

COURS DE CONTROLE OPTIMAL

MAM5A, Polytech Lyon, 2022-2023

Ionel Sorin CIUPERCA

Ce cours est enseigné à l'École d'ingénieurs Polytech Lyon, filière modélisation mathématique (MAM), 5^{ème} année.

Nous avons fait le choix d'une présentation orientée "sciences de l'ingénieurs" plutôt que "mathématiques". On insiste beaucoup sur les idées et les calculs à faire et beaucoup des résultats sont donnés sans démonstration.

Table des matières

1	Introduction et motivation	3
2	Contrôle optimal direct (ou à boucle ouvert) pour les systèmes d'équations différentielles ordinaires (EDO)	8
2.1	Position du problème de contrôle optimal	8
2.2	Quelques notions de minimisation d'une fonction sur un espace vectoriel . .	15
2.3	Le principe de minimum de Pontryagin pour le contrôle optimal des systèmes EDO	20
2.3.1	Le cas sans contraintes sur l'état final	20
2.3.2	Le cas avec contraintes sur l'état final	26
2.3.3	Un résultat de contrôlabilité	27
3	Contrôle optimal des équations aux dérivées partielles (EDP)	32
4	Méthodes numériques	39
4.1	Méthodes basées sur la résolution du système d'optimalité	39
4.2	Méthodes basées sur la minimisation numérique de la fonction coût	41
4.2.1	Rappel sur la notion de projection	41
4.2.2	Algorithmes de minimisation	42
4.2.3	Choix des facteurs ρ_k	43
5	Contrôle optimal en feed-back ou à boucle fermée	45
5.1	Quelques généralités	45
5.2	La fonction valeur	46
5.3	L'équation de Hamilton-Jacobi-Bellman	49
5.4	L'obtention du contrôle en feed-back	51

Chapitre 1

Introduction et motivation

Partout dans ce cours m, n, p désigne des nombres dans \mathbb{N}^* .

Ce cours est consacré à l'étude des systèmes commandés, c'est à dire des **systèmes dynamiques** sur lesquelles on peut agir au moyen d'une **commande** ou **contrôle**.

L'objectif principal dans ce cours est de déterminer un contrôle optimale, c'est à dire un contrôle minimisant un certain critère dépendant du contrôle et de la trajectoire résultant de ce contrôle. Il s'agit du problème de **contrôle optimale**.

Un autre objectif qui sera abordé ici est de voir si on peut amener le système d'un état initial donné à un état final fixé *a priori* (un cible) en respectant éventuellement certaines contraintes. Il s'agit du problème de **contrôlabilité**.

Le champ d'applications est très vaste. On rencontre des problèmes de contrôle optimale et de contrôlabilité dans des domaines très variés comme l'aéronautique, la mécanique en général, l'électronique, la médecine, économie et les finances , etc ...

De point de vue mathématique un système contrôlé est un système dynamique dont l'état est décrit par une fonction inconnue dite **fonction d'état** (ou **variable d'état**), qui vérifie une ou plusieurs lois d'évolution (très souvent ce sont des équations différentielles, mais d'autres lois peuvent être envisagées : équations intégrales, équations aux différences, lois d'évolutions stochastiques, etc ..). On supposera qu'on peut agir sur le système (en fait sur l'état du système) via une ou plusieurs fonctions qu'on appelle des **contrôles** (ou **commandes**).

Un autre type de problème qu'on peut rencontrer, mais qui revient au même sur le plan mathématique, est le fait que le système en lui même peut être mal connu, c'est à dire qu'il y a un ou plusieurs paramètres qui ne sont pas connus, ou qui sont inaccessibles ou difficile à mesurer directement. On cherche alors à déduire ces paramètres en les voyant comme des contrôles et en observant l'état du système ; c'est ce qu'on appelle un **problème d'identification des paramètres** ou **problème inverse**.

Exemples des problèmes de contrôle :

Exemple 1.1. *On considère un véhicule sur une route supposée droite avec une origine 0 et un sens positif. On suppose qu'au moment $t = 0$ le véhicule se trouve à une position $y_0 \in \mathbb{R}$ et il a une vitesse $v_0 \in \mathbb{R}$. On souhaite, grâce à une accélération u donnée au*

véhicule, faire en sorte qu'il s'arrête à un point $y_1 \in \mathbb{R}$ donné.

Ici le système est la véhicule et l'état du système est la fonction à valeurs vectorielles

$$t \in [0, +\infty[\rightarrow \begin{pmatrix} y(t) \\ v(t) \end{pmatrix} \in \mathbb{R}^2$$

où $y(t)$ et $v(t)$ sont respectivement la position et la vitesse au temps t .

Les variables d'état satisfont le système d'équations différentielles

$$\begin{cases} y' = v \\ v' = u \end{cases} \quad (1.1)$$

avec $u =$ l'accélération du véhicule.

Nous avons aussi les conditions initiales

$$\begin{cases} y(0) = y_0 \\ v(0) = v_0. \end{cases} \quad (1.2)$$

On peut voir ce système comme un **système contrôlé** où le contrôle est l'accélération u qui est une fonction du temps

$$t \in [0, +\infty[\rightarrow u(t) \in \mathbb{R}.$$

Il faut faire une hypothèse "raisonnable" sur le contrôle u ; on supposera par exemple que u est une fonction continue, ou continue par morceaux, ou $u \in L^2([0, +\infty[)$, etc ..

On peut aussi supposer que $|u|$ est bornée par une constante donnée $a > 0$, donc

$$|u(t)| \leq a, \quad \forall t \geq 0.$$

On mettra toutes les fonctions u satisfaisant ces conditions dans un ensemble "admissible" que nous notons U_{ad} .

Il faut maintenant introduire un **critère d'optimisation** et il y a plusieurs possibilités qui sont utilisées habituellement dans la pratique :

Possibilité 1 : Trouver "le meilleur" u parmi tous les $u \in U_{ad}$ tel que le véhicule s'arrête au point $y_1 \in \mathbb{R}$ (donné) dans **un temps minimum**. Le problème peut alors s'écrire :

$$\min \{ \tau > 0, \quad y(\tau) = y_1 \text{ et } v(\tau) = 0, \quad (y, v, u) \text{ satisfont (1.1) - (1.2), avec } u \in U_{ad} \}.$$

On peut aussi "relaxer" le problème et considérer pour un $\epsilon > 0$ assez petit donné le problème

$$\min \{ \tau > 0, \quad (y(\tau), v(\tau)) \in [y_1 - \epsilon, y_1 + \epsilon] \times \{0\}, \quad (y, v, u) \text{ satisfont (1.1) - (1.2), } u \in U_{ad} \}.$$

On se contente donc d'amener (y, v) "proche" de $(y_1, 0)$ dans un temps minimum; c'est un problème moins contraignant.

Possibilité 2 : Amener le véhicule à l'arrêt au point y_1 en ayant un coût du déplacement le plus petit possible, ceci dans un temps donné $T > 0$. On introduit alors une fonctionnelle

$E : U_{ad} \rightarrow \mathbb{R}$ qui donne le coût du déplacement en fonction de l'accélération ; l'exemple le plus utilisé est

$$E(u) = \int_0^T u^2(t) dt, \quad \forall u \in U_{ad}$$

(pour cela il faut choisir U_{ad} de telle manière que l'intégrale ci-dessus existe pour tout $u \in U_{ad}$).

Alors le problème de contrôle optimal sera

$$\min \{E(u), (y(T), v(T)) = (y_1, 0), (y, v, u) \text{ satisfont (1.1) - (1.2), avec } u \in U_{ad}\}.$$

ou encore le problème "relaxé" :

$$\min \{E(u), (y(T), v(T)) \in [y_1 - \epsilon, y_1 + \epsilon] \times \{0\}, (y, v, u) \text{ satisfont (1.1) - (1.2), } u \in U_{ad}\}.$$

Il y a une autre manière plus convenable mathématiquement de définir un tel problème de contrôle. Pour $\epsilon > 0$ assez petit on considère la **fonction pénalisée** $E_\epsilon : U_{ad} \rightarrow \mathbb{R}$ définie par

$$E_\epsilon(u) = E(u) + \frac{1}{\epsilon}|y(T) - y_1|^2 + \frac{1}{\epsilon}|v(T)|^2$$

et le problème sera

$$\min \{E_\epsilon(u), (y, v, u) \text{ satisfont (1.1) - (1.2), avec } u \in U_{ad}\}.$$

L'idée de cette méthode est de pénaliser "l'éloignement" de $y(T)$ par rapport à y_1 et celui de $v(T)$ par rapport à 0. Ce problème est bien plus facile à manier mathématiquement, car il n'y a pas des contraintes explicites sur l'état final.

Exemple 1.2. On considère l'équation de la chaleur d'évolution : on se donne $\Omega \subset \mathbb{R}^n$ un ouvert ($n \in \{1, 2, 3\}$) et $T > 0$ et on cherche $y : \Omega \times [0, T] \rightarrow \mathbb{R}$ ou $y(x, t)$ représente la température, solution de l'EDP suivant :

$$\left\{ \begin{array}{ll} \frac{\partial y}{\partial t} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(a \frac{\partial y}{\partial x_i} \right) = f & \text{pour } (x, t) \in \Omega \times [0, T] \\ a \frac{\partial y}{\partial \nu} = g & \text{pour } (x, t) \in \partial\Omega \times [0, T] \\ y(x, 0) = y_0(x) & \text{pour } x \in \Omega. \end{array} \right. \quad (1.3)$$

Dans ce système les donnée sont les fonctions a, g et f , avec : la fonction $a : \Omega \rightarrow]0, +\infty[$ est la conductivité thermique, $g : \partial\Omega \times [0, T] \rightarrow \mathbb{R}$ est le flux rentrant de chaleur sur la frontière $\partial\Omega$ et $f : \Omega \times [0, T] \rightarrow \mathbb{R}$ est la source de chaleur. Le terme $\frac{\partial y}{\partial \nu}$ est égal à $\nabla y \cdot \nu$ où ν est le vecteur normal à la frontière $\partial\Omega$, extérieur à Ω .

Remarque : Le même système d'équations peut être utilisé pour modéliser la concentration

d'une substance dans un mélange (dans ce cas l'inconnue y étant la concentration), ou l'écoulement d'un fluide dans un milieu poreux ($y =$ la pression) ou le potentiel créé par une charge électrique ($y =$ le potentiel).

On peut poser sur le système d'équations (1.3) plusieurs problèmes de contrôle optimal :

1. On peut se proposer d'amener la température $y(x, t)$ le plus proche que possible d'une température donnée $y_d(x)$ au moment $t = T$ avec une consommation la plus petite possible d'énergie, en utilisant comme contrôle u :

- soit le terme source de chaleur, donc $u = f$ (a et g sont données)

- soit le flux rentrant de chaleur, donc $u = g$ (a et f sont données)

En considérant l'expression suivante pour la consommation d'énergie E :

- soit $E(f) = \int_0^T \int_{\Omega} f^2(x, t) dx dt$ si le contrôle est f

- soit $E(g) = \int_0^T \int_{\partial\Omega} g^2(x, t) d\sigma dt$ si le contrôle est g

le problème de contrôle optimal sera

$$\min \left\{ \int_0^T \int_{\Omega} f^2(x, t) dx dt + \frac{1}{\epsilon} \int_{\Omega} [y(x, T) - y_d]^2 dx, \quad (y, f) \text{ satisfont (1.3)}, f \in U_{ad} \right\}$$

respectivement

$$\min \left\{ \int_0^T \int_{\partial\Omega} g^2(x, t) d\sigma dt + \frac{1}{\epsilon} \int_{\Omega} [y(x, T) - y_d]^2 dx, \quad (y, g) \text{ satisfont (1.3)}, g \in U_{ad} \right\}$$

ou l'ensemble U_{ad} est à choisir convenablement dans chaque cas.

2. Un autre problème qu'on peut se poser ici est un **problème inverse**. On suppose que la conductivité $a(x)$ n'est pas bien connue et que f et g sont connues; on verra alors la fonction a comme le contrôle, donc $u = a$. Le but sera alors de trouver $a(x)$ en mesurant la température $y(x, t)$ dans des points $x \in \partial\Omega$ et à des moments $t \in [0, T]$ choisis convenablement.

Supposons par exemple qu'on peut mesurer $y(x, t)$ pour tout $x \in \partial\Omega$ et tout $t \in [0, T]$ et qu'on obtient une fonction $y_m : \partial\Omega \times [0, T] \rightarrow \mathbb{R}$ (fonction donnée). En utilisant la méthode des moindres carrées, le problème de contrôle optimal sera alors

$$\min \left\{ \int_0^T \int_{\partial\Omega} [y(x, t) - y_m(x, t)]^2 d\sigma dt \quad (y, a) \text{ satisfont (1.3)}, a \in U_{ad} \right\}$$

avec encore U_{ad} un ensemble des fonction a à choisir convenablement.

Ce problème veut dire : on cherche a telle que y s'éloigne le moins possible de la fonction mesurée y_m .

Il y a encore d'autres possibilités de définir un problème de contrôle optimal dans ce cadre : il est plus réaliste de supposer que y n'est mesuré que à un nombre fini d'instants $\tau_1, \tau_2, \dots, \tau_k$. Dans ce cas dans le problème de contrôle considéré il faut remplacer $\int_0^T \int_{\partial\Omega} [y(x, t) - y_m(x, t)]^2 d\sigma dt$ par $\sum_{i=1}^k \int_{\partial\Omega} [y(x, \tau_i) - y_m(x, \tau_i)]^2 d\sigma$.

Il est encore plus réaliste de supposer que y est mesuré aussi dans un nombre fini de points $x_1, x_2, \dots, x_m \in \partial\Omega$. Dans ce cas il faut remplacer le terme $\int_0^T \int_{\partial\Omega} [y(x, t) - y_m(x, t)]^2 d\sigma dt$ par $\sum_{i=1}^k \sum_{j=1}^m [y(x_j, \tau_i) - y_m(x_j, \tau_i)]^2$.

Les problèmes considérés dans ces exemples entrent dans la catégorie des **problèmes de contrôle direct** ou **à boucle ouverte**, où le contrôle est déterminé comme une fonction du temps.

Il y a encore une catégorie importante des problèmes de contrôle appelée **problèmes de contrôle en feed-back** ou **à boucle fermée**. Dans ce cas le contrôle u est déterminé comme une fonction de temps mais aussi de l'état y . Autrement dit, l'état du système est pris en compte à chaque instant afin de déterminer "en temps réel" la commande ou le contrôle. L'étude de ce type de contrôle fera l'objet du dernier chapitre du cours.

Chapitre 2

Contrôle optimal direct (ou à boucle ouvert) pour les systèmes d'équations différentielles ordinaires (EDO)

Notations : Partout dans ce cours, pour une fonction $f : \Omega \rightarrow \mathbb{R}^m$ avec Ω sous-ensemble ouvert de \mathbb{R}^n , pour tout $x \in \Omega$ nous notons (quand ça existe) par $J_f(x)$ la **matrice Jacobienne** de f en x . Rappelons que si $f = (f_1, f_2, \dots, f_m)^T$ alors

$$J_f(x) \in \mathcal{M}_{m,n}(\mathbb{R}) \quad \text{et} \quad (J_f(x))_{ij} = \frac{\partial f_i}{\partial x_j}(x), \quad \forall i = 1, 2, \dots, m, j = 1, 2, \dots, n.$$

(la ligne i de la matrice $J_f(x)$ contient les dérivées partielles de f_i en x_1, x_2, \dots, x_n). Nous notons par $\nabla f(x) \in \mathcal{M}_{n,m}(\mathbb{R})$ la transposée de la matrice Jacobienne $J_f(x)$; On peut écrire $\nabla f(x) = (J_f(x))^T$ (la colonne i de la matrice $\nabla f(x)$ contient les dérivées partielles de f_i en x).

On peut alors écrire $J_f(x) = (\nabla f(x))^T = \nabla^T f(x)$.

Remarque : si f est à valeurs scalaires (donc $m = 1$) alors $\nabla f(x)$ s'identifie avec le vecteur colonne des dérivées partielles de f en x .

2.1 Position du problème de contrôle optimal

On se donne deux réels $t_0, t_1 \in \mathbb{R}$ avec $t_0 < t_1$, une fonction $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ et un vecteur $x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ \cdot \\ \cdot \\ x_n^0 \end{pmatrix} \in \mathbb{R}^n$.

On considère le système d'équations différentielles ordinaires (EDO) suivant : pour une

fonction $u : [t_0, t_1] \rightarrow \mathbb{R}^m$ donnée avec $u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ u_m(t) \end{pmatrix}$ (avec certaines propriétés qui

seront précisées plus loin) trouver $x : [t_0, t_1] \rightarrow \mathbb{R}^n$ avec $x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \cdot \\ x_n(t) \end{pmatrix}$ satisfaisant le

problème de Cauchy

$$\begin{cases} x'(t) = f(t, x(t), u(t)), & t \in [t_0, t_1] \\ x(t_0) = x^0. \end{cases} \quad (2.1)$$

Le système (2.1) est appelé **système contrôlé**

$x(t)$ s'appelle **l'état** du système (2.1)

$u(t)$ s'appelle **le contrôle** du système (2.1).

