ad}}$ , on se convainc aisément que les  $\lambda_i$  tels que  $\langle \theta_i, u \rangle < t_i$  doivent nécessairement être nuls. En posant  $\mu_j = \gamma_j - \gamma'_j$  (qui est alors de signe quelconque), on obtient la représentation souhaitée de  $r$  comme un élément de  $C(u)$ .  $\square$

**Corollaire 5.19.** *Si, dans (5.9), toutes les fonctions sont affines continues et si elles définissent un sous-ensemble  $U^{\text{ad}} \neq \emptyset$ , alors une solution  $u^\sharp$  de (5.1) est caractérisée par l'existence de "multiplicateurs"  $\lambda_i^\sharp, i = 1, \dots, m$ , satisfaisant (5.11) ou (5.13) (pour  $u = u^\sharp$ ) et de "multiplicateurs"  $\mu_j^\sharp, j = 1, \dots, p$ , tels que*

$$0 \in \partial J(u^\sharp) + \sum_{i=1}^m \lambda_i^\sharp \theta_i + \sum_{j=1}^p \mu_j^\sharp a_j.$$

Ce corollaire est une conséquence immédiate du Lemme 5.18 et de la caractérisation (5.5b) des solutions. On laisse le lecteur formuler un énoncé analogue utilisant la caractérisation (5.5c).

Pour pouvoir traiter le cas où les  $\Theta_i$  sont convexes mais non nécessairement affines, nous allons introduire une condition supplémentaire, connue sous le nom de *condition de Slater*, qui est aussi une condition de qualification des contraintes, mais en général non minimale, c'est-à-dire que c'est seulement une condition *suffisante* pour assurer (5.14). Cette hypothèse s'énonce ainsi (on rappelle que les contraintes égalité s'écrivent globalement  $A(u) = b$ ) :

$$\exists u^\circ : \Theta_i(u^\circ) < 0, \quad i = 1, \dots, m, \quad A(u^\circ) = b. \quad (5.15)$$

Cette hypothèse dit qu'il existe un point  $u^\circ$  satisfaisant *strictement* les contraintes (le mot "strictement" ne s'applique bien sûr qu'aux contraintes inégalité). Observons que l'existence d'un tel  $u^\circ$  n'était pas assurée dans le cas de l'Exemple 5.17.

Plutôt que de démontrer que (5.15) entraîne (5.14), on va montrer directement qu'elle permet de généraliser le Corollaire 5.19 au cas non affine. Cependant, comme les hypothèses de qualification des contraintes sont indépendantes de la fonction coût  $J$ , il doit être clair que le fait de pouvoir énoncer l'équivalent du Corollaire 5.19 pour tout  $J$  établit que (5.14) est également vérifiée.

**Théorème 5.20.** *Sous les hypothèses suivantes :*

- $J$  est une fonction convexe s.c.i. sous-différentiable ;
- les fonctions  $\Theta_i$  sont convexes, continues et sous-différentiables ;
- les fonctions  $\Omega_j$  sont affines et continues (les contraintes (5.9b) s'écrivent globalement  $A(u) = b$ ) ;
- les conditions (5.4) et (5.15) sont satisfaites.

Alors, une condition nécessaire et suffisante pour que  $u^\sharp$  soit une solution du problème (5.1) est qu'il existe des multiplicateurs  $\lambda_i^\sharp$ ,  $i = 1, \dots, m$ , satisfaisant (5.11) ou (5.13) (pour  $u = u^\sharp$ ) et des multiplicateurs  $\mu_j^\sharp$ ,  $j = 1, \dots, p$ , ( $\mu^\sharp$  désigne le vecteur de  $\mathbb{R}^p$  de coordonnées  $\mu_j^\sharp$ ) tels que

$$0 \in \partial J(u^\sharp) + \sum_{i=1}^m \lambda_i^\sharp \partial \Theta_i(u^\sharp) + A^* \mu^\sharp . \quad (5.16)$$

*Démonstration.* La condition suffisante est évidente puisque si (5.16) est vraie, ce qui veut dire que  $0 \in \partial J(u^\sharp) + C(u^\sharp)$ , alors a fortiori (5.5b) est vraie en raison de la conclusion du Lemme 5.16. Par ailleurs, (5.5b) est une condition suffisante d'optimalité.

Il nous faut maintenant montrer la condition nécessaire. On peut déjà écarter le cas trivial où  $u^\sharp$  lui-même est tel que  $\Theta_i(u^\sharp) < 0$  pour tout  $i$ . En effet, dans ce cas les contraintes inégalité n'interviennent plus et on est ramené à la situation déjà étudiée du Corollaire 5.19. On supposera donc désormais qu'il existe au moins une valeur de  $i$  telle que  $\Theta_i(u^\sharp) = 0$ . On va maintenant considérer la fonction suivante :

$$\Xi(u) \stackrel{\text{def}}{=} \max \left( J(u) - J(u^\sharp), \max_{i=1, \dots, m} \Theta_i(u) \right) .$$

On observe que  $\Xi(u^\sharp) = 0$  (car le premier terme dans le max prend la valeur 0 et les autres termes sont négatifs ou nuls). De plus, sur l'ensemble admissible défini par les contraintes égalité uniquement, cette fonction prend des valeurs non négatives. En effet, s'il existait un  $v$  vérifiant les contraintes égalité et donnant une valeur strictement négative à  $\Xi$ , alors nécessairement les contraintes inégalité seraient aussi vérifiées pour ce  $v$ , mais il faudrait aussi que  $J(v) < J(u^\sharp)$ , ce qui contredirait l'hypothèse que  $u^\sharp$  est optimal. La fonction  $\Xi$  est évidemment convexe et satisfait toutes les hypothèses pour que l'on puisse appliquer le Corollaire 5.19 au problème de sa minimisation sous les contraintes égalité. Puisque  $u^\sharp$  est une solution de ce problème d'après ce qui vient d'être dit, on sait qu'il existe  $\mu \in \mathbb{R}^p$  tel que

$$0 \in \partial \Xi(u^\sharp) + A^* \mu . \quad (5.17)$$

Le calcul du sous-différentiel de  $\Xi$  renvoie à celui sur l'enveloppe supérieure de fonctions. On a évoqué le problème à l'Exercice 4.67 mais la réponse complète se trouve par exemple dans [9]. Puisque le sup (ou le max) porte sur un ensemble fini, le sous-différentiel de l'enveloppe supérieure est la clôture convexe des sous-différentiels des fonctions "actives" en cause (celles qui contribuent à l'enveloppe supérieure en  $u^\sharp$ ), à savoir  $J$  et les  $\Theta_i$  telles que  $\Theta_i(u^\sharp) = 0$ , c'est-à-dire telles que  $i \in I(u^\sharp)$  (on a écarté plus haut le cas trivial où  $I(u^\sharp) = \emptyset$ ). En résumé, pour tout  $r \in \partial \Xi(u^\sharp)$ , il existe  $\alpha \in \Sigma_{m+1}$  avec  $\alpha_i = 0$  si  $i \notin I(u^\sharp)$  ainsi que des  $s_0 \in \partial J(u^\sharp)$ ,  $s_i \in \partial \Theta_i(u^\sharp)$  tels que

$$r = \sum_{i \in \{0, \dots, m\} \cap I(u^\sharp)} \alpha_i s_i .$$

La condition (5.17) se traduit donc par l'existence de  $\alpha_i$  et  $s_i$  ayant les propriétés énoncées, ainsi que  $\mu \in \mathbb{R}^p$ , tels que

$$\sum_{i \in \{0, \dots, m\} \cap I(u^\sharp)} \alpha_i s_i + A^* \mu = 0. \quad (5.18)$$

Si  $\alpha_0 \neq 0$ , la preuve est terminée car il suffit de diviser cette équation par  $\alpha_0$  et de rebaptiser  $\mu/\alpha_0$  en  $\mu^\sharp$  et  $\alpha_i/\alpha_0$  (tous les  $\alpha_i$  sont non négatifs) en  $\lambda_i^\sharp$ .

Montrons donc que  $\alpha_0$  ne peut pas être nul grâce à l'hypothèse de Slater. En effet, si c'était le cas, alors le  $r$  qui appartenait à  $\partial \Xi(u^\sharp)$  pourrait aussi s'interpréter comme appartenant au sous-différentiel de la fonction  $u \mapsto \max_{i=1, \dots, m} \Theta_i(u)$  (ne faisant plus intervenir  $J$ ) et la condition (5.18) impliquerait que cette fonction est minimale (sous les contraintes égalité) en  $u^\sharp$ . Or, avec l'hypothèse de non vacuité de  $I(u^\sharp)$  faite, cette fonction vaut 0 en  $u^\sharp$  alors qu'elle est strictement négative en  $u^\circ$  par l'hypothèse (5.15), d'où une contradiction. Ceci achève la démonstration.  $\square$

**Corollaire 5.21.** *On reprend toutes les hypothèses du Théorème 5.20 mais on suppose de plus ici que les fonctions  $J$  et  $\Theta_i$ ,  $i = 1, \dots, m$ , sont différentiables. Alors, une condition nécessaire et suffisante pour que  $u^\sharp$  soit une solution du problème (5.1) est qu'il existe des multiplicateurs  $\lambda_i^\sharp$ ,  $i = 1, \dots, m$ , satisfaisant (5.11) ou (5.13) (pour  $u = u^\sharp$ ) et des multiplicateurs  $\mu_j^\sharp$ ,  $j = 1, \dots, p$ , tels que*

$$J'(u^\sharp) + \sum_{i=1}^m \lambda_i^\sharp \Theta'_i(u^\sharp) + \sum_{j=1}^p \mu_j^\sharp \Omega'_j(u^\sharp) = 0. \quad (5.19)$$

**Remarque 5.22.** On trouve parfois dans la littérature l'hypothèse supplémentaire suivante concernant les contraintes égalité :

“ $A$  est surjective”

ou bien, puisque  $A$  est constituée des lignes  $a_j$ ,

“les  $a_j$  sont linéairement indépendants”

(la surjectivité de  $A$  implique l'injectivité de  $A^*$  et donc la non existence d'une combinaison linéaire non triviale et nulle des  $a_j$ ). Si  $A$  n'est pas surjective, ou bien  $b \notin \text{im } A$  et alors  $U^{\text{ad}} = \emptyset$ , ou bien  $b \in \text{im } A$  et les contraintes égalité sont alors redondantes (certaines s'expriment comme des combinaisons linéaires des autres). L'hypothèse de surjectivité de  $A$  est donc raisonnable et peu contraignante. Elle évite en particulier, dans le cas où  $U^{\text{ad}} \neq \emptyset$ , donc  $b \in \text{im } A$ , la non unicité et le caractère non borné des multiplicateurs. En effet, comme le vecteur  $\mu^\sharp$  des multiplicateurs intervient sous la forme  $A^* \mu^\sharp$ , si  $A^*$  n'est pas injective, on peut ajouter à un  $\mu^\sharp$  qui convient, n'importe quel élément du noyau de  $A^*$ .

Les conditions de Lagrange-Karush-Kuhn-Tucker ont été étudiées par Lagrange dans le cas de fonctions différentiables et pour des contraintes égalité uniquement. Les généralisations aux autres situations sont partiellement dûes aux autres auteurs cités. En fait, ces conditions constituent des conditions *nécessaires* dans le cas non convexe (sous des hypothèses de qualification des contraintes qui prennent alors des formes diverses), et elles deviennent des conditions suffisantes dans le cas convexe. En tout cas, elles peuvent être considérées comme des conditions “locales” en ce sens qu'elles font intervenir des notions de différentiabilité ou de sous-différentiabilité. Évidemment, dans le cas convexe, on a vu que l'on passe assez facilement des notions locales aux notions globales. Cependant, le chapitre suivant abordera un autre point de vue spécifiquement global (c'est-à-dire où on ne parle pas de dérivées ou de sous-gradients mais seulement d'inégalités) sur les multiplicateurs et la *dualité*. Signalons l'appellation “*variables duales*” pour ce que nous avons appelé “multiplicateurs” (de Lagrange ou de Kuhn et Tucker).

### 5.2.4 Interprétation marginaliste des multiplicateurs : explication géométrique heuristique

On va donner ici une explication heuristique simple de cette interprétation marginaliste des multiplicateurs sur le cas représenté sur la Figure 5.4. On y a représenté dans  $\mathcal{U} = \mathbb{R}^2$  deux contraintes inégalité

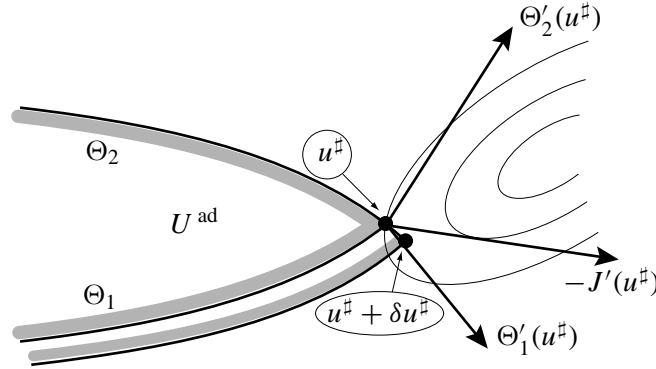


Figure 5.4: Perturbation d'une contrainte

$\Theta_1$  et  $\Theta_2$ , toutes les deux actives à la solution  $u^\#$ , ainsi que les ensembles de niveau de la fonction coût  $J$ . Toutes les fonctions sont supposées différentiables pour simplifier. On suppose qu'une perturbation d'amplitude  $\varepsilon > 0$  "très petite" est appliquée au second membre de la première contrainte, ce qui a pour effet de déplacer la partie correspondante de la frontière de  $U^{\text{ad}}$ . Autrement dit, la contrainte devient  $\Theta_1(u) \leq \varepsilon$  au lieu de  $\Theta_1(u) \leq 0$ . Ceci correspond à un élargissement du sous-ensemble admissible  $U^{\text{ad}}$  et donc à une diminution (en tout cas une "non augmentation") du coût optimal  $J(u^\#)$  (exercice). Si le second membre des contraintes représente une ressource disponible en quantité limitée,<sup>4</sup> la petite perturbation  $\varepsilon > 0$  correspond à un accroissement marginal de la ressource disponible, et encore une fois à une amélioration concomitante du coût optimal. De combien s'améliore le coût optimal ? C'est à cette question que permettent de répondre les multiplicateurs.

Revenons à la Figure 5.4 et observons que puisque la contrainte en  $\Theta_2$  était active avant la perturbation et que cette perturbation n'a pas affecté cette contrainte, la solution optimale *après* perturbation, notée  $u^\# + \delta u^\#$ , a "glissé" le long de la frontière relative à  $\Theta_2$  et donc,  $\langle \Theta'_2(u^\#), \delta u^\# \rangle = 0$  puisque le gradient  $\Theta'_2(u^\#)$  est lui orthogonal à la frontière. Par ailleurs,  $\langle \Theta'_1(u^\#), \delta u^\# \rangle = \varepsilon$  puisque  $\varepsilon$  représente la variation autorisée de  $\Theta_1(u^\#)$  qui peut être évaluée "au premier ordre" si elle est petite.<sup>5</sup> La condition d'optimalité du type (5.16) est ici, pour certains  $\lambda_1^\#$  et  $\lambda_2^\#$  non négatifs :

$$J'(u^\#) + \lambda_1^\# \Theta'_1(u^\#) + \lambda_2^\# \Theta'_2(u^\#) = 0 .$$

En faisant le produit scalaire avec  $\delta u^\#$  et en fonction des explications qui viennent d'être données, on obtient que :

$$\begin{aligned} 0 &= \left\langle J'(u^\#) + \lambda_1^\# \Theta'_1(u^\#) + \lambda_2^\# \Theta'_2(u^\#), \delta u^\# \right\rangle \\ &= \langle J'(u^\#), \delta u^\# \rangle + \lambda_1^\# \varepsilon . \end{aligned}$$

<sup>4</sup>On a toujours pris apparemment 0 comme second membre des contraintes inégalité mais c'est une illusion car un second membre non nul peut être caché dans l'expression des  $\Theta_i$ .

<sup>5</sup>Ce raisonnement suppose bien sûr que la contrainte  $\Theta_1$  qui était active *avant* la perturbation le reste *après* cette perturbation, mais ce sera génériquement le cas si les ensembles de niveau ne sont pas perpendiculaires à la frontière  $\Theta_1$  au point  $u^\#$  — c'est-à-dire si  $\langle J'(u^\#), \Theta'_1(u^\#) \rangle$  n'est pas nul, ce qui n'a aucune raison particulière d'être le cas.

Autrement dit, le produit de la perturbation  $\varepsilon$  du second membre de la contrainte par son multiplicateur associé est égal, au signe près, à la variation du coût optimal.<sup>6</sup> L'interprétation marginaliste des multiplicateurs est finalement la suivante : *c'est au signe près la sensibilité du coût optimal par rapport à une variation "marginale" du second membre de la contrainte correspondante.*

Du point de vue économique, le multiplicateur représente donc le "juste prix" de la ressource correspondante, c'est-à-dire le prix qu'un responsable des approvisionnements de cette ressource serait prêt à payer pour en acquérir une petite quantité  $\varepsilon$  supplémentaire en fonction de l'amélioration espérée de la fonction "coût optimal". Du coup, l'interprétation de la condition d'écart complémentaires devient claire. Si une ressource n'est pas "rare", c'est-à-dire si la quantité disponible n'est pas utilisée à plein à l'optimum (ou encore, si la contrainte correspondante n'est pas active à l'optimum), alors le prix marginal de cette ressource est nul : en acquérir une petite quantité supplémentaire ne changera pas la solution (si la contrainte n'était pas active, déplacer sa frontière ne change pas la solution) et donc le revenu correspondant est nul. Inversement, seules les ressources rares (correspondant à des contraintes saturées à l'optimum) peuvent avoir un prix marginal non nul. Le cas limite entre les deux cas que nous venons d'évoquer, celui où une contrainte est saturée mais son multiplicateur est nul, ce cas peut parfaitement se produire, mais c'est un cas "dégénéré" ou "non générique" (c'est en fait dans notre exemple le cas évoqué dans la note de bas de page 5 : exercice).

Nous reviendrons au chapitre suivant sur l'interprétation marginaliste des multiplicateurs sous la forme plus précise de sous-gradients (au signe près) de la fonction "coût optimal" comme fonction du second membre des contraintes.

### 5.2.5 Retour sur la pénalisation exacte

On revient sur la technique de pénalisation exacte déjà abordée au §5.1.5 mais dans une formulation implicite des contraintes. Avec la description explicite (5.9), on peut introduire des fonctions  $\tau$  et  $\omega$  de  $\mathbb{R}$  dans  $\mathbb{R}_+$  que l'on supposera convexes, continues et sous-différentiables et satisfaisant :

$$\tau(x) \begin{cases} = 0 & \text{si } x \leq 0, \\ > 0 & \text{sinon.} \end{cases}, \quad \omega(x) \begin{cases} = 0 & \text{si } x = 0, \\ > 0 & \text{sinon.} \end{cases}$$

Ainsi,  $\tau(\Theta_i(u))$  constitue une pénalisation de la contrainte inégalité  $\Theta_i(u) \leq 0$  et  $\omega(\Omega_j(u))$  est une pénalisation de la contrainte égalité  $\Omega_j(u) = 0$ . Un exemple de telles fonctions est

$$\tau(x) = \max(0, x), \quad \omega(x) = |x|. \quad (5.20)$$

Observons que ces fonctions ne sont rien d'autre que les fonctions distance aux sous-ensembles de  $\mathbb{R}$  constitués par la demi-droite  $x \leq 0$  et le point  $\{0\}$ , respectivement. On a alors le résultat suivant.

