

THÈSE DE DOCTORAT DE  
L'UNIVERSITÉ PARIS I - PANTHÉON-SORBONNE  
ET DE L'UNIVERSITÉ FÉDÉRALE DE SANTA CATARINA

Spécialité :  
MATHÉMATIQUES APPLIQUÉES

présentée par :

Mme. Elizabeth Wegner KARAS

pour obtenir le grade de DOCTEUR en co-tutelle de thèse de  
L'UNIVERSITÉ PARIS I - PANTHÉON-SORBONNE  
et de L'UNIVERSITÉ FÉDÉRALE DE SANTA CATARINA.

Sujet de la thèse :

Exemples de chemins centraux non réguliers et  
un algorithme de filtre pour l'optimisation non-linéaire.

Soutenue le 28 juin 2002 devant le jury composé de :

<i>président</i>	Mme. Claudia Alejandra SAGASTIZÁBAL
<i>rapporteurs</i>	M. Jean-Bernard BAILLON M. José Mario MARTÍNEZ
<i>directeurs de thèse</i>	M. Jean Charles GILBERT M. Clóvis Caesar GONZAGA

*Pour  
Edilton et Eduardo.*

# Remerciements

Je tiens tout d'abord à remercier Clóvis Caesar Gonzaga et Jean Charles Gilbert, qui m'ont fait l'honneur d'être mes directeurs de thèse.

Travailler avec Clóvis Caesar Gonzaga pendant les années de doctorat a été un grand plaisir, parce qu'il donne le meilleur de lui-même à ses élèves. Avec lui j'ai appris non seulement en mathématique, mais aussi sur la gastronomie, la littérature, la culture générale, les valeurs, bref sur la vie.

L'année de doctorat sandwich en France a été très important pour ma formation. Je remercie Jean Charles Gilbert qui m'a orienté pendant cette période et qui a fait devenir réalité le processus de co-tutelle avec l'Université Paris I, Panthéon - Sorbonne.

Toute ma gratitude s'adresse à Jean Bernard Baillon et José Mario Martínez, je me réjouis qu'ils aient accepté de juger mon travail en tant que rapporteurs. J'adresse également mes remerciements à Cláudia Alejandra Sagastizabal pour avoir présidé les délibérations du jury.

Je salue également Laurent Chauvier, Xavier Jonsson, Márcia Vanti, Luiz Carlos Matioli et Rafael Casali avec qui j'ai passé des moments agréables. Je remercie Laurent Chauvier par la bonne volonté et la compétence avec laquelle il a corrigé la version française de ma thèse. Je remercie Marie-Lou Margaria par l'attention.

Je remercie aussi entre autres institutions : Le Département de Mathématique de l'Université Fédérale du Paraná (UFPR) qui m'a concédé un congé afin de suivre le doctorat ; Le Département de Mathématique de l'Université Fédérale de Santa Catarina (UFSC) où j'ai développé une grande partie de ma thèse ; L'Institut National de Recherche en Automatique et en Informatique (INRIA) où j'ai passé l'année de doctorat sandwich, à la CAPES et au CNPq qui sont subvenus financièrement à ces années d'étude.

À mes parents Ronald (in memorian) e Ernestina, exemples de perseverance, à qui je dois amour et dédication.

Je dédie ce travail au Edilton et à notre fils Eduardo.

# Abstract

We present some examples of ill-behaved central paths in convex optimization. We work with an extremely simple region in  $\mathbb{R}^2$  and construct increasingly complex objective functions. First, we exhibit a continuous convex function giving rise to an “antenna-like” central path, containing an infinite number of horizontal segments of constant length. The second example is obtained by slightly perturbing the first one and results in a zig-zagging central path with infinite variation. We proceed by smoothing both functions, and produce the same behaviors as before for a differentiable objective function, and finally for a  $C^\infty$  function.

In this work we present also a filter algorithm for nonlinear programming and prove its global convergence to stationary points. Each iteration is composed of a restoration phase, which reduces a measure of infeasibility, and an optimality phase, which reduces the objective function in a tangential approximation of the feasible set. These two phases are totally independent, and the only coupling between them is provided by the filter. The method is independent of the internal algorithms used in each iteration, as long as these algorithms satisfy reasonable assumptions on their efficiency. The main algorithm uses no trust regions, and no properties of the linear models. Under standard hypotheses, we show two results: for a filter with minimum size, the algorithm generates a stationary accumulation point; for a slightly larger filter, all accumulation points are stationary.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Exemples de chemins centraux non réguliers en optimisation convexe</b>	<b>2</b>
1.1 Exemples avec fonctions objectifs continues . . . . .	4
1.2 Opérations de régularisation . . . . .	10
1.3 Des exemples continûment différentiables . . . . .	12
1.4 Un exemple de classe $C^\infty$ . . . . .	14
<b>2 Algorithme de filtre pour l'optimisation non-linéaire</b>	<b>17</b>
2.1 L'algorithme . . . . .	21
2.2 Algorithmes internes . . . . .	25
<b>Conclusion</b>	<b>29</b>
<b>A Examples of ill-behaved central paths in convex optimization</b>	<b>30</b>
<b>B A globally convergent filter method for nonlinear programming</b>	<b>64</b>
<b>Referências Bibliográficas</b>	<b>90</b>

# Introduction

Ce texte correspond au résumé en français de la thèse soumise à l'Université Paris I - Panthéon - Sorbonne et à L'Université Fédérale de Santa Catarina, pour obtenir le grade de docteur en co-tutelle de thèse.

Ce travail se divise en deux parties indépendantes, présentant chacune des résultats nouveaux.

Dans la première partie, j'étudie, dans le cadre de l'optimisation convexe, quelques exemples de chemins centraux à l'allure mouvementée. Certains d'entre eux sont formés d'un nombre infini de segments centraux de longueur fixe ; d'autres manifestent des oscillations avec une variation infinie. Ces chemins peuvent se rencontrer même si les données du problème sont infiniment différentiables.

Dans la deuxième partie de ma thèse, j'introduis un algorithme de filtre globalement convergent pour l'optimisation non linéaire.

Je vais faire une description de ces résultats, en omettant les démonstrations. Pour plus de détails, je renvoie le lecteur à deux articles en annexe. L'article "Examples of ill-behaved central paths in convex optimization" a été écrit avec mes directeurs de thèse : Clóvis Caesar Gonzaga et J. Charles Gilbert. Le deuxième article, intitulé "A globally convergent filter method for nonlinear programming", a été écrit en collaboration avec Clóvis Gonzaga et Márcia Vanti.

# Chapitre 1

## Exemples de chemins centraux non réguliers en optimisation convexe

Les approches par pénalisation ont été centrales dans l'évolution des techniques d'optimisation. Des problèmes avec contraintes sont fréquemment résolus par une suite de problèmes pénalisés qui dépendent d'un seul paramètre de pénalisation  $\mu$ . Les solutions des problèmes pénalisés associées à chaque valeur de  $\mu$  définissent un chemin qui, dans les cas favorables, est régulier et conduit à une solution du problème original.

Ceci est bien développé dans l'ouvrage classique de Fiacco et McCormick [26], dans lequel sont décrites les principales méthodes de pénalisation extérieure, intérieure et mixte pour l'optimisation non linéaire. Les méthodes de pénalisation intérieure sont aussi connues sous le nom de méthodes barrières. Elles se sont considérablement développées ces dernières années, après le succès des méthodes de points intérieurs pour les problèmes linéaires et de complémentarité linéaire (voir [22, 42, 46, 64, 73, 76, 86]; voir aussi les extensions pour l'optimisation non linéaire [9, 20, 24, 32, 34, 80]).

Les premières études approfondies de ces chemins de “solutions”, qui sont aujourd'hui appelés chemins centraux, sont dues à Bayer et Lagarias [5] et à Megiddo [57]. Leur caractérisation du chemin primal-dual a fait date. Une introduction aux méthodes de suivi de chemin central est donnée par Gonzaga [39]. On sait que, pour

l'optimisation linéaire, ce chemin est bien défini, qu'il est infiniment différentiable, et qu'il est de longueur bornée et associée à la complexité des méthodes.

L'allure du chemin central au voisinage d'une solution (du problème original) est décrite par Monteiro et Adler [59]. Ils montrent que le chemin se rapproche de l'ensemble optimal avec un angle bien défini. Des résultats similaires sont obtenus pour certains problèmes convexes par McLinden [56] : sous des hypothèses raisonnables (dont la complémentarité stricte), il prouve que le chemin central converge vers le centre analytique de l'ensemble optimal du problème. Vavasis et Ye [81] montrent que pour l'optimisation linéaire dans  $\mathbb{R}^n$ , le chemin central est composé d'une suite d'au plus  $n^2$  courbes alternées de secteurs quasiment rectilignes.

Après ces résultats, on pourrait s'attendre à des propriétés de régularité semblables pour le chemin central en optimisation convexe. Nous allons montrer que ceci n'est pas nécessairement vrai. Nous donnerons des exemples de problèmes convexes pour lesquels le chemin central associé a une allure mouvementée, quelle que soit la pénalité utilisée.

Nous travaillons dans une région très simple de  $\mathbb{R}^2$  et construisons des fonctions objectifs de complexité croissante. La première de ces fonctions est convexe continue; elle conduit à un chemin central en forme d'antenne, avec un nombre infini de segments horizontaux de longueur constante. Dans le deuxième exemple, on perturbe la première fonction. Le résultat est un chemin en forme de zigzag, avec variation infinie. Nous régularisons ensuite ces deux fonctions en gardant leur convexité et sans modifier la disposition des chemins centraux : nous obtenons les mêmes allures pour des fonctions objectifs différentiables, et finalement pour des fonctions de classe  $C^\infty$ .

On conclut de cette étude que la convexité, même avec de la différentiabilité à un ordre arbitraire, ne permet d'assurer que le chemin central soit régulier (et qu'il soit donc intéressant dans le cadre de méthodes de suivi de chemin central).

### Problème.

Soit  $F : z = (x, y) \in \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction convexe. On étudie des problèmes de la forme

$$\begin{aligned} \text{minimiser } & F(x, y) \\ & y \geq 0, \end{aligned} \tag{1.1}$$



et on suppose que leurs solutions se trouvent sur l'axe  $Ox = \{(x, 0) : x \in \mathbb{R}\}$ . Pour cet ensemble admissible très simple, et pour la majorité des fonctions de pénalisation (en particulier, celles étudiées dans [4]), les solutions des problèmes pénalisés forment des ensembles s'écrivant

$$\arg \min \{F(x, y) \mid y = \text{constante}\}. \quad (1.2)$$

En fonction de la pénalisation utilisée, intérieure ou extérieure,  $y$  pourra être à valeurs positives ou négatives. On considère ici le cas de méthodes de points intérieurs, dont le chemin central est paramétré par  $y > 0$  :

$$y \in \mathbb{R}_{++} \mapsto \chi(y) = \arg \min_{x \in \mathbb{R}} F(x, y). \quad (1.3)$$

## 1.1 Exemples avec fonctions objectifs continues

Nous commençons par deux exemples de problèmes avec chemins centraux triviaux.