On se donne aussi U_{ad} un ensemble des fonctions *admissibles* pour le contrôle u (on verra plus tard un exemple de U_{ad}) et aussi un ensemble **cible** pour $x(t_1)$, c'est à dire on se donne un ensemble $C_1 \subset \mathbb{R}^n$ et on demandera $x(t_1) \in C_1$.

Remarque 2.1. *Il y a dans la théorie de contrôle la notion de **contrôlabilité** ; donnons pour l'instant juste quelques définitions :*

- *Si on peut trouver un temps $\tau \in]t_0, t_1]$ et une fonction $u \in U_{ad}$ tels que $x(\tau) \in C_1$ alors on dira que le système (2.1) est **contrôlable pour la cible** C_1 ; dans le cas particulier $C_1 = \{0\}$ (donc $x(\tau) = 0$) on dira simplement que (2.1) est **contrôlable**.*
- *Pour un $T \in]t_0, t_1]$ fixé si on peut trouver $u \in U_{ad}$ tel que $x(T) = 0$ alors on dira que le système (2.1) est **contrôlable en temps** T .*

Revenons à notre problème de contrôle optimal ; on va introduire une fonction S dépendant de x et de u appelée **fonction coût**, qui sera de la forme

$$S(x, u) = \int_{t_0}^{t_1} L(t, x(t), u(t)) dt + \varphi(x(t_1)) \quad (2.2)$$

avec

$$L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$$

et

$$\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$$

des fonctions données.

Le problème de contrôle optimal consiste à trouver parmi toutes les paires de fonctions (x, u) satisfaisant le système (2.1) avec $u \in U_{ad}$ et $x(t_1) \in C_1$ celles qui minimisent la

fonction S donnée par (2.2).

On peut alors écrire le problème sous la forme

$$\min \left\{ \int_{t_0}^{t_1} L(t, x(t), u(t)) dt + \varphi(x(t_1)) \mid x' = f(t, x, u), x(t_0) = x^0, u \in U_{ad}, x(t_1) \in C_1 \right\} \quad (2.3)$$

On peut voir ceci comme un problème de minimisation avec contraintes :

$$\min_{(x,u) \in CTR} S(x, u)$$

avec $CTR =$ un ensemble des contraintes défini par

$$CTR = \{(x, u), x' = f(t, x, u), x(t_0) = x^0, u \in U_{ad}, x(t_1) \in C_1\}.$$

Le problème de contrôle optimal (2.3) est appelé problème en **formulation de Bolza**.

Exemple 2.1. *Rappelons l'Exemple 1.1 avec le contrôle d'un véhicule. L'état de ce système est (y, v) satisfaisant le problème de Cauchy*

$$\begin{cases} y' = v \\ v' = u \\ y(0) = y_0 \\ v(0) = v_0 \end{cases} \quad (2.4)$$

Le premier problème de contrôle considéré était

$$\min \left\{ \int_0^T u^2(t) dt \mid (y, v, u) \text{ satisfait (2.4), } u \in U_{ad}, (y, v)(T) = (y_1, 0) \right\}$$

avec $T > 0, y_0, v_0, y_1 \in \mathbb{R}$ donnés.

C'est un problème du type (2.3) avec

$$t_0 = 0$$

$$t_1 = T$$

$$n = 2$$

$$m = 1$$

$$x = \begin{pmatrix} y \\ v \end{pmatrix}$$

$$x^0 = \begin{pmatrix} y_0 \\ v_0 \end{pmatrix} \in \mathbb{R}^2$$

$f : [0, T] \times \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}^2$ donnée par

$$f(t, (\tilde{x}_1, \tilde{x}_2), \tilde{u}) = \begin{pmatrix} \tilde{x}_2 \\ \tilde{u} \end{pmatrix}, \quad \forall (t, (\tilde{x}_1, \tilde{x}_2), \tilde{u}) \in [0, T] \times \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}^2.$$

$$\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \varphi(z) = 0$$

$$L : [0, T] \times \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}, \quad L(\tilde{t}, \tilde{x}, \tilde{u}) = \tilde{u}^2$$

$$C_1 = \left\{ \begin{pmatrix} y_1 \\ 0 \end{pmatrix} \right\}.$$

On peut aussi considérer un autre problème de contrôle optimal :

$$\min \left\{ \int_0^T u^2(t) dt + \frac{1}{\epsilon} [y(T) - y_1]^2 + \frac{1}{\epsilon} v^2(T) \mid (y, v, u) \text{ satisfait (2.4), } u \in U_{ad} \right\}$$

avec $\epsilon > 0$ assez petit donné. C'est un problème pénalisé du problème précédent et c'est encore du type (2.3) avec les changements suivants :

- prendre $C_1 = \mathbb{R}^2$ (pas de contraintes sur l'état final)
- prendre $\varphi(z) = \frac{1}{\epsilon}(z_1 - y_1)^2 + \frac{1}{\epsilon}z_2^2, \quad \forall z = (z_1, z_2) \in \mathbb{R}^2$.

Remarque 2.2. Dans cet exemple la fonction f est linéaire par rapport aux variables $((\tilde{x}_1, \tilde{x}_2), \tilde{u})$; elle peut s'écrire

$$f((\tilde{x}_1, \tilde{x}_2), \tilde{u}) = A \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} + B\tilde{u}$$

avec

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}) \quad \text{et} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R}).$$

Quelques éléments de vocabulaire :

1. Si $m = 1$ alors le système contrôlé (2.1) est dit **à simple commande** (c'est le cas dans l'Exemple 2.1).
2. Si $m \geq 2$ alors le système est dit **à commande multiple**
3. Si φ est la fonction constante nulle alors on dit que le système est **à observation distribuée** (c'est le cas pour le premier problème de l'Exemple 2.1).
4. Si φ n'est pas la fonction constante nulle alors on dit que le système est **à observation finale** (c'est le cas pour le deuxième problème de l'Exemple 2.1).
5. Si $C_1 = \mathbb{R}^n$ on dit que le système est **sans contrainte sur l'état final** (c'est le cas pour le deuxième problème de l'Exemple 2.1).
6. Si $C_1 \neq \mathbb{R}^n$ on dit que le système est **avec contrainte sur l'état final** (c'est le cas pour le premier problème de l'Exemple 2.1).

Remarque 2.3. Le cas $C_1 = \mathbb{R}^n$ (donc sans contrainte sur l'état final) est beaucoup plus facile de point de vue mathématique que les cas contraire $C_1 \neq \mathbb{R}^n$ (donc avec contrainte sur l'état final).

C'est pourquoi, on cherche très souvent à remplacer un problème avec $C_1 \neq \mathbb{R}^n$ par un problème "approché" avec $C_1 = \mathbb{R}^n$ en ajoutant à la fonction coût un terme de pénalisation

du type $\frac{1}{\epsilon} \text{dist}^2(x(t_1), C_1)$ avec $\epsilon > 0$ très petit et “dist” une “distance” (à définir) entre un point et un ensemble. Par exemple le passage du premier au deuxième problème dans l'Exemple 2.1 a été fait en remplaçant $C_1 = \{(y_1, 0)\}$ par $C_1 = \mathbb{R}^2$ et en ajoutant à la fonction cout le terme de pénalisation $\frac{1}{\epsilon}[y(T) - y_1]^2 + \frac{1}{\epsilon}v^2(T)$ qui n'est autre que la distance au carré entre le point $(y(T), v(T))$ et le singleton $\{(y_1, 0)\}$, multipliée par $1/\epsilon$.

Remarque 2.4. Un cas particulier important et qui a été beaucoup étudié est le cas où la fonction f est linéaire par rapport aux variables (x, u) (on dira dans ce cas que le système contrôlé (2.1) est **linéaire**).

Pour donner une forme générale de f dans ce cas, on définit deux fonctions à valeurs matricielles

$$A : [t_0, t_1] \rightarrow \mathcal{M}_n(\mathbb{R}), \quad A(t) = (A_{ij}(t))_{i,j=1\dots n}$$

et

$$B : [t_0, t_1] \rightarrow \mathcal{M}_{n,m}(\mathbb{R}), \quad B(t) = (B_{ij}(t))_{i=1\dots n, j=1,\dots,m}$$

et une fonction à valeurs vectorielles

$$g : [t_0, t_1] \rightarrow \mathbb{R}^n, \quad g(t) = (g_1(t), g_2(t), \dots, g_n(t))^T$$

Alors f sera sous la forme $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ avec

$$f(t, \tilde{x}, \tilde{u}) = A(t)\tilde{x} + B(t)\tilde{u} + g(t), \quad \forall (t, \tilde{x}, \tilde{u}) \in [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m.$$

Le système contrôlé (2.1) s'écrit alors

$$\begin{cases} x'(t) = A(t)x(t) + B(t)u(t) + g(t), & t \in [t_0, t_1] \\ x(t_0) = x^0 \end{cases} \quad (2.5)$$

On supposera que A et B sont des fonctions continues, ce qui veut dire que chacune des composantes de A et B sont des fonctions continues de $[t_0, t_1]$ dans \mathbb{R} .

Un cas particulier souvent considéré sera le cas où A et B sont des fonctions matricielles constantes, notées encore A et respectivement B .

Le choix de U_{ad}

Rappels espaces de Lebesgue :

Si Ω est un ensemble ouvert et non vide de \mathbb{R}^p et $r \in [1, +\infty[$ on note

$$L^r(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R}, \quad \int_{\Omega} |f(x)|^r dx < +\infty \right\}$$

(avec la remarque que f doit être en plus mesurable et qu'on confond toujours deux fonctions égale presque partout).

On peut considérer la généralisation suivante : pour tout $q \in \mathbb{N}^*$

$$L^r(\Omega; \mathbb{R}^q) = \left\{ f : \Omega \rightarrow \mathbb{R}^q, \quad \int_{\Omega} \|f(x)\|^r dx < +\infty \right\}$$

ou $\|\cdot\|$ désigne une norme quelconque (par exemple la norme Euclidienne) dans \mathbb{R}^q .
Remarquons qu'une définition équivalente pour ce dernier espace est

$$L^r(\Omega; \mathbb{R}^q) = \{f = (f_1, f_2, \dots, f_q)^T, \quad f_i \in L^r(\Omega), \quad \forall i = 1, 2, \dots, q\}$$

Rappelons que $L^r(\Omega; \mathbb{R}^q)$ est un **espace de Banach** avec la norme

$$f \in L^r(\Omega; \mathbb{R}^q) \rightarrow \|f\|_{L^r(\Omega; \mathbb{R}^q)} = \left(\int_{\Omega} \|f(x)\|^r dx \right)^{1/r}.$$

Rappelons aussi que dans le particulier $r = 2$ l'espace $L^2(\Omega; \mathbb{R}^q)$ est un **espace de Hilbert** avec produit scalaire

$$\langle f | g \rangle_{L^2(\Omega; \mathbb{R}^q)} = \int_{\Omega} \langle f(x) | g(x) \rangle_{\mathbb{R}^q} dx = \sum_{i=1}^n \int_{\Omega} f_i(x) g_i(x) dx$$

Nous choisissons comme ensemble des contrôles admissibles

$$U_{ad} = L^2(]t_0, t_1[; \mathbb{R}^m) \quad \text{ou} \quad U_{ad} = \text{un sous-ensemble approprié de } L^2(]t_0, t_1[; \mathbb{R}^m).$$

On montrera dans la suite qu'il y a un cas où le problème de contrôle considéré peut s'écrire de manière plus simple comme un problème de minimisation sur un sous-ensemble de l'espace de Hilbert $L^2(]t_0, t_1[; \mathbb{R}^m)$.

On supposera dans ce chapitre que la fonction f est telle que pour toute fonction u donnée avec $u \in L^2(]t_0, t_1[; \mathbb{R}^m)$ on a l'existence et l'unicité d'une solution $x(t)$ du problème de Cauchy (2.1).

Remarque 2.5. *Il nous reste à préciser clairement qu'est-ce que ça veut dire que x est solution du problème de Cauchy (2.1).*

Dans le cas où les fonctions f et u sont continues alors on peut appliquer le cadre classique : une solution de (2.1) est une fonction x de classe C^1 sur $[t_0, t_1]$ et qui satisfait en même temps

- la première égalité de (2.1) pour tous $t \in [t_0, t_1]$
- la deuxième égalité de (2.1).

Dans le cas où au moins l'une des fonctions f ou u n'est pas continue alors le cadre classique ne s'applique pas. Une possibilité pour donner un sens au fait que x est une solution de (2.1) est de dire que x est continue, qu'elle satisfait la deuxième égalité de (2.1) et que la première égalité de (2.1) est satisfaite au sens des distributions $\mathcal{D}'(]t_0, t_1[)$.

Nous aurons très souvent le cas suivant : la fonction u ainsi que la fonction $t \mapsto f(t, x(t), u(t))$ sont des fonctions continues par morceaux sur $[t_0, t_1]$. Alors sur chaque "morceau ouvert" la première égalité de (2.1) doit être satisfaite au sens classique, alors que dans tous les points de discontinuité de u ou de $t \mapsto f(t, x(t), u(t))$ la solution x doit être continue (la solution x sera alors continue et de classe C^1 par morceaux). On admet que dans ce cas l'équation principale de (2.1) est satisfaite au sens des distributions $\mathcal{D}'(]t_0, t_1[)$.

Un autre cadre mathématique qui est utilisé est de chercher la solution x comme une fonction **absolument continue** sur $[t_0, t_1]$. Alors elle sera continue sur $[t_0, t_1]$ et dérivable pour presque tout $t \in [t_0, t_1]$. On demandera alors que la deuxième égalité de (2.1) soit satisfaite et que la première égalité de (2.1) soit satisfaite pour presque tout $t \in [t_0, t_1]$. Une telle solution s'appelle **solution au sens de Carathéodory** de (2.1).

Rappel :

Il est utile de rappeler la solution d'un système EDO linéaire avec condition initiale. Considérons donc le problème suivant : trouver $y : I \rightarrow \mathbb{R}^n$ solution du système

$$\begin{cases} y'(t) = A(t)y(t) + h(t), & \forall t \in I \\ y(\tau_0) = y^0 \end{cases} \quad (2.6)$$

avec $I \subset \mathbb{R}$ un intervalle, $A(t) \in \mathcal{M}_n(\mathbb{R})$, $h(t) \in \mathbb{R}^n$, $\tau_0 \in I$ et $y^0 \in \mathbb{R}^n$ données. (Remarquons que pour tout contrôle u fixé le système linéaire (2.5) est de la forme (2.6) avec $h = Bu + g$).

1. Dans le cas particulier où A est une matrice constante, la solution de (2.6) est donnée par la **formule de Duhamel** :

$$y(t) = e^{A(t-\tau_0)}y^0 + \int_{\tau_0}^t e^{A(t-s)}h(s) ds, \quad \forall t \in I.$$

(Rappelons que pour toute matrice carrée réelle $M \in \mathcal{M}_n(\mathbb{R})$ l'exponentielle e^M de M se définit comme la matrice de même taille donnée par

$$e^M = \sum_{j=0}^{\infty} \frac{1}{j!} M^j. \quad)$$

2. Dans le cas particulier $n = 1$ (même avec A dépendant du temps) la solution de (2.6) est donnée par

$$y(t) = \exp\left(\int_{\tau_0}^t A(\tau) d\tau\right)y^0 + \int_{\tau_0}^t \exp\left(\int_s^t A(\tau) d\tau\right)h(s) ds, \quad \forall t \in I.$$

3. Dans le cas $n \geq 2$ et A est dépendante du temps il n'y a pas de formule générale explicite pour la solution de (2.6) (on sait seulement qu'une telle solution existe et est unique, par exemple dans le cas où A et h sont continues).

Rappelons donc qu'on suppose que à tout contrôle u on peut faire correspondre un état unique x du système, solution de (2.1). Pour mettre en évidence la dépendance de l'état du contrôle on notera souvent la solution x de (2.1) par x^u .

On observe alors que dans le cas $C_1 = \mathbb{R}^n$ le problème de contrôle (2.3) est équivalent au problème de minimisation plus simple :

$$\min \{J(u), \quad u \in U_{ad}\} \quad (2.7)$$

où $J : L^2([t_0, t_1]; \mathbb{R}^m) \rightarrow \mathbb{R}$ est une fonction donnée par

$$J(u) = \int_{t_0}^{t_1} L(t, x^u(t), u(t)) dt + \varphi(x^u(t_1)), \quad \forall u \in U_{ad}. \quad (2.8)$$

Remarque 2.6. Cette méthode qui consiste à réduire le problème (2.3) au problème plus simple (2.7) ne peut pas s'appliquer en général dans le cas $C_1 \neq \mathbb{R}^n$. Ceci est dû au fait que pour un u arbitraire dans U_{ad} il n'est pas sûr que x^u satisfait aussi la contrainte $x^u(t_1) \in C_1$ sur l'état final.