**Théorème 5.23.** *Soit  $u^\sharp$  une solution du problème (5.1) où  $U^{\text{ad}}$  est défini par les contraintes (5.9) et supposons qu'il existe des multiplicateurs optimaux  $\lambda_i^\sharp$  et  $\mu_j^\sharp$  relatifs à ces contraintes. Si*

$$\forall i = 1, \dots, m, \lambda_i^\sharp \in \partial\tau(0) \quad \text{et} \quad \forall j = 1, \dots, p, \mu_j^\sharp \in \partial\omega(0), \quad (5.21)$$

*alors  $u^\sharp$  est aussi la solution du problème "pénalisé"*

$$\min_{u \in \mathcal{U}} \left( J(u) + \sum_{i=1}^m \tau(\Theta_i(u)) + \sum_{j=1}^p \omega(\Omega_j(u)) \right). \quad (5.22)$$

<sup>6</sup>On a d'ailleurs la confirmation, étant donné que  $\varepsilon$  et  $\lambda_1^\sharp$  sont positifs ou au moins non négatifs, que le coût optimal ne peut que s'améliorer, c'est-à-dire diminuer ou au moins ne pas augmenter

*Démonstration.* Observons d'abord que  $\Omega_j(u^\sharp) = 0$  pour  $j = 1, \dots, p$  et donc on a aussi (d'après (5.21)) que  $\mu_j^\sharp \in \partial\omega(\Omega_j(u^\sharp))$ . Remarquons aussi que  $\Theta_i(u^\sharp) \leq 0$  pour  $i = 1, \dots, m$ , :

- si  $\Theta_i(u^\sharp) < 0$  pour un certain  $i$ , alors d'une part  $\lambda_i^\sharp = 0$  (écarts complémentaires) et d'autre part  $\partial\tau(\Theta_i(u^\sharp)) = \{0\}$  (car  $\tau$  est identiquement nul pour  $x \leq 0$ ) ; on peut donc dire que  $\lambda_i^\sharp \in \partial\tau(\Theta_i(u^\sharp))$  ;
- si  $\Theta_i(u^\sharp) = 0$ , la condition exprimée par (5.21) peut encore s'écrire  $\lambda_i^\sharp \in \partial\tau(\Theta_i(u^\sharp))$ .

Finalement, la condition (5.21) se reformule aussi sous la forme

$$\forall i = 1, \dots, m, \lambda_i^\sharp \in \partial\tau(\Theta_i(u^\sharp)) \quad \text{et} \quad \forall j = 1, \dots, p, \mu_j^\sharp \in \partial\omega(\Omega_j(u^\sharp)).$$

De ce fait, on a

$$\begin{aligned} \tau(\Theta_i(u)) &\geq \tau(\Theta_i(u^\sharp)) + \lambda_i^\sharp(\Theta_i(u) - \Theta_i(u^\sharp)) \\ &= \lambda_i^\sharp\Theta_i(u), \end{aligned} \tag{5.23a}$$

car  $\lambda_i^\sharp\Theta_i(u^\sharp) = 0$  par la condition d'écarts complémentaires. De même,

$$\omega(\Omega_j(u)) \geq \mu_j^\sharp\Omega_j(u). \tag{5.23b}$$

La condition (5.16) s'interprète comme le fait que  $u^\sharp$  réalise le minimum de la fonction

$$J(\cdot) + \sum_i^m \lambda_i^\sharp\Theta_i(\cdot) + \sum_{j=1}^p \mu_j^\sharp\Omega_j(\cdot)$$

(que l'on étudiera comme étant le "Lagrangien" du problème d'optimisation sous contraintes). On a donc, pour tout  $u \in \mathcal{U}$ ,

$$\begin{aligned} J(u^\sharp) &= J(u^\sharp) + \sum_i^m \lambda_i^\sharp\tau(\Theta_i(u^\sharp)) + \sum_{j=1}^p \mu_j^\sharp\omega(\Omega_j(u^\sharp)) \\ &= J(u^\sharp) + \sum_i^m \lambda_i^\sharp\Theta_i(u^\sharp) + \sum_{j=1}^p \mu_j^\sharp\Omega_j(u^\sharp) \\ &\leq J(u) + \sum_i^m \lambda_i^\sharp\Theta_i(u) + \sum_{j=1}^p \mu_j^\sharp\Omega_j(u) \\ &\leq J(u) + \sum_i^m \lambda_i^\sharp\tau(\Theta_i(u)) + \sum_{j=1}^p \mu_j^\sharp\omega(\Omega_j(u)) \end{aligned}$$

la dernière inégalité résultant de l'observation (5.23). Ceci achève la démonstration.  $\square$

Avec les fonctions (5.20), la condition (5.21) est facile à garantir si on connaît un majorant  $M$  de la valeur absolue des tous les multiplicateurs  $\lambda_i^\sharp$  et  $\mu_j^\sharp$ . En effet, puisque le sous-différentiel en 0 de  $\tau$ , respectivement  $\omega$ , est égal à  $[0, 1]$ , respectivement  $[-1, 1]$ , il suffit de considérer les fonctions de pénalisation  $\alpha\tau$ , respectivement  $\alpha\omega$  pour un quelconque  $\alpha \geq M$ .

L'interprétation économique de la technique de pénalisation exacte est alors immédiate. Puisque les multiplicateurs mesurent l'intérêt marginal de déplacer le second membre des contraintes, dès que le prix à payer (via les fonctions de pénalisation) pour aller au delà devient supérieur à ce prix marginal,



on n'a plus intérêt à outrepasser le niveau des contraintes alloué. Du point de vue pratique, on peut le plus souvent avoir une estimation de ce majorant  $M$  des prix marginaux à partir de la connaissance des données économiques du problème traité. Plus  $M$  est choisi grand, moins on a de chance de se tromper dans cette estimation, mais en contrepartie, la non-différentiabilité des fonctions de pénalisation au point 0 est amplifiée, ce qui se traduit par plus de difficulté dans la résolution numérique du problème pénalisé.

### 5.3 Résumons nous

Dans ce chapitre, on a abordé les problèmes d'optimisation convexe sous contraintes. Les contraintes ont d'abord été décrites implicitement ou géométriquement par la donnée d'un ensemble "admissible", puis par une collection finie de contraintes explicites scalaires sous forme d'égalités et d'inégalités à respecter. Alors que l'espace de décision est un espace de Hilbert de dimension éventuellement infinie, lorsque les contraintes sont décrites explicitement, l'espace dual (celui dans lequel on cherchera les multiplicateurs) est ici fini. Cette restriction sera levée ultérieurement.

Le cadre dit "convexe" de tels problèmes est celui où la fonction coût est convexe et l'ensemble admissible est convexe, ce qui implique, dans le cas de contraintes explicites, que les égalités sont définies à l'aide de fonctions affines et les inégalités ( $\leq 0$ ) à l'aide de fonctions convexes.

Sous des hypothèses topologiques supplémentaires (semi-continuité inférieure de la fonction coût, fermeture de l'ensemble admissible, ce qui nécessite des fonctions affines continues pour les égalités et des fonctions convexes s.c.i. pour les inégalités), et sous une hypothèse de "coercivité" destinée à empêcher la solution de se situer à l'infini, on peut garantir l'existence d'une solution. Dans le cas convexe, toute solution "locale" est aussi "globale".

Il existe diverses formes de la caractérisation locale des solutions optimales dans le cadre de contraintes implicites. Géométriquement, elles disent toutes qu'un sous-gradient de la fonction coût à l'optimum appartient à l'opposé du cône orthogonal aux contraintes à cet optimum. Ceci garantit que toute variation admissible autour de la solution (nécessairement dans le cône tangent à l'ensemble admissible) ne peut pas faire décroître la fonction coût. Une expression possible de cette condition revêt la forme d'une *inéquation variationnelle*. On peut, en décomposant la fonction coût en une somme de termes, mélanger dans l'expression des conditions d'optimalité des variations locales ou "au premier ordre" avec des variations exactes.

Lorsque les contraintes prennent une forme explicite, la question centrale est de savoir si le cône orthogonal à l'ensemble admissible à la solution peut lui-même s'exprimer explicitement comme le cône convexe fermé engendré par les sous-gradients des diverses fonctions entrant dans cette expression des contraintes. Malheureusement, ce n'est pas toujours le cas, sauf dans le cas où toutes les contraintes sont affines. Le rôle des hypothèses dites "de qualification des contraintes" est précisément de garantir ce résultat. Il ne faut pas perdre de vue que

- la plupart de ces hypothèses de qualification des contraintes sont des conditions *suffisantes* pour garantir l'existence de multiplicateurs optimaux (dans ce chapitre, on a donné la plus répandue, l'hypothèse dite "de Slater", qui requiert l'existence d'un point admissible satisfaisant strictement les contraintes inégalités) ;
- l'expression même d'un ensemble admissible par une collection de contraintes n'a rien d'intrinsèque et certaines expressions peuvent satisfaire une hypothèse de qualification alors que d'autres ne la satisferont pas (ce qui ne prouve pas d'ailleurs qu'on ne pourra pas malgré tout leur associer des multiplicateurs optimaux).

Ces multiplicateurs optimaux, lorsqu'ils existent, expriment en quelque sorte les "coordonnées" de l'opposé d'un sous-gradient de la fonction coût à la solution sur les sous-gradients de toutes les con-

traintes actives (contraintes égalité et contraintes inégalité saturées à l'optimum). Les contraintes inégalité non saturées ont des multiplicateurs nuls. Ces conditions (les multiplicateurs sont nuls si les contraintes inégalité correspondantes sont non saturées à la solution, ou de façon équivalente, les multiplicateurs non nuls pour les contraintes inégalité correspondent obligatoirement à des contraintes non saturées), dites d'“écarts complémentaires”, sont compatibles avec l'interprétation économique qui a été faite de tous les multiplicateurs (pour les contraintes égalité aussi bien qu'inégalité) : ils mesurent, changés de signe, la sensibilité du coût optimal au changement du second membre des contraintes. De cette interprétation, on peut aussi déduire que les multiplicateurs associés à des contraintes égalité sont de signe quelconque alors que ceux associés à des contraintes inégalité sont non négatifs.

La technique dite de “pénalisation exacte” peut également être liée à cette interprétation marginaliste des multiplicateurs optimaux. Il s'agit de remplacer les contraintes par des termes supplémentaires dans la fonction coût. Ces termes doivent être “neutres” (ils doivent prendre la valeur 0) lorsque les contraintes sont satisfaites en un point, et doivent prendre une valeur positive (ils “pénalisent” si on minimise) dans le cas contraire. L'idée intuitive est que si le coût marginal pour violer une contrainte est en valeur absolue supérieur à la sensibilité fournie par le multiplicateur correspondant, on sera dissuadé de violer cette contrainte. Comme toujours en optimisation, ce qui compte donc ce n'est pas tant les valeurs prises par les fonctions coût ou de pénalité, mais bien leur “sensibilité” exprimée par leur sous-différentiel. Pour que la pénalisation soit effective dès que l'on quitte l'ensemble admissible, il faut que la dérivée directionnelle dans ces directions soit non nulle, et comme elle est par contre nulle vers l'intérieur de cet ensemble admissible, on pressent que la pénalisation ne sera “exacte” que si ces fonctions de pénalisation sont non différentiables au bord de l'ensemble admissible.

## Chapitre 6

# Conditions d'optimalité globales et techniques Lagrangiennes

Le chapitre précédent a permis d'étudier les conditions d'optimalité d'un problème d'optimisation sous contraintes en termes de configuration d'un sous-gradient de la fonction coût par rapport au cône orthogonal à l'ensemble admissible au point solution (cas des contraintes implicites) ou par rapport au cône engendré par les sous-gradients des contraintes actives (cas des contraintes explicites avec qualification des contraintes). Pour cette raison, nous avons qualifié ces conditions de "locales".

Cependant, comme on l'a déjà souligné à plusieurs reprises dans ce cours, les hypothèses de convexité permettent de passer facilement de considérations locales à des considérations globales, et les conditions énoncées précédemment sont alors à la fois nécessaires et suffisantes. Sans hypothèses de convexité, ces conditions locales sont en fait des conditions nécessaires seulement, et elles font appel à des hypothèses de différentiabilité (ou de "sous-différentiabilité", mais en un sens qu'il faudrait préciser car nous n'avons défini cette notion que dans le cas des fonctions convexes).

Le présent chapitre aborde la question des conditions d'optimalité de façon *globale* (en particulier sans aucune référence à des notions de différentiabilité) et en termes de conditions *suffisantes*. On va voir en particulier que la partie primale d'un point selle du Lagrangien du problème (si un tel point selle existe) est une solution du problème d'optimisation sous contraintes indépendamment de toute considération de convexité. Évidemment, dans le cas convexe, cette théorie globale des conditions suffisantes rejoint la théorie locale des conditions nécessaires.

### 6.1 Formulation compacte des contraintes, Lagrangien, point selle et condition suffisante

#### 6.1.1 Contraintes égalité et inégalité en dimension finie ou infinie

Dans le chapitre précédent, on a explicité les contraintes sous forme d'une collection *finie* de contraintes *scalaires* égalité et inégalité (voir (5.9)). On va adopter ici une formulation beaucoup plus compacte et qui permettra de recouvrir dans la même notation les contraintes égalité et inégalité à valeurs dans des espaces de dimension finie ou infinie. À cet effet, on introduit un "espace des contraintes"  $\mathcal{C}$  qui est un espace de Hilbert, et une "fonction contrainte"  $\Theta : \mathcal{U} \rightarrow \mathcal{C}$ . On introduit aussi un cône convexe et saillant  $C$  dans  $\mathcal{C}$  et la relation d'ordre associée (voir (3.8)). Les contraintes seront alors formulées sous la forme  $\Theta(u) \in -C$ , ce qui signifie  $\Theta(u) \leq 0$ , mais nous avons dit au §3.5.3 pourquoi nous n'utiliserons pas cette dernière notation.

Malgré les apparences, nous pouvons par cette formulation traduire aussi des contraintes égalité : il suffit que le cône  $C$  soit réduit à  $\{0\}$  pour que  $\Theta(u) \in -C$  signifie en fait  $\Theta(u) = 0$ . Plus généralement,

on peut considérer un espace produit  $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2$  avec un cône de la forme  $C_1 \times C_2$  où  $C_2 = \{0\}$  pour que l'on puisse traiter simultanément des contraintes inégalité et égalité.

En plus des contraintes explicites  $\Theta(u) \in -C$ , nous allons aussi considérer des contraintes implicites sous la forme  $u \in U^{\text{ad}}$  où  $U^{\text{ad}}$  est un sous-ensemble convexe fermé. Les contraintes explicites seront les contraintes traitées par dualité (c'est-à-dire qu'on va leur associer des multiplicateurs ; on dira aussi qu'on va les "dualiser") alors que les contraintes implicites seront toujours traitées comme des contraintes. Finalement, on considère le problème d'optimisation suivant :

$$\min_{u \in U^{\text{ad}}} J(u) \quad (6.1a)$$

sous

$$\Theta(u) \in -C . \quad (6.1b)$$

On pose

$$U_{\Theta}^{\text{ad}} = \{u \in \mathcal{U} \mid \Theta(u) \in -C\} . \quad (6.2)$$

Autrement dit, le sous-ensemble admissible du problème (6.1) est le sous-ensemble  $U^{\text{ad}} \cap U_{\Theta}^{\text{ad}}$ .

### 6.1.2 Lagrangien

On considère  $\lambda \in \mathcal{C}^*$  et on pose

$$L(u, \lambda) = J(u) + \langle \lambda, \Theta(u) \rangle . \quad (6.3)$$

Cette expression est appelée *Lagrangien* associé au problème d'optimisation (6.1) et la variable  $\lambda$  est appelée *multiplicateur* ou *variable duale*. Par opposition,  $u$  est la *variable primale*.

De même que  $u$  est restreinte au sous-ensemble admissible  $U^{\text{ad}}$ ,  $\lambda$  va être restreinte au sous-ensemble admissible  $C^*$  (voir (3.9)). On observe que cela correspond aux contraintes de positivité sur les multiplicateurs associés aux contraintes inégalité que nous avons rencontrées au chapitre précédent. On note aussi que dans le cas de contraintes égalité, correspondant au choix  $C = \{0\}$ , la définition (3.9) de  $C^*$  donne  $C^* = \mathcal{C}^*$ , ce qui est consistant avec le fait que les multiplicateurs associés à des contraintes égalité ne supportent pas de contraintes de signe.

Le lemme suivant constitue une première justification de l'introduction du Lagrangien.

**Lemme 6.1.** *Soit*

$$P(u) \stackrel{\text{def}}{=} \sup_{\lambda \in C^*} L(u, \lambda) . \quad (6.4)$$

Alors  $P(u) = J(u) + I_{U_{\Theta}^{\text{ad}}}(u)$  (voir (6.2) pour la définition de  $U_{\Theta}^{\text{ad}}$ ) et le problème

$$\min_{u \in U^{\text{ad}}} P(u) \quad (6.5)$$

est équivalent au problème (6.1) (au sens que les valeurs optimales de la fonction coût sont égales et toute solution de l'un est solution de l'autre problème).

*Démonstration.* Observons que si  $u \in U_{\Theta}^{\text{ad}}$ , alors  $\langle \lambda, \Theta(u) \rangle \leq 0$  pour tout  $\lambda \in C^*$  et que la valeur 0 de cette expression est atteinte pour  $\lambda = 0$  (par exemple). Donc,

$$u \in U_{\Theta}^{\text{ad}} \Rightarrow \sup_{\lambda \in C^*} \langle \lambda, \Theta(u) \rangle = 0 .$$

Inversement, du fait que  $C = C^{**}$  (voir Lemme 3.61), si  $\langle \lambda, \Theta(u) \rangle \leq 0$  pour tout  $\lambda \in C^*$ , alors  $\Theta(u) \in -C^{**} = -C$ , et donc, si  $\Theta(u) \notin -C$ , c'est-à-dire  $u \notin U_{\Theta}^{\text{ad}}$ , alors il existe  $\lambda \in C^*$  tel que  $\langle \lambda, \Theta(u) \rangle > 0$ . Puisque  $C^*$  est un cône, il en résulte que

$$u \notin U_{\Theta}^{\text{ad}} \Rightarrow \sup_{\lambda \in C^*} \langle \lambda, \Theta(u) \rangle = +\infty .$$

De tout ceci, on déduit que  $P$  définie par (6.4) est égale à  $J + I_{U_{\Theta}^{\text{ad}}}$ . Il est alors clair que minimiser  $P$  sur  $U^{\text{ad}}$  est équivalent à minimiser  $J$  sur  $U^{\text{ad}} \cap U_{\Theta}^{\text{ad}}$ , ce qui est une autre formulation de (6.1).  $\square$

### 6.1.3 Point selle

#### 6.1.3.1 Définition et caractérisations

**Définition 6.2.** Pour une fonction  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  et des sous-ensembles convexes fermés  $X \subset \mathcal{X}$  et  $Y \subset \mathcal{Y}$ , on dit que  $(x^{\sharp}, y^{\sharp}) \in X \times Y$  est un *point selle* de  $f$  sur  $X \times Y$  si

$$\forall x \in X, \forall y \in Y, f(x^{\sharp}, y) \leq f(x^{\sharp}, y^{\sharp}) \leq f(x, y^{\sharp}) . \tag{6.6a}$$

La Figure 6.1 illustre cette notion (appelée aussi “point col”). L'idée générale est qu'on cherche à

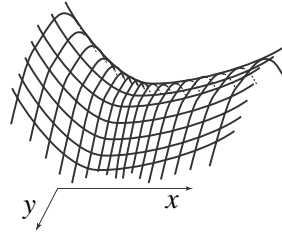


Figure 6.1: Point selle

maximiser  $f$  en  $y$  alors qu'on cherche à la minimiser en  $x$ . Pour cette raison, on peut voir cette notion dans une interprétation “théorie des jeux”, plus précisément un jeu à “somme nulle”, en ce sens que si les joueurs cherchent à minimiser leur propre fonction coût par rapport à la variable de décision qu'ils manipulent, la fonction coût du joueur  $x$  est  $f(x, y)$  alors que celle du joueur  $y$  est  $-f(x, y)$  (dans la mesure où maximiser  $f$  revient à minimiser  $-f$ ), d'où la “somme nulle” des deux fonctions coût (jeu purement antagoniste où ce que gagne l'un des joueurs est exactement ce qui perdu par l'autre).