**Exemple 1** La fonction  $F$  est définie par  $(x, y) \in \mathbb{R}^2 \mapsto F(x, y) = ay + b$ , où  $a > 0$  et  $b \in \mathbb{R}$ . Dans ce cas,  $\chi(y) = \mathbb{R}$  pour tout  $y > 0$ .  $\square$

**Exemple 2** La fonction  $F$  est définie par  $(x, y) \in \mathbb{R}^2 \mapsto F(x, y) = ay + b + \varepsilon x$ , où  $a > 0$ ,  $b$  et  $\varepsilon \in \mathbb{R}$ , avec  $\varepsilon \neq 0$ . Maintenant,  $\chi(y) = \emptyset$  pour tout  $y > 0$ .  $\square$

Perturbons maintenant les fonctions des exemples 1 et 2. Soit  $y_k > 0$  (l'indice  $k \in \mathbb{N}$  sera utile dans la suite). On définit la fonction  $g_k^0$  par

$$z = (x, y) \in \mathbb{R}^2 \mapsto g_k^0(x, y) = \frac{1}{2} (\|z - z_k^-\| + \|z - z_k^+\| - 2), \quad (1.4)$$

où  $\|\cdot\|$  désigne la norme euclidienne,

$$z_k^- = (-1, y_k) \quad \text{et} \quad z_k^+ = (1, y_k).$$

La fonction  $g_k^0$  a les propriétés suivantes :

- (i) elle est convexe ;

(ii) son minimum est 0, il est atteint en tout point du segment

$$[z_k^-, z_k^+] = \{(1-t)z_k^- + tz_k^+ : t \in [0, 1]\};$$

(iii) elle est différentiable en tout point, à l'exception de  $z_k^-$  et  $z_k^+$ .

$$(iv) \quad g_k^0(x, y) \leq \begin{cases} |y - y_k| & \text{si } |x| \leq 1, \\ |x| + |y - y_k| - 1 & \text{dans le cas contraire.} \end{cases}$$

Nous construirons des fonctions convexes  $g_k$  qui régularisent  $g_k^0$  (à un ordre plus ou moins élevé) et qui satisfont

$$0 \leq g_k(z) \leq g_k^0(z), \quad (1.5)$$

pour  $z$  dans un certain domaine.

On peut présenter notre premier exemple non trivial, la fonction  $f_k$ . On additionne à la fonction de l'exemple 2 (avec  $a$ ,  $b$ , et  $\varepsilon$  indexés par  $k$ ), la perturbation convexe  $g_k$  décrite ci-dessus, multipliée par une constante positive  $c_k$ ; et la constante  $-a_k y_k$ .

**Exemple 3** On se donne des paramètres  $a_k, b_k, c_k, \varepsilon_k \in \mathbb{R}$ . La fonction  $f_k$  est définie par

$$z = (x, y) \in \mathbb{R}^2 \mapsto f_k(z) = a_k(y - y_k) + b_k + c_k g_k(z) + \varepsilon_k x, \quad (1.6)$$

où la fonction convexe  $g_k : \mathbb{R}^2 \rightarrow \mathbb{R}$  satisfait (1.5), avec  $g_k^0$  donnée par (1.4).  $\square$

Pour se familiariser avec la fonction  $f_k$ , que l'on va beaucoup utiliser, voici deux cas particuliers avec  $g_k \equiv g_k^0$ .

- Si  $\varepsilon_k = 0$ , nous avons l'exemple 1 avec en plus une perturbation  $c_k g_k^0(\cdot)$ . Supposons que  $c_k$  soit positif, mais suffisamment petit pour que le problème (1.1) ait ses solutions sur l'axe  $Ox$ . On peut facilement voir, en utilisant (1.3), que pour tout  $y > 0$ ,  $\chi(y) = \{0\}$  si  $y \neq y_k$ , et  $\chi(y_k) = [-1, +1]$ . Le chemin central est la croix à gauche dans la figure 1.1.
- Supposons maintenant que  $\varepsilon_k \neq 0$ , mais petit en valeur absolue, disons  $|\varepsilon_k| < c_k$ . Pour  $y \neq y_k$ ,  $x \mapsto f_k(x, y)$  est strictement convexe, avec des ensembles de niveau compacts, donc  $\chi(y)$  est un singleton. Si  $y = y_k$ , on observe que :

$$\begin{aligned}
f'_k(z_k^-; -e^1) &= c_k - \varepsilon_k, \\
f'_k(z_k^-; e^1) &= \varepsilon_k, \\
f'_k(z_k^+; -e^1) &= -\varepsilon_k, \\
f'_k(z_k^+; e^1) &= c_k + \varepsilon_k.
\end{aligned}$$

Par conséquent, l'inégalité  $|\varepsilon_k| < c_k$  implique que  $\chi(y_k) = \{-1\}$  si  $\varepsilon_k > 0$  et  $\chi(y_k) = \{+1\}$  si  $\varepsilon_k < 0$ . Le chemin central est dévié sur le côté, comme le montre la partie droite de la figure 1.1 : le chemin est dévié vers la gauche ou vers la droite, selon que  $\varepsilon_k > 0$  ou  $\varepsilon_k < 0$ .

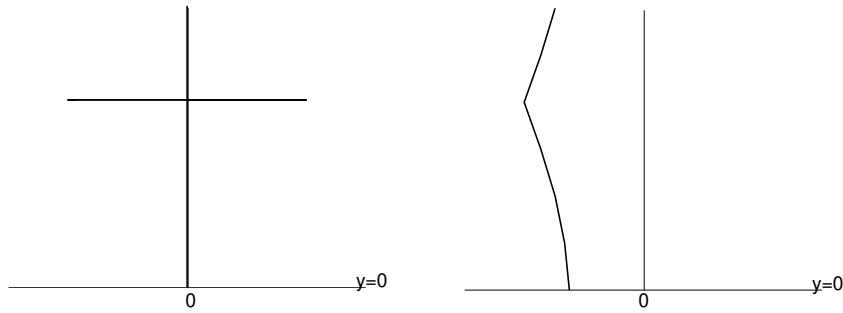


FIG. 1.1 – Chemin central de l'exemple 3 avec  $g_k = g_k^0$  et  $\varepsilon_k = 0$  (à gauche) ou  $\varepsilon_k > 0$

Les principaux exemples de ce travail sont construits à partir de la fonction  $f_k$  définie dans l'exemple 3, avec plusieurs coordonnées de type  $y_k$  et plusieurs coefficients  $a_k$ ,  $b_k$ ,  $c_k$  et  $\varepsilon_k$ . Ces fonctions sont tangentes à une *fonction support*  $(x, y) \in \mathbb{R} \times [-0.5, 1] \mapsto \psi(y)$ , où  $\psi : [-0.5, 1] \rightarrow \mathbb{R}$  a les propriétés suivantes :

$$\begin{cases} \psi \text{ est continûment différentiable,} \\ \psi(y) = 0, \text{ pour tout } y \in [-0.5, 0], \\ y \in [0, 1] \mapsto \psi(y) \text{ est strictement convexe.} \end{cases} \quad (1.7)$$

Clairement, ces propriétés impliquent que  $\psi(y) > 0$  et  $\psi'(y) > 0$  pour tout  $y \in (0, 1]$ . Les fonctions  $f_k$  sont tangentes à la fonction support  $\psi$  dans le sens où les coefficients  $a_k$  et  $b_k$  sont choisis en utilisant la pente et la valeur de  $\psi$  en  $y_k$ , comme ci-dessous :

$$y_k = 2^{-k}, \quad a_k = \psi'(y_k), \quad \text{et} \quad b_k = \psi(y_k). \quad (1.8)$$

Des hypothèses (1.7) pour  $\psi$ , nous déduisons que les suites  $\{y_k\}$ ,  $\{a_k\}$  et  $\{b_k\}$  sont positives, décroissantes, et convergent vers zéro lorsque  $k \rightarrow \infty$ .

Nous avons encore des degrés de liberté pour déterminer chaque  $f_k$ , grâce aux paramètres non spécifiés  $c_k > 0$  et  $\varepsilon_k$ , et grâce à la fonction  $g_k$ . Les paramètres  $c_k$  et  $\varepsilon_k$  seront fixés pour contrôler les positions respectives des graphes de  $f_k$  et  $f_{k-1}$ , et donner des propriétés particulières au chemin central. D'autre part, la fonction  $g_k$  sera une modification appropriée de  $g_k^0$  pour fournir des propriétés de différentiabilité.

Le contrôle fin de la fonction objectif dans les prochains exemples sera fait seulement sur le rectangle fermé de  $\mathbb{R}^2$  défini par

$$\Omega = \{(x, y) \in \mathbb{R}^2 \mid |x| \leq 1.5, -0.5 \leq y \leq 1\}. \quad (1.9)$$

Par exemple, les résultats de régularité ne seront montrés que pour des points de  $\Omega$ . Ces propriétés peuvent facilement être étendues à tout l'espace  $\mathbb{R}^2$ . Des propriétés particulières seront satisfaites dans les bandes horizontales de  $\Omega$ , définies pour  $k = 1, 2, \dots$  par

$$\Omega_k = \{(x, y) \in \Omega \mid y \in [y_k, y_{k-1}]\}. \quad (1.10)$$

Notre principal objectif est maintenant de définir les paramètres de perturbation  $c_k$  et  $\varepsilon_k$ . On va examiner deux fonctions consécutives  $f_{k-1}$  et  $f_k$ , pour un indice  $k \geq 1$  (voir la figure 1.2).

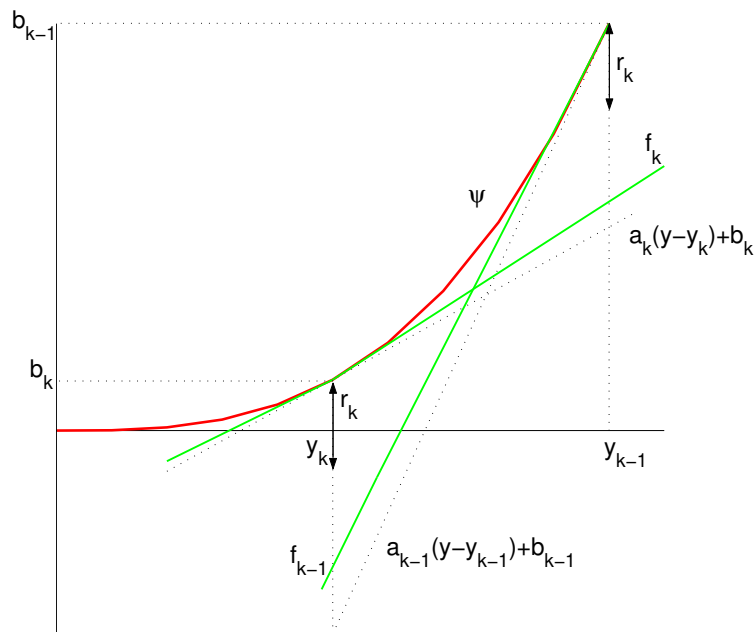
On souhaite que les fonctions  $f_{k-1}$  et  $f_k$  se "croisent" dans  $\Omega_k$ , avec une différence positive d'au moins  $r_k$  entre les modèles linéarisés en  $y_{k-1}$  et  $y_k$  (voir la figure 1.2). L'écart  $r_k$  doit satisfaire

$$\begin{aligned} r_k &\leq b_k - (b_{k-1} + a_{k-1}(y_k - y_{k-1})), \\ r_k &\leq b_{k-1} - (b_k + a_k(y_{k-1} - y_k)). \end{aligned}$$

Ces inégalités sont compatibles avec la stricte positivité de  $r_k$  : grâce à la stricte convexité de  $\psi$  sur  $[0, 1]$ , les termes dans le membre de droite sont positifs. Notre choix pour  $r_k$  ( $k \geq 1$ ) est

$$r_k = \frac{1}{2} \min \{b_k - (b_{k-1} + a_{k-1}(y_k - y_{k-1})), b_{k-1} - (b_k + a_k(y_{k-1} - y_k))\}. \quad (1.11)$$

Par conséquent, la suite  $\{r_k\}$  est positive et tend vers zéro lorsque  $k \rightarrow \infty$ . On considère, pour  $k \geq 0$  (pour démarrer le processus, on prend  $r_0 = +\infty$  et  $\varepsilon_{-1} =$

FIG. 1.2 – Choix des coefficients de la fonction  $f_k$ .

$+\infty$ ) :

$$c_k = \frac{1}{4} \min\{r_k, r_{k+1}\}$$

ou bien  $\varepsilon_k = 0$  pour tout  $k \geq 0$ ,

$$\text{ou bien } \varepsilon_k = (-1)^k |\varepsilon_k| \text{ avec } 0 < |\varepsilon_k| \leq \min\{c_k/4, |\varepsilon_{k-1}|\}, \text{ pour tout } k \geq 0. \quad (1.12)$$

La suite  $\{|\varepsilon_k|\}$  est ou bien identiquement nulle, ou bien positive et décroissante.

