

Avant-propos

Dans le numéro 83 (juillet 2007) de MATAPLI,¹ J.-M. Coron et E. Trélat définissent ainsi en quelques lignes la théorie du contrôle :

Un système de contrôle est un système dynamique sur lequel on peut agir au moyen d'une commande ou contrôle. La théorie du contrôle analyse les propriétés de ce système dans le but de l'amener d'un état initial donné à un certain état final, en respectant éventuellement certains critères : c'est l'étape de réalisation de la commande. Les systèmes abordés sont multiples et leurs origines très diverses (mécanique, électricité, électronique, biologie, chimie, économie...). L'objectif peut être aussi de stabiliser le système pour le rendre insensible à certaines perturbations (stabilisation), ou encore de déterminer des solutions optimales pour un certain critère d'optimisation (contrôle optimal). Pour le modéliser, on peut avoir recours à des équations différentielles, intégrales, fonctionnelles, aux différences finies, aux dérivées partielles, etc. Pour cette raison la théorie du contrôle est à l'interconnexion de nombreux domaines mathématiques.

Tout le monde sait maintenir en équilibre un balai sur son doigt (problème du pendule inversé). En revanche il est beaucoup plus difficile de maintenir en équilibre sur son doigt un double pendule inversé, c'est-à-dire un système composé de deux balais l'un sur l'autre, surtout si l'on ferme les yeux. La théorie du contrôle permet pourtant de le faire. Mais pour réaliser effectivement cet équilibre, mieux vaut disposer d'un bon modèle et savoir résoudre des équations.

L'objectif de ce cours est de présenter des techniques d'analyse et de résolution de problèmes de *contrôle optimal* dans les équations différentielles ordinaires non linéaires. On considère un système dynamique sur lequel on peut agir au moyen d'un contrôle. Résoudre un problème de contrôle optimal consiste à déterminer une trajectoire avec le contrôle associé permettant par exemple de rejoindre une cible finale donnée, en partant d'une configuration initiale. Cette paire trajectoire-contrôle doit être optimale au sens qu'elle doit minimiser un certain critère. La valeur du critère pour une paire trajectoire-contrôle représente ce que coûte d'appliquer tel contrôle le long de la trajectoire. Avant de résoudre un tel problème de contrôle optimal, il faut d'abord savoir si la cible est atteignable. C'est le problème de contrôlabilité. Le Chapitre 1 introduit les notions essentielles de solution et de contrôlabilité des systèmes dynamiques contrôlés. Dans le Chapitre 2, nous donnons une définition générale d'un problème de contrôle optimal et nous présentons quelques exemples avec leurs solutions pour mettre en avant les différences de structure des contrôles optimaux. Dans le Chapitre 3, nous supposons l'absence de contraintes sur le contrôle. Dans ce cas, nous obtenons

1. http://smai.emath.fr/matapli_pdf/matapli83.pdf#page=60

des conditions nécessaires d'optimalité que nous pouvons qualifier de faible. Nous présentons ces conditions ainsi que les preuves s'appuyant sur le Chapitre 1. Nous donnons ensuite des exemples et des exercices dans les cadres classiques du calcul des variations et de la géométrie riemannienne. Dans le Chapitre 4, nous présentons le résultat le plus important du cours, c'est-à-dire le Principe du Maximum de Pontryagin (PMP). Le PMP fournit un ensemble de conditions nécessaires d'optimalité dont on peut tirer profit afin de définir des algorithmes de résolution numérique des problèmes de contrôle optimal. Ainsi, dans le Chapitre 5, nous présentons la méthode indirecte de tir s'appuyant sur le PMP.

Ce support de cours est destiné aux élèves de 2^e année du département Sciences du Numérique de l'ENSEEIH à Toulouse.

Table des matières

Avant-propos	i
I Contrôle optimal	1
Chapitre 1. Systèmes dynamiques contrôlés	3
1.1 Introduction	3
1.2 Solution d'un système contrôlé	5
1.3 Flot d'un système contrôlé	7
1.4 Application entrée/sortie	11
1.5 Ensemble accessible et contrôlabilité	13
1.6 Caractérisation hamiltonienne des contrôles singuliers	14
1.7 Calcul de contrôles singuliers	18
Chapitre 2. Problèmes de contrôle optimal	21
2.1 Introduction	21
2.2 Formulations	22
2.3 Exemples	24
Chapitre 3. Principe du maximum faible	33
3.1 Mayer avec conditions terminales	33
3.2 Lagrange avec conditions aux limites simples	35
3.3 Bolza à temps final libre et conditions terminales	37
3.4 Condition de Legendre et conservation du hamiltonien	41
3.5 Exemples	43
3.5.1 Simple intégrateur et fonction de tir	43
3.5.2 Calcul des variations	45
3.5.3 Principe fondamental de la dynamique	48
3.6 Condition du second ordre faible : cas Lagrange	51
3.7 Exercices	55
Chapitre 4. Principe du maximum de Pontryagin	61
4.1 Le principe du maximum	61
4.2 Exemples de maximisation de hamiltoniens	64
4.3 Exemples de conditions de transversalité	67
4.4 Cadre hamiltonien et origine symplectique	70
4.5 Problèmes affines en la commande	74
4.6 Exercices	76

Chapitre 5. Méthode indirecte de tir	79
5.1 Introduction par l'exemple	80
5.2 La démarche générale : condition aux limites simples	81
5.3 Exemple avec conditions aux limites sous-dimensionnées	83
5.4 Tir multiple de stabilité	85
5.5 Tir simple avec détections des commutations	86
5.6 Difficultés numériques des méthodes de tir simple	89
5.7 Tir multiple de structure	91
Chapitre 6. Corrections des exercices	95
 II Annexes	 99
Chapitre A. Théorèmes d'analyse et optimisation	101
A.1 Théorèmes des accroissements finis et applications	101
A.2 Dérivabilité dans le cas fonctionnel	105
A.3 Théorème du point fixe	109
A.4 Application inversion	110
A.5 Théorème d'inversion locale	113
A.6 Théorème des fonctions implicites	117
A.7 Conditions nécessaires d'optimalité	118
Chapitre B. Calcul des variations	121
B.1 Introduction	121
B.2 Solutions des problèmes variationnels	124
B.3 Équation de Euler-Lagrange	128
B.4 Condition de Legendre	135
B.5 Extrémum libre et contraintes affines	139
B.5.1 Extrémum libre	139
B.5.2 Extrémum lié : contraintes affines	144
B.5.3 "Two-norm discrepancy"	145
B.6 Condition de Jacobi	150
Bibliographie	153
Index	159

Première partie

Contrôle optimal

Systèmes dynamiques contrôlés

1.1	Introduction	3
1.2	Solution d'un système contrôlé	5
1.3	Flot d'un système contrôlé	7
1.4	Application entrée/sortie	11
1.5	Ensemble accessible et contrôlabilité	13
1.6	Caractérisation hamiltonienne des contrôles singuliers	14
1.7	Calcul de contrôles singuliers	18

L'analyse d'un problème de contrôle optimal se déroule par étape : avant de déterminer une trajectoire optimale joignant un état initial à une cible, il faut s'assurer que cette cible soit atteignable (ou accessible). L'ensemble des états atteignables à partir d'un état initial donné est appelé *ensemble accessible*. Si toute cible se trouve dans cet ensemble, le problème est dit *contrôlable*. La *contrôlabilité* d'un système dynamique contrôlé est donc une propriété centrale en contrôle optimal. Elle peut être globale ou locale. De plus, cette notion de contrôlabilité est liée à celle d'*application entrée/sortie* puisque l'ensemble accessible est l'image de l'ensemble des lois de commande (ou contrôle) admissibles par l'application entrée/sortie. Une dernière notion centrale concernant les systèmes dynamiques contrôlés est celle de *contrôle singulier*. En effet, les contrôles singuliers, en partie, déterminent la frontière de l'ensemble accessible. Un point remarquable est que nous pouvons donner une caractérisation hamiltonienne de ces contrôles singuliers, et dans certains cas particuliers fondamentaux, exhiber un véritable hamiltonien et une structure symplectique, responsable de la géométrie sous-jacente. C'est ce que nous faisons dans ce chapitre.

1.1 Introduction

Nous considérons des systèmes dynamiques régis par des équations différentielles ordinaires sur lesquelles nous pouvons agir au cours du temps par l'intermédiaire d'une commande. Ces systèmes que l'on peut commander sont au coeur de l'automatique, mais aussi, et plus généralement, au coeur de la théorie du contrôle optimal.

D'après la description de V. Arnold [5], la théorie des équations différentielles ordinaires (sans commande) est l'un des principaux instruments des mathématiques. Elle permet d'étudier quantité de processus d'évolution *déterministes*, *finis* et *différentiables*. On dit qu'un processus est déterministe si toute son évolution et son passé sont définis de façon unique par son état présent. L'ensemble de tous les états de ce processus s'appelle l'*espace des phases*. Ainsi, en mécanique classique, on étudie le mouvement de systèmes dont l'avenir et le passé sont univoquement définis par les positions et les vitesses initiales de tous les

points du système. L'espace des phases d'un tel système mécanique est un ensemble dont les éléments sont constitués par l'ensemble des positions et vitesses de tous les points du système considéré. En mécanique quantique, le mouvement des particules n'est pas un processus déterministe. Un processus est de dimension finie s'il en est de même de son espace des phases, c'est-à-dire si est fini le nombre des paramètres indispensables à la description de son état. Tel est en mécanique classique par exemple, le cas d'un système constitué d'un nombre fini de points matériels ou de corps solides. Les mouvements des fluides en hydrodynamique, les vibrations de la corde de la membrane, la propagation des ondes en optique et en acoustique sont des exemples de processus qu'il est impossible de décrire dans un espace des phases de dimension finie. Un processus est différentiable si son espace des phases est muni d'une structure de variété différentielle et si ses changements d'états dans le temps, sont définis par des fonctions différentiables. Les coordonnées et les vitesses des points d'un système mécanique sont différentiables par rapport au temps contrairement aux mouvements étudiés en théorie du choc. En mécanique classique donc, le mouvement d'un système peut être décrit par des équations différentielles ordinaires, tandis que d'autres procédés seront nécessaires, en mécanique quantique, en théorie de la chaleur, en hydrodynamique, en théorie de l'élasticité, en optique et en acoustique par exemple.

D'après la description de F. Jean [49], les systèmes contrôlés (ou commandés) peuvent être de natures variées : système mécanique (moteur, robot, drone, satellite, avion), processus chimique (réacteur, colonne de distillation), circuit électrique ou électronique, gestion de portefeuille (économie)... Un tel système résulte d'une relation entrée/sortie où l'entrée u représente le contrôle (ou commande) qui nous permet d'agir sur le système et la sortie y représente ce que l'on observe du système, généralement sous forme de mesures prises par intervalle de temps. On peut représenter cette relation par le schéma suivant.



Les enjeux sont ici d'analyser le comportement du système via la sortie y , en fonction de l'entrée u et de synthétiser les lois de commande à imposer en entrée afin d'obtenir un comportement souhaité. Nous utilisons une description interne du système, décrit par une modélisation mathématique des phénomènes physiques agissant sur le système. Il s'agit de l'approche par représentation d'état, en opposition à l'approche par représentation par fonction de transfert. Nous considérons donc que le système au **temps** t est décrit par son **état** $x(t)$ et nous modélisons l'évolution du vecteur $x(t)$ au cours du temps par une équation différentielle ordinaire commandée, appelée **système (dynamique) commandé**, de la forme :

$$\dot{x}(t) = f(t, x(t), u(t)). \quad (1.1)$$

La **sortie** $y(t)$ est donnée par une relation de la forme $y(t) = g(t, x(t), u(t))$. Typiquement, pour chaque instant t , $x(t)$ est un vecteur de \mathbb{R}^n , $u(t)$ un vecteur de \mathbb{R}^m et $y(t)$ un vecteur de \mathbb{R}^p . On appelle **loi de commande** une fonction $t \mapsto u(t)$, définie sur un intervalle $[t_0, \tau_u[$, avec $t_0 \in \mathbb{R}$ l'instant initial à partir duquel on commence à agir sur le système et $\tau_u \in \mathbb{R} \cup \{+\infty\}$ tel que $t_0 < \tau_u$ un instant limite après lequel la loi de commande n'est plus

définie. à une loi de commande $u(\cdot)$ donnée¹ est associée une équation différentielle ordinaire

$$\dot{x}(t) = f_{u(\cdot)}(t, x(t)), \quad t \in [t_0, \tau_u[, \quad (1.2)$$

avec $f_{u(\cdot)}(t, x) := f(t, x, u(t))$. Ainsi, une fonction d'état $x(\cdot)$ est solution du système commandé (1.1) s'il existe une loi de commande $u(\cdot)$ telle que $x(\cdot)$ est solution de l'équation différentielle (1.2).

Les premières questions essentielles portent sur l'analyse du comportement dynamique du système et d'autre part sur la synthèse des lois de commande. En ce qui concerne le premier point, les questions sont les suivantes :

- Est-il possible de trouver une commande $u(\cdot)$ qui amène un système, d'un état initial x_0 au temps initial t_0 , à un état terminal quelconque x_f , au temps final $t = t_f$? C'est le problème de la contrôlabilité.
- Les connaissances de la sortie $y(t)$ et du contrôle $u(t)$ pour tout $t \in [t_0, \tau_u[$, permet-elle de déterminer l'état $x(\cdot)$ pour tout $t \in [t_0, \tau_u[$ (ou de manière équivalente, l'état initial x_0) ? C'est le problème de l'observabilité.

Pour la synthèse des lois de commande, les premières questions sont :

- Si la réponse à la question de la contrôlabilité est positive, comment calculer la loi de commande réalisant le transfert de l'état initial à la cible, en temps fixé ? C'est la question de la planification de trajectoires.
- Si la réponse à la question de la contrôlabilité est positive, comment calculer la loi de commande, si elle existe, qui réalise le transfert de l'état initial à la cible tout en minimisant un critère donné ? On entre ici dans la théorie du contrôle optimal.
- Comment construire, si possible, un contrôle $u(\cdot)$ qui stabilise asymptotiquement le système autour d'un état d'équilibre x_e , c'est-à-dire tel que, pour toute condition initiale x_0 , potentiellement suffisamment proche de x_e , on ait $\lim_{t \rightarrow +\infty} x(t) = x_e$? C'est la question de la stabilisation ou régulation.

1.2 Solution d'un système contrôlé

Soient Ω un ouvert de \mathbb{R}^n , Π un ouvert de \mathbb{R}^m et une application de classe \mathcal{C}^1

$$\begin{aligned} f: \Omega \times \Pi &\longrightarrow \mathbb{R}^n \\ (x, u) &\longmapsto f(x, u). \end{aligned}$$

On considère le **système dynamique contrôlé**, ou équation différentielle ordinaire contrôlée, de la forme

$$\dot{x}(t) = f(x(t), u(t)). \quad (\Sigma_u)$$

Remarque 1.2.1. Attention, dans la définition de f , les variables x et u sont des vecteurs de \mathbb{R}^n et \mathbb{R}^m respectivement. Alors que dans l'équation différentielle contrôlée, x et u sont des fonctions du temps, c'est $x(t)$ et $u(t)$ qui sont des vecteurs de \mathbb{R}^n et \mathbb{R}^m respectivement. C'est pourquoi, pour ne pas confondre, on note parfois $x(\cdot)$ et $u(\cdot)$ les fonctions d'état et de contrôle qui dépendent du temps.

1. Nous écrivons $u(\cdot)$ et non u pour insister sur le fait que cet objet est une fonction du temps et non pas un vecteur de \mathbb{R}^m . Ce ne sera pas toujours le cas.

Remarque 1.2.2. Nous nous restreignons ici aux systèmes *autonomes*, pour lesquels l'application f ne dépend pas explicitement du temps t , pour faire ressortir l'influence de la commande sur la régularité des solutions de (Σ_u) . Dans ce cas, nous pouvons fixer l'origine des temps, par exemple à 0.

Il est important de préciser à quelle classe de fonctions appartiennent les lois de commande $u(\cdot)$. En pratique, celles-ci peuvent être extrêmement irrégulières, en particulier discontinues (par exemple, si l'action est contrôlée par un interrupteur). La classe choisie doit donc couvrir ces possibilités mais elle doit aussi permettre d'assurer l'existence de solutions pour l'équation associée (1.2). Nous supposons donc qu'une loi de commande $u(\cdot)$ est une fonction définie sur un intervalle $[0, \tau_u[$, avec $\tau_u \in \mathbb{R}_+^* \cup \{+\infty\}$, **mesurable et essentiellement bornée**, c'est à dire dans l'espace² $L^\infty([0, \tau_u[, \mathbb{R}^m)$ et à valeurs dans Π , *i.e.* dans $L^\infty([0, \tau_u[, \Pi)$. On notera

$$\mathcal{U} := \left\{ u(\cdot) \in L^\infty([0, \tau_u[, \Pi) \mid \tau_u \in \mathbb{R}_+^* \cup \{+\infty\} \right\}$$

cet espace, qui contient donc l'ensemble des lois de commandes autorisées pour (Σ_u) . Avec une telle classe de lois de commande, l'équation (1.2) ne vérifie pas les hypothèses classiques du théorème de Cauchy-Lipschitz assurant l'existence et l'unicité de solutions maximales ; en particulier, $t \mapsto f_{u(\cdot)}(t, x)$ n'est pas nécessairement continue. Nous utilisons donc la notion suivante de solution au sens de Carathéodory.³

Définition 1.2.1 – Solution de (Σ_u)

Soit $u(\cdot) \in \mathcal{U}$, définie sur $[0, \tau_u[$. On appelle solution de (Σ_u) tout couple $(I, x(\cdot))$, où I est un sous-intervalle de $[0, \tau_u[$ et $x(\cdot) : I \rightarrow \Omega$ est une fonction **absolument continue**⁴, telle que pour tout t, t_0 dans I :

$$x(t) = x(t_0) + \int_{t_0}^t f(x(s), u(s)) \, ds.$$

En particulier, $x(\cdot)$ satisfait $\dot{x}(t) = f(x(t), u(t))$ **presque partout** sur I . D'autre part, une solution est dite **maximale** si elle n'admet pas de prolongement à un intervalle strictement plus grand.

Remarque 1.2.3. Nous pourrions considérer des lois de commande plus générales, dans L_{loc}^∞ par exemple, c'est-à-dire seulement essentiellement bornées sur tout compact.⁵ Cependant, dans ce manuscrit, nous nous limitons à des problèmes de contrôle optimal définis sur un intervalle de temps compact. Ainsi, les lois de commande considérées sont dans tous les cas essentiellement bornées sur tout l'intervalle de temps considéré, c'est-à-dire dans L^∞ . Quant à la régularité de la fonction d'état $x(\cdot)$, nous pouvons mentionner que sa dérivée est (presque partout) bornée sur chaque sous-intervalle compact, ceci vient du fait que le second membre f de l'équation (Σ_u) , qui est continu, est composé avec des contrôles supposés essentiellement bornés et avec la fonction d'état elle-même continue

2. $L^\infty([0, \tau_u[, \mathbb{R}^m) = \{u : [0, \tau_u[\rightarrow \mathbb{R}^m \mid u \text{ mesurable et } \exists C \geq 0 \text{ t.q. } \|u(t)\| \leq C \text{ p.p. sur } [0, \tau_u[\}$.

3. cf. <https://tinyurl.com/caratheodory-wikipedia>.

4. $x \in AC([a, b], \mathbb{R}^n) \Leftrightarrow x \in L^1([a, b], \mathbb{R}^n)$ et $\exists y \in L^1([a, b], \mathbb{R}^n) : \forall t \in [a, b], x(t) = x(a) + \int_a^t y(s) \, ds$.

5. $\varphi \in L_{\text{loc}}^\infty(\mathcal{I}, \mathbb{R}^n)$ si $\varphi \mathbb{1}_K \in L^\infty(\mathcal{I}, \mathbb{R}^n)$, $\forall K \subset \mathcal{I}$ compact, où $\mathbb{1}_K$ est la fonction caractéristique de K .

(puisque absolument continue). Ainsi, les solutions $x(\cdot)$ sont en fait localement lipschitziennes et pas seulement absolument continues, et ceci marche aussi pour des contrôles dans L_{loc}^∞ . Enfin, nous pouvons ajouter que puisque l'on se restreint à des intervalles de temps compacts, les solutions sont donc lipschitziennes.

Nous avons alors le théorème suivant qui n'est qu'un cas particulier de [67, Lemme 2.6.2], nous assurant l'existence et l'unicité de solution maximale pour (Σ_u) .

Théorème 1.2.2 – Existence et unicité de solutions maximales

Soit $f: \Omega \times \Pi \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 , Ω ouvert de \mathbb{R}^n , Π ouvert de \mathbb{R}^m et soit $x_0 \in \Omega$. Alors, pour toute commande $u(\cdot) \in \mathcal{U}$, définie sur $[0, \tau_u[$, il existe une unique solution maximale $x(\cdot)$ (au sens de la Définition 1.2.1) de $\dot{x}(t) = f(x(t), u(t))$, définie sur un intervalle I inclus dans $[0, \tau_u[$ et contenant 0, telle que $x(0) = x_0$.

1.3 Flot d'un système contrôlé

Soient Ω un ouvert de \mathbb{R}^n , Π un ouvert de \mathbb{R}^m , $x_0 \in \Omega$, et une application de classe \mathcal{C}^1 , $f: \Omega \times \Pi \rightarrow \mathbb{R}^n$, $(x, u) \mapsto f(x, u)$. Ces hypothèses assurent d'après le Théorème 1.2.2, que pour tout contrôle $u(\cdot) \in \mathcal{U}$, défini sur $[0, \tau_u[$, avec $\tau_u \in \mathbb{R}_+^* \cup \{+\infty\}$, il existe une unique solution maximale que l'on notera

$$t \mapsto \Phi(t, x_0, u(\cdot)),$$

définie sur un intervalle I de la forme $[0, t_u[$, avec $t_u \leq \tau_u$, telle que $x(0) = x_0$, et vérifiant presque partout $\dot{x}(t) = f(x(t), u(t)) = f_{u(\cdot)}(t, x(t))$. Si par exemple $\Omega = \mathbb{R}^n$, $\tau_u = +\infty$ et si le temps limite t_u est fini, i.e. $t_u < +\infty$, alors la trajectoire part à l'infini quand t tend vers t_u d'après le critère de maximalité (ou théorème d'échappement), cf. [30, Section V.2.6]. Par exemple, si on considère le système $\dot{x} = x^2 + u$ avec comme condition initiale $x(0) = 0$, alors on voit que pour $u(\cdot) \equiv 1$ défini sur tout \mathbb{R} , i.e. $\tau_u = \infty$, la trajectoire solution associée $t \mapsto \tan t$ explose en $t_u = \pi/2$. Elle n'est donc définie que sur l'intervalle $[0, t_u[$ avec $t_u < \tau_u$.

Remarque 1.3.1. Si nous fixons un temps $t \geq 0$ et que nous nous intéressons aux solutions seulement sur l'intervalle $[0, t]$, alors, nous pouvons considérer les contrôles $u(\cdot)$ dans $L^\infty([0, t], \Pi)$. En effet, nous pouvons toujours prolonger $u(\cdot)$ sur $]t, \infty]$ par 0 par exemple pour avoir un contrôle dans \mathcal{U} .

On introduit la définition suivante qui fait disparaître les difficultés introduites par les temps t_u et τ_u .

Définition 1.3.1 – Ensemble des contrôles admissibles

Soient $t \geq 0$ et $x_0 \in \Omega$. On note \mathcal{U}_{t, x_0} l'ensemble des **contrôles admissibles** sur $[0, t]$ avec x_0 comme condition initiale, c'est-à-dire les contrôles $u(\cdot) \in L^\infty([0, t], \Pi)$ tels que $\Phi(\cdot, x_0, u(\cdot))$ est bien définie sur $[0, t]$.

Ainsi, l'application Φ est définie sur

$$\mathcal{D} := \{(t, x_0, u(\cdot)) \mid t \in \mathbb{R}_+, x_0 \in \Omega, u(\cdot) \in \mathcal{U}_{t, x_0}\}.$$

On définit alors pour $t \geq 0$, le **flot** (au temps t) du système contrôlé (Σ_u) :

$$\begin{aligned} \Phi_t: \quad \mathcal{D}_t &\longrightarrow \Omega \\ (x_0, u(\cdot)) &\longmapsto \Phi_t(x_0, u(\cdot)) := \Phi(t, x_0, u(\cdot)) \end{aligned} \quad (1.3)$$

avec $\mathcal{D}_t := \{(x_0, u(\cdot)) \mid (t, x_0, u(\cdot)) \in \mathcal{D}\}$. Concernant la régularité du flot, nous avons :

Théorème 1.3.2 – Régularité et différentielle du flot

Soit $T \geq 0$. Soit $f: \Omega \times \Pi \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 , avec $\Omega \subset \mathbb{R}^n$ ouvert et $\Pi \subset \mathbb{R}^m$ ouvert. Soit une solution $x(\cdot): [0, T] \rightarrow \Omega$ de (Σ_u) associée à la commande $u(\cdot)$, tels que $(x_0, u(\cdot)) \in \mathcal{D}_T$, $x_0 := x(0)$.

