



## Commande optimale

Joseph Frédéric Bonnans

► **To cite this version:**

| Joseph Frédéric Bonnans. Commande optimale . 2015, pp.29. hal-01247050

**HAL Id: hal-01247050**

**<https://hal.inria.fr/hal-01247050>**

Submitted on 21 Dec 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# COMMANDE OPTIMALE

VERSION PRÉLIMINAIRE D'UN ARTICLE POUR LES TECHNIQUES DE L'INGÉNIEUR

VERSION DE JANVIER 2015

J. FRÉDÉRIC BONNANS

DIRECTEUR DE RECHERCHE INRIA  
INRIA ET CENTRE DE MATHÉMATIQUES APPLIQUÉES,  
ECOLE POLYTECHNIQUE, PALAISEAU

**RÉSUMÉ.** L'objet de la commande optimale est l'optimisation de systèmes dynamiques suivant différents objectifs : atteinte d'une cible en temps ou énergie minimale, maximisation du rendement d'un processus industriel par exemple. Pour cela on joue à la fois sur des paramètres indépendants du temps et sur les commandes qui, elles, dépendent du temps. L'article analyse les conditions d'optimalité du premier et second ordre, et leur résolution par discrétisation temporelle, algorithme de tir, ou programmation dynamique.

**ABSTRACT.** The optimal control theory analyzes how to optimize dynamical systems with various criteria : reach a target in minimal time or minimal energy, maximize the efficiency of an industrial process for instance. This involves the optimization of both time independent parameters, and the control variables that are function of time. The article analyzes the first and second order optimality conditions, and the ways to solve them, by time discretization, the shooting algorithm, or dynamic programming.

**Mots-clés :** Systèmes dynamiques, suivi de trajectoire, temps minimal, conditions d'optimalité du second ordre, algorithme de tir, programmation dynamique.

**Keywords :** Dynamical systems, path following, minimal time, second order optimality conditions, shooting algorithm, dynamic programming.

**Éclairage historique.** La commande optimale est née dans les années 1950, avec les travaux de l'école de L.S. Pontriaguine [19] et de R. Bellman [5]. Pontriaguine a obtenu des conditions d'optimalité d'une trajectoire, qui ont permis la résolution explicite ou numérique de nombreux problèmes, en particulier de trajectoires aériennes ou spatiales. Bellman a énoncé le principe de programmation dynamique, qui s'applique également dans le cadre discret ou stochastique. Les approfondissements théoriques et algorithmiques de la théorie, et l'amélioration des moyens de calcul permettent maintenant de résoudre facilement un grand nombre de problèmes appliqués.

## TABLE DES MATIÈRES

Introduction	2
1. Notations	3
2. Cadre mathématique	4
2.1. Système dynamique commandé et critère associé	4
2.2. Différents types de contraintes	4
2.3. Un format général	5
3. Contraintes sur la commande	6
3.1. Dérivée du critère	6
3.2. Conditions du premier ordre avec contraintes sur la commande	7
3.3. Principe du minimum de Pontriaguine (PMP)	7
4. Principe de Pontriaguine avec contraintes aux deux bouts	9
4.1. Lagrangien et état adjoint	9
4.2. Conservation du préhamiltonien	11
4.3. Paramètres de décision	11
4.4. Horizon variable	11
4.5. Temps minimal, géodésiques	11
5. Algorithme de tir et discrétisation du problème	13
5.1. Problèmes sans contraintes	13

5.2. Problèmes avec contraintes aux deux bouts	13
5.3. Discrétisation du problème	14
6. Conditions d'optimalité du second ordre	15
6.1. Développements du critère réduit	15
6.2. Expression du hessien du critère réduit	15
6.3. Extension au cas de contraintes aux deux bouts	16
6.4. Lien avec l'algorithme de tir	16
7. Problèmes avec contraintes sur l'état	17
7.1. Principe de Pontriaguine	17
7.2. Ordre des contraintes sur l'état et conditions de jonction	18
7.3. Système d'optimalité alternatif	19
8. Problèmes affines en la commande	20
8.1. Cadre : coût final et temps optimal	20
8.2. Arcs bang et singuliers	21
8.3. Cas particuliers - Problèmes de dimension 2 et 3	21
8.4. Contraintes sur l'état	22
8.4.1. Saut de la dérivée de la fonction de commutation	22
8.4.2. Illustration : un problème proie-prédateur	23
9. Programmation dynamique	23
9.1. Principe de programmation dynamique	23
9.2. Equation HJB	24
9.3. Solution de viscosité	25
9.4. Problèmes en horizon infini	26
9.5. Problèmes de temps d'arrêt	26
9.6. Commande impulsionnelle	26
10. Schémas numériques pour l'équation HJB	26
10.1. Différences finies, état scalaire	27
10.2. État multidimensionnel	27
10.3. Problèmes en horizon infini	27
10.4. Problèmes de temps d'arrêt	28
10.5. Algorithmes semi lagrangiens	28
10.6. Bilan sur les méthodes numériques pour l'équation HJB	28
Conclusion / Orientations bibliographiques	28
Références	29

## INTRODUCTION

On dit qu'un système dynamique est commandé s'il est possible d'agir sur lui par des variables dépendant du temps, appelées *commandes*. Illustrons ce concept dans le cas d'un engin spatial, décrit par des variables de position et vitesse (dans  $\mathbb{R}^3$ )  $h$  et  $V$ , et une masse  $m > 0$ , soit 7 variables d'état. La dynamique est, omettant l'argument temps,  $\dot{h} = V$ ,  $m\dot{V} = F(h, V) + u$ , et  $\dot{m} = -c|u|$ . Ici  $c$  est une constante positive et  $F(h, V)$  correspond aux forces de gravité et (le cas échéant) aérodynamiques. La commande est la force appliquée, dont on a noté la norme euclidienne par  $|u|$ , soumise à une contrainte du type  $|u| \leq U$ . Étant donné un point initial fixé, on cherche à atteindre une cible (partie de l'espace d'état) en minimisant un compromis entre temps de parcours et énergie dépensée.

Pour l'implémentation en temps réel d'une commande il est nécessaire de prendre en compte les moyens d'observations et la reconstitution de l'état, prenant en compte des aspects de traitement du signal et le choix de l'électronique de commande. Au contraire, dans cet article, nous ne considérons que l'étude amont, dans laquelle on se donne un cadre déterministe et on calcule donc hors ligne une commande optimale. L'allure de cette dernière pourra guider la conception de la commande en temps réel.

L'exposé suivra d'abord l'approche de Lagrange et Pontriaguine qui consiste à étudier les variations d'une trajectoire optimale pour déterminer les propriétés de cette dernière. On analysera les conditions d'optimalité du premier et second ordre, en lien avec l'algorithme de tir, et avec une attention toute particulière pour les problèmes avec contraintes sur l'état. Dans une seconde partie, on abordera l'approche par programmation dynamique et équation de Hamilton-Jacobi-Bellman (HJB). Après avoir introduit

la notion de solution de viscosité de l'équation HJB, on présentera les problèmes de temps d'arrêt et de commande impulsienne, et on finira par l'étude des schémas de résolution numérique.

## 1. NOTATIONS

On note  $\mathbb{R}^n$  l'espace euclidien de dimension  $n$ , dont les éléments sont des vecteurs verticaux, de norme  $|x| := (\sum_{i=1}^n x_i^2)^{1/2}$ ; son dual  $\mathbb{R}^{n*}$  est identifié à l'ensemble des vecteurs horizontaux. On note  $A^\dagger$  la transposée d'une matrice  $A$ ,  $x \cdot y$  le produit scalaire de  $x$  et  $y$ , ainsi que  $\|x\|_s := (\sum_{i=1}^n |x_i|^s)^{1/s}$ , pour  $s \in [1, \infty[$ , et  $\|x\|_\infty := \max_{i=1}^n |x_i|$ . L'ensemble réduit au zéro de  $\mathbb{R}^n$  est noté  $\{0\}_{\mathbb{R}^n}$ ; l'ensemble des vecteurs de coordonnées toutes positives (resp. négatives) est appelé cône positif (resp. négatif) de  $\mathbb{R}^n$  et noté  $\mathbb{R}_+^n$  (resp.  $\mathbb{R}_-^n$ ). Ainsi, une partie de  $\mathbb{R}^n$  définie par un ensemble fini d'égalités et d'inégalités, soit

$$(1) \quad \left\{ \begin{array}{l} x \in \mathbb{R}^n; \quad g_i(x) = 0, \quad i = 1 \dots, n'; \\ g_i(x) \leq 0, \quad i = n' + 1 \dots, n' + n'' \end{array} \right\}$$

peut se réécrire comme  $g(x) \in K[n', n'']$ , où

$$(2) \quad K[n', n''] := \{0\}_{\mathbb{R}^{n'}} \times \mathbb{R}_-^{n''}.$$

L'ensemble  $K[n', n'']$  est convexe et fermé. On considèrera plus généralement des contraintes du type  $g(x) \in \mathcal{K}$ , où  $\mathcal{K}$  est une partie convexe fermée de  $\mathbb{R}^n$ . Soit  $\bar{x} \in \mathcal{K}$ . On appelle (i) *cône normal* à  $\mathcal{K}$  en  $\bar{x}$  l'ensemble

$$(3) \quad N_{\mathcal{K}}(\bar{x}) := \{y \in \mathbb{R}^n; \quad y \cdot (x - \bar{x}) \leq 0, \quad \text{pour tout } x \in \mathcal{K}\},$$

(ii) *cône des directions radiales* à  $\mathcal{K}$  en  $\bar{x}$  l'ensemble

$$(4) \quad \mathcal{R}_{\mathcal{K}}(\bar{x}) := \{z \in \mathbb{R}^n; \quad \bar{x} + \varepsilon z \in \mathcal{K}, \quad \text{pour } \varepsilon > 0 \text{ assez petit}\},$$

et (iii) *cône tangent* à  $\mathcal{K}$  en  $\bar{x}$  la fermeture notée  $T_{\mathcal{K}}(\bar{x})$  de  $\mathcal{R}_{\mathcal{K}}(\bar{x})$ . Quand le bord de  $\mathcal{K}$  est lisse,  $N_{\mathcal{K}}(\bar{x})$  se réduit à la demi droite engendrée par la normale extérieure en  $\bar{x}$  à  $\mathcal{K}$ , et  $T_{\mathcal{K}}(\bar{x})$  est le demi espace fermé des directions faisant un produit scalaire négatif avec cette normale. Si  $\mathcal{K} = K[n', n'']$ , et  $\bar{x} \in \mathcal{K}$ , on vérifie facilement que

$$(5) \quad N_{\mathcal{K}}(\bar{x}) = \{y \in \mathbb{R}^{n'} \times \mathbb{R}_+^{n''}; \quad y_j \bar{x}_j = 0, \quad j = n' + 1, \dots, n' + n''\},$$

$$(6) \quad \mathcal{R}_{\mathcal{K}}(\bar{x}) = T_{\mathcal{K}}(\bar{x}) = \{0\}_{\mathbb{R}^{n'}} \times \{y \in \mathbb{R}^{n''}; \quad y_j \leq 0 \text{ si } \bar{x}_j = 0, \quad j = n' + 1, \dots, n' + n''\}.$$

Soient  $A$  et  $B$  des matrices symétriques de taille  $n$ . On note  $A \succeq B$  si  $A - B$  est semi définie positive, et  $A \succ B$  si  $A - B$  est définie positive. Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  est dérivable on identifie sa jacobienne, notée  $Df(x)$  ou  $f'(x)$ , en  $x \in \mathbb{R}^n$ , à une matrice de taille  $p \times n$ . Quand  $p = 1$ , on note par  $\nabla f(x) := f'(x)^\dagger$  le gradient de  $f$  en  $x$ ; c'est l'élément de  $\mathbb{R}^n$  de coordonnées égales aux dérivées partielles de  $f$  en  $x$ . Si  $f(x, y)$  est fonction de  $x \in \mathbb{R}^n$  et  $y \in \mathbb{R}^q$  on note par  $D_x f(x, y)$  ou  $f_x(x, y)$  sa dérivée partielle, et par  $\nabla_x f(x, y)$  son gradient partiel. Si  $f$  est à valeurs scalaires,  $f_{xy}(x, y)$  désigne la matrice de taille  $n \times q$  de terme général  $\partial f(x, y) / \partial x_i \partial y_j$ .

On note la dérivée temporelle d'une fonction  $f$  du temps par  $\dot{f}(t)$ . Notons  $\mathcal{D}(0, T)$ , où  $T > 0$ , l'ensemble des fonctions de classe  $C^\infty$  à support compact sur  $]0, T[$ . Soit  $f \in L^1(0, T)$ . S'il existe  $g \in L^1(0, T)$  tel que

$$\int_0^T g(t) \varphi(t) dt = - \int_0^T f(t) \dot{\varphi}(t) dt, \quad \text{pour tout } \varphi \in \mathcal{D}(0, T),$$

on dit que  $g$  est la *dérivée faible* de  $f$ . Si  $f$  est de classe  $C^1$ , alors  $g = \dot{f}$  p.p. (presque partout). Soit  $s \in [0, \infty[$ . On note  $W^{1,s}(0, T)$  l'ensemble des fonction de  $L^s(0, T)$  de dérivée faible dans  $L^s(0, T)$ , de norme

$$\|y\|_{1,s} := \|y\|_{L^s(0,T)} + \|\dot{y}\|_{L^s(0,T)}.$$

On sait que  $W^{1,\infty}(0, T)$  est l'ensemble des fonctions lipschitziennes sur  $[0, T]$ . On définit les espaces de commande et d'état

$$(7) \quad \mathcal{U}_s := L^s(0, T)^m, \quad \mathcal{Y}_s := W^{1,s}(0, T)^n,$$

de norme associées notées  $\|\mathbf{u}\|_s$ ,  $\|\mathbf{y}\|_{1,s}$  (les fonctions du temps telles que commandes et état sont notées en gras), et on pose  $\mathcal{U} := \mathcal{U}_\infty$ ,  $\mathcal{Y} := \mathcal{Y}_\infty$ ,  $\|\mathbf{u}\| := \|\mathbf{u}\|_\infty$ ,  $\|\mathbf{y}\| := \|\mathbf{y}\|_{1,\infty}$ .

## 2. CADRE MATHÉMATIQUE

**2.1. Système dynamique commandé et critère associé.** Considérons des *systèmes dynamiques commandés* du type

$$(8) \quad \text{(i) } \dot{\mathbf{y}}(t) = f(\mathbf{u}(t), \mathbf{y}(t)), \quad t \in [0, T]; \quad \text{(ii) } \mathbf{y}(0) = \mathbf{y}^0.$$

Les données sont l'*horizon* ou *temps final*  $T > 0$ , la *dynamique*  $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ , supposée lipschitzienne et de classe  $C^\infty$ , et la *condition initiale*  $\mathbf{y}^0$ . On appelle  $\mathbf{u}(t)$  et  $\mathbf{y}(t)$  la *commande* et l'*état* à l'instant  $t$ . On choisit  $\mathcal{U}$  et  $\mathcal{Y}$  comme espaces pour la commande et pour l'état. Alors (8)(i) est une égalité dans  $L^\infty(0, T)^n$ , donc aussi p.p. Une variante du théorème de Cauchy-Lipschitz permet de vérifier que, pour tout  $(\mathbf{u}, \mathbf{y}^0) \in \mathcal{U}_1 \times \mathbb{R}^n$ , l'équation d'état (8) a dans  $\mathcal{Y}$  une solution unique notée  $\mathbf{y}[\mathbf{u}, \mathbf{y}^0]$  (ou  $\mathbf{y}[\mathbf{u}]$  si  $\mathbf{y}^0$  est fixé), et d'après le lemme de Gronwall, on a pour un certain  $C_f$  ne dépendant que de la constante de Lipschitz de  $f$  :

$$(9) \quad \|\mathbf{y}[\mathbf{u}', (\mathbf{y}^0)'] - \mathbf{y}[\mathbf{u}, \mathbf{y}^0]\|_\infty \leq C_f (\|\mathbf{u}' - \mathbf{u}\|_1 + |(\mathbf{y}^0)' - \mathbf{y}^0|).$$

De plus  $(\mathbf{u}, \mathbf{y}^0) \mapsto \mathbf{y}[\mathbf{u}, \mathbf{y}^0]$  est de classe  $C^\infty : \mathcal{U} \times \mathbb{R}^n \rightarrow \mathcal{Y}$  (appliquer le théorème des fonctions implicites (TFI) à l'équation d'état).

On dira que  $(\mathbf{u}, \mathbf{y}) \in \mathcal{U} \times \mathcal{Y}$  est une *trajectoire* si  $\mathbf{y} = \mathbf{y}[\mathbf{u}, \mathbf{y}(0)]$ . Dans la suite on effectue l'analyse autour de la *trajectoire nominale*  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$ , et on note

$$(10) \quad \bar{f}(t) := f(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t)), \quad D\bar{f}(t) := f'(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t))$$

la *dynamique nominale* et sa dérivée, et de même pour les dérivées de tous ordres et les autres fonctions de  $(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t))$ , les dérivées partielles étant notées  $\bar{f}_u(t)$  par exemple. La *dérivée directionnelle* de  $\mathbf{y}[\mathbf{u}, \mathbf{y}^0]$  au point  $(\bar{\mathbf{u}}, \bar{\mathbf{y}}^0)$  dans la direction  $(\mathbf{v}, \mathbf{z}^0) \in \mathcal{U} \times \mathbb{R}^n$ , notée

$$(11) \quad \mathbf{z}[\mathbf{v}, \mathbf{z}^0] := \lim_{s \rightarrow 0} (\mathbf{y}[\bar{\mathbf{u}} + s\mathbf{v}, \bar{\mathbf{y}}^0 + s\mathbf{z}^0] - \mathbf{y}[\bar{\mathbf{u}}, \bar{\mathbf{y}}^0])/s$$

est solution unique de l'équation d'état linéarisée

$$(12) \quad \text{(i) } \dot{\mathbf{z}}(t) = D\bar{f}(t)(\bar{\mathbf{v}}(t), \mathbf{z}(t)), \quad t \in [0, T]; \quad \text{(ii) } \mathbf{z}(0) = \mathbf{z}^0.$$

Si  $\mathbf{z}^0 = 0$ , on notera  $\mathbf{z}[\mathbf{v}] := \mathbf{z}[\mathbf{v}, 0]$ . On associe au système commandé (8) le *critère*

$$(13) \quad J(\mathbf{u}, \mathbf{y}) := \int_0^T \ell(\mathbf{u}(t), \mathbf{y}(t)) dt + \varphi(\mathbf{y}(T)),$$

somme d'un *coût intégral*, d'intégrande  $\ell$ , et d'un *coût final*  $\varphi$ ; on suppose  $\ell$  et  $\varphi$  de classe  $C^\infty$ . Alors  $J : \mathcal{U} \times \mathcal{Y} \rightarrow \mathbb{R}$  est aussi de classe  $C^\infty$ , de même que le *critère réduit*

$$(14) \quad F(\mathbf{u}, \mathbf{y}^0) := J(\mathbf{u}, \mathbf{y}[\mathbf{u}, \mathbf{y}^0]).$$

De plus, si  $(\mathbf{v}, \mathbf{z}^0) \in \mathcal{U} \times \mathbb{R}^n$ , notant  $\mathbf{z} = \mathbf{z}[\mathbf{v}, \mathbf{z}^0]$ , la dérivée du critère réduit a pour expression

$$(15) \quad F'(\bar{\mathbf{u}}, \bar{\mathbf{y}}^0)(\mathbf{v}, \mathbf{z}^0) = \int_0^T \bar{\ell}'(t)(\bar{\mathbf{v}}(t), \mathbf{z}(t)) dt + \varphi'(\bar{\mathbf{y}}(T))\mathbf{z}(T).$$

**2.2. Différents types de contraintes.** Les *contraintes mixtes* (*commande et état*) sont du type

$$(16) \quad g(\mathbf{u}(t), \mathbf{y}(t)) \in K_g, \quad \text{p.p. } t \in [0, T],$$

avec  $g : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_g}$  et  $K_g$  partie convexe fermée de  $\mathbb{R}^{n_g}$ . Si  $g$  ne dépend que de la première (resp. seconde) variable on parle de *contraintes sur la commande* (resp. de *contraintes (pures) sur l'état*). On peut aussi considérer des contraintes *abstraites* sur la commande du type

$$(17) \quad \mathbf{u}(t) \in U_{ad}, \quad \text{p.p. } t \in [0, T],$$

où  $U_{ad}$  est un fermé de  $\mathbb{R}^m$ , traduisant par exemple l'intervalle de variation admissible de chaque composante : alors  $U_{ad}$  est de la forme  $\prod_{i=1}^m [a_i, b_i]$ . Un autre cas fréquent est celui où  $U_{ad}$  est après normalisation égal à la boule euclidienne fermée  $\bar{B} := \{u \in \mathbb{R}^3; \sum_{i=1}^3 u_i^2 \leq 1\}$ .

Les *contraintes sur l'état final*, étant donné  $\Phi_F : \mathbb{R}^n \rightarrow \mathbb{R}^{n_\Phi}$ , sont de la forme

$$(18) \quad \Phi_F(\mathbf{y}(T)) \in K_F,$$

avec  $K_F$  partie convexe fermée de  $\mathbb{R}^{n_\Phi}$ . Soit une *subdivision* de  $[0, T]$  :

$$(19) \quad 0 = t_0 < t_1 < \dots < t_{n_{pi}} < t_{n_{pi}+1} = T.$$

On considère les *contraintes de points intérieurs*

$$(20) \quad \Phi^i(\mathbf{y}(t_i)) \in K_i, \quad i = 1, \dots, n_{pi},$$

avec  $K_i$  partie convexe fermée de  $\mathbb{R}^{n_i}$ . Elles peuvent traduire une contrainte d'interpolation, le passage d'un véhicule par des points imposés, etc. Dans certains cas les  $t_i$  sont des variables de décision.

**2.3. Un format général.** À une trajectoire  $(\mathbf{u}, \mathbf{y})$  on peut associer un critère incluant un *coût aux deux bouts* et des *contraintes aux deux bouts* :

$$(21) \quad J^{IF}(\mathbf{u}, \mathbf{y}) := \int_0^T \ell(\mathbf{u}(t), \mathbf{y}(t)) dt + \varphi(\mathbf{y}(0), \mathbf{y}(T)),$$

$$(22) \quad \Phi(\mathbf{y}(0), \mathbf{y}(T)) \in K_\Phi,$$

avec maintenant  $\varphi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\Phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_\Phi}$ , de classe  $C^\infty$ , et  $K_\Phi$  partie convexe fermée de  $\mathbb{R}^{n_\Phi}$ . Considérons le problème :

$$(23) \quad \text{Min}_{(\mathbf{u}, \mathbf{y}) \in \mathcal{U} \times \mathcal{Y} \times \mathbb{R}^n} J^{IF}(\mathbf{u}, \mathbf{y}) \quad \text{soumis à (8), (16), (17), (22)}.$$

Le problème réduit correspondant est obtenu en exprimant l'état comme fonction de la commande et de l'état initial, soit  $\mathbf{y} = \mathbf{y}[\mathbf{u}, y^0]$ . Montrons comment ramener au format (23) plusieurs autres types de problèmes.