On peut à présent décrire notre quatrième exemple de fonction objectif.

**Exemple 4** Soient  $\psi : [-0.5, 1] \rightarrow \mathbb{R}$  une fonction qui satisfait (1.7) et des fonctions  $f_k$  définies pour tout  $k \in \mathbb{N}$  comme dans l'exemple 3, avec  $y_k$  et des coefficients  $a_k$ ,  $b_k$ ,  $c_k$ ,  $\varepsilon_k$  satisfaisant (1.8), (1.11) et (1.12). Les fonctions convexes  $g_k$  ne sont pas spécifiées, mais elles doivent satisfaire (1.5) sur  $\Omega$ . La fonction objectif  $F$  est alors définie pour  $z \in \Omega$  par

$$F(z) = \sup_{k \in \mathbb{N}} f_k(z). \quad (1.13)$$

□

Pour ce choix, on montre que deux fonctions consécutives  $f_{k-1}$  et  $f_k$  se croisent dans la bande  $\Omega_k$ .

**Lemme 1.1.1** *Soient les fonctions  $f_k$  et les paramètres  $r_k$  de l'exemple 4. Alors, pour  $k \geq 1$  et  $(x, y) \in \Omega$  :*

$$f_{k-1}(x, y) \leq f_k(x, y) - r_k \quad \text{si } y \leq y_k \quad (1.14)$$

$$f_k(x, y) \leq f_{k-1}(x, y) - r_k \quad \text{si } y \geq y_{k-1}. \quad (1.15)$$

□

Nous allons donner quelques propriétés intéressantes de la fonction objectif  $F$  introduite dans l'exemple 4. Elles découlent du dernier lemme.

**Lemme 1.1.2** *Considérons les fonctions  $f_k$  et  $F$  définies dans l'exemple 4 et  $z = (x, y) \in \Omega$ . Alors  $\lim_{k \rightarrow \infty} f_k(z) = 0$ . De plus :*

$$(i) \quad F(x, y) = 0, \text{ si } y \leq 0;$$

$$(ii) \quad F(z) = \max\{f_{k-1}(z), f_k(z)\}, \text{ si } z \in \Omega_k;$$

$$(iii) \quad F(x, y_k) = f_k(x, y_k), \text{ si } (x, y_k) \in \Omega.$$

□

Le lemme suivant précise certaines propriétés du chemin central.

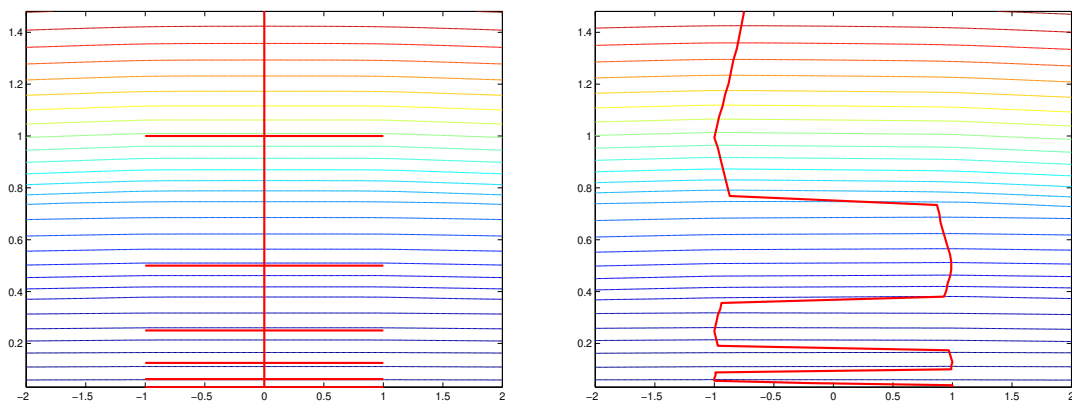
**Lemme 1.1.3** *La fonction objectif  $F$  définie dans l'exemple 4 est convexe et continue sur  $\Omega$ . Soit  $k \geq 1$  et supposons que  $g_k \equiv g_k^0$ . Alors*

$$\arg \min_{|x| \leq 1.5} F(x, y_k) = \begin{cases} [-1, 1] & \text{si } \varepsilon_k = 0, \\ \{(-1)^{k+1}\} & \text{dans le cas contraire.} \end{cases} \quad (1.16)$$

*De plus, pour  $y \in (y_k, y_{k-1})$ ,  $\arg \min_{|x| \leq 1.5} F(x, y)$  est réduit à un seul point. Ce point est  $x = 0$  lorsque  $\varepsilon_k = \varepsilon_{k-1} = 0$ .*

□

Supposons que  $g_k \equiv g_k^0$  dans l'exemple 4. On déduit du lemme 1.1.3 l'allure du chemin central qui est représenté à la figure 1.3. Si  $\varepsilon_k = 0$ , le chemin ressemble à une antenne avec une infinité de branches (à gauche dans la figure). Dans le cas contraire, le chemin passe alternativement par les points  $z_k^-$  lorsque  $k$  est pair et  $z_k^+$  lorsque  $k$  est impair (à droite). On a un zigzag avec variation infinie. On montre aussi que, dans ce dernier cas, le chemin est une courbe continue pour  $\mu > 0$ .

FIG. 1.3 – Antenne et zigzag pour une fonction objectif  $C^0$ .

## 1.2 Opérations de régularisation

Notre objectif est maintenant de régulariser la fonction objectif  $F$  introduite dans l'exemple 4, sans modifier l'allure du chemin central obtenu pour  $\varepsilon_k = 0$  (antenne) et  $\varepsilon_k \neq 0$  (zigzag). Un examen de la construction de  $F$  montre que pour y parvenir, il faut régulariser les fonctions élémentaires  $f_k$  introduites en (1.6), ce qui demande de régulariser  $g_k^0$  et l'opérateur "sup" qui apparaît dans (1.13).

Tout d'abord, occupons-nous des fonctions élémentaires  $f_k$ . On construit  $g_k$  dans la formule (1.6) comme une approximation "lisse" de la fonction continue  $g_k^0$  donnée par (1.4), en maintenant la condition (1.5) et sa positivité. Ainsi, les lemmes de la dernière section portant sur  $g_k$  sont encore vrais.

La deuxième opération de régularisation concerne le "sup" qui apparaît dans (1.13). Le lemme 1.1.3 nous montre que cet opérateur, qui porte a priori sur toutes les fonctions  $f_k$ , ne prend en compte, sur la bande  $\Omega_k$ , que les deux fonctions élémentaires successives  $f_k$  et  $f_{k-1}$ . Par conséquent, il suffit de régulariser le maximum de *deux* fonctions convexes, en des points où ces deux fonctions ont des valeurs proches ; on applique ensuite cette technique bande par bande, pour assurer la régularité de  $F$  en chaque point  $y > 0$ .

### Régularisation du maximum de deux fonctions.

Soient deux fonctions convexes  $f_1$  et  $f_2$  à valeurs réelles définies sur  $\mathbb{R}^n$  et de classe  $C^q$  (ici comme dans la suite,  $q$  est ou bien un entier positif, ou bien  $q = \infty$ ).

Le maximum de  $f_1$  et  $f_2$  est noté  $f_{max}$  :

$$z \in \mathbb{R}^n \mapsto f_{max}(z) = \max\{f_1(z), f_2(z)\}.$$

Soit  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction convexe de classe  $C^q$ , telle que

$$\varphi(w) = |w|, \quad \text{pour } |w| \geq 1. \quad (1.17)$$

Pour obtenir des fonctions ayant cette propriété, on peut par exemple intégrer 2 fois une fonction densité de probabilité, comme proposé par Chen et Mangasarian [15, 16]. La fonction  $\varphi$  est une approximation  $C^q$  de la fonction valeur absolue. Elle ne la modifie pas pour  $|w| \geq 1$ . Comme cette approximation devra être de plus en plus précise, nous considérons un scalaire  $r > 0$  et la fonction convexe  $\varphi_r : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$\varphi_r(w) = r \varphi\left(\frac{w}{r}\right) \quad \text{pour tout } w \in \mathbb{R}. \quad (1.18)$$

Nous avons  $\varphi_r(w) = |w|$  pour  $|w| \geq r$ . Les fonctions  $\varphi_r$  sont des approximations d'autant plus précises de la fonction valeur absolue que  $r \rightarrow 0$ .

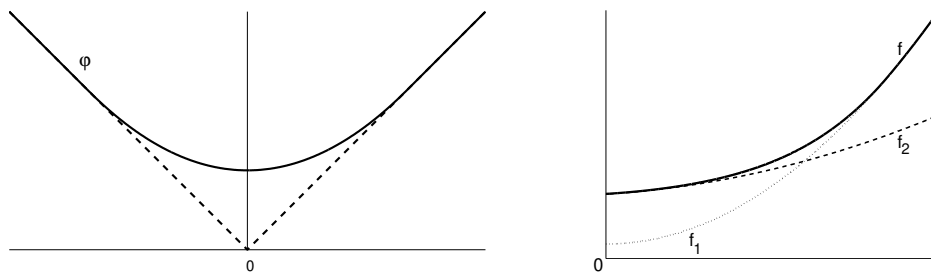


FIG. 1.4 – Régularisation de la fonction valeur absolue et du maximum de deux fonctions convexes.

Étant donné  $r > 0$ , on introduit la fonction  $M_r : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$M_r(t_1, t_2) = \frac{1}{2}(t_1 + t_2 + \varphi_r(t_1 - t_2)), \quad (1.19)$$

qui est une version régularisée de la fonction max. Il est donc naturel d'approcher la fonction  $f_{max}$  par (voir l'image de droite dans la figure 1.4)

$$z \in \mathbb{R}^n \mapsto f(z) = M_r(f_1(z), f_2(z)). \quad (1.20)$$



La fonction  $M_r$  est convexe et croissante. Le lemme suivant montre en quel sens la fonction  $f$  définie par (1.20) est une régularisation convexe de la fonction maximum.

**Lemme 1.2.1** *Soient  $f_1$  et  $f_2$  deux fonctions convexes de classe  $C^q$ , avec  $0 \leq q \leq \infty$ , et  $f_{\max} = \max\{f_1, f_2\}$ . Soit  $f$  la fonction définie par (1.20), où  $M_r$  est construite comme ci-dessus, avec la fonction  $\varphi$  de classe  $C^q$ . Alors*

- (i)  $f$  est convexe,
- (ii)  $f$  est de classe  $C^q$ ,
- (iii) pour tout  $z \in \mathbb{R}^n$  tel que  $|f_1(z) - f_2(z)| \geq r$ , on a  $f(z) = f_{\max}(z)$ ,
- (iv) pour tout  $z \in \mathbb{R}^n$ , on a  $f_{\max}(z) \leq f(z) \leq f_{\max}(z) + r/2$ .

□

### 1.3 Des exemples continûment différentiables

Notre cinquième exemple est construit en spécifiant quelques paramètres qui étaient libres dans l'exemple 4, et en ayant pour objectif de rendre la fonction  $F$  qui en résulte de classe  $C^1$ .

**Exemple 5** La fonction  $F : \Omega \rightarrow \mathbb{R}$  est construite avec les ingrédients suivants :

- la fonction support  $\psi : [-0.5, 1] \rightarrow \mathbb{R}$  définie pour  $y \in [-0.5, 1]$  par

$$\psi(y) = \frac{1}{2} (\max\{0, y\})^2. \quad (1.21)$$

Elle satisfait les conditions (1.7) ;

- les fonctions  $g_k(z) = \frac{1}{4} (g_k^0(z))^2$ , avec  $k \in \mathbb{N}$ , où  $g_k^0$  est donnée par (1.4). Ces fonctions sont convexes, de classe  $C^1$  et satisfont (1.5) ;
- les fonctions élémentaires  $f_k$ ,  $k \in \mathbb{N}$ , définies par (1.6) (comme dans l'exemple 3),  $y_k$  et les coefficients  $a_k$ ,  $b_k$ ,  $c_k$  et  $\varepsilon_k$  satisfaisant (1.8), (1.11) et (1.12) ;
- la fonction de régularisation  $\varphi$  est obtenue par double intégration de la fonction indicatrice.