Soit $t \in [0, T]$. Alors, il existe un voisinage de $(x_0, u(\cdot))$ dans $\Omega \times L^\infty([0, t], \Pi)$ sur lequel l'application Φ_t est bien définie. De plus, Φ_t est \mathcal{C}^1 sur ce voisinage et sa différentielle en $(x_0, u(\cdot))$ est donnée pour tout $\delta x_0 \in \mathbb{R}^n$ et $\delta u(\cdot) \in L^\infty([0, t], \mathbb{R}^m)$ par :

$$\Phi'_t(x_0, u(\cdot)) \cdot (\delta x_0, \delta u(\cdot)) =: \delta x(t),$$

où $\delta x(t)$ est la solution au temps t de

$$\dot{\delta x}(s) = \frac{\partial f}{\partial x}(x(s), u(s)) \cdot \delta x(s) + \frac{\partial f}{\partial u}(x(s), u(s)) \cdot \delta u(s), \quad \delta x(0) = \delta x_0.$$

► Rappelons que l'espace de Sobolev $AC([0, t], \mathbb{R}^n) = W^{1,1}([0, t], \mathbb{R}^n) =: W_n^{1,1}$ et l'espace de Lebesgue $L^\infty([0, t], \mathbb{R}^m) =: L_m^\infty$ munis de leurs normes usuelles sont des espaces de Banach [22]. Nous introduisons aussi la notation $L^1([0, t], \mathbb{R}^n) =: L_n^1$. Définissons quelques applications qui vont nous être utiles.

1. Définissons l'application

$$\begin{aligned} F: \quad W_n^{1,1}([0, t], \Omega) \times L^\infty([0, t], \Pi) &\longrightarrow L_n^1 \\ (x, u) &\longmapsto F(x, u) \end{aligned}$$

telle que $\forall s \in [0, t]$, $F(x, u)(s) := f(x(s), u(s))$. On a F de classe \mathcal{C}^1 car f est \mathcal{C}^1 . Ce fait n'est pas trivial, F est un opérateur de Nemytskii (ou de superposition), cf. [4]. Ce résultat est légèrement plus général que celui de la Proposition A.2.1.

2. Définissons l'application linéaire

$$\begin{aligned} D: \quad W_n^{1,1} &\longrightarrow L_n^1 \\ x &\longmapsto D(x) := \dot{x}. \end{aligned}$$

Cette application est continue car

$$\|D(x)\|_{L_n^1} = \|\dot{x}\|_{L_n^1} \leq \|x\|_{W_n^{1,1}} = \|x\|_{L_n^1} + \|\dot{x}\|_{L_n^1}.$$

Ainsi, $\forall x \in W_n^{1,1}$, $\forall \delta x \in W_n^{1,1}$,

$$D'(x) \cdot \delta x = D(\delta x) = \dot{\delta x}.$$

3. Définissons, pour $s \in [0, t]$, l'application linéaire et continue (évident)

$$\begin{aligned} L_s: \quad W_n^{1,1} &\longrightarrow \mathbb{R}^n \\ x &\longmapsto L_s(x) := x(s). \end{aligned}$$

Ainsi, $\forall x \in W_n^{1,1}, \forall \delta x \in W_n^{1,1}, L'_s(x) \cdot \delta x = L_s(\delta x) = \delta x(s)$.

On définit ensuite l'application

$$\begin{aligned} \Psi: W^{1,1}([0, t], \Omega) \times \Omega \times L^\infty([0, t], \Pi) &\longrightarrow L_n^1 \times \mathbb{R}^n \\ (x, x_0, u) &\longmapsto (D(x) - F(x, u), L_0(x) - x_0), \end{aligned}$$

qui est une application de classe \mathcal{C}^1 puisque D, F et L_0 sont \mathcal{C}^1 . On note \bar{x}, \bar{x}_0 et \bar{u} la solution de (Σ_u) de l'énoncé qui est associée, avec la nouvelle notation, à \bar{x}_0 et \bar{u} . Nous voulons appliquer le théorème des fonctions implicites (A.6.1) à Ψ au point $(\bar{x}, \bar{x}_0, \bar{u})$, sur l'ouvert $W^{1,1}([0, t], \Omega) \times \Omega \times L^\infty([0, t], \Pi)$ de l'espace de Banach $W_n^{1,1} \times \mathbb{R}^n \times L_m^\infty$. Pour cela, nous devons vérifier que $\partial_x \Psi(\bar{x}, \bar{x}_0, \bar{u})$ est inversible. Elle est inversible si, pour tout $(y, v) \in L_n^1 \times \mathbb{R}^n$, il existe un unique δx dans $W_n^{1,1}$ tel que

$$\frac{\partial \Psi}{\partial x}(\bar{x}, \bar{x}_0, \bar{u}) \cdot \delta x = (\dot{\delta x} - \frac{\partial F}{\partial x}(\bar{x}, \bar{u}) \cdot \delta x, \delta x(0)) = (y, v).$$

Il s'agit donc de résoudre l'équation différentielle ordinaire linéaire avec second membre :

$$\begin{cases} \dot{\delta x}(s) = \frac{\partial f}{\partial x}(\bar{x}(s), \bar{u}(s)) \cdot \delta x(s) + y(s), & s \in [0, t], \\ \delta x(0) = v, \end{cases} \quad (1.4)$$

qui admet bien une unique solution dans $W_n^{1,1}$, cf. [67, Section C.3]. D'après le théorème des fonctions implicites, il existe un voisinage V de (\bar{x}_0, \bar{u}) dans $\Omega \times L^\infty([0, t], \Pi)$ et une application $\varphi: V \rightarrow W^{1,1}([0, t], \Omega)$ de classe \mathcal{C}^1 , tels que, pour tout $(x_0, u) \in V$,

$$\Psi(\varphi(x_0, u), x_0, u) = 0,$$

c'est-à-dire que $\varphi(x_0, u)$ est solution sur $[0, t]$ de (Σ_u) et vérifie $\varphi(x_0, u)(0) = x_0$. Autrement dit, $\Phi_t(x_0, u) = \varphi(x_0, u)(t) = (L_t \circ \varphi)(x_0, u)$ et donc Φ_t est bien définie et \mathcal{C}^1 sur V , et sa différentielle en (\bar{x}_0, \bar{u}) contre le vecteur $(\delta x_0, \delta u)$, notée $\delta x(t)$, est donnée par

$$\begin{aligned} \delta x &:= \varphi'(\bar{x}_0, \bar{u}) \cdot (\delta x_0, \delta u) \\ &= -\left(\frac{\partial \Psi}{\partial x}(\bar{x}, \bar{x}_0, \bar{u})\right)^{-1} \cdot \left(\frac{\partial \Psi}{\partial x_0}(\bar{x}, \bar{x}_0, \bar{u}) \cdot \delta x_0 + \frac{\partial \Psi}{\partial u}(\bar{x}, \bar{x}_0, \bar{u}) \cdot \delta u\right) \\ &= \left(\frac{\partial \Psi}{\partial x}(\bar{x}, \bar{x}_0, \bar{u})\right)^{-1} \cdot \left(\frac{\partial F}{\partial u}(\bar{x}, \bar{u}) \cdot \delta u, \delta x_0\right), \end{aligned}$$

qui est la solution de (1.4) avec $y = \partial_u F(\bar{x}, \bar{u}) \cdot \delta u$ et $v = \delta x_0$. Le théorème est démontré. \blacksquare

Corollaire 1.3.3

L'ensemble \mathcal{D}_t est un ouvert de $\Omega \times L^\infty([0, t], \Pi)$ et \mathcal{U}_{t, x_0} est un ouvert de $L^\infty([0, t], \Pi)$.

Nous pouvons donner une explication intuitive du Théorème 1.3.2. Considérons donc une solution $x(\cdot): [0, T] \rightarrow \Omega$ de (Σ_u) associée à la commande $u(\cdot)$ et notons $x_0 := x(0)$.

Supposons que $(x_0, u(\cdot)) \in \mathcal{D}_T$. Soit maintenant une paire $(\delta x_0, \delta u(\cdot)) \in \mathbb{R}^n \times L^\infty([0, T], \mathbb{R}^m)$ telle que $(x_0 + \delta x_0, u(\cdot) + \delta u(\cdot)) \in \mathcal{D}_T$. On note $x(\cdot) + \delta x(\cdot) + \dots$ la solution de condition initiale $x_0 + \delta x_0$ associée au contrôle $u(\cdot) + \delta u(\cdot)$, où $\delta x(\cdot)$ est la variation à l'ordre 1 et le reste regroupe les variations d'ordres supérieurs. Autrement dit, $\delta x(t) = \Phi'_t(x_0, u(\cdot)) \cdot (\delta x_0, \delta u(\cdot))$ pour tout $t \in [0, T]$. Calculons $\delta x(t)$ pour $t \in [0, T]$ fixé. Pour tout $s \in [0, t]$, on a

$$\begin{aligned} \frac{d}{dt}(x + \delta x + \dots)(s) &= f(x(s) + \delta x(s) + \dots, u(s) + \delta u(s)) \\ &= f(x(s), u(s)) + \frac{\partial f}{\partial x}(x(s), u(s)) \cdot \delta x(s) + \frac{\partial f}{\partial u}(x(s), u(s)) \cdot \delta u(s) + \dots \end{aligned}$$

et $x_0 + \delta x_0 = x(0) + \delta x(0) + \dots$. Ainsi, puisque $\dot{x}(s) = f(x(s), u(s))$, on obtient que $\delta x(t)$ est la solution au temps t de

$$\begin{cases} \dot{\delta x}(s) = \frac{\partial f}{\partial x}(x(s), u(s)) \cdot \delta x(s) + \frac{\partial f}{\partial u}(x(s), u(s)) \cdot \delta u(s), \\ \delta x(0) = \delta x_0. \end{cases}$$

Cette équation différentielle étant affine, la solution est donnée par

$$\delta x(t) = R(t, 0) \delta x_0 + \int_0^t R(t, s) B(s) \delta u(s) ds,$$

avec $B(s) := \partial_u f(x(s), u(s))$ et où $R(t, t_0)$ est la résolvante de l'équation différentielle homogène : $\delta \dot{x}(s) = \partial_x f(x(s), u(s)) \cdot \delta x(s)$.

Exercice 1.3.1: Retrouver l'expression de la différentielle de Φ_t à partir de l'équation intégrale

$$\Phi_t(x_0, u(\cdot)) = x_0 + \int_0^t f(\Phi_s(x_0, u(\cdot)), u(s)) ds.$$

Dépendance par rapport à l'instant initial. Considérons un système contrôlé non autonome de la forme

$$\dot{x}(t) = f(t, x(t), u(t)) \tag{1.5}$$

où la fonction f est de classe \mathcal{C}^1 sur $\mathcal{I} \times \Omega \times \Pi$, où \mathcal{I} est un intervalle ouvert de \mathbb{R} , Ω un ouvert de \mathbb{R}^n et Π un ouvert de \mathbb{R}^m . On note $t \mapsto \Phi(t, t_0, x_0, u(\cdot))$ la solution (maximale) du système (1.5), pour une loi de commande $u(\cdot)$ donnée, vérifiant la condition initiale $x(t_0) = x_0$. Pour un quadruplet $(t_1, t_0, x_0, u(\cdot)) \in \mathcal{I} \times \mathcal{I} \times \Omega \times L^\infty(\mathcal{I}, \Pi)$ tel que $t \mapsto \Phi(t, t_0, x_0, u(\cdot))$ est bien définie sur $[t_0, t_1]$, la différentielle de l'application partielle $t_0 \mapsto \Phi(t_1, t_0, x_0, u(\cdot))$ est l'application qui à $\delta t_0 \in \mathbb{R}$ associe

$$\delta x(t_1) := \frac{\partial \Phi}{\partial t_0}(t_1, t_0, x_0, u(\cdot)) \cdot \delta t_0,$$

la solution au temps t_1 de l'équation différentielle linéaire homogène

$$\begin{cases} \dot{\delta x}(t) = \frac{\partial f}{\partial x}(t, x(t), u(t)) \cdot \delta x(t), \\ \delta x(t_0) = -f(t_0, x_0, u(t_0)) \delta t_0, \end{cases}$$

où $x(t) := \Phi(t, t_0, x_0, u(\cdot))$.

Remarque 1.3.2. On adopte pour la suite la notation $\mathcal{U}_{t_1, t_0, x_0}$ pour l'ensemble des contrôles admissibles sur $[t_0, t_1]$ associés au système (1.5) et à la condition initiale $x(t_0) = x_0$. Cette notation est utile dans le cas non autonome.

Exercice 1.3.2: Retrouver de quelle équation différentielle est solution

$$\frac{\partial \Phi}{\partial t_0}(t_1, t_0, x_0, u(\cdot)) \cdot \delta t_0$$

à partir de l'équation intégrale

$$\Phi(t_1, t_0, x_0, u(\cdot)) = x_0 + \int_{t_0}^{t_1} f(t, \Phi(t, t_0, x_0, u(\cdot)), u(t)) dt.$$

Le Théorème 1.3.2 montre que l'étude de la différentielle du flot d'un système commandé revient à l'étude d'une équation différentielle commandée linéaire. Introduisons alors la notion d'équation linéarisée.

Définition 1.3.4 – Équation linéarisée

Soit $x(\cdot) : [0, T] \rightarrow \Omega$ une solution de (Σ_u) associée à la commande $u(\cdot) \in L^\infty([0, T], \Pi)$. L'équation différentielle sur $[0, T]$

$$\dot{\hat{\delta x}}(t) = \frac{\partial f}{\partial x}(x(t), u(t)) \cdot \delta x(t) + \frac{\partial f}{\partial u}(x(t), u(t)) \cdot \delta u(t),$$

pour $\delta u(\cdot) \in L^\infty([0, T], \mathbb{R}^m)$ fixé, est appelé **équation linéarisée** de (Σ_u) autour de $(x(\cdot), u(\cdot))$.

Un cas particulier très intéressant est celui où le système est linéarisé autour d'un couple d'équilibre, c'est-à-dire d'un couple $(x_e, u_e) \in \Omega \times \Pi$ tel que $f(x_e, u_e) = 0$. L'état x_e est alors une solution stationnaire de (Σ_u) associée à la commande constante u_e et le système linéarisé de (Σ_u) autour de (x_e, u_e) est linéaire et autonome si l'on considère une variation de commande de la forme $\delta u(\cdot) := K \delta x(\cdot)$:

$$\dot{\hat{\delta x}}(t) = (A + BK) \delta x(t), \quad A = \frac{\partial f}{\partial x}(x_e, u_e) \quad \text{et} \quad B = \frac{\partial f}{\partial u}(x_e, u_e).$$

Nous obtenons alors le schéma suivant commute :

Remarque 1.3.3. Pour stabiliser asymptotiquement le système non linéaire (Σ_u) autour de l'équilibre (x_e, u_e) , il suffit alors choisir K de sorte que les valeurs propres de $A + BK$ soient à partie réelle strictement négative. C'est ce que l'on appelle le placement de pôles.

1.4 Application entrée/sortie

Pour alléger les notations, nous utilisons la notation u à la place de $u(\cdot)$. Nous revenons à un système autonome.

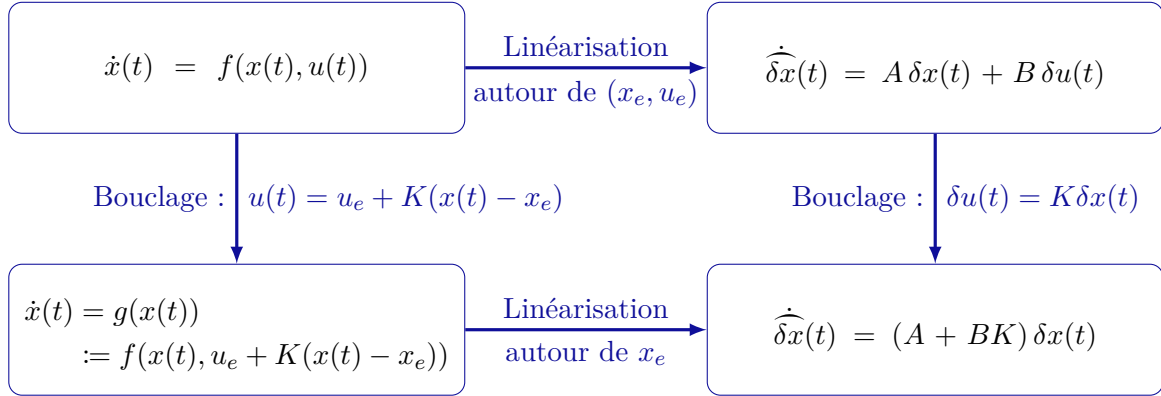


FIGURE 1.1 – Bouclage par retour d'état et linéarisation autour de points d'équilibre.

Définition 1.4.1 – Application entrée/sortie

On appelle **application entrée/sortie** au temps $t \geq 0$ et au point $x_0 \in \Omega$, l'application partielle

$$\begin{aligned} E_{t,x_0}: \mathcal{U}_{t,x_0} &\longrightarrow \Omega \\ u &\longmapsto E_{t,x_0}(u) := \Phi(t, x_0, u). \end{aligned}$$

Cette application est fondamentale en théorie du contrôle optimal [19, 20, 72]. Nous donnons ici une propriété sur la régularité de cette application, comme corollaire du Théorème 1.3.2.

Corollaire 1.4.2 – Différentielle de l'application entrée/sortie

Soient $x_0 \in \Omega$, $t \geq 0$ et $u \in \mathcal{U}_{t,x_0}$. Alors E_{t,x_0} est de classe \mathcal{C}^1 et sa différentielle en u suivant le vecteur $\delta u \in L^\infty([0, t], \mathbb{R}^m)$ est donnée par :

$$E'_{t,x_0}(u) \cdot \delta u = \frac{\partial \Phi_t}{\partial u}(x_0, u) \cdot \delta u = \int_0^t R(t, s) B(s) \delta u(s) ds,$$

où R est la résolvante de l'équation linéarisée de (Σ_u) autour de $(\Phi(\cdot, x_0, u), u)$ et où $B(s) := \partial_u f(\Phi(s, x_0, u), u(s))$.

Exemple 1.4.1. Considérons un système linéaire autonome de la forme

$$\dot{x}(t) = A x(t) + B u(t), \quad A \in \mathbf{M}_n(\mathbb{R}), \quad B \in \mathbf{M}_{n,m}(\mathbb{R}). \quad (\Sigma_{u,L})$$

Alors, on a

$$E_{t,x_0}(u) = \Phi_t(x_0, u) = \Phi(t, x_0, u) = e^{tA} x_0 + \int_0^t e^{(t-s)A} B u(s) ds,$$

(ici $R(t, s) = e^{(t-s)A}$) qui est affine en u et dont la différentielle est donnée par

$$E'_{t,x_0}(u) \cdot \delta u = \int_0^t e^{(t-s)A} B \delta u(s) ds.$$

□

1.5 Ensemble accessible et contrôlabilité

On s'intéresse ici à la question suivante : étant donné le système (Σ_u) et une condition initiale $x_0 \in \Omega$, où peut-on aller en temps fini $t \geq 0$, en faisant varier le contrôle u ? Nous définissons tout d'abord la notion d'ensemble accessible.

Définition 1.5.1 – Ensemble accessible

L'**ensemble accessible** (ou atteignable) $\mathcal{A}(t, x_0)$ en temps $t \geq 0$ depuis $x_0 \in \Omega$ pour le système (Σ_u) est l'ensemble des solutions au temps t pour tout $u \in \mathcal{U}_{t, x_0}$:

$$\mathcal{A}(t, x_0) := \{E_{t, x_0}(u) \mid u \in \mathcal{U}_{t, x_0}\} = \text{Im } E_{t, x_0}.$$

Sous de bonnes hypothèses, voir la notion de “bracket generating system” de [1, Chapitre 8 et Section 1.4], l'ensemble accessible $\mathcal{A}(x_0) := \{\mathcal{A}(t, x_0) \mid t \geq 0\}$ est d'intérieur non vide. Les Figures 1.2 et 1.3 montrent, sous ces hypothèses, de possibles ensembles accessibles tandis que la Figure 1.4 montre des cas impossibles.

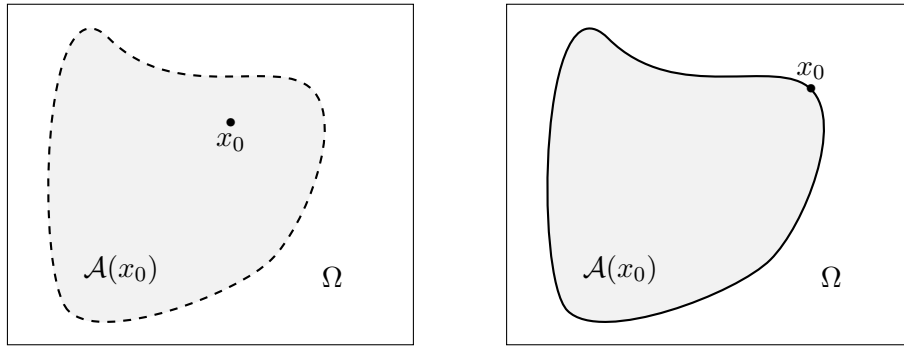


FIGURE 1.2 – Ensembles accessibles possibles : ouvert (gauche) ou à bord lisse (droite).

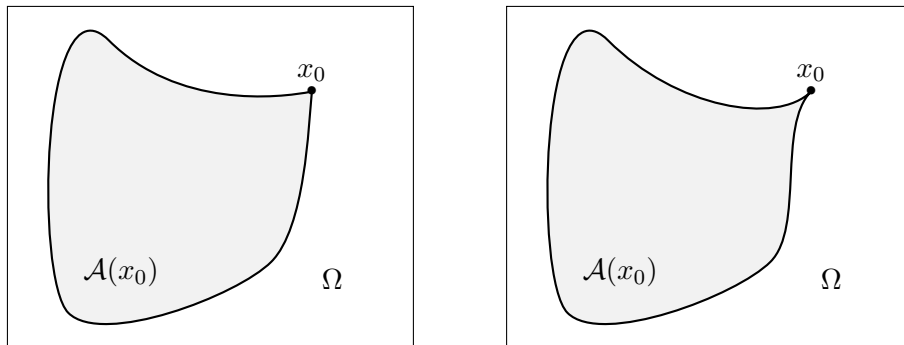


FIGURE 1.3 – Ensembles accessibles possibles ayant des bords non-lisses, anguleux (gauche) ou cuspidal (droite).

On note $\text{Int}(A)$ l'intérieur d'un ensemble A et $\text{Front}(A)$ sa frontière.

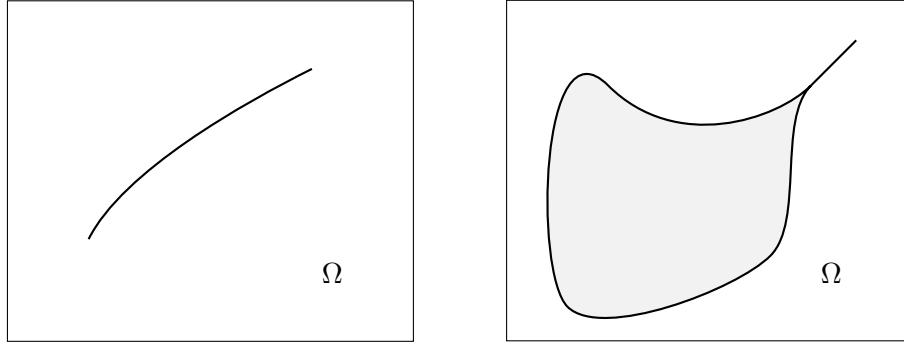


FIGURE 1.4 – Ensembles accessibles non possibles, de dimension non pleine (gauche) ou avec des points frontières isolés (droite).

Définition 1.5.2 – Contrôlabilités

Soit le temps $t > 0$. Le système (Σ_u) est :

- **contrôlable** depuis $x_0 \in \Omega$ en t si $\mathcal{A}(t, x_0) = \Omega$.
- **complètement contrôlable** en t si $\mathcal{A}(t, x_0) = \Omega$ pour tout $x_0 \in \Omega$.
- **localement contrôlable** en $x_0 \in \Omega$ en t autour de $x_1 \in \Omega$ si $x_1 \in \text{Int}(\mathcal{A}(t, x_0))$.

Remarque 1.5.1. Si E_{t,x_0} est surjective alors (Σ_u) est contrôlable depuis x_0 en temps t . Si $E'_{t,x_0}(u)$ est surjective alors (Σ_u) est localement contrôlable autour de $E_{t,x_0}(u)$, cf. Proposition 1.6.2 ci-après.

Nous avons le résultat standard suivant dans le cas linéaire.

Théorème 1.5.3 – Critère de contrôlabilité de Kalman

Le système linéaire $(\Sigma_{u,L})$ est complètement contrôlable en tout temps $t > 0$ ssi

$$\text{rang} [B \ AB \ \cdots \ A^{n-1}B] = n.$$

► Voir par exemple [67, Théorème 3] ou [49, Corollaire 6.1] pour la preuve. ■

Pour plus de détails concernant les notions de contrôlabilité, voir [19, 48, 50, 65, 72].

1.6 Caractérisation hamiltonienne des contrôles singuliers

On considère dans cette sous-section un triplet $(t, x_0, u) \in \mathcal{D}$.

Définition 1.6.1 – Contrôle régulier, singulier

Le contrôle u est dit **régulier** si $E'_{t,x_0}(u)$ est surjective. Sinon, il est dit **singulier**.

La notion de contrôle singulier est centrale en contrôle optimal [19, 20, 72]. Une simple application du Théorème A.7.2 de l'application ouverte non linéaire, nous donne le résultat

suisant. Voir aussi la Figure 1.5 pour une illustration du cas u régulier.

Proposition 1.6.2

Si u est un contrôle régulier pour E_{t,x_0} , alors (Σ_u) est localement contrôlable en x_0 autour de $E_{t,x_0}(u)$ en temps t .