2.2 Quelques notions de minimisation d'une fonction sur un espace vectoriel

Dans ce paragraphe X désigne un **espace vectoriel** sur \mathbb{R} (noter *e.v.*). Rappelons que si on muni X d'une **norme** $\|\cdot\|$ alors X devient un **espace vectoriel normé** (noter *e.v.n.*). Rappelons que dans un *e.v.n.* X nous avons les notions suivantes, qui généralisent des notions connues dans l'espace euclidien \mathbb{R}^n :

1. **convergence** des suites ($x_k \rightarrow x$ en X veut dire : $\|x_k - x\| \rightarrow 0$ en \mathbb{R} , pour $k \rightarrow +\infty$).
2. **boule ouverte** en X centré en $a \in X$ et de rayon $r > 0$ notée $B_X(a, r)$ (ou simplement $B(a, r)$) ; elle est définie par

$$B_X(a, r) = \{y \in X, \quad \|y - a\| < r\}.$$

3. ensemble ouvert : $\Omega \subset X$ est dit **ouvert** si pour tout $a \in \Omega$ il existe $r > 0$ tel que $B_X(a, r) \subset \Omega$.
4. ensemble fermé : $F \subset X$ est dit **fermé** si pour toute suite $(x_k) \subset F$ tel que $x_k \rightarrow x$ en X on a $x \in F$.

Rappelons qu'un *e.v.n.* qui est complet (c'est à dire : toute suite de Cauchy est convergent) s'appelle **espace de Banach**.

Enfin, rappelons qu'un espace de Banach X dont la norme provient d'un **produit scalaire** $\langle \cdot | \cdot \rangle$ sur X (c'est à dire : $\|x\| = \sqrt{\langle x | x \rangle}$, $\forall x \in X$) s'appelle **espace de Hilbert**.

Rappel : Un exemple élémentaire d'espace de Hilbert est l'espace Euclidien \mathbb{R}^n muni du produit scalaire habituel entre deux vecteurs de \mathbb{R}^n :

$$\langle x | y \rangle = \sum_{i=1}^n x_i y_i.$$

Pour un ensemble A non vide et pour une fonction $g : A \rightarrow \mathbb{R}$ nous considérons le problème de minimisation sur A :

$$\min \{g(y), \quad y \in A\} \quad (2.9)$$

c'est à dire, trouver $y^* \in A$ tel que

$$g(y^*) \leq g(y), \quad \forall y \in A. \quad (2.10)$$

Rappelons quelques notions liées à la convexité et à la coercivité :

Définition 2.1. Soit X un e.v. et D un sous ensemble non vide de X .

1. On dit que D est **convexe** si $\forall x, y \in D$ et $\forall \lambda \in [0, 1]$ on a $\lambda x + (1 - \lambda)y \in D$.
2. Soit $D \subset X$ un ensemble convexe et $h : D \rightarrow \mathbb{R}$ une fonction.
On dit que h est **une fonction convexe** sur D si

$$h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y), \quad \forall x, y \in D, \quad \forall \lambda \in [0, 1].$$

On dit que h est **une fonction strictement convexe** sur D si

$$h(\lambda x + (1 - \lambda)y) < \lambda h(x) + (1 - \lambda)h(y), \quad \forall x, y \in D \text{ avec } x \neq y, \quad \forall \lambda \in]0, 1[.$$

Définition 2.2. Soit X un e.v.n. $D \subset X$ un sous ensemble non vide et supposons que D est non bornée (c'est à dire : il existe au moins une suite $(x_k)_{k \in \mathbb{N}} \subset D$ tel que $\|x_k\|_H \rightarrow +\infty$ pour $k \rightarrow +\infty$).

Soit $h : X \rightarrow \mathbb{R}$ une fonction. On dit que h est **coercive** sur D si pour toute suite $(x_k)_{k \in \mathbb{N}} \subset D$ tel que $\|x_k\|_H \rightarrow +\infty$ pour $k \rightarrow +\infty$ on a $h(x_k) \rightarrow +\infty$ pour $k \rightarrow +\infty$.

Rappelons les résultats suivants connus pour $X = \mathbb{R}^p$:

Proposition 2.1. Soit $X = \mathbb{R}^p$, $g : \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction et $A \subset \mathbb{R}^p$ un ensemble **fermé** et **convexe**. Nous avons

1. Supposons que g est continue sur A avec en plus g coercive dans le cas où A est non bornée. Alors il existe au moins une solution du problème de minimisation (2.9).
2. Si en plus g est strictement convexe alors on a aussi l'unicité de la solution de (2.9).
3. Si g est différentiable sur \mathbb{R}^p (sur un voisinage ouvert de A c'est suffisant) et si $y^* \in A$ est une solution de (2.9) (c'est à dire, y^* satisfait (2.10)), alors on a

$$\langle \nabla g(y^*) | y - y^* \rangle \geq 0, \quad \forall y \in A. \quad (2.11)$$

Cette inégalité s'appelle **inéquation d'Euler** et elle donne une **condition nécessaire d'optimalité de premier ordre**.

Remarque 2.7. Si on sait que y^* est dans l'intérieur de A (ce qui est toujours vrai par exemple dans le cas particulier $A = \mathbb{R}^p$), alors (2.11) est équivalente à l'égalité suivante (appelée **l'équation d'Euler**) :

$$\nabla g(y^*) = 0.$$

Sur un espace vectoriel normé (e.v.n.) général ?

Définition 2.3. Soit X un e.v.n., Ω un sous-ensemble ouvert et non vide de X , $h : \Omega \rightarrow \mathbb{R}$ une fonction, $a \in \Omega$ et $v \in X$. On appelle **dérivée directionnelle** de h en a de direction v notée $\frac{dh}{dv}(a)$ la limite réelle suivante (si elle existe)

$$\frac{dh}{dv}(a) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [h(a + \lambda v) - h(a)]. \quad (2.12)$$

Rappel : Si $X = \mathbb{R}^n$ et $h \in C^1(\Omega)$ alors

$$\frac{dh}{dv}(a) = \langle \nabla h(a) | v \rangle, \quad \forall a \in \Omega, \forall v \in \mathbb{R}^n.$$

Remarque 2.8. Dans cette définition l'hypothèse que X est un e.v.n. sert uniquement pour la notion d'ensemble ouvert Ω . On peut généraliser la notion de dérivée directionnelle dans le cas où X est un espace vectoriel (sans aucune norme!), $\Omega \subset X$, $a \in \Omega$, $v \in X$ et $h : \Omega \rightarrow \mathbb{R}$. Il nous faut encore une hypothèse : il existe I un intervalle ouvert en \mathbb{R} qui contient 0 tel que $a + \lambda v \in \Omega$, $\forall \lambda \in I$. Alors on peut encore définir la dérivée directionnelle de h en a de direction v (notée toujours $\frac{dh}{dv}(a)$) par la même limite de (2.12) (si elle existe).

Dans un espace de Hilbert nous avons les notions suivantes :

Définition 2.4. Soit H un espace de Hilbert, Ω un sous-ensemble ouvert et non vide de H et $h : \Omega \rightarrow \mathbb{R}$ une fonction.

1. Soit $a \in \Omega$. On dit que h est **dérivable Gâteaux** en a (on dit aussi que h est **\mathcal{G} - dérivable**) si pour tout $v \in H$ la dérivée directionnelle de h en a et de direction v existe et si en plus il existe $w \in H$ tel que

$$\frac{dh}{dv}(a) = \langle w | v \rangle, \quad \forall v \in H. \quad (2.13)$$

Cet élément $w \in H$ (dont on peut montrer qu'il est unique) s'appelle le **gradient** de h en H et il est noté $\nabla h(a)$.

Donc $\nabla h(a) \in H$ a la propriété

$$\frac{dh}{dv}(a) = \langle \nabla h(a) | v \rangle, \quad \forall v \in \mathbb{R}^n.$$

2. Soit $B \subset \Omega$. On dit que h est **\mathcal{G} - dérivable sur B** si h est **\mathcal{G} - dérivable** en x pour tout $x \in B$.

Remarque 2.9. 1. Dans le cas particulier $H = \mathbb{R}^p$ alors on retrouve le fait que $\nabla h(a)$ est le vecteur des dérivées partielles (car on sait que si h est **\mathcal{G} - dérivable** alors $\frac{dh}{dv}(a) = \langle \nabla h(a) | v \rangle$, $\forall v \in \mathbb{R}^p$ et on conclut par unicité).

2. Si H est un espace de Banach on peut encore définir $\nabla h(a)$ mais comme un élément du dual H' de H (en fait dans le cas particulier où H est un espace de Hilbert le dual H' "s'identifie" à H).

On admet sans preuve le résultat suivant, l'analogie pour un espace de Hilbert général de la Proposition 2.1 :

Proposition 2.2. *Soit H un espace de Hilbert, $g : H \rightarrow \mathbb{R}$ une fonction et $A \subset H$ un ensemble **fermé et convexe**. Nous avons*

1. *Supposons que g est continue et **convexe** sur A , avec en plus g coercive dans le cas où A est non borné. Alors il existe au moins une solution du problème de minimisation (2.9).*
2. *Si en plus g est strictement convexe alors on a aussi l'unicité de la solution de (2.9).*
3. *Si g est dérivable Gâteaux sur H (sur un voisinage ouvert de A c'est suffisant) et si $y^* \in A$ est une solution de (2.9) (c'est à dire, y^* satisfait (2.10)), alors on a l'inéquation d'Euler :*

$$\langle \nabla g(y^*) | y - y^* \rangle \geq 0, \quad \forall y \in A. \quad (2.14)$$

Remarque 2.10. *Si on sait que y^* est dans l'intérieur de A (ce qui est toujours vrai par exemple dans le cas particulier $A = H$), alors (2.14) est équivalente à l'équation d'Euler :*

$$\nabla g(y^*) = 0.$$

Remarque 2.11. *La partie 3. de la Proposition 2.2 peut se généraliser sur un espace vectoriel X quelconque. Supposons qu'on a les ensembles $A \subset \Omega \subset X$, avec A convexe, une fonction $g : \Omega \rightarrow \mathbb{R}$ et un élément y^* qui est un point de minimum de g sur A (donc $g(y^*) \leq g(y)$, $\forall y \in A$). Supposons en plus que la dérivée directionnelle de g en y^* existe pour tout vecteur $v \in X$. On a alors*

$$\frac{dg}{d(y - y^*)}(y^*) \geq 0, \quad \forall y \in A. \quad (2.15)$$

(car $\frac{dg}{d(y - y^*)}(y^*) = \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} [g(y^* + \lambda(y - y^*)) - g(y^*)] \geq 0$ car $g(y^* + \lambda(y - y^*)) - g(y^*) \geq 0$).

Dans la suite on va introduire une notion qui va nous simplifier l'écriture des inégalités variationnelles comme celles de la proposition précédente. Nous avons la définition suivante :

Définition 2.5. *Soit H un espace de Hilbert, $K \subset H$ un sous-ensemble et $x \in K$. On appelle **cône normal** à K en x , noté $N_K(x)$ l'ensemble*

$$N_K(x) = \{w \in H, \quad \langle w | y - x \rangle_H \leq 0, \quad \forall y \in K\}.$$

Remarque 2.12. *Cette notion de cône normal se généralise à des espaces de Banach X , mais dans ce cas la cône normal sera un sous-ensemble de X' (dual de X) au lieu d'un sous-ensemble de X .*

Exemple 2.2. (fondamental)

Si $x \in \overset{\circ}{K}$ (intérieur de K) alors $N_K(x) = \{0\}$.

Démonstration. L'inclusion " \supset " est évidente car $N_K(x)$ contient toujours l'élément 0. L'inclusion " \subset " se démontre par absurd. Supposons qu'il existe $w \in N_K(x)$ avec $w \neq 0$. Comme $x \in \overset{\circ}{K}$ alors il existe $\epsilon > 0$ tel que $B(x, \epsilon) \subset K$. On considère

$$y = x + \frac{\epsilon}{2\|w\|}w \in B(x, \epsilon) \subset K$$

Nous avons

$$\langle w | y - x \rangle_H = \frac{\epsilon}{2\|w\|}\|w\|^2 > 0$$

ce qui est en contradiction avec $\langle w | y - x \rangle_H \leq 0$. □

Conséquence immédiate :

Si H est un espace de Hilbert alors $N_H(x) = \{0\}$, $\forall x \in H$.

Exemple 2.3. (fondamental)

Soit $H = \mathbb{R}$, $K \subset \mathbb{R}$ un intervalle fermé et $x \in K$. Alors

1. Si $x = \min(K)$ alors $N_K(x) = \mathbb{R}_- =]-\infty, 0]$.
2. Si $x = \max(K)$ alors $N_K(x) = \mathbb{R}_+ = [0, +\infty[$.

La preuve est immédiate et laissée en exercice.

Par exemple si $K = [0, +\infty[$ alors on a

$$N_K(x) = \begin{cases}]-\infty, 0] & \text{si } x = 0 \\ \{0\} & \text{si } x > 0 \end{cases}$$

Remarque 2.13. La condition générale d'optimalité (2.14) de la Proposition 2.2 peut s'écrire sous la forme équivalente suivante :

$$-\nabla g(y^*) \in N_A(y^*).$$

Nous avons le résultat suivant :

Lemme 2.1. Soit V un sous-ensemble de l'espace de Hilbert $L^2(]t_0, t_1[; \mathbb{R}^m)$ où V est défini par

$$V = \{v \in L^2(]t_0, t_1[; \mathbb{R}^m), \quad v(t) \in \mathcal{U}, \text{ p.p. } t \in]t_0, t_1[\}.$$

avec $\mathcal{U} \subset \mathbb{R}^m$ un ensemble fermé et convexe. Soit $u \in V$ et $w \in L^2(]t_0, t_1[; \mathbb{R}^m)$. Alors les deux propositions suivantes sont équivalentes :

1.

$$w \in N_V(u) \tag{2.16}$$

2.

$$w(t) \in N_{\mathcal{U}}(u(t)), \quad \text{p.p. } t \in]t_0, t_1[. \tag{2.17}$$

Démonstration. On donne ici seulement l'idée de la preuve de l'implication 2. \implies 1.
Par hypothèse nous avons

$$\langle w(t) \mid y - u(t) \rangle_{\mathbb{R}^m} \leq 0, \quad \text{p.p. } t \in]t_0, t_1[, \quad \forall y \in \mathcal{U}. \quad (2.18)$$

D'autre part pour un $v \in V$ arbitraire on peut prendre $y = v(t)$ en (2.18) pour presque tout $t \in]t_0, t_1[$ arbitraire, ce qui donne

$$\langle w(t) \mid v(t) - u(t) \rangle_{\mathbb{R}^m} \leq 0, \quad \text{p.p. } t \in]t_0, t_1[.$$

En intégrant en t on obtient le résultat. □

Exemple 2.4. Soient $m = 1$, $\mathcal{U} = [0, +\infty[$ donc

$$V = \{v \in L^2(]t_0, t_1[; \mathbb{R}), \quad v(t) \geq 0, \text{ p.p. } t \in]t_0, t_1[\}.$$

Supposons qu'on a $u \in V$ et $w \in L^2(]t_0, t_1[; \mathbb{R})$. Alors $w \in N_V(u)$ si et seulement si pour presque tout $t \in]t_0, t_1[$ on a $w(t) \in N_{[0, +\infty[}(u(t))$, c'est à dire

$$\begin{cases} w(t) = 0 & \text{si } u(t) > 0 \\ w(t) \leq 0 & \text{si } u(t) = 0 \end{cases}$$

2.3 Le principe de minimum de Pontryagin pour le contrôle optimal des systèmes EDO

2.3.1 Le cas sans contraintes sur l'état final

On revient au problème de contrôle (2.3).

Dans la suite, pour faciliter l'écriture, on notera simplement L^2 à la place de $L^2(]t_0, t_1[; \mathbb{R}^m)$. On suppose ici $C_1 = \mathbb{R}^n$, donc il n'y a pas des contraintes sur l'état final; rappelons que dans ce cas notre problème de contrôle optimal s'écrit

$$\min \{J(u), \quad u \in U_{ad}\} \quad (2.19)$$

avec U_{ad} un sous ensemble non vide de $L^2(]t_0, t_1[; \mathbb{R}^m)$.

Rappelons que $J : L^2(]t_0, t_1[; \mathbb{R}^m) \rightarrow \mathbb{R}$ est une fonction donnée par (2.8), c'est à dire

$$J(u) = \int_{t_0}^{t_1} L(t, x^u(t), u(t)) dt + \varphi(x^u(t_1)), \quad \forall u \in L^2 \quad (2.20)$$

avec $x^u = x$ l'unique solution du problème de Cauchy

$$\begin{cases} x'(t) = f(t, x(t), u(t)), & t \in [t_0, t_1] \\ x(t_0) = x^0 \end{cases} \quad (2.21)$$

On supposera partout dans ce chapitre :

U_{ad} est un sous-ensemble convexe et fermé de L^2 .

On aura besoin dans la suite de montrer que J est dérivable Gâteaux et de calculer son gradient. Pour cela on commence par calculer la dérivée directionnelle de J .

On considère u et v arbitraires dans l'espace de Hilbert L^2 ; on doit calculer la limite suivante :

$$\lim_{\lambda \rightarrow 0} \frac{[J(u + \lambda v) - J(u)]}{\lambda}.$$

Nous ferons dans la suite un calcul purement formel, sans preuve rigoureuse.