La fonction objectif  $F$  est construite grâce à la récurrence suivante :

- on prend  $F_0(z) = f_0(z)$ , pour tout  $z \in \Omega$  ;
- pour  $k = 1, 2, \dots$ , on considère

$$F_k(z) = M_{r_k}(f_k(z), F_{k-1}(z)), \quad \text{pour tout } z \in \Omega, \quad (1.22)$$

- où  $r_k$  est défini par (1.11) et  $M_{r_k}$  est définie par (1.18) et (1.19);  
 –  $F$  est la limite point par point des fonctions  $F_k$  :

$$F(z) = \lim_{k \rightarrow \infty} F_k(z), \quad \forall z \in \Omega. \quad (1.23)$$

□

La construction par récurrence n'est pas tout à fait nécessaire. En effet, la fonction  $F$  peut être définie simplement par

$$F(z) = F_k(z) = M_{r_k}(f_k(z), f_{k-1}(z)) \quad \text{si } z \in \Omega_k, \quad k \in \mathbb{N}.$$

Notons  $\bar{f} = \sup_{k \in \mathbb{N}} f_k$  la fonction objectif de l'exemple 4 avec les données de l'exemple 5. Alors on peut montrer que

$$\bar{f}(z) \leq F(z) \leq \bar{f}(z) + r_k/2, \quad \text{pour tout } z \in \Omega_k. \quad (1.24)$$

C'est-à-dire que la fonction  $F$  régularise  $\bar{f}$ . Le processus de régularisation n'a d'effet que dans une bande  $\Omega_k$  à la fois, où il lisse les points de non-différentiabilité, sans affecter les non-différentiabilités des autres bandes.

La fonction  $F$  construite dans l'exemple 5 est en fait convexe et régulière. Analysons son chemin central. Pour  $\varepsilon_k = 0$ , la forme d'antenne du chemin central est préservée, comme dans l'exemple 4. Chaque segment  $[z_k^-, z_k^+]$  appartient au chemin central car  $F(z) = f_k(z)$  lorsque  $z = (x, y_k) \in \Omega$ . Pour  $y \in (y_k, y_{k-1})$ , on sait que les deux fonctions  $f_k(\cdot, y)$  et  $f_{k-1}(\cdot, y)$  ont pour seul minimum  $x = 0$ . Alors, ceci est aussi le cas pour  $F(\cdot, y) = M_{r_k}(f_k(\cdot, y), f_{k-1}(\cdot, y))$ , du fait de la propriété de monotonie de  $M_{r_k}$ .

On considère maintenant le cas où  $\varepsilon_k \neq 0$  est calculé comme en (1.12). Pour  $y \in (y_k, y_{k-1})$ ,  $F(\cdot, y)$  a un seul minimum. Pour  $y = y_k$ ,  $F(\cdot, y_k) = f_k(\cdot, y_k)$ , en utilisant la définition des différentes fonctions, un calcul aisé nous donne

$$\arg \min_{x \in \mathbb{R}} F(x, y_k) = \{(-1)^{k+1}(1 + \tau_k)\},$$

où  $0 < \tau_k = 2|\varepsilon_k|/c_k \leq 0.5$ . Dans l'exemple 5, le chemin central est une courbe continue (pour  $y > 0$ ), formant un zigzag avec des tours encore plus grands. Il est possible de modifier les fonctions utilisées dans l'exemple 5 pour construire une

fonction  $F$  encore plus régulière. Les chemins centraux en forme d'antenne et de zigzag avec une régularité  $C^1$  sont présentés à la figure 1.5 (le zigzag est la courbe en tirets). La courbe continue est le chemin central correspondant à une fonction  $F$  de classe  $C^2$ .

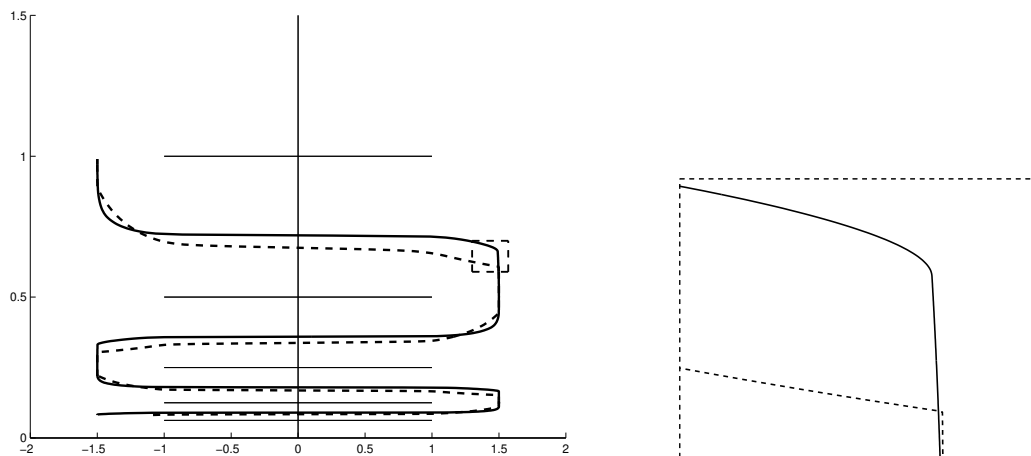


FIG. 1.5 – Chemin central en zigzag pour fonctions objectifs  $C^1$  et  $C^2$ .

Le prochain lemme donne des conditions qui assurent la régularité du chemin central en zigzag.

**Lemme 1.3.1** *Considérons l'exemple 5, dans lequel  $g_k = \gamma \circ g_k^0$ , où  $\gamma : \mathbb{R} \rightarrow \mathbb{R}$  est deux fois différentiable et satisfait  $\gamma'(0) = 0$ , ainsi que  $\gamma'(t) > 0$  et  $\gamma''(t) > 0$  lorsque  $t > 0$ . Supposons également que la fonction de régularisation  $\varphi$  soit deux fois différentiable et que la fonction objectif  $F$  soit convexe et de classe  $C^q$ , avec  $q \geq 2$ . Alors, le chemin central en zigzag est une fonction de classe  $C^{q-1}$  de  $y > 0$ .*

□

## 1.4 Un exemple de classe $C^\infty$

La construction développée dans la dernière section est maintenant étendue au cas d'une régularisation  $C^\infty$ , en spécifiant de nouvelles fonctions  $\psi$ ,  $g_k$ , et  $\varphi$ , sans modifier la forme des chemins centraux.

On utilise pour cela un exemple classique de fonction non analytique de

classe  $C^\infty$  (voir [27, page 51]), définie par

$$w \in \mathbb{R} \mapsto \theta(w) = \begin{cases} e^{-1/w^2} & \text{se } w > 0, \\ 0 & \text{se } w \leq 0. \end{cases} \quad (1.25)$$

**Exemple 6** La fonction  $F : \Omega \rightarrow \mathbb{R}$  est construite de façon analogue à celle de l'exemple 5, avec toutefois les modifications suivantes :

- la fonction support  $\psi : [-0.5, 1] \rightarrow \mathbb{R}$  est maintenant définie, pour  $y \in [-0.5, 1]$ , par

$$\psi(y) = \theta\left(\frac{y}{2}\right), \quad (1.26)$$

où  $\theta$  est la fonction  $C^\infty$  donnée en (1.25) ;

- les fonctions de perturbation  $g_k : \Omega \rightarrow \mathbb{R}$ , pour  $k \in \mathbb{N}$ , sont maintenant obtenues en régularisant, grâce à  $\theta$ , les fonctions  $g_k^0$  définies en (1.4) : pour  $z = (x, y) \in \Omega$ ,

$$g_k(z) = g_k^\infty(z) := \theta\left(\frac{1}{3}g_k^0(z)\right); \quad (1.27)$$

- la fonction de régularisation  $\varphi$  est maintenant une fonction  $C^\infty$ , obtenue par double intégration d'une fonction densité de probabilité définie à partir de la fonction  $\theta$  ci-dessus.

□

La convexité et la régularité des chemins centraux correspondants à la fonction objectif  $F$  de l'exemple 6 sont obtenus à l'aide d'arguments analogues à ceux de la section 1.3. La principale difficulté consiste à montrer que  $F$  est de classe  $C^\infty$ .

### Remarques.

Les minima des exemples 4-6 ne satisfont pas la condition de complémentarité stricte, car  $F(x, y) = 0$  lorsque  $y \leq 0$ . Cependant, les chemins centraux ne seront pas modifiés si on ajoute à  $F$  un terme  $ay$ , avec  $a > 0$ . Dans ce cas, on a complémentarité stricte. On remarque que la forme des chemins centraux est compatible avec le résultat de McLinden [56] selon lequel, lorsque la complémentarité stricte a lieu, les points limites d'une sélection convergente quelconque de points centraux sont dans le centre analytique de l'ensemble optimal. Ici, le centre analytique coïncide avec l'ensemble optimal car l'unique contrainte est toujours active en les solutions.

Il n'est pas possible de construire des chemins centraux ayant l'allure de l'antenne ou du zigzag avec une fonction objectif analytique. Monteiro et Zhou [60]

ont en effet montré que si les fonctions définissant le problème sont analytiques (avec d'autres hypothèses raisonnables), le chemin central primal est une courbe qui converge vers un seul point dans l'ensemble optimal (des résultats analogues sont donnés par Cominetti [19] et Champion [14]). Partant de là, il est remarquable qu'avec la régularité  $C^\infty$ , nous ayons pu construire des chemins centraux d'allure aussi irrégulière.

# Chapitre 2

## Algorithme de filtre pour l'optimisation non-linéaire

Dans la deuxième partie de cette thèse, nous allons étudier le problème d'optimisation non linéaire

$$(P) \quad \begin{aligned} & \text{minimiser } f_0(x) \\ & f_{\mathcal{E}}(x) = 0 \\ & f_{\mathcal{I}}(x) \leq 0. \end{aligned}$$

Les indices des ensembles  $\mathcal{E}$  et  $\mathcal{I}$ , relatifs aux contraintes d'égalité et d'inégalité respectivement, sont supposés former une partition de  $\{1, \dots, m\}$ . Nous supposons également que les fonctions  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  ( $i = 0, 1, \dots, m$ ) sont continûment différentiables. Les matrices jacobiniennes de  $f_{\mathcal{E}}$  et  $f_{\mathcal{I}}$  sont notées  $A_{\mathcal{E}}(\cdot)$  et  $A_{\mathcal{I}}(\cdot)$  respectivement. Définissons la fonction  $f^+ : \mathbb{R}^n \rightarrow \mathbb{R}^m$  par

$$f_i^+(x) = \begin{cases} f_i(x) & \text{si } i \in \mathcal{E} \\ \max\{0, f_i(x)\} & \text{si } i \in \mathcal{I}. \end{cases} \quad (2.1)$$

Considérons une fonction  $x \in \mathbb{R}^n \mapsto h(x)$ , qui mesure la satisfaction des contraintes en  $x$  et pénalise leur violation. Le plus souvent, cette mesure est donnée par

$$h(x) = \|f^+(x)\|, \quad (2.2)$$

où  $\|\cdot\|$  désigne une norme arbitraire. Un algorithme d'optimisation non linéaire doit travailler avec deux critères parfois antagonistes,  $f_0$  et  $h$ , qui doivent être si-

multanément minimisés, sachant que de préférence la mesure  $h$  doit tendre vers zéro.