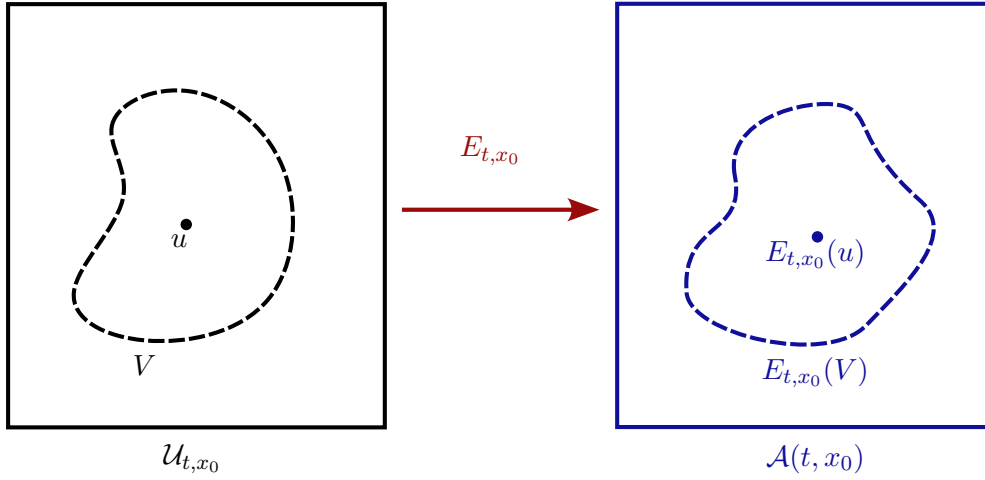


FIGURE 1.5 – Illustration du cas u régulier. L'ensemble V est un voisinage ouvert de u . Dans le cas régulier $E_{t,x_0}(V)$ est un voisinage ouvert de $E_{t,x_0}(u)$.

Le corollaire suivant est immédiat et fondamental.

Corollaire 1.6.3

Si $E_{t,x_0}(u) \in \text{Front}(\mathcal{A}(t, x_0))$, alors le contrôle u est singulier.

Remarque 1.6.1. Un système contrôlé peut être localement contrôlable le long d'une trajectoire singulière (*i.e.* associée à un contrôle singulier) : $\dot{x} = u^3$, $u \equiv 0$ singulier.

Nous pouvons synthétiser les propriétés ci-avant de la façon suivante :

$$\begin{aligned}
 u \text{ régulier} &\Rightarrow E_{t,x_0} \text{ est localement ouverte en } u \\
 &\Rightarrow \mathcal{A}(t, x_0) \text{ est localement ouvert au point } E_{t,x_0}(u) \\
 &\Rightarrow (\Sigma_u) \text{ est localement contrôlable en } x_0 \text{ autour de } E_{t,x_0}(u) \text{ en temps } t.
 \end{aligned}$$

Par contraposée, nous avons

$$E_{t,x_0}(u) \in \text{Front}(\mathcal{A}(t, x_0)) \Rightarrow u \text{ singulier}$$

mais attention, on peut avoir u singulier et $\mathcal{A}(t, x_0)$ localement ouvert en $E_{t,x_0}(u)$.

Avant de donner une caractérisation hamiltonienne des contrôles singuliers, faisons la définition suivante pour introduire ce que l'on appelle le pseudo-hamiltonien associé au système (Σ_u) .

Définition 1.6.4 – Pseudo-hamiltonien

Le **pseudo-hamiltonien** associé au système (Σ_u) est la fonction

$$\begin{aligned} H: \Omega \times (\mathbb{R}^n)^* \times \Pi &\longrightarrow \mathbb{R} \\ (x, p, u) &\longmapsto H(x, p, u) := p \cdot f(x, u). \end{aligned}$$

Nous avons alors la caractérisation suivante des contrôles singuliers.

Théorème 1.6.5 – Caractérisation hamiltonienne des contrôles singuliers

Supposons u singulier sur $[0, t]$ et notons $x(\cdot) := \Phi(\cdot, x_0, u)$ la trajectoire associée. Il existe alors une application $p: [0, t] \rightarrow (\mathbb{R}^n)^* \setminus \{0\}$ absolument continue t.q. pour $s \in [0, t]$ p.p. :

$$\dot{x}(s) = \frac{\partial H}{\partial p}[s], \quad \dot{p}(s) = -\frac{\partial H}{\partial x}[s], \quad 0 = \frac{\partial H}{\partial u}[s],$$

où H est le pseudo-hamiltonien associé au système (Σ_u) et où $[s] := (x(s), p(s), u(s))$. On appelle $p(s)$ le **vecteur adjoint** à $x(s)$ (ou covecteur).

► Par définition $E'_{t,x_0}(u)$ n'est pas surjective. Ainsi, $\exists \lambda \in (\mathbb{R}^n)^* \setminus \{0\}$ orthogonal au sous-espace vectoriel $\text{Im } E'_{t,x_0}(u)$, i.e. t.q. $\forall \delta u \in L^\infty([0, t], \mathbb{R}^m)$

$$\lambda \cdot (E'_{t,x_0}(u) \cdot \delta u) \stackrel{\text{cf. 1.4.2}}{=} \lambda \cdot \int_0^t R(t, s) B(s) \delta u(s) \, ds = 0,$$

avec $B(s) := \partial_u f(x(s), u(s))$ et où $R(t, s)$ est la résolvante de l'équation différentielle linéaire homogène $\dot{X}(\tau) = \partial_x f(x(\tau), u(\tau)) \cdot X(\tau)$, $X(s) = I_n$. Posons

$$p(s) := \lambda R(t, s) \in (\mathbb{R}^n)^*$$

pour tout $s \in [0, t]$.⁶ Alors, tout d'abord $p(s) \neq 0$ car $\lambda \neq 0$. Ensuite, on a :

$$\dot{p}(s) = \lambda \frac{\partial R}{\partial t_0}(t, s) = -\lambda R(t, s) A(s) = -p(s) A(s) = -\frac{\partial H}{\partial x}[s].$$

cf. Remarque 1.6.2, qui s'appuie sur la différentielle d'une solution d'une équation différentielle par rapport à l'instant initial. Et enfin on a :

$$\int_0^t p(s) B(s) \delta u(s) \, ds = 0,$$

et comme ceci est vrai pour tout $\delta u \in L^\infty([0, t], \mathbb{R}^m)$ on peut conclure :

$$\Rightarrow p(\cdot) B(\cdot) \perp L^\infty([0, t], \mathbb{R}^m) := L_m^\infty, \text{ car } L_m^\infty \subset L_m^2.$$

$$\Rightarrow p(\cdot) B(\cdot) = 0, \text{ car } \overline{L_m^\infty} = L_m^2. \text{ Voir [75, Proposition 2.32.1 et Corollaire 3.28.7].}$$

$$\Rightarrow \forall s \in [0, t] \text{ p.p., } p(s) B(s) = p(s) \partial_u f(x(s), u(s)) = \partial_u H[s] = 0. \quad \blacksquare$$

Remarque 1.6.2. Faisons le rappel suivant. Soit $R(t, s)$ la résolvante solution de

$$\dot{X}(\tau) = A(\tau)X(\tau) =: g(X(\tau)), \quad X(s) = I_n.$$

Alors $\partial_{t_0} R(t, s) = -R(t, s)A(s)$ car c'est la solution au temps t de l'équation linéarisée

$$\begin{cases} \widehat{\delta X}(\tau) = g'(X(\tau)) \cdot \delta X(\tau) = A(\tau)\delta X(\tau), \\ \delta X(s) = -g(X(s)) = -A(s)X(s) = -A(s). \end{cases}$$

Remarque 1.6.3. On peut considérer p comme un vecteur de \mathbb{R}^n et utiliser le produit scalaire usuel de \mathbb{R}^n . Le pseudo-hamiltonien s'écrit alors $H(x, p, u) = (p | f(x, u))$ et le système pseudo-hamiltonien devient

$$\dot{x}(s) = \nabla_p H(x(s), p(s), u(s)),$$

$$\dot{p}(s) = -\nabla_x H(x(s), p(s), u(s)),$$

ce qui peut s'écrire, en posant $z := (x, p)$, sous la forme :

$$\dot{z}(s) = J \nabla H(z(s), u(s)) =: \vec{H}(z(s), u(s)), \quad J := \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix},$$

J étant la matrice symplectique [55] canonique de dimension $2n$.

Interprétation géométrique du vecteur adjoint. Soit un triplet $(T, x_0, u) \in \mathcal{D}$. Supposons u singulier pour E_{T, x_0} . Alors, il est aussi singulier⁷ pour E_{t, x_0} pour tout $t \in [0, T]$ et de plus $p(t)$ est orthogonal à $\text{Im } E'_{t, x_0}(u)$. En particulier, $\text{Im } E'_{t, x_0}(u)$ est un sous-espace vectoriel de \mathbb{R}^n de codimension supérieure ou égale à 1. En effet, on a pour tout contrôle $\delta u \in L^\infty([0, t], \mathbb{R}^m)$:

$$p(t) \cdot (E'_{t, x_0}(u) \cdot \delta u) = p(t) \cdot \int_0^t R(t, s) B(s) \delta u(s) ds,$$

avec $p(t) = \lambda R(T, t)$, d'où en prolongeant δu par 0 sur $]t, T]$,

$$\begin{aligned} p(t) \cdot (E'_{t, x_0}(u) \cdot \delta u) &= \lambda \cdot \int_0^T R(T, s) B(s) \delta u(s) ds \\ &= \lambda \cdot (E'_{T, x_0}(u) \cdot \delta u) \\ &= 0. \end{aligned}$$

Nous donnons Figure 1.6 une illustration en dimension deux de cette interprétation géométrique pour un contrôle u qui de plus vérifie $E_{t, x_0}(u) \in \text{Front}(\mathcal{A}(t, x_0))$ pour tout $t \in [0, T]$.

6. Le vecteur adjoint est donc défini de manière rétrograde.

7. Ou plutôt sa restriction à $[0, t]$.

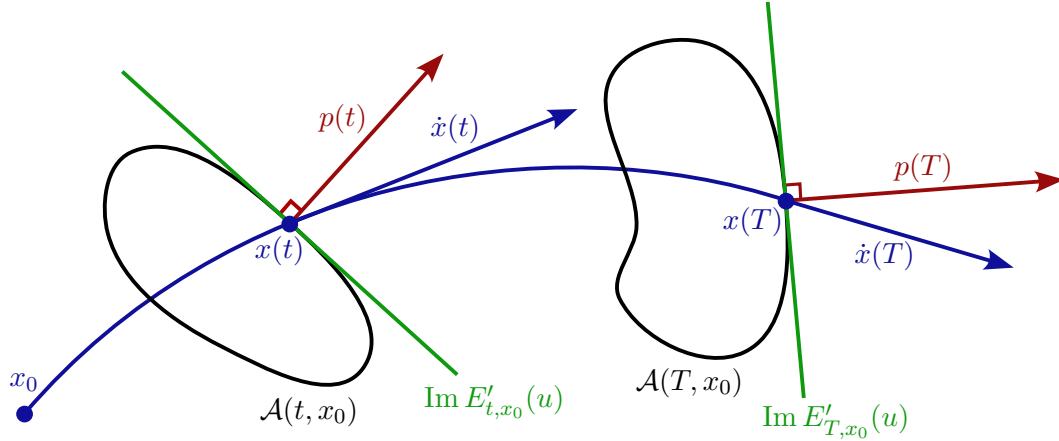


FIGURE 1.6 – Interprétation géométrique du vecteur adjoint en dimension deux dans le cas particulier où $E_{t,x_0}(u) \in \text{Front}(\mathcal{A}(t, x_0))$ pour tout $t \in [0, T]$.

1.7 Calcul de contrôles singuliers

On s'intéresse à la résolution de l'équation $\partial_u H(x, p, u) = 0$. Pour $(\bar{x}, \bar{p}, \bar{u})$ donné, si $\partial_u H(\bar{x}, \bar{p}, \bar{u}) = 0$ et si $\partial_{uu}^2 H(\bar{x}, \bar{p}, \bar{u})$ est inversible, alors d'après le théorème des fonctions implicites A.6.1, on peut trouver une fonction implicite, notée $u_s(x, p)$, telle que localement

$$\frac{\partial H}{\partial u}(x, p, u_s(x, p)) = 0,$$

et vérifiant $u_s(\bar{x}, \bar{p}) = \bar{u}$.

Exemple 1.7.1. Considérons deux exemples⁸ pour lesquels $\partial_{uu}^2 H(x, p, u)$ est inversible et calculons le contrôle singulier.

- Considérons un pseudo-hamiltonien de la forme

$$H(x, p, u) := H_0(x, p) + u p_1 + 0.5 u^2 p_2$$

avec x, p dans \mathbb{R}^2 et u dans \mathbb{R} . On a alors $\partial_u H(x, p, u) = 0$ ssi $p_1 + u p_2 = 0$ donc le contrôle singulier est de la forme

$$u_s(x, p) = -p_1/p_2 \quad \text{si } p_2 \neq 0.$$

- Considérons un pseudo-hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + \sum_{i=1}^m u_i H_i(x, p) + 0.5 \|u\|^2 p_n$$

avec $x \in \mathbb{R}^n$, $p \in (\mathbb{R}^n)^*$ et $u \in \mathbb{R}^m$. Posons

$$\Phi := (H_1, \dots, H_m).$$

On a alors $\partial_u H(x, p, u) = 0$ ssi $\Phi + p_n u = 0$ donc le contrôle singulier est de la forme

$$u_s(x, p) = -\Phi(x, p)/p_n \quad \text{si } p_n \neq 0.$$

□

8. Les deux pseudo-hamiltoniens sont en fait associés à des systèmes augmentés, cf. Définition 2.2.3.

Exercice 1.7.1 (solution p. 95) : Donner le contrôle singulier en fonction de l'état et du vecteur adjoint pour les pseudo-hamiltoniens ci-après. On supposera respectivement $p_2 \neq 0$ et $q \neq 0$.

1. Soit

$$H(x, p, u) = p_1 (-x_1 u) + \frac{p_2}{2} u^2$$

avec $x \in \mathbb{R}^2$, $p \in \mathbb{R}^2$ et $u \in \mathbb{R}$;

2. Soit

$$H((x, y), (p, q), u) = (p | Ax + Bu) + \frac{q}{2} (u^T Ru + x^T Qx)$$

avec $x \in \mathbb{R}^n$, $y \in \mathbb{R}$, $p \in \mathbb{R}^n$, $q \in \mathbb{R}$, $u \in \mathbb{R}^m$ et A , B , R et Q des matrices de tailles adaptées. On suppose R symétrique et inversible.

Donnons un exemple où nous n'avons pas l'inversibilité de la matrice $\partial_{uu}^2 H(x, p, u)$, en considérant un pseudo-hamiltonien affine en u .

Exemple 1.7.2. Considérons le pseudo-hamiltonien $H(x, p, u) = p_1 x_2^2 / 2 + p_2 u$ avec $u \in \mathbb{R}$, $x \in \mathbb{R}^2$ et $p \in \mathbb{R}^2$. Supposons que $\partial_u H(x(t), p(t), u(t)) = p_2(t) = 0$ presque partout sur un intervalle d'intérieur non vide et considérons les hypothèses du Théorème 1.6.5 vraies. Alors pour presque tout temps, on a

$$\frac{d}{dt} p_2(t) = \dot{p}_2(t) = -\frac{\partial H}{\partial x_2}(x(t), p(t), u(t)) = -p_1(t) x_2(t) = 0.$$

En dérivant $t \mapsto p_1(t) x_2(t)$, on obtient

$$\frac{d}{dt} (p_1(t) x_2(t)) = \dot{p}_1(t) x_2(t) + p_1(t) \dot{x}_2(t) = \alpha u(t) = 0, \quad \alpha = p_1(t) \neq 0.$$

Le contrôle singulier est donc $u_s \equiv 0$. □

Dans le cas précédent, nous pouvons faire des calculs de manière plus systématique. Nous avons pour cela besoin des définitions suivantes [55].

Définition 1.7.1 – Crochet de Poisson et système hamiltonien

Soient f et g deux fonctions lisses sur $\Omega \times (\mathbb{R}^n)^*$, $\Omega \subset \mathbb{R}^n$ ouvert. On note pour $z := (x, p) \in \Omega \times (\mathbb{R}^n)^*$:

$$\vec{f}(z) := \left(\frac{\partial f}{\partial p}(z), -\frac{\partial f}{\partial x}(z) \right),$$

$$\{f, g\}(z) := g'(z) \cdot \vec{f}(z) = \sum_{i=1}^n \frac{\partial f}{\partial p_i}(z) \frac{\partial g}{\partial x_i}(z) - \frac{\partial f}{\partial x_i}(z) \frac{\partial g}{\partial p_i}(z).$$

Le crochet $\{f, g\}$ s'appelle le **crochet de Poisson** de f et g , et \vec{f} le **système hamiltonien** (ou gradient symplectique) associé à f .

Remarque 1.7.1. Si on considère p comme un vecteur de \mathbb{R}^n , le crochet de Poisson s'écrit

$$\begin{aligned}\{f, g\}(z) &= (\nabla g(z) \mid J \nabla f(z)) = (\nabla g(z))^T J \nabla f(z) \\ &= (\nabla_x g(z) \mid \nabla_p f(z)) - (\nabla_p g(z) \mid \nabla_x f(z)),\end{aligned}$$

où J est définie dans la Remarque 1.6.3.

Proposition 1.7.2

Le crochet de Poisson est bilinéaire et possède les propriétés suivantes.

- *Antisymétrie* : $\{f, g\} = -\{g, f\}$.
- *Dérivation (ou règle de Leibniz)* : $\{fh, g\} = f\{h, g\} + h\{f, g\}$.
- *Identité de Jacobi* : $\{\{f, g\}, h\} + \{\{h, f\}, g\} + \{\{g, h\}, f\} = 0$.

► Voir par exemple [1, Prop. 11.8] pour l'identité de Jacobi. ■

Exemple 1.7.3. Considérons un pseudo-hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + u H_1(x, p),$$

avec $u \in \mathbb{R}$, $x \in \mathbb{R}^n$, $p \in (\mathbb{R}^n)^*$ et où H_0 et H_1 sont des fonctions lisses de (x, p) . Ainsi,

$$\frac{\partial H}{\partial u}(x, p, u) = H_1(x, p).$$

On note $z := (x, p)$. Si $H_1(z(t)) = 0$ sur un intervalle de temps I non réduit à un point alors pour tout temps $t \in I$, toutes les dérivées (on suppose qu'elles existent) de $t \mapsto H_1(z(t))$ sont égales à 0. En dérivant suffisamment de fois, nous allons faire apparaître le contrôle, ce qui nous donnera sa valeur. Soit $t \in I$, alors

$$\begin{aligned}\frac{d}{dt} H_1(z(t)) &= H'_1(z(t)) \cdot \dot{z}(t) \\ &= \{H, H_1\}(z(t)) && \text{d'après la Définition 1.7.1} \\ &= \{H_0, H_1\}(z(t)) + u(t) \{H_1, H_1\}(z(t)) && \text{car le crochet est linéaire} \\ &= \{H_0, H_1\}(z(t)) && \text{car le crochet est anti-symétrique} \\ &:= H_{01}(z(t)).\end{aligned}$$

Le contrôle n'apparaît pas, on dérive donc une seconde fois :

$$\begin{aligned}\frac{d}{dt} H_{01}(z(t)) &= H'_{01}(z(t)) \cdot \dot{z}(t) \\ &= \{H, H_{01}\}(z(t)) \\ &= \{H_0, H_{01}\}(z(t)) + u(t) \{H_1, H_{01}\}(z(t)) && \text{par linéarité} \\ &:= H_{001}(z(t)) + u(t) H_{101}(z(t)).\end{aligned}$$

Ainsi, le contrôle singulier est de la forme

$$u_s(z(t)) = -\frac{H_{001}(z(t))}{H_{101}(z(t))}$$

si $H_{101}(z(t)) \neq 0$. Si $H_{101}(z(t)) = 0$, il faut dériver encore (un nombre pair de fois). □

Problèmes de contrôle optimal

2.1	Introduction	21
2.2	Formulations	22
2.3	Exemples	24

Maintenant, en plus d'un système contrôlé, on se donne un problème de minimisation. Donnons un exemple introductif. Parmi toutes les solutions du système $\dot{x}(t) = f(x(t), u(t))$ reliant une condition initiale à une cible finale, en un temps t_f donné, nous cherchons à trouver une trajectoire et son contrôle qui minimisent un certain critère. Une telle paire, si elle existe, est dite optimale. L'existence de solutions optimales dépend de la régularité du système, du critère et de la classe des contrôles considérée. Dans ce chapitre, nous introduisons la formulation générale d'un problème de contrôle optimal et nous donnons quelques exemples.

2.1 Introduction

D'un point de vue global, un problème de contrôle optimal se formule sur une variété différentielle [48, 52, 63, 68], mais notre point de vue est local et on travaille sur un ouvert Ω de \mathbb{R}^n . Introduisons la problématique du contrôle optimal sur un exemple. Soient Ω un ouvert de \mathbb{R}^n , Π un ouvert de \mathbb{R}^m , $x_0 \in \Omega$, et une application de classe \mathcal{C}^1

$$\begin{aligned} f : \Omega \times \Pi &\longrightarrow \mathbb{R}^n \\ (x, u) &\longmapsto f(x, u). \end{aligned}$$

Ces hypothèses assurent d'après le Théorème 1.2.2, que pour tout temps $t_f > 0$, toute condition initiale $x_0 \in \Omega$ et tout contrôle $u \in \mathcal{U}_{t_f, x_0}$, il existe une unique solution définie sur $[0, t_f]$ vérifiant la condition initiale $x(0) = x_0$ et vérifiant presque partout sur $[0, t_f]$ l'équation différentielle contrôlée $\dot{x}(t) = f(x(t), u(t))$. Cette solution est simplement la restriction de la solution maximale $t \mapsto \Phi(t, x_0, u)$ à l'intervalle $[0, t_f]$.

Soit $t_f > 0$. Soient de plus f^0 une fonction de classe \mathcal{C}^1 sur $\Omega \times \Pi$ et g une fonction de classe \mathcal{C}^1 sur Ω , toutes deux à valeurs dans \mathbb{R} . Pour toute paire état-contrôle $(x, u) \in AC([0, t_f], \Omega) \times L^\infty([0, t_f], \Pi)$, on définit son coût par un critère de la forme :

$$J(x, u) := g(x(t_f)) + \int_0^{t_f} f^0(x(t), u(t)) \, dt.$$

Soit une condition initiale donnée $x_0 \in \Omega$. Soient maintenant X_f un sous-ensemble de Ω (par exemple une sous-variété de \mathbb{R}^n incluse dans Ω) que l'on cherche à atteindre et un sous-ensemble $U \subset \Pi$ fermé représentant les contraintes imposées sur le contrôle. Un exemple de problème de contrôle optimal est le suivant : on souhaite déterminer le contrôle $u \in \mathcal{U}_{t_f, x_0}$ à

valeurs dans U tel que la trajectoire associée $x(\cdot) := \Phi(\cdot, x_0, u)$ vérifie la condition terminale $x(t_f) \in X_f$ et tel que la paire (x, u) minimise le critère J . Ce problème de contrôle optimal (à temps final et condition initiale fixés) peut se réécrire de manière synthétique sous la forme suivante :

$$(P_1) \quad \begin{cases} \min J(x, u) \\ \dot{x}(t) = f(x(t), u(t)), \quad u(t) \in U, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) \in X_f, \end{cases}$$

ou encore sous la forme équivalente¹ :

$$\min \left\{ J(x, u) \mid u \in \mathcal{U}_{t_f, x_0} \cap U^{\mathbb{R}}, \quad x(t_f) \in X_f, \quad x(\cdot) = \Phi(\cdot, x_0, u) \right\}.$$

Puisque $x(\cdot) = \Phi(\cdot, x_0, u)$ est définie de manière unique par la donnée de u , car $t_0 = 0$ et x_0 est fixé, nous pouvons réécrire le problème sous la forme réduite suivante

$$\min \left\{ F(u) \mid u \in \mathcal{U}_{t_f, x_0} \cap U^{\mathbb{R}}, \quad E_{t_f, x_0}(u) \in X_f \right\},$$

avec $F(u) := J(\Phi(\cdot, x_0, u), u)$ et sachant que $E_{t_f, x_0}(u) = \Phi(t_f, x_0, u)$. Une solution (globale) au problème (P_1) est définie ainsi. Une paire état-contrôle (\bar{x}, \bar{u}) est dite admissible pour (P_1) si $\bar{u} \in \mathcal{U}_{t_f, x_0} \cap U^{\mathbb{R}}$, si $\bar{x}(\cdot) = \Phi(\cdot, x_0, \bar{u})$ et si $\bar{x}(t_f) \in X_f$. Elle est solution de (P_1) si pour toute autre paire état-contrôle (x, u) admissible : $J(\bar{x}, \bar{u}) \leq J(x, u)$.

Remarque 2.1.1. Nous pouvons parler de paire / solution admissible ou optimale. Une solution est sous-entendue optimale.