Il existe plusieurs autres caractérisations des points selle et des propriétés qu'il est important de connaître. Remarquons d'abord qu'une définition plus brève du point selle est la suivante (exercice) :  $x^{\sharp} \in X, y^{\sharp} \in Y$  et

$$f(x^{\sharp}, y) \leq f(x, y^{\sharp}), \forall x \in X, \forall y \in Y . \tag{6.6b}$$

En second lieu, un point selle peut-être vu comme le *point fixe* d'une multi-application. Considérons les multi-applications  $X^{\sharp}$  et  $Y^{\sharp}$  suivantes :

$$X^{\sharp}(y) = \arg \min_{x \in X} f(x, y) ; \quad Y^{\sharp}(x) = \arg \max_{y \in Y} f(x, y) ,$$

et la multi-application  $F^{\sharp} : \mathcal{X} \times \mathcal{Y} \rightarrow 2^{\mathcal{X} \times \mathcal{Y}}$  définie par  $F^{\sharp}(x, y) = X^{\sharp}(y) \times Y^{\sharp}(x)$ .

**Exercice 6.3.** Montrer que  $(x^{\sharp}, y^{\sharp})$  est un point selle de  $f$  sur  $X \times Y$  si et seulement si c'est un point fixe de la multi-application  $F^{\sharp}$ , c'est-à-dire que

$$(x^{\sharp}, y^{\sharp}) \in F^{\sharp}(x^{\sharp}, y^{\sharp}) .$$

Ce point de vue “point fixe” est à la base de théorèmes d’existence de points selle à partir de théorèmes d’existence de points fixes.

Le lemme suivant propose une autre caractérisation des points selle. Auparavant, mentionnons le résultat général suivant.

**Exercice 6.4.** Montrer que, d’une manière générale,

$$\sup_y \inf_x f(x, y) \leq \inf_x \sup_y f(x, y) . \quad (6.7)$$

**Lemme 6.5.** *Le point  $(x^\sharp, y^\sharp)$  est un point selle de  $f$  sur  $X \times Y$  si et seulement si*

$$f(x^\sharp, y^\sharp) = \sup_{y \in Y} f(x^\sharp, y) = \min_{x \in X} \sup_{y \in Y} f(x, y) = \max_{y \in Y} \inf_{x \in X} f(x, y) = \inf_{x \in X} f(x, y^\sharp) , \quad (6.8)$$

autrement dit  $\inf_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \inf_{x \in X} f(x, y)$  et les inf et sup les plus extérieurs sont atteints en  $x^\sharp$ , respectivement  $y^\sharp$ .

*Démonstration.* Soit  $(x^\sharp, y^\sharp)$  vérifiant (6.6a). L’inégalité de droite montre que

$$f(x^\sharp, y^\sharp) = \min_{x \in X} f(x, y^\sharp) ,$$

et évidemment,

$$\inf_{x \in X} f(x, y^\sharp) \leq \sup_{y \in Y} \inf_{x \in X} f(x, y) ,$$

donc

$$f(x^\sharp, y^\sharp) \leq \sup_{y \in Y} \inf_{x \in X} f(x, y) .$$

En raisonnant de façon symétrique avec l’inégalité de gauche de (6.6a), on montre de façon similaire que

$$f(x^\sharp, y^\sharp) = \max_{y \in Y} f(x^\sharp, y) \geq \inf_{x \in X} \sup_{y \in Y} f(x, y) .$$

Étant donné l’inégalité inverse (6.7), les égalités (6.8) sont établies.

Réciproquement, supposons que (6.8) ait lieu. On a que

$$\forall y \in Y, \quad f(x^\sharp, y) \leq \sup_{y \in Y} f(x^\sharp, y) = f(x^\sharp, y^\sharp) .$$

De façon symétrique, on a que

$$\forall x \in X, \quad f(x, y^\sharp) \geq \inf_{x \in X} f(x, y^\sharp) = f(x^\sharp, y^\sharp) ,$$

et donc (6.6a) est établie. □

L’exercice suivant montre que les points selle forment un sous-ensemble “rectangulaire” de  $X \times Y$ .

**Exercice 6.6.** © Si  $(x_1^\sharp, y_1^\sharp)$  et  $(x_2^\sharp, y_2^\sharp)$  sont deux points selle de  $f$  sur  $X \times Y$ , alors  $(x_1^\sharp, y_2^\sharp)$  et  $(x_2^\sharp, y_1^\sharp)$  sont deux autres points selle et

$$f(x_1^\sharp, y_1^\sharp) = f(x_2^\sharp, y_2^\sharp) = f(x_1^\sharp, y_2^\sharp) = f(x_2^\sharp, y_1^\sharp) . \quad (6.9)$$

### 6.1.3.2 Cas convexe-concave

On dit que  $f : X \times Y \rightarrow \mathbb{R}$  est *convexe-concave* si l'application  $x \mapsto f(x, y)$  est convexe pour tout  $y$  et l'application  $y \mapsto f(x, y)$  est concave pour tout  $x$ . En supposant de plus  $f$  sous-différentiable en  $x$  et en  $y$ ,<sup>1</sup> les inégalités de droite et de gauche de (6.6a) sont équivalentes, respectivement, à

$$\exists r^\sharp \in \partial_x f(x^\sharp, y^\sharp) : \forall x \in X, \langle r^\sharp, x - x^\sharp \rangle \geq 0, \quad (6.10a)$$

$$\exists s^\sharp \in \partial_y f(x^\sharp, y^\sharp) : \forall y \in Y, \langle s^\sharp, y - y^\sharp \rangle \leq 0. \quad (6.10b)$$

On peut réécrire la seconde inéquation variationnelle en changeant  $s^\sharp$  en  $t^\sharp = -s^\sharp$  et en changeant le sens de l'inégalité :

$$\exists t^\sharp \in -\partial_y f(x^\sharp, y^\sharp) : \forall y \in Y, \langle t^\sharp, y - y^\sharp \rangle \geq 0. \quad (6.10c)$$

L'intérêt de cette manipulation est le suivant. Posons  $z = (x, y) \in \mathcal{Z} \stackrel{\text{def}}{=} X \times Y$ ,  $Z = X \times Y \subset \mathcal{Z}$ , et définissons l'opérateur (éventuellement multivoque)

$$A : z = (x, y) \in \mathcal{Z} \mapsto \partial_x f(x, y) \times (-\partial_y f(x, y)) \subset X \times Y. \quad (6.11)$$

Alors, on peut résumer les deux inéquations variationnelles (6.10a)–(6.10c) en une seule :

$$\exists a^\sharp \in A(z^\sharp) : \forall z \in Z, \langle a^\sharp, z - z^\sharp \rangle \geq 0. \quad (6.12)$$

En plus de ramener un couple d'inéquations variationnelles (6.10a)–(6.10b) à une seule, l'intérêt de cette manipulation est renforcé par l'énoncé suivant.

**Exercice 6.7.** ☺ Pour une fonction convexe-concave et sous-différentiable, montrer que l'opérateur  $A$  défini par (6.11) est monotone.

On a donc ramené dans le cas convexe-concave le problème de point selle (6.6) à celui d'une inéquation variationnelle (6.12) avec un opérateur  $A$  *monotone*. Cependant, cet opérateur  $A$  “ne dérive pas d'un potentiel” (voir Remarque 5.5) car il n'est pas en général le sous-différentiel d'une fonction (qui serait nécessairement convexe puisque  $A$  est monotone) sauf dans le cas particulier où  $f$  serait de la forme  $f(x, y) = g(x) + h(y)$  (alors  $A$  est le sous-différentiel de  $(x, y) \mapsto g(x) - h(y)$ ).

### 6.1.4 Condition suffisante d'optimalité

On montre ici que si un point selle  $(u^\sharp, \lambda^\sharp)$  existe pour le Lagrangien (6.3) associé au problème (6.1) sur le sous-ensemble  $U^{\text{ad}} \times C^\star$ , alors la partie primale  $u^\sharp$  du point selle est une solution du problème (6.1). Ce que nous résumons un peu brièvement en disant que “le point selle est une condition *suffisante* d'optimalité du problème d'optimisation sous contrainte”. Notons que ce résultat est démontré sans aucune hypothèse de convexité et que c'est un résultat très général. On pouvait l'anticiper à partir du Lemme 6.1, qui nous a montré que la solution  $u^\sharp$  du problème (6.1) est aussi la solution du problème de min sup du Lagrangien, et du Lemme 6.5, qui nous a montré que, lorsqu'un point selle existe, la solution  $u^\sharp$  du min sup coïncide avec la partie primale du point selle.

**Théorème 6.8.** *Si  $(u^\sharp, \lambda^\sharp)$  est un point selle du Lagrangien (6.3) sur  $U^{\text{ad}} \times C^\star$ , alors  $u^\sharp$  est solution du problème (6.1).*

<sup>1</sup>Dans le cas d'une fonction concave  $g$ , on appelle “sous-différentiel” l'opposé du sous-différentiel de  $-g$  qui est convexe. Il faudrait peut-être parler dans ce cas de “sur-différentiel” puisque, si  $r$  appartient au “sur-différentiel” de  $g$  en  $x$ , donc  $-r \in \partial(-g)(x)$ , alors, pour tout  $z$ ,  $(-g)(z) - (-g)(x) \geq \langle -r, z - x \rangle$  et donc  $g(z) - g(x) \leq \langle r, z - x \rangle$ .

*Démonstration.* On suppose les inégalités suivantes, pour tout  $\lambda \in C^*$  et pour tout  $u \in U^{\text{ad}}$ ,

$$J(u^\sharp) + \langle \lambda, \Theta(u^\sharp) \rangle \leq J(u^\sharp) + \langle \lambda^\sharp, \Theta(u^\sharp) \rangle \leq J(u) + \langle \lambda^\sharp, \Theta(u) \rangle . \quad (6.13)$$

On considère d'abord l'inégalité de gauche ci-dessus avec les choix  $\lambda = 0$  et  $\lambda = 2\lambda^\sharp$ . Ceci permet de montrer que la quantité  $\langle \lambda^\sharp, \Theta(u^\sharp) \rangle$  est à la fois non négative et non positive, et donc qu'elle est nulle. Alors, avec la même inégalité, on en déduit que  $\langle \lambda, \Theta(u^\sharp) \rangle \leq 0$  pour tout  $\lambda \in C^*$ , et donc  $\Theta(u^\sharp) \in -C^{**} = -C$ .

On sait maintenant que le terme central dans (6.13) est égal à  $J(u^\sharp)$  et le terme de droite est inférieur ou égal à  $J(u)$  pour tout  $u \in U^{\text{ad}}$  tel que  $\Theta(u) \in -C$  (car alors  $\langle \lambda^\sharp, \Theta(u) \rangle \leq 0$ ). On a donc montré que  $u^\sharp$  réalise le minimum de  $J$  parmi tous les  $u$  satisfaisant ces conditions. C'est donc une solution de (6.1).  $\square$

## 6.2 Le cas convexe, condition nécessaire et suffisante

### 6.2.1 Hypothèses de convexité

On va introduire, pour la formulation (6.1), les hypothèses du “cas convexe”. La fonction coût  $J$  est supposée comme au chapitre précédent convexe, s.c.i. et sous-différentiable.

En ce qui concerne la fonction  $\Theta : \mathcal{U} \rightarrow \mathcal{C}$ , on va faire une hypothèse de convexité adaptée au fait que cette application n'est pas à valeurs réelles, mais à valeurs dans un espace de Hilbert  $\mathcal{C}$  muni d'un “cône positif”  $C$  induisant une relation d'ordre.

**Définition 6.9.** On dit que l'application  $\Theta : \mathcal{U} \rightarrow \mathcal{C}$  est  $C$ -convexe si, pour tout  $u$  et  $v$  dans  $\mathcal{U}$  et tout  $\alpha$  dans  $[0, 1]$ ,

$$\Theta(\alpha u + (1 - \alpha)v) - \alpha\Theta(u) - (1 - \alpha)\Theta(v) \in -C . \quad (6.14)$$

On comparera cette définition à la définition (4.2) des fonctions convexes à valeurs réelles. On vérifie facilement que la contrainte (6.1b) définit bien un ensemble convexe de  $u$  admissibles, et qui plus est, fermé si  $\Theta$  est continue, ce que nous supposons.

En nous tournant maintenant vers les questions de sous-différentiabilité, on appellera “ $C$ -sous-gradient” de  $\Theta$  au point  $u$  tout élément  $\theta$  de  $\mathcal{L}(\mathcal{U}, \mathcal{C})$  (application linéaire continue de  $\mathcal{U}$  dans  $\mathcal{C}$ ) telle que

$$\forall v, \quad \Theta(v) - \Theta(u) - \theta(v - u) \in C , \quad (6.15)$$

à comparer à (4.15). On notera  $\partial\Theta(u)$  l'ensemble de ces applications linéaires que l'on appellera “sous-différentiel” de  $\Theta$  en  $u$ .

On observe que la définition (6.14) signifie que  $\Theta$  est *affine* lorsque  $C = \{0\}$ , ce que nous avons indiqué comme un moyen d'obtenir des contraintes égalité. Enfin, on note que lorsque  $\lambda \in C^*$ , la fonction à valeurs réelles  $u \mapsto \langle \lambda, \Theta(u) \rangle$  est convexe au sens ordinaire (voir (4.2)). En supposant toujours que  $\lambda \in C^*$ , il est facile de montrer que

$$\partial_u \langle \lambda, \Theta(u) \rangle \supset \{r \in \mathcal{C} \mid r = \theta^*(\lambda), \theta \in \partial\Theta(u)\} . \quad (6.16)$$

Lorsque l'égalité a lieu, on dit que  $\Theta$  est “régulièrement sous-différentiable” (voir [2, Chap. 3, Remark 1.8] à ce sujet). Dans le cas où  $\Theta$  est différentiable, il n'y a par contre aucune difficulté.



## 6.2.2 Point selle du Lagrangien et multiplicateurs de Kuhn et Tucker

L'existence de multiplicateurs de Lagrange-Kuhn-Tucker est intimement liée à l'existence d'un point selle du Lagrangien. C'est ce que précise le lemme suivant.

**Lemme 6.10.** *L'inégalité de gauche dans (6.13) est équivalente au fait que*

$$\lambda^\sharp \in C^* , \quad \Theta(u^\sharp) \in -C , \quad \langle \lambda^\sharp , \Theta(u^\sharp) \rangle = 0 . \quad (6.17a)$$

*Sous l'hypothèse que  $\Theta$  est continue et régulièrement sous-différentiable, l'inégalité de droite dans (6.13) est équivalente à*

$$u^\sharp \in U^{\text{ad}} , \quad \exists r^\sharp \in \partial J(u^\sharp) , \quad \exists \theta^\sharp \in \partial \Theta(u^\sharp) : \quad \forall u \in U^{\text{ad}} , \quad \langle r^\sharp + (\theta^\sharp)^*(\lambda^\sharp) , u - u^\sharp \rangle \geq 0 . \quad (6.17b)$$

*Démonstration.* Le fait que l'inégalité de gauche du point selle entraîne (6.17a) a déjà été démontrée dans la première partie de la démonstration du Théorème 6.8. La réciproque est immédiate puisque, si (6.17a) est vraie, alors, pour tout  $\lambda \in C^*$ ,  $\langle \lambda , \Theta(u^\sharp) \rangle \leq 0 = \langle \lambda^\sharp , \Theta(u^\sharp) \rangle$ .

La fonction  $u \mapsto \langle \lambda^\sharp , \Theta(u) \rangle$  étant continue (grâce aux hypothèses faites), le sous-différentiel de la somme de cette fonction et de  $J$  est égal à la somme des sous-différentiels ; de plus, le sous-différentiel de  $\langle \lambda^\sharp , \Theta(\cdot) \rangle$  est obtenu par le second membre de (6.16). Alors, l'équivalence de l'inégalité de droite de (6.13) avec (6.17b) découle immédiatement des conditions d'optimalité du type (5.5d).  $\square$

**Remarque 6.11.** En se référant aux conditions d'optimalité du type (5.7), on peut remplacer (6.17b) par

$$u^\sharp \in U^{\text{ad}} , \quad \exists r^\sharp \in \partial J(u^\sharp) : \quad \forall u \in U^{\text{ad}} , \quad \langle r^\sharp , u - u^\sharp \rangle + \langle \lambda^\sharp , \Theta(u) - \Theta(u^\sharp) \rangle \geq 0 , \quad (6.17c)$$

ce qui a l'avantage de ne faire référence à aucune notion de sous-différentiabilité de  $\Theta$ .

Les conditions (6.17a) sont bien sûr à mettre en rapport avec (5.13a) et (6.17b) est à mettre en correspondance avec (5.16) (ou (5.19) dans le cas différentiable). Par conséquent, l'existence d'une solution au problème d'optimisation sous contrainte et de multiplicateurs de Lagrange-Kuhn-Tucker d'une part, d'un point selle du Lagrangien associé d'autre part (sur les sous-ensembles appropriés) sont deux choses équivalentes (dans le cas de la programmation convexe évidemment). Cette observation a l'implication suivante.

**Corollaire 6.12.** *On suppose que le problème convexe (6.1) admet une solution  $u^\sharp$  et un multiplicateur  $\lambda^\sharp$  associé à la contrainte (6.1b), c'est-à-dire vérifiant avec  $u^\sharp$  les conditions de Kuhn-Tucker (6.17). Si la solution primale  $u^\sharp$  n'est pas unique, alors l'ensemble des multiplicateurs  $\lambda^\sharp$  (eux-mêmes possible-ment non uniques) ne dépend pas de la solution  $u^\sharp$  particulière choisie.*

*Démonstration.* On vient de voir que si  $(u^\sharp, \lambda^\sharp)$  est une paire primale-duale optimale pour (6.1), elle est aussi un point-selle du Lagrangien associé sur le sous-ensemble  $U^{\text{ad}} \times C^*$ . L'Exercice 6.6 a montré que les parties primales et duales du point selle (lorsque celui-ci n'est pas unique) peuvent être appariées d'une façon quelconque. Par conséquent, les multiplicateurs optimaux pour une solution primale particulière sont aussi les multiplicateurs optimaux (c'est-à-dire, vérifient aussi les conditions de Kuhn-Tucker) pour toute autre solution de (6.1).  $\square$

### 6.2.3 Application d'un théorème général d'existence de point selle

On a vu au §6.1.4, que, sans aucune hypothèse de convexité, l'existence d'un point selle du Lagrangien constitue une condition *suffisante* d'optimalité du problème d'optimisation sous contraintes en ce qui concerne la partie primale de ce point selle. Par ailleurs, au §6.2.2, l'existence d'une solution primale et de multiplicateurs de Lagrange-Kuhn-Tucker pour un problème d'optimisation sous contraintes a été mise en correspondance directe, dans le cas convexe uniquement, avec l'existence d'un point selle du Lagrangien. Par conséquent, dans le cas convexe, lorsqu'une solution primale existe, et lorsque l'existence de multiplicateurs peut être assurée, l'existence d'un point selle du Lagrangien devient aussi une condition *nécessaire* d'optimalité.

Cependant, l'existence de multiplicateurs associés à une solution primale n'a été établie au Théorème 5.20 que pour le cas d'un nombre fini de contraintes égalité (affines) et inégalité (convexes). Pour traiter le cas plus général de contraintes (6.1b) en dimension infinie, on peut suivre diverses voies. L'une d'entre elles passe par le recours à des théorèmes généraux d'existence de point selle de fonctions convexes-concaves sur des ensembles convexes fermés. C'est ce type de démarche que nous allons examiner ici (pour en mesurer les limites).

**Théorème 6.13.** *Soit  $f : X \times Y \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$  une fonction convexe-concave s.c.i.-s.c.s..<sup>2</sup> Soit  $X$  et  $Y$  deux sous-ensembles convexes fermés de  $X$  et  $Y$ , respectivement. On suppose aussi qu'il existe  $x^\circ \in X$  tel que*

$$\lim f(x^\circ, y) = -\infty \text{ lorsque } \|y\| \rightarrow +\infty \text{ et } y \in Y, \quad (6.18a)$$

*et qu'il existe  $y^\circ \in Y$  tel que*

$$\lim f(x, y^\circ) = +\infty \text{ lorsque } \|x\| \rightarrow +\infty \text{ et } x \in X, \quad (6.18b)$$

*ou bien  $Y$ , respectivement  $X$ , est borné. Alors, il existe un point selle de  $f$  sur  $X \times Y$ .*

Voir par exemple [2, Chap. 2, Corollary 3.8] pour cet énoncé.