L'optimalité et l'admissibilité peuvent être combinées en utilisant des fonctions de pénalisation ou des lagrangiens augmentés, ou peuvent être considérées plus ou moins indépendantes l'une de l'autre. Les méthodes étudiées dans ce travail traitent  $f_0$  et  $h$  comme deux objectifs indépendants. Chaque itération de ces méthodes est composée de deux phases : une phase de admissibilité, qui fait décroître  $h$ , suivie d'une phase d'optimisation, qui fait décroître  $f_0$ .

De telles méthodes s'apparentent aux méthodes de gradient projeté de Rosen [74] et GRG de Abadie et Carpentier [1]. Elles ont été étudiées, plus récemment, par Martínez et Pilotta [53]. En combinant des idées de l'optimisation quadratique successive et des algorithmes de régions de confiance pour problèmes avec contraintes d'égalité seulement, Celis, Dennis et Tapia [13] ont ouvert une direction de recherche qui a abouti à la méthode de Byrd et Omojokum [8, 66] : à chaque itération de cette méthode, on considère une région de confiance centrée sur l'itéré courant  $x^k$ , et on calcule un pas *normal* (de admissibilité) suivi d'un pas *tangent* (vers l'optimalité), qui maintiennent  $x^{k+1}$  dans la région de confiance. Le pas tangentiel est dans le noyau de la jacobienne des contraintes en  $x^k$ .

Les phases de admissibilité et d'optimisation sont devenues plus indépendantes dans les algorithmes de *restauration inexacte* proposés par Martínez [51] et par Martínez et Pilotta [52, 53]. Leur région de confiance est centrée sur le point obtenu après la phase de admissibilité. Diverses méthodes peuvent être utilisées pour réduire  $h$ . Ils décrivent un algorithme pour problèmes avec contraintes d'égalité non linéaires et contraintes de bornes. La phase de admissibilité peut aussi utiliser des idées de Byrd, Gilbert et Nocedal [9] et de Byrd, Hribar et Nocedal [10], qui récrivent le problème en n'utilisant que des contraintes d'égalité et des variables d'écart strictement positives. Dans ces algorithmes, le progrès vers la solution est habituellement mesuré par une *fonction de mérite*  $\psi = f_0 + \nu h$ , où  $\nu$  est un poids positif. A l'itération  $k$ , les points de  $\{x \in \mathbb{R}^n \mid \psi(x) \geq \psi(x^k)\}$  sont *interdits*, et le pas doit permettre de diminuer la valeur de  $\psi$ . Le choix de  $\nu$  est délicat : de petites valeurs de  $\nu$  peuvent interdire des solutions optimales ; de grandes valeurs de  $\nu$  peuvent rendre l'algorithme trop lent. En règle générale,  $\nu$  doit croître lorsque l'algorithme progresse en augmentant l'importance de  $h$  dans  $\psi$ . Il fait habituellement interagir le pas de admissibilité et le pas tangentiel (et réduit leur indépendance). Chaque

itération ne dépend que de l'itéré courant.

### Algorithmes de filtre.

Les algorithmes de filtre définissent une *région interdite* d'une façon différente et astucieuse, en mémorisant certains couples  $(f_0(x^k), h(x^k))$  au cours des itérations précédentes, et en rejetant un point s'il est dominé par les  $x^k$  correspondants, au sens de la règle usuelle de domination de Pareto :

$$\text{“ } x \text{ domine } y \text{ si et seulement si } f_0(y) \geq f_0(x) \text{ et } h(y) \geq h(x)\text{”}.$$

Nous ne pouvons pas construire l'ensemble des points interdits, mais il est facile de vérifier qu'un point en fait partie, en faisant un petit nombre de comparaisons dans  $\mathbb{R}^2$ . Ces méthodes ont été introduites par Fletcher et Leyffer dans un article fondamental [29], et une preuve de convergence globale a été obtenue par Fletcher, Gould, Leyffer, Toint et Wächter [28]. Dans ces papiers, chaque phase de admissibilité doit réduire  $h$  jusqu'à ce qu'une propriété dite de *compatibilité* soit vérifiée. Elle dépend du rayon de la région de confiance et du modèle linéaire des contraintes.

Notre méthode est un algorithme de restauration dans le sens de Martínez et Pilotta [52], qui utilisent également un filtre. La méthode a les caractéristiques suivantes :

- chaque itération commence avec un filtre et sa région interdite associée ;
- les phases de admissibilité et d'optimisation sont totalement indépendantes, et peuvent faire intervenir n'importe quels algorithmes (qui doivent toutefois satisfaire quelques hypothèses raisonnables). Le seul lien entre les deux phases est qu'elles ne doivent pas permettre de générer des points interdits ;
- aucune compatibilité n'est exigée après la restauration, il est seulement demandé que  $h$  décroisse d'un ratio fixe ;
- dans notre premier algorithme, le nombre de couples  $(h(x^k), f_0(x^k))$  introduits dans le filtre peut-être soit le minimum possible pour assurer l'existence d'un point d'accumulation stationnaire.

### Hypothèses.

On développe des algorithmes qui génèrent des suites  $(x^k)$  et  $(z^k)$  dans  $\mathbb{R}^n$ . Les hypothèses générales utilisées dans ce travail sont les suivantes :



- (H1) les itérés  $(x^k)$  et  $(z^k)$  restent dans une région compacte et convexe  $X \subset \mathbb{R}^n$  ;
- (H2) les fonctions  $f_i(\cdot)$ , pour  $i = 0, 1, \dots, m$ , sont Lipschitz continûment différentiables dans un ensemble ouvert contenant  $X$  ;
- (H3) tout point d'accumulation admissible  $\bar{x} \in X$  de la suite  $(x^k)$  satisfait la condition de qualification de Mangasarian-Fromovitz (M-F) : les gradients  $\nabla f_i(\bar{x})$ ,  $i \in \mathcal{E}$ , sont linéairement indépendants, et il existe une direction  $d \in \mathbb{R}^n$  telle que  $A_{\mathcal{E}}(\bar{x})d = 0$  et  $A_{\bar{\mathcal{I}}}(\bar{x})d < 0$ , avec  $\bar{\mathcal{I}} = \{i \in \mathcal{I} \mid f_i(\bar{x}) = 0\}$ .

### Les ensembles linéarisés.

Nous associons à chaque  $z \in \mathbb{R}^n$  et à l'ensemble

$$\{x \in \mathbb{R}^n \mid f_{\mathcal{E}}(x) = f_{\mathcal{E}}(z), f_{\mathcal{I}}(x) \leq f_{\mathcal{I}}^+(z)\}$$

l'ensemble linéarisé

$$L(z) = \{x \in \mathbb{R}^n \mid A_{\mathcal{E}}(z)(x - z) = 0, f_{\mathcal{I}}(z) + A_{\mathcal{I}}(z)(x - z) \leq f_{\mathcal{I}}^+(z)\}. \quad (2.3)$$

On vérifie facilement que :

- la condition de Mangasarian-Fromowitz en un point admissible  $z$  est équivalente à la condition suivante : les lignes de  $A_{\mathcal{E}}(z)$  sont linéairement indépendantes et l'ensemble  $L(z)$  a un point intérieur, au sens où  $y \in L(z)$  et  $f_{\mathcal{I}}(z) + A_{\mathcal{I}}(z)(y - z) < f_{\mathcal{I}}^+(z)$  (condition de Slater) ;
- les conditions de Karush-Kuhn-Tucker (KKT) pour (P) en  $z$  coïncident avec les conditions de KKT en  $z$  pour le problème de minimisation de  $f_0(\cdot)$  sur  $L(z)$ . Ces conditions sont aussi équivalentes au fait qu'il n'existe pas de direction de descente admissible à partir de  $z$  dans  $L(z)$ .

### Conditions d'optimalité.

Voici quelques commentaires sur les conditions d'optimalité et le sens que nous donnerons à l'expression *point stationnaire*. La *direction de Cauchy projetée* ou *direction du gradient projeté* associée à chaque  $z \in \mathbb{R}^n$  est

$$d_c(z) = P_{L(z)}(z - \nabla f_0(z)) - z, \quad (2.4)$$

où  $P_{\Gamma}(w)$  désigne la projection orthogonale de  $w \in \mathbb{R}^n$  sur l'ensemble fermé  $\Gamma \subset \mathbb{R}^n$ .

La direction du gradient projeté est bien connue. Voir par exemple Bertsekas [6]. Elle satisfait  $d_c(z) = 0$  si et seulement s'il n'existe pas de direction de descente admissible à partir de  $z$  dans  $L(z)$ . Nous en déduisons qu'en un point admissible  $z$ , les conditions de KKT sont équivalentes à  $d_c(z) = 0$ . Si  $d_c(z) \neq 0$ , alors  $\nabla f_0(z)^T d_c(z) < 0$ . Cette direction est le principal outil utilisé par Martínez et Svaiter [55] pour définir une condition d'optimalité à mi-chemin entre les conditions de KKT et de Fritz-John. Un point admissible  $\bar{x}$  satisfait la condition d'optimalité de Martínez-Svaiter si et seulement si

$$\liminf_{x \rightarrow \bar{x}} \|d_c(x)\| = 0. \quad (2.5)$$

Cette condition d'optimalité est constructive : nos algorithmes produisent des points admissibles qui satisfont (2.5). Ces points seront appelés *stationnaires*.

Nous avons deux possibilités : ou on considère ceci dans notre travail et nous n'utilisons pas la condition Fromowitz, ou on l'utilise et on rappelle que dans ce cas là KKT et Martínez-Svaiter sont des conditions équivalentes. On choisit la deuxième.

## 2.1 L'algorithme

Nous allons présenter le schéma de base de la méthode, sans spécifier les algorithmes internes utilisés pour calculer les pas de admissibilité et d'optimisation. Ensuite, nous ferons quelques hypothèses sur ces pas et montrerons que toute la suite générée par l'algorithme a un point d'accumulation stationnaire. Nous montrons dans la section suivante que des méthodes classiques satisfont ces hypothèses.

**Algorithme 2.1.1** *Algorithme de filtre.*

DONNÉES

$x^0 \in \mathbb{R}^n$ ,  $F_0 = \emptyset$ ,  $\mathcal{F}_0 = \emptyset$ ,  $\alpha \in (0, 1)$ .

$k = 0$ .

RÉPÉTER

$(\tilde{f}_0, \tilde{h}) = (f_0(x^k) - \alpha h(x^k), (1 - \alpha)h(x^k))$ .

Construire l'ensemble  $\tilde{F}_k = F_k \cup \{(\tilde{f}_0, \tilde{h})\}$ .

Définir l'ensemble  $\tilde{\mathcal{F}}_k = \mathcal{F}_k \cup \{x \in \mathbb{R}^n \mid f_0(x) \geq \tilde{f}_0, h(x) \geq \tilde{h}\}$ .

*Phase de admissibilité :*

- si  $h(x^k) = 0$  alors prendre  $z^k = x^k$ ,
- sinon calculer  $z^k \notin \bar{\mathcal{F}}_k$  tel que  $h(z^k) < (1 - \alpha) h(x^k)$ ;
- si impossible, arrêt de l'algorithme : échec.

*Phase d'optimalité :*

- si  $z^k$  est stationnaire, arrêt de l'algorithme : succès;
- sinon calculer  $x^{k+1} \notin \bar{\mathcal{F}}_k$  tel que
- $x^{k+1} \in L(z^k)$  et  $f_0(x^{k+1}) \leq f_0(z^k)$ .