2.2 Formulations

Dans ce manuscrit, nous nous intéressons principalement à la forme (la plus générale) suivante de problème de contrôle optimal :²

$$(OCP) \quad \begin{cases} \min J(x, u, t_0, t_f) := g(t_0, x(t_0), t_f, x(t_f)) + \int_{t_0}^{t_f} f^0(t, x(t), u(t)) \, dt \\ \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U, \quad t \in [t_0, t_f] \text{ p.p.}, \\ c(t_0, x(t_0), t_f, x(t_f)) = 0_{\mathbb{R}^p}, \end{cases}$$

où f , f^0 , g et c sont au moins de classe \mathcal{C}^1 , avec

- $f: \mathcal{I} \times \Omega \times \Pi \rightarrow \mathbb{R}^n$,
- $f^0: \mathcal{I} \times \Omega \times \Pi \rightarrow \mathbb{R}$,
- $g: \mathcal{I} \times \Omega \times \mathcal{I} \times \Omega \rightarrow \mathbb{R}$ et
- $c: \mathcal{I} \times \Omega \times \mathcal{I} \times \Omega \rightarrow \mathbb{R}^p$, $p \leq 2(n+1)$,

1. L'ensemble $U^{\mathbb{R}}$ est l'ensemble des applications de \mathbb{R} dans U .

2. On trouve tout de même dans le manuscrit des problèmes avec des contraintes sur l'état mais ceux-ci sont bien plus complexes à résoudre et les conditions nécessaires bien plus difficile à obtenir, c'est pourquoi, dans un souci de présentation, nous nous limitons à une formulation sans contraintes sur l'état.

où \mathcal{I} est un intervalle ouvert de \mathbb{R} , Ω un ouvert de \mathbb{R}^n , Π un ouvert de \mathbb{R}^m , $U \subset \Pi$ fermé et où la fonction c est une **submersion** en tout point de $c^{-1}(\{0_{\mathbb{R}^p}\})$, c'est-à-dire $c'(t_0, x_0, t_f, x_f)$ est surjective pour tout point (t_0, x_0, t_f, x_f) vérifiant $c(t_0, x_0, t_f, x_f) = 0$.

Le problème de contrôle optimal est dit à **à temps final libre** si t_f n'est pas fixé, sinon on parle de problème **à temps final fixé**. De même, la temps initial t_0 peut être une variable d'optimisation et on parle alors de problème **à temps initial libre**. Si t_0 est fixé on parle de problème à **à temps initial fixé**. On définit une solution (globale) au Problème (OCP) ainsi :

Définition 2.2.1 – Solution de (OCP)

Un 4-uplet $(\bar{x}, \bar{u}, \bar{t}_0, \bar{t}_f)$ est dit **admissible** pour (OCP) si

- $(\bar{t}_0, \bar{t}_f) \in \mathcal{I}^2$,
- $\bar{u} \in L^\infty([\bar{t}_0, \bar{t}_f], \Pi)$, $u(t) \in U$ sur $[\bar{t}_0, \bar{t}_f]$,
- $\bar{x} \in AC([\bar{t}_0, \bar{t}_f], \Omega)$, $\bar{x} = \Phi(\cdot, \bar{t}_0, \bar{x}(t_0), \bar{u})$ et
- $c(\bar{t}_0, \bar{x}(\bar{t}_0), \bar{t}_f, \bar{x}(\bar{t}_f)) = 0$.

Il est de plus **optimal** si pour tout autre 4-uplet (x, u, t_0, t_f) admissible :

$$J(\bar{x}, \bar{u}, \bar{t}_0, \bar{t}_f) \leq J(x, u, t_0, t_f).$$

Définition 2.2.2

La formulation du problème de contrôle optimal (OCP) est dite de **Bolza**. Il existe deux autres formulations ; celle de **Lagrange** ($g \equiv 0$) et celle de **Mayer** ($f^0 \equiv 0$).

Le problème important du **temps minimal** correspond à $f^0 \equiv 1$ et $g \equiv 0$ ou bien à $f^0 \equiv 0$ et $g(t_0, x_0, t_f, x_f) = t_f$. Le temps final t_f étant bien entendu libre dans ce cas. On peut toujours se ramener à un problème à temps final fixé en normalisant l'intervalle de temps à $[0, 1]$ par exemple, puis en ajoutant t_f aux variables d'état. Si on note $s \in [0, 1]$ le nouveau temps, alors la dynamique pour la nouvelle variable d'état est $\dot{t}_f(s) = 0$ et on laisse $t_f(0)$ et $t_f(1)$ libres. Un autre point de vue est le suivant. Une fois l'intervalle de temps normalisé, on laisse t_f comme un paramètre à optimiser mais cette formulation à paramètres n'est pas incluse dans (OCP).

Exercice 2.2.1: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u) := \int_0^{t_f} f^0(x(t), u(t)) dt \\ \dot{x}(t) = f(x(t), u(t)), \quad u(t) \in U, \quad t \in [0, t_f] \text{ p.p.,} \\ x(0) = x_0, \quad c(x(t_f)) = 0, \end{cases}$$

sous les hypothèses habituelles. Montrer que l'on peut ramener le problème sur l'intervalle de temps $[0, 1]$ au lieu de $[0, t_f]$, en faisant un changement de variable adapté. Attention, le changement de variable peut avoir des conséquences sur le critère, l'équation différentielle et la contrainte $c = 0$.

D'une formulation de Lagrange à Mayer. Supposons $g \equiv 0$ et considérons :

$$(P_2) \quad \begin{cases} \min J(x, u, t_0, t_f) := \int_{t_0}^{t_f} f^0(t, x(t), u(t)) \, dt \\ \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U, \quad t \in [t_0, t_f] \text{ p.p.}, \\ c(t_0, x(t_0), t_f, x(t_f)) = 0, \end{cases}$$

Introduisons la définition suivante fondamentale pour la suite.

Définition 2.2.3 – Système augmenté

On note $\tilde{x} := (x, x^0)$ l'*état augmenté* et $\tilde{f} := (f, f^0)$ le *système augmenté* tels que

$$\tilde{f}(t, \tilde{x}, u) := (f(t, x, u), f^0(t, x, u)).$$

Le Problème (P_2) peut alors se réécrire sous la forme de Mayer particulière suivante :

$$(P_3) \quad \begin{cases} \min \tilde{J}(\tilde{x}, u, t_0, t_f) := x^0(t_f) \\ \dot{\tilde{x}}(t) = \tilde{f}(t, \tilde{x}(t), u(t)), \quad u(t) \in U, \quad t \in [t_0, t_f] \text{ p.p.}, \\ \tilde{c}(t_0, \tilde{x}(t_0), t_f, \tilde{x}(t_f)) = 0, \end{cases}$$

avec $\tilde{c}(t_0, \tilde{x}_0, t_f, \tilde{x}_f) := (c(t_0, x_0, t_f, x_f), x_0^0)$. Dans ce cas, la valeur initiale du critère est fixée à 0. Le pseudo-hamiltonien (voir Définition 1.6.4) associé au système augmenté s'écrit

$$H(t, \tilde{x}, \tilde{p}, u) := \tilde{p} \cdot \tilde{f}(t, \tilde{x}, u) = p \cdot f(t, x, u) + p^0 f^0(t, x, u),$$

où $\tilde{p} := (p, p^0)$ et puisque ce pseudo-hamiltonien ne dépend pas de la variable x^0 nous pouvons l'écrire $H(t, x, p, p^0, u)$. Nous donnons donc la définition suivante :

Définition 2.2.4 – Pseudo-hamiltonien associé au Problème (OCP)

Le *pseudo-hamiltonien* associé au problème de contrôle optimal (OCP) est :

$$\begin{aligned} H : \mathcal{I} \times \Omega \times (\mathbb{R}^n)^* \times \mathbb{R} \times \Pi &\longrightarrow \mathbb{R} \\ (t, x, p, p^0, u) &\longmapsto H(t, x, p, p^0, u) := p \cdot f(t, x, u) + p^0 f^0(t, x, u). \end{aligned}$$

2.3 Exemples

Exemple 2.3.1 (Transfert orbital autour de la Terre). L'une des activités historiques en contrôle optimal de l'équipe Algorithmes Parallèles et Optimisation (APO) de l'Institut de Recherche en Informatique de Toulouse (IRIT), est l'analyse de problèmes de transfert orbital de satellites. L'un des problèmes qui fut étudiés en collaboration avec le Centre National d'études Spatiales (CNES) de Toulouse, est le transfert d'un satellite autour de la Terre d'une orbite basse vers une orbite géostationnaire à l'aide d'un moteur à poussée faible. Les moteurs chimiques qui ont une poussée forte, permettent de réaliser le transfert en quelques

manoeuvres. L'utilisation de moteurs plus récents à propulsion électro-ionique à poussée faible permet la réduction significative des coûts liés à la consommation de carburant.

Dans ce problème, on connaît les orbites initiale et finale ainsi que la dynamique du satellite. L'objectif est alors de déterminer le contrôle (ou loi de commande), c'est-à-dire la poussée du moteur, à tous les instants, permettant de réaliser le transfert en minimisant la consommation de carburant. Le Principe du Maximum de Pontryagin (cf. Chapitre 4) indique alors que le contrôle optimal est de type **Bang-Bang**, c'est-à-dire ici que le moteur est par intervalle soit éteint, soit à pleine poussée.

Modélisation du problème. Le satellite est assimilé à un point matériel de masse m , soumis à l'attraction terrestre et à une force de propulsion F , dépendante du temps t . En première approximation, le système s'écrit

$$\ddot{r}(t) = -\mu \frac{r(t)}{\|r(t)\|^3} + \frac{F(t)}{m(t)},$$

où $r(t)$ désigne le vecteur position du satellite à l'instant t dans un référentiel dont l'origine est le centre de la Terre, μ la constante de gravitation de la Terre et $\|\cdot\|$ la norme euclidienne. Le système libre $F \equiv 0$ correspond aux équations de Kepler. Pratiquement, la poussée est limitée, *i.e.* $\|F\| \leq F_{\max}$ et on peut changer son orientation. La propulsion se fait par éjection de matière, à vitesse v_e et il faudrait rajouter au système l'équation de la masse $\dot{m}(t) = -F(t)/v_e$, ce que nous ne faisons pas dans cet exemple. Nous considérons la masse constante. L'état du système est donc $x := (r, v)$. Il vérifie à chaque instant t :

$$\begin{aligned}\dot{r}(t) &= v(t), \\ \dot{v}(t) &= -\mu \frac{r(t)}{\|r(t)\|^3} + \frac{F(t)}{m}\end{aligned}$$

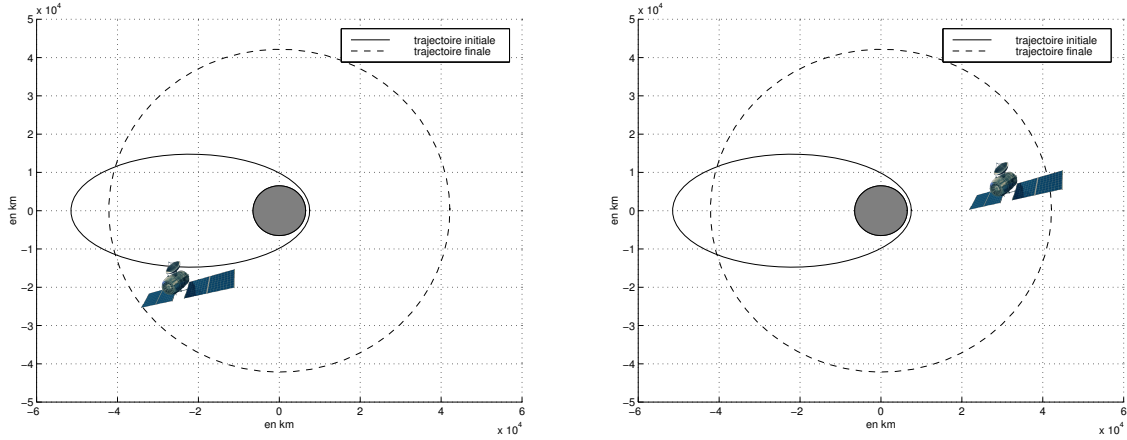
et on contrôle la poussée de l'engin F au cours du temps. Puisque nous considérons des poussées faibles, nous réécrivons $F := \varepsilon u$, de sorte que $u := (u_1, u_2, u_3)$ est soumis à la contrainte $\|u\| \leq 1$ et tel que $\varepsilon > 0$ petit soit la poussée maximale.

Deux questions fondamentales.

1. Soient x_0 une orbite initiale et x_f une orbite terminale, existe-t-il un contrôle $u(\cdot)$ tel que la trajectoire associée joigne x_0 à x_f en un temps fini t_f ? **C'est le problème de contrôlabilité.**

2. Si la condition de la question précédente est remplie, existe-t-il un contrôle joignant x_0 à x_f en temps fini qui de plus minimise la consommation de carburant ? Et s'il existe, comment le calculer ? **C'est le problème de contrôle optimal.**

Un résultat. Les données sont les suivantes. Le satellite a une masse $m = 1500$ Kg. Il doit être transféré autour de la Terre ($\mu = 398600.47 \text{ km}^3\text{s}^{-2}$) d'une orbite initiale elliptique basse, excentrique et inclinée (le satellite est lancé par une fusée) à une orbite finale géostationnaire dans le plan équatorial, en utilisant une poussée maximale de $\varepsilon = 10$ N (Newton), tout en minimisant la consommation de carburant.



Le critère ou encore la fonction coût à minimiser (*i.e.* la consommation de carburant) dépend ici du contrôle $u(\cdot)$ appliqué et du temps de transfert t_f . La fonction coût est simplement la norme L^1 du contrôle :

$$\int_0^{t_f} \|u(t)\| dt \rightarrow \min,$$

le temps final t_f étant fixé. Le contrôle optimal (de type Bang-Bang) ainsi que sa trajectoire associée se trouve Figure 2.1.

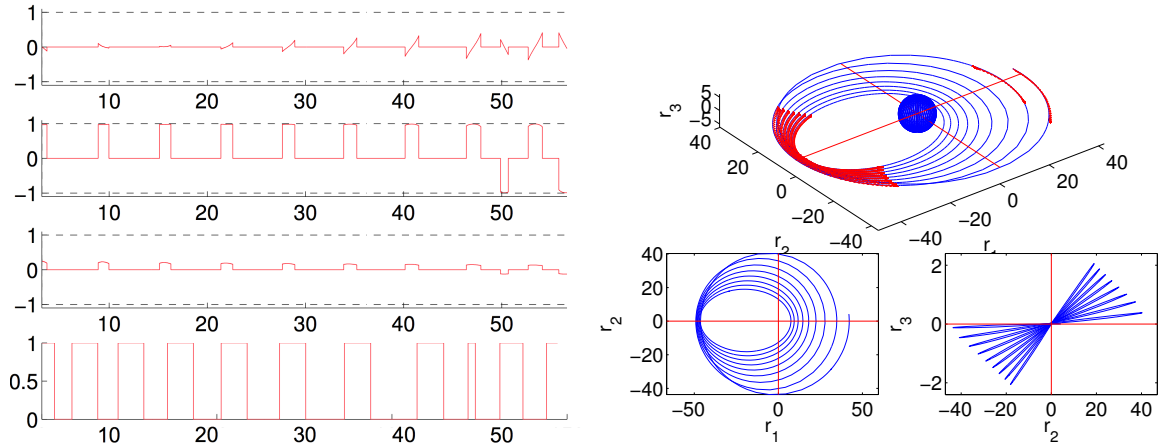


FIGURE 2.1 – Le contrôle optimal (graphes de gauche) est donné en fonction du temps pour les trois composantes u_1 , u_2 , u_3 et la norme $\|u\|$, respectivement de haut en bas. La trajectoire optimale est représentée sur les graphes de droite. On retrouve la trajectoire (en bleu) autour de la Terre en trois dimensions (en haut), et les projections sur les plans (r_1, r_2) et (r_2, r_3) (en bas). Les flèches rouges indique l'activité du contrôle qui est clairement localisée autour des apogées et des deux derniers périgées.

□

Exemple 2.3.2 (Problème temps-minimal linéaire). Voici un exemple de problème linéaire, où l'on cherche à minimiser le temps final et avec des conditions aux limites simples :

$$\begin{cases} \min J(x, u, t_f) := t_f \\ \dot{x}(t) = A x(t) + B u(t), \quad u(t) \in U, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

où $x(t) \in \mathbb{R}^n$, U est un polytope convexe compact de \mathbb{R}^m , A et B sont des matrices constantes d'ordre respectivement $n \times n$ et $n \times m$, et où x_0 et x_f sont fixés. Un tel problème est appelé un **problème temps-minimal linéaire**. \square

Exemple 2.3.3 (Problème linéaire-quadratique). Les problèmes linéaire-quadratique sont très présents en automatique par exemple. On peut par exemple chercher à stabiliser un pendule inversé à la verticale, ou encore un robot Lego segway, cf. Figure 2.2.

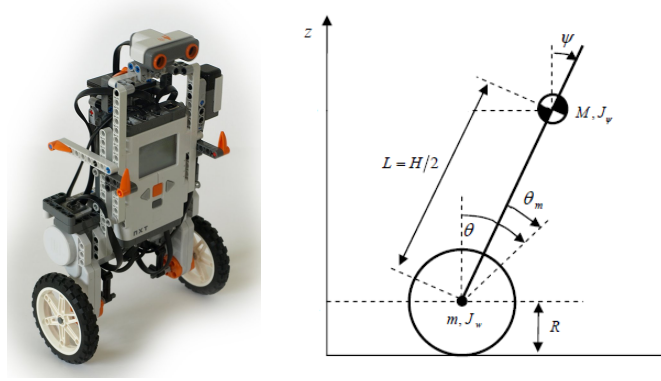


FIGURE 2.2 – Robot Lego segway.

Ces problèmes sont définis par un système linéaire, un coût quadratique, et en général il n'y a pas de contraintes terminales :

$$\begin{cases} \min J(x, u) := \frac{1}{2} \int_0^{t_f} (R u(t) | u(t)) + (P x(t) | u(t)) + (Q x(t) | x(t)) \\ \dot{x}(t) = A x(t) + B u(t), \quad u(t) \in \mathbb{R}^m, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \end{cases}$$

où $x(t) \in \mathbb{R}^n$, x_0 et t_f sont fixés, où A , B , P , Q et R sont des matrices constantes de dimensions adaptées, R et Q sont symétriques, et $(\cdot | \cdot)$ représente le produit scalaire usuel. Un tel problème est appelé un **problème linéaire-quadratique**. \square

Exemple 2.3.4 (Calcul des variations [25, 35]). Considérons un système contrôlé sans contraintes sous sa forme la plus simple $\dot{x}(t) = u(t)$, avec $\Pi = \mathbb{R}^n$. Supposons que les conditions aux limites soient elles aussi simples, i.e. $x(0) = x_0$ et $x(t_f) = x_f$, avec $t_f > 0$ fixé. Ce problème de contrôle optimal se réduit en un problème standard de **calcul des variations** :

$$\min_x \int_0^{t_f} L(x(t), \dot{x}(t)) dt, \quad x(0) = x_0, \quad x(t_f) = x_f,$$

où $L := f^0$ est un ce que l'on appelle un **lagrangien**. D'un point de vue contrôle optimal, il s'écrit :

$$\begin{cases} \min J(x, u) := \int_0^{t_f} L(x(t), u(t)) dt \\ \dot{x}(t) = u(t), \quad u(t) \in \mathbb{R}^n, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f. \end{cases}$$

Exemple de lagrangien : $L(x, u) = T(u) - V(x) = \frac{1}{2}\|u\|^2 - V(x)$, V un potentiel. \square

Exemple 2.3.5 (Géométrie riemannienne [31, 51]). Supposons que l'espace d'état soit une variété différentielle M de dimension n munie d'une **métrique riemannienne** g . Autrement dit, (M, g) est une variété riemannienne. Prenons le cas $M = \mathbb{R}^n$ pour simplifier. Dans, ce cas, la métrique peut s'écrire sous la forme suivante. Pour tout $x \in M = \mathbb{R}^n$ et tout vecteur tangent $v \in T_x M = \mathbb{R}^n$, on a :

$$g_x(v) = v^T A(x) v,$$

où $A(x)$ est une matrice définie positive et l'application $x \mapsto A(x)$ est lisse. L'objectif ici est la recherche de plus courts chemins entre deux points x_0 et x_f donnés de \mathbb{R}^n . Ceci revient à résoudre le problème de calcul des variations suivant :

$$\min \int_0^{t_f} \sqrt{g_{x(t)}(\dot{x}(t))} dt, \quad x(0) = x_0, \quad x(t_f) = x_f.$$

Une illustration sur un ellipsoïde de révolution oblat est donné Figure 2.3.

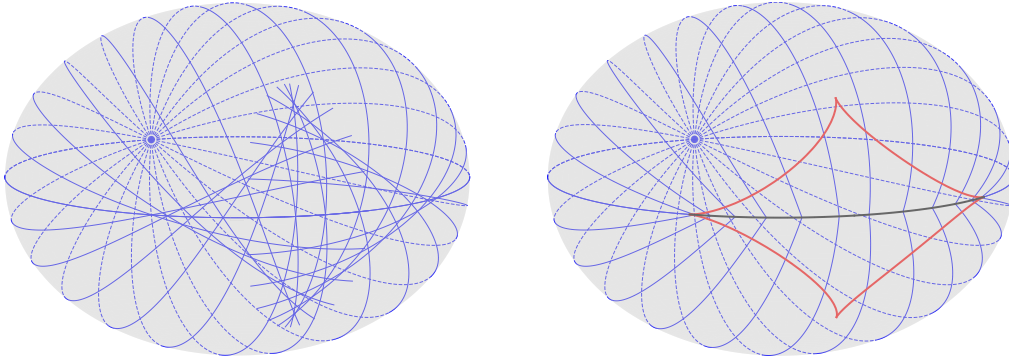
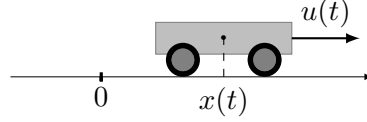


FIGURE 2.3 – Les trajectoires en bleu sont des géodésiques (minimisantes sur la figure de gauche), c'est-à-dire des trajectoires en ligne droite pour la métrique g , partant d'un même point x_0 et recouvrant l'ellipsoïde. Ici la métrique est la métrique usuelle de \mathbb{R}^3 projetée sur l'ellipsoïde. La courbe noire (à gauche) est ce que l'on appelle le lieu de coupure depuis x_0 , voir [31], c'est l'ensemble des points où les géodésiques perdent leur optimalité (globale). La caustique [7] rouge en forme d'astroïde, cf. <https://tinyurl.com/astroid-wikipedia>, est le lieu conjugué depuis x_0 , c'est l'ensemble des points où les géodésiques perdent leur optimalité locale et se replient. Ce lieu forme une enveloppe du flot géodésique que l'on peut distinguer sur la figure de droite.

\square

Exemple 2.3.6 (Chariot sur rail ou double intégrateur). Considérons un chariot se déplaçant sur un rail, dont on peut contrôler l'accélération par l'intermédiaire d'une force u . On choisit de représenter l'état q du chariot par la position de son centre x et sa vitesse v .



On suppose que la masse du chariot est constante et unitaire, et qu'il n'y a pas de frottements. La dynamique s'écrit : $\dot{x}(t) = v(t)$, $\dot{v}(t) = u(t)$. Cherchons à minimiser l'énergie du transfert : $\int_0^{t_f} u^2(t) dt$, avec comme conditions aux limites $x(0) = -1$, $v(0) = 0$, $x(t_f) = 0$ et $v(t_f) = 0$, en prenant $t_f = 1$. Dans ce cas, le contrôle optimal est lisse. Le résultat est donné Figure 2.4.

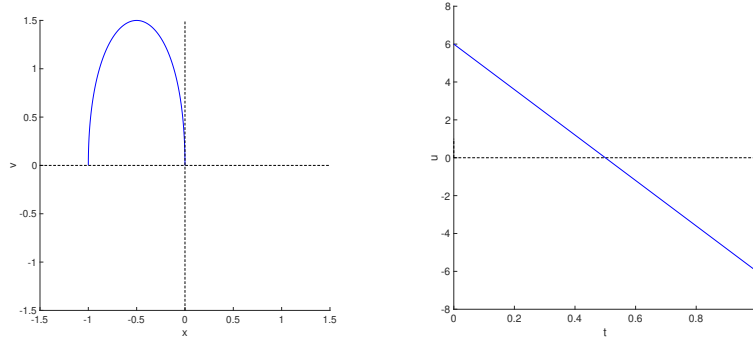


FIGURE 2.4 – Pour le problème à énergie minimale, le contrôle optimal est lisse.

Remplaçons le critère énergie par le temps de transfert. Il nous faut alors ajouter une contrainte sur la force u et le problème à temps minimal s'écrit :

$$(P_4) \quad \begin{cases} \min J(x, u, t_f) := t_f \\ \dot{x}(t) = v(t), \quad \dot{v}(t) = u(t), \quad |u(t)| \leq 1, \quad t \in [0, t_f] \text{ p.p.}, \\ (x(0), v(0)) = (-1, 0), \quad (x(t_f), v(t_f)) = (0, 0). \end{cases}$$

Dans ce cas, le contrôle optimal n'est pas lisse. Il possède une discontinuité, cf. Figure 2.5.

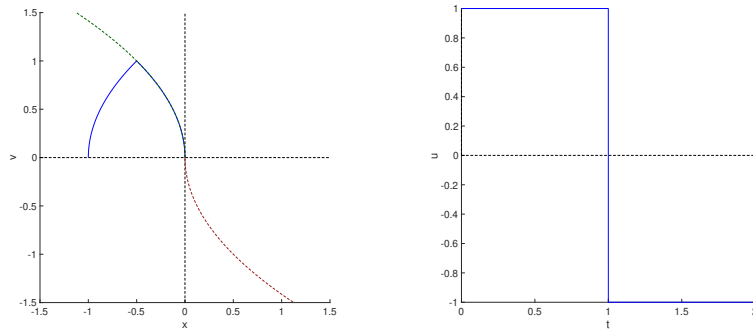


FIGURE 2.5 – Pour le problème à temps minimal, le contrôle optimal (à droite) est Bang-Bang ($u = +1$ puis $u = -1$). Sur la figure de gauche : (Bleu) Trajectoire dans le plan de phase. (Rouge) Demie parabole $x = \frac{1}{2}v^2$, $v \leq 0$. (Vert) Demie parabole $x = -\frac{1}{2}v^2$, $v \leq 0$. \square

Exemple 2.3.7 (Contrôle bang-bang (problème de Rayleigh)). Considérons le problème de temps minimal, *i.e.* $\min t_f$, pour le système contrôlé de Rayleigh [60] :

$$\dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = -x_1(t) + x_2(t)(1.4 - 0.14x_2(t)^2) + u_1(t),$$

avec les conditions aux limites $x_1(0) = x_2(0) = -5$, $x_1(t_f) = x_2(t_f) = 0$ et avec la contrainte sur le contrôle $|u_1(\cdot)| \leq 4$. Ce problème possède une unique solution, de nature Bang-Bang ; $u_1(\cdot)$ est discontinu et commute entre les valeurs -4 et $+4$. On remarque que $|u_1(t)| = 4$, pour tout $t \in [0, t_f]$, $t_f \approx 3.67$. La trajectoire et le contrôle solutions sont présentés Figure 2.6.