Nous avons d'abord

$$J(u + \lambda v) - J(u) = E_1 + E_2$$

avec

$$E_1 = \int_{t_0}^{t_1} [L(t, x^{u+\lambda v}(t), u(t) + \lambda v(t)) - L(t, x^u(t), u(t))] dt$$

et

$$E_2 = \varphi(x^{u+\lambda v}(t_1)) - \varphi(x^u(t_1)).$$

D'autre part on a

$$E_1 = E_3 + E_4$$

avec

$$E_3 = \int_{t_0}^{t_1} [L(t, x^u(t), u(t) + \lambda v(t)) - L(t, x^u(t), u(t))] dt$$

et

$$E_4 = \int_{t_0}^{t_1} [L(t, x^{u+\lambda v}(t), u(t) + \lambda v(t)) - L(t, x^u(t), u(t) + \lambda v(t))] dt.$$

En utilisant le Théorème d'Accroissements Finis (TAF) on a

$$E_3 = \int_{t_0}^{t_1} \langle \nabla_u L(t, x^u(t), u(t) + \theta(t)[(u + \lambda v)(t) - u(t)]) | [(u + \lambda v)(t) - u(t)] \rangle dt$$

avec la notation

$$\nabla_u L(t, x, u) = \begin{pmatrix} \frac{\partial L}{\partial u_1} \\ \frac{\partial L}{\partial u_2} \\ \cdot \\ \cdot \\ \frac{\partial L}{\partial u_m} \end{pmatrix}$$

(c'est le **gradient** de L par rapport à u).

On montre alors (encore sous des "bonnes" hypothèses sur la fonction L) :

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} E_3 = \int_{t_0}^{t_1} \langle \nabla_u L(t, x^u(t), u(t)) | v(t) \rangle dt$$

Ensuite on a

$$E_4 = \int_{t_0}^{t_1} \langle \nabla_x L(t, x^u(t) + \theta(t)[x^{u+\lambda v}(t) - x^u(t)], u(t) + \lambda v(t)) \mid (x^{u+\lambda v}(t) - x^u(t)) \rangle dt$$

avec $\theta(t) \in]0, 1[$.

Nous avons utilisé la notation

$$\nabla_x L(t, x, u) = \begin{pmatrix} \frac{\partial L}{\partial x_1} \\ \frac{\partial L}{\partial x_2} \\ \cdot \\ \cdot \\ \frac{\partial L}{\partial x_n} \end{pmatrix}$$

(c'est le **gradient** de L par rapport à u).

On montre alors (sous des "bonnes" hypothèses sur la fonction L) :

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} E_4 = \int_{t_0}^{t_1} \langle \nabla_x L(t, x^u(t), u(t)) \mid z(t) \rangle dt$$

où on pose

$$z(t) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [x^{u+\lambda v}(t) - x^u(t)].$$

Nous admettons (formellement) que cette limite existe, ce qui implique qu'on a

$$x^{u+\lambda v}(t) \rightarrow x^u(t), \quad \text{pour } \lambda \rightarrow 0.$$

En utilisant encore (TAF) on obtient aussi

$$E_2 = \langle \nabla \varphi(x^u(t_1) + \theta[x^{u+\lambda v}(t_1) - x^u(t_1)]) \mid [x^{u+\lambda v}(t_1) - x^u(t_1)] \rangle$$

ce qui donne pour φ de classe C^1 :

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} E_2 = \langle \nabla \varphi(x^u(t_1)) \mid z(t_1) \rangle.$$

Nous avons alors obtenu (formellement)

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [J(u + \lambda v) - J(u)] &= \int_{t_0}^{t_1} \langle \nabla_u L(t, x^u(t), u(t)) \mid v(t) \rangle dt + \\ &\int_{t_0}^{t_1} \langle \nabla_x L(t, x^u(t), u(t)) \mid z(t) \rangle dt + \langle \nabla \varphi(x^u(t_1)) \mid z(t_1) \rangle \end{aligned} \quad (2.22)$$

Nous voulons en fait mettre cette expression sous la forme

$$\int_{t_0}^{t_1} \langle w(t) \mid v(t) \rangle dt \quad \text{c'est à dire } \langle w \mid v \rangle_{L^2}$$

avec $w \in L^2$ à trouver.

Pour cela on a besoin de connaître la fonction z . On pose

$$z_\lambda = \frac{1}{\lambda} (x^{u+\lambda v} - x^u)$$

et nous avons

$$\begin{cases} (x^{u+\lambda v})' = f(t, x^{u+\lambda v}(t), (u + \lambda v)(t)) \\ x^{u+\lambda v}(t_0) = x^0 \end{cases}$$

et

$$\begin{cases} (x^u)' = f(t, x(t), u(t)) \\ x^u(t_0) = x^0 \end{cases}$$

L'idée est de faire la différence entre ces deux systèmes équations, de diviser par λ et de passer à la limite $\lambda \mapsto 0$.

Il est difficile de faire ceci dans le cas général avec une fonction f arbitraire ; on fera les calculs dans le cas d'un système contrôlé linéaire et on admettra le résultat dans le cas général.

Nous considérons alors pour la suite du calcul une fonction f comme dans la Remarque 2.4, donc de la forme

$$f(t, \tilde{x}, \tilde{u}) = A(t)\tilde{x} + B(t)\tilde{u} + g(t).$$

Alors les deux systèmes ci-dessus s'écrivent

$$\begin{cases} (x^{u+\lambda v})' = Ax^{u+\lambda v} + B(u + \lambda v) + g \\ x^{u+\lambda v}(t_0) = x^0 \end{cases}$$

et

$$\begin{cases} (x^u)' = Ax^u + Bu + g \\ x^u(t_0) = x^0 \end{cases}$$

En faisant la différence entre ces deux systèmes et en divisant par λ on trouve

$$\begin{cases} z'_\lambda = Az_\lambda + Bv \\ z_\lambda(t_0) = 0 \end{cases}$$

On observe que z_λ est en fait indépendante de λ et donc on peut la noter par z ; la limite pour $\lambda \rightarrow 0$ de z_λ est alors z . On a donc obtenu : z est l'unique solution du système

$$\begin{cases} z' = Az + Bv \\ z(t_0) = 0 \end{cases} \quad (2.23)$$

Nous faisons le produit scalaire de (2.23) avec une fonction $p(t)$, où $p : [t_0, t_1] \rightarrow \mathbb{R}^n$ est une fonction à déterminer ("avec les propriétés de régularité qu'il faut") et on intègre entre t_0 et t_1 . En faisant une intégration par parties on obtient

$$\langle z(t_1) | p(t_1) \rangle - 0 - \int_{t_0}^{t_1} \langle z(t) | p'(t) \rangle dt = \int_{t_0}^{t_1} \langle z(t) | A^T p(t) \rangle dt + \int_{t_0}^{t_1} \langle v(t) | B^T p(t) \rangle dt$$

(on a utilisé $\langle Az | p \rangle = \langle z | A^T p \rangle$ et $\langle Bv | p \rangle = \langle v | B^T p \rangle$), ce qui donne

$$\langle z(t_1) | p(t_1) \rangle - \int_{t_0}^{t_1} \langle z(t) | p'(t) + A^T p(t) \rangle dt = \int_{t_0}^{t_1} \langle v(t) | B^T p(t) \rangle dt.$$

Alors si on choisit p tel que

$$\begin{cases} p' + A^T p = -\nabla_x L(t, x^u, u) \\ p(t_1) = \nabla \varphi(x^u(t_1)) \end{cases}$$

alors on aura de (2.22) :

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [J(u + \lambda v) - J(u)] = \int_{t_0}^{t_1} \langle \nabla_u L(t, x^u(t), u(t)) | v(t) \rangle dt + \int_{t_0}^{t_1} \langle B^T p(t) | v(t) \rangle dt.$$

On a alors "démonstré" le résultat suivant :

Lemme 2.2. *Dans le cas linéaire (donc avec f comme dans la Remarque 2.4), sous des hypothèses appropriées pour L et si φ est de classe C^1 , la fonction J est dérivable Gâteaux sur L^2 et on a*

$$\nabla J(u)(t) = B^T(t) p(t) + \nabla_u L(t, x^u(t), u(t)), \quad \forall u \in L^2 \quad (2.24)$$

avec $p : [t_0, t_1] \rightarrow \mathbb{R}^n$ l'unique solution du problème appelé **problème adjoint** :

$$\begin{cases} p'(t) = -A^T(t) p(t) - \nabla_x L(t, x^u(t), u(t)) \\ p(t_1) = \nabla \varphi(x^u(t_1)) \end{cases} \quad (2.25)$$

et avec $x = x^u$ l'unique solution de l'équation d'état (2.5).

Revenons maintenant au cas général d'un système contrôlé qui n'est plus forcément linéaire, donc avec une fonction f arbitraire.

On admettra qu'on obtient un résultat analogue au résultat dans le cas linéaire; simplement, au lieu d'avoir $B^T(t)$ en (2.24) on aura $\nabla_u f(t, x^u(t), u(t))$ et au lieu de $A^T(t)$ en (2.25) on aura $\nabla_x f(t, x^u(t), u(t))$. Remarquer que dans le cas linéaire on retrouve

$$\nabla_u f(t, x^u(t), u(t)) = B^T(t) \quad \text{et} \quad \nabla_x f(t, x^u(t), u(t)) = A^T(t).$$

Nous admettons donc le résultat suivant :

Lemme 2.3. *Sous des hypothèses appropriées pour f et L et si φ est de classe C^1 , la fonction J est dérivable Gâteaux sur L^2 et on a*

$$\nabla J(u)(t) = \nabla_u f(t, x^u(t), u(t)) p(t) + \nabla_u L(t, x^u(t), u(t)), \quad \forall u \in L^2 \quad (2.26)$$

avec $p : [t_0, t_1] \rightarrow \mathbb{R}^n$ l'unique solution du **problème adjoint** :

$$\begin{cases} p'(t) = -\nabla_x f(t, x^u(t), u(t)) p(t) - \nabla_x L(t, x^u(t), u(t)) \\ p(t_1) = \nabla \varphi(x^u(t_1)) \end{cases} \quad (2.27)$$

où $x = x^u$ est l'unique solution de l'équation d'état (2.1). La fonction p s'appelle la **variable adjointe** pour le problème de contrôle optimal considéré.

Remarque 2.14. 1. Ce résultat permet d'utiliser une méthode de gradient pour résoudre numériquement le problème de contrôle optimale (voir Chapitre 4).

2. Le problème adjoint est un problème de Cauchy avec inconnue p avec $p(t_1)$ donné. Alors si on connaît u on peut calculer x^u et ensuite p , ce qui nous donne $\nabla J(u)$.

Comme conséquence immédiate de ce lemme, de la Proposition 2.2 et de la Remarque 2.13 nous avons le résultat suivant qu'on appelle le **principe de minimum de Pontryagin (PMP)** :

Théorème 2.1. *Sous des hypothèses appropriées pour f et L et si φ est de classe C^1 , alors toute solution u^* du problème d'optimisation (2.19) satisfait le système suivant : il existe $x^*, p^* : [t_0, t_1] \rightarrow \mathbb{R}^n$ tels que*

$$(x^*)'(t) = f(t, x^*(t), u^*(t)) \quad (2.28)$$

$$x^*(t_0) = x^0 \quad (2.29)$$

$$(p^*)'(t) = -\nabla_x f(t, x^*(t), u^*(t)) p^*(t) - \nabla_x L(t, x^*(t), u^*(t)) \quad (2.30)$$

$$p^*(t_1) = \nabla \varphi(x^*(t_1)) \quad (2.31)$$

$$-\nabla_u f(t, x^*, u^*) p^* - \nabla_u L(t, x^*, u^*) \in N_{U_{ad}}(u^*) \quad (2.32)$$

où $N_{U_{ad}}(u^*)$ est le cône normal en u^* à U_{ad} dans l'espace L^2 (la fonction x^* n'est autre que x^{u^*}).

Remarque 2.15. 1. Dans le cas d'un système contrôlé linéaire (donc avec f comme dans la Remarque 2.4), le principe de minimum de Pontryagin s'écrit

$$(x^*)' = Ax^* + Bu^* + g \quad (2.33)$$

$$x^*(t_0) = x^0 \quad (2.34)$$

$$(p^*)' = -A^T p^* - \nabla_x L(t, x^*, u^*) \quad (2.35)$$

$$p^*(t_1) = \nabla \varphi(x^*(t_1)) \quad (2.36)$$

$$-B^T p^* - \nabla_u L(t, x^*, u^*) \in N_{U_{ad}}(u^*). \quad (2.37)$$

2. Le système (2.28) - (2.32) donne une condition **nécessaire** pour que u^* soit solution du problème d'optimisation (2.19). La(les) solution(s) qu'on cherche se trouve(nt) parmi les solutions de (2.28) - (2.32). Dans le cas où on sait qu'on a existence et unicité d'une solution de (2.19) et si le système d'optimalité (2.28) - (2.32) nous donne une seule solution, alors c'est la solution recherchée. Plus généralement, si on sait qu'il existe au moins une solution de (2.19) et si le système d'optimalité (2.28) - (2.32) a une seule solution, alors il existe une unique solution de (2.19) qui est la solution de (2.28) - (2.32).

3. Dans le cas particulier très fréquent

$$U_{ad} = \{v \in L^2, \quad v(t) \in \mathcal{U}, \text{ p.p. } t \in]t_0, t_1[\}$$

avec $\mathcal{U} \subset \mathbb{R}^m$ un ensemble fermé et convexe, alors la condition (2.32) du Théorème 2.1 est équivalente à la condition

$$-\nabla_u f(t, x^*(t), u^*(t)) p^*(t) - \nabla_u L(t, x^*(t), u^*(t)) \in N_{\mathcal{U}}(u^*(t)), \quad p.p. t \in]t_0, t_1[. \quad (2.38)$$

où $N_{\mathcal{U}}(u^*(t))$ représente le cône normal à \mathcal{U} en $u^*(t)$ dans l'espace de Hilbert \mathbb{R}^m (conséquence du Lemme 2.1).

4. Dans le cas encore plus particulier

$$U_{ad} = L^2$$

alors la condition (2.32) du Théorème 2.1 est équivalente à la condition

$$-\nabla_u f(t, x^*(t), u^*(t)) p^*(t) - \nabla_u L(t, x^*(t), u^*(t)) = 0, \quad p.p. t \in]t_0, t_1[.$$

5. Le fait de chercher le contrôle u dans un sous-ensemble U_{ad} de L^2 n'est pas la seule possibilité pour aborder un problème de contrôle optimal de ce type.

Il est par exemple possible de travailler dans un espace de Banach comme $L^q(]t_0, t_1[, \mathbb{R}^m)$ avec $q \geq 1$.

Une possibilité qui est très utilisée dans la pratique est de chercher u dans un sous-ensemble U_{ad} de l'espace vectoriel normé qu'on peut noter $CM([t_0, t_1], ; \mathbb{R}^m)$ qui est l'ensemble des fonctions continues par morceaux sur $[t_0, t_1]$, avec la norme "sup". Plus généralement on peut chercher le contrôle u dans l'espace vectoriel des fonction mesurables. Il faut alors utiliser la dérivée directionnelle et l'inégalité (2.15) de la Remarque 2.11.

Nous admettons que dans tous ces cas on arrive au même système d'optimalité (2.28) - (2.31) et (2.38) mais l'obtention de ce système est beaucoup plus difficile.

2.3.2 Le cas avec contraintes sur l'état final

On considère ici le cas $C_1 \subset \mathbb{R}^n$ mais $C_1 \neq \mathbb{R}^n$; on va supposer que C_1 est **fermé** en \mathbb{R}^n . On supposera aussi que

$$U_{ad} = \{v \in L^2, \quad v(t) \in \mathcal{U}, \quad p.p. t \in]t_0, t_1[\}.$$

avec $\mathcal{U} \subset \mathbb{R}^m$ un sous-ensemble fermé et convexe. Nous rappelons que le problème de minimisation (2.3) ne peut pas se réduire à un problème de minimisation sur la variable u seulement.

Le principe de minimum de Potryagin dans ce cas est le suivant (résultat admis) :

Théorème 2.2. *Sous des hypothèses appropriées pour f et L et si φ est de classe C^1 , alors toute solution u^* du problème d'optimisation (2.3) satisfait le système suivant : il existe*

$x^*, p^* : [t_0, t_1] \rightarrow \mathbb{R}^n$ tels que

$$(x^*)' = f(t, x^*(t), u^*(t)) \quad (2.39)$$

$$x^*(t_0) = x^0 \quad (2.40)$$

$$(p^*)' = -\nabla_x f(t, x^*(t), u^*(t))p^* - \nabla_x L(t, x^*, u^*) \quad (2.41)$$

$$p^*(t_1) - \nabla \varphi(x^*(t_1)) \in N_{C_1}(x^*(t_1)) \quad (2.42)$$

$$-\nabla_u f(t, x^*(t), u^*(t))p^*(t) - \nabla_u L(t, x^*(t), u^*(t)) \in N_{\mathcal{U}}(u^*(t)), \quad p.p. t \in]t_0, t_1[\quad (2.43)$$

où $N_{C_1}(x^*(t_1))$ est le cône normal en $x^*(t_1)$ à C_1 dans \mathbb{R}^n et $N_{\mathcal{U}}(u^*(t))$ est le cône normal en $u^*(t)$ à \mathcal{U} dans \mathbb{R}^m .