- (a) *Paramètres de décision.* Soit le problème de minimisation par rapport à  $(\mathbf{u}, \mathbf{y}, \pi)$ , où  $\pi \in \mathbb{R}^{n_\pi}$ , l'état initial  $y^0$  étant donné, du critère

$$(24) \quad J(u, y, \pi) := \int_0^T \ell(\mathbf{u}(t), \mathbf{y}(t), \pi) dt + \varphi(\mathbf{y}(T), \pi),$$

avec ici  $\ell : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n_\pi} \rightarrow \mathbb{R}$  et  $\varphi : \mathbb{R}^n \times \mathbb{R}^{n_\pi} \rightarrow \mathbb{R}$ , sous la contrainte

$$(25) \quad \dot{\mathbf{y}}(t) = f(\mathbf{u}(t), \mathbf{y}(t), \pi), \quad \text{p.p.}, \quad \mathbf{y}(0) = y^0.$$

On se ramène au format (23) en interprétant  $\pi$  comme un état  $\boldsymbol{\pi}$  de dérivée nulle. On a ainsi un état augmenté  $\mathbf{y}^a := (\mathbf{y}^\dagger, \boldsymbol{\pi}^\dagger)$ . L'équation d'état et les conditions aux deux bouts s'écrivent

$$(26) \quad \dot{\mathbf{y}}^a(t) = \begin{pmatrix} f(\mathbf{u}(t), \mathbf{y}(t), \boldsymbol{\pi}(t)) \\ 0 \end{pmatrix} \quad \text{p.p.}; \quad \mathbf{y}(0) - y^0 = 0.$$

Le critère associé est

$$(27) \quad J(\mathbf{u}, \mathbf{y}, \boldsymbol{\pi}) := \int_0^T \ell(\mathbf{u}(t), \mathbf{y}(t), \boldsymbol{\pi}(t)) dt + \varphi(\mathbf{y}(T), \boldsymbol{\pi}(T)).$$

- (b) *Horizon variable.* Si  $T$  est une variable de décision, notons  $\tau \in [0, 1]$ , le *temps normalisé*,  $\mathbf{y}^N(\tau) := \mathbf{y}(\tau T)$  l'état en temps normalisé, et de même pour les autres fonctions du temps. Alors l'équation d'état (8) se traduit par

$$(28) \quad \dot{\mathbf{y}}^N(\tau) = T f(\mathbf{u}^N(\tau), \mathbf{y}^N(\tau)), \quad \tau \in [0, 1]; \quad \mathbf{y}^N(0) = y^0,$$

et le critère (21) peut s'exprimer comme

$$(29) \quad J^N(\mathbf{u}^N, \mathbf{y}^N) := T \int_0^1 \ell(\mathbf{u}^N(\tau), \mathbf{y}^N(\tau)) d\tau + \varphi(y^0, \mathbf{y}^N(1)).$$

Dans cette écriture en temps normalisé, l'horizon apparaît comme un paramètre de décision. On a vu au point (a) comment se ramener au format (23) en augmentant la taille de l'état.

- (c) *Points intérieurs* (voir (20)). Nous ne donnons que le principe. Reprenant l'analyse qui précède, on peut ramener chaque intervalle  $]t_i, t_{i+1}[$  à une durée 1, en introduisant comme paramètre la durée réelle  $T_i := t_{i+1} - t_i$ . On intègre alors en temps fictif  $\tau \in [0, 1]$ , en parallèle, les états dans chaque intervalle, sous contrainte de recollement : la valeur finale de l'état sur un intervalle doit coïncider avec la valeur initiale de l'état sur l'intervalle suivant.

### 3. CONTRAINTES SUR LA COMMANDE

**3.1. Dérivée du critère.** Définissons le *préhamiltonien*  $H : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  par

$$(30) \quad H(u, y, p) := \ell(u, y) + p \cdot f(u, y).$$

On adopte la notation  $\bar{H}(t)$  dans l'esprit de (10). À la trajectoire  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  on associe l'*état adjoint*  $\bar{\mathbf{p}} \in \mathcal{Y}$ , solution unique de l'équation *rétrograde* dite *équation adjointe*

$$(31) \quad \begin{cases} \text{(i)} & -\dot{\bar{\mathbf{p}}}(t) = \nabla_y \bar{H}(t), \quad \text{p.p. } t \in [0, T]; \\ \text{(ii)} & \bar{\mathbf{p}}(T) = \nabla \varphi(\bar{\mathbf{y}}(T)). \end{cases}$$

Noter que  $\nabla_y \bar{H}(t) = \nabla_y \bar{\ell}(t) + \bar{f}_y(t)^\dagger \bar{\mathbf{p}}(t)$ , et donc

$$(32) \quad \nabla_y \bar{H}(t) \cdot \mathbf{z}(t) = \nabla_y \bar{\ell}(t) \cdot \mathbf{z}(t) + \bar{\mathbf{p}}(t)^\dagger \bar{f}_y(t) \mathbf{z}(t).$$

**Lemme 3.1.** *Le critère réduit  $F$ , défini en (14), a en  $(\bar{\mathbf{u}}, \bar{y}^0)$  une dérivée caractérisée par*

$$(33) \quad F'(\bar{\mathbf{u}}, \bar{y}^0)(\mathbf{v}, z^0) = \bar{\mathbf{p}}(0) \cdot z^0 + \int_0^T \nabla_u \bar{H}(t) \cdot \mathbf{v}(t) dt.$$

**Démonstration.** Il vient, utilisant (12) et (31)-(32) :

$$\begin{aligned} \nabla \varphi(\bar{\mathbf{y}}(T)) \cdot \mathbf{z}(T) &= \bar{\mathbf{p}}(T) \cdot \mathbf{z}(T) \\ &= \bar{\mathbf{p}}(0) \cdot \mathbf{z}(0) + \int_0^T \frac{d}{dt} (\bar{\mathbf{p}}(t) \cdot \mathbf{z}(t)) dt \\ &= \bar{\mathbf{p}}(0) \cdot \mathbf{z}(0) + \int_0^T (\dot{\bar{\mathbf{p}}}(t) \cdot \mathbf{z}(t) + \bar{\mathbf{p}}(t) \cdot \dot{\mathbf{z}}(t)) dt \\ &= \bar{\mathbf{p}}(0) \cdot \mathbf{z}(0) + \int_0^T (\bar{\mathbf{p}}(t)^\dagger \bar{f}_y(t) \mathbf{v}(t) - \nabla_y \bar{\ell}(t) \cdot \mathbf{z}(t)) dt, \end{aligned}$$

d'où le résultat, utilisant (15) et l'expression de  $\nabla_u \bar{H}(t)$ , similaire à (32). ■

On dira que  $(\bar{\mathbf{u}}, \bar{y}^0)$  est un *point de minimum local* de  $F$  dans  $\mathcal{U} \times \mathbb{R}^n$  si

$$(34) \quad F(\bar{\mathbf{u}}, \bar{y}^0) \leq F(\mathbf{u}, y^0), \quad \text{si } \|\bar{\mathbf{u}} - \mathbf{u}\| + |\bar{y}^0 - y^0| \text{ est assez petit.}$$

On définira de même les minima locaux de fonctions quelconques, en notant bien que la définition dépend de la norme choisie. Dans le cas de l'espace  $\mathcal{U} \times \mathbb{R}^n$  on parlera de manière équivalente de *minimum faible*. Comme la dérivée d'une fonction est (si elle est définie) nulle en un point de minimum local, on déduit du lemme 3.1 que :

**Théorème 3.2.** (i) *Si  $\bar{\mathbf{u}}$  est un minimum faible de  $F(\cdot, y^0)$ , alors*

$$(35) \quad H_u(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) = 0, \quad \text{p.p. sur } [0, T].$$

(ii) *Si  $\bar{y}^0$  est un minimum local de  $F(\bar{\mathbf{u}}, \cdot)$ , alors  $\bar{\mathbf{p}}(0) = 0$ .*

**Exemple du régulateur quadratique.** On se donne la dynamique  $f(u, y) := Ay + Bu$ , l'intégrande du critère  $\ell(u, y) := \frac{1}{2}(u^\dagger Ru + y^\dagger Qy)$ , et le coût final  $\varphi(y) := \frac{1}{2}y^\dagger Q_T y$ , avec état initial  $y^0$  fixé comme seule contrainte. Les matrices  $R, Q, Q_T$  sont symétriques. L'équation adjointe a pour expression

$$(36) \quad \begin{cases} -\dot{\bar{\mathbf{p}}}(t) &= A^\dagger \bar{\mathbf{p}}(t) + Q \bar{\mathbf{y}}(t), \quad \text{p.p. sur } [0, T], \\ \bar{\mathbf{p}}(T) &= Q_T \bar{\mathbf{y}}(T), \end{cases}$$

et la condition (35) d'optimalité locale de  $\bar{\mathbf{u}}$  se traduit par

$$(37) \quad R \bar{\mathbf{u}}(t) + B^\dagger \bar{\mathbf{p}}(t) = 0, \quad \text{p.p. sur } [0, T].$$

Si  $R$  est inversible, éliminant la commande par l'équation précédente, on vérifie que  $(\bar{\mathbf{y}}, \bar{\mathbf{p}})$  est solution du *système aux deux bouts*

$$(38) \quad \begin{cases} \dot{\bar{\mathbf{y}}}(t) &= A \bar{\mathbf{y}}(t) - B R^{-1} B^\dagger \bar{\mathbf{p}}(t), \quad \text{p.p. sur } [0, T], \\ -\dot{\bar{\mathbf{p}}}(t) &= A^\dagger \bar{\mathbf{p}}(t) + Q \bar{\mathbf{y}}(t), \quad \text{p.p. sur } [0, T], \\ \bar{\mathbf{y}}(0) &= y^0; \quad \bar{\mathbf{p}}(T) = Q_T \bar{\mathbf{y}}(T). \end{cases}$$

**3.2. Conditions du premier ordre avec contraintes sur la commande.** On suppose la condition initiale fixée, et on ne considère que des contraintes sur la commande du type (17). Le problème s'écrit donc (supprimant l'état initial comme argument du critère réduit)

$$(39) \quad \underset{\mathbf{u}}{\text{Min}} F(\mathbf{u}) \quad \text{soumis à (17).}$$

Supposons que l'ensemble définissant la contrainte sur la commande vérifie

$$(40) \quad U_{ad} \text{ est un ensemble convexe fermé.}$$

On vérifie alors facilement que l'ensemble des commandes admissibles (c'est à dire vérifiant les contraintes du problème), soit

$$\{\mathbf{u} \in \mathcal{U}; \mathbf{u}(t) \in U_{ad} \text{ p.p.}\}$$

est une partie convexe et fermée de  $\mathcal{U}$ .

**Proposition 3.3.** *Soit  $\bar{\mathbf{u}}$  un minimum faible. Alors*

$$(41) \quad \bar{H}_u(t)(u - \bar{\mathbf{u}}(t)) \geq 0, \quad \text{pour tout } u \in U_{ad}, \text{ p.p. } t \in [0, T].$$

**Démonstration.** (a) Soit  $\{u^k\}$  une suite dense dans  $U_{ad}$ . Vérifions d'abord que (41) équivaut à

$$(42) \quad \bar{H}_u(t)(u^k - \bar{\mathbf{u}}(t)) \geq 0, \quad \text{pour tout } k \in \mathbb{N}, \text{ p.p. } t \in [0, T].$$

Évidemment (41) implique (42). Réciproquement, posons  $I_k := \{t \in [0, T]; \bar{H}_u(t)(u^k - \bar{\mathbf{u}}(t)) < 0\}$ . Si (42) est satisfait, l'ensemble  $I := \cup_k I_k$  est de mesure nulle car union dénombrable de tels ensembles, et sur  $[0, T] \setminus I$ , on a  $\bar{H}_u(t)(u^k - \bar{\mathbf{u}}(t)) \geq 0$  pour tout  $k$ , donc  $\bar{H}_u(t)(u - \bar{\mathbf{u}}(t)) \geq 0$  pour tout  $u \in U_{ad}$  par densité des  $u^k$ , ce qui établit (41).

(b) Soient  $\mathbf{u}$  admissible et  $s \in ]0, 1[$ . Alors  $\mathbf{u}^s := \bar{\mathbf{u}} + s(\mathbf{u} - \bar{\mathbf{u}})$  est admissible car combinaison convexe de  $\bar{\mathbf{u}}$  et  $\mathbf{u}$ , donc pour  $s$  assez petit,  $F(\bar{\mathbf{u}}) \leq F(\mathbf{u}^s)$ . Il vient, utilisant le lemme 3.1 :

$$(43) \quad 0 \leq \lim_{s \downarrow 0} \frac{F(\mathbf{u}^s) - F(\bar{\mathbf{u}})}{s} = F'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) = \int_0^T \bar{H}_u(t)(\mathbf{u}(t) - \bar{\mathbf{u}}(t)) dt.$$

(c) Si (42) n'est pas satisfait, il existe une partie mesurable  $E$  de  $[0, T]$  de mesure positive et  $k \in \mathbb{N}$  tels que  $\bar{H}_u(t)(u^k - \bar{\mathbf{u}}(t)) < 0$  p.p. sur  $E$ . Définissons  $\mathbf{u}$  par  $\mathbf{u}(t) = u^k$  si  $t \in E$ , et  $\mathbf{u}(t) = \bar{\mathbf{u}}$  sinon. Alors

$$(44) \quad F'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) = \int_E \bar{H}_u(t)(u^k - \bar{\mathbf{u}}(t)) dt < 0$$

ce qui contredit (43), d'où la conclusion. ■

*Remarque 3.4.* (i) Si  $U_{ad} = \mathbb{R}^m$ , (41) se réduit à (35).

(ii) La relation (41) équivaut à, et est souvent exprimé sous la forme

$$(45) \quad -\nabla_u \bar{H}(t) \in N_{U_{ad}}(\bar{\mathbf{u}}(t)), \quad \text{p.p. } t \in [0, T].$$

On peut aussi dire que  $\bar{\mathbf{u}}(t)$  est p.p. solution du problème de minimisation de  $\bar{H}_u(t)u$  pour  $u \in U_{ad}$ .

Explicitons (41) sur quelques exemples.

*Exemple 3.5.* Contraintes de borne :  $U_{ad} = \prod_{i=1}^m [a_i, b_i]$  avec  $b_i > a_i$ . On a p.p. en  $t$ , pour  $i = 1$  à  $m$  :  $\bar{H}_{u_i}(t) \geq 0$  si  $\bar{\mathbf{u}}_i(t) = a_i$ ,  $\bar{H}_{u_i}(t) \leq 0$  si  $\bar{\mathbf{u}}_i(t) = b_i$ ,  $\bar{H}_{u_i}(t) = 0$  si  $\bar{\mathbf{u}}_i(t) \in ]a_i, b_i[$ .

*Exemple 3.6.* Contraintes sur la norme euclidienne :  $U_{ad} = \bar{B}$  (boule euclidienne fermée). On a p.p. en  $t$ , si  $\bar{H}_u(t) \neq 0$ , alors  $\bar{\mathbf{u}}(t) = -\bar{H}_u(t)/|\bar{H}_u(t)|$ .

**3.3. Principe du minimum de Pontriaguine (PMP).** On reprend le cadre du problème (39), avec maintenant  $U_{ad}$  fermé non vide, convexe ou non. On dit que la trajectoire  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est une *extrémale de Pontriaguine* si l'état adjoint associé  $\bar{\mathbf{p}}$  satisfait l'inégalité hamiltonienne

$$(46) \quad H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \leq H(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \quad \text{pour tout } u \in U_{ad}, \text{ p.p. sur } [0, T],$$

ou de manière équivalente

$$(47) \quad H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) = \inf_{u \in U_{ad}} H(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \quad \text{p.p. sur } [0, T].$$

Si la relation précédente est satisfaite on dit aussi que  $(\bar{\mathbf{u}}, \bar{\mathbf{y}}, \bar{\mathbf{p}})$  (ou  $(\bar{\mathbf{y}}, \bar{\mathbf{p}})$  s'il n'y a pas d'ambiguïté) est une *biextrémale de Pontriaguine*. Notons que, si  $U_{ad}$  est convexe, une extrémale de Pontriaguine satisfait la condition du premier ordre (41). Le PMP s'énonce ainsi :



**Théorème 3.7.** Soit  $\bar{\mathbf{u}}$  solution de (39). Posons  $\bar{\mathbf{y}} := \mathbf{y}[\bar{\mathbf{u}}]$ . Alors  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est une extrémale de Pontriaguine.

La démonstration nécessite le lemme suivant. Dans la suite,  $M \geq \|\bar{\mathbf{u}}\|$ ; posons

$$(48) \quad \mathcal{U}^M := \{\mathbf{u} \in \mathcal{U}; \|\mathbf{u}\| \leq M\}.$$

**Lemme 3.8.** Il existe  $c_M > 0$  tel que, pour tout  $\mathbf{u} \in \mathcal{U}^M$ , il existe  $r(\mathbf{u}) \in \mathbb{R}$  satisfaisant :

$$(49) \quad F(\mathbf{u}) = F(\bar{\mathbf{u}}) + \int_0^T (H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) - \bar{H}(t))dt + r(\mathbf{u}), \quad \text{et} \quad |r(\mathbf{u})| \leq c_M \|\mathbf{u} - \bar{\mathbf{u}}\|_1^2.$$

**Démonstration.** Intégrant par parties et utilisant l'équation adjointe, il vient

$$(50) \quad F(\mathbf{u}) = F(\mathbf{u}) + \int_0^T \bar{\mathbf{p}}(t) \cdot (f(\mathbf{u}(t), \mathbf{y}(t)) - \dot{\mathbf{y}}(t))dt = \Delta_1(\mathbf{u}) + \Delta_2(\mathbf{u}),$$

avec  $\mathbf{y} = \mathbf{y}[u]$  et

$$\begin{aligned} \Delta_1(\mathbf{u}) &:= \varphi(\mathbf{y}(T)) - \varphi'(\bar{\mathbf{y}}(T))\mathbf{y}(T) + \bar{\mathbf{p}}(0) \cdot \mathbf{y}(0), \\ \Delta_2(\mathbf{u}) &:= \int_0^T (H(\mathbf{u}(t), \mathbf{y}(t), \bar{\mathbf{p}}(t)) - \bar{H}_y(t)\mathbf{y}(t))dt. \end{aligned}$$

En conséquence,  $F(\mathbf{u}) - F(\bar{\mathbf{u}}) = \Delta_1(\mathbf{u}) - \Delta_1(\bar{\mathbf{u}}) + \Delta_2(\mathbf{u}) - \Delta_2(\bar{\mathbf{u}})$ . Comme  $\mathbf{u}(t)$  et donc  $\mathbf{y}(t)$  restent dans un ensemble borné, les dérivées de tout ordre de la dynamique et des fonctions de coût sont lipschitziennes sur cet ensemble. De là, notant  $\delta\mathbf{y}(t) := \mathbf{y}(t) - \bar{\mathbf{y}}(t)$ , qui est tel que  $\delta\mathbf{y}(0) = 0$  :

$$(51) \quad \begin{aligned} \Delta_1(\mathbf{u}) - \Delta_1(\bar{\mathbf{u}}) &= \varphi(\mathbf{y}(T)) - \varphi(\bar{\mathbf{y}}(T)) - \varphi'(\bar{\mathbf{y}}(T))\delta\mathbf{y}(T) \\ &= O(|\delta\mathbf{y}(T)|)^2. \end{aligned}$$

Par ailleurs

$$\Delta_2(\mathbf{u}) - \Delta_2(\bar{\mathbf{u}}) = \int_0^T (H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) - \bar{H}(t))dt + \Delta_3,$$

où

$$\Delta_3 := \int_0^T (H(\mathbf{u}(t), \mathbf{y}(t), \bar{\mathbf{p}}(t)) - H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) - \bar{H}_y(t)\delta\mathbf{y}(t))dt.$$

Comme

$$(52) \quad H(\mathbf{u}(t), \mathbf{y}(t), \bar{\mathbf{p}}(t)) - H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) = \left( \int_0^1 H_y(\mathbf{u}(t), \bar{\mathbf{y}}(t) + s\delta\mathbf{y}(t), \bar{\mathbf{p}}(t))ds \right) \delta\mathbf{y}(t),$$

on a  $|\Delta_3| = O((\|\mathbf{u} - \bar{\mathbf{u}}\|_1 + \|\delta\mathbf{y}\|_\infty) \|\delta\mathbf{y}\|_\infty)$ . On conclut en notant que  $\|\delta\mathbf{y}\|_\infty = O(\|\mathbf{u} - \bar{\mathbf{u}}\|_1)$ . ■

**Démonstration du théorème 3.7.** (a) On montre d'abord que, si la trajectoire  $(\mathbf{u}, \mathbf{y})$  est admissible, alors

$$(53) \quad H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \leq H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \quad \text{p.p. sur } [0, T].$$

En effet, dans le cas contraire il existe  $\varepsilon > 0$  et une partie mesurable  $I$  de  $[0, T]$  de mesure strictement positive tels que

$$H(\mathbf{u}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \leq H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) - \varepsilon \quad \text{p.p. sur } I.$$

Soient  $M \geq \max(\|\mathbf{u}\|, \|\bar{\mathbf{u}}\|)$ ,  $I'$  une partie mesurable de  $I$ , et  $\mathbf{u}' \in \mathcal{U}$  tel que  $\mathbf{u}'(t) = \mathbf{u}(t)$  si  $t \in I'$ , et  $\mathbf{u}'(t) = \bar{\mathbf{u}}(t)$  sinon. Comme  $\|\mathbf{u}' - \bar{\mathbf{u}}\|_1 \leq 2M \text{mes}(I')$ , le lemme 3.8 implique que

$$(54) \quad F(\mathbf{u}') \leq F(\bar{\mathbf{u}}) - \varepsilon \text{mes}(I') + 4c_M M^2 \text{mes}(I')^2,$$

ce qui donne une contradiction en choisissant la mesure de  $I'$  strictement positive et assez petite.

(b) Soient  $u^k$  une suite dense dans  $U_{ad}$ , et  $\mathbf{u}^k$  la suite de  $\mathcal{U}$  définie par  $\mathbf{u}^0 := \bar{\mathbf{u}}$ , et

$$(55) \quad \mathbf{u}^k(t) = u^k \text{ si } H(u^k, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) < H(\mathbf{u}^{k-1}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)), \text{ et } \mathbf{u}^k(t) = \mathbf{u}^{k-1}(t) \text{ sinon.}$$

La densité de  $u^k$  implique que  $H(\mathbf{u}^k(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \rightarrow \inf_{u \in U_{ad}} H(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$  p.p. Prenant  $\mathbf{u} = \mathbf{u}^k$  dans (53), on conclut en appliquant le théorème de convergence monotone. ■

On vérifie facilement que si  $U_{ad} = \mathbb{R}^m$ , tout extrémale de Pontriaguine (en particulier quand  $\bar{\mathbf{u}}$  est solution de (39)) satisfait la condition de Legendre-Clebsch suivante :

$$(56) \quad \bar{H}_{uu}(t) \succeq 0 \quad \text{p.p. sur } [0, T].$$

Introduisons une nouvelle notion de minimum.

*Définition 3.9.* On dit qu'une suite  $\mathbf{u}^k$  de  $\mathcal{U}$  converge vers  $\bar{\mathbf{u}}$  au sens de Pontriaguine s'il existe  $M > 0$  tel que  $\|\mathbf{u}^k\| \leq M$ , et que  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  p.p. (par convergence dominée,  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  aussi dans  $L^s(0, T)^m$  pour tout  $s \in [1, \infty[$ ). On dit que  $\bar{\mathbf{u}}$  est un *minimum de Pontriaguine* si  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  au sens de Pontriaguine implique  $F(\bar{\mathbf{u}}) \leq F(\mathbf{u}^k)$  pour  $k$  assez grand.