Examinons comment on peut utiliser un tel théorème d'existence de point selle pour le Lagrangien associé au problème (6.1) (avec évidemment  $X = U^{\text{ad}}$  et  $Y = C^*$ ). Le Lagrangien (6.3) est une fonction affine, donc concave, de  $\lambda$ . Avec les hypothèses de convexité faites au §6.2.1 sur  $J$  et  $\Theta$ , on a vu que c'est une fonction convexe de  $u$  pour tout  $\lambda \in C^*$ . Pour l'application du Théorème 6.13, on complètera donc la définition (6.3) de  $L$  de la façon suivante :

$$L(u, \lambda) \stackrel{\text{def}}{=} \begin{cases} J(u) + \langle \lambda, \Theta(u) \rangle & \text{si } \lambda \in C^*, \\ -\infty & \text{sinon.} \end{cases} \quad (6.19)$$

En ce qui concerne la continuité, les hypothèses que  $J$  est s.c.i. et que  $\Theta$  est continue sont suffisantes pour assurer le caractère s.c.i. par rapport à  $u$ . Par rapport à  $\lambda$ , on a affaire à une fonction affine continue sur le convexe fermé  $C^*$  et le passage à  $-\infty$  lorsque  $\lambda$  sort de  $C^*$  assure le caractère s.c.s.. Il nous reste à examiner comment on peut assurer les hypothèses (6.18).

- Si  $U^{\text{ad}}$  est bornée, l'hypothèse (6.18b) est assurée. Sinon, on fait l'hypothèse que  $J$  est coercive sur  $U^{\text{ad}}$  (voir Définition 5.2). Notons que dans le contexte du problème (6.1), cette hypothèse est plus forte que dans celui du problème (5.1). En effet, l'ensemble admissible est ici  $U^{\text{ad}} \cap U_{\Theta}^{\text{ad}}$  (voir (6.2)) et supposer  $J$  coercive sur  $U^{\text{ad}}$  est donc plus fort que ce que nécessite l'existence d'une solution du problème (6.1). Mais c'est le prix que l'on paie pour la technique Lagrangienne qui en quelque sorte "élimine" les contraintes (6.1b). Moyennant cette hypothèse, (6.18b) se trouve vérifiée avec le choix  $\lambda^\circ = 0$ .

<sup>2</sup>On rappelle que "s.c.s." est l'abréviation pour "semi-continue supérieurement" et qu'une fonction est s.c.s. si son opposée est s.c.i.. Ici, "s.c.i.-s.c.s." signifie s.c.i. par rapport à la variable  $x$  et s.c.s. par rapport à la variable  $y$ .

- Comme  $C^*$  n'est pas borné, il faut assurer l'existence d'un  $u^\circ$  tel que l'équivalent de la condition (6.18a) soit assurée. Cela impose d'abord que  $u^\circ \in \mathcal{T}(J)$ . On suppose de plus que  $u^\circ$  vérifie l'hypothèse de Slater :

$$\exists u^\circ \in U^{\text{ad}} \cap \mathcal{T}(J) : \Theta(u^\circ) \in -\check{C}, \quad (6.20)$$

qui est la contre-partie de (5.15). Il est facile de voir qu'alors  $L(u^\circ, \lambda)$  tend vers  $-\infty$  lorsque  $\lambda$  part à l'infini en restant dans  $C^*$ .

Cependant, l'hypothèse (6.20) n'a de sens que lorsque  $\check{C}$  est non vide, ce qui exclut de pouvoir inclure dans les contraintes globalement représentées par (6.1b) des contraintes égalité (voir la discussion à ce sujet au §6.1.1).

Moyennant donc les hypothèses ci-dessus dont nous venons de mesurer le caractère un peu restrictif, le Théorème 6.13 assure l'existence d'un point selle du Lagrangien sur  $U^{\text{ad}} \times C^*$ . La discussion ci-dessus a au moins l'intérêt de montrer que l'hypothèse de Slater (et donc de "qualification des contraintes") joue en quelque sorte pour les variables duales un rôle symétrique de l'hypothèse de coercivité pour les variables primales.

### 6.3 Un point de vue géométrique sur la dualité

La section précédente a montré la difficulté de l'approche de la dualité qui consiste à essayer de prouver l'existence d'un point selle du Lagrangien associé au problème (6.1) en s'appuyant sur les théorèmes généraux d'existence de points selle. On va revenir ici à une démarche plus intrinsèquement liée aux problèmes d'optimisation sous contraintes (dans leur formulation (6.1b)) et également plus géométrique. L'intérêt de cette démarche est de conduire à l'introduction de deux fonctions qui jouent un rôle fondamental dans cette théorie de la dualité en optimisation. L'existence de multiplicateurs (associés à une solution primale supposée exister) et leur interprétation marginaliste en découleront naturellement.

#### 6.3.1 La fonction "perturbation"

On introduit le sous-ensemble suivant :

$$S \stackrel{\text{def}}{=} \{(\theta, y) \in \mathcal{C} \times \mathbb{R} \mid \exists u \in U^{\text{ad}}, \exists c \in C, \exists z \in \mathbb{R}_+ : \theta = \Theta(u) + c, y = J(u) + z\}. \quad (6.21)$$

On va voir que ce sous-ensemble est l'épigraphe d'une certaine fonction appelée "fonction perturbation". Pour le voir, on doit se placer à une "abscisse"  $\theta$  fixée dans  $\mathcal{C}$  (donc  $\Theta(u) = \theta - c$ , pour un certain  $c \in C$  et un certain  $u \in U^{\text{ad}}$ , autrement dit  $\Theta(u) \in \theta - C$ ), puis chercher le point le plus bas verticalement. Cela correspond évidemment à prendre  $z$  (intervenant dans (6.21)) égal à 0 et à choisir le  $u \in U^{\text{ad}}$  tel que  $\Theta(u) \in \theta - C$  et tel que  $J(u)$  est le plus petit possible. Autrement dit, cela revient à résoudre le problème (6.1), si toutefois  $\theta = 0$ , et le problème "perturbé" par  $\theta$  au second membre des contraintes (6.1b) en général. Ce "bord inférieur" du sous-ensemble  $S$  est le graphe de la fonction "perturbation" ( $S$  étant son épigraphe) et cette fonction, notée  $\varphi$ , peut être ainsi définie :

$$\varphi(\theta) \stackrel{\text{def}}{=} \min \{J(u) \mid u \in U^{\text{ad}}, \Theta(u) \in \theta - C\}. \quad (6.22)$$

La résolution de (6.1) revient donc à calculer  $\varphi(0)$ .

**Exemple 6.14.** Considérons le problème "perturbé"

$$\min_{u \in \mathbb{R}} (u + 1)^2 \text{ sous } u = \theta \text{ ou } u \leq \theta.$$

La Figure 6.2 montre l'ensemble  $S$  et la fonction  $\varphi$  pour ces deux cas (à gauche pour le cas égalité, à droite pour le cas inégalité). On observe que la fonction  $\varphi$  est convexe et qu'elle est non décroissante dans le cas inégalité. C'est ce que va montrer le lemme suivant.

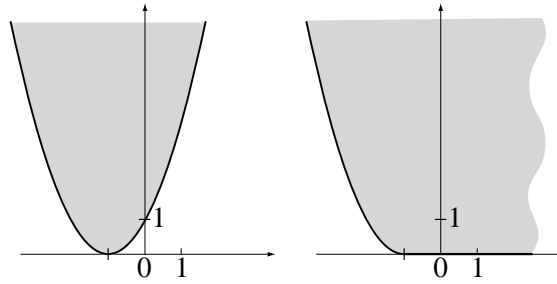


Figure 6.2: La fonction perturbation  $\varphi$  et son épigraphe  $S$  pour des contraintes égalité ou inégalité

**Lemme 6.15.** *Sous les hypothèses de convexité et de semi-continuité inférieure, respectivement de continuité, formulées au §6.2.1 pour  $J$  et  $\Theta$  (ainsi évidemment que le caractère convexe fermé de  $U^{\text{ad}}$  et  $C$ ), le sous-ensemble  $S$  est convexe fermé et donc la fonction  $\varphi$  est convexe s.c.i.. Elle est de plus  $C$ -non croissante, ce qui signifie que*

$$\theta' - \theta \in C \Rightarrow \varphi(\theta') \leq \varphi(\theta), \quad (6.23)$$

et en conséquence, tout  $r \in \partial\varphi(\theta)$  appartient à  $-C^*$ .<sup>3</sup>

*Démonstration.* On ne montrera que la convexité de  $S$ , le caractère fermé pouvant être montré en s'inspirant de la démonstration du Théorème 4.30. Soit  $(\theta_i, y_i) \in S$ ,  $i = 1, 2$ , il faut montrer que toute combinaison convexe de ces deux points appartient encore à  $S$ . Il existe  $u_i \in U^{\text{ad}}$ ,  $c_i \in C$ ,  $z_i \geq 0$  tels que  $\theta_i = \Theta(u_i) + c_i$ ,  $y_i = J(u_i) + z_i$ . Alors, pour  $\alpha \in [0, 1]$ ,

$$\alpha\theta_1 + (1 - \alpha)\theta_2 = \Theta(\alpha u_1 + (1 - \alpha)u_2) + \alpha c_1 + (1 - \alpha)c_2 + t$$

avec

$$t = \alpha\Theta(u_1) + (1 - \alpha)\Theta(u_2) - \Theta(\alpha u_1 + (1 - \alpha)u_2).$$

On observe que  $t \in C$  par (6.14), et qu'alors  $\alpha\theta_1 + (1 - \alpha)\theta_2$  est de la forme  $\Theta(u_3) + c_3$ , avec  $u_3 = \alpha u_1 + (1 - \alpha)u_2 \in U^{\text{ad}}$  (par convexité de  $U^{\text{ad}}$ ) et  $c_3 = \alpha c_1 + (1 - \alpha)c_2 + t$  du fait que  $C$  est un cône convexe fermé (voir en particulier (3.7)).

Le même type de raisonnement s'applique à  $\alpha y_1 + (1 - \alpha)y_2$  pour montrer qu'il est de la forme  $J(u_3) +$  un scalaire non négatif. En conséquence,  $\alpha(\theta_1, y_1) + (1 - \alpha)(\theta_2, y_2)$  appartient bien à  $S$ .

Pour montrer la  $C$ -non croissance de  $\varphi$ , il suffit de montrer que si  $(\theta, y)$  appartient à  $\mathfrak{U}(\varphi) = S$ , alors  $(\theta', y)$  lui appartient aussi dès que  $\theta' - \theta \in C$ , ce qui est immédiat d'après la définition de  $S$ . Enfin, par l'inégalité du sous différentiel, si  $r \in \partial\varphi(\theta)$ ,  $\langle r, \theta' - \theta \rangle \leq 0$  dès que  $\theta' - \theta \in C$ , ce qui montre que  $r \in -C^*$ .  $\square$

En appliquant le Lemme 6.1 au problème d'optimisation sous contrainte définissant  $\varphi$  (voir (6.22)), on peut donner l'expression alternative suivante de cette fonction :

$$\varphi(\theta) = \inf_{u \in U^{\text{ad}}} \sup_{\lambda \in C^*} (J(u) + \langle \lambda, \Theta(u) - \theta \rangle). \quad (6.24)$$

<sup>3</sup>On notera que lorsque  $C = \{0\}$ , ces dernières affirmations sont triviales et ne nous apprennent rien.

### 6.3.2 Multiplicateurs et sous-différentiel de la fonction perturbation

On va d'abord montrer que  $\lambda^\sharp$  est un multiplicateur optimal associé à la contrainte (6.1b) pour le problème (6.1) (supposé admettre une solution  $u^\sharp$ ) si et seulement si  $-\lambda^\sharp \in \partial\varphi(0)$ . Cela confirme en particulier l'interprétation marginaliste des multiplicateurs évoquée au §5.2.4.

**Lemme 6.16.** *Soit  $(u^\sharp, \lambda^\sharp)$  un point selle du Lagrangien associé à (6.1) sur  $U^{\text{ad}} \times C^*$ . Alors,  $-\lambda^\sharp$  est un sous-gradient de la fonction  $\varphi$  au point 0. Réciproquement, si  $\partial\varphi(0)$  est non vide (et cela suppose bien sûr que  $0 \in \nabla(\varphi)$ , donc qu'il existe une solution  $u^\sharp$  au problème (6.1), que  $\varphi(0) = J(u^\sharp)$  et que  $\Theta(u^\sharp) \in -C$ ), tout élément dans  $-\partial\varphi(0)$  constitue avec  $u^\sharp$  un point selle du Lagrangien associé.*

*Démonstration.* Considérons l'inégalité de droite du point selle :

$$\forall u \in U^{\text{ad}}, \quad J(u^\sharp) + \langle \lambda^\sharp, \Theta(u^\sharp) \rangle \leq J(u) + \langle \lambda^\sharp, \Theta(u) \rangle.$$

D'une part on sait que, grâce à l'autre inégalité du point selle (ou les écarts complémentaires), le membre de gauche de l'inégalité ci-dessus est égal à  $J(u^\sharp)$  qui n'est autre que  $\varphi(0)$ . D'autre part, considérons dans le membre de droite les  $u \in U^{\text{ad}}$  satisfaisant  $\Theta(u) \in \theta - C$ , donc  $\Theta(u) = \theta - c$  pour  $c \in C$ . Notons que  $\langle \lambda^\sharp, c \rangle \geq 0$  et donc,

$$\varphi(0) \leq J(u) + \langle \lambda^\sharp, \theta - c \rangle \leq J(u) + \langle \lambda^\sharp, \theta \rangle.$$

En prenant dans le dernier membre le minimum pour tous les  $u \in U^{\text{ad}}$  tels que  $\Theta(u) \in \theta - C$ , on fait apparaître  $\varphi(\theta)$  et on obtient l'inégalité

$$\varphi(0) \leq \varphi(\theta) + \langle \lambda^\sharp, \theta \rangle, \quad (6.25)$$

vraie pour tout  $\theta$ , y compris si  $\theta$  est tel que le problème d'optimisation sous contrainte perturbée par  $\theta$  n'a pas de solution (car alors  $\varphi(\theta) = +\infty$ ). L'inégalité prouve que  $-\lambda^\sharp \in \partial\varphi(0)$ .

Réciproquement, supposons (6.25) que l'on peut réécrire, en fonction de (6.24) et de l'hypothèse que  $\varphi(0) = J(u^\sharp)$  de la façon suivante :

$$J(u^\sharp) \leq \inf_{u \in U^{\text{ad}}} \sup_{\lambda \in C^*} (J(u) + \langle \lambda, \Theta(u) - \theta \rangle) + \langle \lambda^\sharp, \theta \rangle.$$

Cette inégalité implique que, pour tout  $u \in U^{\text{ad}}$  et tout  $\theta \in \mathcal{C}$ ,

$$J(u^\sharp) \leq \sup_{\lambda \in C^*} (J(u) + \langle \lambda, \Theta(u) - \theta \rangle) + \langle \lambda^\sharp, \theta \rangle.$$

En prenant  $\theta = \Theta(u)$ , on obtient que, pour tout  $u \in U^{\text{ad}}$ ,

$$J(u^\sharp) \leq J(u) + \langle \lambda^\sharp, \Theta(u) \rangle. \quad (6.26)$$

Par ailleurs, on sait d'après le Lemme 6.15 que  $-\lambda^\sharp \in \partial\varphi(0)$  appartient à  $-C^*$ , donc  $\lambda^\sharp \in C^*$  et comme évidemment  $\Theta(u^\sharp) \in -C$ , alors  $\langle \lambda^\sharp, \Theta(u^\sharp) \rangle \leq 0$ . Mais, en prenant  $u = u^\sharp$  dans (6.26), on montre aussi l'inégalité inverse, donc finalement  $\langle \lambda^\sharp, \Theta(u^\sharp) \rangle = 0$ . Cette égalité d'écarts complémentaires ainsi que l'inégalité (6.26) et le fait que  $\Theta(u^\sharp) \in -C$  et  $\lambda^\sharp \in C^*$  établissent la propriété de point selle (voir en particulier (6.17a)).  $\square$

On va maintenant s'intéresser, moyennant en particulier une nouvelle hypothèse de qualification des contraintes, à l'existence d'un sous-gradient de  $\varphi$  en 0 et par conséquent, en vertu du Lemme 6.16, à l'existence de multiplicateurs optimaux pour le problème (6.1). Pour montrer l'existence d'un sous-gradient de  $\varphi$  en 0, on va s'appuyer sur la condition nécessaire et suffisante énoncée au Lemme 4.50, et donc étudier le cône normal à  $S = \Psi(\varphi)$  au point  $(0, J(u^\sharp))$ .

**Théorème 6.17.** *On suppose qu'il existe une solution  $u^\sharp$  au problème (6.1)<sup>4</sup> et on fait l'hypothèse de qualification des contraintes suivante :*

$$0 \in \overbrace{\Theta(U^{\text{ad}}) + C}^{\checkmark} . \quad (6.27a)$$

Alors il existe un multiplicateur optimal  $\lambda^\sharp$  et donc un point selle au Lagrangien.

**Remarque 6.18.** Dans le cas de contraintes égalité (c'est-à-dire lorsque  $C = \{0\}$ ), (6.27a) prend la forme

$$0 \in \overbrace{\Theta(U^{\text{ad}})}^{\checkmark} . \quad (6.27b)$$

D'une manière générale, (6.27) exprime que la contrainte (6.1b) peut être satisfaite avec un  $u \in U^{\text{ad}}$  non seulement sous sa forme originale ( $\theta = 0$ ) mais pour toute perturbation  $\theta \in \mathcal{C}$  suffisamment petite et dans toutes les directions de l'espace  $\mathcal{C}$ .

*Démonstration.* On a vu au Lemme 3.54 que le cône orthogonal à un convexe fermé en un point frontière n'est pas réduit à 0 si on est en dimension finie, et qu'en dimension infinie (voir commentaire après le lemme) il faut de plus supposer que l'intérieur du convexe n'est pas vide. Ici on s'intéresse à l'ensemble  $S$  et au point frontière  $(0, \varphi(0))$  avec  $\varphi(0) = J(u^\sharp)$ . Sous l'hypothèse (6.27), il est clair que  $\check{S} \neq \emptyset$  (il suffit de considérer le produit cartésien d'une boule ouverte de  $\mathcal{C}$  centrée en 0 et contenue dans  $\Theta(U^{\text{ad}}) + C$  avec la demi-droite "verticale" ouverte  $]J(u^\sharp), +\infty[$  : cet ouvert est contenu dans  $S$ ).

Pour montrer que  $\partial\varphi(0)$  n'est pas vide, il faut et il suffit, selon le Lemme 4.50, de montrer que le cône normal dont on vient de parler contient un vecteur avec une composante "verticale"  $-1$ . Pour cela, il suffit de montrer qu'il ne contient pas de vecteurs de la forme  $(\mu, 0) \in \mathcal{C} \times \mathbb{R}$  et que tous les vecteurs qu'il contient appartiennent à  $(-C^*) \times \mathbb{R}_-$  : dans ces conditions, n'étant pas vide, il contient obligatoirement un vecteur de la forme  $(-\lambda^\sharp, -1)$  avec  $\lambda^\sharp \in C^*$  (par normalisation) et la démonstration sera complète.

Un vecteur  $(\mu, \eta) \in \mathcal{C} \times \mathbb{R}$  dans  $S_{(0, J(u^\sharp))}^\perp$  satisfait par définition :

$$\forall (\theta, y) \in S, \quad \langle \mu, \theta \rangle + \eta(y - J(u^\sharp)) \leq 0 ,$$

soit

$$\forall u \in U^{\text{ad}}, \forall c \in C, \forall z \geq 0, \quad \langle \mu, \Theta(u) + c \rangle + \eta(J(u) + z - J(u^\sharp)) \leq 0 .$$

Si  $C = \{0\}$ , l'affirmation que  $\mu \in -C^* = \mathcal{C}^*$  est triviale, et sinon, la supposition que  $\mu \notin -C^*$  conduit à une contradiction car alors le premier membre peut prendre des valeurs arbitrairement grandes et positives par des choix appropriés de  $c \in C$ . De même, la supposition que  $\eta > 0$  conduit à une contradiction pour le même type de raison en choisissant bien  $z$ . Donc nécessairement  $\mu \in -C^*$  et  $\eta \leq 0$ .