Remarque 2.16. 1. Dans le cas particulier $C_1 = \mathbb{R}^n$ l'égalité (2.42) devient (2.31).

2. L'idée de la preuve du Théorème 2.2 consiste à voir le problème de contrôle comme un problème de minimisation avec contraintes sur (x, u) et où le problème de Cauchy (2.1) avec $x(t_1) \in C_1$ sont des contraintes. Le fait d'écrire les conditions d'optimalité pour un point de minimum (x^*, u^*) mène au système d'optimalité du Théorème 2.2.

3. En fait la formulation du Théorème 2.2 n'est pas tout à fait exacte. En général il y a encore une inconnue $p_0 \geq 0$ telle que $\exists t \in [t_0, t_1]$ avec $\|p(t)\| + p_0 > 0$. Mais dans le cas où les contraintes du problème sont "qualifiés" (sans préciser ce que ça veut dire; admettons que c'est très souvent le cas pour ce type de problème et que dans tous les exemples de ce cours cette condition sera satisfaite!) alors on peut prendre $p_0 = 1$ et le système d'optimalité peut se réduire à celui énoncé dans le Théorème 2.2.

2.3.3 Un résultat de contrôlabilité

On donnera ici un résultat de contrôlabilité qui est surtout nécessaire quand on considère des problèmes de contrôle optimal avec des contraintes sur l'état final.

Le système contrôlé qu'on considère ici est de la forme

$$\begin{cases} x'(t) = Ax(t) + Bu(t) + g(t), & t \in [t_0, t_1] \\ x(t_0) = x^0 \end{cases} \quad (2.44)$$

avec $A \in \mathcal{M}_n(\mathbb{R})$, $B \in \mathcal{M}_{n,m}(\mathbb{R})$ des matrices constantes et $g :]t_0, t_1[\rightarrow \mathbb{R}^n$ une fonction de $L^1(]t_0, t_1[, \mathbb{R}^n)$.

Il s'agit donc du cas particulier du système contrôlé linéaire à coefficients constants.

Définition 2.6. On appelle **ensemble des points accessibles** de x^0 en temps t_1 pour (2.44) l'ensemble

$$Acc(x^0, t_1) = \{x^u(t_1), u \in L^2\} \subset \mathbb{R}^n$$

avec x^u solution du système (2.44).

Définition 2.7. On dit que le système (2.44) est **contrôlable** en temps t_1 si $Acc(x^0, t_1) = \mathbb{R}^n$ (c'est à dire, si on a : $\forall z \in \mathbb{R}^n \exists u \in L^2$ tel que $x^u(t_1) = z$).

Nous avons le résultat suivant :

Théorème 2.3. *Le système (2.44) est contrôlable pour tous $x^0 \in \mathbb{R}^n$ et $t_1 > t_0$ si et seulement si la matrice suivante donnée par blocs :*

$$K = (B, AB, A^2B, \dots, A^{n-1}B)$$

satisfait la condition

$$\text{rang}(K) = n. \quad (2.45)$$

La matrice K s'appelle **matrice de Kalman** et la condition (2.45) s'appelle **condition de Kalman**.

Remarque 2.17. 1. *Chaque bloc de la forme $A^j B$ de K est une matrice à n lignes et m colonnes; alors K est une matrice à n lignes et mn colonnes.*

2. *Le fait que le système (2.44) soit contrôlable ou non ne dépend pas de x^0 , t_1 ou de la fonction g .*

Idée preuve Théorème (2.3) :

On considère l'application $\phi : L^2 \rightarrow \mathbb{R}^n$ donnée par

$$u \mapsto \phi(u) = \int_{t_0}^{t_1} e^{(t_1-t)A} B u(t) dt$$

Remarquons que ϕ est une application linéaire.

Nous avons la formule suivante :

$$x(t_1) = e^{(t_1-t_0)A} x^0 + \int_{t_0}^{t_1} e^{(t_1-t)A} g(t) dt + \int_{t_0}^{t_1} e^{(t_1-t)A} B u(t) dt.$$

Alors $\text{Acc}(x^0, t_1) = \mathbb{R}^n$ si et seulement si ϕ est une application surjective. Tout revient alors à démontrer :

$$\text{rang}(K) = n \Leftrightarrow \phi \text{ surjective.}$$

Comme le rang maximal possible de la matrice K est n (le nombre de lignes) alors tout revient à démontrer

$$\text{rang}(K) < n \Leftrightarrow \phi \text{ non surjective.}$$

Montrons \Rightarrow :

On suppose donc $\text{rang}(K) < n$ et on doit montrer que ϕ n'est pas surjective. Il existe alors une combinaison linéaire nulle des lignes de K avec des coefficients pas tous nuls; si on note K_1, K_2, \dots, K_n les lignes de K il existe donc un vecteur ligne $s = (s_1, \dots, s_n)$ non nul tel que $s_1 K_1 + s_2 K_2 + \dots + s_n K_n = 0$, c'est à dire $sK = 0$. En écrivant cela par blocs on a

$$sB = sAB = sA^2B = \dots = sA^{n-1}B = 0. \quad (2.46)$$

D'autre part le Théorème de Cayley-Hamilton nous dit

$$P_A(A) = 0$$

où P_A est le polynome caractéristique de A défini par

$$P_A(\lambda) = \det(A - \lambda I_n), \quad \forall \lambda \in \mathbb{C}.$$

De Théorème de Cayley-Hamilton on déduit une relation de la forme

$$A^n + b_{n-1}A^{n-1} + b_{n-2}A^{n-2} + \dots + b_1A + b_0I_n = 0, \quad \text{avec } b_0, b_1 \dots b_{n-1} \in \mathbb{R}. \quad (2.47)$$

On multiplie cette égalité à gauche par s et à droite par B et à l'aide de (2.46) on obtient $sA^nB = 0$. En multipliant (2.47) à gauche par sA et à droite par B et en utilisant de nouveau (2.46) et le fait que $sA^nB = 0$ on déduit $sA^{n+1}B = 0$. Par récurrence on montre facilement

$$sA^k B = 0, \quad \forall k \in \mathbb{N}.$$

De la définition de e^{tA} on déduit alors

$$se^{tA}B = 0, \quad \forall t \in [t_0, t_1].$$

Nous obtenons alors

$$s \int_{t_0}^{t_1} e^{(t_1-t)A} B u(t) dt, \quad \forall u \in L^2$$

c'est à dire

$$s\phi(u) = 0, \quad \forall u \in L^2.$$

Ceci nous dit qu'il n'existe pas $u \in L^2$ tel que $\phi(u) = s^T$, car avec l'égalité précédente cela donnerait $\|s\|^2 = 0$ donc $s = 0$, contradiction. Nous avons donc montré la non surjectivité de ϕ .

Montrons \Leftarrow :

On suppose que ϕ est non surjective et on doit montrer que $\text{rang}(K) < n$. Comme ϕ n'est pas surjective alors $\phi(L^2)$ est un sous-espace vectoriel de \mathbb{R}^n strictement inclu dans \mathbb{R}^n . Il existe alors un vecteur ligne $s = (s_1, \dots, s_n)$ dans \mathbb{R}^n tel que s^T soit orthogonal à $\phi(L^2)$. Ceci donne

$$s\phi(u) = 0, \quad \forall u \in L^2$$

c'est à dire

$$\int_{t_0}^{t_1} se^{(t_1-t)A} B u(t) dt, \quad \forall u \in L^2$$

donc

$$\langle (se^{(t_1-t)A} B)^T | u \rangle = 0, \quad \forall u \in L^2$$

où $\langle \cdot | \cdot \rangle$ désigne le produit scalaire en L^2 . Ceci donne

$$se^{(t_1-t)A} B = 0, \quad \forall t \in [t_0, t_1]. \quad (2.48)$$

En prenant $t = t_1$ en (2.48) on obtient

$$sB = 0.$$

En dérivant (2.48) en t et en prenant $t = t_1$ on obtient

$$sAB = 0.$$

Par dérivation de (2.48) en t à l'ordre $k \in \mathbb{N}$ et en prenant ensuite $t = t_1$ on obtient

$$sA^k B = 0, \quad \forall k \in \mathbb{N}.$$

Ceci nous donne l'égalité matricielle

$$s(B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B) = 0 \quad \text{donc} \quad sK = 0$$

ce qui nous dit $\text{rang}(K) < n$ et finit la preuve.

Quelques exemples :

1. Si $B = 0$ alors la matrice de Kalman est la matrice 0 qui est de rang 0. Comme $0 < n$ alors le système (2.44) n'est pas contrôlable.

On peut aussi montrer directement que ce système n'est pas contrôlable : pour tout $u \in L^2$ la solution de (2.44) est donnée par

$$x^u(t_1) = e^{(t_1-t_0)A}x^0 + \int_{t_0}^{t_1} e^{A(t_1-t)}g(s) ds$$

donc $\text{Acc}(x^0, t_1)$ se réduit au singleton $\{e^{(t_1-t_0)A}x^0 + \int_{t_0}^{t_1} e^{A(t_1-t)}g(s) ds\}$ en \mathbb{R}^n . Alors le système (2.44) n'est pas contrôlable.

2. Supposons que $m = n = 1$, donc $A, B \in \mathbb{R}$. Alors la matrice de Kalman K n'est autre que B et il est clair que

$$\text{rang}(K) = 1 \quad \iff \quad B \neq 0.$$

Donc si $B \neq 0$ le Théorème 2.3 nous dit que le système (2.44) est contrôlable.

Montrons maintenant directement que ce système est contrôlable ; il faut donc montrer

$$\text{Acc}(x^0, t_1) = \mathbb{R}, \quad \forall x^0 \in \mathbb{R}, \quad t_1 > t_0.$$

On procède par double inclusion ; l'inclusion " \subseteq " est évidente. Pour montrer " \supseteq " on fixe $z \in \mathbb{R}$ arbitraire et on doit trouver $u \in L^2$ tel que $x(t_1) = z$ avec

$$x(t_1) = e^{(t_1-t_0)A}x^0 + \left(\int_{t_0}^{t_1} e^{(t_1-t)A} u(t) dt \right) B + \int_{t_0}^{t_1} e^{A(t_1-t)}g(s) ds. \quad (2.49)$$

Ceci revient à

$$\int_{t_0}^{t_1} e^{(t_1-t)A} u(t) dt = r$$

où nous notons

$$r = \frac{z - e^{(t_1-t_0)A} x^0 - \int_{t_0}^{t_1} e^{A(t_1-s)} g(s) ds}{B}.$$

On va prendre u constante réelle notée \hat{u} tel que

$$\hat{u} \int_{t_0}^{t_1} e^{(t_1-t)A} dt = r$$

car $\int_{t_0}^{t_1} e^{(t_1-t)A} dt \neq 0$ pour tout $A \in \mathbb{R}$ (facile à vérifier : distinguer les cas $A = 0$ et $A \neq 0$). Donc le système (2.44) est contrôlable.

3. Pour l'exemple donnée au début du cours sur le contrôle d'une voiture (système (2.4) de l'Exemple 2.1) nous avons

$$n = 2, \quad m = 1, \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

La matrice de Kalman de ce système contrôlé est donnée par

$$K = (B \quad AB).$$

On a

$$AB = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

ce qui nous donne

$$K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Il est clair que $\text{rang}(K) = 2$ ce qui nous dit que le système (2.4) est contrôlable.

Il est difficile de vérifier directement en utilisant la définition que ce système est contrôlable.

Remarque 2.18. 1. *Supposons que nous avons un problème de contrôle optimal avec des contraintes sur l'état final pour le système contrôlé (2.44).*

Alors si ce système est contrôlable, nous sommes sûrs qu'il y a au moins un contrôle $u \in L^2$ pour lequel les contraintes sont satisfaites.

Dans le cas où le système (2.44) n'est pas contrôlable il est possible qu'il n'existe aucun $u \in L^2$ tel que les contraintes sur l'état final soient satisfaites. Dans ce cas il n'y a aucune solution du problème de contrôle optimal.

2. *Supposons que nous avons un problème de contrôle optimal avec contraintes sur l'état final pour un système (2.44) qui est contrôlable. Considérons un problème pénalisé associé, qui sera sans contraintes sur l'état final (voir Remarque 2.3). Alors on peut montrer (admis!) que la solution de ce problème pénalisé converge vers la solution du problème avec contraintes sur l'état final.*

Chapitre 3

Contrôle optimal des équations aux dérivées partielles (EDP)

On fera dans cette section uniquement le contrôle d'une EDP parabolique, qui est l'équation de la chaleur d'évolution. On va considérer une simplification du modèle décrit dans l'Introduction dans l'Exemple 1.2. On se donne Ω un ouvert dans \mathbb{R}^n , $n = 1, 2, 3$ et deux temps $t_0, t_1 \in \mathbb{R}$ avec $t_0 < t_1$.

La variable d'état du système sera la température $y : \Omega \times [t_0, t_1] \rightarrow \mathbb{R}$, $y(x, t)$.

Le contrôle du système sera la source de chaleur $u : \Omega \times [t_0, t_1] \rightarrow \mathbb{R}$, $u(x, t)$. Nous avons une température initiale donnée $y_0 : \Omega \rightarrow \mathbb{R}$ et le but est d'arriver au moment t_1 à une température proche d'une autre température donnée $y_f : \Omega \rightarrow \mathbb{R}$, ceci en consommant le minimum d'énergie.

Alors l'équation d'état est le système suivant :

$$\left\{ \begin{array}{ll} \frac{\partial y}{\partial t} - \Delta_x y = u & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial y}{\partial \nu} = g & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ y(x, t_0) = y_0(x) & \text{pour } x \in \Omega. \end{array} \right. \quad (3.1)$$

ou $g : \partial\Omega \times [t_0, t_1] \rightarrow \mathbb{R}$ est une fonction donnée représentant les pertes de chaleur à travers la frontière $\partial\Omega$. On peut formuler le problème suivant de contrôle optimal :

$$\min \left\{ \int_{t_0}^{t_1} \int_{\Omega} u^2(x, t) dx dt + \alpha \int_{\Omega} |y(x, t_1) - y_f(x)|^2 dx, \quad (y, u) \text{ satisfont (3.1), } u \in U_{ad} \right\} \quad (3.2)$$

où U_{ad} est un ensemble des contrôles à préciser et α est une constante strictement positive. Nous pouvons choisir α assez "grande", afin de pénaliser l'éloignement de la fonction $y(\cdot, t_1)$ par rapport à la fonction y_f .

Nous considérons dans la suite l'ensemble ouvert $Q = \Omega \times]t_0, t_1[\subset \mathbb{R}^{n+1}$ et nous choisissons

U_{ad} comme un sous-ensemble fermé et convexe de l'espace de Hilbert $L^2(Q)$. Pour simplifier, on fera toujours dans cette section l'hypothèse supplémentaire :

$$U_{ad} = \{v \in L^2(Q), \quad v(x, t) \in \mathcal{U}, \text{ p.p. } (x, t) \in Q\}$$

où \mathcal{U} est un intervalle fermé dans \mathbb{R} .

Nous admettons que pour tout $u \in L^2(Q)$ il y a existence et unicité d'une solution y du système (3.1), que nous notons aussi $y = y^u$ (sans donner un sens précis de ce que veut dire solution de (3.1)). Comme dans les sections précédentes sur le contrôle des EDO nous pouvons écrire le problème de contrôle (3.2) sous la forme équivalent plus simple :

$$\min \{J(u), \quad u \in U_{ad}\} \quad (3.3)$$

avec $J : U_{ad} \rightarrow \mathbb{R}$ définie par

$$J(u) = \int_{t_0}^{t_1} \int_{\Omega} u^2(x, t) dxdt + \alpha \int_{\Omega} |y^u(x, t_1) - y_f(x)|^2 dx, \quad \forall u \in L^2(Q).$$

Pour simplicité on supposera dans la suite $\alpha = 1$.

Nous admettons l'existence et l'unicité d'une solution u^* (contrôle optimal) du problème (3.3).

Pour écrire pour ce problème un principe de minimum de Pontryagin nous avons besoin de montrer que J est dérivable Gâteaux sur $L^2(Q)$ et de calculer le gradient de J . On fera un calcul analogue à celui des sections précédentes pour les EDO.