*Remarque 3.10.* On vérifiera aisément que la démonstration du théorème 3.7. est encore valide quand  $\bar{\mathbf{u}}$  est un minimum de Pontriaguine. À tout minimum de Pontriaguine correspond donc une extrémale de Pontriaguine.

#### 4. PRINCIPE DE PONTRIAGUINE AVEC CONTRAINTES AUX DEUX BOUTS

**4.1. Lagrangien et état adjoint.** Considérons le problème avec contraintes sur la commande et aux deux bouts (le critère a été défini dans (21)) :

$$(57) \quad \text{Min}_{\mathbf{u}, \mathbf{y}} J^{IF}(\mathbf{u}, \mathbf{y}); \quad \text{soumis à (8), (17) et (22).}$$

Le *lagrangien* est défini comme une combinaison linéaire du critère et des contraintes par des coefficients appelés *multiplicateurs de Lagrange*. L'approche lagrangienne des conditions d'optimalité consiste à former et à exprimer la stationnarité du lagrangien par rapport aux variables à optimiser. Redéfinissons le *préhamiltonien*, sous la forme dite *non qualifiée*,  $H : \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$  par

$$(58) \quad H(\beta, u, y, p) := \beta \ell(u, y) + p \cdot f(u, y).$$

Notons  $\Psi \in \mathbb{R}^{n_\Psi}$  le multiplicateur associé aux contraintes aux deux bouts. On définit le *lagrangien aux deux bouts*

$$(59) \quad L^{IF}(\beta, y^0, y_T, \Psi) := \beta \varphi(y^0, y_T) + \Psi \cdot \Phi(y^0, y_T).$$

Le lagrangien du problème (57) est par définition

$$(60) \quad \begin{aligned} L(\beta, \mathbf{u}, \mathbf{y}, \mathbf{p}, \Psi) &:= \beta J^{IF}(\mathbf{u}, \mathbf{y}) + \int_0^T \mathbf{p}(t) \cdot (f(\mathbf{u}(t), \mathbf{y}(t)) - \dot{\mathbf{y}}(t)) dt + \Psi \cdot \Phi(\mathbf{y}(0), \mathbf{y}(T)) \\ &= \int_0^T (H(\beta, \mathbf{u}(t), \mathbf{y}(t), \mathbf{p}(t)) - \mathbf{p}(t) \cdot \dot{\mathbf{y}}(t)) dt + L^{IF}(\beta, \mathbf{y}(0), \mathbf{y}(T), \Psi). \end{aligned}$$

Cherchons les multiplicateurs  $(\bar{\beta}, \bar{\mathbf{p}}, \bar{\Psi})$  rendant le lagrangien stationnaire (c'est-à-dire de dérivée partielle nulle) par rapport à l'état, sur la trajectoire nominale. Ceci revient après intégration par parties (supposant  $\bar{\mathbf{p}}$  assez régulier pour cela) à vérifier, pour tout  $\mathbf{z} \in \mathcal{Y}$  :

$$(61) \quad \begin{aligned} \int_0^T (\nabla_{\mathbf{y}} H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) + \dot{\bar{\mathbf{p}}}(t)) \cdot \mathbf{z}(t) dt \\ + (\nabla_{y^0} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}) + \bar{\mathbf{p}}(0)) \cdot \mathbf{z}(0) \\ + (\nabla_{y_T} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}) - \bar{\mathbf{p}}(T)) \cdot \mathbf{z}(T) = 0. \end{aligned}$$

On vérifie que ceci équivaut à la nullité des coefficients ci-dessus de  $\mathbf{z}(t)$ ,  $\mathbf{z}(0)$  et  $\mathbf{z}(T)$ . Ceci conduit à définir l'*équation de l'état adjoint* comme

$$(62) \quad \begin{cases} \text{(i)} & -\dot{\bar{\mathbf{p}}}(t) = \nabla_{\mathbf{y}} H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)), \quad \text{p.p. } t \in [0, T], \\ \text{(ii)} & -\bar{\mathbf{p}}(0) = \nabla_{y^0} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}), \\ \text{(iii)} & \bar{\mathbf{p}}(T) = \nabla_{y_T} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}). \end{cases}$$

*Définition 4.1.* On appelle *multiplicateur de Pontriaguine*, associé à la trajectoire nominale  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$ , tout triplet  $\bar{\lambda} := (\bar{\beta}, \bar{\Psi}, \bar{\mathbf{p}})$  vérifiant (62),  $\bar{\beta} \geq 0$ ,  $\bar{\Psi} \in N_{K_\Phi}(\Phi(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T)))$ , la relation de non nullité

$$(63) \quad \bar{\beta} + |\bar{\Psi}| > 0,$$

ainsi que l'*inégalité hamiltonienne*, qui généralise (47) :

$$(64) \quad H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) = \inf_{u \in U_{ad}} H(\bar{\beta}, u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \quad \text{p.p. sur } [0, T].$$

On dit que  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est une *extrémale de Pontriaguine* s'il existe au moins un multiplicateur de Pontriaguine associé.

*Définition 4.2.* Si  $U_{ad}$  est convexe, on définit de manière similaire un *multiplicateur de Lagrange*, en changeant l'inégalité hamiltonienne (64) par la condition du premier ordre correspondante (comparer à (41)), soit

$$(65) \quad H_u(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))(u - \bar{\mathbf{u}}(t)) \geq 0, \quad \text{pour tout } u \in U_{ad}, \text{ p.p. sur } [0, T].$$

L'ensemble  $\Lambda_P(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  des multiplicateurs de Pontriaguine est donc un cône (il est stable par multiplication par un nombre strictement positif). Noter que la relation (63) empêche un multiplicateur d'être nul. Si  $\Lambda_P(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  n'est pas vide, il contient donc un élément avec, soit  $\bar{\beta} = 0$  (multiplicateur singulier), soit  $\bar{\beta} = 1$  (multiplicateur régulier).

**Théorème 4.3.** *Toute solution de (57) est une extrémale de Pontriaguine.*

On trouvera une démonstration courte de ce résultat dans [13] (avec état initial fixé, mais l'extension est facile), basée sur le "principe variationnel d'Ekeland".

*Exemple 4.4.* Le problème avec état initial fixé et sans contraintes sur l'état final correspond au cas où  $\varphi$  n'est fonction que de l'état final,  $\Phi(y^0, y_T) = y^0$ , et  $K_\Phi$  est réduit à un singleton  $\{\bar{y}^0\}$ . Dans l'équation adjointe, les conditions initiale et finale se réduisent resp. à  $\bar{\mathbf{p}}(T) = \bar{\beta}\nabla\varphi(\bar{\mathbf{y}}(T))$  et  $\bar{\mathbf{p}}(0) = -\bar{\Psi}$ ; cette dernière relation exprime  $\bar{\Psi}$  en fonction de  $\bar{\mathbf{p}}$ . Si  $\bar{\beta} = 0$ , la condition finale sur l'adjoint est nulle et la dynamique de l'état adjoint étant linéaire,  $\bar{\mathbf{p}}$  est nul donc aussi  $\bar{\Psi}$ , ce qui contredit (63). Il existe donc un multiplicateur avec  $\bar{\beta} = 1$ ; on retrouve ainsi la conclusion du théorème 3.7.

*Exemple 4.5.* Plus généralement, si l'état initial est fixé, on peut écrire  $\varphi$  comme fonction de  $y_T$  seul, et supposer les contraintes aux deux bouts de la forme  $(y^0, \Phi_F(\mathbf{y}(T))) \in \{\hat{y}^0\}_{\mathbb{R}^n} \times K_F$ , avec  $K_F$  convexe fermé. On vérifie facilement qu'on obtient un système équivalent au PMP en remplaçant les conditions aux deux bouts de l'équation adjointe par

$$(66) \quad \begin{cases} \bar{\mathbf{p}}(T) &= \bar{\beta}\nabla\varphi(\bar{\mathbf{y}}(T)) + D\Phi_F(\bar{\mathbf{y}}(T))^\dagger\bar{\Psi}_F; \\ \bar{\Psi}_F &\in N_{K_F}(\Phi_F(\bar{\mathbf{y}}(T))), \end{cases}$$

et la relation de non nullité (63) par

$$(67) \quad \bar{\beta} + |\bar{\Psi}_F| > 0.$$

*Définition 4.6.* On appelle *condition de qualification* toute hypothèse sur les contraintes, assurant la positivité stricte de  $\bar{\beta}$ .

Donnons-en un exemple simple (rappelons que  $\mathbf{z}[\mathbf{v}]$  est la solution de (12) avec  $z^0 = 0$ ).

*Définition 4.7.* On dira que la contrainte finale  $\Phi_F(\mathbf{y}(T)) \in K_F$  est *fortement qualifiée* sur la trajectoire nominale si

$$(68) \quad \text{l'application } \mathbf{v} \mapsto \Phi'_F(\bar{\mathbf{y}}(T))\mathbf{z}[\mathbf{v}](T) \text{ est surjective.}$$

En l'absence de contraintes sur la commande et avec un état initial fixé, cette hypothèse assure bien que  $\bar{\beta} \neq 0$ , car dans le cas contraire on aurait  $\bar{\Psi}_F \neq 0$  d'après (67), donc notant  $\mathbf{z} = \mathbf{z}[\mathbf{v}]$ , pour un choix convenable de  $\mathbf{v}$  :

$$(69) \quad 0 \neq \bar{\Psi}_F^\dagger\Phi'_F(\bar{\mathbf{y}}(T))\mathbf{z}[\mathbf{v}](T) = \bar{\mathbf{p}}(T) \cdot \mathbf{z}(T) = \int_0^T (\dot{\bar{\mathbf{p}}}(t) \cdot \mathbf{z}(t) + \bar{\mathbf{p}}(t) \cdot \dot{\mathbf{z}}(t))dt = \int_0^T \bar{H}_u(t)\mathbf{v}(t)dt.$$

Or le dernier terme est nul car  $\bar{H}_u(t) = 0$  p.p. en raison du PMP. Par le même type de calcul on vérifie que *la forte qualification entraîne l'unicité du multiplicateur* pour lequel  $\bar{\beta} = 1$ . Si  $\Phi_F$  est l'identité, la qualification forte s'interprète comme la *contrôlabilité de l'état final du système linéarisé*.

*Exemple 4.8.* Par un raisonnement analogue on vérifie plus généralement que, si les seules contraintes sont initiales-finales, et si

$$(70) \quad (\mathbf{v}, z^0) \mapsto \Phi'(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T))(z^0, \mathbf{z}[\mathbf{v}](T)) \text{ est surjective,}$$

on a nécessairement  $\bar{\beta} > 0$  et on peut donc se ramener à la forme qualifiée ( $\bar{\beta} = 1$ ), pour laquelle on a l'unicité du multiplicateur.

*Remarque 4.9.* Dans le cas où les contraintes finales modélisent un ensemble d'égalités et d'inégalités, comme l'analyse est locale, on peut éliminer de la formulation du problèmes les contraintes d'inégalités non actives sur la trajectoire de référence, ce qui conduit à une condition de qualification plus faible. Voir d'autres conditions de qualification dans [10, Ch. 2].

**4.2. Conservation du préhamiltonien.** Soient  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  une trajectoire et  $\bar{\mathbf{p}} \in \mathcal{Y}$  satisfaisant (62)(i). On dit que  $(\bar{\mathbf{u}}, \bar{\mathbf{y}}, \bar{\mathbf{p}})$  est une extrémale de Pontriaguine sur  $]0, T[$  si l'inégalité hamiltonienne (64) est vérifiée. Posons  $h(t) := H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$ .

**Lemme 4.10.** *Soit  $(\bar{\mathbf{u}}, \bar{\mathbf{y}}, \bar{\mathbf{p}})$  une extrémale de Pontriaguine sur  $]0, T[$ . Alors  $h(t)$  est (à un ensemble de mesure nulle près) constante sur  $]0, T[$ .*

**Démonstration.** a) Supposons d'abord  $\bar{\mathbf{u}}(t)$  dérivable en tout point, et l'absence de contraintes sur la commande. Alors  $H_u(t) = 0$  p.p. d'après le PMP, et  $h(t)$  a pour dérivée, utilisant les équation d'état et d'état adjoint :

$$\dot{h}(t) = H_y(t)\dot{\bar{\mathbf{y}}}(t) + H_p(t)\dot{\bar{\mathbf{p}}}(t) = -\dot{\bar{\mathbf{p}}}(t) \cdot \dot{\bar{\mathbf{y}}}(t) + \dot{\bar{\mathbf{y}}}(t) \cdot \dot{\bar{\mathbf{p}}}(t) = 0,$$

d'où le résultat.

b) Pour le cas général, adapter les arguments de [9, Partie II, Section 1.4]. ■

**4.3. Paramètres de décision.** Reprenons le problème de minimisation de (24) sous la contrainte (25), et sa reformulation (26)-(27). À l'état augmenté  $\mathbf{y}^a := (\mathbf{y}, \boldsymbol{\pi})$  associons l'état adjoint  $\mathbf{p}^a := (\mathbf{p}, \mathbf{p}_\pi)$ . On écrit le système d'optimalité sous forme qualifiée ( $\bar{\beta} = 1$ ). Posons  $H(u, y, \pi, p) := \ell(u, y, \pi) + p \cdot f(u, y, \pi)$ . Comme  $\bar{\boldsymbol{\pi}}(t) = \bar{\pi}$  pour tout  $t$ , l'équation adjointe (62) implique

$$(71) \quad \begin{cases} -\dot{\bar{\mathbf{p}}}(t) &= \nabla_y H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\pi}, \bar{\mathbf{p}}(t)), \quad \text{p.p. } t \in [0, T], \\ -\dot{\bar{\mathbf{p}}}_\pi(t) &= \nabla_\pi H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\pi}, \bar{\mathbf{p}}(t)), \quad \text{p.p. } t \in [0, T], \\ \bar{\mathbf{p}}_\pi(0) &= 0, \\ \bar{\mathbf{p}}(T) &= \nabla_{y^a} \varphi(\bar{\mathbf{y}}(T), \bar{\pi}), \\ \bar{\mathbf{p}}_\pi(T) &= \nabla_\pi \varphi(\bar{\mathbf{y}}(T), \bar{\pi}). \end{cases}$$

Comme  $\bar{\mathbf{p}}_\pi(0) = -\int_0^T \dot{\bar{\mathbf{p}}}_\pi(t) dt + \bar{\mathbf{p}}_\pi(T)$ , on en déduit que

$$0 = \bar{\mathbf{p}}_\pi(0) = \int_0^T \nabla_\pi H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\pi}, \bar{\mathbf{p}}(t)) dt + \nabla_\pi \varphi(\bar{\mathbf{y}}(T), \bar{\pi}),$$

qui n'est rien d'autre que la condition de stationnarité du lagrangien (du problème (24)-(25)) par rapport à  $\pi$ , et on a d'autre part l'inégalité hamiltonienne comme conséquence du PMP : p.p. en  $t$  on a

$$H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\pi}, \bar{\mathbf{p}}(t)) = \inf_{u \in U_{ad}} H(u, \bar{\mathbf{y}}(t), \bar{\pi}, \bar{\mathbf{p}}(t)).$$

**4.4. Horizon variable.** On reprend le problème (57), mais avec ici un horizon libre. Nous avons vu comment se ramener au format standard en passant en temps normalisé, voir (28)-(29), dans lequel l'horizon apparaît comme un paramètres de décision. D'après ce qui précède on sait que la condition d'optimalité est identique à celle du cas à horizon fixé (théorème 4.3), avec en plus la condition de stationnarité du lagrangien (du problème en temps normalisé) par rapport à l'horizon, ce qui équivaut à  $\int_0^T H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) dt = 0$ . Or le préhamiltonien est constant d'après le lemme 4.10 ; on retiendra donc que *le préhamiltonien est nul si l'horizon est libre*. Ceci ne vaut, bien entendu, que si l'horizon n'intervient pas dans les dynamique, coûts et contraintes du problème.

*Exemple 4.11.* Supposons que l'horizon intervienne dans la contrainte initiale-finale (rendez-vous avec un engin spatial) qui s'écrit alors  $\Phi(\mathbf{y}(0), \mathbf{y}(T), T) = 0$ . Notant par  $\bar{h}$  la valeur constante du préhamiltonien, on voit que la condition de stationnarité du lagrangien par rapport à  $T$  donne alors

$$\bar{h} + \bar{\Psi} \cdot D_T \Phi(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), T) = 0.$$

**4.5. Temps minimal, géodésiques.** C'est le cas particulier du problème précédent dans lequel le coût final est nul, et l'intégrande du coût intégral est  $\ell(u, y) = 1$ , avec état initial fixé et contrainte finale  $\Phi_F(\bar{\mathbf{y}}(T)) \in K_F$ . Le préhamiltonien  $\beta + p \cdot f(u, y)$  est nul le long d'une extrémale de Pontriaguine car l'horizon est libre, et la dynamique adjointe est (voir (66)-(67)) :

$$\begin{cases} -\dot{\bar{\mathbf{p}}}(t) &= \bar{f}_y(t)^\dagger \bar{\mathbf{p}}(t) \text{ p.p.}; \\ \bar{\mathbf{p}}(T) &= D\Phi_F(\bar{\mathbf{y}}(T))^\dagger \bar{\Psi}_F, \end{cases}$$

avec  $\bar{\Psi}_F \in N_{K_F}(\bar{\mathbf{y}}(T))$ , de sorte que  $\bar{\mathbf{p}}$  s'annule en un instant ssi il est identiquement nul. Le résultat essentiel est :

**Lemme 4.12.** *Si  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est solution du problème en temps minimal, alors  $\bar{\Psi}_F \neq 0$ , et si  $\Phi'_F(\bar{\mathbf{y}}(T))$  est surjective, alors  $\bar{\mathbf{p}}$  ne s'annule en aucun instant.*

**Démonstration.** On a  $\bar{\Psi}_F \neq 0$ , sinon  $\bar{\mathbf{p}}(t) = 0$  pour tout  $t$  d'après l'équation adjointe, et comme le temps final est libre, le préhamiltonien est nul donc  $\bar{\beta} = 0$ , ce qui contredit (67). Si  $\Phi'_F(\bar{\mathbf{y}}(T))$  est surjective, sa transposée est injective, donc  $\bar{\mathbf{p}}(T) = \Phi'_F(\bar{\mathbf{y}}(T))^\dagger \bar{\Psi}_F$  n'est pas nul, d'où le résultat. ■

L'inégalité hamiltonienne (64) se ramène à

$$\bar{\mathbf{p}}(t) \cdot f(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t)) = \inf_{u \in U_{ad}} \bar{\mathbf{p}}(t) \cdot f(u, \bar{\mathbf{y}}(t)),$$

le membre de gauche ci-dessus étant négatif, et nul ssi  $\bar{\beta} = 0$  (hamiltonien nul et  $\bar{\beta} \geq 0$ ).

*Remarque 4.13.* Une alternative à la formulation avec coût intégral ci-dessus est la formulation avec coût final  $T$  et coût intégral nul. Nous laissons le lecteur vérifier qu'elle aboutit aux mêmes conditions d'optimalité. Par préhamiltonien d'un problème de temps minimal, on entend souvent celui de cette seconde formulation, qui est  $p \cdot f(u, y)$ .

*Exemple 4.14.* Cas du système dynamique linéaire :  $f(u, y) = Ay + Bu$  avec  $A$  et  $B$  matrices de dimension convenable. Alors  $-\dot{\bar{\mathbf{p}}}(t) = A^\dagger \bar{\mathbf{p}}(t)$  donc  $\bar{\mathbf{p}}(t) = e^{(T-t)A^\dagger} \bar{\mathbf{p}}(T)$  est une fonction analytique. Notons  $\bar{\mathbf{p}}(t)^{(k)}$  la dérivée  $k$ ème de  $\bar{\mathbf{p}}(t)$ . La commande doit minimiser sur  $U_{ad}$  la fonction  $\bar{\mathbf{p}}(t)^\dagger Bu$ . Si celle-ci est constamment nulle, alors pour  $k = 0$  à  $n - 1$ , on a  $0 = B^\dagger \bar{\mathbf{p}}(t)^{(k)} = (-1)^k B^\dagger (A^\dagger)^k \bar{\mathbf{p}}(t)$  et donc  $\bar{\mathbf{p}}(t)$  est orthogonal à  $A^{(k)}B$ . Si le système est commandable, ce qui signifie que la famille des  $A^{(k)}B$ ,  $k = 0$  à  $n - 1$ , est de rang  $n$ ,  $\bar{\mathbf{p}}$  est identiquement nul : on sait que ceci est impossible si  $\Phi'_F(\bar{\mathbf{y}}(T))$  est surjective.

Si  $m = 1$  et  $U_{ad} = [-1, 1]$ , supposant le système commandable et  $\Phi'_F(\bar{\mathbf{y}}(T))$  surjective, utilisant le fait qu'une fonction analytique non constante a un nombre fini de changements de signes sur un intervalle borné, on en déduit que la commande est bang-bang (elle vaut  $\pm 1$  p.p.), plus précisément égale à  $-1$  quand  $B^\dagger \bar{\mathbf{p}}(t) > 0$ , et à  $1$  quand  $B^\dagger \bar{\mathbf{p}}(t) < 0$ , avec un nombre fini d'instantants de commutations.

*Exemple 4.15 (Optique géométrique).* La dynamique nominale est  $\dot{\bar{\mathbf{y}}}(t) = F(\bar{\mathbf{y}}(t))\bar{\mathbf{u}}(t)$ , avec  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^1$  à valeurs strictement positives (indice variable du milieu) et  $\bar{\mathbf{u}}(t) \in \bar{B}$  (boule unité euclidienne fermée) p.p. On suppose  $\Phi'_F(\bar{\mathbf{y}}(T))$  surjective, donc l'état adjoint ne s'annule pas. Le préhamiltonien  $\bar{\beta} + F(\bar{\mathbf{y}}(t))\bar{\mathbf{p}}(t) \cdot \bar{\mathbf{u}}(t)$  atteint son minimum sur  $\bar{B}$  en  $-\bar{\mathbf{p}}(t)/|\bar{\mathbf{p}}(t)|$ ; comme il est nul ceci impose  $\bar{\beta} > 0$  (on peut donc supposer que  $\bar{\beta} = 1$ ). Le couple état-état adjoint a donc pour dynamique, supprimant le temps en argument :

$$(72) \quad \dot{\bar{\mathbf{y}}} = -F(\bar{\mathbf{y}})\bar{\mathbf{p}}/|\bar{\mathbf{p}}|; \quad \dot{\bar{\mathbf{p}}} = |\bar{\mathbf{p}}|\nabla F(\bar{\mathbf{y}}).$$

Si sur une partie ouverte  $F$  est constant,  $\bar{\mathbf{p}}$  et  $\bar{\mathbf{u}}$  le sont aussi et la trajectoire est localement une ligne droite.