Supposons alors que  $\eta = 0$ . Par conséquent, on soit supposer que  $\mu \neq 0$ . On aurait donc

$$\forall u \in U^{\text{ad}}, \forall c \in C, \quad \langle \mu, \Theta(u) + c \rangle \leq 0 . \quad (6.28)$$

Or, grâce à (6.27), on sait qu'il existe  $\varepsilon$  assez petit tel que la boule ouverte  $B(0, \varepsilon) \subset \Theta(U^{\text{ad}}) + C$ , autrement dit, pour tout vecteur  $d \in \mathcal{C}$  de norme inférieure à  $\varepsilon$ , il existe un  $u \in U^{\text{ad}}$  et un  $c \in C$  tel que  $d = \Theta(u) + c$ . On peut utiliser cette affirmation pour  $d$  égal à  $\alpha\mu/\|\mu\|$  pour tout  $\alpha < \varepsilon$ , et on fait apparaître une contradiction dans (6.28) puisqu'alors

$$\alpha \|\mu\| \leq 0$$

(on rappelle que  $\mu \neq 0$ ). Ceci achève la preuve. □

<sup>4</sup>Cela nécessite donc que  $\neg(J) \cap U^{\text{ad}} \cap U_{\Theta}^{\text{ad}} \neq \emptyset$ .

On retiendra donc que sous l'hypothèse de qualification des contraintes (6.27), la fonction de pénalisation du problème (6.1) est sous-différentiable en 0 et que tout sous-gradient, changé de signe, constitue, avec une solution primale un point selle du Lagrangien associé. On étudie maintenant le rapport étroit qui existe entre la fonction perturbation et le Lagrangien.

### 6.3.3 Fonction perturbation, transformée de Fenchel et Lagrangien

Le théorème ci-après établit le lien entre fonction perturbation et Lagrangien. Plus exactement, on va faire apparaître une nouvelle fonction "duale" liée au Lagrangien et qui joue, avec la fonction perturbation, un rôle fondamental dans la théorie de la dualité. Il s'agit de la fonction  $\psi : \mathcal{C} \rightarrow \mathbb{R} \cup \{-\infty\}$  ainsi définie :

$$\psi(\lambda) \stackrel{\text{def}}{=} \begin{cases} \inf_{u \in U^{\text{ad}}} L(u, \lambda) & \text{si } \lambda \in C^*, \\ -\infty & \text{sinon.} \end{cases} \quad (6.29)$$

Notons que la deuxième clause de cette définition est inutile si on adopte la définition (6.19) du Lagrangien. Notons aussi que la formule (6.29) définit en tout état de cause une fonction concave et s.c.s. (comme enveloppe inférieure d'une famille fonctions affines). Notons enfin que lorsque le Lagrangien admet un point selle  $(u^\sharp, \lambda^\sharp)$  sur  $U^{\text{ad}} \times C^*$ , le coût optimal du problème (6.1) qui est toujours égal au  $\min_{u \in U^{\text{ad}}} \sup_{\lambda \in C^*} L(u, \lambda)$  (voir Lemme 6.1) est aussi égal au  $\max_{\lambda \in C^*} \inf_{u \in U^{\text{ad}}} L(u, \lambda)$  qui n'est autre que  $\max_{\lambda \in \mathcal{C}} \psi(\lambda)$ . Ce maximum est atteint évidemment pour tout  $\lambda^\sharp$  (partie duale du point selle). On va confirmer toutes ces observations et le fait que  $-\lambda^\sharp \in \partial\varphi(0)$  comme corollaire du théorème suivant.

**Théorème 6.19.** *On a la formule :*

$$\forall \lambda \in \mathcal{C}, \quad \varphi^F(\lambda) = -\psi(-\lambda). \quad (6.30)$$

*Démonstration.* La transformée de Fenchel de  $\varphi$  est égale par définition à

$$\begin{aligned} \varphi^F(\lambda) &= \sup_{\theta \in \mathcal{C}} (\langle \lambda, \theta \rangle - \varphi(\theta)) \\ &= \sup_{\theta \in \mathcal{C}} \left( \langle \lambda, \theta \rangle - \min \{ J(u) \mid u \in U^{\text{ad}}, \Theta(u) \in \theta - C \} \right) \end{aligned}$$

d'après (6.22),

$$= \max_{u \in U^{\text{ad}}} \sup_{\theta \in C + \Theta(u)} (\langle \lambda, \theta \rangle - J(u))$$

parce que  $-\min(\cdot) = \max(-\cdot)$  et parce que les sup ou max commutent entre eux,

$$= \begin{cases} +\infty & \text{si } \lambda \notin -C^*, \\ \sup_{u \in U^{\text{ad}}} (\langle \lambda, \Theta(u) \rangle - J(u)) & \text{si } \lambda \in -C^*, \end{cases}$$

car dans le premier cas,  $\theta$  peut partir à l'infini dans la "bonne" direction en faisant croître arbitrairement l'expression, et dans le second cas, on a au contraire intérêt à le prendre le plus petit possible, c'est-à-dire égal à  $\Theta(u)$ . Une nouvelle utilisation de  $\max(-\cdot) = -\min(\cdot)$  et le rapprochement avec permet (6.29) de montrer finalement la formule (6.30).  $\square$

De ce théorème résultent en corollaire plusieurs résultats qui recourent en grande partie ce que nous savions déjà.

- D'après la formule (4.8), on a pour tout  $\lambda \in C^*$  (c'est même vrai, mais trivialement, pour  $\lambda \in \mathcal{C}^*$ ) :

$$J(u^\sharp) - \psi(\lambda) = \varphi(0) + \varphi^F(-\lambda) \geq 0 ,$$

ce qui signifie que  $\psi$  est majorée par la valeur optimale du problème (6.1). Si  $\lambda^\sharp$  est un multiplicateur optimal pour ce problème, l'égalité est atteinte, donc  $\psi$  atteint son maximum en tout  $\lambda^\sharp$ .

- Lorsque l'égalité a lieu, on déduit d'après le Théorème 4.71 que  $-\lambda^\sharp \in \partial\varphi(0)$ , ce que nous savions déjà.
- Du même théorème, on déduit dualement que  $0 \in \partial\varphi^F(-\lambda^\sharp)$ . D'après (6.30), le sous-différentiel de  $\varphi^F$  en un certain  $\lambda$  est égal à l'opposé du sous-différentiel de  $-\psi$  (qui est aussi convexe) au point  $-\lambda$ , ce qui est aussi par définition, le "sur-différentiel" de la fonction concave  $\psi$  au même point  $-\lambda$ . De tout ceci, on déduit que  $0$  appartient au "sur-différentiel" de  $\psi$  au point  $\lambda^\sharp$ , ce qui n'a rien de surprenant puisque l'on vient de voir que  $\psi$  atteint son maximum en  $\lambda^\sharp$ .
- La dernière observation concerne l'expression du "sur-différentiel" de  $\psi$ . Afin d'éviter les difficultés de langage, on va plutôt s'intéresser au sous-différentiel de  $\Psi \stackrel{\text{def}}{=} -\psi$  ( $\Psi$  est convexe et est en fait telle que  $\Psi(\lambda) = \varphi^F(-\lambda)$ ). Avec (6.29), on voit que

$$\Psi = \overline{\Psi} + I_{C^*} \text{ avec } \overline{\Psi}(\lambda) \stackrel{\text{def}}{=} - \inf_{u \in U^{\text{ad}}} L(u, \lambda), \quad \forall \lambda \in \mathcal{C}^* .$$

La formule (4.24) conduit à s'intéresser aux sous-différentiels des deux fonctions suivantes.

- L'application  $\lambda \mapsto \overline{\Psi}(\lambda) = \sup_{u \in U^{\text{ad}}} (-L)(u, \lambda)$  est une fonction marginale du type de celle étudiée à l'Exercice 4.67. En vertu des résultats de cet exercice, on peut dire que, parmi les éléments du sous-différentiel de  $\overline{\Psi}$ , on trouve les éléments de la forme

$$-\Theta(\widehat{u}(\lambda)) = \frac{\partial(-L)(\widehat{u}(\lambda), \lambda)}{\partial\lambda} \text{ avec } \widehat{u}(\lambda) \in \arg \min_{u \in U^{\text{ad}}} L(u, \lambda) .$$

- Par ailleurs, l'Exercice 4.52 a montré que  $\partial I_{C^*}(\lambda)$  est égal au cône orthogonal à  $C^*$  au point  $\lambda$  et l'Exercice 3.25 nous indique que ce cône orthogonal est égal au cône orthogonal en  $0$ , c'est-à-dire  $-C$ , intersecté avec les éléments orthogonaux à  $\lambda$ , donc de produit scalaire nul avec  $\lambda$ .

Observons que si  $(u^\sharp, \lambda^\sharp)$  est un point selle de  $L$ ,

- l'inégalité de droite du point selle dit précisément que  $u^\sharp = \widehat{u}(\lambda^\sharp)$ , donc  $-\Theta(u^\sharp)$  est un sous-gradient de  $\overline{\Psi}$  au point  $\lambda^\sharp$ ,
- et l'inégalité de gauche du point selle, équivalente à (6.17a), indique que  $\Theta(u^\sharp)$  est un sous-gradient de  $I_{C^*}$  au point  $\lambda^\sharp$ .

Autrement dit, on a retrouvé, par les deux inégalités du point selle, que

$$0 \in \partial\overline{\Psi}(u^\sharp) + \partial I_{C^*}(\lambda) \subset \partial(\overline{\Psi} + I_{C^*})(u^\sharp) = \partial\Psi(\lambda^\sharp)$$

et donc que  $\Psi$  est minimale (donc  $\psi$  est maximale) en  $\lambda^\sharp$ .



### 6.3.4 Explication géométrique des résultats

#### 6.3.4.1 Le cas convexe

Considérons, dans l'espace  $\mathcal{C} \times \mathbb{R}$ , l'équation  $y + \langle \lambda, \theta \rangle = a$ , où les variables sont  $\theta \in \mathcal{C}$  et  $y \in \mathbb{R}$ , où  $\lambda \in \mathcal{C}$  (ou  $\mathcal{C}^*$ ) est un paramètre, et où  $a \in \mathbb{R}$  est un autre paramètre. Cette équation décrit un hyperplan et on peut appeler  $-\lambda$  sa "pente" (ou plus généralement le vecteur de ses "coefficients directeurs") tandis que  $a$  est l'"ordonnée à l'origine" (valeur de  $y$  lorsque  $\theta = 0$ ).

L'espace  $\mathcal{C} \times \mathbb{R}$  est celui dans lequel a été défini le sous-ensemble  $S$  (voir (6.21)) et on peut s'intéresser à l'intersection de l'hyperplan avec ce sous-ensemble. Cela revient à ne considérer dans l'équation de l'hyperplan que les points de la forme  $\theta = \Theta(u) + c$  et  $y = J(u) + z$  pour  $u \in U^{\text{ad}}$ ,  $c \in \mathcal{C}$  et  $z \geq 0$ . On va s'intéresser en fait à "abaissier" le plus possible l'hyperplan en le déplaçant parallèlement à lui-même (donc en gardant  $\lambda$  fixe mais en cherchant l'ordonnée à l'origine  $a$  la plus petite possible) sous la contrainte de conserver une intersection non vide avec  $S$ . La situation est représentée sur la Figure 6.3 (partie gauche) dans le cas où  $\mathcal{C} = \mathbb{R}$ .

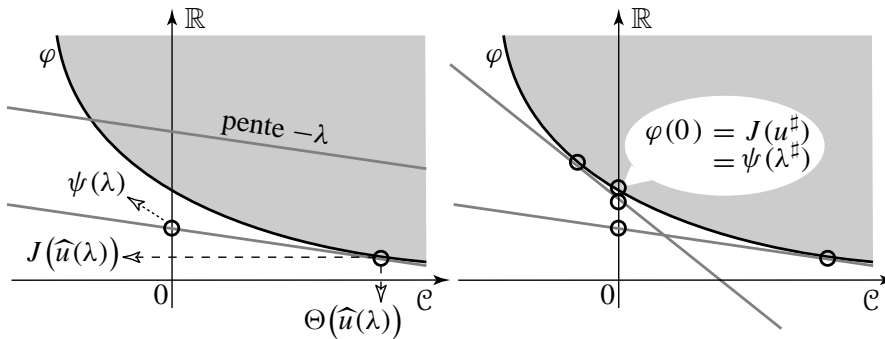


Figure 6.3: La dualité par l'image

Lorsque le bord inférieur de  $S$  (c'est-à-dire le graphe de  $\varphi$ ) a l'allure décroissante représentée sur la figure (lorsque  $C = \mathbb{R}_+$  dans le cas où  $\mathcal{C} = \mathbb{R}$  dessiné ici, par opposition au cas où  $C = \{0\}$  — revoir la Figure 6.2), on voit que la plus petite ordonnée à l'origine possible reste finie (n'est pas égale à  $-\infty$ ) si la "pente" n'est pas positive, donc si  $\lambda \in \mathbb{R}_+$  (plus généralement si  $\lambda \in \mathcal{C}^*$ ). Si de plus  $S$  a une forme convexe (notamment quand les hypothèses du §6.2.1 sont satisfaites), la plus petite ordonnée à l'origine sera obtenue dans la position limite où la droite (plus généralement l'hyperplan) "tangente" l'ensemble  $S$  sur son bord inférieur (plus généralement, constitue un hyperplan d'appui de  $S$ ), le point de contact étant de la forme  $(\Theta(u), J(u))$  (c'est-à-dire qu'on a intérêt à prendre  $z = 0$  et  $c = 0$  pour minimiser  $a$  en restant dans  $S$  à partir du moment où  $\lambda$  a le bon signe ou la bonne orientation). Autrement dit, le plus petit  $a$  associé à un  $\lambda$  donné est obtenu en minimisant  $y + \langle \lambda, \theta \rangle = J(u) + \langle \lambda, \Theta(u) \rangle$  pour  $u \in U^{\text{ad}}$ . Finalement, calculer la plus petite ordonnée à l'origine correspondant à un  $\lambda$  fixé, c'est minimiser  $L(u, \lambda)$  pour  $u \in U^{\text{ad}}$  et c'est donc calculer  $\psi(\lambda)$ . Cette constatation est d'ailleurs conforme avec le fait démontré précédemment que  $\psi$  a un rapport avec  $\varphi^F$ .

On va chercher maintenant à obtenir l'ordonnée à l'origine  $\psi(\lambda)$  de l'hyperplan d'appui de  $S = \mathcal{U}(\varphi)$  la plus élevée possible en agissant sur la "pente"  $-\lambda$ . On peut voir sur la partie droite de la Figure 6.3 que la plus haute ordonnée à l'origine que l'on peut obtenir par ce procédé est la valeur  $\varphi(0)$  (qui est égale, comme on le sait à  $J(u^\sharp)$ ). Cela se démontre facilement : puisque  $y = \langle -\lambda, \theta \rangle + \psi(\lambda)$  définit un hyperplan d'appui de  $\mathcal{U}(S)$ , le graphe de  $\varphi$  se trouve au dessus de tout hyperplan d'appui et donc

$$\forall \theta, \forall \lambda, \varphi(\theta) \geq \langle -\lambda, \theta \rangle + \psi(\lambda) \Rightarrow \forall \lambda, \varphi(0) \geq \psi(\lambda).$$

Pour obtenir la valeur maximale de l'ordonnée à l'origine  $\psi(\lambda)$ , il faut donc choisir  $\lambda = \lambda^\sharp$  telle que  $\psi(\lambda^\sharp) = \varphi(0) = J(u^\sharp)$ . Le point de contact de cet hyperplan avec  $S$  a comme coordonnées  $(0, J(u^\sharp))$ ,

c'est-à-dire que  $\langle \lambda^\sharp, \Theta(u^\sharp) \rangle = 0$ .

Enfin, la récapitulation de toutes ces opérations montre que l'on a réalisé  $\max_{\lambda \in C^*} \inf_{u \in U^{\text{ad}}} L(u, \lambda)$  et que l'on ne parvient à atteindre  $\varphi(0) = J(u^\sharp) = \min_{u \in U^{\text{ad}}} \sup_{\lambda \in C^*} L(u, \lambda)$  que dans le cas où ces deux valeurs sont égales, ce qui se réalise dans le cas convexe.

### 6.3.4.2 Cas non convexes : remarques diverses sur le saut de dualité

Il est donc intéressant de visualiser ce qui peut se passer dans les cas non convexes. En fait, toutes les explications concernant l'interprétation du calcul de  $\max_{\lambda \in C^*} \inf_{u \in U^{\text{ad}}} L(u, \lambda)$  comme l'ordonnée à l'origine la plus grande possible d'un hyperplan d'appui de  $S$  restent valables, sauf que les hyperplans d'appui ne peuvent suivre que le contour de  $\overline{\text{co}}(S)$ . En premier lieu, si  $\varphi$  n'admet pas de minorante affine, l'enveloppe convexe de  $S$  risque d'être triviale (exercice : faire un dessin d'un tel cas). En admettant que cette minorante affine existe, la Figure 6.4 montre deux situations possibles :

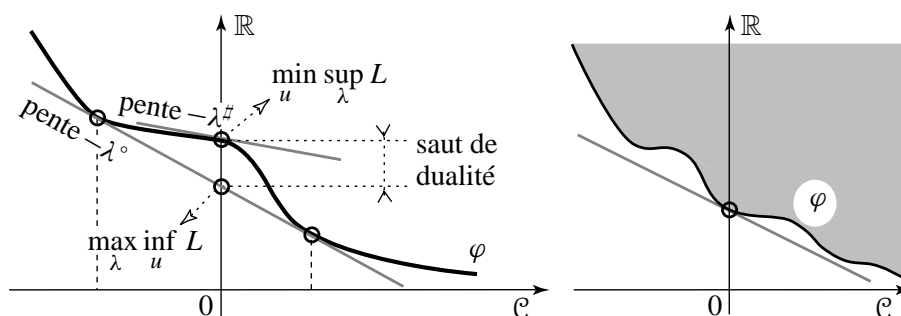


Figure 6.4: Saut de dualité possible, mais pas certain, dans les cas non convexes

- à gauche,  $\varphi$  ne coïncide pas en 0 avec son enveloppe convexe ; dans ce cas, la plus grande ordonnée à l'origine des hyperplans d'appui ne parvient pas à atteindre la valeur  $\varphi(0)$  ; l'une de ces valeurs étant le  $\max_{\lambda} \inf_u L$  et l'autre le  $\min_u \sup_{\lambda} L$ , on visualise le “saut de dualité” qui est l'écart entre ces deux valeurs ;
- à droite, bien que  $\varphi$  soit non convexe, elle est localement convexe autour de 0, ce qui est une condition nécessaire, mais pas suffisante, pour que  $\varphi$  coïncide en 0 avec son enveloppe convexe ; en tout cas, cette dernière condition est suffisante pour que le saut de dualité soit nul, et qu'il existe donc, malgré l'absence de convexité globale, un point selle au Lagrangien.

Toutes ces constatations graphiques peuvent faire l'objet de démonstrations rigoureuses, de même que la remarque suivante. Dans le cas de la partie gauche de la Figure 6.4 où le saut de dualité est non nul, on voit que l'hyperplan d'appui “optimal” (ayant la plus grande ordonnée à l'origine) touche l'ensemble  $S$  ou le graphe de  $\varphi$  en “deux” points (en général, on peut simplement affirmer que le point de contact ne sera pas unique). On a vu (Figure 6.3) que les points de contact d'un hyperplan d'appui de “pente”  $-\lambda$  sont déterminés par les  $\widehat{u}(\lambda) \in \arg \min_{u \in U^{\text{ad}}} L(u, \lambda)$ . Par conséquent, dans le cas considéré, cet  $\arg \min$  n'est pas réduit à un point. Plus exactement, les “abscisses”  $\Theta(\widehat{u}(\lambda))$  des points de contact ne sont pas réduites à une seule valeur. Or on a vu (reconsidérer la formule (4.26) ou plus exactement la formule duale pour le cas concave) que chaque  $\Theta(\widehat{u}(\lambda))$  constitue un “sur-gradient” de  $\psi$  en  $\lambda$ . On en déduit donc que dans le cas d'un saut de dualité, la fonction  $\psi$  n'est pas différentiable en son maximum  $\lambda^\circ$ ,  $-\lambda^\circ$  étant la pente du “meilleur” hyperplan d'appui.