Nous fixons $u, v \in L^2(Q)$ et nous cherchons à calculer (et montrer qu'elle existe)

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [J(u + \lambda v) - J(u)].$$

Nous avons

$$J(u + \lambda v) - J(u) = E_1 + E_2$$

avec

$$E_1 = \int_Q [(u + \lambda v)^2 - u^2] dxdt$$

et

$$E_2 = \int_{\Omega} \left\{ [y^{u+\lambda v}(x, t_1) - y_f(x)]^2 - [y^u(x, t_1) - y_f(x)]^2 \right\} dx.$$

En développant, nous avons

$$E_1 = \int_Q (2\lambda uv + \lambda^2 v^2) dxdt$$

ce qui donne

$$\frac{E_1}{\lambda} \rightarrow 2 \int_Q u(x, t)v(x, t) dxdt, \quad \text{pour } \lambda \rightarrow 0. \quad (3.4)$$

D'autre part nous avons

$$E_2 = \int_{\Omega} [y^{u+\lambda v}(x, t_1) - y^u(x, t_1)] \cdot [y^{u+\lambda v}(x, t_1) + y^u(x, t_1) - 2y_f(x)] dx.$$

donc

$$\frac{E_2}{\lambda} = \int_{\Omega} z_{\lambda}(x, t_1) \cdot [y^{u+\lambda v}(x, t_1) + y^u(x, t_1) - 2y_f(x)] dx. \quad (3.5)$$

où on note

$$z_{\lambda}(x, t) = \frac{1}{\lambda} [y^{u+\lambda v}(x, t) - y^u(x, t)], \quad \forall (x, t) \in Q.$$

Nous procédons comme pour les EDO pour trouver le système satisfait par z_{λ} ; nous avons

$$\left\{ \begin{array}{ll} \frac{\partial y^{u+\lambda v}}{\partial t} - \Delta_x y^{u+\lambda v} = u + \lambda v & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial y^{u+\lambda v}}{\partial \nu} = g & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ y^{u+\lambda v}(x, t_0) = y_0(x) & \text{pour } x \in \Omega \end{array} \right.$$

et aussi

$$\left\{ \begin{array}{ll} \frac{\partial y^u}{\partial t} - \Delta_x y^u = u & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial y^u}{\partial \nu} = g & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ y^u(x, t_0) = y_0(x) & \text{pour } x \in \Omega. \end{array} \right.$$

En faisant la différence entre ces deux systèmes et en divisant par λ , on trouve par linéarité

$$\left\{ \begin{array}{ll} \frac{\partial z_{\lambda}}{\partial t} - \Delta_x z_{\lambda} = v & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial z_{\lambda}}{\partial \nu} = 0 & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ z_{\lambda}(x, t_0) = 0 & \text{pour } x \in \Omega. \end{array} \right.$$

Comme le problème satisfait par z_{λ} est indépendant de λ on déduit que z_{λ} est indépendant de λ , donc $z_{\lambda} = z \quad \forall \lambda \in \mathbb{R}$, avec $z(x, t)$ l'unique solution du système

$$\left\{ \begin{array}{ll} \frac{\partial z}{\partial t} - \Delta_x z = v & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial z}{\partial \nu} = 0 & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ z(x, t_0) = 0 & \text{pour } x \in \Omega. \end{array} \right. \quad (3.6)$$

En passant à la limite dans (3.5) on trouve

$$\lim_{\lambda \rightarrow 0} \frac{E_2}{\lambda} = 2 \int_{\Omega} z(x, t_1) \cdot [y^u(x, t_1) - y_f(x)] dx. \quad (3.7)$$

De (3.4) et (3.7) on déduit que la dérivée directionnelle de J en u de direction v est donnée par

$$\frac{dJ}{dv}(u) = 2 \int_Q u(x, t)v(x, t) dxdt + 2 \int_{\Omega} z(x, t_1) [y^u(x, t_1) - y_f(x)] dx \quad (3.8)$$

Nous allons écrire cette expression sous la forme d'un produit scalaire en $L^2(Q)$ entre v et une fonction à trouver. Pour cela on multiplie l'équation principale de (3.6) par $p(x, t)$ avec p à choisir convenablement, et on intègre sur Q . Ceci donne

$$\int_Q \frac{\partial z}{\partial t} p dxdt - \int_Q \Delta_x z p dxdt = \int_Q vp dxdt$$

En faisant des intégrations par parties et en utilisant la formule de Green, nous obtenons :

$$\begin{aligned} & \int_{\Omega} z(x, t_1)p(x, t_1) dx - \int_{\Omega} z(x, t_0)p(x, t_0) dx + \int_Q \left(-z \frac{\partial p}{\partial t} + \langle \nabla_x z | \nabla_x p \rangle \right) dxdt \\ & - \int_{t_0}^{t_1} \int_{\partial\Omega} \frac{\partial z}{\partial \nu} p d\sigma dt = \int_Q vp dxdt \end{aligned}$$

En utilisant les condition limite et initiale de (3.6) pour z et encore une fois la formule de Green, nous obtenons

$$\int_{\Omega} z(x, t_1)p(x, t_1) dx - \int_Q z \left(\frac{\partial p}{\partial t} + \Delta_x p \right) dxdt + \int_{t_0}^{t_1} \int_{\partial\Omega} z \frac{\partial p}{\partial \nu} d\sigma dt = \int_Q vp dxdt \quad (3.9)$$

Nous choisissons $p : Q \rightarrow \mathbb{R}$ solution du système suivant :

$$\left\{ \begin{array}{ll} \frac{\partial p}{\partial t} + \Delta_x p = 0 & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial p}{\partial \nu} = 0 & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ p(x, t_1) = 2[y^u(x, t_1) - y_f(x)] & \text{pour } x \in \Omega. \end{array} \right. \quad (3.10)$$

Ce système s'appelle le **problème adjoint** associé au problème de contrôle optimal (3.2) ou (3.3).

Remarque 3.1. *Le système (3.10) n'est pas sous la forme habituelle pour une équation parabolique d'évolution (on a ici $\frac{\partial p}{\partial t} + \Delta_x p = 0$ au lieu d'avoir $\frac{\partial p}{\partial t} - \Delta_x p = 0$). Mais on a ici une condition au moment final $t = t_1$ au lieu de l'avoir pour le moment initial $t = t_0$. Pour montrer que le problème (3.10) est bien posé (c'est à dire, on a existence et unicité d'une*

solution) on peut se ramener à la forme habituelle d'une équation parabolique d'évolution en faisant le changement des variables suivant pour le temps t :

$$\tilde{t} = t_0 + t_1 - t.$$

En posant $\tilde{p}(x, t) = p(x, \tilde{t})$ alors le problème satisfait par \tilde{p} sera sous la forme habituelle

$$\left\{ \begin{array}{ll} \frac{\partial \tilde{p}}{\partial t} - \Delta_x \tilde{p} = 0 & \text{pour } (x, t) \in \Omega \times [t_0, t_1] \\ \frac{\partial \tilde{p}}{\partial \nu} = 0 & \text{pour } (x, t) \in \partial\Omega \times [t_0, t_1] \\ \tilde{p}(x, t_0) = 2[y^u(x, t_1) - y_f(x)] & \text{pour } x \in \Omega. \end{array} \right.$$

Comme on a existence et unicité d'une solution \tilde{p} de ce système, on a alors existence et unicité d'une solution p du problème adjoint (3.10).

On déduit de (3.9) et (3.10) :

$$2 \int_{\Omega} z(x, t_1)[y^u(x, t_1) - y_f(x)] dx = \int_Q vp dxdt$$

ce qui avec (3.8) donne

$$\frac{dJ}{dv}(u) = \int_Q (2u + p)v dxdt = \langle 2u + p \mid v \rangle_{L^2(Q)}$$

et ceci pour tous $u, v \in L^2(Q)$.

Nous avons donc montré :

Lemme 3.1. J est dérivable Gâteaux sur $L^2(Q)$ et on a

$$\nabla J(u) = 2u + p$$

où p est l'unique solution du problème adjoint (3.10).

Ceci nous permet d'utiliser une méthode numérique du type gradient pour trouver une approximation numérique du problème du contrôle optimal (voir Chapitre 4).

D'autre part, ce résultat nous donne les conditions nécessaires d'optimalité suivantes (c'est le **principe de minimum de Pontryagin** pour le problème de contrôle optimal (3.2) ou (3.3)) :

Théorème 3.1. *Si u^* est une solution de (3.2) ou (3.3) alors elle satisfait le système d'optimalité suivant : il existe $y^*, p^* : Q \rightarrow \mathbb{R}$ tels que*

$$\frac{\partial y^*}{\partial t} - \Delta_x y^* = u^* \quad \text{pour } (x, t) \in Q \quad (3.11)$$

$$\frac{\partial y^*}{\partial \nu} = g \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.12)$$

$$y^*(x, t_0) = y_0(x) \quad \text{pour } x \in \Omega \quad (3.13)$$

$$\frac{\partial p^*}{\partial t} + \Delta_x p^* = 0 \quad \text{pour } (x, t) \in Q \quad (3.14)$$

$$\frac{\partial p^*}{\partial \nu} = 0 \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.15)$$

$$p^*(x, t_1) = 2[y^*(x, t_1) - y_f(x)] \quad \text{pour } x \in \Omega \quad (3.16)$$

$$-2u^* - p^* \in N_{U_{ad}}(u^*) \quad (3.17)$$

où $N_{U_{ad}}(u^*)$ est le cône normal à U_{ad} en u^* dans l'espace de Hilbert $L^2(Q)$.

Remarque 3.2. *Comme nous avons fait l'hypothèse*

$$U_{ad} = \{v \in L^2(Q), \quad v(x, t) \in \mathcal{U}, \text{ p.p. } (x, t) \in Q\}$$

où \mathcal{U} est un intervalle fermé dans \mathbb{R} alors on peut montrer exactement comme dans la partie contrôle optimale des EDO que (3.17) est équivalent à :

$$-2u^*(t) - p^*(t) \in N_{\mathcal{U}}(u^*(t)), \quad \text{p.p. } t \in]t_0, t_1[$$

où $N_{\mathcal{U}}(u^*(t))$ est le cône normal à \mathcal{U} en $u^*(t)$ dans l'espace de Hilbert \mathbb{R} .

Remarque 3.3. *Dans le cas particulier $\mathcal{U} = \mathbb{R}$ (donc $U_{ad} = L^2(Q)$) on a que (3.17) devient*

$$2u^* + p^* = 0.$$

On peut alors exprimer p^* en fonction de u^* par la relation

$$p^* = -2u^*$$

et en remplaçant dans le système d'optimalité (3.11) - (3.17) nous obtenons

$$\frac{\partial y^*}{\partial t} - \Delta_x y^* = u^* \quad \text{pour } (x, t) \in Q \quad (3.18)$$

$$\frac{\partial y^*}{\partial \nu} = g \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.19)$$

$$y^*(x, t_0) = y_0(x) \quad \text{pour } x \in \Omega \quad (3.20)$$

$$\frac{\partial u^*}{\partial t} + \Delta_x u^* = 0 \quad \text{pour } (x, t) \in Q \quad (3.21)$$

$$\frac{\partial u^*}{\partial \nu} = 0 \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.22)$$

$$u^*(x, t_1) = y_f(x) - y^*(x, t_1) \quad \text{pour } x \in \Omega. \quad (3.23)$$

Ce système peut se résoudre numériquement en utilisant une méthode de point fixe.
 Une autre méthode de résolution est d'éliminer u^* en remplaçant u^* donné par (3.18) dans les autres équation pour obtenir le système suivant d'inconnue y :

$$\frac{\partial^2 y^*}{\partial t^2} - \Delta_x^2 y^* = 0 \quad \text{pour } (x, t) \in Q \quad (3.24)$$

$$\frac{\partial y^*}{\partial \nu} = g \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.25)$$

$$\frac{\partial}{\partial \nu} \frac{\partial y^*}{\partial t} - \frac{\partial}{\partial \nu} \Delta_x y^* = 0 \quad \text{pour } (x, t) \in \partial\Omega \times]t_0, t_1[\quad (3.26)$$

$$y^*(x, t_0) = y_0(x) \quad \text{pour } x \in \Omega \quad (3.27)$$

$$\frac{\partial y^*}{\partial t}(x, t_1) - \Delta_x y^*(x, t_1) + y^*(x, t_1) = y_f(x) \quad \text{pour } x \in \Omega. \quad (3.28)$$

Ce dernier système peut aussi se résoudre par des méthodes numériques.

Chapitre 4

Méthodes numériques

Nous donnons dans ce chapitre des méthodes pour l'approximation numérique de la(les) solution(s) du problème de contrôle optimal associé à un système EDO général, sans contraintes sur l'état final (voir problème (2.19)).

Dans la pratique il y a deux types de méthodes qui sont utilisées :

1. Méthodes basées sur la résolution du système d'optimalité
2. Méthodes basées sur la minimisation de la fonction coût J .

4.1 Méthodes basées sur la résolution du système d'optimalité

On suppose dans cette section que

$$U_{ad} = \{v \in L^2, \quad v(t) \in \mathcal{U}, \text{ p.p. } t \in]t_0, t_1[\}$$

avec $\mathcal{U} \subset \mathbb{R}^m$ un ensemble fermé et convexe.

Rappelons que nous avons le système d'optimalité (2.28) - (2.32) pour le problème de contrôle (2.19).

Rappelons que grâce à l'hypothèse faite ici sur U_{ad} (2.32) peut être remplacé par (2.38) ; nous supposons dans cette section que (2.38) permet de trouver pour presque tout $t \in]t_0, t_1[$ le contrôle $u^*(t)$ en fonction de $t, x^*(t)$ et $p^*(t)$, c'est à dire , on peut trouver une fonction $\Phi : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ telle qu'on a l'équivalence

(2.38) \iff (4.1) avec

$$u^*(t) = \Phi(t, x^*(t), p^*(t)), \quad \forall t \in]t_0, t_1[. \quad (4.1)$$

Alors en remplaçant u^* donné par (4.1) en (2.28) - (2.31) on arrive au système (pour

simplicité on va rénoter x^* par x et p^* par p) :

$$x'(t) = f_1(t, x(t), p(t)), \quad t \in [t_0, t_1] \quad (4.2)$$

$$p'(t) = g_1(t, x(t), p(t)), \quad t \in [t_0, t_1] \quad (4.3)$$

$$x(t_0) = x^0 \quad (4.4)$$

$$p(t_1) = \nabla\varphi(x(t_1)) \quad (4.5)$$

avec $f_1 : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ donnée par

$$f_1(t, \tilde{x}, \tilde{p}) = f(t, \tilde{x}, \Phi(t, \tilde{x}, \tilde{p})), \quad \forall (t, \tilde{x}, \tilde{p}) \in [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$$

et $g_1 : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ donnée par

$$g_1(t, \tilde{x}, \tilde{p}) = -\nabla_x f(t, \tilde{x}, \Phi(t, \tilde{x}, \tilde{p})) \cdot \tilde{p} - \nabla_x L(t, \tilde{x}, \Phi(t, \tilde{x}, \tilde{p})), \quad \forall (t, \tilde{x}, \tilde{p}) \in [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n.$$

Le but est de résoudre numériquement le système (4.2) - (4.5). Observons qu'on a $2n$ équations différentielles avec $2n$ inconnues mais que nous avons un "mélange" entre des conditions initiales pour x et des conditions finales pour p . Ce n'est pas un problème de Cauchy et on ne peut pas utiliser des algorithmes (ou logiciels) classiques pour la résolution numérique de ce système.

On utilisera une méthode appelée **méthode de tir**. L'idée de cette méthode est la suivante : pour un $\eta \in \mathbb{R}^n$ arbitraire on va remplacer la condition "finale" en p qui est (4.5) par une condition initiale artificielle en p qui est

$$p(t_0) = \eta. \quad (4.6)$$

On va considérer alors le problème de Cauchy (4.2), (4.3), (4.4) et (4.6). Comme la solution de ce problème dépend de la variable η on va noter les inconnues de ce système par x^η et p^η . Nous supposons qu'on sait résoudre numériquement ce système et donc qu'on peut calculer x^η et p^η pour tout $\eta \in \mathbb{R}^n$. Il restera alors de trouver $\eta \in \mathbb{R}^n$ tel que l'égalité (4.5) soit satisfaite, ce qui nous donne l'équation algébrique avec inconnue η :

$$p^\eta(t_1) - \nabla\varphi(x^\eta(t_1)) = 0 \quad (4.7)$$

En fait l'équation (4.7) est un système avec n équations et n inconnues (les composantes de η). Pour la résolution numérique de (4.7) on peut utiliser une méthode connue comme par exemple la dichotomie (uniquement pour $n = 1$), la méthode des approximations successives ou la méthode de Newton.

Remarque 4.1. *Ce type de méthode basé sur la résolution du système d'optimalité peut s'étendre au cas avec contraintes sur l'état final (voir sous-section 2.3.2). Le cas le plus fréquent est le cas où la condition $x(t_1) \in C_1$ consiste en un nombre $p \leq n$ d'équations algébriques satisfaites par $x(t_1)$. Dans ce cas la condition (2.42) consiste en $n - p$ équations satisfaites par $x(t_1)$ et $p(t_1)$.*

4.2 Méthodes basées sur la minimisation numérique de la fonction coût

4.2.1 Rappel sur la notion de projection

Rappelons d'abord la notion de projection dans un espace de Hilbert.