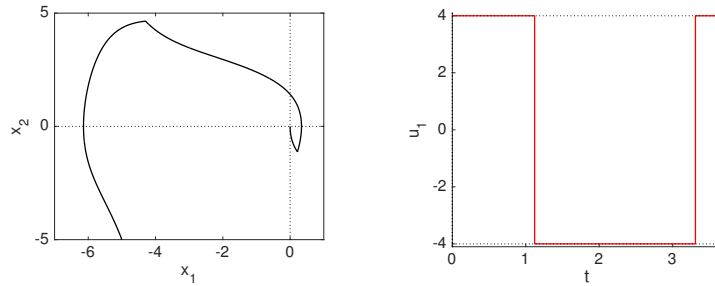


FIGURE 2.6 – Problème de Rayleigh. (Gauche) Trajectoire. (Droite) Contrôle. \square

Exemple 2.3.8 (Convergence presque-partout du contrôle régularisé (problème de Rayleigh)). Nous allons maintenant régulariser le problème de Rayleigh afin de rendre le contrôle optimal lisse. Nous utilisons la régularisation présentée dans [66]. On introduit un nouveau contrôle $u_2(\cdot)$ et un paramètre $\varepsilon > 0$ petit :

$$\dot{x}_1^\varepsilon(t) = x_2^\varepsilon(t) + \varepsilon u_2^\varepsilon(t), \quad \dot{x}_2^\varepsilon(t) = -x_1^\varepsilon(t) + x_2^\varepsilon(t)(1.4 - 0.14x_2^\varepsilon(t)^2) + u_1^\varepsilon(t),$$

avec les mêmes conditions aux limites que pour l'exemple précédent. Le contrôle $u^\varepsilon(\cdot) := (u_1^\varepsilon(\cdot), u_2^\varepsilon(\cdot))$ satisfait la contrainte $u_1^\varepsilon(\cdot)^2 + u_2^\varepsilon(\cdot)^2 \leq 16$. D'après [66, Théorème 1], la solution de ce problème notée $(x^\varepsilon(\cdot), u^\varepsilon(\cdot), t_f^\varepsilon)$ est telle que :

- t_f^ε converge vers t_f ,
- $x^\varepsilon(\cdot)$ converge uniformément vers $x(\cdot)$,
- $u_1^\varepsilon(\cdot)$ converge vers $u_1(\cdot)$ et $u_2^\varepsilon(\cdot)$ converge vers 0 presque partout sur $[0, t_f]$,

où $(x(\cdot), u_1(\cdot), t_f)$ est la solution du problème de Rayleigh. On peut observer ceci Figure 2.7.

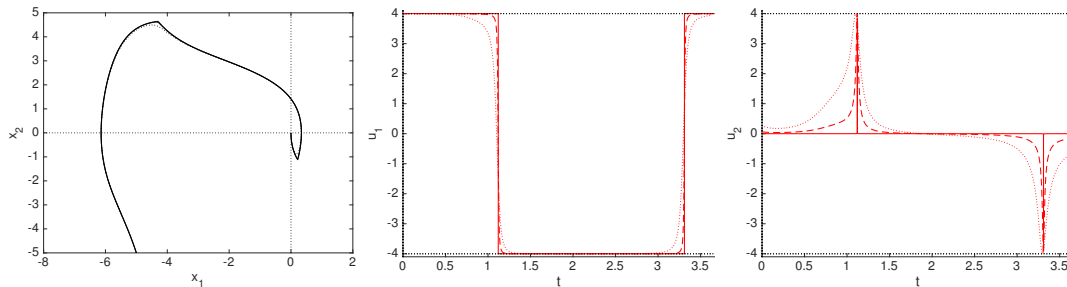


FIGURE 2.7 – Problème de Rayleigh régularisé, avec $\varepsilon = 1e^{-5}$, 0.01 et 0.05. Le temps final est normalisé à t_f^0 pour l'affichage. (Gauche) Trajectoire. (Milieu) $u_1^\varepsilon(\cdot)$. (Droite) $u_2^\varepsilon(\cdot)$. \square

Exemple 2.3.9 (Problème de Goddard et contrôle Bang-Singulier-Bang). Nous nous intéressons maintenant au problème centenaire de Goddard [37]. L'objectif est de maximiser l'altitude finale d'une fusée en vol vertical en un temps donné :

$$(P_G) \quad \begin{cases} \max h(t_f) \\ \dot{h}(t) = v(t), \\ \dot{v}(t) = \frac{1}{m(t)} (c u(t) - D(v(t), h(t))) - g_0, \\ \dot{m}(t) = -u(t), \\ 0 \leq u(t) \leq u_{\max}, \\ h(t_0) = h_0, \quad v(t_0) = v_0, \quad m(t_0) = m_0, \\ m(t_f) = m_f, \end{cases}$$



avec $t_0 = 0$, $q_0 := (h_0, v_0, m_0) = (0, 0, m_0)$, et où $h(t)$ est l'altitude au temps t , $v(t)$ est la vitesse, $m(t)$ la masse de la fusée, $D(h, v)$ la force de traînée, g_0 la constante de gravité, c l'impulsion spécifique et $u(t)$ la poussée du moteur. La structure optimale dépend du temps final fixé t_f . Nous reprenons les conditions expérimentales de l'article [59]. On sait alors que pour $t_f = 20$, la structure de la solution est Bang-Bang avec deux arcs. La poussée est maximale durant le premier arc, *i.e.* $u(\cdot) \equiv u_{\max}$, puis lors du deuxième arc, $u(\cdot) \equiv 0$. La trajectoire au cours du temps avec la loi de commande est donnée Figure 2.8. Pour $t_f \approx 206$, la structure optimale est Bang-Singulière-Bang avec trois arcs, le premier $u(\cdot) \equiv u_{\max}$, puis $u(\cdot) \in]0, u_{\max}[$ et enfin $u(\cdot) \equiv 0$, cf. Figure 2.9.

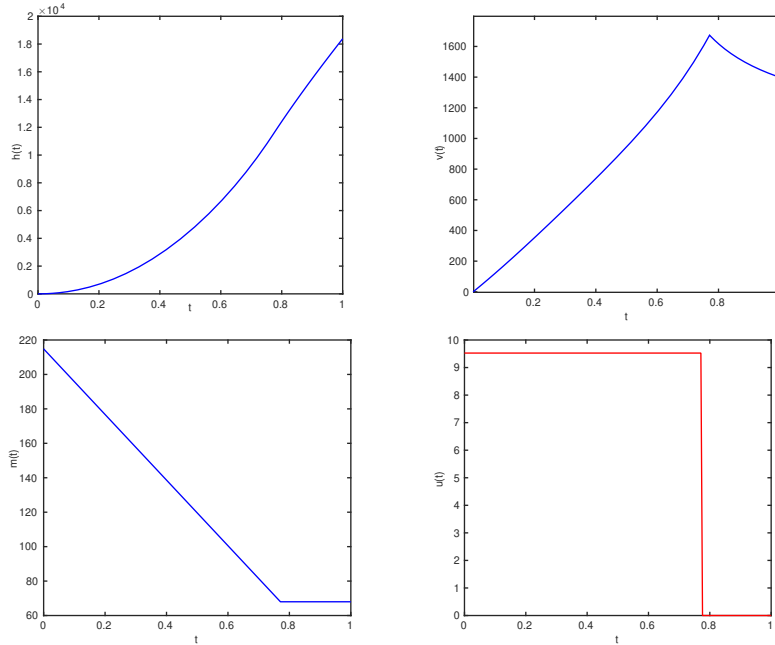


FIGURE 2.8 – Altitude, vitesse et masse de la fusée et poussée du moteur en fonction du temps normalisé. Ici $t_f = 20$.

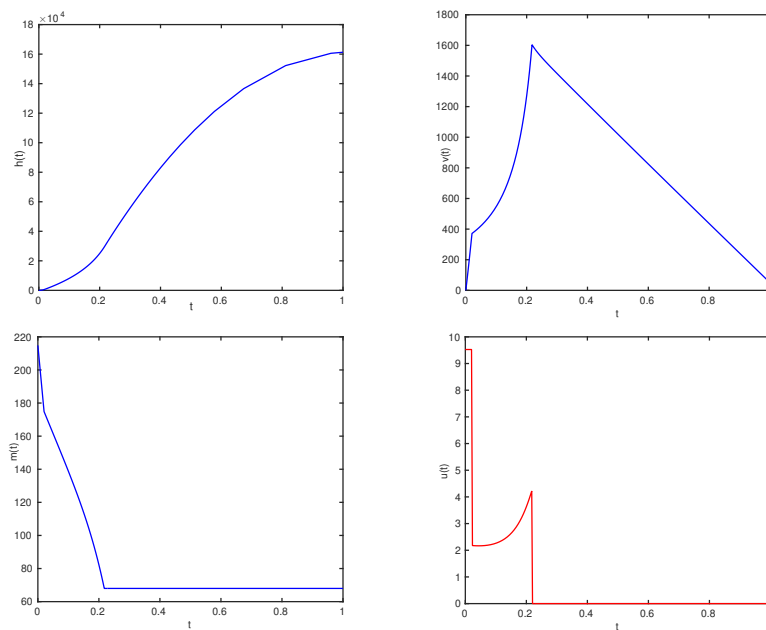


FIGURE 2.9 – Altitude, vitesse et masse de la fusée et poussée du moteur en fonction du temps normalisé. Ici $t_f \approx 206$.

□

Principe du maximum faible

3.1	Mayer avec conditions terminales	33
3.2	Lagrange avec conditions aux limites simples	35
3.3	Bolza à temps final libre et conditions terminales	37
3.4	Condition de Legendre et conservation du hamiltonien	41
3.5	Exemples	43
3.5.1	Simple intégrateur et fonction de tir	43
3.5.2	Calcul des variations	45
3.5.3	Principe fondamental de la dynamique	48
3.6	Condition du second ordre faible : cas Lagrange	51
3.7	Exercices	55

Dans cette section, nous allons donner des conditions nécessaires d'optimalité que l'on peut qualifier de faible par comparaison avec le Principe du Maximum de Pontryagin présenté par la suite. Pour le principe faible, nous supposons que $U = \Pi$ dans (OCP), c'est-à-dire qu'il n'y a pas de contraintes sur le contrôle. Autrement dit, U est un ouvert de \mathbb{R}^m .

3.1 Mayer avec conditions terminales

Nous commençons avec un problème sous la forme de Mayer et des conditions terminales seulement. La condition initiale, le temps initial et le temps final sont fixés. Nous considérons un système autonome donc nous pouvons fixer $t_0 = 0$.

Hypothèse H1. Soit un système contrôlé autonome $\dot{x}(t) = f(x(t), u(t))$, où f est une application lisse de $\Omega \times \Pi$ dans \mathbb{R}^n , Ω un ouvert de \mathbb{R}^n et Π un ouvert de \mathbb{R}^m . Considérons de plus deux applications g et c lisses de Ω dans, respectivement, \mathbb{R} et \mathbb{R}^p , $p \leq n$. On suppose que c est une submersion sur $c^{-1}(\{0_{\mathbb{R}^p}\})$. On considère enfin $t_f \geq 0$ et $x_0 \in \Omega$.

Remarque 3.1.1. Le qualificatif lisse veut dire suffisamment régulier. On peut supposer que les applications sont au moins de classe \mathcal{C}^1 .

On définit le problème de contrôle optimal de Mayer avec conditions terminales par :

$$(P_M) \quad \begin{cases} \min J(x, u) := g(x(t_f)) \\ \dot{x}(t) = f(x(t), u(t)), \quad t \in [0, t_f] \text{ p.p.,} \\ x(0) = x_0, \quad c(x(t_f)) = 0, \end{cases}$$

Nous pouvons réécrire le problème sous la forme suivante :

$$(P_M) \equiv \min \{ (g \circ E_{t_f, x_0})(u) \mid u \in \mathcal{U}_{t_f, x_0}, (c \circ E_{t_f, x_0})(u) = 0 \},$$

où \mathcal{U}_{t_f, x_0} et E_{t_f, x_0} sont définis Sections 1.3 et 1.4 respectivement. Rappelons que \mathcal{U}_{t_f, x_0} est un ouvert de l'ensemble des contrôles admissibles, cf. Corollaire 1.3.3. Nous avons alors les conditions nécessaires d'optimalité suivantes.

Proposition 3.1.1 – Conditions nécessaires, problème de Mayer

Si (x, u) est solution du Problème (P_M) sous les Hypothèses (H1), alors il existe un vecteur adjoint $p \in AC([0, t_f], (\mathbb{R}^n)^*)$, un réel $p^0 \in \{-1, 0\}$ et une forme linéaire $\lambda \in (\mathbb{R}^p)^*$, tels que $(p, p^0) \neq (0, 0)$, et les équations suivantes sont vérifiées pour $t \in [0, t_f]$ p.p. :

$$\begin{aligned}\dot{x}(t) &= \frac{\partial H}{\partial p}(x(t), p(t), u(t)), \\ \dot{p}(t) &= -\frac{\partial H}{\partial x}(x(t), p(t), u(t)), \\ 0 &= \frac{\partial H}{\partial u}(x(t), p(t), u(t)),\end{aligned}$$

où $H(x, p, u) := p \cdot f(x, u)$ est le pseudo-hamiltonien associé à (P_M) . Les conditions terminales $c(x(t_f)) = 0$ sont vérifiées et on a la **condition de transversalité** :

$$p(t_f) = (p^0 g'(x(t_f)) + \lambda \circ c'(x(t_f))) .$$

► Le Problème (P_M) est équivalent au problème

$$\min \{ (g \circ E)(u) \mid u \in \mathcal{U}_{t_f, x_0}, (c \circ E)(u) = 0_{\mathbb{R}^p} \}$$

où on a noté $E := E_{t_f, x_0}$ et où \mathcal{U}_{t_f, x_0} est un ouvert de $L^\infty([0, t_f], \mathbb{R}^m)$, cf. Corollaire 1.3.3, sachant que $L^\infty([0, t_f], \mathbb{R}^m)$ est un espace de Banach. Posons

$$\begin{aligned}L: \mathcal{U}_{t_f, x_0} \times (\mathbb{R}^p)^* \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (u, \lambda, \lambda^0) &\longmapsto L(u, \lambda, \lambda^0) := \lambda^0 (g \circ E)(u) + \lambda \cdot (c \circ E)(u).\end{aligned}$$

D'après le Théorème A.7.3, si u est optimal, alors il existe une paire $(\lambda, \lambda^0) \in (\mathbb{R}^p)^* \times \mathbb{R}$ telle que $(\lambda, \lambda^0) \neq (0, 0)$ et on peut par homogénéité fixer $\lambda^0 \in \{-1, 0\}$. De plus, pour tout $\delta u \in L^\infty([0, t_f], \mathbb{R}^m)$:

$$\frac{\partial L}{\partial u}(u, \lambda, \lambda^0) \cdot \delta u = (\lambda^0 g'(E(u)) + \lambda \circ c'(E(u))) \cdot (E'(u) \cdot \delta u) =: \Psi \cdot \delta x(t_f) = 0,$$

où l'on a défini $\Psi := \lambda^0 g'(E(u)) + \lambda \circ c'(E(u))$ et $\delta x(t_f) := E'(u) \cdot \delta u$. On pose $p^0 := \lambda^0$.

* Voyons pour les trois premières équations de la proposition. Si $\Psi = 0$ alors elles sont vérifiées en posant $p(t) = 0$. Si $\Psi \neq 0$, alors u est singulier pour E et on applique le Théorème 1.6.5. On a dans les deux cas, $p(t) = \Psi R(t_f, t)$, où R est la résolvante associée à l'équation linéarisée.

* Pour le reste de la proposition, quelque soit la valeur de Ψ , on a $p(t_f) = \Psi$ par définition de p et $(p, p^0) \neq (0, 0)$ car sinon $(p(t_f), p^0) = (0, 0)$ donc $\lambda \circ c'(E(u)) = 0$, ce qui implique alors que $\lambda = 0$, car c est une submersion sur $c^{-1}(\{0\})$, et ceci est impossible car alors on aurait $(\lambda, \lambda^0) = (0, 0)$, ce qui n'est pas. ■

Donnons deux exemples où $\Psi = p(t_f) = 0$, Ψ étant définie dans la démonstration de la Proposition 3.1.1. Remarquons que dire que $\Psi = 0$ revient à dire que $x(t_f)$ vérifie la condition nécessaire d'optimalité à l'ordre 1 (CN1) du problème :

$$\min_{x \in \Omega} \{g(x) \mid c(x) = 0\}. \quad (3.1)$$

Exemple 3.1.1. Soit le problème de Mayer défini par $\Omega = \mathbb{R}$, $\Pi = \mathbb{R}$, $g(x) = x^2$, $f(x, u) = u$, $c(x) \equiv 0$ et $x_0 = -1$. Le pseudo-hamiltonien est $H(x, p, u) = pu$. Ainsi, $\partial_u H(x, p, u) = p = 0$ implique que $p(t) = 0$, donc $p^0 = -1$. La condition de transversalité est $p(t_f) = -2x(t_f)$. Puisque $p(t_f) = 0$, on trouve que $x(t_f) = 0$, ce qui est bien le minimum de g sur \mathbb{R} . Remarquons que l'on peut choisir n'importe quel contrôle nous amenant en 0 en temps t_f . Tous ces contrôles admissibles ont le même coût. \square

Exemple 3.1.2. Soit le problème de Mayer défini par $\Omega = \mathbb{R}^2$, $\Pi = \mathbb{R}^2$, $g(x) = x_1$, $f(x, u) = u$, $c(x) = \|x\|^2 - 1$ et $x_0 = 0$. Le pseudo-hamiltonien est $H(x, p, u) = pu$. Ainsi, $\partial_u H(x, p, u) = p = 0$ implique que $p(t) = 0$. La condition de transversalité est $p(t_f) = (-1, 0) + 2\lambda x(t_f)$. Puisque $p(t_f) = 0$, on trouve que $x_2(t_f) = 0$ et $2\lambda x_1(t_f) = 1$. La condition finale $c(x(t_f)) = 0$ implique que $x_1(t_f) = \pm 1$, c-à-d qu'il y a deux solutions à la CN1 de (3.1), mais puisque l'on cherche à minimiser $g(x) = x_1$, la solution est $x_1(t_f) = -1$. Au final, on obtient que $x(t_f) = (-1, 0)$ et $\lambda = -1/2$. Comme dans l'exemple précédent, n'importe quel contrôle amenant l'état en $(-1, 0)$ en temps t_f est optimal. \square

3.2 Lagrange avec conditions aux limites simples

Nous donnons maintenant les conditions nécessaires d'optimalité faibles pour un problème de Lagrange avec conditions aux limites simples, sans contraintes sur le contrôle et où t_0 , x_0 et t_f sont fixés. Rappelons que par la suite, le Principe du Maximum de Pontryagin généralisera toutes les conditions nécessaires que nous donnons dans ce chapitre.

Hypothèse H2. Soit un système contrôlé autonome $\dot{x}(t) = f(x(t), u(t))$, où f est une application lisse de $\Omega \times \Pi$ dans \mathbb{R}^n , Ω un ouvert de \mathbb{R}^n et Π un ouvert de \mathbb{R}^m . Soient une application f^0 lisse de $\Omega \times \Pi$ dans \mathbb{R} , un scalaire $t_f \geq 0$ et deux points x_0, x_f dans Ω .

On définit le problème de Lagrange avec conditions aux limites simples par :

$$(P_L) \quad \begin{cases} \min J(x, u) := \int_0^{t_f} f^0(x(t), u(t)) dt \\ \dot{x}(t) = f(x(t), u(t)), \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

Nous pouvons écrire le problème sous la forme suivante :

$$(P_L) \equiv \min \left\{ (\pi_{x_0} \circ \tilde{E}_{t_f, \tilde{x}_0})(u) \mid u \in \tilde{\mathcal{U}}_{t_f, \tilde{x}_0}, E_{t_f, x_0}(u) = x_f \right\},$$

où $\tilde{\mathcal{U}}_{t_f, \tilde{x}_0}$ est l'ensemble des contrôles admissibles pour le système augmenté, cf. Définition 2.2.3, et où $\tilde{E}_{t_f, \tilde{x}_0}$ est l'application entrée/sortie associée au système augmenté. Nous

avons aussi introduit la projection $\pi_{x^0}(\tilde{x}) = x^0$, avec $\tilde{x} = (x, x^0) \in \mathbb{R}^n \times \mathbb{R}$. Introduisons l'ensemble accessible augmenté $\tilde{\mathcal{A}}(t, \tilde{x}_0) := \text{Im } \tilde{E}_{t, \tilde{x}_0}$. Alors, nécessairement

$$\tilde{E}_{t_f, \tilde{x}_0}(u) \in \text{Front}(\tilde{\mathcal{A}}(t_f, \tilde{x}_0)),$$

car sinon, nous pourrions baisser le coût, voir l'illustration Figure 3.1. En effet sinon, il existerait un voisinage du point $\tilde{x}(t_f) = \tilde{E}_{t_f, \tilde{x}_0}(u) = (x(t_f), x^0(t_f))$ dans $\tilde{\mathcal{A}}(t_f, \tilde{x}_0)$ contenant un point (y, y^0) tel que $y^0 < x^0(t_f)$, ce qui contredirait l'optimalité du contrôle u . Ainsi, d'après le Corollaire 1.6.3, u est un contrôle singulier pour l'application entrée/sortie augmentée. On peut alors appliquer le Théorème 1.6.5 mais sur le système augmenté.

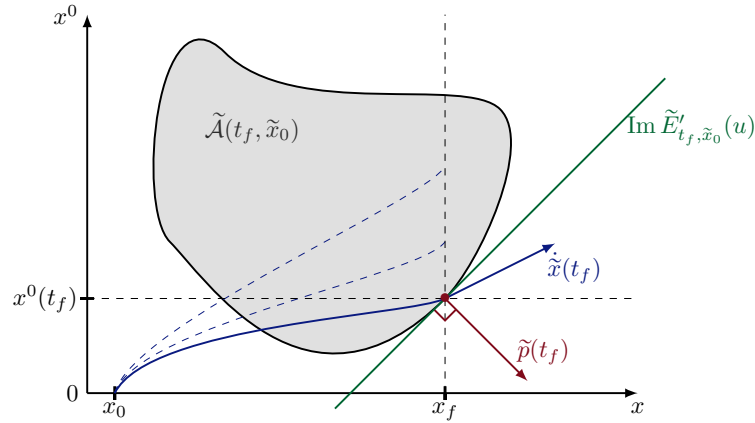


FIGURE 3.1 – Sur cet exemple, on a $p^0 \neq 0$.

Ceci nous mène aux conditions nécessaires d'optimalité suivantes.

Proposition 3.2.1 – Conditions nécessaires, problème de Lagrange

Si (x, u) est une solution du Problème (P_L) sous les Hypothèses (H2), alors il existe un vecteur adjoint $p \in AC([0, t_f], (\mathbb{R}^n)^*)$, un réel $p^0 \in \{-1, 0\}$, tels que $(p, p^0) \neq (0, 0)$, et tels que les équations suivantes sont vérifiées pour $t \in [0, t_f]$ p.p. :

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H}{\partial p}(x(t), p(t), p^0, u(t)), \\ \dot{p}(t) &= -\frac{\partial H}{\partial x}(x(t), p(t), p^0, u(t)), \\ 0 &= \frac{\partial H}{\partial u}(x(t), p(t), p^0, u(t)), \end{aligned}$$

où $H(x, p, p^0, u) := p \cdot f(x, u) + p^0 f^0(x, u)$ est le pseudo-hamiltonien associé à (P_L) .

Remarque 3.2.1. Nous pouvons remarquer qu'il n'y a pas de conditions de transversalité dans la Proposition 3.2.1, simplement parce que la condition de transversalité de la Proposition 3.1.1 appliquée au problème de Mayer associé s'écrit $\tilde{p}(t_f) = (\lambda, p^0)$ et n'apporte donc aucune information pertinente.

Exercice 3.2.1: La preuve de la Proposition 3.2.1 est laissée en exercice. On pourra utiliser le passage d'une formulation de Lagrange à Mayer.

Remarque 3.2.2. Le choix de fixer $p^0 \leq 0$ est arbitraire. On aurait très bien pu choisir $p^0 \geq 0$. Cela changerait alors simplement l'orientation du vecteur adjoint p .

Remarque 3.2.3. Si $E'_{t_f, x_0}(u)$ est surjective, i.e. si u n'est pas un contrôle singulier pour l'application entrée/sortie non augmentée, alors $p^0 \neq 0$ (on a la qualification des contraintes). Si $E_{t_f, x_0}(u) \in \text{Front}(\mathcal{A}(t_f, x_0))$, alors u est singulier pour E_{t_f, x_0} et $p^0 = 0$, cf. Figure 3.2.