*Mise à jour du filtre :*

- si  $f_0(x^{k+1}) < f_0(x^k)$  alors
- $F_{k+1} = F_k, \quad \mathcal{F}_{k+1} = \mathcal{F}_k, \quad (\text{itération-}f_0)$
- sinon
- $F_{k+1} = \bar{F}_k, \quad \mathcal{F}_{k+1} = \bar{\mathcal{F}}_k. \quad (\text{itération-}h)$

$k \leftarrow k + 1.$

Un point clé de l'algorithme est la construction du filtre : au début de chaque itération, le couple  $(f_0(x^k) - \delta, h(x^k) - \delta)$ , avec  $\delta = \alpha h(x^k)$ ,  $\alpha \in (0, 1)$ , est *temporairement* introduit dans le filtre. A la fin de l'itération, on ne garde ce couple que si l'on n'a pas réduit  $f_0$ . L'algorithme utilise le filtre et l'ensemble interdit qui lui est associé. Nous devons avoir à l'esprit que l'ensemble interdit n'est jamais construit, il nous aide simplement à comprendre le processus.

### Arrêt de l'algorithme.

Deux situations peuvent se produire :

- (i) un point stationnaire est obtenu. L'algorithme a normalement convergé;
- (ii) l'algorithme de admissibilité échoue. En fonction de la méthode utilisée pour calculer le pas de admissibilité, ceci peut se produire du fait de l'existence d'un point stationnaire  $\bar{x}$  de  $h(\cdot)$ , avec  $h(\bar{x}) \neq 0$ .

Désormais on va supposer que l'algorithme génère des suites infinies  $(x^k)$  et  $(z^k)$ , que les hypothèses (H1-H3) sont satisfaites, et de plus que les conditions suivantes (sur les algorithmes internes) sont respectées.

### L'hypothèse principale.

Étant donnée une itéré  $x^k$ , on définit *l'écart du filtre* en  $x^k$  par

$$H_k = \min \{1, \min \{h^j \mid (f_0^j, h^j) \in F_k, f_0^j \leq f_0(x^k)\}\}. \quad (2.6)$$

Cette quantité est illustrée par la figure 2.1.

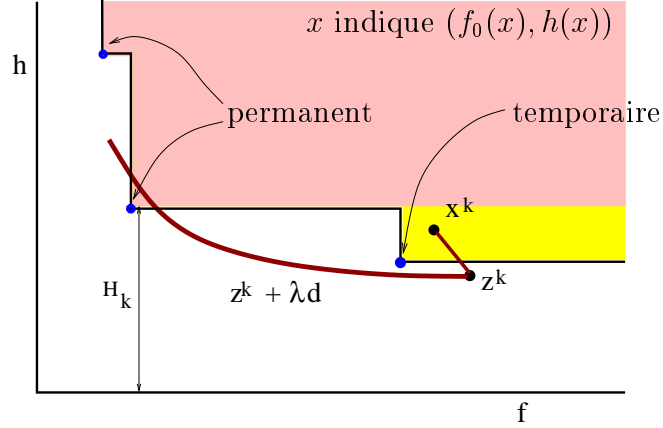


FIG. 2.1 – Exemple d'ensemble  $\bar{F}_k$  et de la quantité  $H_k$ .

Notre hypothèse principale est :

(H4) Étant donné un point admissible non stationnaire  $\bar{x} \in X$ , il existe un voisinage  $V$  de  $\bar{x}$  tel que pour tout  $x^k \in V$ ,

$$f_0(x^k) - f_0(x^{k+1}) = \Omega(\sqrt{H_k}). \quad (2.7)$$

On note que (H4) est une condition locale. L'égalité (2.7) signifie qu'il existe  $\gamma > 0$ , dépendant de  $\bar{x}$ , tel que si  $x^k$  est proche de  $\bar{x}$ , alors  $f_0(x^k) - f_0(x^{k+1}) \geq \gamma \sqrt{H_k}$ .

Compte-tenu de nos hypothèses et de la construction de l'algorithme, nous avons les propriétés suivantes :

- $x^{k+p} \notin \mathcal{F}_{k+1}$  pour tous  $k \in \mathbb{N}$  et  $p \geq 1$  ;
- étant donné  $k \in \mathbb{N}$ , au moins une des deux situations suivantes est vérifiée :

$$\begin{aligned} (i) \quad & h(x^{k+1}) \leq (1 - \alpha) h(x^k), \\ (ii) \quad & f_0(x^{k+1}) \leq f_0(x^k) - \alpha h(x^k); \end{aligned}$$

- $h^j > 0$  pour tous  $k \in \mathbb{N}$  et  $j \in \mathbb{N}$  tel que  $(f_0^j, h^j) \in F_k$ . Et encore,  $H_k > 0$  pour tous  $k \in \mathbb{N}$ .

Dans l'Algorithme 2.1.1, le couple  $(\tilde{f}_0, \tilde{h})$  est inclut dans le filtre à la fin de l'itération si et seulement s'il s'agit d'une itération- $h$ . Si  $\tilde{h} = h(x^k) = 0$ , alors  $z^k = x^k$  et  $f_0(x^{k+1}) < f_0(z^k)$ , donc l'itération  $k$  est une itération- $f_0$ .

Voici la suite des résultats.

**Lemme 2.1.2** *Soit  $\bar{x} \in X$  un point non stationnaire. Alors il existe un voisinage  $V$  de  $\bar{x}$  tel que pour tout  $x^k \in V$ , l'itération  $k$  est une itération- $f_0$ .*

□

**Lemme 2.1.3** *Supposons que  $(x^k)_{k \in \mathbb{N}}$  n'ait pas de point d'accumulation stationnaire. Alors, pour  $k$  suffisamment grand, l'itération  $k$  est une itération- $f_0$ .*

□

**Théorème 2.1.4** *La suite  $(x^k)$  a un point d'accumulation stationnaire.*

□

**L'hypothèse (H4).**

On postule cette hypothèse pour chaque itération complète. Quoique cela peut être difficile à vérifier pour des algorithmes spécifiques, son interprétation est simple : proche d'un point non stationnaire, le pas d'optimalité domine, et la réduction de  $f_0$  est grand. L'écart du filtre  $H_k$  indique combien  $h$  peut augmenter dans le pas tangentiel, et étant tangentiel, ou s'attend à ce que  $h$  change avec le carré de la variation de  $x$ . Dans un pas tangentiel efficace  $f_0$  peut varier linéairement avec la variation de  $x$ , et alors (H4) sera avéré.

Maintenant nous montrons comment (H4) peut être remplacé par des hypothèses plus simples faites séparément pour les pas d'admissibilité et d'optimalité.

**Condition sur le pas de admissibilité.**

**(H5)** A chaque itération, la phase de admissibilité permet de satisfaire

$$h(x^k) - h(z^k) = \Omega(\|z^k - x^k\|). \quad (2.8)$$

**Condition sur le pas d'optimalité.**

**(H6)** Étant donné un point admissible non stationnaire  $\bar{x} \in X$ , il existe un voisinage  $V$  de  $\bar{x}$  tel que pour tout  $x^k \in V$ ,

$$f_0(z^k) - f_0(x^{k+1}) = \Omega(\sqrt{H_k}). \quad (2.9)$$

On montre deux résultats par rapport ces hypothèses.

**Lemme 2.1.5** *(H5) et (H6) impliquent (H4).*

**Lemme 2.1.6** *Supposons que (H5-H7) soient vérifiées. Alors étant donné un point admissible non stationnaire  $\bar{x} \in X$ , il existe un voisinage  $V$  de  $\bar{x}$  tel que pour tout*

itéré  $x^k \in V$ ,

$$f_0(x^k) - f_0(x^{k+1}) = \Omega(\|x^{k+1} - x^k\|).$$

## 2.2 Algorithmes internes

Dans cette section, on s'intéresse aux pas internes utilisés à chaque itération de l'algorithme principal. Nous supposons que l'Algorithme 2.1.1 a généré des suites infinies  $(x^k)$  et  $(z^k)$  et que les Hypothèses (H1-H3) sont satisfaites.

### Calcul du pas de admissibilité.

Le but de la phase de admissibilité est de trouver un point  $z^k \notin \bar{\mathcal{F}}_k$  tel que  $h(z^k) < (1 - \alpha)h(x^k)$ . Le procédé mis en œuvre peut être n'importe quel algorithme itératif qui réduise  $h$  (en un nombre fini d'étapes) car comme nous avons vu ci-dessus, pour toutes les entrées du filtre  $(f_0^j, h^j) \in F_k$ , on a  $h^j > 0$ . Le pas de admissibilité étudié par Martínez [51] satisfait l'hypothèse (H5) et il peut être appliqué directement dans notre cadre. Nous n'allons donc pas le détailler. L'algorithme de admissibilité peut échouer si  $h(\cdot)$  a un point stationnaire qui n'est pas admissible pour  $(P)$ . Dans ce cas, la méthode échoue.

### Calcul du pas d'optimalité.

Le pas tangentiel doit permettre de trouver  $x^{k+1}$  dans l'ensemble linéarisé  $L(z^k)$  tel que  $f_0(x^{k+1}) \leq f_0(z^k)$ , et de sorte que  $x^{k+1} \notin \bar{\mathcal{F}}_k$ . On va décrire une méthode à régions de confiance générale, et montrer que le pas qui en résulte satisfait les hypothèses (H6) et (H7).

### Le modèle quadratique.

Soit  $z^k \in X$  généré par l'Algorithme 2.1.1 dans la phase de admissibilité. L'algorithme de régions de confiance associe à  $z^k$  un modèle quadratique de  $f_0$ ,

$$x \in \mathbb{R}^n \mapsto m_k(x) = f_0(z^k) + \nabla f_0(z^k)^T (x - z^k) + \frac{1}{2}(x - z^k)^T B_k (x - z^k), \quad (2.10)$$

où  $B_k$  est une matrice  $n \times n$  symétrique. Cette matrice peut être une approximation de  $\nabla^2 f_0(z^k)$ , ou tout autre matrice vérifiant l'hypothèse (H8) ci-dessous. Normalement,  $B_k$  est une approximation du hessien d'une fonction lagrangienne, auquel cas  $m_k$  n'est pas un modèle de  $f_0$ ; nous ne nous sommes pas intéressés à ce point.

**(H8)** Il existe  $\beta > 0$  tel que le modèle quadratique (2.10) satisfasse  $\|B_k\| \leq \beta$  pour tout  $k \in \mathbb{N}$ .

L'algorithme de régions de confiance utilise un rayon  $\Delta > 0$  et il calcule un pas  $d(z^k, \Delta) \in \mathbb{R}^n$  tel que  $\|d(z^k, \Delta)\| \leq \Delta$ . Nous définissons la *prédiction de réduction* pour le pas  $d(z^k, \Delta)$  comme

$$\text{pred}(z^k, \Delta) = m_k(z^k) - m_k(z^k + d(z^k, \Delta)), \quad (2.11)$$

et la *réduction véritable* comme

$$\text{ared}(z^k, \Delta) = f_0(z^k) - f_0(z^k + d(z^k, \Delta)). \quad (2.12)$$

Dans le calcul du pas tangentiel qui sera discuté ci-dessous, on fait les choix suivants pour simplifier le procédé :

- (1) chaque itération commence avec un rayon  $\Delta \geq \Delta_{min}$  où  $\Delta_{min} > 0$  est fixé. La choix de ce  $\Delta$  n'est pas important pour la théorie, et on prend celui qui est issu de l'itération précédente ;
- (2) un pas  $d(z^k, \Delta)$  n'est accepté que si la condition de décroissance

$$\text{ared}(z^k, \Delta) > \eta \text{pred}(z^k, \Delta) \quad (2.13)$$

est satisfaite, pour  $\eta \in (0, 1)$  donné ;

- (3) on résout approximativement le problème

$$\begin{aligned} & \text{minimiser } m_k(x) \\ & x \in L(z^k) \\ & \|x - z^k\| \leq \Delta, \end{aligned} \quad (2.14)$$

où  $\|\cdot\|$  est une norme quelconque sur  $\mathbb{R}^n$ .

Voici ce que nous entendons par "résoudre approximativement" le problème (2.14). Étant donné  $z \in X$  et l'ensemble  $L(z)$ , la direction du gradient projeté est définie par

$$d_c(z) = P_{L(z)}(z - \nabla f_0(z)) - z. \quad (2.15)$$

Définissons

$$\varphi(z) = - \nabla f_0(z)^T \frac{d_c(z)}{\|d_c(z)\|}.$$

Alors  $\varphi$  donne la vitesse de décroissance de  $f_0$  au long de  $d_c$ .