*Remarque 4.16.* Le problème de *transfert en énergie minimale* consiste à minimiser  $\frac{1}{2} \int_0^T \mathbf{u}(t)^2 dt$  avec les contraintes finales et dynamique précédentes. Le préhamiltonien est  $\frac{1}{2}\beta u^2 + F(y)p \cdot u$ . On peut vérifier comme précédemment que  $\bar{\beta} > 0$  et donc supposer  $\bar{\beta} = 1$ . La commande minimisant le hamiltonien est  $\bar{\mathbf{u}}(t) = -F(\bar{\mathbf{y}}(t))\bar{\mathbf{p}}(t)$ , et le couple état-état adjoint a pour dynamique

$$\dot{\bar{\mathbf{y}}} = -F(\bar{\mathbf{y}})^2 \bar{\mathbf{p}}; \quad \dot{\bar{\mathbf{p}}} = F(\bar{\mathbf{y}})|\bar{\mathbf{p}}|^2 \nabla F(\bar{\mathbf{y}}).$$

Comme cette dynamique est proportionnelle à celle de (72), les lieux des trajectoires associées de  $(\bar{\mathbf{y}}, \bar{\mathbf{p}})$  sont identiques, donc *les extrémales de Pontriaguine de l'optique géométrique coïncident avec celles du transfert en énergie minimale*. Par ailleurs, dans le problème de transfert en énergie minimale, comme le préhamiltonien est constant, on vérifie que  $|\bar{\mathbf{u}}(t)| = F(\bar{\mathbf{y}}(t))|\bar{\mathbf{p}}(t)|$  l'est aussi : *le transfert à énergie minimale se fait à vitesse constante*.

*Exemple 4.17 (Géométrie riemannienne).* Soit le problème de transfert en temps minimal avec la dynamique nominale  $\dot{\bar{\mathbf{y}}} = F(\bar{\mathbf{y}})\bar{\mathbf{u}}$ , où la matrice symétrique  $F(y)$ , de taille  $n$ , est inversible pour tout  $y$ , et fonction de classe  $C^\infty$  de  $y$ , et la contrainte  $\bar{\mathbf{u}}(t) \in \bar{B}$  p.p., ce qui revient à imposer  $\dot{\bar{\mathbf{y}}}^\dagger F(\bar{\mathbf{y}})^{-2} \dot{\bar{\mathbf{y}}} \leq 1$ . On peut encore se ramener à  $\bar{\beta} = 1$  et montrer que  $\bar{\mathbf{p}}$  ne peut s'annuler. Le préhamiltonien  $1 + p^\dagger F(y)u$  atteint sur une extrémale de Pontriaguine son minimum en  $-F(\bar{\mathbf{y}})\bar{\mathbf{p}}/|F(\bar{\mathbf{y}})\bar{\mathbf{p}}|$ , donc  $\dot{\bar{\mathbf{y}}} = -F^2(\bar{\mathbf{y}})\bar{\mathbf{p}}/|F(\bar{\mathbf{y}})\bar{\mathbf{p}}|$ . De plus, pour  $k = 1$  à  $n$ , notant  $F_k(y) := \partial F(y)/\partial y_k$ , on a  $\dot{\bar{\mathbf{p}}}_k = \bar{\mathbf{p}}^\dagger F_k(\bar{\mathbf{y}})F(\bar{\mathbf{y}})\bar{\mathbf{p}}/|F(\bar{\mathbf{y}})\bar{\mathbf{p}}|$ .

Le problème correspondant en énergie minimale consiste à minimiser  $\frac{1}{2} \int_0^T \mathbf{u}(t)^2 dt$  sous les dynamique et contrainte précédentes. Le préhamiltonien  $\frac{1}{2}u^2 + p^\dagger F(y)u$  atteint son minimum en  $u = -F(y)p$ , et donc  $\dot{\bar{\mathbf{y}}} = -F^2(\bar{\mathbf{y}})\bar{\mathbf{p}}$  et  $\dot{\bar{\mathbf{p}}}_k = \bar{\mathbf{p}}^\dagger F_k(\bar{\mathbf{y}})F(\bar{\mathbf{y}})\bar{\mathbf{p}}$ . Ici encore, cette dynamique est proportionnelle à celle du temps minimal : les lieux de  $(\bar{\mathbf{y}}, \bar{\mathbf{p}})$  pour ces deux problèmes coïncident donc. On appelle *géodésiques* les courbes donnant le lieu de l'état pour toute paire d'état initial et final possible. Le préhamiltonien du problème en énergie minimale étant constant, la vitesse  $|\bar{\mathbf{u}}(t)| = |F(\bar{\mathbf{y}}(t))\bar{\mathbf{p}}(t)|$  l'est aussi.

Nous reverrons des exemples de problème en temps minimal en section 8.1.

**5.1. Problèmes sans contraintes.** On suppose ici que  $\mathbf{u} \mapsto F(\mathbf{u})$  a un minimum faible  $\bar{\mathbf{u}}$ , fonction continue du temps, satisfaisant la *condition de Legendre-Clebsch forte* (uniforme en temps) suivante, où  $I_d$  est la matrice identité (comparer à la condition nécessaire (56)) : Il existe  $\alpha > 0$  tel que

$$(73) \quad \bar{H}_{uu}(t) \succeq \alpha I_d, \quad \text{p.p. sur } [0, T].$$

Appliquant à la relation  $H_u(u, y, p) = 0$  le théorème des fonctions implicites, on déduit que, pour  $\tau \in [0, T]$  et  $(u, y, p)$  voisin de  $(\bar{\mathbf{u}}(\tau), \bar{\mathbf{y}}(\tau), \bar{\mathbf{p}}(\tau))$ , il existe  $\Upsilon : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  de classe  $C^\infty$  tel que la relation  $H_u(u, y, p) = 0$  équivaut à  $u = \Upsilon(y, p)$ . Ajoutant si nécessaire le temps comme état supplémentaire, on peut “recoller” les fonctions  $\Upsilon$  pour différents instants et on obtient ainsi que pour un certain  $\varepsilon > 0$ , si  $|y - \bar{\mathbf{y}}(t)| + |p - \bar{\mathbf{p}}(t)| < \varepsilon$ , alors

$$H_u(u, y, p) = 0 \text{ ssi } u = \Upsilon(y, p).$$

On notera  $\Upsilon(t) := \Upsilon(\bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$ .

*Remarque 5.1.* Quand dans la suite on fera référence à (73), il faudra comprendre  $\bar{H}_{uu}(t)$  comme la dérivée seconde par rapport à la commande du préhamiltonien spécifique de chaque problème.

*Remarque 5.2.* Différenciant en temps la relation  $\nabla_u H(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) = 0$ , il vient

$$\bar{H}_{uu}(t)\dot{\Upsilon}(t) + \bar{H}_{uy}(t)\dot{\bar{\mathbf{y}}}(t) - \dot{\bar{\mathbf{f}}}_u(t)^\dagger \nabla_y \bar{H}(t) = 0,$$

ce qui permet d’exprimer  $\dot{\bar{\mathbf{u}}}(t) = \dot{\Upsilon}(t)$  en fonction de  $(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$ .

Éliminant la commande dans l’équation d’état et d’état adjoint on voit que  $(\bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$  est solution de l’équation différentielle autonome

$$(74) \quad \begin{cases} \dot{\bar{\mathbf{y}}}(t) &= f(\Upsilon(t), \bar{\mathbf{y}}(t)), \\ -\dot{\bar{\mathbf{p}}}(t) &= \nabla_y H(\Upsilon(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)), \end{cases} \quad \text{p.p. } t \in [0, T],$$

avec les conditions aux deux bouts

$$(75) \quad \bar{\mathbf{y}}(0) = \bar{\mathbf{y}}^0; \quad \bar{\mathbf{p}}(T) = \nabla \varphi(\bar{\mathbf{y}}(T)).$$

Dans le cas particulier d’un problème quadratique, on retrouve les relations (38). Introduisons le paramètre  $p^0 \in \mathbb{R}^n$ , et notons (si elle existe) par  $(\mathbf{y}[p^0], \mathbf{p}[p^0])$  la solution de (74) avec la condition initiale  $(\bar{\mathbf{y}}^0, p^0)$ . Alors le *système au deux bouts* (74)-(75) se ramène à trouver un zéro de la *fonction de tir*  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  définie par

$$(76) \quad \mathcal{T}(p^0) := \nabla \varphi(\mathbf{y}[p^0](T)) - \mathbf{p}[p^0](T).$$

Une telle équation est habituellement résolue par une variante de la méthode de Newton. Cette dernière nécessite l’évaluation du jacobien de  $\mathcal{T}$ . Donnons son expression quand  $p^0 = \bar{\mathbf{p}}(0)$  et donc  $(\mathbf{y}[p^0], \mathbf{p}[p^0]) = (\bar{\mathbf{y}}, \bar{\mathbf{p}})$ . On vérifie que

$$(77) \quad D\mathcal{T}(p^0)q^0 = D^2\varphi(\bar{\mathbf{y}}(T))\mathbf{z}(T) - \mathbf{q}(T),$$

où  $(\mathbf{z}(T), \mathbf{q}(T))$  est solution de la linéarisation de (74) avec la condition initiale  $(\mathbf{z}(0), \mathbf{q}(0)) = (0, q^0)$ . Explicitons les relations satisfaites par cette linéarisation. On a, pour un certain  $\mathbf{v} \in \mathcal{U}$  :

$$(78) \quad \begin{cases} \dot{\mathbf{z}}(t) &= \bar{f}'(t)(\mathbf{v}(t), \mathbf{z}(t)), \\ -\dot{\mathbf{q}}(t) &= \dot{\bar{f}}_y(t)^\dagger \mathbf{q}(t) + \bar{H}_{yu}(t)\mathbf{v}(t) + \bar{H}_{yy}(t)\mathbf{z}(t) \\ 0 &= \bar{H}_{uu}(t)\mathbf{v}(t) + \bar{H}_{uy}(t)\mathbf{z}(t) + \dot{\bar{f}}_u(t)^\dagger \mathbf{q}(t). \end{cases} \quad \text{p.p. } t \in [0, T],$$

**5.2. Problèmes avec contraintes aux deux bouts.** On se restreint au cas de la minimisation de  $J^{IF}$  avec contraintes d’égalités aux deux bouts, et donc  $K_\Phi = \{0\}_{\mathbb{R}^{n_\Phi}}$ , sous la condition de qualification (70) qui permet de fixer  $\bar{\beta}$  à 1. La fonction de tir a pour paramètres  $(y^0, p^0, \Psi)$ , les valeurs initiales de l’état et de l’état adjoint, ainsi que le multiplicateur associé aux contraintes. Elle exprime les conditions aux deux bouts pour l’état adjoint ainsi que les contraintes :

$$(79) \quad \mathcal{T}^{IF}(y^0, p^0, \Psi) := \begin{pmatrix} p^0 + \nabla_{y^0} L^{IF}(1, y^0, \mathbf{y}(T), \Psi) \\ -\mathbf{p}(T) + \nabla_{y_T} L^{IF}(1, y^0, \mathbf{y}(T), \Psi) \\ \Phi(y^0, \mathbf{y}(T)) \end{pmatrix},$$

où  $(\mathbf{y}, \mathbf{p})$  est le résultat de l’intégration de (74) avec la condition initiale  $(y^0, p^0)$ , et  $L^{IF}$  défini en (59). Noter qu’on a autant d’équations que d’inconnues.

On dira que  $(\bar{\mathbf{y}}^0, \bar{\mathbf{p}}(0), \bar{\Psi})$  est un *zéro régulier* de  $\mathcal{T}^{IF}$  si le jacobien de  $\mathcal{T}^{IF}$  en ce point est inversible. On sait que cela assure la convergence locale quadratique de la méthode de Newton (et, en raison du TFI,

la stabilité par rapport à des perturbations du zéro de la fonction de tir). Nous verrons en section 6.4 que cette propriété est étroitement liée aux conditions d'optimalité du second ordre, et permet d'assurer la convergence du problème discrétisé.

**5.3. Discrétisation du problème.** Discrétisons l'équation d'état (8) par la méthode d'Euler avec un pas de temps constant, supposant la commande constante sur chaque pas de temps, soit

$$(80) \quad \mathbf{y}^{k+1} = \mathbf{y}^k + hf(\mathbf{u}^k, \mathbf{y}^k), \quad k = 0, \dots, N-1; \quad \mathbf{y}^0 = y^0,$$

avec  $h = T/N$ ,  $N$  entier strictement positif,  $\mathbf{u}^k \in \mathbb{R}^m$  et  $\mathbf{y}^k \in \mathbb{R}^n$ . L'espace des commandes discrètes est  $\mathcal{U}^N := \mathbb{R}^{Nm}$ , celui des états discrets est  $\mathcal{Y}^N := \mathbb{R}^{(N+1)n}$ . À chaque  $\mathbf{u} \in \mathcal{U}^N$  est associé un unique état  $\mathbf{y}[\mathbf{u}, N] \in \mathcal{Y}^N$  solution de (80). La discrétisation correspondante du critère (13) est

$$(81) \quad J^N(\mathbf{u}, \mathbf{y}) := h \sum_{k=0}^{N-1} \ell(\mathbf{u}^k, \mathbf{y}^k) + \varphi(\mathbf{y}^N).$$

Le problème discret sans contrainte (avec  $y^0$  fixé) est donc

$$(82) \quad \text{Min } J^N(\mathbf{u}, \mathbf{y}); \quad \text{soumis à (80)}.$$

Écrivant le lagrangien sous la forme

$$(83) \quad \varphi(\mathbf{y}^N) + \sum_{k=0}^{N-1} (h \ell(\mathbf{u}^k, \mathbf{y}^k) + \mathbf{p}^{k+1} \cdot (\mathbf{y}^k + hf(\mathbf{u}^k, \mathbf{y}^k) - \mathbf{y}^{k+1})) + q \cdot (y^0 - \mathbf{y}^0),$$

on obtient l'équation adjointe discrétisée en rendant le lagrangien stationnaire par rapport à l'état, soit

$$(84) \quad \mathbf{p}^k = \mathbf{p}^{k+1} + h (\nabla_y \ell(\mathbf{u}^k, \mathbf{y}^k) + f_y(\mathbf{u}^k, \mathbf{y}^k)^\dagger \mathbf{p}^{k+1}), \quad k = 0, \dots, N-1; \quad \mathbf{p}^N = \nabla \varphi(\mathbf{y}^N).$$

Le critère réduit est  $F^N(\mathbf{u}) := J^N(\mathbf{u}, \mathbf{y}[\mathbf{u}, N])$ . Opérant de manière analogue au cas du temps continu, et notant  $\mathbf{y} = \mathbf{y}[\mathbf{u}, N]$ , on peut vérifier qu'il satisfait

$$(85) \quad DF^N(\mathbf{u})\mathbf{v} = \sum_{k=0}^{N-1} H_u(\mathbf{u}^k, \mathbf{y}^k, \mathbf{p}^{k+1})\mathbf{v}^k.$$

Une condition nécessaire d'optimalité locale est donc

$$(86) \quad H_u(\mathbf{u}^k, \mathbf{y}^k, \mathbf{p}^{k+1}) = 0, \quad k = 0, \dots, N-1.$$

Posons  $t_k = kh$ ,  $k = 0$  à  $N$ , et supposons la condition de Legendre-Clebsch forte (73) satisfaite. Si

$$|\bar{\mathbf{u}}(t_k) - \mathbf{u}^k| + |\bar{\mathbf{y}}(t_k) - \mathbf{y}^k|$$

est pour tout  $k$  assez petit, on peut (ajoutant si nécessaire le temps comme variable d'état) éliminer la commande comme fonction de l'état et de l'état adjoint :

$$(87) \quad \mathbf{u}^k = \Upsilon(\mathbf{y}^k, \mathbf{p}^{k+1})$$

ce qui permet, abrégeant  $\Upsilon(\mathbf{y}^k, \mathbf{p}^{k+1})$  en  $\Upsilon^k$  de se ramener au système dynamique implicite en  $(\mathbf{y}, \mathbf{p})$ , pour  $k = 0, \dots, N-1$  :

$$(88) \quad \begin{cases} \mathbf{y}^{k+1} &= \mathbf{y}^k + hf(\Upsilon^k, \mathbf{y}^k), \\ \mathbf{p}^k &= \mathbf{p}^{k+1} + h (\nabla_y \ell(\mathbf{u}^k, \mathbf{y}^k) + f_y(\Upsilon^k, \mathbf{y}^k)^\dagger \mathbf{p}^{k+1}), \end{cases}$$

avec les conditions aux deux bouts

$$(89) \quad \mathbf{y}^0 = y^0; \quad \mathbf{p}^N = \nabla \varphi(\mathbf{y}^N).$$

Prenant  $p^0 = \mathbf{p}^0$  comme paramètre et notant  $\mathbf{y}[p^0]$ ,  $\mathbf{p}[p^0]$  la solution de (88) avec la condition initiale  $(y^0, p^0)$ , on obtient la *fonction de tir discrète*

$$(90) \quad \mathcal{T}^N(p^0) := \nabla \varphi(\mathbf{y}^N[p^0]) - \mathbf{p}^N[p^0].$$

Quand  $h$  est assez petit, sous l'hypothèse (73),  $\mathcal{T}^N$  est bien définie et converge uniformément (ainsi que ses dérivées de tout ordre) sur un voisinage de  $\bar{\mathbf{p}}(0)$ . On peut alors montrer que  $\mathcal{T}^N$  a dans ce voisinage une solution unique à laquelle est associée une solution  $(\mathbf{u}_h, \mathbf{y}_h, \mathbf{p}_h)$  des conditions d'optimalité, et que l'ordre de convergence est 1 au sens où on a

$$(91) \quad \max_{0 \leq k \leq N-1} (|\bar{\mathbf{u}}(t_k) - \mathbf{u}_h^k| + |\bar{\mathbf{y}}(t_k) - \mathbf{y}_h^k| + |\bar{\mathbf{p}}(t_k) - \mathbf{p}_h^{k+1}|) + |\mathbf{y}(T) - \mathbf{y}_h^N| = O(h).$$

## 6. CONDITIONS D'OPTIMALITÉ DU SECOND ORDRE

**6.1. Développements du critère réduit.** Analysons un problème sans contrainte autre que l'état initial fixé. Soit  $\bar{\mathbf{u}} \in \mathcal{U}$ . Nous avons noté en section 2.1 que l'application  $\bar{\mathbf{u}} \mapsto F(\bar{\mathbf{u}})$  est de classe  $C^\infty$  sur  $\mathcal{U}$ . Notons par  $Q_0(\mathbf{v}) := F''(\bar{\mathbf{u}})(\mathbf{v}, \mathbf{v})$  la form quadratique associée au hessien (dérivée seconde) de  $F$ . Le développement de  $F$  au second ordre, en  $\mathbf{v} \in \mathcal{U}$ , s'écrit :

$$F(\bar{\mathbf{u}} + \mathbf{v}) = F(\bar{\mathbf{u}}) + F'(\bar{\mathbf{u}})\mathbf{v} + \frac{1}{2}Q_0(\mathbf{v}) + R_2(\mathbf{v}),$$

où le terme de reste

$$R_2(\mathbf{v}) = \int_0^1 (1 - \sigma)(F''(\bar{\mathbf{u}} + \sigma\mathbf{v}) - F''(\bar{\mathbf{u}}))(\mathbf{v}, \mathbf{v})d\sigma$$

vérifie  $|R_2(\mathbf{v})| = O(\|\mathbf{v}\|_\infty^3)$ . Une *condition nécessaire* d'optimalité locale dans  $\mathcal{U}$  est donc

$$F'(\bar{\mathbf{u}}) = 0 \text{ et } Q_0(\mathbf{v}) \geq 0, \text{ pour tout } \mathbf{v} \in \mathcal{U}.$$

Or, du lemme 6.4 ci-dessous on déduit aisément que  $|R_2(\mathbf{v})| = O(\|\mathbf{v}\|_\infty \|\mathbf{v}\|_2^2)$ , d'où la *condition suffisante* d'optimalité locale suivante :

**Théorème 6.1.** *Si  $F'(\bar{\mathbf{u}}) = 0$ , et si*

$$(92) \quad \text{Il existe } \alpha > 0 \text{ tel que } Q_0(\mathbf{v}) \geq 2\alpha\|\mathbf{v}\|_2^2 \text{ pour tout } \mathbf{v} \in \mathcal{U},$$

*alors  $\bar{\mathbf{u}}$  est un minimum faible de  $F$ , et vérifie la condition de croissance quadratique suivante : si  $\|\mathbf{v}\|_\infty$  est assez petit, alors  $F(\bar{\mathbf{u}} + \mathbf{v}) \geq F(\bar{\mathbf{u}}) + \frac{1}{2}\alpha\|\mathbf{v}\|_2^2$ .*

Soit  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  au sens de Pontriaguine (définition 3.9). Si  $E \subset [0, T]$ , est mesurable, on note  $\mathbf{1}_E$  sa fonction caractéristique (qui vaut 1 sur  $E$  et 0 en dehors), et  $|E|$  sa mesure. Posons  $\mathbf{v}^k := \mathbf{u}^k - \bar{\mathbf{u}}$ ,  $\varepsilon_k := \|\mathbf{v}^k\|_1^{1/2}$ ,  $B_k := \{t \in [0, T]; |\mathbf{v}^k(t)| \geq \varepsilon_k\}$ ,  $A_k := [0, T] \setminus B_k$ ,  $\mathbf{v}_A^k := \mathbf{1}_{A_k}\mathbf{v}^k$ . Alors  $\|\mathbf{v}_A^k\|_\infty \leq \varepsilon_k$ ,  $|B_k| \leq \int_0^T (|\mathbf{v}^k(t)|/\varepsilon_k)dt = \varepsilon_k$ , et donc  $|B_k| \rightarrow 0$ . On a le *principe de décomposition* suivant.