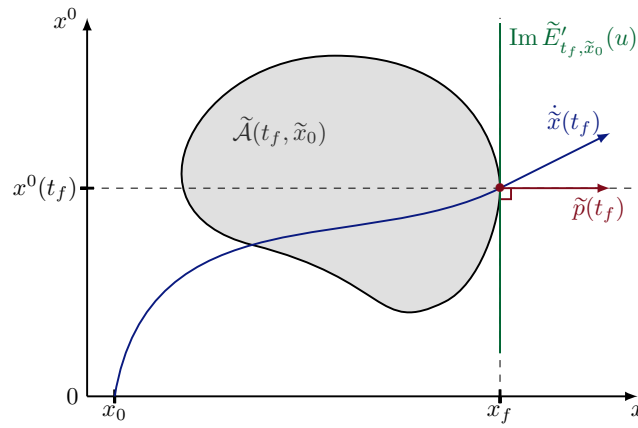


FIGURE 3.2 – Sur cet illustration, on a $p^0 = 0$.

3.3 Bolza à temps final libre et conditions terminales

Généralisons un peu plus le problème de contrôle optimal. Considérons pour cela un système non autonome, un problème sous la forme de Bolza et fixons seulement t_0 et x_0 , c'est-à-dire que le temps final t_f est libre.

Hypothèse H3. Soit un système contrôlé non autonome $\dot{x}(t) = f(t, x(t), u(t))$ où f est une application lisse de $\mathcal{I} \times \Omega \times \Pi$ dans \mathbb{R}^n , \mathcal{I} un intervalle ouvert de \mathbb{R} , Ω un ouvert de \mathbb{R}^n et Π un ouvert de \mathbb{R}^m . Considérons de plus une application f^0 lisse sur $\mathcal{I} \times \Omega \times \Pi$ et à valeurs dans \mathbb{R} et deux applications g et c lisses de $\mathcal{I} \times \Omega$ dans, respectivement, \mathbb{R} et \mathbb{R}^p , $p \leq n + 1$. On notera (t_f, x_f) l'argument des applications g et c . On suppose enfin que c est une submersion sur $c^{-1}(\{0_{\mathbb{R}^p}\})$. Soit $t_0 \in \mathcal{I}$ et $x_0 \in \Omega$.

On définit le problème de Bolza à temps final libre et conditions terminales par :

$$(P_B) \quad \begin{cases} \min J(x, u, t_f) := g(t_f, x(t_f)) + \int_{t_0}^{t_f} f^0(t, x(t), u(t)) dt \\ \dot{x}(t) = f(t, x(t), u(t)), \quad t \in [t_0, t_f] \text{ p.p.}, \\ x(t_0) = x_0, \quad c(t_f, x(t_f)) = 0, \end{cases}$$

Nous pouvons écrire le Problème (P_B) sous la forme suivante :

$$\min \{ F(u, t_f) \mid t_f \in \mathcal{I}, u \in \mathcal{U}_{t_0, t_f, x_0}, c(t_f, \Phi(t_f, t_0, x_0, u)) = 0 \},$$

où

$$F(u, t_f) := J(\Phi(\cdot, t_0, x_0, u), u, t_f),$$

où $\Phi(\cdot, t_0, x_0, u)$ est la solution maximale du système $\dot{x}(t) = f(t, x(t), u(t))$ de condition initiale $x(t_0) = x_0$ et où $\mathcal{U}_{t_0, t_f, x_0}$ est l'ensemble des contrôles $u \in L^\infty([t_0, t_f], \Pi)$ tels que $\Phi(\cdot, t_0, x_0, u)$ est bien définie sur $[t_0, t_f]$. Pour faire apparaître explicitement les dépendances aux variables à optimiser, on peut écrire le critère sous la forme :

$$F(u, t_f) = g(t_f, \Phi(t_f, t_0, x_0, u)) + \pi_{x^0}(\tilde{\Phi}(t_f, t_0, \tilde{x}_0, u)),$$

où $\tilde{\Phi}$ est le flot associé au système augmenté et $\pi_{x^0}(\tilde{x}) = x^0$ pour $\tilde{x} = (x, x^0) \in \mathbb{R}^n \times \mathbb{R}$.

Supposons que le cas trivial $t_f = t_0$ ne soit pas admissible, c'est-à-dire qu'il n'est pas possible de vérifier la contrainte $c = 0$ pour $t_f = t_0$. Dans ce cas, le Problème (P_B) est encore équivalent au problème :

$$\min \{ F(u, t_f) \mid (u, t_f) \in O, h(u, t_f) = 0 \},$$

où $h(u, t_f) := c(t_f, \Phi(t_f, t_0, x_0, u))$ et où

$$O := \{ (u, t_f) \mid t_f \in \mathcal{I} \cap]t_0, \infty[, u \in \mathcal{U}_{t_0, t_f, x_0} \}.$$

Nous voulons appliquer le Théorème (A.7.3). Pour cela, nous aimerions montrer que O est un ouvert d'un espace de Banach. Adoptons une stratégie différente. Supposons que le Problème (P_B) admette une solution (\bar{u}, \bar{t}_f) . Alors, il existe $\varepsilon > 0$ tel que $\bar{u} \in \mathcal{U}_{t_0, \bar{t}_f + \varepsilon, x_0}$, cf. [67, Lemme 2.6.2], et pour lequel $\bar{t}_f + \varepsilon \in \mathcal{I}$. Nous pouvons alors considérer le problème équivalent (à démontrer) suivant :

$$\min \{ F(u, t_f) \mid (u, t_f) \in O_\varepsilon, h(u, t_f) = 0 \},$$

où $O_\varepsilon := \mathcal{U}_{t_0, \bar{t}_f + \varepsilon, x_0} \times]t_0, \bar{t}_f + \varepsilon[$ est un ouvert de l'espace de Banach $L^\infty([t_0, \bar{t}_f + \varepsilon], \Pi) \times \mathbb{R}$, puisque $\mathcal{U}_{t_0, \bar{t}_f + \varepsilon, x_0}$ est un ouvert de $L^\infty([t_0, \bar{t}_f + \varepsilon], \Pi)$ d'après [67, Théorème 1]. Nous pouvons donc appliquer le Théorème (A.7.3).

Puisque le temps final est libre, c'est donc une variable d'optimisation. Dans les conditions nécessaires, nous nous attendons donc à avoir une nouvelle condition à vérifier. Voyons seulement d'où vient cette nouvelle condition. Introduisons la notation $x(t_f, u) := \Phi(t_f, t_0, x_0, u)$, t_0 et x_0 étant fixés. On pose

$$L(u, t_f, p^0, \lambda) := p^0 F(u, t_f) + \lambda \cdot c(t_f, x(t_f, u))$$

le lagrangien associé au problème. Alors la solution doit vérifier (on omet les arguments)

$$\begin{aligned} 0 &= \frac{\partial L}{\partial t_f} = p^0 \frac{\partial F}{\partial t_f} + \lambda \circ \left(\frac{\partial c}{\partial t_f} + \frac{\partial c}{\partial x_f} \cdot \dot{x} \right) \\ &= p^0 \frac{\partial g}{\partial t_f} + \lambda \circ \frac{\partial c}{\partial t_f} + p^0 f^0 + \left(p^0 \frac{\partial g}{\partial x_f} + \lambda \circ \frac{\partial c}{\partial x_f} \right) \cdot f \\ &= \frac{\partial \xi}{\partial t_f} + p^0 f^0 + p(t_f) \cdot f, \end{aligned}$$

car $p(t_f) = \partial_{x_f} \xi$, avec $\xi = p^0 g + \lambda \cdot c$. Nous obtenons les conditions suivantes.

Proposition 3.3.1 – Conditions nécessaires, problème de Bolza

Si (x, u, t_f) est une solution du Problème (P_B) sous les Hypothèses (H3), alors il existe un vecteur adjoint $p \in AC([t_0, t_f], (\mathbb{R}^n)^*)$, un réel $p^0 \in \{-1, 0\}$ et une forme linéaire $\lambda \in (\mathbb{R}^p)^*$ tels que $(p, p^0) \neq (0, 0)$ et tels que les équations suivantes sont vérifiées pour $t \in [t_0, t_f]$ p.p. :

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H}{\partial p}(t, x(t), p(t), p^0, u(t)), \\ \dot{p}(t) &= -\frac{\partial H}{\partial x}(t, x(t), p(t), p^0, u(t)), \\ 0 &= \frac{\partial H}{\partial u}(t, x(t), p(t), p^0, u(t)), \end{aligned} \quad (3.2)$$

où $H(t, x, p, p^0, u) := p \cdot f(t, x, u) + p^0 f^0(t, x, u)$ est le pseudo-hamiltonien associé à (P_B) . Les conditions terminales $c(t_f, x(t_f)) = 0$ sont vérifiées et on a en plus les conditions de transversalité suivantes :

$$p(t_f) = p^0 \frac{\partial g}{\partial x_f}(t_f, x(t_f)) + \sum_{i=1}^p \lambda_i \frac{\partial c_i}{\partial x_f}(t_f, x(t_f)), \quad (3.3)$$

Enfin, puisque le temps final est libre, si $u(\cdot)$ est continu en t_f , alors on a :

$$H[t_f] = - \left(p^0 \frac{\partial g}{\partial t_f}(t_f, x(t_f)) + \sum_{i=1}^p \lambda_i \frac{\partial c_i}{\partial t_f}(t_f, x(t_f)) \right), \quad (3.4)$$

où $[t] := (t, x(t), p(t), p^0, u(t))$.

Définition 3.3.2

- Une **extrémale (faible)** du Problème (P_B) est un quadruplet (x, p, p^0, u) solution des **équations pseudo-hamiltoniennes contraintes** (3.2).
- On parle de **BC-extrémale** (BC vient de “Boundary Conditions”) si l’extrémale vérifie en plus les **conditions aux limites** $c(t_f, x(t_f)) = 0$, $x(t_0) = x_0$, la **condition de transversalité** (3.3) et la **condition sur le hamiltonien** (3.4).
- Une extrémale (x, p, p^0, u) est dite **anormale** si $p^0 = 0$ et **normale** dans le cas contraire, i.e. $p^0 = -1$.

Remarque 3.3.1. Lorsque l’on résout en pratique les équations données par la Proposition 3.3.1, on explicite rarement la valeur de λ car si l’on détermine le vecteur adjoint, alors on peut retrouver facilement le multiplicateur λ . Ceci explique pourquoi dans la définition d’une BC-extrémale λ n’apparaît pas.

Remarque 3.3.2. La Figure 3.2 ci-avant nous donne une illustration du cas anormal.

Exemple 3.3.1. Soit un problème de Lagrange avec une contrainte terminale de la forme $c(x(t_f)) = 0_{\mathbb{R}^p}$, $p < n$ et c une submersion. Alors, la condition de transversalité $p(t_f) = \lambda \circ c'(x(t_f))$ signifie que $p(t_f) \perp T_{x(t_f)}M$, où $M := c^{-1}(\{0_{\mathbb{R}^p}\})$ est une sous-variété de \mathbb{R}^n .

Nous aurions donc pu définir les conditions de transversalité de manière plus générale en considérant une sous-variété M de \mathbb{R}^n et en demandant que $p(t_f) \perp T_{x(t_f)}M$. Cependant, dans les applications de ce manuscrit et dans de nombreuses applications pratiques, l'ensemble terminal M est défini **globalement** par une équation de la forme $c(x) = 0_{\mathbb{R}^p}$ et les conditions de transversalité données à partir de c sont plus pratiques à utiliser. \square

Exemple 3.3.2. Reprenons l'Exemple 3.1.2 et ajoutons au critère, une partie intégrale de la forme

$$\frac{1}{2} \int_0^{t_f} \|u(t)\|^2 dt.$$

Soit donc le problème de Bolza défini par $\Omega = \mathbb{R}^2$, $U = \mathbb{R}^2$, $g(x) = x_1$, $f^0(x, u) = \|u\|^2/2$, $f(x, u) = u$, $c(x) = \|x\|^2 - 1$ et $x_0 = 0 \in \mathbb{R}^2$. Fixons de plus $t_f = 1$. Puisque le système est complètement contrôlable, nous devons retrouver la même valeur pour $x(t_f)$ qu'à l'Exemple 3.1.2. Ce qui change dans ce nouveau cas, c'est que le contrôle optimal est unique. Voyons cela.

Le pseudo-hamiltonien est $H(x, p, u) = pu + p^0\|u\|^2/2$. Ainsi, la condition de stationarité du pseudo-hamiltonien nous donne $\partial_u H(x, p, u) = p + p^0 u = 0$. Si $p^0 = 0$ alors $p = 0$ aussi ce qui est impossible. On a donc $p^0 = -1$ et le contrôle optimal vérifie donc $u = p$. La dynamique sur le vecteur adjoint est $\dot{p} = 0$ donc p est constant. Notons w cette constante. La condition de transversalité est donc

$$p(t_f) = (-1, 0) + 2\lambda x(t_f) = w.$$

En résumé, nous sommes amenés à résoudre le système suivant

$$\begin{aligned} \dot{x}(t) &= u(t) = p(t), \\ \dot{p}(t) &= 0, \\ x(0) &= 0, \\ \|x(t_f)\| &= 1 \\ p_1(t_f) &= -1 + 2\lambda x_1(t_f), \\ p_2(t_f) &= 2\lambda x_2(t_f). \end{aligned}$$

Puisque $p(t) = w$ et $x(0) = 0$, on obtient $x(t) = wt$. Avec $t_f = 1$, on obtient $\|x(t_f)\| = \|w\| = 1$. La condition de transversalité devient $-1 + 2\lambda w_1 = w_1$ et $2\lambda w_2 = w_2$ qui se réécrit $w_1(1 - 2\lambda) = -1$ et $w_2(1 - 2\lambda) = 0$. Ceci implique que $w_2 = 0$. Puisque $\|w\| = 1$, on a $w_1 = \pm 1$. Il y a toujours 2 possibilités mais celle qui minimise le coût est $w_1 = -1$. En conclusion, on a $u(t) = p(t) = (-1, 0)$, $x(t) = (-t, 0)$, $p^0 = -1$ et $\lambda = 0$. \square

Exemple 3.3.3. Dans l'Exemple 3.1.2 la condition de transversalité nous donnait

$$p(t_f) = -g'(x(t_f)) + \lambda c'(x(t_f)) = (-1, 0) + 2\lambda x(t_f) = 0$$

et nous avons remarqué que ceci était la condition nécessaire d'optimalité à l'ordre 1 du problème de minimisation (3.1). En ajoutant un terme intégral dans le coût nous avons vu dans l'exercice précédent que la condition de transversalité devenait

$$p(t_f) = -g'(x(t_f)) + \lambda c'(x(t_f)) = (-1, 0) + 2\lambda x(t_f) = w.$$

Nous pouvons alors nous demander à quel problème de minimisation cette condition correspond. Dans l'exemple précédent, le terme intégral était

$$\frac{1}{2} \int_0^1 \|u(t)\|^2 dt.$$

Le contrôle optimal vérifiait $u(t) = w$ et la trajectoire $x(t) = wt$. Introduisons le critère

$$F(w) := g(x(t_f)) + \frac{1}{2} \int_0^1 \|u(t)\|^2 dt = w_1 + \frac{1}{2} \|w\|^2.$$

Posons le problème $\min_{w \in \mathbb{R}^n} \{F(w) \mid c(w) = 1\}$. La CN1 de ce problème s'écrit

$$0 = -F'(w) + \lambda c'(w) = (-1, 0) - w + 2\lambda w,$$

ce qui correspond exactement à la condition de transversalité. \square

3.4 Condition de Legendre et conservation du hamiltonien

Condition de Legendre. Le long d'une extrémale est vérifiée la condition nécessaire à l'ordre 1 suivante, de stationarité du hamiltonien par rapport à la variable u :

$$\frac{\partial H}{\partial u}(t, x(t), p(t), p^0, u(t)) = 0.$$

Appelons **condition d'Euler-Hamilton** cette condition. La condition d'Euler-Hamilton est obtenue à partir de la différentielle de l'application entrée/sortie, non-augmentée dans la formulation de Mayer et augmentée dans le cas Lagrange ou Bolza, combinée au Théorème (A.7.3). Le point clé de ce résultat est le Théorème (A.7.2) d'application ouverte non linéaire à l'ordre 1. D'une manière analogue, à partir d'un résultat à l'ordre 2 d'application ouverte non linéaire, voir [1, Théorème 20.3] ou Théorème 3.6.1, nous pouvons obtenir une condition nécessaire à l'ordre 2, en utilisant la différentielle d'ordre 2 de l'application entrée/sortie. Cette condition, voir [1, Proposition 20.11], est appelée **condition de Legendre** (ou de Legendre-Clebsch) et implique que :

$$\frac{\partial^2 H}{\partial u^2}(t, x(t), p(t), p^0, u(t)) \cdot (v, v) \leq 0, \quad \forall v \in \mathbb{R}^m, \quad \forall t \in [t_0, t_f] \text{ p.p.}$$

Voir aussi Théorème (B.4.2) pour le cas particulier du calcul des variations. Soit $t \in [t_0, t_f]$, on note $F_t(u) := H(t, x(t), p(t), p^0, u)$. Dire que les conditions d'Euler-Hamilton et celle de Legendre sont vérifiées au temps t , c'est dire que $u(t)$ vérifie les conditions nécessaires aux ordres 1 et 2 du problème de maximisation sans contraintes :

$$\max_{u \in \Pi} F_t(u). \tag{3.5}$$

Si de plus, la condition de Legendre est stricte, alors $u(t)$ vérifie la condition suffisante à l'ordre 2 d'optimalité locale stricte pour le Problème (3.5). **Attention**, même si la condition suffisante est vérifiée sur tout l'intervalle $[t_0, t_f]$ p.p., i.e. si

$$\frac{\partial^2 H}{\partial u^2}(t, x(t), p(t), p^0, u(t)) \cdot (v, v) < 0, \quad \forall v \in \mathbb{R}^m, \quad \forall t \in [t_0, t_f] \text{ p.p.}$$

en aucun cas, cela nous donne une condition suffisante d'optimalité locale pour le problème de contrôle optimal ! Il faut une condition supplémentaire appelée **condition de Jacobi**. Voir Définition 3.6.5 ou [1] pour plus de détails. Ceci même en calcul des variations. Voir Section B.6 ou [65] pour plus de détails.

Discussion sur la régularité des extrémales. Nous allons voir que dans le cas où la condition de Legendre est stricte, alors la régularité des applications f et f^0 nous donne la régularité des extrémales. Supposons f et f^0 de classe \mathcal{C}^k , $k \in \mathbb{N}^*$. Alors,

$$(t, z, p^0, u) \mapsto H(t, z, p^0, u) = p \cdot f(t, x, u) + p^0 f^0(t, x, u)$$

est aussi de classe \mathcal{C}^k , avec $z := (x, p)$. Supposons $k \geq 2$, fixons p^0 et considérons un triplet $(\bar{t}, \bar{z}, \bar{u})$ tel que $\partial_u H(\bar{t}, \bar{z}, p^0, \bar{u}) = 0$ et $\partial_{uu}^2 H(\bar{t}, \bar{z}, p^0, \bar{u})$ est inversible. Si l'extrémale vérifie la condition de Legendre, alors $\partial_{uu}^2 H(\bar{t}, \bar{z}, p^0, \bar{u})$ est inversible si et seulement si l'extrémale vérifie la condition de Legendre stricte. Sous ces hypothèses, d'après le théorème des fonctions implicites, on a au voisinage de $(\bar{t}, \bar{z}, \bar{u})$ une fonction implicite $u(t, z)$ de classe \mathcal{C}^{k-1} telle que sur ce voisinage $\partial_u H(t, z, p^0, u(t, z)) = 0$. Considérons maintenant une extrémale $(z(\cdot), p^0, u(\cdot))$ et supposons que le contrôle vérifie $u(t) = u(t, z(t))$. Alors, le long de l'extrémale, on a

$$\dot{z}(t) = \vec{H}(t, z(t), u(t)) = \vec{H}(t, z(t), u(t, z(t)))$$

où l'on a introduit la notation

$$\vec{H}(t, z, u) := \left(\frac{\partial H}{\partial p}(t, z, p^0, u), -\frac{\partial H}{\partial x}(t, z, p^0, u) \right).$$

Posons

$$F(t, z) := \vec{H}(t, z, u(t, z))$$

et on se demande quelle est la régularité de la fonction F . Par nos hypothèses, H est \mathcal{C}^k donc \vec{H} est \mathcal{C}^{k-1} . Par composition avec $(t, z) \mapsto u(t, z)$, F est donc \mathcal{C}^{k-1} . Puisque, $\dot{z}(t) = F(t, z(t))$, alors la solution $t \mapsto z(t)$ est de classe \mathcal{C}^k .

Conservation du hamiltonien. Nous allons voir que dans le cadre précédent, le pseudo-hamiltonien est constant le long de l'extrémale si le système augmenté est autonome. Par nos hypothèses, H est \mathcal{C}^k , donc

$$t \mapsto M(t) := H(t, z(t), p^0, u(t, z(t)))$$

est de classe \mathcal{C}^{k-1} , car même si $z(\cdot)$ est de classe \mathcal{C}^k , $(t, z) \mapsto u(t, z)$ est lui de classe \mathcal{C}^{k-1} . On rappelle que l'on suppose $k \geq 2$. On a M de classe \mathcal{C}^{k-1} , on peut donc calculer sa dérivée :

$$\begin{aligned} M'(t) &= \frac{\partial H}{\partial t}(t, z(t), p^0, u(t, z(t))) + \frac{\partial H}{\partial z}(t, z(t), p^0, u(t, z(t))) \cdot \dot{z}(t) \\ &\quad + \frac{\partial H}{\partial u}(t, z(t), p^0, u(t, z(t))) \cdot \left(\frac{\partial u}{\partial t}(t, z(t)) + \frac{\partial u}{\partial z}(t, z(t)) \cdot \dot{z}(t) \right) \\ &= \frac{\partial H}{\partial t}(t, z(t), p^0, u(t, z(t))) + \frac{\partial H}{\partial z}(t, z(t), p^0, u(t, z(t))) \cdot \dot{z}(t) \\ &= \frac{\partial H}{\partial t}(t, z(t), p^0, u(t, z(t))), \end{aligned}$$

donc si le système augmenté est autonome, *i.e.* si H ne dépend pas de t , alors H est constant le long d'une extrémale.

3.5 Exemples

3.5.1 Simple intégrateur et fonction de tir

Considérons le problème de contrôle optimal suivant.

$$\begin{cases} \min J(x, u) := \frac{1}{2} \int_0^1 u^2(t) dt \\ \dot{x}(t) = -x(t) + u(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, 1] \text{ p.p.}, \\ x(0) = -1, \quad x(1) = 0, \end{cases}$$

avec $\forall t \in [0, 1], x(t) \in \mathbb{R}$. Appliquons les conditions nécessaires, cf. Proposition 3.2.1. Notons

$$H(x, p, p^0, u) := p(-x + u) + \frac{1}{2} p^0 u^2,$$

le pseudo-hamiltonien associé au problème étudié. Si u est optimal (avec x la trajectoire associée) alors il existe un vecteur adjoint p et un scalaire $p^0 \in \{0, 1\}$ tels que $(p, p^0) \neq 0$ et les équations suivantes sont vérifiées pour $t \in [0, 1]$ p.p. :

$$\begin{cases} \dot{x}(t) &= \partial_p H[t] = -x(t) + u(t), \\ \dot{p}(t) &= -\partial_x H[t] = p(t), \\ 0 &= \partial_u H[t] = p(t) + p^0 u(t), \end{cases}$$

où $[t] := (x(t), p(t), p^0, u(t))$. On ne peut pas avoir $p^0 = 0$ car sinon $p \equiv 0$ ce qui est impossible. Ainsi $p^0 = 1$ et $u(t) = -p(t)$. On obtient ainsi le problème aux deux bouts suivant à résoudre :

$$\begin{cases} \dot{x}(t) &= -x(t) + p(t), \\ \dot{p}(t) &= p(t), \\ x(0) &= -1 =: x_0, \quad x(1) = 0 =: x_f, \end{cases}$$

et le contrôle a disparu. L'inconnue de ce problème aux deux bouts est le vecteur adjoint initial. En effet si l'on fixe $p_0 := p(0)$ alors d'après le théorème de Cauchy-Lipschitz, il existe une unique solution maximale $z := (x, p)$ vérifiant la dynamique (sur x et p) et la condition initiale $z(0) = (x_0, p_0)$. Le problème est donc de trouver p_0 tel que $x(1) = x_f$. Calculons la solution.

$$\dot{p}(t) = p(t) \implies p(t) = e^t p_0 \implies x(t) = (0.5 p_0 e^{2t} + C) e^{-t}.$$

Or $x(0) = x_0$ donc $x(t) = (0.5 p_0 (e^{2t} - 1) + x_0) e^{-t} = p_0 \text{sh}(t) + x_0 e^{-t}$ et finalement, puisque $x(t_f) = x_f$ avec $t_f = 1$,

$$p_0 = \frac{x_f - x_0 e^{-t_f}}{\text{sh}(t_f)} = \frac{e^{-1}}{\text{sh}(1)} = \frac{2}{e^2 - 1} \approx 0.3130.$$

Remarque 3.5.1. Il faut bien comprendre qu'avec p_0 , on a tout ! On est donc passé de la recherche d'une loi de commande (ou contrôle) $u(\cdot)$ à un vecteur de \mathbb{R}^n . Ici $n = 1$. En effet, on a bien tout, car avec (x_0, p_0) , on peut calculer $(x(t), p(t))$ pour tout temps t et donc on a aussi le contrôle $u(t) = p(t)$ à chaque instant.