Comme d'habitude, on dénote  $d_c^k = d_c(z^k)$ ,  $\varphi^k = \varphi(z^k)$ . Comme nous avons dit dans la introduction, au début de cette section, si  $z$  est un point admissible non stationnaire alors  $\varphi(z) > 0$ .

Maintenant on va utiliser des résultats connus sur minimisation de  $m_k(\cdot)$  le long d'une direction – vois la discussion du point de Cauchy en [65]. Définissons le point de Cauchy généralisé comme le minimiseur de  $m_k(\cdot)$  le long de  $d_c$ , dans la région de confiance  $\{x \in \mathbb{R}^n \mid \|x - z^k\| \leq \Delta\}$ ,

$$x_c = \operatorname{argmin} \{m_k(x) \mid \|x - z^k\| \leq \Delta, x = z^k + \lambda d_c^k, \lambda \in [0, 1]\}.$$

On sait que

$$m_k(z^k) - m_k(x_c) \geq \frac{\xi \varphi^k}{2} \min \left\{ \frac{\varphi^k}{\|B_k\|}, \|d_c^k\|, \Delta \right\}$$

où  $\xi$  dépend de la norme utilisée.

Utilisant la Hypothèse (H8), ça peut être réécrit comme

$$m_k(z^k) - m_k(x_c) \geq \frac{\xi \varphi^k}{2} \min \left\{ \frac{\varphi^k}{\beta}, \|d_c^k\|, \Delta \right\}. \quad (2.16)$$

On accepte comme une solution approchée de (2.14), quelque solution admissible pour ce problème qui satisfait (2.16).

Après établir le pas de région de confiance on va étudier ses propriétés.

### Algorithme 2.2.1 Pas d'optimalité.

DONNÉES

$\eta \in (0, 1)$ ,  $\Delta_{min} > 0$ ,  $z^k \notin \bar{\mathcal{F}}_k$ ,  $\Delta = \Delta^0 \geq \Delta_{min}$ .

RÉPÉTER

Calculer  $d = d(z^k, \Delta)$  tel que  $\|d\| \leq \Delta$ ,  $z^k + d \in L(z^k)$  et

$$\operatorname{pred}(z^k, \Delta) \geq \frac{\xi \varphi^k}{2} \min \left\{ \frac{\varphi^k}{\|B_k\|}, \|d_c^k\|, \Delta \right\}.$$

Calculer  $\operatorname{ared}(z^k, \Delta) = f_0(z^k) - f_0(z^k + d)$ .

si  $z^k + d \notin \bar{\mathcal{F}}_k$  et  $\operatorname{ared}(z^k, \Delta) > \eta \operatorname{pred}(z^k, \Delta)$  alors

prendre  $x^{k+1} = z^k + d$ ,  $\Delta_k = \Delta$

arrêt de l'algorithme : succès ;

sinon  $\Delta = \Delta/2$ .



Maintenant on va étudier le pas d'optimalité proche d'un point admissible non stationnaire  $\bar{x} \in X$ . Le premier lemme dit que si on ignore le filtre, alors le pas de région de confiance est grand proche de  $\bar{x}$ .

**Lemme 2.2.2** *Soit  $\bar{x} \in X$  un point admissible non stationnaire qui satisfait une condition M-F. Alors il existe un voisinage  $\tilde{V}$  de  $\bar{x}$ ,  $\tilde{\Delta} \in (0, \Delta_{min})$  et une constante  $\tilde{c} > 0$  telle que pour tout  $z^k \in \tilde{V}$ ,*

$$(i) \text{ pour tout } \Delta > 0, \text{ pred}(z^k, \Delta) \geq \tilde{c} \min\{\Delta, \tilde{\Delta}\},$$

$$(ii) \text{ pour tout } \Delta \in (0, \tilde{\Delta}), \text{ ared}(z^k, \Delta) > \eta \text{ pred}(z^k, \Delta) \geq \eta \tilde{c} \Delta.$$

□

Algorithme 2.2.1 dans l'itération  $k$  commence avec  $\Delta^0 \geq \Delta_{min}$  et itère en faisant  $\Delta^j = 2^{-j} \Delta^0$ ,  $j = 0, 1, \dots$  et calculant pour chaque  $\Delta^j$  le pas  $d(z^k, \Delta^j)$ . Si  $\Delta^j < \tilde{\Delta}$ , alors la condition  $\text{ared}(z^k, \Delta^j) > \eta \text{ pred}(z^k, \Delta^j)$  est satisfaite : le rayon  $\Delta^j$  peut seulement être rejeté si  $z^k + d(z^k, \Delta^j) \in \bar{\mathcal{F}}_k$ .

Par construction,  $z^k \notin \bar{\mathcal{F}}_k$ , car  $h(z^k) < (1 - \alpha)h(x^k)$ . Comme  $\bar{\mathcal{F}}_k$  est un ensemble fermé, nous avons que pour  $\Delta^j$  suffisamment petit,  $z^k + d(z^k, \Delta^j) \notin \bar{\mathcal{F}}_k$ . Ceci montre que l'algorithme toujours termine.

**Lemme 2.2.3** *Soit  $\bar{x} \in X$  un point admissible non stationnaire satisfaisant une condition M-F, et supposons que (2.8) vaut. Alors il existe un voisinage  $V$  de  $\bar{x}$  tel que pour  $x^k \in V$ ,*

$$f_0(z^k) - f_0(x^{k+1}) = \Omega(\sqrt{H_k}), \quad (2.17)$$

$$f_0(z^k) - f_0(x^{k+1}) = \Omega(\|x^{k+1} - z^k\|), \quad (2.18)$$

où  $x^{k+1} = z^k + d(z^k, \Delta)$  est calculé par l'Algorithme 2.2.1.

### Améliorations.

On présente trois améliorations pour nos algorithmes. D'abord, nous améliorons l'analyse de convergence, en montrant que sous les hypothèses (H5,H6) les valeurs de la fonction objectif convergent toujours, limitant ainsi beaucoup la possibilité d'atteindre des points d'accumulation non stationnaire, spécialement quand (H7) est ajouté. Alors nous montrons comment un petit changement dans l'algorithme principal exclue la possibilité de gérer des points d'accumulation non stationnaires. Après, on discute un pas d'optimalité simplifié, en utilisant les matrices jacobiennes déjà calculées dans la phase de admissibilité en place de les recalculer en  $z^k$ .

# Conclusion

Dans la première partie, j'étudie, dans le cadre de l'optimisation convexe, quelques exemples de chemins centraux à l'allure mouvementée. Certains d'entre eux sont formés d'un nombre infini de segments centraux de longueur fixe ; d'autres manifestent des oscillations avec une variation infinie. Ces chemins peuvent se rencontrer même si les données du problème sont infiniment différentiables. Nous concluons que convexité, même avec différentiabilité à quelque degré, encore que les fonctions qui parties au problème soient  $C^\infty$ , n'est pas suffisante pour empêcher les formes pathologiques de la trajectoire centrale. Le résultat est en effet surprenant, puisque Monteiro et Zhou [60] garantissent que, si les fonctions qui entre dans la définition du problème sont analytiques, conjointement avec d'autres hypothèses raisonnables, alors la trajectoire centrale primale est une courbe qui converge vers un seul point dans l'ensemble optimal.

Nous désirons étudier, dans des travaux futures, du point de vue de la complexité, quelles sont les conséquences algorithmiques de nos exemples sur les méthodes de points intérieurs pour l'optimisation non linéaire.

Dans la deuxième partie de ma thèse, j'ai introduit un algorithme de filtre globalement convergent pour l'optimisation non linéaire. Dans le format actuel de cet algorithme, l'effet Maratos [49] peut se produire. Il serait souhaitable d'inclure dans l'algorithme une technique pour éviter l'effet Maratos (troncature du pas). J'ai aussi l'intention de proposer un algorithme de points intérieurs pour l'optimisation non linéaire qui utilise la technique de filtre, sous la forme présentée dans cette thèse, et les idées de l'algorithme proposé dans [9, 10].

## **Annexe A**

### **Examples of ill-behaved central paths in convex optimization**

## **Annexe B**

# **A globally convergent filter method for nonlinear programming**

# Bibliographie

- [1] J. Abadie and J. Carpentier. Generalization of the Wolfe reduced-gradient method to the case of nonlinear constraints. In R. Fletcher, editor, *Optimization*, pages 37–47. Academic Press, New York, 1968.
- [2] I. Adler and R. D. C. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50 :29–51, 1991.
- [3] K. M. Anstreicher. Linear programming in  $O([n^3/\ln n]L)$  operations. *SIAM Journal on Optimization*, 9 :803–812, 1999.
- [4] A. Auslender, R. Cominetti, and M. Haddou. Asymptotic analysis for penalty and barrier methods in convex and linear programming. *Mathematics of Operations Research*, 22 :43–62, 1997.
- [5] D. A. Bayer and J. C. Lagarias. The nonlinear geometry of linear programming, Part I : Affine and projective scaling trajectories. *Transactions of the American Mathematical Society*, 314(2) :499–526, 1989.
- [6] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, Massachusetts, 1995.
- [7] P. T. Boggs and J. W. Tolle. Sequential quadratic programming. *Acta Numer.*, 4 :1–51, 1996.
- [8] R. H. Byrd. Robust trust region methods for constrained optimization. Third SIAM Conference on Optimization, 1987.
- [9] R. H. Byrd, J. C. Gilbert, and J. Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1) :149–185, 2000.

- [10] R. H. Byrd, M. E. Hribar, and J. Nocedal. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, 9(4) :877–900, 1999.
- [11] R. H. Byrd, R. B. Schnabel, and G. A. Schultz. A trust region algorithm for nonlinearity constrained optimization. *SIAM Journal on Optimization*, 24 :1152–1170, 1987.
- [12] T. J. Carpenter, I. J. Lustig, J. M. Mulvey, and D. F. Shanno. Higher order predictor-corrector interior point methods with application to quadratic objectives. *SIAM Journal on Optimization*, 3 :696–725, 1993.
- [13] M. R. Celis, J. E. Dennis, and R. A. Tapia. A trust region strategy for nonlinear equality constrained optimization. In P. T. Boggs, R. H. Byrd, and R. B. Schnabel, editors, *Numerical Optimization 1984*, pages 71 – 82. SIAM, Philadelphia, 1985.
- [14] T. Champion. *Asymptotic Convergence of Penalty Trajectories in Convex Programming with Multiple Solutions*. PhD thesis, Département de Mathématique, Université de Montpellier II, France, 2001.
- [15] C. Chen and O. L. Mangasarian. Smoothing methods for convex inequalities and linear complementary problems. *Mathematical Programming*, 71 :1–112, 1995.
- [16] C. Chen and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementary problems. *Computational Optimization and Applications*, 5 :97–138, 1996.
- [17] T. F. Coleman and Y. Li. On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Mathematical Programming*, 67 :189–224, 1994.
- [18] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optimization*, 6(2) :418–445, May 1996.
- [19] R. Cominetti. Nonlinear averages and convergence of penalty trajectories in convex programming. In Rainer Tichatschke Michel Théra, editor, *Ill-posed variational problems and regularization techniques*, volume 477 of *Lecture Notes in Economics and Mathematical System*, pages 65–78. Springer Verlag, Berlin, 1999.