**Lemme 6.2.** *Soit  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  au sens de Pontriaguine. Alors*

$$F(\mathbf{u}^k) \geq F(\bar{\mathbf{u}}) + Q_0(\mathbf{v}_A^k) + \int_{B_k} (H(\mathbf{u}^k(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) - \bar{H}(t))dt + o(\|\mathbf{v}^k\|_2^2).$$

On dira que le préhamiltonien satisfait une hypothèse de *croissance quadratique* si, pour tout  $M > 0$ , il existe  $\alpha_M > 0$  tel que si  $|u| < M$ , alors

$$H(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) \geq \bar{H}(t) + 2\alpha_M|u - \bar{\mathbf{u}}(t)|^2.$$

On peut déduire du lemme précédent le résultat suivant :

**Théorème 6.3.** *Supposons (92) satisfaite ainsi que l'hypothèse de croissance quadratique du préhamiltonien. Soit  $\mathbf{u}^k \rightarrow \bar{\mathbf{u}}$  au sens de Pontriaguine. Alors il existe  $\alpha > 0$ , fonction de  $M := \sup_k \|\mathbf{u}^k\|_\infty$ , telle que pour  $k$  assez grand :*

$$F(\mathbf{u}^k) \geq F(\bar{\mathbf{u}}) + \alpha\|\mathbf{u}^k - \bar{\mathbf{u}}\|_2^2.$$

**6.2. Expression du hessien du critère réduit.** Considérons le problème de minimisation, sans contrainte autre que l'état initial fixé et l'équation d'état (8), du critère (13). Le lagrangien associé est, cf. (60) :

$$\begin{aligned} J(\mathbf{u}, \mathbf{y}) &+ \int_0^T \mathbf{p}(t) \cdot (f(\mathbf{u}(t), \mathbf{y}(t)) - \dot{\mathbf{y}}(t))dt \\ &= \int_0^T H(\mathbf{u}(t), \mathbf{y}(t), \mathbf{p}(t))dt + \varphi(\mathbf{y}(T)) - \int_0^T \mathbf{p}(t) \cdot \dot{\mathbf{y}}(t)dt. \end{aligned}$$

Notons par  $D^2\bar{H}(t)$  la dérivée seconde du hamiltonien par rapport à  $(\mathbf{u}, \mathbf{y})$ , au point  $(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$ . La dérivée seconde du lagrangien dans la direction  $(\mathbf{v}, \mathbf{z}) \in \mathcal{U} \times \mathcal{Y}$  est

$$\Omega_0(\mathbf{v}, \mathbf{z}) := \int_0^T D^2\bar{H}(t)(\mathbf{v}(t), \mathbf{z}(t))^2 dt + \varphi''(\bar{\mathbf{y}}(T))(\mathbf{z}(T))^2.$$

On peut vérifier que, notant par  $\mathbf{z}[\mathbf{v}]$  la solution de (12) avec  $z^0 = 0$  :

**Lemme 6.4.** *On a  $Q_0(\mathbf{v}) = \Omega_0(\mathbf{v}, \mathbf{z}[\mathbf{v}])$ , pour tout  $\mathbf{v} \in \mathcal{U}$ .*



**6.3. Extension au cas de contraintes aux deux bouts.** On considère le problème suivant, avec contraintes initiales-finales :

$$(93) \quad \min_{(\mathbf{u}, \mathbf{y}) \in \mathcal{U} \times \mathcal{Y}} J^{IF}(\mathbf{u}, \mathbf{y}); \text{ soumis à (8) et (22),}$$

où  $J^{IF}$  est définie en (21), avec  $K_\Phi = K[n'_\Phi, n''_\Phi]$  (nombre fini d'égalités et d'inégalités). Le lagrangien a pour expression (60). Sa dérivée seconde (hessien) en  $(\beta, \bar{\mathbf{u}}, \bar{\mathbf{y}}, \mathbf{p}, \Psi)$ , dans la direction  $(\mathbf{v}, \mathbf{z}) \in \mathcal{U} \times \mathcal{Y}$  est, le préhamiltonien étant sous forme non qualifiée (58), et posant  $\lambda := (\beta, \mathbf{p}, \Psi)$  :

$$\Omega_0[\lambda](\mathbf{v}, \mathbf{z}) := \int_0^T D^2 \bar{H}(t)(\mathbf{v}(t), \mathbf{z}(t))^2 dt + D^2 L^{IF}(\beta, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \Psi)(\mathbf{z}(0), \mathbf{z}(T))^2,$$

où par  $D^2 L^{IF}$  on entend les dérivées secondes par rapport à  $(y^0, y_T)$ . Posons

$$Q[\lambda](\mathbf{v}, z^0) := \Omega[\lambda](\mathbf{v}, \mathbf{z}[\mathbf{v}, z^0]).$$

Comme cette forme quadratique est continue dans la norme de  $\mathcal{U}_2 \times \mathbb{R}^n$ , on peut l'étendre à cet espace.

Les conditions du second ordre font intervenir le *cône critique*, dont les éléments sont appelés *directions critiques*, défini comme suit, où  $\mathbf{z} = \mathbf{z}[\mathbf{v}, z^0]$  (on a défini le cône tangent  $T_{K_\Phi}(\cdot)$  en section 1) :

$$C_2(\bar{\mathbf{u}}, \bar{\mathbf{y}}) := \left\{ \begin{array}{l} (\mathbf{v}, z^0) \in \mathcal{U}_2 \times \mathbb{R}^n; \varphi'(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T))(\mathbf{z}(0), \mathbf{z}(T)) \leq 0 \\ \Phi'(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T))(\mathbf{z}(0), \mathbf{z}(T)) \in T_{K_\Phi}(\Phi(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T))) \end{array} \right\}.$$

**Théorème 6.5.** *Si  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est solution faible (resp. de Pontriaguine) de (93), à toute direction critique  $(\mathbf{v}, z^0)$  est associée un multiplicateur de Lagrange (resp. de Pontriaguine) tel que  $Q[\lambda](\mathbf{v}, z^0) \geq 0$ .*

Donnons maintenant des conditions suffisantes. Rappelons la notation  $\Lambda_P(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  de l'ensemble des multiplicateurs de Pontriaguine. Nous utiliserons les hypothèses suivantes :

$$(94) \quad \left\{ \begin{array}{l} \text{Pour tout } M > \|\bar{\mathbf{u}}\|, \text{ il existe } \hat{\lambda}_M \in \Lambda_P(\bar{\mathbf{u}}, \bar{\mathbf{y}}), \hat{\lambda}_M = (\hat{\beta}, \hat{\mathbf{p}}, \hat{\Psi}), \text{ et } \alpha_M > 0 \text{ tels que} \\ H(\hat{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \hat{\mathbf{p}}(t)) + \alpha_M |u - \bar{\mathbf{u}}(t)|^2 \leq H(\hat{\beta}, u, \bar{\mathbf{y}}(t), \hat{\mathbf{p}}(t)), \text{ si } |u| \leq M, \text{ p.p. } t. \end{array} \right.$$

$$(95) \quad \left\{ \begin{array}{l} \text{Il existe } \alpha > 0 \text{ tel que pour tout } (\mathbf{v}, z^0) \in C_2(\bar{\mathbf{u}}, \bar{\mathbf{y}}), \\ \text{il existe } \lambda \in \Lambda_P(\bar{\mathbf{u}}, \bar{\mathbf{y}}) \text{ tel que } Q[\lambda](\mathbf{v}, z^0) \geq \alpha(\|\mathbf{v}\|_2^2 + |z^0|^2). \end{array} \right.$$

**Théorème 6.6.** *Soit  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  point admissible du problème (93), satisfaisant (94)-(95). Alors  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est un minimum de Pontriaguine.*

**6.4. Lien avec l'algorithme de tir.** On reprend le cadre de la section 5.2 (les seules contraintes sont des égalités aux deux bouts) avec  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  minimum faible du problème. La fonction de tir  $\mathcal{T}^{IF}$  est définie en (79). Nous avons vu en section 5.2 l'intérêt de vérifier si le jacobien de cette fonction est inversible à la solution. Ceci se ramène à vérifier si le noyau du jacobien est réduit à  $\{0\}$ . On sait que la condition de qualification (70) assure l'unicité du multiplicateur noté  $\bar{\lambda}$  (avec la normalisation  $\beta = 1$ ) et (95) équivaut alors pour un certain  $\alpha > 0$  à

$$(96) \quad Q[\bar{\lambda}](\mathbf{v}, z^0) \geq \alpha(\|\mathbf{v}\|_2^2 + |z^0|^2), \text{ pour tout } (\mathbf{v}, z^0) \in C_2(\bar{\mathbf{u}}, \bar{\mathbf{y}}).$$

On peut en déduire la validité de la condition de Legendre-Clebsch forte (73).

**Théorème 6.7.** *On suppose satisfaits la condition de qualification (70) et (96). Alors le jacobien de  $\mathcal{T}^{IF}$  en  $(\bar{y}^0, \bar{p}^0, \bar{\Psi})$  est inversible.*

**Démonstration.** Un élément  $(z^0, q^0, \delta\Psi)$  de  $\text{Ker } D\mathcal{T}^{IF}(\bar{y}^0, \bar{p}^0, \bar{\Psi})$  est, d'après (79), caractérisé par

$$\begin{aligned} q^0 + L_{y^0 y^0}^{IF}(\cdot) z^0 + L_{y^0 y_T}^{IF}(\cdot) \mathbf{z}(T) + \Phi_{y^0}(y^0, \bar{\mathbf{y}}(T))^\dagger \delta\Psi &= 0, \\ -\bar{\mathbf{q}}(T) + L_{y_T y^0}^{IF}(\cdot) z^0 + L_{y_T y_T}^{IF}(\cdot) \mathbf{z}(T) + \Phi_{y_T}(y^0, \bar{\mathbf{y}}(T))^\dagger \delta\Psi &= 0, \\ D\Phi(y^0, \bar{\mathbf{y}}(T))(z^0, \mathbf{z}(T)) &= 0, \end{aligned}$$

avec  $(\mathbf{v}, \mathbf{z}, \mathbf{q})$  solution de (78). On reconnaît les conditions d'optimalité du problème quadratique de minimisation de  $Q[\bar{\lambda}](\mathbf{v}, z^0)$  sous les condition d'équation d'état linéarisée (première condition de (78)). D'après la condition nécessaire du second ordre (théorème 6.5), ce problème a un critère convexe sur le domaine admissible qui est un sous espace vectoriel, et  $(\mathbf{v}, z^0)$  minimise donc  $Q[\bar{\lambda}](\cdot)$  sur le cône critique.

Si (96) est satisfait alors  $Q[\bar{\lambda}](\cdot)$  est strictement convexe sur le cône critique, et a sa condition d'optimalité satisfaite en zéro, donc  $(\mathbf{v}, z^0) = 0$ , et de là  $\mathbf{z} = 0$ . La condition de qualification permet alors de vérifier que  $(\mathbf{q}, \delta\Psi) = 0$ , d'où l'injectivité de la fonction de tir linéarisée, comme il fallait le montrer. ■

## 7. PROBLÈMES AVEC CONTRAINTES SUR L'ÉTAT

**7.1. Principe de Pontriaguine.** Prenons maintenant en compte la présence de contraintes sur l'état :

$$(97) \quad g_j(\mathbf{y}(t)) \leq 0, \quad j = 1, \dots, n_g, \quad \text{p.p. } t \in [0, T],$$

avec  $g$  de classe  $C^\infty$ . De telles contraintes sont presque toujours présentes et il est donc important de savoir les analyser. Considérons donc le problème suivant :

$$(98) \quad \underset{(\mathbf{u}, \mathbf{y}) \in \mathcal{U} \times \mathcal{Y}}{\text{Min}} \quad J^{IF}(\mathbf{u}, \mathbf{y}) \quad \text{soumis à (8), (17), (22), (97).}$$

On choisit  $C([0, T])^{n_g}$  comme espace de Banach pour la contrainte sur l'état. On sait que toute forme linéaire continue sur  $C([0, T])$  est du type  $\mathbf{h} \mapsto \int_0^T \mathbf{h}(t) d\boldsymbol{\mu}(t)$ , où  $d\boldsymbol{\mu} \in M(0, T)$ , l'espace des mesures de Borel sur  $[0, T]$ . On peut identifier l'action d'une telle mesure  $d\boldsymbol{\mu}$  à l'intégrale de Stieltjes d'une fonction à *variation bornée* nulle au temps  $T$ , notée  $\boldsymbol{\mu}$ , qu'on suppose nulle en  $T$ . Rappelons que la définition de la variation d'une fonction  $\boldsymbol{\mu}$  sur  $[0, T]$  est

$$(99) \quad \text{var}(\boldsymbol{\mu}) := \sup \left\{ \sum_{i=0}^{n_{pi}} |\boldsymbol{\mu}(t_{i+1}) - \boldsymbol{\mu}(t_i)|; \quad \text{où } (t_i) \text{ est une subdivision de } [0, T] \right\},$$

(sur la notion de subdivision, voir (19)), et que l'intégrale de Stieltjes de  $\mathbf{h} \in C([0, T])$  associé à la fonction à variation bornée  $\boldsymbol{\mu}$  est

$$(100) \quad \int_0^T \mathbf{h}(t) d\boldsymbol{\mu}(t) := \lim \sum_{i=0}^{n_{pi}} \mathbf{h}(\tau_i) (\boldsymbol{\mu}(t_{i+1}) - \boldsymbol{\mu}(t_i))$$

où  $\tau_i \in [t_i, t_{i+1}]$ , la limite étant prise sur les subdivisions d'incrément maximum tendant vers 0.

Le lagrangien du problème (98) est identique à celui du problème (57) (défini en (60)), à l'ajout près de la contribution  $\Delta^g : \mathcal{Y} \times M(0, T)^{n_g} \rightarrow \mathbb{R}$  de la contrainte sur l'état, soit

$$\Delta^g(\mathbf{y}, \boldsymbol{\mu}) := \sum_{j=1}^{n_g} \int_0^T g_j(\mathbf{y}(t)) d\boldsymbol{\mu}_j(t).$$

Notons la relation, où  $\mathbf{z} \in \mathcal{Y}$  :

$$\Delta_y^g(\mathbf{y}, \boldsymbol{\mu}) \mathbf{z} = \sum_{j=1}^{n_g} \int_0^T g'_j(\mathbf{y}(t)) \mathbf{z}(t) d\boldsymbol{\mu}_j(t).$$

On obtient encore l'équation adjointe en rendant le lagrangien stationnaire par rapport à l'état. On cherche maintenant l'état adjoint dans l'espace  $\mathcal{P} := BV(0, T)^n$  des fonctions à variation bornée, ce qui permet d'utiliser la formule d'*intégration par parties*

$$(101) \quad \int_0^T \mathbf{p}(t) \cdot \dot{\mathbf{y}}(t) dt = \mathbf{p}(T) \cdot \mathbf{y}(T) - \mathbf{p}(0) \cdot \mathbf{y}(0) - \sum_{i=1}^n \int_0^T \mathbf{y}_i(t) d\mathbf{p}_i(t).$$

L'équation de l'état adjoint a alors pour expression,  $L^{IF}(\cdot)$  étant défini en (59) :

$$(102) \quad \begin{cases} -d\bar{\mathbf{p}}(t) &= \nabla_y H(\bar{\beta}, \bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)) dt + \sum_{j=1}^{n_g} \nabla g_j(\bar{\mathbf{y}}(t)) d\bar{\boldsymbol{\mu}}_j(t), \quad \text{p.p. } t \in [0, T], \\ -\bar{\mathbf{p}}(0) &= \nabla_{y^0} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}), \\ \bar{\mathbf{p}}(T) &= \nabla_{y^T} L^{IF}(\bar{\beta}, \bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T), \bar{\Psi}). \end{cases}$$

On appelle multiplicateur de Pontriaguine, associé à la trajectoire nominale, tout  $\lambda = (\bar{\beta}, \bar{\Psi}, \bar{\boldsymbol{\mu}}, \bar{\mathbf{p}})$  vérifiant (102),  $\bar{\Psi} \in N_{K_\Phi}(\Phi(\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(T)))$ ,  $\bar{\beta} \geq 0$ , l'inégalité hamiltonienne (64), et les relations de complémentarité

$$(103) \quad d\bar{\boldsymbol{\mu}} \geq 0; \quad \int_0^T g'_j(\bar{\mathbf{y}}(t)) d\bar{\boldsymbol{\mu}}_j(t) = 0, \quad j = 1, \dots, n_g,$$

et de non nullité

$$(104) \quad \bar{\beta} + |\bar{\Psi}| + \|\bar{\boldsymbol{\mu}}\| > 0.$$

**Théorème 7.1.** *Si  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est solution de (39), c'est une extrémale de Pontriaguine.*

On peut prouver ce résultat en étendant la technique de [13].

**7.2. Ordre des contraintes sur l'état et conditions de jonction.** Pour simplifier l'exposé on suppose l'absence de contraintes sur la commande et l'extrémale qualifiée ( $\bar{\beta} = 1$ ). Notons les sauts des fonctions du temps ayant des limites à droite et à gauche (c'est le cas des fonctions à variation bornée) à l'instant  $\tau \in ]0, T[$ , par un crochet, ainsi  $[\bar{\mathbf{p}}(\tau)] := \bar{\mathbf{p}}(\tau_+) - \bar{\mathbf{p}}(\tau_-)$ , et posons  $\bar{\nu}_j(\tau) := [\bar{\boldsymbol{\mu}}_j(\tau)]$ . En raison de (102), on a pour  $\tau \in [0, T]$  :

$$(105) \quad -[\bar{\mathbf{p}}(\tau)] = \sum_{j=1}^{n_g} \bar{\nu}_j(\tau) \nabla g_j(\bar{\mathbf{y}}(\tau)).$$

On appelle *dérivée totale* de  $g$  (suivant la dynamique  $f$ ) la fonction  $g^{(1)}(u, y) := g'(y)f(u, y)$ . Le long d'une trajectoire  $(\mathbf{u}, \mathbf{y})$ ,  $g^{(1)}(\mathbf{u}, \mathbf{y})$  est la dérivée de  $g(\mathbf{y})$ . Si  $g^{(1)}(u, y)$  dépend de  $u$ , au sens où  $g_u^{(1)}(u, y)$  n'est pas identiquement nulle, on dit que  $g$  est une *contrainte du premier ordre* sur l'état. Dans le cas contraire on peut écrire  $g^{(1)}$  comme fonction de  $y$  seul, et former la dérivée totale à l'ordre deux, omettant les arguments de  $f = f(u, y)$  :

$$g^{(2)}(u, y) := Dg^{(1)}(y)f(u, y) = g''(y)(f, f) + g'(y)f_y f.$$

Si  $g^{(2)}(u, y)$  dépend de  $u$ , on dit que  $g$  est d'ordre 2 et ainsi de suite pour un ordre quelconque. Si  $g$  est vectorielle, on note  $q_i$  l'ordre supposé bien défini de sa  $i$ ème composante,  $I_j$  l'ensemble des contraintes d'ordre  $j$ , et  $I_j(t)$  celles actives à l'instant  $t$ . Considérons l'hypothèse

$$(106) \quad \begin{cases} \bar{\mathbf{u}} \text{ est continu et } \{\nabla_u g_j^{(1)}(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t)), j \in I_1(t)\} \\ \text{est une famille libre, pour tout } t \in [0, T]. \end{cases}$$

**Lemme 7.2.** Soient  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  une extrémale de Pontriaguine et  $(\bar{\beta}, \bar{\mathbf{p}}, \bar{\boldsymbol{\mu}})$  un multiplicateur de Pontriaguine associé. Alors (i) Si  $H(\cdot, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$  est (uniformément en temps) fortement convexe,  $\bar{\mathbf{u}}$  est continue. (ii) Si (106) est satisfait, les composantes de  $\bar{\boldsymbol{\mu}}$  associées aux contraintes du premier ordre sont continues sur  $]0, T[$ .

Le point (ii) est conséquence de la relation suivante, qui utilise (105) :

$$0 = [\bar{H}_u(t)] = [\bar{\mathbf{p}}(t)] \cdot \bar{f}_u(t) = \sum_{j=1}^{n_g} \bar{\nu}_j(t) \nabla_u g_j^{(1)}(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t)).$$

Montrons maintenant le lien entre le saut de la dérivée de la commande et les sauts de  $\boldsymbol{\mu}$  ou de sa dérivée.

**Lemme 7.3.** Si (106) est satisfait, alors (i)  $\bar{H}_{uu}(t)$  est continue, et (ii) si les sauts ci-dessous sont bien définis, on a pour tout  $\tau \in ]0, T[$  :

$$(107) \quad \bar{H}_{uu}(\tau)[\dot{\bar{\mathbf{u}}}(\tau)] = \sum_{j \in I_1(\tau)} [\dot{\bar{\boldsymbol{\mu}}}_j(\tau)] \nabla_u g_j^{(1)}(\tau) - \sum_{j \in I_2(\tau)} \bar{\nu}_j(\tau) \nabla_u g_j^{(2)}(\tau).$$

**Démonstration.** (i) Si  $j \in I_1(\tau)$ , alors  $\bar{\nu}_j(\tau) = 0$  d'après le lemme 7.2 (ii), et sinon,  $g_j^{(1)}(u, y)$  ne dépend pas de  $u$ , de sorte que, pour  $1 \leq i, k \leq m$  :

$$-[\bar{H}_{u_i u_k}(\tau)] = -[\bar{\mathbf{p}}(\tau)] \cdot \bar{f}_{u_i u_k}(\tau) = \sum_{j=1}^{n_g} \bar{\nu}_j(\tau) \bar{g}'_j(\tau) \bar{f}_{u_i u_k}(\tau) = \sum_{j=1}^{n_g} \bar{\nu}_j(\tau) D_{u_i u_k}^2 g_j^{(1)}(\bar{\mathbf{u}}(\tau), \bar{\mathbf{y}}(\tau)) = 0.$$

(ii) Soit  $i \in \{1, \dots, m\}$ . On a, supposant pour simplifier  $\ell = 0$  et omettant le temps en argument :

$$\frac{d}{dt} H_{u_i} = \frac{d}{dt} (\bar{\mathbf{p}} \cdot \bar{f}_{u_i}) = \bar{H}_{u_i u} \dot{\bar{\mathbf{u}}} + \bar{\mathbf{p}}^\dagger \bar{f}_{u_i y} \bar{f} + \dot{\bar{\mathbf{p}}} \cdot \bar{f}_{u_i}.$$

Utilisant l'équation adjointe et le fait que  $\bar{g}'_j \bar{f}_u = 0$  si  $j \notin I_1(\tau)$ , il vient :

$$(108) \quad \frac{d}{dt} H_{u_i} = \bar{H}_{u_i u} \dot{\bar{\mathbf{u}}} + \bar{\mathbf{p}} \cdot (\bar{f}_{u_i y} \bar{f} - \bar{f}_y \bar{f}_{u_i}) - \sum_{j \in I_1} \dot{\bar{\boldsymbol{\mu}}}_j \bar{g}'_j \bar{f}_{u_i}.$$

Comme  $\bar{H}_{u_i} = \dot{\bar{H}}_{u_i} = 0$  p.p., on a  $[\dot{\bar{H}}_{u_i}] = 0$ . La commande et  $\bar{H}_{uu}$  étant continus, ceci implique

$$(109) \quad \bar{H}_{u_i u}[\dot{\bar{\mathbf{u}}}] = -[\bar{\mathbf{p}}] \cdot (\bar{f}_{u_i y} \bar{f} - \bar{f}_y \bar{f}_{u_i}) + \sum_{j \in I_1} [\dot{\bar{\boldsymbol{\mu}}}_j] \bar{g}'_j \bar{f}_{u_i}.$$

Or, si  $j \notin I_1(\tau)$ , on a  $0 = \left( \frac{\partial}{\partial y} (\bar{g}'_j \bar{f}_{u_i}) \right) \bar{f} = \bar{g}''_j(\bar{f}, \bar{f}_{u_i}) + \bar{g}'_j \bar{f}_{u_i} \bar{f}$ , et donc

$$(110) \quad \begin{aligned} g_{j,u_i}^{(2)} &= D_{u_i}(\bar{g}''_j(\bar{f}, \bar{f}) + \bar{g}'_j \bar{f}_y \bar{f}) \\ &= 2\bar{g}''_j(\bar{f}, \bar{f}_{u_i}) + \bar{g}'_j \bar{f}_{u_i} \bar{f} + \bar{g}'_j \bar{f}_y \bar{f}_{u_i} \\ &= \bar{g}'_j(\bar{f}_y \bar{f}_{u_i} - \bar{f}_{u_i} \bar{f}_y). \end{aligned}$$

On conclut en remplaçant  $[\bar{\mathbf{p}}] = -\sum_{j=1}^{n_g} \bar{v}_j \nabla \bar{g}_j$  dans (109), et en notant que  $\bar{v}_j = 0$  quand  $q_j = 1$ , et  $\bar{g}_{j,u_i}^{(2)} = 0$  si  $q_j > 2$ . ■

*Exemple 7.4.* Supposons (106) satisfait,  $\bar{H}_{uu}(\tau)$  inversible, et la contrainte sur l'état scalaire. Si elle est du premier ordre et  $g_u^{(1)}(\tau) \neq 0$ , on obtient  $\bar{H}_{uu}(\tau)[\dot{\bar{\mathbf{u}}}(\tau)] = [\dot{\bar{\boldsymbol{\mu}}}(\tau)] \nabla_u g^{(1)}(\tau)$  et donc  $\dot{\bar{\mathbf{u}}}(\tau)$  est discontinue ssi  $\dot{\bar{\boldsymbol{\mu}}}$  l'est. Si la contrainte est du second ordre et  $g_u^{(2)}(\tau) \neq 0$ , on obtient  $\bar{H}_{uu}(\tau)[\dot{\bar{\mathbf{u}}}(\tau)] = -\bar{v}(\tau) \nabla_u g^{(2)}(\tau)$  et  $\dot{\bar{\mathbf{u}}}(\tau)$  est discontinue ssi  $\dot{\bar{\boldsymbol{\mu}}}$  l'est. Si la contrainte n'est pas d'ordre un ou deux, alors  $\dot{\bar{\mathbf{u}}}(\tau)$  est continue. On voit que le comportement aux jonctions dépend d'une façon essentielle de l'ordre de la contrainte.