Remarque 3.5.2. Pour calculer p_0 , on a résolu le système affine en p_0 à 1 équation $S(p_0) = 0$, avec

$$\begin{aligned} S: \mathbb{R} &\longrightarrow \mathbb{R} \\ p_0 &\longmapsto S(p_0) := \Pi_x(z(t_f, x_0, p_0)) - x_f = p_0 \operatorname{sh}(t_f) + x_0 e^{-t_f} - x_f, \end{aligned}$$

où $z(\cdot, x_0, p_0)$ est la solution du problème de Cauchy

$$\dot{z}(t) = (-x(t) + p(t), p(t)), \quad z(0) = (x_0, p_0),$$

avec $z = (x, p)$ et où $\Pi_x(z) = x$. La fonction S est appelée **fonction de tir** (shooting en anglais ce qui explique la notation S).

On peut retrouver Figure 3.3 le flot entre les temps 0 et 1 du système hamiltonien $\dot{x} = -x + p$, $\dot{p} = p$, restreint à $x(0) = -1$, avec quelques fronts d'onde en vert et l'extrémale solution en rouge. On remarque que les fronts d'onde sont linéaires. On ajoute un paramètre α dans le système contrôlé permettant de modéliser des non-linéarités venant du frottement par exemple. On considère donc le système $\dot{x} = -x + \alpha x^2 + u$ ce qui nous donne comme nouveau système hamiltonien $\dot{x} = -x + \alpha x^2 + p$, $\dot{p} = (1 - 2\alpha x)p$. On retrouve Figure 3.4 le flot entre les temps 0 et 1 de ce système hamiltonien restreint à $x(0) = -1$ pour $\alpha = 1.5$. On remarque que les fronts d'onde ne sont pas linéaires.

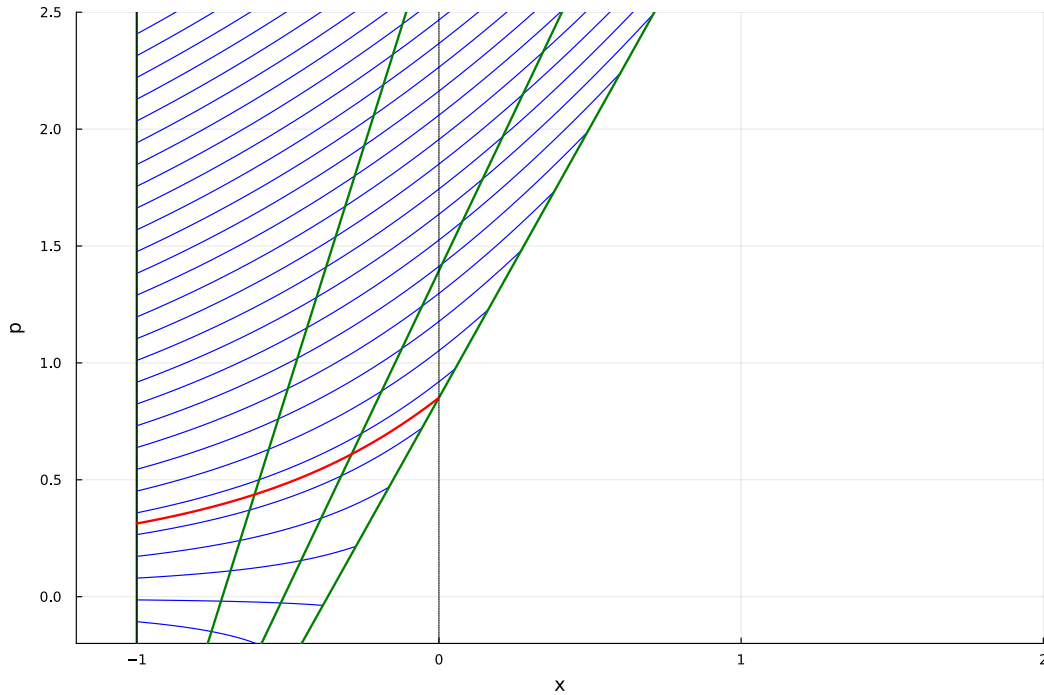


FIGURE 3.3 – Flot entre les temps 0 et 1 du système $\dot{x} = -x + p$, $\dot{p} = p$, restreint à $x(0) = -1$. Les courbes bleues sont des extrémales pour différentes valeurs de $p(0)$ entre -0.2 et 2.5 . La courbe rouge est l'extrémale solution. Les courbes vertes sont les fronts d'onde du flot aux temps 0, $1/3$, $2/3$ et 1.

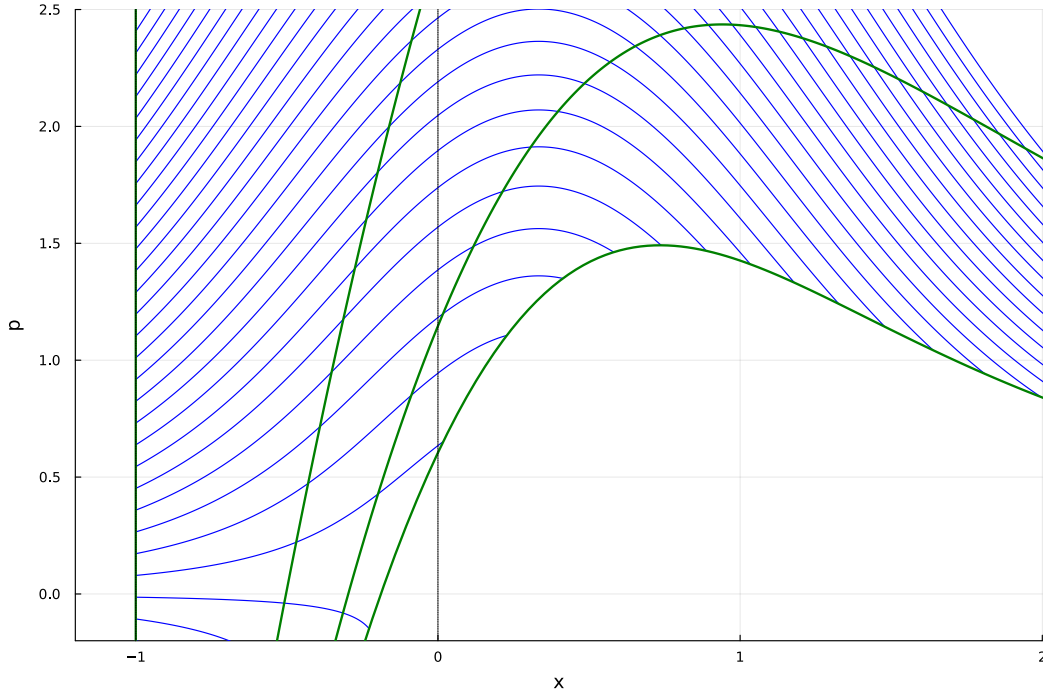


FIGURE 3.4 – Flot entre les temps 0 et 1 du système $\dot{x} = -x + \alpha x^2 + p$, $\dot{p} = (1 - 2\alpha x)p$, restreint à $x(0) = -1$, pour $\alpha = 1.5$. Les courbes bleues sont des extrémales pour différents $p(0)$ entre -0.2 et 2.5 . Les courbes vertes sont les fronts d'onde aux temps 0, $1/3$, $2/3$ et 1.

3.5.2 Calcul des variations

Considérons le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u) := \int_0^{t_f} L(x(t), u(t)) \, dt \\ \dot{x}(t) = u(t), \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

où L est une application lisse appelée lagrangien, où $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^n$ et $t_f > 0$ est fixé. Notons

$$H(x, p, p^0, u) := p u + p^0 L(x, u)$$

le pseudo-hamiltonien associé au problème. Si u est solution, avec x la trajectoire associée, alors d'après la Proposition 3.2.1 il existe un covecteur p absolument continu et un scalaire $p^0 \in \{-1, 0\}$ tels que $(p, p^0) \neq 0$ et les équations suivantes sont vérifiées pour presque tout $t \in [0, t_f]$:

$$\begin{cases} \dot{x}(t) &= \frac{\partial H}{\partial p}[t] = u(t), \\ \dot{p}(t) &= -\frac{\partial H}{\partial x}[t] = -p^0 \frac{\partial L}{\partial x}(x(t), u(t)), \\ 0 &= \frac{\partial H}{\partial u}[t] = p(t) + p^0 \frac{\partial L}{\partial u}(x(t), u(t)), \end{cases}$$

où $[t] := (x(t), p(t), p^0, u(t))$. Si $p^0 = 0$, alors $p \equiv 0$, or $(p, p^0) \neq (0, 0)$, donc $p^0 = -1$. On a donc, puisque $\dot{x}(t) = u(t)$,

$$\dot{p}(t) = \frac{\partial L}{\partial x}(x(t), \dot{x}(t)) \quad \text{et} \quad p(t) = \frac{\partial L}{\partial u}(x(t), \dot{x}(t)).$$

On retrouve ce que l'on appelle les **équations de Euler-Lagrange** :

$$\frac{\partial L}{\partial x}(x(t), \dot{x}(t)) - \frac{d}{dt} \left(\frac{\partial L}{\partial u}(x(t), \dot{x}(t)) \right) = 0, \quad (3.6)$$

qui sont les conditions nécessaires, voir Annexe B, du problème de calcul des variations dans sa formulation la plus classique :

$$\text{minimiser } \int_0^{t_f} L(x(t), \dot{x}(t)) dt : x \in \mathcal{C}^1([0, t_f], \Omega), x(0) = x_0, x(t_f) = x_f.$$

Supposons que l'on puisse appliquer le théorème des fonctions implicites à l'équation $p - \partial_u L(x, u) = 0$, pour obtenir u en fonction de (x, p) et notons $u(x, p)$ la fonction implicite. Introduisons le hamiltonien

$$\mathbf{H}(x, p) := H(x, p, -1, u(x, p)) = p u(x, p) - L(x, u(x, p))$$

et la notation $z = (x, p)$, alors, les équations de Euler-Lagrange sont équivalentes aux **équations de Hamilton**

$$\dot{z}(t) = \vec{\mathbf{H}}(z(t)).$$

Exercices. Les exercices de cette sous-section concernent la résolution de problèmes de calcul des variations. On pourra utiliser les équations de Euler-Lagrange (3.6). On appelle ici extrémale, une trajectoire x solution des équations de Euler-Lagrange (3.6). Une BC-extrémale est une extrémale qui vérifie les conditions aux limites $x(0) = x_0$ et $x(t_f) = x_f$.

Exercice 3.5.1 (solution p. 95) : (cf. [27].)

1. Montrer que les extrémales de $L(t, x, v) := \sqrt{1 + v^2}$ sont affine (*i.e.* $\ddot{x} = 0$).
2. Montrer que l'équation de Euler-Lagrange pour le lagrangien $L(t, x, v) := x^2 + v^2$ est donnée par $\ddot{x} - x = 0$.
3. Trouver l'unique extrémale admissible (*i.e.* BC-extrémale) pour le problème :

$$\inf_x \left\{ J(x) := \int_0^1 (\dot{x}^2(t) + x^2(t)) dt \mid x \in \mathcal{C}^1([0, 1], \mathbb{R}), x(0) = 0, x(1) = 1 \right\}.$$

Exercice 3.5.2 (solution p. 96) : Soit $L(x, v) := \frac{1}{2} \|v\|^2$, avec $\|\cdot\|$ la norme euclidienne. Soient A, B dans \mathbb{R}^n et $T > 0$ dans \mathbb{R} . On considère le problème variationnel

$$\inf_x \left\{ J(x) := \int_0^T L(x(t), \dot{x}(t)) dt \mid x \in \mathcal{C}^1([0, T], \mathbb{R}^n), x(0) = A, x(T) = B \right\}.$$

1. Trouver l'unique BC-extrémale.
2. Calculer pour la BC-extrémale notée x :

$$l(x) := \int_0^T \|\dot{x}(t)\| dt.$$

3. Que pensez-vous que $l(x)$ représente pour les points A et B ?

Exercice 3.5.3: (Paradoxe des jumeaux de Langevin). Soit $c > 0$ une constante réelle, on introduit le lagrangien

$$L(q, u) := c^2 u_2^2 - u_1^2$$

où l'on a noté $u := (u_1, u_2) \in \mathbb{R}^2$ le vecteur vitesse et $q := (x, t) \in \mathbb{R}^2$ les coordonnées.

Attention, ici t est une variable d'état. Plus classiquement, en Physique, les composantes de u sont notées dx et dt .

On introduit le problème variationnel :

$$\sup_q \left\{ J(q) := \int_0^1 L(q(s), \dot{q}(s)) ds \mid q \in \mathcal{C}^1([0, 1], \mathbb{R}^2), q(0) = A, q(1) = B \right\}, \quad (3.7)$$

où $A := (0, 0)$ et $B := (\Delta x, \Delta t)$, et où Δx , et Δt sont deux constantes dans \mathbb{R}_+^* .

1. Le lagrangien L est-il autonome ?
2. Calculer l'unique BC-extrémale de (3.7).
3. Soient q la BC-extrémale et $v \in]0, c[$. On suppose $\Delta x = v\Delta t$. Calculer la **longueur** de q ,

$$l(q) := \int_0^1 \sqrt{L(q(s), \dot{q}(s))} ds,$$

en fonction de c , Δt et du **facteur de Lorentz**

$$\gamma := \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

4. On note $\tau := \frac{l(q)}{c}$. Que vaut $\frac{\tau}{\Delta t}$? A quel intervalle appartient $\frac{\tau}{\Delta t}$?
5. On note $C := (0, 2\Delta t)$. Soit $\bar{q} := (\bar{x}, \bar{t})$ la solution du problème variationnel (3.7) où l'on remplace la condition finale par $q(1) = C$. Donner $\bar{q}(s)$ et calculer la longueur $l(\bar{q})$.
6. **Paradoxe des jumeaux.** Igor part du point A et rejoint B à la vitesse constante $v \in]0, c[$, c'est-à-dire en suivant la trajectoire q . Il fait demi-tour en B à la vitesse $-v$ (on précise que x est une coordonnée d'espace tandis que t est une coordonnée de temps) jusqu'à rejoindre le point C . Pendant ce temps, son frère jumeau Grichka

est resté en $x = 0$, il a donc suivi la trajectoire \bar{q} . Sachant que l'âge de chacun des deux frères est la longueur du chemin qu'il parcourt divisée par c , quels sont les âges de Igor (en fonction de τ) et Grichka (en fonction de Δt) lorsqu'ils se retrouvent au point C ?

7. Qui est le plus âgé au point C ?

8. Soient $v = \frac{4}{5}c$ et $\Delta t = 25$ années. Quels âges ont Igor et Grichka en C ?

Remarque 3.5.3. L'exercice précédent est placé ici dans le cadre du calcul des variations mais c'est en fait un problème de géométrie, non riemannienne, mais lorentzienne, issu de la théorie de la relativité restreinte [39]. La différence vient du fait que le tenseur métrique,¹ dans l'exercice le tenseur métrique est L , n'est pas définie positif mais indéfini, ici de signature $(+, -)$. Dans ce cas, il existe des vecteurs isotropes non nuls, c'est-à-dire tels que $\|v\|_g = 0$, $v \neq 0$. C'est le cas des vecteurs de la forme $(\lambda, c\lambda)$, $\lambda \in \mathbb{R}^*$. Ces vecteurs sont appelés de *genre lumière* et n'ont donc pas d'âge (l'âge est appelé temps propre en relativité restreinte), dans le sens qu'il vaut 0.

3.5.3 Principe fondamental de la dynamique

Considérons un point matériel de masse m évoluant dans un champ de force dérivant d'un potentiel. Le lagrangien associé est la fonction

$$L(x, v) := T(v) - V(x),$$

où $T(v) := \frac{1}{2}m\|v\|^2$ est *l'énergie cinétique* et $V(x)$ est *l'énergie potentielle* du point matériel. La position-vitesse du point matériel étant donnée par le couple (x, v) . Rappelons que l'énergie mécanique est donnée par :

$$E(x, v) := T(v) + V(x).$$

Voici quelques exemples d'énergies potentielles où la variable de position est respectivement r , α et x :

- énergie potentielle gravitationnelle (déplacement d'un satellite autour de la Terre) :

$$E_{pg}(r) = -G \frac{mM_T}{r} + \text{cste},$$

où G est la constante de gravitation, M_T la masse de la Terre et r la distance au centre de la Terre ;

- énergie potentielle de pesanteur (pendule) :

$$E_{pp}(\alpha) = -mgl \cos(\alpha) + \text{cste},$$

où g est la pesanteur, l la longueur du pendule et α l'angle avec la verticale ;

1. En géométrie riemannienne, le tenseur métrique est appelé plus simplement métrique tandis qu'en géométrie lorentzienne, on parle plutôt de pseudo-métrique. Dans notre cas, le tenseur métrique est celui de Minkowski qui s'écrit en général $g = u_2^2 - u_1^2$ et on considère des coordonnées de la forme (x, ct) .

- énergie potentielle élastique (k est la raideur du ressort) :

$$E_{\text{pe}}(x) = \frac{1}{2}kx^2 + \text{cste.}$$

Hamilton observa que le mouvement de ce point matériel est solution d'un problème variationnel.

Proposition 3.5.1 – Principe Fondamental de la Dynamique

Les équations du Principe Fondamental de la Dynamique (PFD) :

$$m\ddot{x}(t) = -\nabla V(x(t)),$$

qui gouvernent le mouvement coïncident avec les équations de Euler-Lagrange (3.6) associé au lagrangien $L = T - V$.

► On a

$$\begin{aligned} \nabla_x L(x(t), \dot{x}(t)) - \frac{d}{dt} (\nabla_v L(x(t), \dot{x}(t))) &= -\nabla V(x(t)) - \frac{d}{dt} (\nabla T(\dot{x}(t))) \\ &= -\nabla V(x(t)) - \frac{d}{dt} (m \dot{x}(t)), \end{aligned}$$

donc les équations de Euler-Lagrange correspondent aux équations du PFD. ■

Le vecteur adjoint (en colonne) est donné par $p = \nabla_v L(x, v) = mv$ donc celui-ci correspond à ce que l'on appelle la quantité de mouvement. Le hamiltonien correspond ici à l'énergie mécanique :

$$\begin{aligned} \mathbf{H}(x, p) &= (p | v) - L(x, v) = \frac{1}{m} \|p\|^2 - L(x, \frac{1}{m}p) \\ &= \frac{1}{2m} \|p\|^2 + V(x) = T(v) + V(x) = E(x, v). \end{aligned}$$

Les équations de Euler-Lagrange nous donne le système hamiltonien

$$\begin{cases} \dot{x}(t) = v(t) = \frac{1}{m}p(t) = \nabla_p \mathbf{H}(x(t), p(t)), \\ m\dot{v}(t) = \dot{p}(t) = -\nabla V(x(t)) = -\nabla_x \mathbf{H}(x(t), p(t)). \end{cases}$$

Enfin, on retrouve la conservation du hamiltonien (on note $[t] := (x(t), p(t))$) :

$$\begin{aligned} \frac{d}{dt} \mathbf{H}[t] &= \frac{\partial}{\partial x} \mathbf{H}[t] \cdot \dot{x}(t) + \frac{\partial}{\partial p} \mathbf{H}[t] \cdot \dot{p}(t) \\ &= (\nabla_x \mathbf{H}[t] | \nabla_p \mathbf{H}[t]) - (\nabla_p \mathbf{H}[t] | \nabla_x \mathbf{H}[t]) \\ &= \{\mathbf{H}, \mathbf{H}\}[t] \\ &= 0. \end{aligned}$$

Exercice 3.5.4 (solution p. 96) : Soit le lagrangien

$$L(\alpha, \dot{\alpha}) := T(\dot{\alpha}) - V(\alpha),$$

avec

$$T(\dot{\alpha}) := \frac{1}{2}ml^2\dot{\alpha}^2 \quad \text{et} \quad V(\alpha) := -mgl \cos \alpha,$$

où m , g et l sont des constantes strictement positives. Nous considérons donc le problème du pendule non amorti, non linéaire. Voir les Figures 3.5, 3.6 et 3.7 pour plus de détails.

1. Ecrire le hamiltonien $H(\alpha, p)$ associé à L en fonction de $I := ml^2$, $\omega := \sqrt{g/l}$, α et p .

2. Rappeler pourquoi H est constant le long de toute extrémale.

3. Soit (α, p) une extrémale. Supposons que $H = h \in \mathbb{R}$ le long de cette extrémale. Montrer que :

$$\forall t : -\omega^2 I \cos \alpha(t) \leq h.$$

4. Le cas limite entre les rotations et les oscillations est donné par $h = \omega^2 I$. Soit α_0 l'angle initial, c'est-à-dire $\alpha(0) = \alpha_0$. Donner $\dot{\alpha}_0$ et p_0 , respectivement les vitesse et adjoint initiaux dans le cas limite.

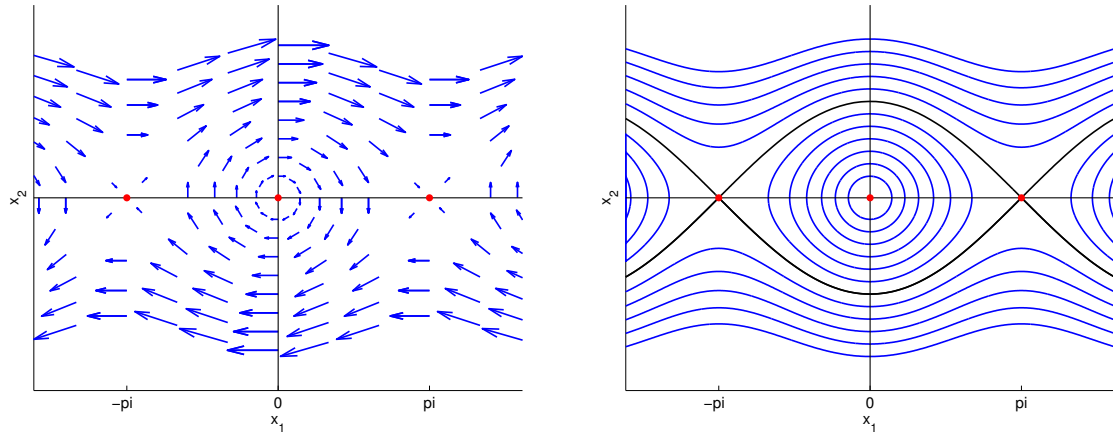


FIGURE 3.5 – Le champ de vitesses de phase du pendule est visible sur le sous-graphe de gauche. Les courbes de phase (ou orbites) en bleu sont données sur le sous-graphe de droite. Les courbes de phase fermées correspondent à des oscillations tandis que les autres correspondent aux rotations. Les courbes en noir sont appelées *séparatrices*. Enfin, on distingue trois points d'équilibre en rouge, 1 stable et 2 instables.

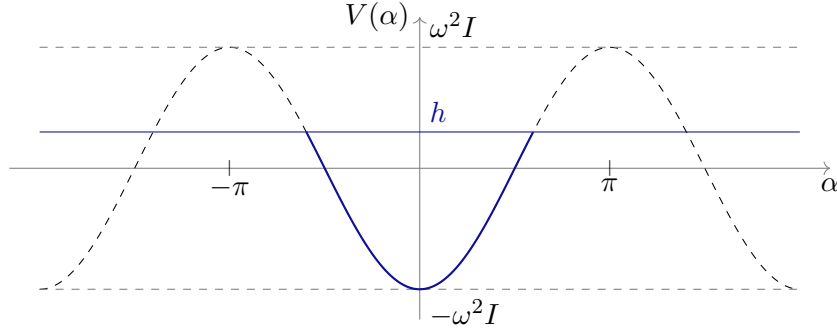


FIGURE 3.6 – Potentiel du pendule et oscillation.

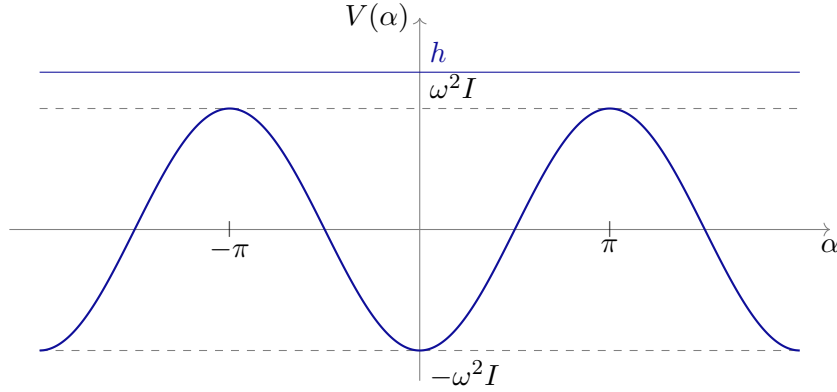


FIGURE 3.7 – Potentiel du pendule et rotation.

3.6 Condition du second ordre faible : cas Lagrange

Nous considérons le Problème (P_L) sous les Hypothèses (H2). Comme nous l'avons vu, si \bar{u} est solution, alors, nécessairement

$$\tilde{E}_{t_f, \tilde{x}_0}(\bar{u}) \in \text{Front}(\tilde{\mathcal{A}}(t_f, \tilde{x}_0)),$$

et donc nous ne sommes pas seulement \bar{u} est singulier pour l'application entrée/sortie augmentée, mais aussi celle-ci n'est pas localement ouverte en \bar{u} . Le fait que \bar{u} soit singulier nous assure l'existence d'une forme linéaire $\bar{\lambda} \in (\mathbb{R}^{n+1})^*$ non nulle orthogonale à $\text{Im } \tilde{E}'_{t_f, \tilde{x}_0}(\bar{u})$, c'est-à-dire

$$\bar{\lambda} \in \left(\text{Im } \tilde{E}'_{t_f, \tilde{x}_0}(\bar{u}) \right)^\perp, \quad \bar{\lambda} \neq 0.$$

Le fait que celle-ci ne soit pas localement ouverte en \bar{u} va nous permettre d'obtenir de nouvelles conditions nécessaires d'optimalité. Pour cela, nous avons besoin du théorème suivant, analogue au Théorème A.7.2, mais à l'ordre 2.