- [20] A. R. Conn, N. I. M. Gould, D. Orban, and P. L. Toint. A primal-dual trust region algorithm for minimizing a non-convex function subject to general inequality and linear equality constraints. Technical Report RAL-TR-1999-054, Computation Science and Engineering Department, Atlas Centre, Rutherford Appleton Laboratory, Oxon, OX11 0QX, 1999.
- [21] G. Dantzig. Maximization of a linear function of variables subject to linear inequalities. In T. C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 339–347. John Wiley, New York, 1951.
- [22] D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming*. Mathematics and its Applications 277. Kluwer Academic Publishers, Dordrecht, 1992.
- [23] D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming. Algorithms and Complexity*. PhD thesis, Faculty of Mathematics and Informatics, September 1992.
- [24] J. E. Dennis, M. Heinkenschloss, and L. N. Vicente. Trust-region interior-point SQP algorithms for a class of nonlinear programming problems. Technical Report 94-45, Department of Computational and Applied Mathematics, Rice University, 1994.
- [25] A. S. El-Bakry, R. A. Tapia, T. Tsuchya, and Y. Zhang. On the formulation and theory of the Newton interior-point method for nonlinear programming. *Journal of Optimization Theory and Applications*, 89 :507–541, 1996.
- [26] A. V. Fiacco and G. P. McCormick. *NonLinear Programming : Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968. Reprint : Volume 4 of *SIAM Classics in Applied Mathematics*, SIAM Publications, Philadelphia, PA 19104–2688, USA, 1990.
- [27] W. H. Fleming. *Functions of Several Variables*. Addison-Wesley Publishing Company, United States, 1965.
- [28] R. Fletcher, N. Gould, S. Leyffer, P. Toint, and A. Wächter. Global convergence of trust-region and SQP-filter algorithms for general nonlinear programming. *SIAM Journal on Optimization*, 2002. to appear.
- [29] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. Technical Report NA/171, Dundee University, Dept. of Mathematics, 1997. Revised 2000.

- [30] R. Fletcher, S. Leyffer, and P. L. Toint. On the global convergence of an SLP-filter algorithm. Technical Report NA/183, Dundee University, Dept. of Mathematics, 1999.
- [31] R. Fletcher, S. Leyffer, and P. L. Toint. On the global convergence of a filter-SQP algorithm. *SIAM Journal on Optimization*, 13(1) :44–59, 2002.
- [32] A. Forsgren and P. E. Gill. Primal-dual interior methods for nonconvex nonlinear programming. *SIAM Journal on Optimization*, 8(4) :1132–1152, 1998.
- [33] K. R. Frisch. The logarithmic potential method for convex programming. Unpublished manuscript, Institute of Economics, University of Oslo, Oslo, Norway, May 1955.
- [34] D. M. Gay, M. L. Overton, and M. H. Wright. A primal-dual interior method for nonconvex nonlinear programming. In Yuan Y, editor, *Advances in Nonlinear Programming*, pages 31–56. Kluwer Academic Publishers, Dordrecht, 1998.
- [35] J. C. Gilbert, C. C. Gonzaga, and E. W. Karas. Examples of ill-behaved central paths in convex optimization. Technical Report 4179, INRIA, Institut National de Recherche en Informatique et en Automatique, 2001.
- [36] F. A. M. Gomes, M. C. Maciel, and J. M. Martínez. Nonlinear programming algorithms using trust regions and augmented Lagrangians with nonmonotone penalty parameters. *Mathematical Programming*, 84(1) :161–200, January 1999.
- [37] J. Gondzio. Multiple centrality corrections in a primal dual method for linear programming. *Comp. Opt. and Appl.*, 6 :137–156, 1996.
- [38] C. C. Gonzaga. An algorithm for solving linear programming problems in  $O(n^3L)$  operations. In N. Megiddo, editor, *Progress in Mathematical Programming : Interior Point and Related Methods*, pages 1–28. Springer Verlag, New York, 1989.
- [39] C. C. Gonzaga. Path following methods for linear programming. *SIAM Review*, 34(2) :167–227, 1992.
- [40] C. C. Gonzaga, E. W. Karas, and M. Vanti. A globally convergent filter method for nonlinear programming. Optimization online, 2001.
- [41] J-B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, New York, 1993.
- [42] B. Jansen. *Interior Point Techniques in Optimization*. Applied Optimization 6. Kluwer Academic Publishers, Dordrecht, 1997.



- [43] N. K. Karmarkar. A new polynomial–time algorithm for linear programming. *Combinatorica*, 4 :373–395, 1984.
- [44] L. G. Khachiyan. A polynomial algorithm for linear programming. *Soviet Mathematics Doklady*, 20 :191–194, 1979.
- [45] V. Klee and G. Minty. How good is the simplex algorithm ? In O. Sisha, editor, *Inequalities III*. Academic Press, New York, 1972.
- [46] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. A unified approach to interior point algorithms for linear complementarity problems. In *Lecture Notes in Computer Science*, volume 538, Springer Verlag, Berlin, 1991.
- [47] M. Kojima, S. Mizuno, and A. Yoshise. A primal–dual interior point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming : Interior Point and Related Methods*, pages 29–47. Springer Verlag, New York, 1989.
- [48] K. Levenberg. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.*, 2 :164–168, 1944.
- [49] N. Maratos. *Exact Penalty Function Algorithms for Finite Dimensional and Control Optimization Problems*. PhD thesis, Imperial College Science Technology, University of London, 1978.
- [50] D. W. Marquardt. An algorithm for least squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11 :431–441, 1963.
- [51] J. M. Martínez. Inexact-restoration method with Lagrangian tangent decrease and a new merit function for nonlinear programming. *Journal of Optimization Theory and Applications*, 2001. To appear.
- [52] J. M. Martínez and E. A. Pilotta. Inexact restoration algorithms for constrained optimization. *Journal of Optimization Theory and Applications*, 104 :135–163, 2000.
- [53] J. M. Martínez and E. A. Pilotta. Inexact restoration methods for nonlinear programming : Advances and perspectives. In Qi, Teo, and Yang, editors, *Optimization and Control with Applications*. Kluwer, 2001. To appear.
- [54] J. M. Martínez and S. A. Santos. Métodos computacionais de otimização. 20.<sup>0</sup> Colóquio Brasileiro de Matemática - IMPA, July 1995. In Portuguese.

- [55] J. M. Martínez and B. F. Svaiter. A practical optimality condition without constraint qualifications for nonlinear programming. Technical report, Institute of Mathematics, University of Campinas, Brazil, 2001.
- [56] L. McLinden. An analogue of Moreau's proximation theorem, with application to the nonlinear complementarity problem. *Pacific Journal of Mathematics*, 88(1) :101–161, 1980.
- [57] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming : Interior Point and Related Methods*, pages 131–158. Springer Verlag, New York, 1989. Identical version in : *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, pages 1–35, 1986.
- [58] S. Mehrotra. On the implementation of a primal–dual interior point method. *SIAM Journal on Optimization*, 2(4) :575–601, 1992.
- [59] R. D. C. Monteiro and I. Adler. Interior path following primal–dual algorithms : Part I : Linear programming. *Mathematical Programming*, 44 :27–41, 1989.
- [60] R. D. C. Monteiro and F. Zhou. On the existence and convergence of the centra path for convex programming and some duality results. *Computational Optimization and Applications*, 10 :51–77, 1998.
- [61] J. J. Moré. Recent developments in algorithms and software for trust region methods. In A. Bachem, M. Grötschel, and B. Korte, editors, *Mathematical Programming, The Sate of the Art*. Springer Verlag, 1983.
- [62] J. J. Moré and D. C. Sorensen. Computing a trust region step. *SIAM Journal on Sci. Statist. Comput.*, 4 :553–572, 1983.
- [63] H. Mukai and E. Polak. On the use of approximations in algorithm for optimization problems with equality and inequality constraints. *SIAM Journal on Numerical Analysis*, 15 :674–693, 1978.
- [64] Y. E. Nesterov and A. S. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics 13*. SIAM, Philadelphia, 1994.
- [65] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, 1999.

- [66] E. Omojokun. *Trust Region Algorithms for Optimization with Nonlinear Equality and Inequality Constraints*. PhD thesis, Dept. of Computer Science, University of Colorado, 1991.
- [67] M. J. D. Powell. A hybrid method for nonlinear equations. In P. Rabinowitz, editor, *Numerical Methods for Nonlinear Algebraic Equations*. Gordon and Breach, London, 1970.
- [68] M. J. D. Powell. Convergence properties of a class of minimization algorithms. In O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, editors, *Nonlinear Programming 2*, pages 1–27. Academic Press, New York, 1975.
- [69] M. J. D. Powell and Y. Yuan. A trust region algorithm for equality constrained optimization. *Mathematical Programming*, 49 :189–211, 1991.
- [70] J. Renegar. A polynomial-time algorithm, based on Newton’s method, for linear programming. *Mathematical Programming*, 40 :59–93, 1988.
- [71] R. T. Rockafellar and R. Wets. *Variational Analysis*. Springer, 1998.
- [72] M. Rom and M. Avriel. Properties of the sequential gradient-restoration algorithm (SGRA), part 2 : Convergence analysis. *Journal of Optimization Theory and Applications*, 62 :99–126, 1989.
- [73] C. Roos, T. Terlaky, and J. P. Vial. *Theory and Algorithms for Linear Optimization : An Interior Point Approach*. John Wiley & Sons, Chichester, 1997.
- [74] J. B. Rosen. The gradient projection method for nonlinear programming, part 1, linear constraints. *SIAM Journal on Applied Mathematics*, 8 :181–217, 1960.
- [75] D. F. Shanno and R. J. Vanderbei. Interior-point methods for nonconvex nonlinear programming : Orderings and higher-order methods. *Mathematical Programming*, 87(2) :303–316, 2000.
- [76] T. Terlaky, editor. *Interior Point Methods of Mathematical Programming*. Kluwer Academic Press, 1996.
- [77] M. J. Todd. Potential-reduction methods in mathematical programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1995.
- [78] M. Ulbrich, S. Ulbrich, and L. N. Vicente. A globally convergent primal-dual interior-point filter method for nonconvex nonlinear programming. Technical Report 00-11, Dep. de Matemática, Universidade de Coimbra, April 2000.

- [79] P. M. Vaidya. An algorithm for linear programming which requires  $O((m+n)n^2 + (m+n)^{1.5}nL)$  arithmetic operations. *Mathematical Programming*, 47 :175–201, 1990. Condensed version in : *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 29–38, 1987.
- [80] R. J. Vanderbei and D. F. Shanno. An interior-point algorithm for nonconvex nonlinear programming. *Computational Optimization and Applications*, 13 :231–252, 1999.
- [81] S. A. Vavasis and Y. Ye. A primal-dual interior point method whose running time depends only on the constraint matrix. *Mathematical Programming*, 74 :79–120, 1996.
- [82] A. Wächter and L. T. Biegler. Global and local convergence of line search filter methods for nonlinear programming. Technical Report B-01-09, CAPD, Department of Chemical Engineering Carnegie Mellon University Pittsburgh, August 2001.
- [83] H. Whitney. Functions differentiable on the boundaries of regions. *Ann. of Math.*, 35 :482–485, 1934.
- [84] R. B. Wilson. *A Simplicial Algorithm for Convex Programming*. PhD thesis, Harvard University, Graduate School of Business Administration, 1963.
- [85] C. Witzgall, P. T. Boggs, and P. D. Domich. On the convergence behavior of trajectories for linear programming. In J. C. Lagarias and M. J. Todd, editors, *Mathematical Developments Arising from Linear Programming : Proceedings of a Joint Summer Research Conference held at Bowdoin College, Brunswick, Maine, USA, June/July 1988*, volume 114 of *Contemporary Mathematics*, pages 161–187. American Mathematical Society, Providence, Rhode Island, USA, 1990.
- [86] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM Publication, Philadelphia, 1997.
- [87] H. Yamashita. A globally convergent primal-dual interior point method for constrained optimization. Technical report, Mathematical Systems Institute Inc., Tokyo, Japan, 1992. Revised in March 1994.
- [88] H. Yamashita, H. Yabe, and T. Tanabe. A globally and superlinearly convergent primal- dual interior point trust region method for large scale constrained op-

timization. Technical report, Mathematical Systems Inc., Shinjuku-ku, Tokyo, Japan, July 1997.