Définition 4.1. Soit H un espace de Hilbert sur \mathbb{R} , $M \subset H$ un sous ensemble et $x \in H$. On dit que $x^* \in M$ est la **projection** de x sur M (on va noter $x^* = \mathcal{P}_M(x)$) si

$$\|x - x^*\|_H = \min_{y \in M} \|x - y\| \quad \text{c'est à dire : } \|x - x^*\|_H \leq \|x - y\|, \quad \forall y \in M.$$

On admet le résultat suivant :

Proposition 4.1. Si M est un sous ensemble fermé et convexe de H alors on a l'existence et l'unicité de $x^* \in M$ tel que $x^* = \mathcal{P}_M(x)$ avec en plus

$$\langle x - x^* | y - x^* \rangle \leq 0, \quad \forall y \in M.$$

Remarque 4.2. De la définition et de la proposition ci-dessus on déduit

1.

$$x \in M \Leftrightarrow \mathcal{P}_M(x) = x.$$

2.

$$x^* = \mathcal{P}_M(x) \Leftrightarrow x - x^* \in N_M(x^*).$$

Exemple 4.1. Soit $H = \mathbb{R}$ et M un intervalle fermé de \mathbb{R} ; notons $a = \min(M) \in [-\infty, +\infty[$ et $b = \max(M) \in]-\infty, +\infty]$. Nous pouvons montrer facilement (exercice) que

$$\mathcal{P}_M(x) = \begin{cases} a & \text{si } x < a \quad (\text{cette partie est inexistente si } a = -\infty) \\ x & \text{si } a \leq x \leq b \\ b & \text{si } x > b \quad (\text{cette partie est inexistente si } b = +\infty). \end{cases}$$

Par exemple si $M = [0, +\infty[$ alors $\mathcal{P}_M(x) = x^+ = \max\{x, 0\}$, $\forall x \in \mathbb{R}$.

Exemple 4.2. Soit $H = \mathbb{R}^p$ et $M = M_1 \times M_2 \times \dots \times M_p$ avec M_1, M_2, \dots, M_p des intervalles fermés de \mathbb{R} . On admet le résultat suivant : pour tout $x = (x_1, x_2, \dots, x_p) \in \mathbb{R}^p$ on a

$$\mathcal{P}_M(x) = (P_{M_1}(x_1), P_{M_2}(x_2), \dots, P_{M_p}(x_p))$$

où \mathcal{P}_{M_j} représente la projection sur M_j en \mathbb{R} pour tout $j = 1, 2, \dots, p$.

Exemple 4.3. Soit $H = L^2(]t_0, t_1[, \mathbb{R}^m)$ et

$$M = \{u \in L^2(]t_0, t_1[, \mathbb{R}^m), u(t) \in \mathcal{U}, \text{ p.p. } t \in]t_0, t_1[\}$$

avec \mathcal{U} un sous-ensemble fermé et convexe de \mathbb{R}^m . On admet le résultat suivant : pour tout $v \in L^2(]t_0, t_1[, \mathbb{R}^m)$ on a

$$\mathcal{P}_M(v)(t) = \mathcal{P}_\mathcal{U}(v(t)), \quad \text{p.p. } t \in]t_0, t_1[.$$

où \mathcal{P}_M désigne la projection sur M en L^2 et $\mathcal{P}_\mathcal{U}$ désigne la projection sur \mathcal{U} en \mathbb{R}^m .

Par exemple si $M = \{u \in L^2(]t_0, t_1[), u(t) \geq 0 \text{ p.p. } t \in]t_0, t_1[\}$ alors pour tout $v \in L^2(]t_0, t_1[)$ on a $\mathcal{P}_M(v)(t) = v^+(t)$, p.p. $t \in]t_0, t_1[$ (car $\mathcal{P}_M(v)(t) = \mathcal{P}_{[0, +\infty[}(v(t)) = v^+(t)$).

4.2.2 Algorithmes de minimisation

Nous considérons encore notre problème de contrôle optimal qui s'écrit (2.19) avec J donnée par (2.20).

La méthode proposée ici consiste à minimiser numériquement la fonction J sur l'espace de Hilbert L^2 . On utilisera la **méthode de gradient** pour la minimisation sur un espace de Hilbert, qui est l'extension naturelle de la même méthode sur \mathbb{R}^n . Rappelons que nous avons l'expression de $\nabla J(u)$, voir Lemme 2.3.

L'idée est de construire par récurrence une suite $u_k \in U_{ad}$ qui va approcher la solution u^* . On va traiter séparément les cas où U_{ad} est l'espace entier L^2 et le cas contraire.

Cas1. $U_{ad} = L^2$.

On se donne $u_0 \in L^2$ (par exemple $u_0 = 0$) et pour tout u_k donné avec $k \in \mathbb{N}$ on construit u_{k+1} par la relation

$$u_{k+1} = u_k - \rho_k \nabla J(u_k). \quad (4.8)$$

avec un test de convergence (tester si u_{k+1} et "proche" de u_k) et un test d'arrêt avec "Échec" dans le cas où k est trop grand. Ici $\rho_k > 0$ est un facteur à choisir suivant la méthode utilisée (on va détailler ceci ultérieurement).

L'algorithme général est le suivant :

pas 1. On choisit $u_0 \in L^2$ et on pose $k = 0$.

pas 2. Résoudre (numériquement) le problème de Cauchy : trouver x_k solution de

$$\begin{cases} x'_k & = f(t, x_k, u_k) \\ x_k(t_0) & = x^0 \end{cases} \quad (4.9)$$

ensuite résoudre le problème de Cauchy : trouver p_k solution de

$$\begin{cases} p'_k & = -\nabla_x f(t, x_k, u_k)p_k - \nabla_x L(t, x_k, u_k) \\ p_k(t_1) & = \nabla \varphi(x_k(t_1)) \end{cases} \quad (4.10)$$

et poser

$$g_k = \nabla_u f(t, x_k, u_k) p_k + \nabla_u L(t, x_k, u_k)$$

(g_k représente $\nabla J(u_k)$).

pas 3. On teste d'abord si $g_k = 0$ (en fait on teste si g_k est "proche" de 0); si OUI, on arrête l'algorithme avec "*Succes*" u_k étant la solution recherchée.

Si $g_k \neq 0$ alors on pose

$$u_{k+1} = u_k - \rho_k g_k$$

avec $\rho_k > 0$ à choisir (voir les détails dans la sous-section 4.2.3).

pas 4. Test : si u_{k+1} est "proche" de u_k alors sortir de l'algorithme avec "*Succes*", u_{k+1} étant la solution recherchée

sinon, faire $k = k + 1$ et retour au **pas 2**

(Il faut prévoir ici la possibilité de sortir de l'algorithme avec message d'erreur du type "algorithme divergent" si k est trop grand).

Cas 2. $U_{ad} = \{u \in L^2(]t_0, t_1[, \mathbb{R}^m), \quad u(t) \in \mathcal{U}, \text{ p.p. } t \in]t_0, t_1[\}$
avec $\mathcal{U} \subset \mathbb{R}^m$ un ensemble convexe et fermé.

On peut utiliser dans ce cas **la méthode du gradient projeté**. L'algorithme associé consiste à reprendre l'algorithme précédent (celui du **Cas 1.**) avec 2 changements :

- au **pas 1.** remplacer " $u_0 \in L^2$ " par " $u_0 \in U_{ad}$ ".

- au **pas 3.** remplacer " $u_{k+1} = u_k - \rho_k g_k$ " par " $u_{k+1} = \mathcal{P}_{U_{ad}}(u_k - \rho_k g_k)$ " où $\mathcal{P}_{U_{ad}}$ est la projection sur U_{ad} dans l'espace de Hilbert L^2 .

Comme conséquence de l'Exemple 4.3 nous avons alors

$$u_{k+1}(t) = \mathcal{P}_{\mathcal{U}}(u_k(t) - \rho_k g_k(t)), \quad \text{p.p. } t \in]t_0, t_1[$$

où $\mathcal{P}_{\mathcal{U}}$ représente la projection sur \mathcal{U} en \mathbb{R}^m .

4.2.3 Choix des facteurs ρ_k

On donnera ici trois possibilités pour choisir les facteurs ρ_k dans le **pas 3** de l'algorithme précédent.

Pour simplifier l'écriture, pour tout $k \in \mathbb{N}$ fixé pour lequel on a calculé u_k et g_k on introduit une nouvelle fonction $J_k : [0, +\infty[\rightarrow \mathbb{R}$ donnée par

$$J_k(r) = J(u_k - r g_k), \quad \forall r \geq 0.$$

On admet que J_k est de classe C^1 et que sa dérivée est donnée par

$$J'_k(r) = - \langle g_k | \nabla J(u_k - r g_k) \rangle_{L^2}, \quad \forall r > 0.$$

Possibilité 1. Poser $\rho_k = \rho$, constante indépendante de k , avec ρ en général petit (par exemple $\rho = 1/10$). C'est l'algorithme du gradient à **pas fixe**. Le risque pour cette méthode est de choisir ρ trop grand, ce qui ferait que l'algorithme peut ne pas converger. D'autre part, si ρ est trop petit, l'algorithme peut être trop "couteux" en nombre d'opérations.

Possibilité 2. On fait une “recherche linéaire” pour trouver un $\rho_k > 0$ tel que $J_k(\rho_k) < J_k(0)$. On a alors l’algorithme suivant :

choisir $r > 0$ (par exemple $r = 1$)
 tant que ($J_k(r) \geq J_k(0)$)
 faire $r = r/2$
 fin “tant que”
 $\rho_k = r$.

Remarque 4.3. Comme $J'_k(0) = -\|g_k\|_{L^2}^2$ alors $J'_k(0) < 0$ si $g_k \neq 0$ (donc si l’algorithme n’a pas encore convergé). Alors il existe $r_0 > 0$ tel que $J_k(r) < J_k(0) \quad \forall r \in]0, r_0]$, donc l’algorithme ci-dessus va s’arrêter en un nombre fini de pas.

Possibilité 3. C’est une amélioration de la possibilité précédente. Il s’agit de chercher $r > 0$ et $s \in \mathbb{N}$ avec $s \geq 2$ tels que

$$J_k(0) > J_k(r) > J_k(2r) > \dots > J_k(sr) \quad \text{et} \quad J_k(sr) \leq J_k((s+1)r).$$

On posera alors $\rho_k = sr$. L’algorithme en est le suivant :

choisir $r > 0$ (par exemple $r = \frac{1}{10}$)
 tant que ($J_k(0) \leq J_k(r)$ ou $J_k(r) \leq J_k(2r)$)
 faire $r = r/2$
 fin “tant que”
 $\rho = 2r$.
 tant que ($J_k(\rho + r) < J_k(\rho)$)
 faire $\rho = \rho + r$
 fin “tant que”
 $\rho_k = \rho$.

Cet algorithme est une variante simplifiée de l’algorithme de **Armijo-Goldstein**.

Remarque 4.4. L’inconvénient principal de tous les algorithmes du type gradient est le fait qu’ils peuvent s’arrêter dans un “point” de minimum local pour J alors que nous voulons un minimum global. Si J est une fonction convexe (ce qui est difficile à montrer en général, sauf pour des cas très particuliers, comme par exemple f du type linéaire et L et φ du type quadratique) alors cet inconvénient disparaît, car dans ce cas tout point de minimum local est aussi global.

Chapitre 5

Contrôle optimal en feed-back ou à boucle fermée

5.1 Quelques généralités

Nous considérons ici le problème général de contrôle optimal sans contraintes sur l'état final, comme dans le Chapitre 2.

Rappelons le cadre mathématique : on se donne

$$L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$$

$$f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$$

$$\varphi : \mathbb{R}^n \mapsto \mathbb{R}, \quad x^0 \in \mathbb{R}^n, \quad \mathcal{U} \subset \mathbb{R}^m$$

et on cherche $x : [t_0, t_1] \rightarrow \mathbb{R}^n$ et $u : [t_0, t_1] \rightarrow \mathbb{R}^m$ tels que

$$\begin{cases} x'(t) = f(t, x(t), u(t)), & t \in [t_0, t_1] \\ x(t_0) = x^0 \end{cases} \quad (5.1)$$

et qui minimise

$$\int_{t_0}^{t_1} L(t, x(t), u(t)) dt + \varphi(x(t_1)). \quad (5.2)$$

Rappelons aussi la notation : $L^2 = L^2(]t_0, t_1[, \mathbb{R}^m)$. On note encore

$$U_{ad} = \{u \in L^2, \quad u(t) \in \mathcal{U} \text{ p.p. } t \in]t_0, t_1[\}.$$

avec $\mathcal{U} \subset \mathbb{R}^m$.

On suppose que pour tout $u \in U_{ad}$ on a l'existence et l'unicité d'une solution $x(t)$ (que nous pouvons noter $x^u(t)$) de (5.1).

On peut alors écrire notre problème de contrôle sous la forme

$$\min_{u \in U_{ad}} J(u) \quad (5.3)$$

avec

$$J(u) = \int_{t_0}^{t_1} L(t, x^u(t), u(t)) dt + \varphi(x^u(t_1)).$$

But de ce chapitre : supposons qu'on a un contrôle optimal u^* pour le problème (5.3) avec x^* l'état optimal correspondant. On veut alors trouver une fonction $\phi : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ telle que u^* s'écrit sous la forme

$$u^*(t) = \phi(t, x^*(t)), \quad \forall t \in [t_0, t_1].$$

On dira alors que le contrôle optimal u^* est donnée comme un **feed-back** par rapport à l'état optimal x^* .

Remarque : En général le principe de minimum de Pontryagin nous permet de trouver $u^*(t)$ comme une fonction de $t, x^*(t)$ mais aussi de l'état adjoint $p^*(t)$; on déduit que si on peut exprimer $p^*(t)$ comme une fonction de t et $x^*(t)$ alors on aura la représentation en feed-back souhaitée. Mais on ne pourra pas toujours exprimer $p^*(t)$ en fonction de t et $x^*(t)$, d'où la nécessité de donner une méthode générale pour avoir la représentation en feed-back.

Notation : Dans ce chapitre, pour tout $u \in U_{ad}$, $r \in [t_0, t_1]$ et $y \in \mathbb{R}^m$ nous notons par $x^{u,r,y}$ la solution de l'équation principale de (5.1) avec donnée "initiale" y en r ; donc $x^{u,r,y} : [t_0, t_1] \rightarrow \mathbb{R}^n$ est l'unique solution du problème de Cauchy :

$$\begin{cases} (x^{u,r,y})' = f(t, x^{u,r,y}, u), & t \in [t_0, t_1] \\ x(r) = y \end{cases} \quad (5.4)$$

Remarque 5.1. 1. Pour $r = t_0$ on peut utiliser la notation plus simple x^u au lieu de x^{u,t_0,x^0} .

2. Pour tous $r, s \in [t_0, t_1]$ nous avons sur $[t_0, t_1]$ l'égalité

$$x^{u,r,y} = x^{u,s,v} \quad \text{avec} \quad v = x^{u,r,y}(s)$$

(car les deux fonctions satisfont la même équation différentielle avec la même condition "initiale" : $x^{u,r,y}(s) = x^{u,s,v}(s) = v$).

3. Si u est définie uniquement sur un sous-intervalle de $[t_0, t_1]$ contenant r alors $x^{u,r,y}$ sera défini sur ce sous-intervalle.

Nous utilisons dans la suite la notation suivante :

$$U_{ad,r,s} = \{u|_{[r,s]}, \quad u \in U_{ad}\} \quad \text{pour tous } r, s \text{ avec } t_0 \leq r \leq s \leq t_1.$$

5.2 La fonction valeur

Définition 5.1. 1. On appelle **fonction valeur** pour le problème de contrôle (5.3) la fonction $V : [t_0, t_1] \times \mathbb{R}^m \rightarrow \mathbb{R}$ définie par

$$(r, y) \in [t_0, t_1] \times \mathbb{R}^m \rightarrow V(r, y) = \min_{u \in U_{ad}} \left\{ \int_r^{t_1} L(t, x^{u,r,y}(t), u(t)) dt + \varphi(x^{u,r,y}(t_1)) \right\}$$

(on supposera qu'un tel minimum existe et est un élément de \mathbb{R}). Donc $V(r, y)$ est le minimum de notre fonction coût, mais uniquement sur l'intervalle $[r, t_1]$, avec donnée initiale y en r .

On peut aussi écrire :

$$V(r, y) = \min_{u \in U_{ad}} \left\{ \int_r^{t_1} L(t, x(t), u(t)) dt + \varphi(x(t_1)), \quad x' = f(t, x, u) \text{ sur } [r, t_1], x(r) = y \right\}.$$

2. On dit que $u^* : [r, t_1] \rightarrow \mathbb{R}^m$ est une **réalisation** de $V(r, y)$ si $u^* \in U_{ad, r, t_1}$ et si pour tout $u \in U_{ad, r, t_1}$ on a

$$\int_r^{t_1} L(t, x^{u^*, r, y}(t), u^*(t)) dt + \varphi(x^{u^*, r, y}(t_1)) \leq \int_r^{t_1} L(t, x^{u, r, y}(t), u(t)) dt + \varphi(x^{u, r, y}(t_1)).$$

Remarque 5.2. 1. Si u^* est le contrôle optimal recherché pour le problème (5.3) alors

$$J(u^*) = V(t_0, x^0)$$

donc le contrôle u^* est une réalisation de $V(t_0, x^0)$.