Observons enfin que ce  $\lambda^\circ$ , qui ne constitue plus avec  $u^\sharp$  un point selle de  $L$ , mais qui continue à résoudre le problème du  $\max_{\lambda} \inf_u L$ , perd aussi son interprétation marginaliste. La pente  $-\lambda^\circ$  sur la gauche de la Figure 6.4 n'a aucun rapport avec  $-\lambda^\sharp$ , pente de la fonction  $\varphi$  autour de 0, et c'est bien

sûr cette dernière qui représente la variation marginale du coût optimal par rapport à une perturbation  $\theta$  du second membre des contraintes. On peut montrer que le (ou les) multiplicateur(s)  $\lambda^\#$  qui garde(nt) l'interprétation marginaliste est (sont) celui (ceux) qui intervien(nen)t dans les conditions d'optimalité de Kuhn et Tucker (théorie locale), ces conditions restant des conditions nécessaires (mais pas suffisantes en général dans les cas non convexes).

## 6.4 Résumons nous

Dans ce chapitre, on a abordé les problèmes de minimisation d'une fonction coût sous un "nombre (potentiellement) infini de contraintes", c'est-à-dire plus précisément, en introduisant des "fonctions contraintes" à valeurs dans un espace de Hilbert avec la contrainte que ces valeurs appartiennent à un "cône négatif" (éventuellement réduit à  $\{0\}$  pour obtenir des contraintes égalité).

On a aussi introduit un outil fondamental, le *Lagrangien*, qui fait intervenir les variables de décision du problème d'optimisation (dites "variables primales") et des multiplicateurs des contraintes "dualisées" (dites "variables duales"). Sans aucune hypothèse particulière, notamment de convexité, on démontre facilement que si le Lagrangien admet un point selle sur les sous-ensembles appropriés, alors toute partie primale de ce point selle constitue une solution du problème d'optimisation sous contraintes. En fait, on démontre dans ce cadre très général que la résolution du problème de minimisation sous contraintes est équivalent à un min sup du Lagrangien, et l'existence du point selle permet de ramener ce problème de min sup à celui d'un max inf.

Revenant maintenant au cadre convexe, on peut chercher à établir une réciproque de la condition suffisante d'optimalité (existence d'un point selle), c'est-à-dire une condition nécessaire d'optimalité en termes d'existence de multiplicateurs associés qui constitueraient, avec la solution primale du problème d'optimisation sous contraintes, un point selle du Lagrangien. Une première voie qui a été explorée à cet effet est le recours à des théorèmes généraux d'existence de point selle pour des fonctions convexes-concaves, avec des hypothèses techniques permettant d'assurer en quelque sorte la *coercivité* de ces fonctions du côté du min et du côté du max. On a vu que cela réussit partiellement pour le cas du Lagrangien en ce sens que la présence de contraintes égalité (conduisant à un cône négatif d'intérieur vide) est difficilement couvert par cette approche (la coercivité du côté dual étant obtenue par une hypothèse de *qualification des contraintes* "à la Slater" impossible à formuler dans le cas de contraintes égalité).

On s'est alors tourné vers une approche plus géométrique et plus directement liée au problème d'optimisation sous contraintes en introduisant un deuxième outil fondamental : la "fonction perturbation" qui revient à perturber le second membre des contraintes dualisées et à s'intéresser au coût optimal correspondant. On établit que les multiplicateurs optimaux sont, *au signe près*, les sous-gradients de cette fonction perturbation au point 0 (perturbation nulle, donc problème original). Établir l'existence de multiplicateurs optimaux revient à établir la non vacuité de ce sous-différentiel, ce qui demande une hypothèse technique qui redonne une hypothèse de *qualification des contraintes* cette fois-ci plus intrinsèque et valable pour un mélange arbitraire de contraintes égalité et inégalité.

De plus, cette approche a l'avantage de procurer de façon naturelle un fondement mathématique à l'*interprétation marginaliste* des multiplicateurs déjà évoquée au chapitre précédent de façon plus heuristique.

Dans la dernière partie du chapitre, on établit un lien direct entre les deux outils fondamentaux introduit précédemment, à savoir, la fonction perturbation d'une part, et le Lagrangien, ou plus exactement le min de ce Lagrangien par rapport aux variables primales — une fonction de variables duales donc — d'autre part. C'est la transformée de Fenchel qui permet de relier (avec au passage quelques signes —) ces deux fonctions. Alors, par toute une série de remarques utilisant abondamment les résultats des chapitres précédents sur les ensembles et fonctions convexes, on retrouve la plupart des résultats établis plus tôt dans ce chapitre. Ceci suggère d'ailleurs qu'il y aurait de nombreuses autres façons de rédiger ce chapitre et de choisir un cheminement vers les résultats présentés ici.

On termine par la compréhension géométrique de ces résultats en termes d'hyperplans d'appui de la fonction perturbation, compréhension qui doit aller au delà du cas convexe en permettant de visualiser les "sauts de dualité" susceptibles de se produire dans les cas non convexes. Le chapitre suivant évoquera un remède possible à ces sauts de dualité.

## 6.5 Corrigé des exercices

**Corrigé de l'Exercice 6.6** On suppose les inégalités suivantes, pour tout  $x \in X$  et  $y \in Y$  :

$$f(x_1^\sharp, y) \leq f(x_1^\sharp, y_1^\sharp) \leq f(x, y_1^\sharp), \quad (6.31a)$$

$$f(x_2^\sharp, y) \leq f(x_2^\sharp, y_2^\sharp) \leq f(x, y_2^\sharp). \quad (6.31b)$$

Alors,

$$\begin{aligned} f(x_1^\sharp, y_2^\sharp) &\leq f(x_1^\sharp, y_1^\sharp) && \text{(par (6.31a)–gauche)} \\ &\leq f(x_2^\sharp, y_1^\sharp) && \text{(par (6.31a)–droite)} \\ &\leq f(x_2^\sharp, y_2^\sharp) && \text{(par (6.31b)–gauche)} \\ &\leq f(x_1^\sharp, y_2^\sharp) && \text{(par (6.31b)–droite)} \end{aligned}$$

ce qui établit (6.9). Ensuite, pour tout  $x \in X$ ,  $y \in Y$ ,

$$\begin{aligned} f(x_1^\sharp, y) &\leq f(x_1^\sharp, y_1^\sharp) && \text{(par (6.31a)–gauche)} \\ &= f(x_1^\sharp, y_2^\sharp) && \text{(par (6.9))} \\ &= f(x_2^\sharp, y_2^\sharp) && \text{(par (6.9))} \\ &\leq f(x, y_2^\sharp) && \text{(par (6.31b)–droite)} \end{aligned}$$

ce qui montre que  $(x_1^\sharp, y_2^\sharp)$  est un point selle de  $f$  sur  $X \times Y$ . On démontre de la même façon que  $(x_2^\sharp, y_1^\sharp)$  est aussi un point selle.

**Corrigé de l'Exercice 6.7** Soit  $a_i \in A(z_i)$ ,  $i = 1, 2$ . Si  $z_i = (x_i, y_i)$ , alors  $a_i = (r_i, -s_i)$  avec  $r_i \in \partial_x f(x_i, y_i)$  et  $s_i \in \partial_y f(x_i, y_i)$ . On a donc

$$\begin{aligned} f(x_2, y_1) - f(x_1, y_1) &\geq \langle r_1, x_2 - x_1 \rangle, \\ f(x_1, y_2) - f(x_2, y_2) &\geq \langle r_2, x_1 - x_2 \rangle, \\ -f(x_1, y_2) + f(x_1, y_1) &\geq \langle -s_1, y_2 - y_1 \rangle, \\ -f(x_2, y_1) + f(x_2, y_2) &\geq \langle -s_2, y_1 - y_2 \rangle. \end{aligned}$$

L'addition de ces quatre inégalités conduit à

$$\begin{aligned} \langle a_2 - a_1, z_2 - z_1 \rangle &= \langle r_2 - r_1, x_2 - x_1 \rangle - \langle s_2 - s_1, y_2 - y_1 \rangle \\ &\geq 0. \end{aligned}$$

# Chapitre 7

## Lagrangien augmenté

### 7.1 Justification géométrique

Poursuivons ici la discussion de la fin du chapitre précédent.

#### 7.1.1 Cas non convexes

Revenons à la Figure 6.4. On constate que si, au lieu d'utiliser des objets "plats" comme les hyperplans, on avait utilisé des objets "concaves", on aurait eu plus de chance d'atteindre la valeur  $\varphi(0)$  "cachée dans un creux" (voir Figure 7.1). Pour mettre en œuvre cette idée, il suffit de remplacer les hyperplans par des

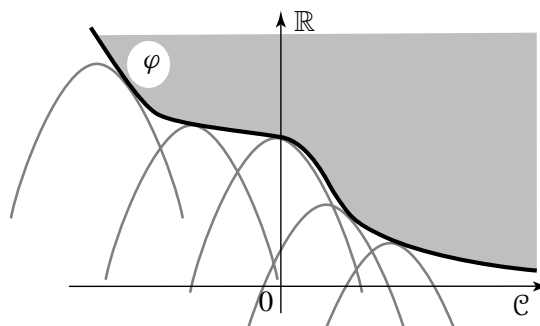


Figure 7.1: Exploration de  $\varphi$  par des paraboles

"paraboles renversées" d'équation (dans l'espace  $\mathbb{C} \times \mathbb{R}$ ) :

$$y = a - \langle \lambda, \theta \rangle - b \|\theta\|^2 / 2 . \quad (7.1)$$

Dans cette équation,

- $b$  est une constante positive qui règle la concavité du paraboléide (plus  $b$  est grand, plus ce paraboléide est "pointu"),
- $a$  est à nouveau l'ordonnée à l'origine (ordonnée atteinte en  $\theta = 0$ ),
- $-\lambda \in \mathbb{C}^*$  est désormais la "pente" à l'origine (valeur de la dérivée première en  $\theta = 0$ ).

On procède comme au chapitre précédent : pour  $b$  fixé,

- on commence par fixer  $\lambda$  et on cherche l'ordonnée à l'origine la plus petite possible telle que la parabole garde au moins un point commun avec  $S = \Psi(\varphi)$  ; cela revient à calculer :

$$\begin{aligned} a_\lambda &= \inf_{u \in U^{\text{ad}}} \inf_{c \in C} \inf_{z \geq 0} (J(u) + z + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2) \\ &= \inf_{u \in U^{\text{ad}}} \inf_{c \in C} (J(u) + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2) ; \end{aligned} \quad (7.2)$$

- on cherche ensuite à faire *monter* cette ordonnée à l'origine  $a_\lambda$  le plus possible en ajustant la pente à l'origine  $-\lambda$ .

La Figure 7.2 montre la position de deux paraboles, correspondant deux valeurs distinctes de  $b$ , à l'issue de ce processus, l'une pour laquelle un saut de dualité subsiste ( $b$  pas assez grand) et l'autre pour laquelle le saut de dualité a disparu.

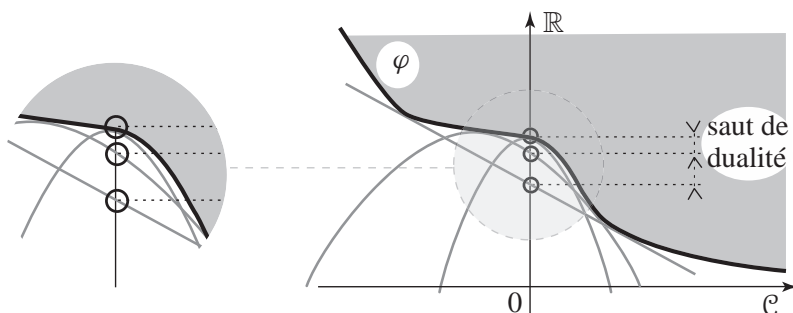


Figure 7.2: Réduction ou disparition du saut de dualité

D'une manière générale, on ne parviendra à réduire le saut de dualité à 0 que si  $\varphi$  admet une minorante quadratique et si  $b$  a été choisi assez grand. Mais ces conditions ne sont pas suffisantes. La Figure 7.3 représente ce qui peut se passer par exemple pour un problème "en nombres entiers", c'est-à-dire lorsque  $u$  parcourt un ensemble discret. Dans ce cas,  $\varphi$  peut présenter des décrochements verticaux dus au saut de la solution optimale du problème d'optimisation sous contraintes d'une valeur entière à une autre lorsque le second membre des contraintes varie. Cependant, même dans ce cas, le saut de

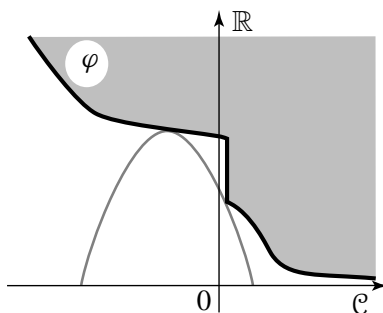


Figure 7.3: Cas de problèmes en variables entières ou mixtes

dualité se trouve réduit, sinon annulé, par l'utilisation de cette technique.

Une autre cause d'échec peut tenir au fait que  $\varphi$  n'admet pas de minorante quadratique, par exemple parce que sa décroissance à l'infini est plus rapide que quadratique. Dans ce cas, on peut utiliser une technique de "localisation" : ceci revient à restreindre l'ensemble  $U^{\text{ad}}$  à  $U^{\text{ad}} \cap \overline{B}(u^\circ, \rho)$ , où  $u^\circ$  est une estimation de la solution, et  $\rho$  est choisi pour que la boule  $\overline{B}(u^\circ, \rho)$  contienne (avec une bonne probabilité) une solution  $u^\dagger$ . L'effet de cette restriction de  $U^{\text{ad}}$  est de réduire l'ensemble  $S$  en "coupant les branches infinies". On récupère alors (éventuellement) une minorante quadratique (voir Figure 7.4).

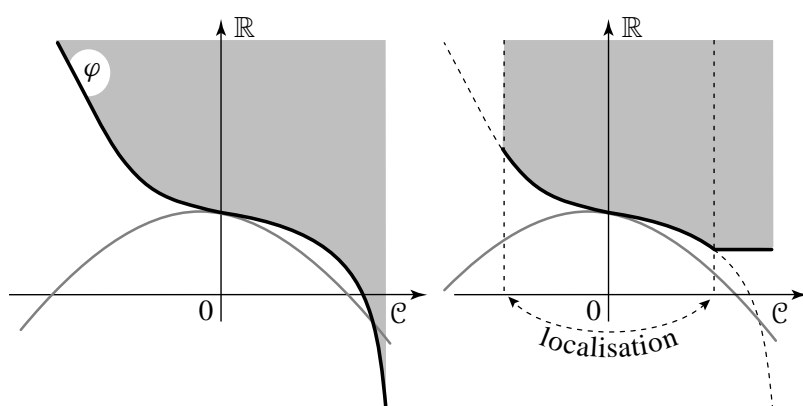


Figure 7.4: Effet de la localisation : après localisation, il existe une minorante quadratique

### 7.1.2 Cas convexe

Même dans le cas convexe où le problème du saut de dualité ne se pose pas, la technique précédente conserve un intérêt. Pour le comprendre, il faut considérer des cas convexes mais pas *strictement* convexes. Imaginons qu’au voisinage de 0,  $\varphi$  présente une partie “plate” (voir Figure 7.5). Si on utilise des hyperplans d’appui, les points de contact seront multiples et cela a deux conséquences.

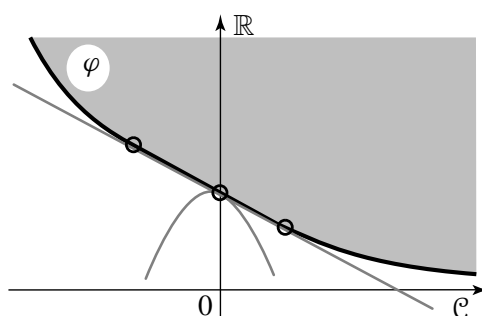


Figure 7.5: Multiplicité ou unicité des points de contact dans le cas non strictement convexe

- D’une part, comme on l’a expliqué au chapitre précédent, cela signifie que l’arg  $\min_{u \in U^{\text{ad}}} L(u, \lambda^\sharp)$  n’est pas unique, et même que la valeur prise par  $\Theta(\widehat{u}(\lambda^\sharp))$  lorsque  $\widehat{u}(\lambda^\sharp)$  parcourt cet arg min n’est pas unique, ce qui signifie que la fonction  $\psi$  (voir (6.29)) n’est pas différentiable en son maximum  $\lambda^\sharp$ . Du point de vue algorithmique,<sup>1</sup> c’est un inconvénient sérieux.
- D’autre part, et ceci est encore plus grave, ces valeurs multiples de  $\Theta(\widehat{u}(\lambda^\sharp))$  signifient probablement que tous les éléments de l’arg  $\min_u L(u, \lambda^\sharp)$  ne sont pas des solutions, et même ne sont pas des solutions admissibles (c’est-à-dire respectant la contrainte (6.1b)), même si  $\lambda^\sharp$  est un multiplicateur optimal. L’exemple ci-après va montrer que ce phénomène est en effet possible. Donc, si l’on imagine que l’on a trouvé, par un procédé quelconque, une valeur optimale  $\lambda^\sharp$  du multiplicateur, on ne pourra pas trouver une solution primale  $u^\sharp$  en se contentant de calculer *n’importe quelle* solution du problème  $\min_{u \in U^{\text{ad}}} L(u, \lambda^\sharp)$ .

**Exemple 7.1.** On considère le problème dans  $\mathbb{R}$  :

$$\min_{u \in [-1, 1]} (-u) \text{ sous } u \leq 0 .$$

<sup>1</sup>Question qui n’est pas abordée dans ce cours, mais on peut imaginer une méthode cherchant à maximiser  $\psi$  sur  $C^*$

Le Lagrangien est égal à  $L(u, \lambda) = -u + \lambda u = (\lambda - 1)u$ . La solution de ce problème est unique et c'est  $u^\sharp = 0$  et l'unique multiplicateur optimal est  $\lambda^\sharp = 1$ .

**Exercice 7.2.** Démontrer toutes ces affirmations en se servant par exemple de l'interprétation marginaliste des multiplicateurs et vérifier les inégalités du point selle.

Or, pour  $\lambda^\sharp = 1$ ,  $\arg \min_{u \in [-1, 1]} L(u, \lambda^\sharp)$  est égal à tout l'intervalle  $[-1, 1]$  et pourtant seule  $u = 0$  est solution.

**Exercice 7.3.** Calculer la fonction  $\psi$  pour cet exemple et montrer qu'elle n'est pas différentiable en  $\lambda^\sharp = 1$ .

Toujours dans le cadre de l'Exemple 7.1, imaginons que l'on soit capable de construire, par un algorithme itératif, une suite  $\{\lambda^k\}$  convergeant vers la valeur optimale 1. Les itérées  $\lambda^k$  seront peut-être de temps en temps inférieures à 1 et de temps en temps supérieures à 1 (tout en convergeant vers 1). Il est facile de voir que, contrairement à ce qui se passe pour  $\lambda = 1$ , pour tout autre valeur de  $\lambda$ , l' $\arg \min_{u \in [-1, 1]} L(u, \lambda)$  est réduit à une seule valeur. Plus précisément, on a :

$$\arg \min_{u \in [-1, 1]} L(u, \lambda) = \arg \min_{u \in [-1, 1]} (\lambda - 1)u = \begin{cases} -1 & \text{si } \lambda > 1, \\ 1 & \text{si } \lambda < 1. \end{cases}$$

On réalise donc que la suite  $\{u^k\}$  que l'on pourra calculer par minimisation du Lagrangien pour les itérées  $\lambda^k$  aura comme points d'accumulation  $-1$  et  $1$ . Or on cherche à converger vers  $u^\sharp = 0$  !

Ce genre d'inconvénient ne se produirait pas si l'on pouvait garantir la propriété suivante.