*Exemple 7.5.* Soit le problème de calcul de l'équilibre d'une *corde élastique pesante* de longueur  $T$ , soumise à un obstacle et fixée à ses extrémités. Plus spécifiquement, on minimise la somme des énergies élastique et potentielle :  $E(y) := \int_0^T (\frac{1}{2} \dot{\bar{\mathbf{y}}}(t)^2 + \bar{\mathbf{y}}(t)) dt$ , où  $\bar{\mathbf{y}}(t)$  représente la déformation verticale, sous les contraintes  $\bar{\mathbf{y}}(0) = \bar{\mathbf{y}}(T) = 1$  et  $\bar{\mathbf{y}}(t) \geq 0$  pour tout  $t$ . On se ramène au cadre de la commande optimale avec l'équation d'état  $\dot{\bar{\mathbf{y}}}(t) = \bar{\mathbf{u}}(t)$ , un critère intégral d'intégrande  $\ell(u, y) = \frac{1}{2} u^2 + y$ , et la contrainte sur l'état  $g(y) = -y$ . Quand  $T$  est assez grand, la solution optimale est, pour un certain  $t_0 \in ]0, \frac{1}{2}T[$ , nulle sur  $[t_0, t_1]$  avec  $t_1 = T - t_0$ , et strictement positive en dehors. Le préhamiltonien est  $H(u, y, p) = \frac{1}{2} u^2 + y + pu$ , et la dynamique adjointe est  $-d\bar{\mathbf{p}}(t) = dt - d\bar{\boldsymbol{\mu}}(t)$ , donc (comme  $\bar{\boldsymbol{\mu}}(T) = 0$ ),  $\bar{\mathbf{p}}(t) = T - t + \bar{\boldsymbol{\mu}}(t) + \bar{\mathbf{p}}(T)$ . L'inégalité hamiltonienne donne  $\bar{\mathbf{u}}(t) = -\bar{\mathbf{p}}(t)$ . Sur  $]t_0, t_1[$ , on a donc  $0 = \bar{\mathbf{u}}(t) = -\bar{\mathbf{p}}(t)$ , ce qui implique  $\dot{\bar{\boldsymbol{\mu}}} = 1$ . Sur chaque arc (intervalle de temps) non contraint, on a  $\bar{\mathbf{p}}(t) = c - t$  pour un certain  $c \in \mathbb{R}$  et donc  $\bar{\mathbf{u}}(t) = t - c$ . D'après le lemme 7.2, la commande ainsi que  $\bar{\boldsymbol{\mu}}$  sont continus en  $t_0$  et  $t_1$ , donc la commande vaut  $t - t_0$  sur  $[0, t_0[$  et  $t - t_1$  sur  $[t_1, T[$ . Le saut de  $\dot{\bar{\mathbf{u}}}$  vaut -1 en  $t_0$  et 1 en  $t_1$ , et (comme  $\dot{\bar{\boldsymbol{\mu}}} = 0$  hors de  $[t_0, t_1]$ ) le saut de  $\dot{\bar{\boldsymbol{\mu}}}$  vaut 1 en  $t_0$  et -1 en  $t_1$ , de sorte que (107) est bien satisfait.

*Exemple 7.6.* Soit le problème de calcul de l'équilibre d'une *poutre élastique pesante* de longueur  $T$ , soumise à un obstacle et fixée à ses extrémités. On minimise la somme des énergies élastique et potentielle :  $E(x) := \int_0^T (\frac{1}{2} \ddot{\bar{\mathbf{x}}}(t)^2 + \bar{\mathbf{x}}(t)) dt$ , où  $\bar{\mathbf{x}}(t)$  représente la déformation verticale, sous les contraintes  $\bar{\mathbf{x}}(0) = \bar{\mathbf{x}}(T) = 1$  et  $\bar{\mathbf{x}}(t) \geq 0$  pour tout  $t$ . On se ramène au cadre de la commande optimale avec l'équation d'état  $\dot{\bar{\mathbf{y}}}_1(t) = \bar{\mathbf{y}}_2(t)$ ,  $\dot{\bar{\mathbf{y}}}_2(t) = \bar{\mathbf{u}}(t)$ , un critère intégral d'intégrande  $\ell(u, y) = \frac{1}{2} u^2 + y_1$ , et la contrainte sur l'état  $g(y) = -y_1$ , et les contraintes aux extrémités, par exemple  $\bar{\mathbf{y}}_1(0) = \bar{\mathbf{y}}_1(T) = 1$ . Quand  $T$  est assez grand, la solution optimale est, pour un certain  $t_0 \in ]0, \frac{1}{2}T[$ , nulle sur  $[t_0, t_1]$  avec  $t_1 = T - t_0$ , et strictement positive en dehors. Le préhamiltonien est  $H(u, y, p) = \frac{1}{2} u^2 + y_1 + p_1 y_2 + p_2 u$ , et la dynamique adjointe est  $-d\bar{\mathbf{p}}_1(t) = dt - d\bar{\boldsymbol{\mu}}(t)$ ,  $-\dot{\bar{\mathbf{p}}}_2(t) = \bar{\mathbf{p}}_1(t)$ , donc (comme  $\bar{\boldsymbol{\mu}}(T) = 0$ ),  $\bar{\mathbf{p}}_1(t) = T - t + \bar{\boldsymbol{\mu}}(t) + \bar{\mathbf{p}}_1(T)$ . L'inégalité hamiltonienne donne  $\bar{\mathbf{u}}(t) = -\bar{\mathbf{p}}_2(t)$ . Sur  $]t_0, t_1[$ , on a donc

$$0 = \bar{\mathbf{u}}(t) = -\bar{\mathbf{p}}_2(t) = -\dot{\bar{\mathbf{p}}}_2(t) = \bar{\mathbf{p}}_1(t),$$

ce qui implique  $\dot{\bar{\boldsymbol{\mu}}} = 1$ . Sur chaque arc non contraint, on a  $\bar{\mathbf{p}}_1(t) = c_1 - t$  pour un certain  $c_1$ , donc  $\bar{\mathbf{p}}_2(t) = \frac{1}{2} t^2 - c_1 t - c_2$  pour un certain  $c_2$ , et finalement  $\bar{\mathbf{u}}(t) = c_1 t + c_2 - \frac{1}{2} t^2$ . D'après le lemme 7.2, la commande est continue en  $t_0$  et  $t_1$ , et est donc de la forme  $-c'_1(t - t_0) - \frac{1}{2}(t - t_0)^2$  sur  $[0, t_0[$  et  $c'_1(t - t_1) - \frac{1}{2}(t - t_1)^2$  sur  $[t_1, T[$ . Comme à l'équilibre la courbure de la poutre (donc la commande) est positive, on a  $[\dot{\bar{\mathbf{u}}}(t_0)] = c'_1 \geq 0$  et  $[\dot{\bar{\mathbf{u}}}(t_1)] = c'_1 \geq 0$ . La contrainte sur l'état est du second ordre,  $g_{j,u}^{(1)} = 0$  et  $g_{j,u}^{(2)}(t) = -1$ , donc  $[\dot{\bar{\mathbf{u}}}(t_i)] = \bar{v}(t_i) \geq 0$ ,  $i = 1, 2$ , d'après (107). On retrouve bien la positivité des sauts de  $\dot{\bar{\mathbf{u}}}$  en  $t_0$  et  $t_1$ .

**7.3. Système d'optimalité alternatif.** On se limitera au cas où  $\ell = 0$ , la contrainte sur l'état étant scalaire et du premier ordre. On introduit le *préhamiltonien alternatif*  $H^1 : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  tel que

$$H^1(u, y, p^1, \mu^1) := p^1 \cdot f(u, y) + \mu^1 g^{(1)}(u, y).$$

Définissons le multiplicateur et état adjoint alternatifs  $(\bar{\boldsymbol{\mu}}^1, \bar{\mathbf{p}}^1)$  par  $\bar{\boldsymbol{\mu}}^1(t) := -\bar{\boldsymbol{\mu}}(t)$ ,  $\bar{\mathbf{p}}^1(t) := \bar{\mathbf{p}}(t) + \bar{\boldsymbol{\mu}}(t) \nabla g(\bar{\mathbf{y}}(t))$  pour tout  $t$ . Alors  $\bar{\mathbf{p}}^1$  est solution de l'équation

$$-\dot{\bar{\mathbf{p}}^1}(t) = \nabla_y H^1(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}^1(t), \bar{\boldsymbol{\mu}}^1(t)) \quad \text{p.p.}$$

Omettant les arguments, on a donc la forme hamiltonienne  $\frac{d}{dt}(\bar{\mathbf{y}}(t), -\bar{\mathbf{p}}^1(t)) = (H_{p^1}^1, -H_y^1)$ . De plus, on a pour tout  $u \in \mathbb{R}^m$

$$H^1(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}^1(t), \bar{\boldsymbol{\mu}}^1(t)) = (\bar{\mathbf{p}}^1(t) + \bar{\boldsymbol{\mu}}^1(t) \nabla g(\bar{\mathbf{y}}(t))) \cdot f(u, \bar{\mathbf{y}}(t)) = H(u, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t)).$$

La stationnarité ou minimalité de  $H(\cdot, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$  au point  $\bar{\mathbf{u}}(t)$  équivaut donc à la même propriété pour  $H^1(\cdot, \bar{\mathbf{y}}(t), \bar{\mathbf{p}}^1(t), \bar{\boldsymbol{\mu}}^1(t))$ . L'avantage de la formulation alternative est que, sur un intervalle de temps où la contrainte sur l'état est active, les variables "algébriques"  $(\bar{\mathbf{u}}, \bar{\boldsymbol{\mu}}^1)$  peuvent dans certains cas s'obtenir par élimination dans les deux équations algébriques (supposant ici que  $U_{ad} = \mathbb{R}^m$ ) :

$$0 = \nabla_u H^1(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t), \bar{\mathbf{p}}^1(t), \bar{\boldsymbol{\mu}}^1(t)) = f_u(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t))^\dagger \bar{\mathbf{p}}^1(t) + \bar{\boldsymbol{\mu}}^1(t) \nabla_u g^{(1)}(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t))$$

et sur un arc où la contrainte sur l'état est active

$$0 = \frac{d}{dt} g(\bar{\mathbf{y}}(t)) = g^{(1)}(\bar{\mathbf{u}}(t), \bar{\mathbf{y}}(t)).$$

En effet le jacobien de cette paire d'équation par rapport au variables algébriques  $(\bar{\mathbf{u}}, \bar{\boldsymbol{\mu}}^1)$  est, omettant les arguments

$$(111) \quad \begin{pmatrix} \bar{H}_{uu}^1 & \nabla_u \bar{g}^{(1)} \\ \bar{g}_u^{(1)} & 0 \end{pmatrix}.$$

Ce jacobien est inversible, si par exemple la commande est continue, la condition de Legendre-Clebsch forte (73) est satisfaite, et si  $\bar{g}_u^{(1)}(t) \neq 0$ , pour tout  $t \in [0, T]$ . À partir de ce type de propriété on peut construire une extension des algorithmes de tir pour les problèmes avec contraintes sur l'état.

*Exemple 7.7.* On reprend l'exemple 7.5, par l'approche alternative. Le préhamiltonien alternatif est  $\frac{1}{2}u^2 + y + (p^1 - \mu^1)u$ . De là  $-\dot{\bar{\mathbf{p}}}(t) = 1$ , d'où  $\bar{\mathbf{p}}^1(t) = T - t + \bar{\mathbf{p}}(T)$ . L'inégalité hamiltonienne « alternative » donne  $\bar{\mathbf{u}}(t) + \bar{\mathbf{p}}^1(t) - \bar{\boldsymbol{\mu}}^1(t) = 0$ . Sur l'arc contraint le jacobien (111) est inversible,  $\bar{\mathbf{u}}(t) = 0$ , et  $\bar{\boldsymbol{\mu}}^1(t) = \bar{\mathbf{p}}^1(t) = T - t + \bar{\mathbf{p}}(T)$ .

## 8. PROBLÈMES AFFINES EN LA COMMANDE

**8.1. Cadre : coût final et temps optimal.** Considérons des problèmes avec *coût final*, commande scalaire, et une dynamique de la forme  $f(u, y) = f_0(y) + u f_1(y)$ , où  $f_0$  et  $f_1$  sont deux champs de vecteurs (applications de  $\mathbb{R}^n$  dans lui même), avec état initial fixé. L'équation d'état est donc

$$(112) \quad \dot{\mathbf{y}}(t) = f_0(\mathbf{y}(t)) + \mathbf{u}(t) f_1(\mathbf{y}(t)) \quad \text{p.p., } \mathbf{y}(0) = \mathbf{y}^0,$$

et on a une contrainte finale

$$(113) \quad \Phi_F(\mathbf{y}(T)) \in K_F$$

et une contrainte sur la commande

$$(114) \quad \mathbf{u}(t) \in [-1, 1] \quad \text{p.p.}$$

Le problème de commande optimale est

$$\text{Min}_{(\mathbf{u}, \mathbf{y})} \varphi(\mathbf{y}(T)); \quad \text{soumis à (112)-(114).}$$

On suppose toutes les fonctions de classe  $C^\infty$ . Le préhamiltonien est  $p \cdot (f_0(y) + u f_1(y))$ . D'après le théorème 4.3 et l'exemple 4.5, une solution  $(\bar{\mathbf{u}}, \bar{\mathbf{y}})$  est une extrémale de Pontriaguine : il existe un multiplicateur  $(\bar{\beta}, \bar{\Psi}_F) \in \mathbb{R}_+ \times N_{K_F}(\Phi_F(\bar{\mathbf{y}}(T)))$  non nul tel que l'état adjoint solution de

$$\begin{cases} -\dot{\bar{\mathbf{p}}}(t) &= (f_0'(\bar{\mathbf{y}}(t)) + \bar{\mathbf{u}}(t) f_1'(\bar{\mathbf{y}}(t)))^\dagger \bar{\mathbf{p}}(t) \quad \text{p.p.,} \\ \bar{\mathbf{p}}(T) &= \bar{\beta} \nabla \varphi(\bar{\mathbf{y}}(T)) + D\Phi_F(\bar{\mathbf{y}}(T))^\dagger \bar{\Psi}_F \end{cases}$$

vérifie l'inégalité hamiltonienne (64), qui se ramène ici à

$$(115) \quad (\bar{\mathbf{u}}(t) - u) \bar{\mathbf{p}}(t) \cdot f_1(\bar{\mathbf{y}}(t)) \leq 0, \quad \text{pout tout } u \in [-1, 1].$$

*Transfert en temps minimal.* La théorie qui suit s'applique aussi si on remplace le coût final par la minimisation de l'horizon ; dans la dynamique adjointe, il faut alors remplacer la condition finale par

$$\bar{\mathbf{p}}(T) = D\Phi(\bar{\mathbf{y}}(T))^\dagger \bar{\Psi}_F.$$

En raison du lemme 4.12,  $\bar{\Psi}_F \neq 0$ .

*Non nullité de l'état adjoint.* La dynamique de l'état adjoint fait que ce dernier est nul à un instant donné ssi il est identiquement nul, et alors l'inégalité hamiltonienne ne donne pas d'information. On peut exclure ce cas : (i) dans le problème avec coût final, comme  $(\bar{\beta}, \bar{\Psi}_F) \neq 0$ , si la famille  $\{\nabla \varphi(\bar{\mathbf{y}}(T)), \nabla \Phi_i(\bar{\mathbf{y}}(T)), i =$

$1, \dots, n_\Phi\}$  est linéairement indépendante, et (ii) pour le problème de transfert en temps minimal si  $D\Phi(\bar{\mathbf{y}}(T))$  est surjective (lemme 4.12). Nous ferons donc dans certains cas l'hypothèse d'un état adjoint non nul.

**8.2. Arcs bang et singuliers.** Notant  $\Xi(t) := \bar{\mathbf{p}}(t) \cdot f_1(\bar{\mathbf{y}}(t))$  la *fonction de commutation*, on peut exprimer l'inégalité hamiltonienne (115), pour  $t \in [0, T]$ , comme

$$(116) \quad \begin{cases} \bar{\mathbf{u}}(t) = 1 & \text{si } \Xi(t) < 0, \\ \bar{\mathbf{u}}(t) = -1 & \text{si } \Xi(t) > 0, \\ \bar{\mathbf{u}}(t) \in [-1, 1] & \text{si } \Xi(t) = 0. \end{cases}$$

Soient  $0 \leq a < b \leq T$ . On dira que  $(a, b)$  est un *arc singulier* (arc  $S$ ) si  $\bar{\mathbf{u}}(t) \in ]-1, 1[$  p.p. sur  $]a, b[$ , et un *arc bang* si  $\bar{\mathbf{u}}$  est constant, égal à  $\pm 1$  sur  $]a, b[$  (on parlera d'arc  $B_-$  s'il vaut  $-1$ , d'arc  $B_+$  sinon).

Soient  $X$  et  $Y$  deux champs de vecteurs. Leur *crochet de Lie* est le champ de vecteurs défini par

$$[X, Y](y) := X'(y)Y(y) - Y'(y)X(y),$$

avec  $X'_{ij}(y) = \partial X_i(y)/\partial y_j$ , soit

$$[X, Y]_i(y) := \sum_{j=1}^n X'_{ij}(y)Y_j(y) - Y'_i(y)X_j(y).$$

L'application  $(X, Y) \mapsto [X, Y]$  est bilinéaire et antisymétrique et, si  $X$  est un champ de vecteurs, alors

$$\frac{d}{dt}(\bar{\mathbf{p}}(t) \cdot X(\bar{\mathbf{y}}(t))) = \bar{\mathbf{p}}(t) \cdot [X, f_0](\bar{\mathbf{y}}(t)) + \bar{\mathbf{u}}(t)\bar{\mathbf{p}}(t) \cdot [X, f_1](\bar{\mathbf{y}}(t)).$$

Comme  $[f_1, f_1] = 0$ , prenant  $X = f_1$ , il vient successivement

$$\dot{\Xi}(t) = \bar{\mathbf{p}}(t) \cdot [f_1, f_0](\bar{\mathbf{y}}(t)),$$

$$\ddot{\Xi}(t) = \bar{\mathbf{p}}(t) \cdot [[f_1, f_0], f_0](\bar{\mathbf{y}}(t)) + \bar{\mathbf{u}}(t)\bar{\mathbf{p}}(t) \cdot [[f_1, f_0], f_1](\bar{\mathbf{y}}(t)).$$

Sur un arc singulier,  $\ddot{\Xi}(t) = 0$ , donc si  $\bar{\mathbf{p}}(t) \cdot [[f_1, f_0], f_1](\bar{\mathbf{y}}(t)) \neq 0$ , on peut exprimer la commande en fonction de l'état et de l'état adjoint, soit, utilisant l'antisymétrie du crochet :  $\bar{\mathbf{u}}(t) = \Gamma(\bar{\mathbf{y}}(t), \bar{\mathbf{p}}(t))$ , où

$$(117) \quad \Gamma(y, p) := p \cdot [[f_1, f_0], f_0](y) / p \cdot [[f_0, f_1], f_1](y).$$

Il faut quelquefois dériver la fonction de commutation plus de deux fois pour obtenir une expression de la commande en fonction de l'état et de l'état adjoint, et cette expression est toujours obtenue après un nombre pair de dérivations, comme l'illustre l'exemple suivant.

*Exemple 8.1.* Soit le problème de Fuller [15] : minimisation de  $\frac{1}{2} \int_0^T \mathbf{y}_1(t)^2 dt$ , avec équation d'état  $\dot{\mathbf{y}}_1(t) = \mathbf{y}_2(t)$ ,  $\dot{\mathbf{y}}_2(t) = \mathbf{u}(t)$ , état initial fixé, et  $\mathbf{u}(t) \in [-1, 1]$ . Pour se ramener à un coût final, on introduit l'état  $\mathbf{y}_3$  de dynamique  $\dot{\mathbf{y}}_3(t) = \frac{1}{2}\mathbf{y}_1(t)^2$ , avec  $\mathbf{y}_3(0) = 0$ , et le coût devient  $\mathbf{y}_3(T)$ . Le préhamiltonien est  $p_1 y_2 + p_2 u + \frac{1}{2} p_3 y_1^2$ , et l'équation adjointe est  $-\dot{\bar{\mathbf{p}}}_1(t) = \bar{\mathbf{p}}_3(t)\bar{\mathbf{y}}_1(t)$ ,  $-\dot{\bar{\mathbf{p}}}_2(t) = \bar{\mathbf{p}}_1(t)$ ,  $-\dot{\bar{\mathbf{p}}}_3(t) = 0$ , avec la condition finale  $\bar{\mathbf{p}}(T) = (0 \ 0 \ 1)^\dagger$ , donc  $\bar{\mathbf{p}}_3(t) = 1$  pour tout  $t$ . Il faut dériver quatre fois la fonction de commutation  $\Xi(t) = \bar{\mathbf{p}}_2(t)$  pour obtenir  $\Xi^{(4)}(t) = \bar{\mathbf{u}}(t)$ . L'état et la commande sont nuls sur un arc singulier. On généralise facilement cet exemple à celui d'un intégrateur  $n$  fois, pour lequel il faut dériver  $2n$  fois la fonction de commutation pour retrouver une expression de la commande. Notons que la commande optimale comporte en général un nombre infini de commutations entre  $\pm 1$ .

**8.3. Cas particuliers - Problèmes de dimension 2 et 3.** On dira que  $t \in [0, T]$  est *critique* si  $\Xi(t) = 0$ , et *bicritique* si de plus  $\dot{\Xi}(t) = 0$ . Notons  $\mathcal{T}$  l'ensemble des instants critiques. Comme  $\dot{\Xi}(\cdot)$  est continue, un instant critique mais non bicritique est un point isolé du fermé  $\mathcal{T}$ . En particulier, s'il n'existe aucun instant bicritique, alors  $[0, T] \setminus \mathcal{T}$  est une union finie d'arcs bang. Si  $t$  est bicritique, alors  $\bar{\mathbf{p}}(t)$  est orthogonal à  $f_1(t)$  et  $[f_1, f_0](t)$ .