2. Nous avons :

$$V(t_1, y) = \min_{u \in U_{ad}} \left\{ \int_{t_1}^{t_1} L(t, x^{u, r, y}(t), u(t)) dt + \varphi(x^{u, t_1, y}(t_1)) \right\} = \min_{u \in U_{ad}} \{\varphi(y)\}$$

donc

$$V(t_1, y) = \varphi(y), \quad \forall y \in \mathbb{R}^n. \quad (5.5)$$

Nous avons le résultat suivant :

Lemme 5.1. (principe de la programmation dynamique) : Soient r, s avec $t_0 \leq r \leq s \leq t_1$. Si $u^* \in U_{ad, r, t_1}$ est une réalisation de $V(r, y)$ alors $u^*|_{[s, t_1]}$ est une réalisation de $V(s, x^{u^*, r, y}(s))$.

Démonstration. Notons $z^* = x^{u^*, r, y}(s) \in \mathbb{R}^n$. On suppose par absurd que le résultat est faux. Il existe alors $\tilde{u} : [s, t_1] \rightarrow \mathcal{U}$ avec $\tilde{u} \in U_{ad, s, t_1}$ tel que

$$\int_s^{t_1} L(t, x^{\tilde{u}, s, z^*}(t), \tilde{u}(t)) dt + \varphi(x^{\tilde{u}, s, z^*}(t_1)) < \int_s^{t_1} L(t, x^{u^*, s, z^*}(t), u^*(t)) dt + \varphi(x^{u^*, s, z^*}(t_1)). \quad (5.6)$$

Remarquons que $x^{u^*, s, z^*} = x^{u^*, r, y}$ sur $[r, t_1]$.

On considère la fonction $\tilde{u}_* : [r, t_1] \rightarrow \mathbb{R}^n$ donnée par l'expression

$$\tilde{u}_*(t) = \begin{cases} u^*(t) & \text{si } t \in [r, s] \\ \tilde{u}(t) & \text{si } t \in]s, t_1]. \end{cases}$$

On observe qu'on a

$$x^{\tilde{u}_*, r, y}(t) = \begin{cases} x^{u^*, r, y}(t) & \text{si } t \in [r, s] \\ x^{\tilde{u}, s, z^*}(t) & \text{si } t \in]s, t_1]. \end{cases}$$

En ajoutant $\int_r^s L(t, x^{u^*, r, y}(t), t) dt$ aux deux cotés de (5.6) on obtient

$$\int_r^{t_1} L(t, x^{\tilde{u}_*, r, y}(t), \tilde{u}_*(t)) dt + \varphi(x^{\tilde{u}_*, r, y}(t_1)) < \int_r^{t_1} L(t, x^{u^*, r, y}(t), u^*(t)) dt + \varphi(x^{u^*, r, y}(t_1))$$

ce qui contredit le fait que u^* est une réalisation de $V(r, y)$. Ceci finit la preuve. \square

Comme conséquence de ce résultat nous avons :

Corollaire 5.1. *Soient r, s tels que $t_0 \leq r \leq s \leq t_1$. Si u^* est une réalisation de $V(r, y)$ alors on a*

$$V(r, y) = \int_r^s L(t, x^{u^*, r, y}(t), u^*(t)) dt + V(s, x^{u^*, r, y}(s)).$$

Démonstration. On décompose $V(r, y)$ sous la forme

$$V(r, y) = \int_r^s L(t, x^{u^*, r, y}(t), u^*(t)) dt + \int_s^{t_1} L(t, x^{u^*, r, y}(t), u^*(t)) dt + \varphi(x^{u^*, r, y}(t_1)).$$

Comme du Lemme 5.1 $u^*_{|[s, t_1]}$ est une réalisation de $V(s, z^*)$ (avec encore $z^* = x^{u^*, r, y}(s)$) alors $\int_s^{t_1} L(t, x^{u^*, r, y}(t), u^*(t)) dt + \varphi(x^{u^*, r, y}(t_1)) = V(s, z^*)$ (car $x^{u^*, r, y} = x^{u^*, s, z^*}$) ce qui donne le résultat. \square

On a le théorème suivant :

Théorème 5.1. *(principe d'optimalité de Bellman) Pour tous r, s avec $t_0 \leq r \leq s \leq t_1$ et pour tout $y \in \mathbb{R}^n$ on a*

$$V(r, y) = \inf_{u \in U_{ad, r, s}} \left\{ \int_r^s L(t, x^{u, r, y}(t), u(t)) dt + V(s, x^{u, r, y}(s)) \right\}. \quad (5.7)$$

Démonstration. On fera la preuve dans le cas particulier où il existe une réalisation $u^* \in U_{ad, r, t_1}$ de $V(r, y)$.

Le Corollaire 5.1 nous donne

$$V(r, y) = \int_r^s L(t, x^{u^*, r, y}(t), u(t)) dt + V(s, x^{u^*, r, y}(s))$$

ce qui nous donne immédiatement la partie " \geq " de (5.7).

Montrons maintenant la partie " \leq " de (5.7). Nous avons

$$V(r, y) \leq \int_r^{t_1} L(t, x^{u, r, y}(t), u(t)) dt + \varphi(x^{u, r, y}(t_1)), \quad \forall u \in U_{ad, r, t_1}$$

donc pour tout $u \in U_{ad,r,t_1}$ on a

$$V(r, y) \leq \int_r^s L(t, x^{u,r,y}(t), u(t)) dt + \int_s^{t_1} L(t, x^{u,s,z}(t), u(t)) dt + \varphi(x^{u,s,z}(t_1)) \quad (5.8)$$

où on a noté $z = x^{u,r,y}(s)$ (on utilise le fait que $x^{u,r,y} = x^{u,s,z}$). On fixe d'abord $u|_{]r,s[}$ et on passe à l'infimum en (5.8) par rapport à $u|_{]s,t_1[}$, ce qui nous donne

$$V(r, y) \leq \int_r^s L(t, x^{u,r,y}(t), u(t)) dt + V(s, x^{u,r,y}(s)).$$

En passant, dans cette dernière inégalité à l'infimum par rapport à $u|_{]r,s[}$ on obtient le résultat. \square

5.3 L'équation de Hamilton-Jacobi-Bellman

On a la définition suivante :

Définition 5.2. 1. On appelle **pre-Hamiltonian** du problème de contrôle (5.3) la fonction $\underline{H} : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{U} \rightarrow \mathbb{R}$ définie par

$$(t, \tilde{x}, \tilde{p}, \tilde{u}) \rightarrow \underline{H}(t, \tilde{x}, \tilde{p}, \tilde{u}) = L(t, \tilde{x}, \tilde{u}) + \langle f(t, \tilde{x}, \tilde{u}) | \tilde{p} \rangle.$$

2. On appelle **Hamiltonian** du problème de contrôle (5.3) la fonction (quand elle existe)

$$H : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \text{ définie par}$$

$$(t, \tilde{x}, \tilde{p}) \rightarrow H(t, \tilde{x}, \tilde{p}) = \inf_{\tilde{u} \in \mathcal{U}} \underline{H}(t, \tilde{x}, \tilde{p}, \tilde{u})$$

Remarque : parfois dans la littérature on appelle Hamiltonien ce qui est appelé pre-Hamiltonian dans ce cours.

Nous avons le résultat suivant :

Théorème 5.2. Supposons que le Hamiltonien donné en Définition 5.2 existe et supposons que la fonction valeur $V(t, y)$ est de classe C^1 sur $[t_0, t_1] \times \mathbb{R}^n$. Alors V satisfait l'EDP suivante, appelée **équation de Hamilton-Jacobi-Bellman** (on l'appelle aussi l'équation HJB) :

$$\frac{\partial V}{\partial t} + H(t, y, \nabla_y V) = 0. \quad (5.9)$$

Démonstration. On fait la preuve dans le cas particulier où il existe toujours une réalisation de la fonction valeur. Soient $(t, y) \in [t_0, t_1] \times \mathbb{R}^n$ fixées. Montrons d'abord la partie " \leq " de (5.9). Soit u une réalisation de $V(t, y)$. Le Corollaire 5.1 nous permet d'écrire pour tout $\Delta t > 0$ "petit" :

$$V(t, y) = \int_t^{t+\Delta t} L(\tau, x^{u,t,y}(\tau), u(\tau)) d\tau + V(t + \Delta t, x^{u,t,y}(t + \Delta t)).$$

En faisant un développement limité par rapport à Δt à l'ordre 1 on obtient

$$V(t, y) = L(t, y, u(t))\Delta t + V(t, y) + \left[\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | (x^{u,t,y})'(t) \rangle \right] \Delta t + o(\Delta t)$$

c'est à dire

$$V(t, y) = L(t, y, u(t))\Delta t + V(t, y) + \left[\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | f(t, y, u(t)) \rangle \right] \Delta t + o(\Delta t)$$

On simplifie par $V(t, y)$, on divise par Δt et on passe à la limite $\Delta t \rightarrow 0$, ce qui donne

$$\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | f(t, y, u(t)) \rangle + L(t, y, u(t)) = 0.$$

Comme $\langle \nabla_y V(t, y) | f(t, y, u(t)) \rangle + L(t, y, u(t)) = \underline{H}(t, y, \nabla_y V(t, y), u(t)) \geq H(t, y, \nabla_y V(t, y))$ on obtient la partie " \leq " de (5.9).

Pour obtenir " \geq " nous fixons $u_0 \in \mathcal{U}$ arbitraire et nous écrivons, comme conséquence du Théorème 5.1 :

$$V(t, y) \leq \int_t^{t+\Delta t} L(s, x^{u_0,t,y}(s), u_0) ds + V(t + \Delta t, x^{u_0,t,y}(t + \Delta t))$$

(il faut voir ici u_0 comme une fonction contrôle constante sur l'intervalle $[t, t + \Delta t]$). Si on développe encore la partie de droite en Δt on obtient

$$V(t, y) \leq L(t, y, u_0)\Delta t + V(t, y) + \left[\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | (x^{u_0,t,y})'(t) \rangle \right] \Delta t + o(\Delta t).$$

c'est à dire

$$V(t, y) \leq L(t, y, u_0)\Delta t + V(t, y) + \left[\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | f(t, y, u_0) \rangle \right] \Delta t + o(\Delta t).$$

On simplifie par $V(t, y)$, on divise par Δt et on fait $\Delta t \rightarrow 0$, ce qui donne

$$\frac{\partial V}{\partial t}(t, y) + \langle \nabla_y V(t, y) | f(t, y, u_0) \rangle + L(t, y, u_0) \geq 0$$

et ceci est vraie pour tout $u_0 \in \mathcal{U}$. En passant à l'infimum sur u_0 , on obtient la partie " \geq " de (5.9), ce qui finit la preuve. \square

Remarque 5.3. On a vu que la fonction valeur V satisfait aussi (5.5). Alors V est une solution du système suivant (équation HJB et condition "finale" en $t = t_1$) :

$$\begin{cases} \frac{\partial V}{\partial t} + H(t, y, \nabla_y V) = 0, & \forall (t, y) \in]t_0, t_1[\times \mathbb{R}^n \\ V(t_1, y) = \varphi(y), & \forall y \in \mathbb{R}^n. \end{cases} \quad (5.10)$$

5.4 L'obtention du contrôle en feed-back

Le résultat suivant nous dit comment obtenir un contrôle en feed-back si on connaît une solution du système (5.10).

Théorème 5.3. *Soit $V(t, y)$ une solution de classe C^1 du système (5.10). Supposons que pour tout $(t, y) \in [t_0, t_1] \times \mathbb{R}^n$ il existe $\hat{U}(t, y) \in \mathcal{U}$ solution du problème d'optimisation*

$$\min_{w \in \mathcal{U}} \{L(t, y, w) + \langle \nabla_y V(t, y) | f(t, y, w) \rangle\}$$

c'est à dire, $\hat{U}(t, y) \in \mathcal{U}$ est tel que

$$L(t, y, \hat{U}(t, y)) + \langle \nabla_y V(t, y) | f(t, y, \hat{U}(t, y)) \rangle \leq L(t, y, w) + \langle \nabla_y V(t, y) | f(t, y, w) \rangle, \quad \forall w \in \mathcal{U}.$$

(donc $H(t, y, \nabla_y V(t, y)) = \underline{H}(t, y, \nabla_y V(t, y), \hat{U}(t, y))$). Alors \hat{U} est un contrôle en feed-back du problème de contrôle (5.3), c'est à dire si on note par x^ l'unique solution du problème de Cauchy*

$$\begin{cases} (x^*)'(t) = f(t, x^*(t), \hat{U}(t, x^*(t))), & t \in [t_0, t_1] \\ x^*(t_0) = x^0 \end{cases} \quad (5.11)$$

et on pose $u^(t) = \hat{U}(t, x^*(t))$, $\forall t \in]t_0, t_1[$ alors x^* est une trajectoire optimale de (5.1) avec u^* comme contrôle optimal. On a donc $x^* = x^{u^*, t_0, x^0}$ et*

$$J(u^*) \leq J(u), \quad \forall u \in U_{ad}.$$

Démonstration. Nous avons

$$J(u) = \int_{t_0}^{t_1} L(t, x^u(t), u(t)) dt + \varphi(x^u(t_1))$$

(rappel notation : $x^u = x^{u, t_0, x^0}$). Mais nous avons

$$\varphi(x^u(t_1)) = V(t_1, x^u(t_1)) = V(t_0, x^0) + \int_{t_0}^{t_1} \frac{d}{dt} V(t, x^u(t)) dt$$

car $x^u(t_0) = x^0$. Nous avons aussi

$$\frac{d}{dt} V(t, x^u(t)) = \frac{\partial V}{\partial t}(t, x^u(t)) + \langle \nabla_y V(t, x^u(t)) | (x^u)'(t) \rangle$$

donc

$$\frac{d}{dt} V(t, x^u(t)) = \frac{\partial V}{\partial t}(t, x^u(t)) + \langle \nabla_y V(t, x^u(t)) | f(t, x^u(t), u(t)) \rangle.$$

Ceci nous donne

$$J(u) = V(t_0, x^0) + \int_{t_0}^{t_1} \left[\frac{\partial V}{\partial t}(t, x^u(t)) + L(t, x^u(t), u(t)) + \langle \nabla_y V(t, x^u(t)) \mid f(t, x^u(t), u(t)) \rangle \right] dt$$

c'est à dire

$$J(u) = V(t_0, x^0) + \int_{t_0}^{t_1} \left[\frac{\partial V}{\partial t}(t, x^u(t)) + \underline{H}(t, x^u(t), \nabla_y V(t, x^u(t)), u(t)) \right] dt.$$

Pour tout $u \in U_{ad}$ nous introduisons la fonction $\psi_u : [t_0, t_1] \rightarrow \mathbb{R}$ donnée par

$$\psi_u(t) = \frac{\partial V}{\partial t}(t, x^u(t)) + \underline{H}(t, x^u(t), \nabla_y V(t, x^u(t)), u(t))$$

et on observe que l'égalité précédente s'écrit

$$J(u) = V(t_0, x^0) + \int_{t_0}^{t_1} \psi_u(t) dt. \quad (5.12)$$

Par la définition du Hamiltonian, nous avons

$$\psi_u(t) \geq \frac{\partial V}{\partial t}(t, x^u(t)) + H(t, x^u(t), \nabla_y V(t, x^u(t))) = 0$$

(car V satisfait l'équation de Hamilton-Jacobi-Bellman (5.10)).

On obtient alors de (5.12)

$$J(u) \geq V(t_0, x^0), \quad \forall u \in U_{ad}. \quad (5.13)$$

D'autre part nous avons

$$J(u^*) = V(t_0, x^0) + \int_{t_0}^{t_1} \psi_{u^*}(t) dt \quad (5.14)$$

Comme $u^*(t) = \hat{U}(t, x^*(t))$ alors par définition de \hat{U} on a

$$\underline{H}(t, x^*(t), \nabla_y V(t, x^*(t)), u^*(t)) = H(t, x^*(t), \nabla_y V(t, x^*(t)))$$

ce qui donne

$$\psi_{u^*}(t) = \frac{\partial V}{\partial t}(t, x^*(t)) + H(t, x^*(t), \nabla_y V(t, x^*(t))) = 0, \quad \forall t \in [t_0, t_1]$$

car V satisfait l'équation de Hamilton-Jacobi-Bellman. On en déduit avec (5.14)

$$J(u^*) = V(t_0, x^0)$$

ce qui avec (5.13) nous donne le résultat. □

Bibliographie

- [1] J. B. Hiriart-Urruty, *Les mathématiques du mieux faire, Vol. 2 : La commande optimale pour les débutants*, Opuscules
- [2] M. Bergounioux, *Optimisation et contrôle des systèmes linéaires*, Dunod, Paris, 2001
- [3] A.V. Fursikov, *Optimal Control of Distributed systems. Theory and applications*, Translations of Mathematical Monographs, 2000
- [4] E.B. Lee et L. Markus, *Foundations of Optimal Control Theory*, John Wiley and Sons, 1968
- [5] R. Pytlak, *Numerical Methods for Optimal Control Problems with State Constraints*, Lecture Notes in Mathematics, 1707.
- [6] E. Trélat, *Contrôle optimal : Théorie et applications* , Vuibert, 2005.