**Définition 7.4.** On dit qu'un Lagrangien admettant l'ensemble de points selle  $U^\sharp \times \Lambda^\sharp$  sur  $U^{\text{ad}} \times C^*$  est *stable en  $u$*  si

$$\forall \lambda^\sharp \in \Lambda^\sharp, \arg \min_{u \in U^{\text{ad}}} L(u, \lambda^\sharp) = U^\sharp.$$

D'une manière générale, on peut seulement affirmer l'inclusion du second membre dans le premier.

Tous ces ennuis vont être surmontés en utilisant des paraboles à la place des hyperplans. En retournant à la Figure 7.5, on voit que l'utilisation d'une parabole (objet strictement concave) élimine les points de contact "parasites" pour ne garder que le bon. Unicité du point de contact signifie unicité des valeurs  $\Theta(\hat{u}(\lambda^\sharp))$  (mais pas nécessairement des  $\arg \min \hat{u}(\lambda^\sharp)$  eux-mêmes). Mais on verra que cela résoud la plupart des difficultés évoquées ci-dessus et que notamment la "stabilité en  $u$ " du Lagrangien augmenté sera assurée.

## 7.2 Lagrangien augmenté et régularisation duale

Le remplacement des hyperplans par des paraboloides nous a conduit au calcul (7.2). On appelle donc "Lagrangien augmenté" associé au problème (6.1) l'expression suivante :

$$L_b(u, \lambda) \stackrel{\text{def}}{=} \inf_{c \in C} (J(u) + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2). \quad (7.3a)$$

En effet, ayant posé cette définition, le calcul de la plus petite ordonnée à l'origine  $a_\lambda$  associée à une pente à l'origine  $-\lambda$  revient à minimiser  $L_b(u, \lambda)$  en  $u \in U^{\text{ad}}$ , puis il faut réaliser le max en  $\lambda$ . On va voir que contrairement au cas des Lagrangiens "ordinaires"<sup>2</sup>, il n'est pas utile de limiter  $\lambda$  à  $C^*$  (on peut le laisser varier dans tout  $\mathbb{C}^*$ ). On peut en effet voir sur les figures que le cas des "pentes à l'origine" positives (donc des  $\lambda$  négatifs) n'a pas un comportement différent de celui des pentes de signe opposé, ce qui n'est pas du tout le cas avec des hyperplans.

<sup>2</sup>Nous appellerons désormais ainsi ceux qui ne sont pas "augmentés"



## 7.2.1 Diverses expressions et propriétés du Lagrangien augmenté

On va maintenant donner des expressions variées du Lagrangien augmenté  $L_b$ .

**Lemme 7.5.** *Le Lagrangien augmenté (7.3a) admet les expressions suivantes :*

$$L_b(u, \lambda) = J(u) + \zeta_b(\Theta(u), \lambda), \quad (7.3b)$$

avec

$$\zeta_b(\theta, \lambda) = \inf_{c \in \mathcal{C}} (\langle \lambda, \theta + c \rangle + b \|\theta + c\|^2 / 2) \quad (7.3c)$$

$$= \sup_{\mu \in C^*} (\langle \mu, \theta \rangle - \|\lambda - \mu\|^2 / 2b) \quad (7.3d)$$

$$= (\|\Pi_{C^*}(\lambda + b\theta)\|^2 - \|\lambda\|^2) / 2b \quad (7.3e)$$

$$= \langle \lambda, \theta \rangle + b \|\theta\|^2 / 2 - \|\Pi_{(-C)}(\lambda + b\theta)\|^2 / 2b. \quad (7.3f)$$

La fonction  $\zeta_b$  est convexe s.c.i. en  $\theta$  et concave s.c.s. en  $\lambda$  ; elle est aussi  $C$ -non décroissante en  $\theta$  ; enfin, elle est différentiable et l'on a

$$(\zeta_b)'_{\theta}(\theta, \lambda) = \Pi_{C^*}(\lambda + b\theta), \quad (7.3g)$$

$$(\zeta_b)'_{\lambda}(\theta, \lambda) = (\Pi_{C^*}(\lambda + b\theta) - \lambda) / b. \quad (7.3h)$$

*Démonstration.* Avec (7.3c), (7.3b) est une simple réécriture de (7.3a). L'expression (7.3c) est un problème de minimisation en  $c$  sous la contrainte  $-c \in -C$ . En vertu du Lemme 6.1, en dualisant cette contrainte avec un multiplicateur  $\mu \in C^*$ , on obtient :

$$\begin{aligned} \zeta_b(\theta, \lambda) &= \inf_{c \in \mathcal{C}} \sup_{\mu \in C^*} (\langle \lambda, \theta + c \rangle + b \|\theta + c\|^2 / 2 - \langle \mu, c \rangle) \\ &= \sup_{\mu \in C^*} \inf_{c \in \mathcal{C}} (\langle \lambda, \theta + c \rangle + b \|\theta + c\|^2 / 2 - \langle \mu, c \rangle), \end{aligned}$$

par inversion de l'inf et du sup (la contrainte  $-c \in -C$  étant affine est qualifiée — voir Lemme 5.18 — et donc l'existence d'un multiplicateur optimal et donc d'un point selle — cas convexe — est assurée). Le passage de cette dernière expression à l'expression (7.3d) résulte d'un calcul très simple puisque la minimisation en  $c \in \mathcal{C}$  est celle d'une fonction quadratique.

Pour obtenir l'expression (7.3e), on résoud par exemple le problème (7.3d) dont une condition d'optimalité est

$$\langle \theta - (\hat{\mu} - \lambda) / b, \mu - \hat{\mu} \rangle \leq 0, \quad \forall \mu \in C^*,$$

ce qui donne la solution explicite (voir (3.14)) :

$$\hat{\mu} = \Pi_{C^*}(\lambda + b\theta). \quad (7.4)$$

On reporte ensuite cette solution dans la fonction coût :

$$\begin{aligned} \langle \Pi_{C^*}(\lambda + b\theta), \theta \rangle - \|\lambda - \Pi_{C^*}(\lambda + b\theta)\|^2 / 2b &= \\ \langle \Pi_{C^*}(\lambda + b\theta), \lambda + b\theta \rangle / b - \|\lambda\|^2 / 2b - \|\Pi_{C^*}(\lambda + b\theta)\|^2 / 2b. \end{aligned}$$

Le premier terme est égal à deux fois le troisième terme en vertu de (3.18a). On obtient finalement l'expression (7.3e). L'expression (7.3f) peut être obtenue de façon similaire en conduisant des calculs

analogues à partir du problème de minimisation sous contrainte (7.3c) (la solution est  $\widehat{c} = -\Pi_{(-C)}(\lambda + b\theta)/b$ ), ou bien encore en se servant de (7.3e) et de (3.20).

En vertu de (7.3c), la fonction  $\zeta_b(\theta, \cdot)$  est l'enveloppe inférieure d'une famille de fonctions affines en  $\lambda$  : elle est donc concave s.c.s. en cette variable. En vertu de (7.3d), la fonction  $\zeta_b(\cdot, \lambda)$  est l'enveloppe supérieure d'une famille de fonctions affines en  $\theta$  : elle est donc convexe s.c.i. en cette variable.

Toujours à partir de (7.3d), la fonction sous le sup est de façon évidente une fonction  $C$ -non décroissante en  $\theta$  dès que  $\mu \in C^*$ . Il en découle immédiatement la même propriété pour  $\zeta_b(\cdot, \lambda)$ .

La différentiabilité de  $\zeta_b$  peut se voir de diverses façons à partir de ses diverses expressions. Elle tient essentiellement à la différentiabilité des fonctions sous l'inf ou le sup dans (7.3c) et (7.3d), et à l'unicité des solutions optimales  $\widehat{c}$  et  $\widehat{\mu}$  correspondantes. on peut par exemple utiliser la formule (4.29) à partir de l'expression (7.3c) pour calculer

$$\begin{aligned} (\zeta_b)'_{\theta} &= \partial(\langle \lambda, \theta + \widehat{c} \rangle + b \|\theta + \widehat{c}\|^2 / 2) / \partial \theta \\ &= \lambda + b(\theta + \widehat{c}) \\ &= \lambda + b\theta - \Pi_{(-C)}(\lambda + b\theta) \\ &= \Pi_{C^*}(\lambda + b\theta) . \end{aligned}$$

De la même façon, en utilisant les mêmes résultats transposés au cas concave à partir de l'expression (7.3d), on obtient :

$$\begin{aligned} (\zeta_b)'_{\lambda} &= \partial(\langle \widehat{\mu}, \theta \rangle - \|\lambda - \widehat{\mu}\|^2 / 2b) / \partial \lambda \\ &= (\widehat{\mu} - \lambda) / b \\ &= (\Pi_{C^*}(\lambda + b\theta) - \lambda) / b . \end{aligned}$$

□

**Remarque 7.6.** On observe la relation suivante :

$$(\zeta_b)'_{\theta}(\theta, \lambda) = b(\zeta_b)'_{\lambda}(\theta, \lambda) + \lambda . \quad (7.5)$$

**Corollaire 7.7.** *Sous les hypothèses de convexité du §6.2.1, le Lagrangien augmenté est une fonction convexe en  $u$  et concave en  $\lambda$ .*

*Démonstration.* La concavité en  $\lambda$  de  $L_b$  est une conséquence directe de celle de  $\zeta_b$ . La dépendance en  $u$  passe par celle de  $J$  supposée convexe et par celle de la composition de la fonction  $\Theta$  supposée  $C$ -convexe avec  $\zeta_b$  qui est  $C$ -non décroissante en  $\theta$  et convexe. La convexité du résultat de cette composition est une généralisation d'un résultat sur la composition de fonctions énoncé au §4.4. □

**Remarque 7.8.** Dans le cas des contraintes égalité pour lequel  $C = \{0\}$  et  $C^* = \mathcal{C}^*$ , et donc  $\Pi_{(-C)}$  est identiquement nulle et  $\Pi_{C^*}$  est l'identité, les formules pour  $L_b$  prennent la forme plus simple :

$$\begin{aligned} L_b(u, \lambda) &= L(u, \lambda) + b \|\Theta(u)\|^2 / 2 \\ &= J(u) + \langle \lambda, \Theta(u) \rangle + b \|\Theta(u)\|^2 / 2 , \end{aligned} \quad (7.6)$$

ce qui fait souvent dire que la technique du Lagrangien augmenté est un mélange de technique Lagrangienne et de pénalisation quadratique des contraintes. Cependant, cette interprétation tourne court dans le cas des contraintes inégalité, car si elle était vraie, la formule pour ce cas devrait être

$$L_b(u, \lambda) = L(u, \lambda) + b \|\Pi_{C^*}(\Theta(u))\|^2 / 2 ,$$

ce qui ne coïncide pas avec les formules (7.3).

## 7.2.2 Quelques propriétés de la régularisation de Yosida-Moreau

Dans cette section, on passe en revue quelques propriétés d'une transformation appelée "régularisation de Yosida-Moreau". On va étudier le cas convexe comme c'est traditionnel, bien que l'application des résultats qui sera faite dans la suite concerne le cas concave (il faudra donc transposer les résultats pour ce cas).

**Définition 7.9.** Étant donné une constante  $b$  positive, on appelle *régularisée de Yosida-Moreau*  $f_b$  d'une fonction  $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ , convexe, le résultat de l'inf-convolution de cette fonction avec la fonction quadratique  $bQ$  où  $Q$  est la forme quadratique bâtie sur l'identité ( $Q(x) = \|x\|^2/2$ ). Plus précisément, on a :

$$f_b(x) = \inf_{y \in \mathcal{X}} (f(y) + b \|x - y\|^2 / 2) .$$

**Théorème 7.10.** Soit  $f$  une fonction convexe s.c.i. de domaine non vide et sous-différentiable en un point. Alors sa régularisée de Yosida-Moreau  $f_b$  est convexe, partout continue et différentiable, et sa dérivée est Lipschitzienne de constante  $1/b$  (au moins).

*Démonstration.* Comme on l'a vu à l'Exercice 4.17, l'inf-convolution de deux fonctions convexes admettant une minorante affine commune est convexe. Dans le cas particulier étudié ici, et puisque  $\inf_{x \in \mathcal{X}} (b \|x\|^2 / 2 - \langle a, x \rangle)$  est fini pour tout  $a \in \mathcal{X}$ , toute "pente"  $a$  convient pour obtenir une minorante affine de  $bQ$ . Si la fonction  $f$  est sous-différentiable en au moins un point, elle admet elle-même une minorante affine qui est donc, moyennant éventuellement une translation "verticale" finie appropriée, une minorante affine de la fonction  $bQ$  également. On a donc bien trouvé une minorante affine commune. L'inf-convolution est donc une fonction convexe.

Le domaine de  $f_b$  est égal à tout  $\mathcal{X}$  dès que le domaine de  $f$  n'est pas vide (ce qui est évidemment impliqué par les hypothèses). En effet, dès que  $x_0 \in \mathcal{T}(f)$ , la fonction  $f$  est majorée par  $x \mapsto I_{\{x_0\}}(x) + f(x_0)$  et c'est donc aussi le cas pour  $f_b$  qui est majorée par la régularisée de cette fonction, à savoir, la fonction  $x \mapsto f(x_0) + b \|x - x_0\|^2 / 2$ . Comme le domaine de cette dernière fonction est tout  $\mathcal{X}$ , alors  $\mathcal{T}(f_b) = \mathcal{X}$ . En particulier,  $\widehat{\mathcal{T}}(f_b) \neq \emptyset$ , donc  $f_b$  est continue dans cet intérieur (Théorème 4.27) qui est tout  $\mathcal{X}$ .

Comme  $f_b$  est un cas particulier de fonction marginale (voir (4.7)), on peut lui appliquer la formule (4.29) pour calculer son sous-différentiel. Dans notre cas particulier,  $f_b$  est la fonction marginale de  $(x, y) \mapsto b \|x - y\|^2 / 2 + f(y)$  qui est, par rapport la variable  $y$ , strictement convexe, s.c.i. et coercive, puisque  $bQ$  est strictement convexe et qu'elle tend à l'infini vers  $+\infty$  de façon quadratique, alors que  $f$ , qui admet une minorante affine, ne peut partir à  $-\infty$  que de façon linéaire au plus. La minimisation en  $y$  est donc réalisée en une valeur unique  $\widehat{y}(x)$  pour tout  $x$ , et elle est caractérisée par :

$$0 \in b(\widehat{y}(x) - x) + \partial f(\widehat{y}(x)) .$$

L'application de la formule (4.29) conduit à

$$\partial f_b(x) = \{f'_b(x)\} = \{x - \widehat{y}(x)\} .$$

Considérons deux valeurs  $x_i$ ,  $i = 1, 2$ , les valeurs  $y_i \stackrel{\text{def}}{=} \widehat{y}(x_i)$  correspondantes et les valeurs  $r_i \in \partial f(y_i)$  telles que  $0 = b(y_i - x_i) + r_i$  (d'après la condition d'optimalité sur  $y_i$ ). Alors,

$$0 = \langle b(y_1 - x_1) + r_1 - b(y_2 - x_2) - r_2, y_2 - y_1 \rangle ,$$

d'où

$$b \langle y_1 - x_1 - y_2 + x_2, y_2 - y_1 \rangle = \langle r_1 - r_2, y_1 - y_2 \rangle \geq 0$$

puisque  $\partial f$  est monotone, puis

$$\|y_1 - x_1 - y_2 + x_2\|^2 \leq \langle y_1 - x_1 - y_2 + x_2, x_1 - x_2 \rangle$$

ce qui s'interprète comme

$$\|f'_b(x_1) - f'_b(x_2)\|^2 \leq (1/b) \langle f'_b(x_1) - f'_b(x_2), x_1 - x_2 \rangle,$$

et par l'inégalité de Cauchy-Schwarz (voir (2.15))

$$\|f'_b(x_1) - f'_b(x_2)\|^2 \leq (1/b) \|f'_b(x_1) - f'_b(x_2)\| \|x_1 - x_2\|,$$

d'où la propriété de Lipschitz de  $f'_b$  avec la constante  $1/b$ . □

**Théorème 7.11.** *La fonction  $f_b$  est partout inférieure ou égale à  $f$  mais les valeurs infimales de ces deux fonctions sont égales et leurs arg min sont égaux.*

*Démonstration.* On a

$$f_b(x) = \inf_{y \in \mathcal{X}} (b \|x - y\|^2 / 2 + f(y)) \leq f(x)$$

en prenant la valeur particulière  $y = x$ . De plus,

$$\begin{aligned} \inf_{x \in \mathcal{X}} f_b(x) &= \inf_{x \in \mathcal{X}} \inf_{y \in \mathcal{X}} (b \|x - y\|^2 / 2 + f(y)) \\ &= \inf_{y \in \mathcal{X}} \inf_{x \in \mathcal{X}} (b \|x - y\|^2 / 2 + f(y)) \end{aligned}$$

puisque'on peut commuter les deux inf,

$$= \inf_{y \in \mathcal{X}} f(y)$$

puisque l' $\inf_{x \in \mathcal{X}}$  est évidemment atteint en  $x = y$ .

Si  $x^\sharp \in \arg \min_{x \in \mathcal{X}} f(x)$ , on a donc, en fonction de ce qui vient d'être montré, que

$$f(x^\sharp) = \inf_{x \in \mathcal{X}} f(x) = \inf_{x \in \mathcal{X}} f_b(x) \leq f_b(x^\sharp) \leq f(x^\sharp),$$

donc l'inf de  $f_b$  est bien atteint en  $x^\sharp$ . Inversement, si  $x^\sharp \in \arg \min_{x \in \mathcal{X}} f_b(x)$ , on a

$$\begin{aligned} 0 &= f'_b(x^\sharp) = x^\sharp - \widehat{y}(x^\sharp) \quad (\text{optimalité de } x^\sharp) \\ \text{avec } 0 &\in \widehat{y}(x^\sharp) - x^\sharp + \partial f(\widehat{y}(x^\sharp)) \quad (\text{optimalité de } \widehat{y}(x^\sharp)), \end{aligned}$$

d'où il découle évidemment que  $0 \in \partial f(x^\sharp)$ , donc  $f$  réalise son minimum en  $x^\sharp$ . □

On voit donc que si l'on doit résoudre le problème  $\min_{x \in \mathcal{X}} f(x)$ , on peut le remplacer par le problème  $\min_{x \in \mathcal{X}} f_b(x)$  qui fournira les mêmes solutions et la même valeur optimale. De plus,  $f_b$  étant une fonction partout différentiable et de dérivée Lipschitzienne, on aurait grandement intérêt à manipuler, sur le plan des algorithmes numériques, la fonction  $f_b$  plutôt que la fonction  $f$ . Malheureusement, le calcul de  $f_b$  à partir de  $f$  est loin d'être "gratuit" puisqu'il consiste lui-même à résoudre un problème d'optimisation où  $f$  intervient déjà. Donc, sur le plan pratique, on ne semble pas avoir progressé. On verra cependant, dans le cas des Lagrangiens augmentés, que le calcul de la régularisée est analytique et que cette idée conserve de l'intérêt.