*Cas de deux états ( $n = 2$ ).* On suppose satisfaite l'hypothèse d'état adjoint non nul. Si  $t \in [0, T]$  est tel que  $f_1(t)$  et  $[f_1, f_0](t)$  sont linéairement indépendants,  $t$  ne peut être bicritique, et donc :

- Si  $f_1(t)$  et  $[f_1, f_0](t)$  sont linéairement indépendants pour tout  $t \in [0, T]$ , alors  $[0, T] \setminus \mathcal{T}$  est une union finie d'arcs bang.
- Sinon, les états aux instants bicritiques (et en particulier les arcs singuliers) sont contenus dans la *variété singulière* définie par (notant par  $\wedge$  le produit vectoriel)

$$\{y \in \mathbb{R}^2; f_1(y) \wedge [f_1, f_0](y) = 0\}.$$

*Exemple 8.2.* Soit le problème de minimiser  $\frac{1}{2} \int_0^T \bar{\mathbf{x}}(t)^2 dt$  sous la contrainte  $\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{u}}(t) \in [-1, 1]$  (qui se résout facilement!). On se ramène à un coût final en posant  $\bar{\mathbf{y}}(t) \in \mathbb{R}^2$ ,  $\dot{\bar{\mathbf{y}}}_1(t) = u(t)$ ,  $\dot{\bar{\mathbf{y}}}_2(t) = \frac{1}{2} \bar{\mathbf{y}}_1(t)^2$ , de dynamique  $f_0(\bar{\mathbf{y}}) = (0 \ \frac{1}{2} \bar{\mathbf{y}}_1^2)^\dagger$ ,  $f_1(\bar{\mathbf{y}}) = (1 \ 0)^\dagger$ . Il vient

$$[f_1, f_0](\bar{\mathbf{y}}) = -f_0'(\bar{\mathbf{y}}) f_1(\bar{\mathbf{y}}) = (0 \ -y_1)^\dagger,$$

et finalement  $f_1(\bar{\mathbf{y}}) \wedge [f_1, f_0](\bar{\mathbf{y}}) = -y_1$ . La variété singulière est donc l'axe vertical  $y_1 = 0$  (sur lequel la commande est nulle).

*Cas de trois états ( $n = 3$ ).* Posons  $G(\bar{\mathbf{y}}) := f_1(\bar{\mathbf{y}}) \wedge [f_1, f_0](\bar{\mathbf{y}})$ . Si  $t$  est un instant bicritique tel que  $f_1(t)$  et  $[f_1, f_0](t)$  sont linéairement indépendants, alors l'état adjoint est colinéaire à  $G(\bar{\mathbf{y}}(t))$ . En particulier, quand  $\bar{\mathbf{p}}(t) \cdot [[f_0, f_1], f_1](\bar{\mathbf{y}}(t))$  ne s'annule pas sur un arc singulier, comme seule la direction de l'adjoint intervient dans l'expression (117) de la commande, on obtient l'expression de la commande fonction de l'état :  $\bar{\mathbf{u}}(t) = \Gamma(\bar{\mathbf{y}}(t), G(\bar{\mathbf{y}}(t)))$ .

Si de plus l'horizon est libre, le préhamiltonien s'annule le long de la trajectoire optimale. Sur un arc singulier, comme  $\bar{\mathbf{p}}(y) \cdot f_1(\bar{\mathbf{y}}(t)) = 0$ , on en déduit que  $\bar{\mathbf{p}}(y) \cdot f_0(\bar{\mathbf{y}}(t)) = 0$ . Si l'état adjoint ne s'annule pas, les états aux instants bicritiques (et en particulier les arcs singuliers) sont contenus dans la *variété singulière*  $\{y \in \mathbb{R}^n; G(y) \wedge f_0(y) = 0\}$ .

*Exemple 8.3 (Problème de Goddard).* Il s'agit d'une variante « monodimensionnelle » de l'exemple de l'introduction. On désire maximiser l'altitude finale d'un vol vertical de fusée. Les variables d'état sont l'altitude  $h$ , la vitesse  $V$  et la masse  $m$ , de dynamique  $\dot{\mathbf{h}} = V$ ,  $\mathbf{m}\dot{V} = c\mathbf{u} - D(\mathbf{h}, V)$ ,  $\dot{\mathbf{m}} = -\mathbf{u}$  pour un certain  $c > 0$ ,  $D(\mathbf{h}, V) \geq 0$  représentant la traînée aérodynamique, avec contrainte  $\mathbf{u}(t) \in [0, 1]$  pour tout  $t$ , et  $\mathbf{m}(T) \geq m_T$ . On maximise donc  $h(T)$  avec horizon libre. On peut appliquer à cet exemple la théorie précédente (nous ne détaillerons pas les calculs), ce qui permet d'exprimer la commande fonction de l'état sur l'arc singulier, et d'explicitier l'équation de la variété singulière.

**8.4. Contraintes sur l'état.** En présence de contraintes sur l'état de type (97), le théorème 7.1 s'applique et l'expression (105) des sauts de l'état adjoint reste valable. Nous allons reprendre l'étude des conditions de jonction, en tenant compte de la contrainte sur la commande ainsi que de la possibilité d'un saut de cette dernière. Si  $\tau \in ]0, T[$ , notons d'abord que (116) implique  $[H_u(\tau)][\bar{\mathbf{u}}(\tau) \leq 0$ . On note, pour  $t \in [0, T]$ , l'ensemble de contraintes du premier ordre régulières en  $\tau$  par :

$$I_1^+(t) := \{j \in I_1(t); g_j'(\bar{\mathbf{y}}(t)) f_1(\bar{\mathbf{y}}(t)) \neq 0\}.$$

**Lemme 8.4.** *Si la commande est discontinue en  $\tau \in ]0, T[$ , alors  $\bar{\nu}_j := [\bar{\boldsymbol{\mu}}_j(\tau)]$  est nul, pour tout  $j \in I_1^+(\tau)$ .*

**Démonstration.** Comme  $[\bar{H}_u(\tau)] = [\bar{\mathbf{p}}(\tau)] \cdot \bar{f}_1(\tau) = -\sum_j \nu_j \bar{g}_j'(\tau) \bar{f}_1(\tau)$ , et que  $\bar{g}_j'(\tau) \bar{f}_1(\tau) = 0$  si  $j \notin I_1^+(\tau)$ , de  $[H_u(\tau)][\bar{\mathbf{u}}(\tau)] \leq 0$ , on déduit que

$$(118) \quad \sum_{j \in I_1^+(\tau)} \nu_j [\bar{\mathbf{u}}(\tau)] \bar{g}_j'(\tau) \bar{f}_1(\tau) \geq 0.$$

Or si  $\nu_j \neq 0$ ,  $\bar{g}_j(\cdot)$  atteint un maximum local en  $\tau$ , donc  $0 \geq [\frac{d}{dt} \bar{g}_j(\tau)] = [\bar{\mathbf{u}}(\tau)] \bar{g}_j'(\tau) \bar{f}_1(\tau)$ . Comme  $\nu_j \geq 0$ , (118) est une somme positive de termes négatifs, donc chaque terme est nul, d'où la conclusion. ■

**8.4.1. Saut de la dérivée de la fonction de commutation.** D'après la dynamique de l'état adjoint, on a sur chaque arc

$$\dot{\Xi}(t) = \bar{\mathbf{p}}(t) \cdot [\bar{f}_1, \bar{f}_0](t) - \sum_j \dot{\bar{\boldsymbol{\mu}}}_j(t) \bar{g}_j'(t) \bar{f}_1(t),$$

et donc pour  $\tau \in ]0, T[$  :

$$[\dot{\Xi}(\tau)] = - \sum_j (\nu_j \bar{g}_j'(\tau) \cdot [\bar{f}_1, \bar{f}_0](\tau) + [\dot{\bar{\boldsymbol{\mu}}}_j(\tau)] \bar{g}_j'(\tau) \bar{f}_1(\tau)).$$

*Exemple 8.5.* Soit  $\tau$  un point de jonction entre un arc singulier et un arc où la contrainte sur l'état (et non celle sur la commande) est active. Alors  $[\dot{\Xi}(\tau)] = 0$ . Supposons  $g$  scalaire, et envisageons deux cas.

(i) Si  $\bar{g}'(\tau) \bar{f}_1(\tau) \neq 0$ , la commande étant discontinue à l'instant  $\tau$ , alors  $\nu = 0$  d'après le lemme 8.4, et

$$0 = [\dot{\Xi}(\tau)] = -[\dot{\bar{\boldsymbol{\mu}}}(\tau)] \bar{g}'(\tau) \bar{f}_1(\tau),$$

donc  $\dot{\bar{\boldsymbol{\mu}}}$  a pour limite zéro en  $\tau$ .

(ii) Si  $\bar{g}'(\tau) \bar{f}_1(\tau) = 0$ , alors

$$0 = [\dot{\Xi}(\tau)] = -\nu \bar{g}'(\tau) \cdot [\bar{f}_1, \bar{f}_0](\tau).$$

Si  $\bar{g}'(\tau) \cdot [\bar{f}_1, \bar{f}_0](\tau) \neq 0$ , on en déduit que  $\bar{\mu}$  est continue à l'instant  $\tau$ .

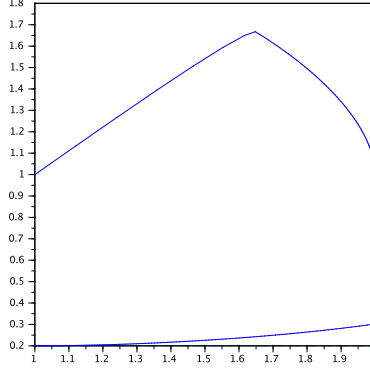


FIGURE 1. Trajectoire proie-prédateur

8.4.2. *Illustration : un problème proie-prédateur.* Dans ce modèle inspiré de [17] (on a ajouté la contrainte sur l'état),  $\mathbf{y}_1$  est le nombre de proies et  $\mathbf{y}_2$  est le nombre de prédateurs, de dynamique  $\dot{\mathbf{y}}_1 = (1 - \mathbf{y}_2)\mathbf{y}_1 - 0.9\mathbf{y}_1\mathbf{u}$ ,  $\dot{\mathbf{y}}_2 = (\mathbf{y}_1 - 1)\mathbf{y}_2$ . On minimise  $\int_0^T (1 + \mathbf{u}(t))dt$ , sous les contraintes  $\mathbf{u} \geq 0$ ,  $\mathbf{y}_1 \leq 2$ , et  $\mathbf{y}(T) = (1, 1)^\dagger$  (point invariant de la dynamique non commandée). L'horizon est libre et la contrainte sur l'état est du premier ordre. La trajectoire calculée avec le logiciel Bocop (développé dans l'équipe de l'auteur) est constituée de 4 arcs : les premier et troisième avec commande nulle, le second avec contrainte d'état active, et le dernier est un arc singulier. La commande ne semble discontinue qu'à la première jonction.

## 9. PROGRAMMATION DYNAMIQUE

9.1. **Principe de programmation dynamique.** On considère une famille de systèmes dynamiques commandés paramétrée par  $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ , l'état et l'instant initial : la dynamique est

$$(119) \quad \dot{\mathbf{y}}(s) = f(\mathbf{u}(s), \mathbf{y}(s)), \quad \text{p.p. } s \in [t, T]; \quad \mathbf{y}(t) = x,$$

avec  $T > 0$  fixé,  $t \in [0, T]$ , et les contraintes sur la commande

$$(120) \quad \mathbf{u}(s) \in U_{ad}, \quad \text{p.p. } s \in [t, T].$$

On supposera dans le reste de l'article  $U_{ad}$  compact, et  $f$  lipschitzienne. Considérons la *valeur de Bellman* du problème de minimisation suivant, paramétré par l'état et le temps initiaux :

$$(121) \quad V(x, t) := \inf_{\mathbf{u}, \mathbf{y}} \int_t^T \ell(\mathbf{u}(s), \mathbf{y}(s)) ds + \varphi(\mathbf{y}(T)) \quad \text{soumis à (119)-(120),}$$

où  $\ell$  et  $\varphi$  sont continues. Ici  $\mathbf{u} \in L^\infty(t, T)^m$  et  $\mathbf{y} \in W^{1, \infty}(t, T)^n$ . En particulier,  $V(x, T) = \varphi(x)$ . On peut vérifier que  $V$  est localement lipschitzienne. Posons

$$\mathcal{U}_{ad}[t, \tau] := \{\mathbf{u} \in L^\infty(t, \tau)^m; \mathbf{u}(s) \in U_{ad} \text{ p.p.}\}.$$

Le *principe de programmation dynamique* s'énonce ainsi, pour tout  $\tau \in ]t, T]$  :

**Théorème 9.1.** *Soient  $0 \leq t \leq \tau \leq T$ . Notant par  $\mathbf{y}[\mathbf{u}]$  la solution de l'équation d'état sur  $[t, \tau]$ , avec condition initiale  $x$ , on a :*

$$(122) \quad V(x, t) = \inf_{\mathbf{u} \in \mathcal{U}_{ad}[t, \tau]} \left( \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}[\mathbf{u}](s)) ds + V(\mathbf{y}[\mathbf{u}](\tau), \tau) \right).$$

**Démonstration.** Notons  $V'(x, t)$  le membre de droite de (122). Soit  $\mathbf{u} \in \mathcal{U}_{ad}[t, \tau]$ ; posons  $\mathbf{y} = \mathbf{y}[\mathbf{u}]$  et  $J(\mathbf{u}, \mathbf{y}) := \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}(s)) ds + \varphi(\mathbf{y}(T))$ . Alors  $J(\mathbf{u}, \mathbf{y}) = \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}(s)) ds + \Delta'$ , avec

$$\Delta' := \int_\tau^T \ell(\mathbf{u}(s), \mathbf{y}(s)) ds + \varphi(\mathbf{y}(T)) \geq V(\mathbf{y}[\mathbf{u}](\tau), \tau),$$



et donc

$$J(\mathbf{u}, \mathbf{y}) \geq \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}(s)) ds + V(\mathbf{y}[\mathbf{u}](\tau), \tau).$$

Minimisant en  $\mathbf{u} \in \mathcal{U}_{ad}[t, \tau]$ , on obtient  $V(x, t) \geq V'(x, t)$ . Réciproquement, soient  $\varepsilon > 0$  et  $\mathbf{u} \in \mathcal{U}_{ad}[t, \tau]$  tel que

$$V'(x, t) + \varepsilon \geq \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}[\mathbf{u}](s)) ds + V(\mathbf{y}[\mathbf{u}](\tau), \tau).$$

Par définition de  $V(\mathbf{y}[\mathbf{u}](\tau), \tau)$ , il existe  $\mathbf{u}' \in \mathcal{U}_{ad}[\tau, T]$  tel que, notant par  $\mathbf{y}'$  la solution associée de l'équation d'état, partant du point  $\mathbf{y}[\mathbf{u}](\tau)$  à l'instant  $\tau$  :

$$V(\mathbf{y}[\mathbf{u}](\tau), \tau) + \varepsilon \geq \int_\tau^T \ell(\mathbf{u}'(s), \mathbf{y}'(s)) ds + \varphi(\mathbf{y}'(T)).$$

Notant par  $\mathbf{u}''$  le recollement de  $\mathbf{u}$  et  $\mathbf{u}'$  (fonction égale à  $\mathbf{u}$  sur  $[t, \tau[$  et à  $\mathbf{u}'$  sur  $[\tau, T]$ ), commande admissible dont l'état associé  $\mathbf{y}''$  est le recollement de  $\mathbf{y}$  et  $\mathbf{y}'$ , on vérifie en additionnant les inégalités ci-dessus que

$$V'(x, t) + 2\varepsilon \geq J(\mathbf{u}'', \mathbf{y}'') \geq V(x, t).$$

On conclut en faisant tendre  $\varepsilon$  vers 0. ■

**9.2. Equation HJB.** Rappelons que le préhamiltonien a pour expression  $H(u, y, p) := \ell(u, y) + p \cdot f(u, y)$ .

**Lemme 9.2.** *Si  $V$  est différentiable en  $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ , alors*

$$V_t(x, t) + \inf_{u \in U_{ad}} H(u, x, \nabla_x V(x, t)) = 0.$$

**Démonstration.** Comme  $\ell$  et  $f$  sont lipschitziennes, du principe de programmation dynamique (122), on déduit facilement les relations

$$(123) \quad V(x, t) = \inf_{\mathbf{u} \in \mathcal{U}_{ad}[t, \tau]} \left( \int_t^\tau \ell(\mathbf{u}(s), x) ds + V(\mathbf{y}[\mathbf{u}](\tau), \tau) \right) + o(\tau - t),$$

et

$$(124) \quad \mathbf{y}[\mathbf{u}](\tau) = x + \int_t^\tau f(\mathbf{u}(t), x) dt + o(\tau - t),$$

avec ici et dans la suite,  $o(\tau - t)/(\tau - t) \rightarrow 0$  quand  $\tau \downarrow t$ , uniformément par rapport à la commande. Comme  $V$  est différentiable en  $(x, t)$ , on a donc

$$(125) \quad V(\mathbf{y}[\mathbf{u}](\tau), \tau) = V(x, t) + V_t(x, t)(\tau - t) + \int_t^\tau \nabla_x V(x, t) \cdot f(\mathbf{u}(t), x) ds + o(\tau - t).$$

Combinant avec (123) et divisant par  $\delta := \tau - t$ , il vient

$$(126) \quad V_t(x, t) + \delta^{-1} \inf_{u \in U_{ad}[t, \tau]} \left\{ \int_t^\tau [\ell(\mathbf{u}(t), x) + \nabla_x V(x, t) \cdot f(\mathbf{u}(t), x)] ds \right\} = o(1).$$

Comme l'infimum ci-dessus est atteint en minimisant séparément pour chaque  $t$ , il vaut  $\delta \inf_{u \in U_{ad}} H(u, x, \nabla_x V(x, t))$ , d'où le résultat. ■

On appellera *équation de Hamilton-Jacobi-Bellman* (HJB), l'équation aux dérivées partielles (EDP) du premier ordre suivante :

$$(127) \quad \begin{cases} \text{(i)} & v_t(x, t) + \inf_{u \in U_{ad}} H(u, x, \nabla_x v(x, t)) = 0, \quad \text{pour tout } t \in [0, T] \text{ et } x \in \mathbb{R}^n, \\ \text{(ii)} & v(x, T) = \varphi(x), \quad \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

*Exemple 9.3.* Soit la dynamique  $\dot{y}(t) = u(t)$ , avec  $|u(t)| \leq 1$ , et un coût intégral nul. L'équation HJB a pour expression  $v_t(x, t) - |\nabla_x v(x, t)| = 0$ . Si le coût final est  $\varphi(x) := e^{-|x|}$ , la stratégie optimale est  $u(t) = y(t)/|y(t)|$  quand  $y(t) \neq 0$ , et donc  $V(x, t) = e^{t-T-|x|}$ , qui vérifie bien l'équation HJB aux points où elle est dérivable (hors de 0).

**9.3. Solution de viscosité.** On sait que  $V(x, t)$  est localement lipschitzienne mais en général non différentiable partout. Une fonction lipschitzienne étant différentiable p.p. (théorème de Rademacher), d'après le lemme 9.2,  $V(x, t)$  vérifie la relation (127)(i) p.p. et bien entendu (127)(ii), mais ce n'est pas la seule fonction qui vérifie cette propriété. Il est donc nécessaire de donner un sens plus restreint qui assure l'unicité de la solution ; c'est l'objet de la théorie des EDP au sens de viscosité. Pour nous conformer aux conventions de signe de cette théorie, réécrivons l'équation HJB sous la forme

$$(128) \quad \begin{cases} \text{(i)} & -v_t(x, t) + \mathcal{H}(x, \nabla_x v(x, t)) = 0 \quad \text{pour tout } t \in [0, T] \text{ et } x \in \mathbb{R}^n, \\ \text{(ii)} & v(x, T) = \varphi(x), \quad \text{pour tout } x \in \mathbb{R}^n, \end{cases}$$

le hamiltonien  $\mathcal{H}(\cdot)$  étant défini par

$$\mathcal{H}(x, p) := \sup_{u \in U_{ad}} (-\ell(u, x) - p \cdot f(u, x)).$$

*Définition 9.4.* Soit  $v$  une fonction continue de  $\mathbb{R}^n \times \mathbb{R}$  vers  $\mathbb{R}$ . On dit que  $v$  est *sous solution* (resp. *sur solution*) de (128)(i) au sens de viscosité si, pour tout  $(x, t) \in \mathbb{R}^n \times ]0, T[$ , et toute fonction  $\Theta : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^1$ , telle que  $v - \Theta$  a un point de maximum (resp. minimum) en  $(x, t)$ , on a  $-\Theta_t(x, t) + \mathcal{H}(x, \nabla_x \Theta(x, t)) \leq 0$  (resp.  $-\Theta_t(x, t) + \mathcal{H}(x, \nabla_x \Theta(x, t)) \geq 0$ ). On dit que  $v$  est solution de (128)(i) au sens de viscosité si c'est une sur et sous solution de (128).

*Remarque 9.5.* Si  $v(x, t)$  est de classe  $C^1$  au voisinage de  $(x_0, t_0)$ , et vérifie l'équation HJB en ce point, alors elle vérifie aussi en ce point les inégalités de la définition ci-dessus.

**Théorème 9.6.** *La fonction valeur  $V$  est solution de viscosité de (128)(i).*

**Démonstration.** (a) Montrons que  $V$  est une sous solution. Soient  $(\bar{x}, \bar{t})$  et  $\Theta$  comme dans la définition de sous solution ; on peut supposer que  $V(x, t) = \Theta(x, t)$ , et donc  $V(x', t') \leq \Theta(x', t')$  si  $(x', t')$  est voisin de  $(x, t)$ . Soit  $\tau \in ]t, T[$ . Si  $\delta = \tau - t$  est assez petit, comme  $U_{ad}$  est compact, on a  $V(\mathbf{y}[\mathbf{u}](\tau), \tau) \leq \Theta(\mathbf{y}[\mathbf{u}](\tau), \tau)$  pour toute commande  $\mathbf{u} \in U_{ad}[t, \tau]$ , donc d'après le principe de programmation dynamique

$$(129) \quad \Theta(x, t) = V(x, t) \leq \inf_{\mathbf{u} \in U_{ad}[t, \tau]} \left( \int_t^\tau \ell(\mathbf{u}(s), \mathbf{y}[\mathbf{u}](s)) ds + \Theta(\mathbf{y}[\mathbf{u}](\tau), \tau) \right).$$

Comme  $\Theta$  est de classe  $C^1$ , reprenant le principe des calculs de la preuve du lemme 9.2, mais avec  $\Theta$  à la place de  $V$  et des inégalités au lieu d'égalités, on déduit que  $-\Theta_t(x, t) + \mathcal{H}(x, \nabla_x \Theta(x, t)) \leq 0$  comme il fallait le montrer. La preuve de sur solution est analogue. ■

Soient  $\mathcal{F}_1$  et  $\mathcal{F}_2$  deux ensembles de fonctions sur  $\mathbb{R}^n \times [0, T]$ . On dit que l'équation HJB vérifie un *principe de comparaison* relatif à  $(\mathcal{F}_1, \mathcal{F}_2)$  si, pour toute sous solution  $u \in \mathcal{F}_1$  et toute sur solution  $v \in \mathcal{F}_2$  de (128)(i), avec  $u(x, T) \leq v(x, T)$  pour tout  $x \in \mathbb{R}^n$ , on a  $u(x, t) \leq v(x, t)$ , pour tout  $(x, t) \in \mathbb{R}^n \times [0, T]$ . Comparant les éléments de  $\mathcal{F}_1$  et  $\mathcal{F}_2$  à  $V$ , on obtient :

**Lemme 9.7.** *On suppose que l'équation HJB vérifie un principe de comparaison relatif à  $(\mathcal{F}_1, \mathcal{F}_2)$ , avec  $V \in \mathcal{F}_1 \cap \mathcal{F}_2$ . Alors  $V$  est la seule solution au sens de viscosité de (128)(i), égale à  $\varphi$  au temps  $T$ , dans l'ensemble  $\mathcal{F}_1 \cap \mathcal{F}_2$ .*

Reste à établir des principes de comparaison. Donnons l'idée motivant une telle propriété en considérant le cas d'un sous solution  $u$  et d'une sur solution  $v$ , toutes deux de classe  $C^1$ . Posons  $\Phi(x, t) := e^t(u(x, t) - v(x, t))$ . Si  $u$  n'est pas toujours inférieure à  $v$ , le supremum de  $\Phi$  est strictement positif ; supposons-le atteint en  $(x, t)$ . Alors  $t < T$  et  $\Phi$  a une dérivée nulle en  $(x, t)$ , soit  $u_x(x, t) = v_x(x, t)$  et

$$(130) \quad u(x, t) - v(x, t) + u_t(x, t) - v_t(x, t) = 0.$$

D'après la remarque 9.5, on a

$$v_t(x, t) \leq \mathcal{H}(x, \nabla_x v(x, t)) = \mathcal{H}(x, \nabla_x u(x, t)) \leq u_t(x, t),$$

donc avec (130),  $u(x, t) \leq v(x, t)$ , ce qui donne la contradiction recherchée.