**Remarque 7.12.** On peut se poser la question de savoir ce que deviennent les résultats précédents si le problème à résoudre est un problème sous contrainte, à savoir

$$\min_{x \in X^{\text{ad}}} f(x) .$$

Pour répondre à cette question, il suffit d'appliquer la théorie précédente à  $f + I_{X^{\text{ad}}}$ . On est donc amené à définir en quelque sorte une "régularisée de  $f$  sur  $X^{\text{ad}}$ ", à savoir,

$$\begin{aligned} f_b^{X^{\text{ad}}}(x) &\stackrel{\text{def}}{=} (f + I_{X^{\text{ad}}})_b(x) \\ &= \inf_{y \in \mathcal{X}} (b \|x - y\|^2 / 2 + f(y) + I_{X^{\text{ad}}}(y)) \\ &= \inf_{y \in X^{\text{ad}}} (b \|x - y\|^2 / 2 + f(y)) . \end{aligned}$$

Ensuite, le Théorème 7.11 s'applique évidemment, ce qui signifie que

$$\inf_{x \in \mathcal{X}} f_b^{X^{\text{ad}}}(x) = \inf_{y \in X^{\text{ad}}} f(y) .$$

Il est très important de noter que c'est la régularisation qui se fait sur  $X^{\text{ad}}$  puis la minimisation de la fonction régularisée sur tout  $\mathcal{X}$ . On pourra se convaincre sur des exemples que le schéma inverse :

- calcul de la régularisée de  $f$  sur tout  $\mathcal{X}$ ,
- puis minimisation de cette régularisée sur  $X^{\text{ad}}$ ,

*ne fonctionne pas.*

### 7.2.3 Régularisation du Lagrangien par rapport $\lambda$ et conséquences

Revenons à l'expression (7.3d) de la fonction  $\zeta_b$ . Elle montre que la fonction  $\lambda \mapsto \zeta_b(\theta, \lambda)$  peut être considérée comme la régularisée de la fonction  $\lambda \mapsto \langle \lambda, \theta \rangle$ , vue comme une fonction *concave*,<sup>3</sup> pour la constante  $1/b$  et sur le sous-ensemble  $C^*$ . Une conséquence immédiate de cette observation est que le Lagrangien augmenté  $L_b$  est aussi la régularisée du Lagrangien ordinaire  $L$  par rapport à la variable duale  $\lambda$  pour la constante  $1/b$  et sur le sous-ensemble  $C^*$ . Cette remarque va avoir de multiples et intéressantes conséquences.

**Théorème 7.13.** *Le Lagrangien ordinaire  $L$  sur  $U^{\text{ad}} \times C^*$  et le Lagrangien augmenté  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$  ont les mêmes points selle.*

<sup>3</sup>Par conséquent, il faut transposer toutes les définitions de la régularisation au cas concave.

*Démonstration.* Considérons un point selle  $u^\sharp, \lambda^\sharp$  de  $L$  sur  $U^{\text{ad}} \times C^*$  :

$$\forall \lambda \in C^*, L(u^\sharp, \lambda) \leq L(u^\sharp, \lambda^\sharp), \quad (7.7a)$$

$$\forall u \in U^{\text{ad}}, L(u^\sharp, \lambda^\sharp) \leq L(u, \lambda^\sharp). \quad (7.7b)$$

L'inégalité (7.7a) indique que la fonction  $\lambda \mapsto L(u^\sharp, \lambda)$  atteint son maximum sur  $C^*$  au point  $\lambda^\sharp$ . Le Théorème 7.11 nous apprend que l'on peut en dire autant pour sa régularisée  $\lambda \mapsto L_b(u^\sharp, \lambda)$ , la maximisation portant maintenant sur tout  $\mathcal{C}^*$ . En particulier, les valeurs maximales  $L(u^\sharp, \lambda^\sharp)$  et  $L_b(u^\sharp, \lambda^\sharp)$  sont égales. Ceci nous fournit donc l'inégalité de gauche du point selle de  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$ .

Considérons maintenant l'inégalité (7.7b). Le fait que  $\zeta_b(\theta, \cdot)$  soit la régularisée de  $\langle \cdot, \theta \rangle - I_{C^*}(\cdot)$  (cas concave), ou directement la formule (7.3d), montre que  $\zeta_b(\theta, \lambda) \geq \langle \lambda, \theta \rangle - I_{C^*}(\lambda)$  et en particulier

$$\forall u, \zeta_b(\Theta(u), \lambda^\sharp) \geq \langle \lambda^\sharp, \Theta(u) \rangle$$

puisque  $I_{C^*}(\lambda^\sharp) = 0$ . Il découle de cette inégalité et de (7.7b) l'inégalité de droite pour le point selle de  $L_b$ . Finalement,  $(u^\sharp, \lambda^\sharp)$  est un point selle de  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$ .

Réciproquement, soit  $(u^\sharp, \lambda^\sharp)$  un point selle de  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$ . Pour obtenir (7.7a), on utilise exactement l'argument symétrique du précédent comme l'autorise le Théorème 7.11 qui fonctionne dans les deux sens. On observe aussi qu'en raison du fait que  $\lambda \mapsto L_b(u^\sharp, \lambda)$  atteint son maximum en  $\lambda$  sur tout  $\mathcal{C}^*$  au point  $\lambda^\sharp$  et du fait que cette fonction est différentiable en  $\lambda$ , sa dérivée en  $\lambda$  au point  $\lambda^\sharp$  est nulle. Alors,

$$(L_b)'_{\lambda}(u^\sharp, \lambda^\sharp) = (\zeta_b)'_{\lambda}(\Theta(u^\sharp), \lambda^\sharp) = 0 \Rightarrow (\zeta_b)'_{\theta}(\Theta(u^\sharp), \lambda^\sharp) = \lambda^\sharp,$$

la dernière implication résultant de la relation (7.5). Cette remarque débouche sur l'autre inégalité du point selle pour  $L$ . En effet, admettons pour l'instant (voir Exercice 7.14 ci-après) que l'inégalité de droite du point selle pour  $L_b$  soit équivalente à la condition d'optimalité :

$$\forall u \in U^{\text{ad}}, J(u) - J(u^\sharp) + \langle (\zeta_b)'_{\theta}(\Theta(u^\sharp), \lambda^\sharp), \Theta(u) - \Theta(u^\sharp) \rangle \geq 0.$$

Si on tient compte dans cette inéquation variationnelle de la remarque précédente, on obtient l'inégalité de droite du point selle pour  $L$  et la démonstration est complète.  $\square$

**Exercice 7.14.** À titre de généralisation des conditions d'optimalité du §5.1.4, considérer le problème

$$\min_{u \in U^{\text{ad}}} \left( J(u) + G(H(u)) \right),$$

où  $U^{\text{ad}} \subset \mathcal{U}$  est convexe fermé,  $J : \mathcal{U} \rightarrow \overline{\mathbb{R}}$  est une fonction convexe s.c.i.  $H : \mathcal{U} \rightarrow \mathcal{C}$  est une application à valeurs dans  $\mathcal{C}$  muni d'un "cône positif"  $C$  (convexe fermé),  $H$  étant  $C$ -convexe et continue,  $G : \mathcal{C} \rightarrow \mathbb{R}$  est une fonction convexe,  $C$ -non décroissante, continue et différentiable, et montrer qu'une solution optimale  $u^\sharp$  (supposée exister) peut être caractérisée par

$$\forall u \in U^{\text{ad}}, J(u) - J(u^\sharp) + \langle G'(H(u^\sharp)), H(u) - H(u^\sharp) \rangle \geq 0.$$

*Indication :* on commencera par transformer le problème en le problème équivalent (à démontrer !)

$$\min_{u \in U^{\text{ad}}, c \in \mathcal{C}} \left( J(u) + G(c) \right) \text{ sous } H(u) - c \in -C,$$

puis on écrira les conditions d'optimalité de ce problème de façon appropriée.

Du fait que  $L_b(u, \cdot)$  est la régularisée de  $L(u, \cdot)$  (plus précisément de  $L - I_{C^*}(\cdot)$ ) avec la constante  $1/b$ , et étant donné que, d'après (6.29), la fonction  $\psi$  est définie par minimisation en  $u$  de  $L(u, \cdot) - I_{C^*}(\cdot)$ , cela suggère que si l'on définit :

$$\forall \lambda \in \mathcal{C}^*, \quad \psi_b(\lambda) \stackrel{\text{def}}{=} \inf_{u \in U^{\text{ad}}} L_b(u, \lambda), \quad (7.8)$$

alors  $\psi_b$  devrait être la régularisée de  $\psi$  avec la constante  $1/b$  (régularisation sur tout l'espace). C'est ce que nous montrons maintenant.

**Théorème 7.15.** *La fonction  $\psi_b$  définie par (7.8) est la régularisée concave de  $\psi$  définie par (6.29) avec la constante  $1/b$ .*

*Démonstration.* Si on attaque la démonstration directement à partir des définitions de  $\psi$  et  $\psi_b$  et de l'expression (7.3d) qui a permis d'interpréter  $\zeta_b(\theta, \cdot)$  elle-même comme la régularisation de  $\langle \cdot, \theta \rangle$ , on doit commuter un  $\inf_{u \in U^{\text{ad}}}$  et un  $\sup_{\mu \in C^*}$  pour parvenir au résultat. La fonction de  $u, \mu$  en cause (à savoir  $L(u, \mu) - \|\lambda - \mu\|^2 / 2b$  avec  $\lambda$  comme paramètre) est bien convexe-concave, s.c.i.-s.c.s. et même coercive en  $\mu$  grâce au terme quadratique. Mais, pour appliquer le Théorème 6.13 d'existence de point selle qui permettrait cette commutation, il nous manque une hypothèse de coercivité en  $u$  que nous n'avons pas faite, préférant simplement postuler l'existence d'une solution  $u^\sharp$  au problème (6.1).

C'est pourquoi, une autre preuve possible consiste à utiliser plutôt la formule initiale (7.3c) pour  $\zeta_b$ . On a donc à calculer

$$\begin{aligned} \psi_b(\lambda) &= \inf_{u \in U^{\text{ad}}} \left( J(u) + \zeta_b(\Theta(u), \lambda) \right) \\ &= \inf_{u \in U^{\text{ad}}} \inf_{c \in C} \left( J(u) + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2 \right) \\ &= \inf_{u \in U^{\text{ad}}} \inf_{c \in C} \sup_{\mu \in C^*} \left( J(u) + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2 - \langle \mu, c \rangle \right) \end{aligned}$$

par dualité comme dans la démonstration du Lemme 7.5,

$$= \sup_{\mu \in C^*} \inf_{u \in U^{\text{ad}}} \inf_{c \in C} \left( J(u) + \langle \lambda, \Theta(u) + c \rangle + b \|\Theta(u) + c\|^2 / 2 - \langle \mu, c \rangle \right)$$

car la contrainte  $c \in C$  étant affine est qualifiée,

$$= \sup_{\mu \in C^*} \left( \inf_{u \in U^{\text{ad}}} \left( J(u) + \langle \mu, \Theta(u) \rangle \right) - \|\lambda - \mu\|^2 / 2b \right)$$

par minimisation explicite en  $c$

$$\begin{aligned} &= \sup_{\mu \in C^*} \left( \inf_{u \in U^{\text{ad}}} L(u, \mu) - \|\lambda - \mu\|^2 / 2b \right) \\ &= \sup_{\mu \in C^*} \left( \psi(\mu) - \|\lambda - \mu\|^2 / 2b \right), \end{aligned}$$

où l'on reconnaît la régularisée de  $\psi$  sur  $C^*$ . □

On énonce maintenant un résultat fondamental qui peut être considéré comme l'aboutissement et le but ultime de toute la démarche du Lagrangien augmenté.

**Théorème 7.16.** *Le Lagrangien augmenté est stable en  $u$  (voir Définition 7.4).*

*Démonstration.* Il nous faut montrer que si  $U^\sharp \times \Lambda^\sharp$  est l'ensemble des points selle de  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$ ,<sup>4</sup> alors, pour tout  $\lambda^\sharp \in \Lambda^\sharp$  et tout  $\widehat{u} \in \arg \min_{u \in U^{\text{ad}}} L_b(u, \lambda^\sharp)$ ,  $\widehat{u}$  est aussi dans  $U^\sharp$ .

Si  $\lambda^\sharp \in \Lambda^\sharp$ , alors  $\lambda^\sharp$  appartient aussi à  $\arg \max_{\lambda \in \mathcal{C}^*} \psi_b(\lambda)$ . Or cette fonction est différentiable (comme régularisée de  $\psi$ ) et puisqu'elle atteint son maximum sur tout  $\mathcal{C}^*$  en  $\lambda^\sharp$ , alors sa dérivée en  $\lambda^\sharp$  est nulle. Pour tout  $\widehat{u} \in \arg \min_{u \in U^{\text{ad}}} L_b(u, \lambda^\sharp)$ , on sait que  $(L_b)'_{\lambda}(\widehat{u}, \lambda^\sharp)$  constitue un sur-gradient de  $\psi_b$  au point  $\lambda^\sharp$ , compte tenu de la définition (7.8) de  $\psi_b$  et de l'analogie de (4.26) pour les fonctions concaves. Mais  $\psi_b$  étant dérivable, cette expression est en fait la dérivée de  $\psi_b$  en  $\lambda^\sharp$ , et cette dérivée est nulle ; donc

$$(L_b)'_{\lambda}(\widehat{u}, \lambda^\sharp) = 0 .$$

Cette égalité montre que l'application  $\lambda \mapsto L_b(\widehat{u}, \lambda)$  est aussi maximale en  $\lambda^\sharp$ , et ceci n'est rien d'autre que l'inégalité de gauche du point selle de  $L_b$  sur  $U^{\text{ad}} \times \mathcal{C}^*$  pour le couple  $(\widehat{u}, \lambda^\sharp)$ . L'inégalité de droite pour le même couple provient tout simplement de la définition même de  $\widehat{u}$ . Par conséquent, ce couple est un point selle de  $L_b$  et donc  $\widehat{u} \in U^\sharp$ .  $\square$

### 7.3 Conclusion

Revenons à l'Exemple 7.1 pour comprendre l'intérêt fondamental de ce résultat. On a vu que, dans la situation décrite par cet exemple, non seulement la fonction  $\psi$  que l'on peut chercher à maximiser en  $\lambda \in \mathcal{C}^*$  est le plus souvent non différentiable à l'optimum, mais en plus, même si l'on parvient à construire une suite  $\{\lambda^k\}$  convergeant vers un  $\lambda^\sharp$  optimal, il est peu probable qu'une suite correspondante  $\{u^k\}$ , avec  $u^k \in \arg \min_{u \in U^{\text{ad}}} L(u, \lambda^k)$ , converge vers une solution primale  $u^\sharp$ . L'utilisation du Lagrangien augmenté dans ces circonstances a de nombreux avantages.

- La fonction  $\psi_b$  qui remplace alors la fonction  $\psi$  est différentiable (et qui plus est, de gradient Lipschitzien), ce qui est évidemment préférable du point de vue numérique.
- Elle doit être maximiser sur tout l'espace plutôt que sur le cône positif dual uniquement. Ceci est un autre avantage : les algorithmes dits de “gradient conjugué” ne s'accomodent pas de la présence de contraintes inégalité (la projection, non linéaire, sur ces contraintes détruit la géométrie que ces méthodes cherchent à reconstruire).
- Grâce à la “stabilité en  $u$ ”, on peut désormais s'attendre à ce que si la suite  $\{\lambda^k\}$  converge vers  $\lambda^\sharp \in \Lambda^\sharp$ , alors la suite des  $\widehat{u}(\lambda^k)$  converge vers un point dans  $U^\sharp$ , ce qui n'était pas le cas en l'absence de stabilité en  $u$  du Lagrangien ordinaire. C'est ce que montrent les études de convergence d'algorithmes avec Lagrangien augmenté [7].
- Un autre avantage tient au “conditionnement” de la fonction duale  $\psi_b$ . Cette notion n'a pas été abordée dans ce cours parce que nous n'avons pas parlé d'algorithmes numériques, mais on peut montrer que la facilité de convergence d'algorithmes de type gradient est liée à cette notion de conditionnement, et que par ailleurs ce conditionnement est amélioré par l'opération de régularisation [7]. Une convergence plus rapide de la suite “maximisante”  $\{\lambda^k\}$  peut donc être espérée lorsque cette suite est construite à partir du Lagrangien augmenté plutôt qu'à partir du Lagrangien ordinaire. C'est ce que confirme toute l'expérience numérique accumulée à ce jour avec cette méthode sur des problèmes très variés, et c'est ce qui doit inciter à utiliser cette technique même lorsque le Lagrangien ordinaire est stable en  $u$  et qu'il fournit déjà une fonction  $\psi$  différentiable (cas où  $J$  est strictement, voire fortement, convexe par exemple).

<sup>4</sup>Il est inutile de mettre un indice  $b$  aux ensembles  $U^\sharp$  et  $\Lambda^\sharp$  puisqu'on sait qu'ils sont égaux aux ensembles analogues pour  $L$  d'après le Théorème 7.13.



On peut même dire que plus la constante  $b$  est choisie grande, meilleur est le conditionnement de la fonction  $\psi_b$  et meilleure est donc en principe la convergence des variables duales. Cependant, dans cette matière, il existe un compromis (que montre une étude soignée de convergence [7]) car si  $b$  est choisi trop grand, c'est le conditionnement du problème primal  $\min_{u \in U^{\text{ad}}} L(u, \lambda)$  qui se dégrade et donc sa difficulté de résolution qui augmente.

- Enfin, dans le cas non convexe, l'intérêt de l'utilisation du Lagrangien augmenté est encore plus flagrant, comme le suggère la discussion du début de ce chapitre. En effet, dans ce cas, c'est l'existence même du point selle qui est en cause, et on a vu comment on peut espérer récupérer un saut de dualité grâce à cette technique (au moins "localement"). Pour les cas non convexes, on peut consulter [4, 5].

**Exercice 7.17.** Reprendre l'Exemple 7.1. Utiliser la technique du Lagrangien augmenté. Étudier la fonction  $\psi_b$  et vérifier qu'elle est bien la régularisée de la fonction  $\psi$  calculée à l'Exercice 7.3. Vérifier sa différentiabilité. Calculer l'arg  $\min_{u \in U^{\text{ad}}} L_b(u, \lambda)$  et comparer au cas du Lagrangien ordinaire.

## 7.4 Résumons nous

Par une intuition géométrique de ce qu'est le saut de dualité dans le cas non convexe, on comprend que l'utilisation de paraboloides concaves à la place d'hyperplans pour "ausculter" l'épigraphe de la fonction "perturbation" peut présenter un grand intérêt. Mais même dans le cas limite de problèmes convexes mais pas *strictement* ou *fortement* convexes, on réalise qu'il y a un certain nombre de difficultés (non différentiabilité de la fonction duale, non "stabilité" du Lagrangien par rapport aux variables primales) à manipuler ce Lagrangien dans ce cas, et on pressent, géométriquement, que le Lagrangien augmenté doit permettre de pallier ces difficultés. Dans le cas convexe, cette construction intuitive du Lagrangien augmenté rejoint, de façon remarquable, une construction très classique de l'analyse convexe, à savoir la "régularisation de Yosida-Moreau". En l'occurrence, cette régularisation porte sur la fonction duale qui récupère alors sa différentiabilité (avec une dérivée Lipschitzienne), mais aussi, et pour la même raison, on récupère la *stabilité* du Lagrangien augmenté. Ces avantages sont déjà décisifs sur le plan d'une résolution numérique du problème d'optimisation sous contraintes, mais de plus, la régularisation signifie aussi "meilleure conditionnement" de la fonction duale, et donc meilleure convergence des algorithmes du côté dual. Enfin, dans le cas non convexe, l'utilisation des Lagrangiens augmenté est encore plus vitale dans le cas d'un "saut de dualité".



# Bibliographie

- [1] J.-P. Aubin, *Applied Abstract Analysis*, Wiley-Interscience, New-York, 1977.
- [2] V. Barbu, Th. Precupanu, *Convexity and optimization in Banach spaces*. D. Reidel Publishing Company, Bucarest, 1986.
- [3] D.P. Bertsekas, Necessary and sufficient conditions for a penalty method to be exact. *Mathematical Programming*, Vol. 9, pp. 87–99, 1975.
- [4] D.P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*, Academic Press, New York, 1982.
- [5] D.P. Bertsekas, *Nonlinear Programming*. Athena Scientific, Belmont, USA, 2nd Ed., 1999.
- [6] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland/American Elsevier, Amsterdam/New York, 1973.
- [7] G. Cohen, *Optimisation de grands systèmes*. Cours de DEA MMME, Université de Paris-I, 1999.
- [8] I. Ekeland, R. Temam, *Convex analysis and variational problems*, North-Holland/American Elsevier, Amsterdam/New York, 1976.
- [9] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Tomes I et II, Springer-Verlag, Berlin, 1993.
- [10] A. Kolmogorov, S. Fomine, *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*, Éditions Mir, Moscou, 1974.
- [11] P.-J. Laurent, *Approximation et optimisation*, Hermann, Paris, 1972.
- [12] H. Moulin, F. Fogelman-Soulié, *La convexité dans les Mathématiques de la décision*, Hermann, Paris, 1979.
- [13] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.