Dans les faits les semi solutions ne sont pas  $C^1$  et  $\Phi(x, t)$  n'atteint pas nécessairement son maximum. Nous admettrons donc le résultat suivant, renvoyons à [3] pour les détails. On note  $BUC(\mathbb{R}^n)$  l'espace des fonctions bornées et uniformément continues sur  $\mathbb{R}^n$ , et on supposera que

$$(131) \quad f, \ell, \varphi \text{ sont bornées et lipschitziennes.}$$

**Théorème 9.8.** *Si (131) est satisfait, l'équation HJB vérifie un principe de comparaison avec  $\mathcal{F}_1 = \mathcal{F}_2 = BUC(\mathbb{R}^n)$ , et donc, la fonction valeur  $V$  est l'unique élément de  $BUC(\mathbb{R}^n)$ , solution de viscosité de (128)(i), égal à  $\varphi$  quand  $t = T$ .*

9.4. **Problèmes en horizon infini.** Soit le problème en horizon infini de fonction valeur

$$(132) \quad V(x) := \inf_{\mathbf{u} \in \mathcal{U}_{ad}[0, \infty]} \int_0^\infty \ell(\mathbf{u}(t), \mathbf{y}[\mathbf{u}, x](t)) e^{-\lambda t} dt.$$

Ici  $\lambda > 0$  est le coefficient d'actualisation, et  $\mathbf{y}[\mathbf{u}, x]$  est solution de l'équation différentielle (bien définie quand (131) est satisfait)

$$\dot{\mathbf{y}}(t) = f(\mathbf{u}(t), \mathbf{y}(t)), \quad \text{p.p. } t \in ]0, \infty[; \quad \mathbf{y}(0) = x.$$

**Théorème 9.9.** *Si (131) est satisfait, la fonction valeur du problème en horizon infini est l'unique élément de  $BUC(\mathbb{R}^n)$ , solution de viscosité de l'équation*

$$(133) \quad \lambda v(x) + \mathcal{H}(x, \nabla_x v(x)) = 0, \quad \text{pour tout } x \in \mathbb{R}^n.$$

9.5. **Problèmes de temps d'arrêt.** On considère le problème de temps d'arrêt de fonction valeur

$$(134) \quad V(x) := \inf_{(\mathbf{u}, \theta)} \left[ \int_0^\theta \ell(\mathbf{u}(t), \mathbf{y}[\mathbf{u}, x](t)) e^{-\lambda t} dt + \varphi(\mathbf{y}[\mathbf{u}, x](\theta)) e^{-\lambda \theta} \right].$$

Ici  $\lambda > 0$  est le coefficient d'actualisation,  $\theta \in ]0, \infty[$  est le temps d'arrêt,  $\varphi(\cdot)$  est le coût d'arrêt,  $\mathbf{u} \in \mathcal{U}_{ad}[0, \theta]$  est la commande, et  $\mathbf{y}[\mathbf{u}, x]$  est solution de

$$\dot{\mathbf{y}}(t) = f(\mathbf{u}(t), \mathbf{y}(t)), \quad \text{p.p. } t \in ]0, \theta[; \quad \mathbf{y}(0) = x.$$

**Théorème 9.10.** *Si (131) est satisfait, la fonction valeur du problème de temps d'arrêt est l'unique élément de  $BUC(\mathbb{R}^n)$ , solution de viscosité du problème d'obstacle*

$$(135) \quad \max(\lambda v(x) + \mathcal{H}(x, \nabla_x v(x)), v(x) - \varphi(x)) = 0, \quad \text{pour tout } x \in \mathbb{R}^n.$$

Le terme d'obstacle est dû au fait que certains problèmes d'obstacle en mécanique se modélisent par des équations similaires. Faisant  $\theta \downarrow 0$  dans (134), on vérifie que  $v(x) \leq \varphi(x)$ ; donc (135) (pris au sens ponctuel, et non de viscosité) revient à dire que

$$\lambda v(x) + \mathcal{H}(x, \nabla_x v(x)) \leq 0, \quad \text{avec égalité si } v(x) < \varphi(x).$$

Intuitivement, la première (seconde) relation exprime la sous optimalité (l'optimalité) de ne pas faire d'arrêt au point  $x$ .

9.6. **Commande impulsionnelle.** Dans cette modélisation la décision combine à une commande, la possibilité de sauts de l'état, de coût positif. Plus précisément, on se donne une suite strictement croissante, et tendant vers l'infini, de temps d'arrêt  $\theta_i > 0, i \geq 1$ ; on pose  $\theta_0 := 0$ . La dynamique est

$$(136) \quad \dot{\mathbf{y}}(t) = f(\mathbf{u}(t), \mathbf{y}(t)), \quad \theta_i < t < \theta_{i+1}, \quad \mathbf{y}(\theta_i^+) = \mathbf{y}(\theta_i^-) + \xi_i, \quad i \geq 1, \quad \mathbf{y}(0) = x.$$

La décision est donc  $(\mathbf{u}, \theta, \xi)$  avec  $\mathbf{u} \in \mathcal{U}_{ad}[0, \infty]$ , et la valeur associé est

$$(137) \quad V(x) = \inf_{(\mathbf{u}, \theta, \xi)} \left[ \int_0^\infty \ell(\mathbf{u}(t), \mathbf{y}(t)) e^{-\lambda t} dt + \sum_{i \geq 1} (c_0 + c(\xi)) e^{-\lambda \theta_i} \right].$$

On suppose encore  $\lambda > 0$  et  $f, \ell$  lipschitziennes et bornées, et de plus  $c_0 > 0$ , et  $c : \mathbb{R}^n \rightarrow \mathbb{R}_+$  telle que  $c(0) = 0$  et

$$c(\xi_1 + \xi_2) \leq c(\xi_1) + c(\xi_2), \quad \text{pour tout } \xi_1, \xi_2 \text{ dans } \mathbb{R}^n.$$

La relation ci-dessus (sachant que  $c_0 > 0$ ) fait qu'il vaut mieux faire un seul saut que deux immédiatement à la suite l'un de l'autre et est donc naturelle. Notons la valeur optimale obtenue après un saut à l'instant 0 par

$$\mathcal{M}[v](x) := c_0 + \inf_{\xi} (c(\xi) + v(x + \xi)).$$

On peut voir le problème de commande impulsionnelle comme un problème d'obstacle dans lequel la valeur d'arrêt serait  $\mathcal{M}[v](x)$ , ce qui motive l'énoncé suivant.

**Théorème 9.11.** *Si (131) est satisfait, la fonction valeur du problème de commande impulsionnelle est l'unique élément de  $BUC(\mathbb{R}^n)$ , solution de viscosité du problème*

$$(138) \quad \max(\lambda v(x) + \mathcal{H}(x, \nabla_x v(x)), v(x) - \mathcal{M}[v](x)) = 0, \quad \text{pour tout } x \in \mathbb{R}^n.$$

## 10. SCHÉMAS NUMÉRIQUES POUR L'ÉQUATION HJB

On suppose dans cette section les fonctions de coût et dynamique lipschitziennes et bornées. Nous commençons par la discrétisation du problème à horizon fini (121).

**10.1. Différences finies, état scalaire.** On se donne un nombre de pas de temps  $N$  et donc un pas de temps  $h_0 := T/N$ ; les instants de discrétisation sont les  $t_k := kh_0$ ,  $k = 0 \text{ à } N$ . Le pas d'espace est  $h_1 > 0$ , et les points de la grille spatiale sont les  $x_j := jh_1$ ,  $j \in \mathbb{Z}$ . On approxime  $V(x_j, t_k)$  par la quantité  $v_j^k$ , calculée de manière rétrograde; la condition finale naturelle est

$$v_j^N = \varphi(x_j), \quad j \in \mathbb{Z}.$$

Comme l'équation HJB (127) est l'expression différentielle du principe de programmation dynamique, qui nécessite d'évaluer la valeur au point obtenu par intégration du système dynamique, il est naturel d'approcher la dérivée en espace par décentrage à droite si  $f(u, x_j)$  est positif, et par décentrage à gauche sinon, ce qui donne le *schéma explicite* suivant, notant  $a_+ := \max(a, 0)$  et  $a_- := \min(a, 0)$ , pour  $k = N$  à 1 :

$$(139) \quad \frac{v_j^k - v_j^{k-1}}{h_0} + \min_{u \in U_{ad}} \left( \ell(u, x_j) + f(u, x_j)_+ \frac{v_{j+1}^k - v_j^k}{h_1} + f(u, x_j)_- \frac{v_j^k - v_{j-1}^k}{h_1} \right) = 0.$$

Réordonnant cette expression suivant les valeurs de  $v^k$ , il vient

$$(140) \quad v_j^{k-1} = \min_{u \in U_{ad}} \left( h_0 \ell(u, x_j) + (1 - h_0 h_1^{-1} |f(u, x_j)|) v_j^k + h_0 h_1^{-1} f(u, x_j)_+ v_{j+1}^k - h_0 h_1^{-1} f(u, x_j)_- v_{j-1}^k \right).$$

Si la condition suivante dite de *monotonie* est satisfaite

$$(141) \quad h_0 \|f\|_\infty \leq h_1,$$

les coefficients des composantes de  $v^k$  dans (140) sont tous positifs (et de somme un) et  $v^{k-1}$  est donc fonction monotone de  $\ell$  et  $v^k$ , donc par récurrence fonction monotone de  $(\ell, \varphi)$  : on parle alors d'un *schéma monotone*, et on vérifie facilement que  $\sup_j |v_j^{k-1}| \leq h_0 \|\ell\|_\infty + \sup_j |v_j^k|$ , d'où par récurrence l'estimation de stabilité en norme  $L^\infty$  :

$$(142) \quad \|v\|_\infty \leq T \|\ell\|_\infty + \|\varphi\|_\infty.$$

**10.2. État multidimensionnel.** Les quantités  $N, h_0, t_k$  sont comme précédemment. On note  $e_i$  le  $i$ ème élément de la base naturelle de  $\mathbb{R}^n$ , et  $h_i > 0$  le pas d'espace dans la direction  $e_i$ ,  $i = 1 \text{ à } n$ . À  $j \in \mathbb{Z}^n$  est associé le point de la grille  $x_j := \sum_{i=1}^n j_i h_i e_i$ . On approxime  $V(x_j, t_k)$  par la quantité  $v_j^k$ , avec condition finale

$$v_j^N = \varphi(x_j), \quad j \in \mathbb{Z}^n.$$

On reprend le principe du décentrage, composante par composante, ce qui donne le schéma suivant :

$$(143) \quad \frac{v_j^k - v_j^{k-1}}{h_0} + \min_{u \in U_{ad}} \left( \ell(u, x_j) + \sum_{i=1}^n \left( f_i(u, x_j)_+ \frac{v_{j+e_i}^k - v_j^k}{h_i} + f_i(u, x_j)_- \frac{v_j^k - v_{j-e_i}^k}{h_i} \right) \right) = 0.$$

Réordonnant cette expression suivant les valeurs de  $v^k$ , il vient

$$(144) \quad v_j^{k-1} = \min_{u \in U_{ad}} \left( h_0 \ell(u, x_j) + \left( 1 - h_0 \sum_{i=1}^n h_i^{-1} |f_i(u, x_j)| \right) v_j^k + h_0 \sum_{i=1}^n h_i^{-1} (f_i(u, x_j)_+ v_{j+e_i}^k - f_i(u, x_j)_- v_{j-e_i}^k) \right).$$

Le schéma est donc monotone si

$$h_0 \sum_{i=1}^n h_i^{-1} \|f_i\|_\infty \leq 1,$$

et dans ce cas on a encore l'estimation (142).

**10.3. Problèmes en horizon infini.** On reprend le problème (132), quand  $n = 1$ . Opérant comme précédemment, notant par  $v_j$  une approximation de  $V(x_j)$ , on obtient le schéma

$$(145) \quad \lambda v_j = \min_{u \in U_{ad}} \left( \ell(u, x_j) - h_1^{-1} |f(u, x_j)| v_j + h_1^{-1} f(u, x_j)_+ v_{j+1} - h_1^{-1} f(u, x_j)_- v_{j-1} \right).$$

Multipliant par  $h_0$ , ajoutant  $v_j$  de chaque côté et posant  $\beta := (1 + h_0 \lambda)^{-1}$ , on obtient la *forme de point fixe*

$$(146) \quad v_j = \beta \min_{u \in U_{ad}} \left( h_0 \ell(u, x_j) + (1 - h_0 h_1^{-1} |f(u, x_j)|) v_j + h_0 h_1^{-1} f(u, x_j)_+ v_{j+1} - h_0 h_1^{-1} f(u, x_j)_- v_{j-1} \right).$$

On suppose (141) satisfaite. Notons  $B(\mathbb{Z})$  l'espace des fonctions bornées sur  $\mathbb{Z}$ , muni de la norme de la convergence uniforme. On peut alors vérifier que le membre de droite ci-dessus définit un opérateur contractant (de facteur  $\beta < 1$ ) de  $B(\mathbb{Z})$  sur lui même.

Le schéma a donc une solution unique (qui ne dépend pas de  $h_0$ ) qu'on peut calculer en itérant cet opérateur de point fixe : c'est l'approche par *itération sur les valeurs*. On peut également résoudre (146) en alternant le calcul du minimum en  $u$ , pour  $v$  fixé, et la résolution du système linéaire à  $(u_j)$ ,  $j \in \mathbb{Z}$  fixé, soit

$$v_j = \beta \left( h_0 \ell(u_j, x_j) + (1 - h_0 h_1^{-1} |f(u_j, x_j)|) v_j + h_0 h_1^{-1} f(u_j, x_j)_+ v_{j+1} - h_0 h_1^{-1} f(u_j, x_j)_- v_{j-1} \right),$$

pour tout  $j \in \mathbb{Z}$ . C'est l'approche par *itération sur les stratégies*, ou *algorithme de Howard*. Cette seconde approche nécessite (après troncature du domaine de calcul) la résolution de grands systèmes linéaires.

**10.4. Problèmes de temps d'arrêt.** On reprend le problème (134), quand  $n = 1$ . Soient encore  $h_0 > 0$  et  $\beta := (1 + h_0 \lambda)^{-1}$ . L'équation HJB (135) peut se réécrire comme

$$\begin{aligned} v(x) &= \min(\beta(v(x) - h_0 \mathcal{H}(x, \nabla_x v(x, t))), \varphi(x)), \\ &= \min(\beta(v(x) + h_0 \inf_{u \in U_{ad}} H(u, x, \nabla_x v(x, t))), \varphi(x)). \end{aligned}$$

Discrétisant cette relation dans l'esprit du cas de l'horizon infini, on obtient le schéma (147)

$$v_j = \min \left( \beta \min_{u \in U_{ad}} \left( h_0 \ell(u, x_j) + (1 - \frac{h_0}{h_1} f(u, x_j)) v_j + \frac{h_0}{h_1} f(u, x_j)_+ v_{j+1} - \frac{h_0}{h_1} f(u, x_j)_- v_{j-1} \right), \varphi(x_j) \right).$$

Le membre de droite correspond encore à un point fixe contractant quand (141) est satisfait, et le schéma est donc bien posé.

**10.5. Algorithmes semi lagrangiens.** Introduisons la discrétisation temporelle suivante du principe de programmation dynamique (122) :

$$\tilde{v}^k(x) = \inf_{u \in U_{ad}} (h_0 \ell(u, x) + \tilde{v}^{k+1}(x + h_0 f(u, x))).$$

Ici  $N$ ,  $h_0$ , et les  $t_k$  sont comme précédemment,  $\tilde{v}^k : \mathbb{R}^n \rightarrow \mathbb{R}$  est une approximation de  $V(\cdot, t_k)$ ,  $k = 0$  à  $N$ , et  $\tilde{v}^N(x) = \varphi(x)$  pour tout  $x \in \mathbb{R}^n$ . On introduit ensuite comme dans l'approche par différences finies la grille des points  $x_j$ ,  $j \in \mathbb{Z}^n$ , et l'approximation  $v_j^k$  de  $V(x_j, t_k)$ , solution de

$$v_j^k = \inf_{u \in U_{ad}} (h_0 \ell(u, x_j) + \mathcal{I}^{k+1}(x_j + h_0 f(u, x_j))),$$

où  $\mathcal{I}^{k+1}$  est une formule d'interpolation des valeurs de  $v^{k+1}$  aux points de la grille. Si

$$\mathcal{I}^{k+1}(x) = \sum_{q \in \mathbb{Z}^n} \alpha_q(x) v_q^{k+1}$$

avec les  $\alpha_q(x)$  positifs et de somme un, le schéma est monotone et (142) est encore satisfait. On peut raffiner le schéma en introduisant des schémas d'intégration temporelle et d'interpolation spatiale d'ordre plus élevés, au risque de perdre la monotonie.

**10.6. Bilan sur les méthodes numériques pour l'équation HBJ.** La minimisation par rapport à  $u$  s'effectue habituellement en discrétisant les valeurs des commandes. Ces algorithmes convergent uniformément sur tout compact vers la fonction valeur. Cependant ils deviennent coûteux quand la dimension (nombre de variables d'état et de commandes) augmente. On les appliquera donc si possible à une version simplifiée du problème de commande optimale, dont la dimension est réduite.

## CONCLUSION / ORIENTATIONS BIBLIOGRAPHIQUES

On pourra consulter [10] pour les bases théoriques de l'optimisation, [9] pour une introduction à l'automatique et à la commande optimale, et [6], [7] pour les aspects algorithmique. Pour les applications : en mécanique spatiale (et en mécanique quantique dans le second cas) voir [11], [12], en médecine la partie IV de [18], en écologie [16], en chimie [7], et sur les bioprocédés, par exemple [4]. Mentionnons aussi l'ouvrage plus avancé [20], l'article [8] sur les conditions du second ordre, et [1] sur l'approche de tir. Concernant l'approche HJB, les livres de référence sont [2] et [3] pour la partie théorique, et [14] en ce qui concerne les méthodes numériques.

**Remerciements** L'auteur remercie en particulier S. Aronna (IMPA), X. Dupuis (Inria), A. Festa (Ensta), L. Pfeiffer (U. Graz), et F. Silva (U. Limoges), pour leurs lectures attentives d'une première version de cet article.

#### RÉFÉRENCES

- [1] M.S. Aronna, J.F. Bonnans, and P. Martinon. A shooting algorithm for optimal control problems with singular arcs. *J. Optim. Theory Appl.*, 158(2) :419–459, 2013.
- [2] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Birkhäuser, Boston, MA, 1997.
- [3] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*. Springer-Verlag, Paris, 1994.
- [4] T. Bayen, A. Rapaport, and M. Sebbah. Minimal time control of the two tanks gradostat model under a cascade inputs constraint. *SIAM J. Control Optim.*, 2014. In print.
- [5] R. Bellman. *Dynamic programming*. Princeton University Press, Princeton, N. J., 1957.
- [6] J.T. Betts. *Practical methods for optimal control and estimation using nonlinear programming*. SIAM, Philadelphia, PA, second edition, 2010.
- [7] L.T. Biegler. *Nonlinear programming*. SIAM, Philadelphia, PA, 2010.
- [8] J.F. Bonnans, X. Dupuis, and L. Pfeiffer. Second-order sufficient conditions for strong solutions to optimal control problems. *ESAIM-COCV*, 20(3) :704–724, 2013.
- [9] J.F. Bonnans and P. Rouchon. *Commande et optimisation de systèmes dynamiques*. Editions de l'Ecole Polytechnique, Palaiseau, 2005.
- [10] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.
- [11] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux*. Springer-Verlag, Berlin, 2006.
- [12] B. Bonnard and D. Sugny. *Optimal control with applications in space and quantum dynamics*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2012.
- [13] I. Ekeland. Nonconvex minimization problems. *Bull. Amer. Math. Soc. (N.S.)*, 1(3) :443–474, 1979.
- [14] M. Falcone and R. Ferretti. *Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations*. SIAM Publications, 2013.
- [15] A. T. Fuller. Study of an optimum non-linear control system. *J. Electronics Control (1)*, 15 :63–71, 1963.
- [16] B.S. Goh. *Management and analysis of biological populations*. Elsevier, 1980.
- [17] B.S. Goh, G. Leitmann, and T.L. Vincent. Optimal control of a prey-predator system. *Math. Biosci.*, 19 :263–286, 1974.
- [18] U. Ledzewicz, H. Schättler, A. Friedman, and E. Kashdan, editors. *Mathematical methods and models in biomedicine*. Springer, New York, 2013.
- [19] L.S. Pontryagin, V.G. Boltyanskiï, R.V. Gamkrelidze, and E.F. Mishchenko. *The mathematical theory of optimal processes*. Gordon & Breach Science Publishers, New York, 1986. Reprint of the 1962 English translation.
- [20] H. Schättler and U. Ledzewicz. *Geometric optimal control*. Springer, New York, 2012.

#### DANS LES TECHNIQUES DE L'INGÉNIEUR

- [21] BARATCHART (L.) ET BERNHARD (P.). - *Automatique et systèmes* [A 1370]. Base documentaire "Mathématiques pour l'ingénieur" (1992).
- [22] BORNE (P.) ET ROTELLA (F.). - *Commande optimale* [R 7427]. Base documentaire "Régulation et commande des systèmes asservis" (1996).
- [23] BREZINSKI (C.). - *Aspects numériques du contrôle linéaire* [AF 1400]. Base documentaire "Mathématiques pour l'ingénieur" (2007).
- [24] DACOROGNA (B.). - *Calcul des Variations* [AF 111]. Base documentaire "Mathématiques pour l'ingénieur" (2007).
- [25] LÉVINE (J.) ET ROUCHON (P.). *Systèmes dynamique et commande* [S 7430]. Base documentaire "Mathématiques pour l'ingénieur" (2008).
- [26] YVON (J.P.). - *Contrôle des systèmes à paramètres distribués* [AF 1372]. Base documentaire "Mathématiques pour l'ingénieur" (2009).

#### QUELQUES LOGICIELS LIBRES

Acado : <http://sourceforge.net/p/acado/wiki/Home/>  
Bocop : [www.bocop.org](http://www.bocop.org)  
Hampath : <http://apo.enseiht.fr/hampath>  
Psopt : <https://sites.google.com/a/psopt.org/psopt/>