Théorème 3.6.1 – de l'application ouverte non linéaire à l'ordre 2

Soit $F: E \rightarrow \mathbb{R}^n$ une application lisse définie sur un espace de Banach E . Soit u un point singulier de corang un ($\text{codim Im } F'(u) = 1$). Soit $\lambda \in (\mathbb{R}^n)^* \setminus 0$ dans $(\text{Im } F'(u))^\perp$. Si $\lambda F''(u)$ est indéfinie sur $\text{Ker } F'(u)$, alors F est localement ouverte en u .

La forme bilinéaire

$$\lambda F''(u) \in \mathcal{L}(E, \mathcal{L}(E, \mathbb{R})) \simeq \mathcal{L}_2(E \times E, \mathbb{R})$$

est appelée la **dérivée seconde intrinsèque** de F en u et est définie à un scalaire près dans le cas de corang un. Par la suite, on suppose que le contrôle \bar{u} est de corang un. Ainsi, la trajectoire associée \bar{x} admet un unique (à un scalaire près) relèvement $(\bar{x}, \bar{p}, p^0, \bar{u})$ sur $[0, t_f]$, que l'on suppose être normal ($p^0 \neq 0$). Dans notre situation, on applique le théorème précédent à l'application entrée/sortie augmentée. On a donc

$$\tilde{\lambda} \tilde{E}_{t_f, \tilde{x}_0}''(\bar{u}) \text{ semi-définie sur } \text{Ker } \tilde{E}_{t_f, \tilde{x}_0}'(\bar{u}).$$

Remarque 3.6.1. Faisons le lien avec le point de vue de l'optimisation. Considérons le Problème (P_L) sous sa forme réduite et définissons le lagrangien (on omet les indices des applications entrée/sortie) :

$$L(u, \tilde{\lambda}) = \lambda^0 \pi_{x^0}(\tilde{E}(u)) + \lambda \cdot (E(u) - x_f), \quad \tilde{\lambda} = (\lambda, \lambda^0) \in (\mathbb{R}^n)^* \times \mathbb{R}_-.$$

Notons $x^0(u) := \pi_{x^0}(\tilde{E}(u))$ et $x(u) := E(u)$ de telle sorte que $\tilde{E}(u) = (x(u), x^0(u))$. D'un point de vue optimisation, la condition nécessaire d'optimalité à l'ordre 2 est que $\partial_{uu}^2 L(\bar{u}, \tilde{\lambda}) = \tilde{\lambda} \tilde{E}''(\bar{u})$ soit semi-définie négative sur l'espace tangent aux contraintes. Dans le cas où l'hypothèse de qualification des contraintes est vérifiée ($\lambda^0 < 0$), l'espace tangent est donné par $\text{Ker } x'(\bar{u})$ et non par $\text{Ker } \tilde{E}'(\bar{u})$. Mais finalement ces deux noyaux sont égaux puisque pour tout $v \in \text{Ker } x'(\bar{u})$, on a d'après la CN1 (A.7.3) :

$$0 = \frac{\partial L}{\partial u}(\bar{u}, \tilde{\lambda}) \cdot v = \lambda^0 x^{0'}(\bar{u}) \cdot v + \lambda \cdot (x'(\bar{u}) \cdot v) = \lambda^0 x^{0'}(\bar{u}) \cdot v$$

et donc $x^{0'}(\bar{u}) \cdot v = 0$ puisque $\lambda^0 \neq 0$. Ainsi, $\text{Ker } x'(\bar{u}) \subset \text{Ker } \tilde{E}'(\bar{u})$. De plus, $\text{Ker } \tilde{E}'(\bar{u}) \subset \text{Ker } x'(\bar{u})$ puisque $\tilde{E}(u) = (x(u), x^0(u))$. Finalement, $\text{Ker } \tilde{E}'(\bar{u}) = \text{Ker } x'(\bar{u})$ et donc la condition nécessaire d'optimalité à l'ordre 2 est bien que $\tilde{\lambda} \tilde{E}''(\bar{u})$ soit semi-définie (négative) sur $\text{Ker } \tilde{E}'(\bar{u})$.

Pour $t \in [0, t_f]$, on définit la forme bilinéaire symétrique

$$B_t := \tilde{\lambda} \tilde{E}_{t, \tilde{x}_0}''(\bar{u})$$

et on introduit $K_t := \text{Ker } \tilde{E}_{t, \tilde{x}_0}'(\bar{u})$. Nous avons le résultat suivant.

Proposition 3.6.2 – [21]

Si l'extrémale $(\bar{x}, \bar{p}, -1, \bar{u})$ vérifie la condition forte de Legendre, alors il existe $\varepsilon > 0$ tel que $B_t|_{K_t}$ est définie négative pour tout $t \in [0, \varepsilon]$.

Clairement, si $s \leq t$, alors $B_t|_{K_t} \prec 0$ (i.e. définie négative) implique $B_s|_{K_s} \prec 0$, d'où la définition suivante. On définit le **premier instant conjugué** t_{1c} , le long d'une extrémale normale vérifiant la condition forte de Legendre, comme le supremum des temps t tels que B_t soit définie négative :

$$t_{1c} = \sup \{t > 0 \mid B_t|_{K_t} \prec 0\}.$$

D'après [1], alors $B_{t_{1c}}|_{K_{t_{1c}}}$ a un noyau non trivial. De fait, les temps conjugués sont définis comme les temps t_c tels que $B_{t_c}|_{K_{t_c}}$ est dégénérée.

Reprenons notre extrémale de référence $(\bar{x}, \bar{p}, -1, \bar{u})$ et supposons qu'elle vérifie la condition forte de Legendre :

$$\frac{\partial^2 H}{\partial u^2}(\bar{x}(t), \bar{p}(t), -1, \bar{u}(t)) \cdot (v, v) < 0, \quad \forall v \in \mathbb{R}^m, \quad \forall t \in [0, t_f] \text{ p.p..}$$

Sous cette condition, l'équation $\partial_u H = 0$ peut être résolue dans un voisinage de l'extrémale et on peut donc définir le contrôle comme une fonction de l'état et du vecteur adjoint, notée $u(x, p)$. En définissant sur ce voisinage le hamiltonien $\mathbf{H}(z) := H(z, -1, u(z))$, $z := (x, p)$, on obtient

$$\mathbf{H}'(z) = \frac{\partial H}{\partial z}(z, -1, u(z)) + \frac{\partial H}{\partial u}(z, -1, u(z)) u'_s(z) = \frac{\partial H}{\partial z}(z, -1, u(z))$$

puisque $\partial_u H(z, -1, u(z)) = 0$. Ainsi, sur ce voisinage, le système

$$\dot{x}(t) = \frac{\partial H}{\partial p}[t], \quad \dot{p}(t) = -\frac{\partial H}{\partial x}[t], \quad 0 = \frac{\partial H}{\partial u}[t], \quad [t] := (x(t), p(t), -1, u(t)),$$

de la Proposition 3.2.1, est équivalent au **système hamiltonien**

$$\dot{z}(t) = \vec{\mathbf{H}}(z(t)), \quad \vec{\mathbf{H}}(z) = \left(\frac{\partial \mathbf{H}}{\partial p}(z), -\frac{\partial \mathbf{H}}{\partial x}(z) \right).$$

Définition 3.6.3

Une solution de l'équation linéarisée le long de z , appelée **équation de Jacobi**,

$$\widehat{\delta z}(t) = \vec{\mathbf{H}}'(z(t)) \cdot \delta z(t), \tag{3.8}$$

est appelée un **champ de Jacobi**. Un champ de Jacobi $\delta z = (\delta x, \delta p) \in \mathbb{R}^n \times (\mathbb{R}^n)^*$, est dit **vertical au temps** t si $\delta x(t) = 0$.

Nous pouvons maintenant donner une caractérisation géométrique du premier instant conjugué.

Proposition 3.6.4

Un instant $t_c \in]0, t_f]$ est un temps conjugué le long d'une extrémale normale vérifiant la condition forte de Legendre si et seulement si il existe un champ de Jacobi $\delta z = (\delta x, \delta p)$ vertical en 0 et t_c , tel que $\delta x \not\equiv 0$ sur $[0, t_c]$.

Définition 3.6.5 – Condition de Jacobi

Pour une extrémale définie sur l'intervalle $[a, b]$, on dit que la **condition faible de Jacobi** est vérifiée si l'intervalle ouvert $]a, b[$ ne contient pas de temps conjugué. On dit que la **condition forte de Jacobi** est vérifiée si l'intervalle semi-ouvert $]a, b]$ ne contient pas de temps conjugué.

Remarque 3.6.2. Dans la proposition précédente, c'est bien un δx et non $\delta \tilde{x}$ qui est considéré. En effet, le long d'une extrémale normale (vérifiant Legendre fort), on a $\delta \tilde{x} = (\delta x, \delta x^0)$ est vertical en t si et seulement si δx est vertical en t . Pour le voir, il suffit de remarquer que $\delta x(t) = E'(u) \cdot \delta u = x'(u) \cdot \delta u$ pour un certain δu donné, et où on a repris les notations de la remarque précédente. Ainsi, par la CN1 :

$$0 = \lambda^0 x^{0'}(u) \cdot \delta u + \lambda \cdot (x'(u) \cdot \delta u) =: \lambda^0 \delta x^0(t) + \lambda \cdot \delta x(t)$$

et donc, puisque $\lambda^0 \neq 0$, on a $\delta x(t) = 0$ si et seulement si $\delta \tilde{x}(t) = (\delta x(t), \delta x^0(t)) = 0$.

La question que l'on se pose maintenant est : est-ce que la trajectoire est nécessairement non optimale après le premier instant conjugué ? Pour ce faire, la forme quadratique Q_t associée à $B_t|_{K_t}$ doit être indéfinie pour $t > t_{1c}$. Cependant, il peut arriver dans des cas dégénérés, que Q_t reste semi-définie sur un intervalle $[t_{1c}, t_{1c} + \eta]$, $\eta > 0$. Dans le cadre analytique, cela ne peut pas arriver et nécessairement pour $t > t_{1c}$, Q_t est indéfinie. On obtient ainsi la condition nécessaire suivante, d'optimalité locale faible à l'ordre deux dans le cadre analytique de corang un.

Théorème 3.6.6 – [1]

Pour une extrémale normale vérifiant la condition forte de Legendre, dont le contrôle analytique associé est de corang un, la condition faible de Jacobi est une condition nécessaire d'optimalité locale pour la topologie L^∞ (sur u).

Remarque 3.6.3. Ce résultat est à comparer avec le Théorème B.6.1. On parle d'optimalité locale faible dans un sens analogue à la Définition B.2.2. Voir [1, 18] pour plus de détails sur ces notions d'optimalité locale faible et forte dans le cadre du contrôle optimal.

3.7 Exercices

Exercice 3.7.1 (solution p. 97) : (Problème de contrôle optimal à conditions aux limites simples.) Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(q, u) := \frac{1}{2} \int_0^{t_f} \|u(t)\|^2 dt, \\ \dot{q}(t) = u(t), \quad u(t) \in \mathbb{R}^2, \quad t \in [0, t_f] \text{ p.p.}, \\ q(0) = q_0, \quad q(t_f) = q_1, \end{cases}$$

où $\|\cdot\|$ est la norme euclidienne dans \mathbb{R}^2 , où $q := (x, y) \in \mathbb{R}^2$, et avec q_0 et q_1 dans \mathbb{R}^2 donnés, et $t_f > 0$ fixé.

1. Calculer la solution données par la Proposition 3.2.1.
2. Donner, pour cette solution, la valeur de $\int_0^{t_f} \|u(t)\| dt$.
3. Qu'avez vous (re)-démontré ?

Exercice 3.7.2 (solution p. 97) : (Problème de contrôle optimal à conditions finales sous-dimensionnées) Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(q, u) := \frac{1}{2} \int_0^{t_f} \|u(t)\|^2 dt, \\ \dot{q}(t) = u(t), \quad u(t) \in \mathbb{R}^2, \quad t \in [0, t_f] \text{ p.p.}, \\ q(0) = q_0, \quad q(t_f) \in q_1 + \mathbb{R}v, \end{cases}$$

où $\|\cdot\|$ est la norme euclidienne dans \mathbb{R}^2 , où $q := (x, y) \in \mathbb{R}^2$, et avec $q_0 := (0, 0)$, $q_1 := (1, 0)$, $v := (0, 1)$, et $t_f > 0$ fixé.

1. Donner le 5-uplet $(u(\cdot), q(\cdot), p(\cdot), p^0, \lambda)$ solution des conditions nécessaires issues de la Proposition 3.3.1.
2. Faire un dessin dans le plan de la trajectoire solution avec les contraintes terminales.

Exercice 3.7.3: (Problème de contrôle optimal à conditions finales sous-dimensionnées) Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(q, u) := \frac{1}{2} \int_0^{t_f} \|u(t)\|^2 dt, \\ \dot{q}(t) = u(t), \quad u(t) \in \mathbb{R}^2, \quad t \in [0, t_f] \text{ p.p.}, \\ q(0) = q_0, \quad q(t_f) \in q_1 + \mathbb{S}^1, \end{cases}$$

où $\|\cdot\|$ est la norme euclidienne dans \mathbb{R}^2 , $q := (x, y) \in \mathbb{R}^2$, et avec $q_0 := (0, 0)$, $q_1 :=$

$(2, 0)$, et $t_f > 0$ fixé. On rappelle que \mathbb{S}^1 est la sphère de centre $(0, 0)$ et de rayon 1.

1. Donner les deux 5-uplets $(u(\cdot), q(\cdot), p(\cdot), p^0, \lambda)$ solutions des conditions nécessaires issues de la Proposition 3.3.1.

2. Faire un dessin dans le plan des contraintes terminales, ainsi que de la meilleure des deux trajectoires.

Exercice 3.7.4: On note $c(t_f, x_f) = x_f - t_f^2$. Calculer la BC-extrémale du problème

$$\begin{cases} \min J(x, u, t_f) := \int_{-2}^{t_f} \sqrt{1 + u^2(t)} dt, & t_f > -2, \\ \dot{x}(t) = u(t), & u(t) \in \mathbb{R}, \quad t \in [-2, t_f] \text{ p.p.}, \\ x(-2) = 0, & c(t_f, x(t_f)) = 0. \end{cases}$$

Exercice 3.7.5: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u, t_f) := \int_0^{t_f} u^2(t) dt, \\ \dot{x}(t) = u(t), & u(t) \in \mathbb{R}, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = 0, & c(t_f, x(t_f)) = 0, \end{cases}$$

avec $c(t_f, x_f) = x_f - t_f - 10$ et où $t_f > 0$ est libre. Calculer la BC-extrémale.

Exercice 3.7.6:

Soit le problème de contrôle optimal suivant :

$$(P_5) \quad \begin{cases} \min J(x, u) := \int_0^1 \left(\frac{u^2(t)}{x^2(t)} - \ln(x^2(t)) \right) dt, \\ \dot{x}(t) = u(t), & u(t) \in \mathbb{R}, \quad t \in [0, 1] \text{ p.p.}, \\ x(0) = 1, & x(1) = e. \end{cases}$$

On fera l'hypothèse : $\forall t \in [0, 1], x(t) \neq 0$. On ne considère que le cas normal et on fixe la variable duale du coût : $p^0 = -\frac{1}{2}$. Soit $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u}(\cdot))$ une BC-extrémale du problème (P_5) .

1. Donner le pseudo-hamiltonien $H(x, p, u)$ du problème (P_5) , l'équation différentielle vérifiée par $(\bar{x}(\cdot), \bar{p}(\cdot))$ et le contrôle optimal $\bar{u}(\cdot)$.

2. Montrer que pour tout $t \in [0, 1]$ on peut écrire $\bar{p}(t) \bar{x}(t)$ sous la forme $a t + b$, avec a et b deux réels à déterminer ou donner en fonction de $\bar{p}(0)$.

3. Donner les expressions de $\bar{x}(\cdot)$ et $\bar{p}(\cdot)$ en fonction de t et $\bar{p}(0)$ et en déduire $\bar{p}(0)$.

Exercice 3.7.7: Soit le problème linéaire-quadratique suivant :

$$(LQR) \quad \begin{cases} \min J(x, u) := \frac{1}{2} \int_0^{t_f} (u(t)^T R u(t) + x(t)^T Q x(t)) dt, \\ \dot{x}(t) = A x(t) + B u(t), \quad u(t) \in \mathbb{R}^m, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \end{cases}$$

où $x(t) \in \mathbb{R}^n$, où $x_0 \in \mathbb{R}^n$ et $t_f > 0$ sont fixés, où A , B , R et Q sont des matrices constantes de dimensions adaptées, et où R est **symétrique définie positive** et Q est **symétrique semi-définie positive**.

1. Donner le pseudo-hamiltonien $H(x, p, p^0, u)$ associé au problème (LQR).
2. Donner la dynamique adjointe (issue de la CN1) vérifiée par p .
3. Donner les conditions de transversalité vérifiée en t_f .
4. Montrer que pour tout extrémale $(x(\cdot), p(\cdot), p^0, u(\cdot))$, $p^0 \neq 0$. On fixera ensuite $p^0 = -1$.
5. Montrer que le contrôle optimal vérifie $\bar{u}(x, p) = R^{-1} B^T p^T$ (p est un vecteur ligne).
6. Montrer que pour tout extrémale, le contrôle $t \mapsto u(t) = \bar{u}(x(t), p(t))$ est continu.

Soit le problème à valeur terminale (TVP) suivant :

$$\dot{S}(t) = S(t) B R^{-1} B^T S(t) - (S(t) A + A^T S(t) + Q), \quad S(t_f) = 0.$$

Nous faisons l'hypothèse suivante.

Hypothèse H4. Il existe une unique solution à (TVP) notée $t \mapsto S(t) \in \mathbb{R}^{n \times n}$, $S(t)$ **symétrique** et définie sur tout l'intervalle $[0, t_f]$.

7. Calculer

$$\frac{d}{dt} (x(t)^T S(t) x(t))$$

en fonction de $x(t)$, $S(t)$, $\dot{S}(t)$, A , B et $u(t)$. On pourra omettre l'argument t dans ce calcul et les calculs suivants.

8. Montrer que l'on peut écrire $J(x, u)$ sous la forme suivante (on omet t) :

$$J(u) = \frac{1}{2} \int_0^{t_f} (x^T (S B R^{-1} B^T S) x + u^T R u + 2 u^T B^T S x) dt + \frac{1}{2} x(0)^T S(0) x(0).$$

9. Montrer que

$$x^T (S B R^{-1} B^T S) x + u^T R u + 2 u^T B^T S x = (u + R^{-1} B^T S x)^T R (u + R^{-1} B^T S x).$$

10. En déduire l'expression du contrôle optimal $u^*(t, x)$ en fonction de R , B , $S(t)$ et x et en déduire la valeur minimale du critère.

Exercice 3.7.8: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u, t_f) := t_f + \frac{1}{2} \int_0^{t_f} (x^2(t) + u^2(t)) \, dt, \\ \dot{x}(t) = u(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = 0, \quad x(t_f) = x_f, \end{cases}$$

avec $x_f > 0$ donné. Calculer les BC-extrémales du problème, pour le cas normal : $p^0 = -1$.

Rappel : La solution de l'équation différentielle $\ddot{y}(t) - y(t) = 0$, s'écrit sous la forme

$$y(t) = \alpha \cosh(t) + \beta \sinh(t), \quad (\alpha, \beta) \in \mathbb{R}^2,$$

ou bien de manière équivalente

$$y(t) = \lambda e^t + \mu e^{-t}, \quad (\lambda, \mu) \in \mathbb{R}^2,$$

car

$$\cosh(t) = \frac{e^t + e^{-t}}{2} \quad \text{et} \quad \sinh(t) = \frac{e^t - e^{-t}}{2}.$$

Pour mémoire on a également

$$\operatorname{argsh}(z) = \ln(z + \sqrt{z^2 + 1}), \quad z \in \mathbb{R}.$$

Exercice 3.7.9: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u) := x(t_f), \\ \dot{x}(t) = -x(t)u(t) + \frac{1}{2}u^2(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \end{cases}$$

avec $x_0 > 0$ et $t_f > 0$ donnés. Calculer les BC-extrémales du problème. On montrera que $p(t) < 0$, $t \in [0, t_f]$, où p désigne l'état adjoint associé à x .

Exercice 3.7.10: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u, t_f) := \int_0^{t_f} x^2(t) u^2(t) \, dt, \quad t_f > 0, \\ \dot{x}(t) = u(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = \varepsilon, \quad x^2(t_f) - t_f^2 = 3 + \varepsilon^2, \end{cases}$$

où $\varepsilon > 0$ est fixé, et où t_f est libre.

1. écrire le pseudo-hamiltonien $H(x, p, p^0, u)$ du problème.
2. Montrer que $p^0 = 0 \Rightarrow p(\cdot) \equiv 0$. En déduire $p^0 \neq 0$. On fixe $p^0 = -1$.
3. Montrer que pour tout $t \in [0, t_f]$ tel que $x(t) \neq 0$, on a

$$\dot{x}(t) = \frac{c}{2x(t)}, \quad \dot{p}(t) = c \dot{x}(t),$$

où c est une constante à déterminer en fonction de $x(t)$ et $p(t)$.

4. Montrer que $c = 4t_f$.
5. Calculer $x^2(t)$ et $p^2(t)$, $t \in [0, t_f]$, en fonction de t , t_f , ε et $p(0)$.
6. Montrer que $t_f = 1$ et en déduire la valeur de $p^2(0)$.

Exercice 3.7.11: Soit le problème de contrôle optimal suivant :

$$\begin{cases} \min J(x, u) := \int_0^{t_f} (2a x(t) u(t) + a^2 u^2(t)) \, dt \\ \dot{x}(t) = \frac{1}{a} x(t) + u(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, t_f], \\ x(0) = 0, \quad x(t_f) = x_f, \end{cases}$$

où $a > 0$ est fixé, x_f est donné et où $t_f = \pi a/2$. Calculer les BC-extrémales du problème.

Rappel : La solution de l'équation différentielle : $\ddot{y}(t) + c y(t) = 0$, $c > 0$, s'écrit sous la forme : $y(t) = \lambda \cos(\beta t) + \mu \sin(\beta t)$, $(\lambda, \mu, \beta) \in \mathbb{R}^3$.

Exercice 3.7.12: On note $q := (x, y)$ l'état, $u := (u_1, u_2)$ le contrôle et on notera $p := (p_1, p_2)$ le vecteur adjoint. Soit le système contrôlé

$$\dot{q}(t) = (\dot{x}(t), \dot{y}(t)) = f(q(t), u(t)) := (u_1(t), 1 + u_2(t)). \quad (3.9)$$

Soit le pseudo-hamiltonien $H(q, p, u) := (p | f(q, u)) - \frac{1}{2} \|u\|^2$, où $\|\cdot\|$ est la norme euclidienne et $(\cdot | \cdot)$ est le produit scalaire usuel sur \mathbb{R}^2 .

1. Donner le système pseudo-hamiltonien

$$\dot{q}(t) = \nabla_p H(q(t), p(t), u(t)), \quad \dot{p}(t) = -\nabla_q H(q(t), p(t), u(t)).$$

2. Donner $u[p]$ vérifiant $\nabla_u H(q, p, u[p]) = 0$, pour $(q, p) \in \mathbb{R}^2 \times \mathbb{R}^2$.
3. Calculer $(q(t, \alpha, \beta), p(t, \alpha, \beta))$ la solution du système pseudo-hamiltonien bouclé

$$\dot{q}(t) = \nabla_p H(q(t), p(t), u[p(t)]), \quad \dot{p}(t) = -\nabla_q H(q(t), p(t), u[p(t)]),$$

ayant comme condition initiale $q(0, \alpha, \beta) = (0, 0)$ et $p(0, \alpha, \beta) = (\alpha, \beta)$.

4. Soit $t_f > 0$. Donner l'ensemble des paires (α, β) telles que $x(t_f, \alpha, \beta) = 1$.

5. Calculer $F(\alpha, \beta) := \frac{1}{2} \int_0^{t_f} \|u[p(t, \alpha, \beta)]\|^2 dt$.

6. Donner $(\tilde{\alpha}, \tilde{\beta}) := \arg \min \{F(\alpha, \beta) \mid (\alpha, \beta) \in \mathbb{R}^2 \text{ et } x(t_f, \alpha, \beta) = 1\}$.

7. Donner $q(t_f, \tilde{\alpha}, \tilde{\beta})$, $\dot{q}(t_f, \tilde{\alpha}, \tilde{\beta})$ et $p(t_f, \tilde{\alpha}, \tilde{\beta})$ en fonction de t_f .

8. Soit le problème de contrôle optimal

$$\begin{cases} \min J(q, u) := \frac{1}{2} \int_0^1 \|u(t)\|^2 dt, \\ \dot{q}(t) = f(q(t), u(t)), \quad u(t) \in \mathbb{R}^2, \quad t \in [0, 1] \text{ p.p.}, \\ q(0) = (0, 0), \quad x(1) = 1. \end{cases}$$

8.1. Soit $h(q, p, p^0, u) := (p \mid f(q, u)) + \frac{1}{2} p^0 \|u\|^2$ le pseudo-hamiltonien associé au problème de contrôle. Justifier que $p^0 = -1$.

8.2. Rappeler la condition de transversalité en $t_f = 1$ vérifiée par le vecteur adjoint. On notera λ le multiplicateur associé à la contrainte terminale. Que vaut λ ?

Principe du maximum de Pontryagin

4.1	Le principe du maximum	61
4.2	Exemples de maximisation de hamiltoniens	64
4.3	Exemples de conditions de transversalité	67
4.4	Cadre hamiltonien et origine symplectique	70
4.5	Problèmes affines en la commande	74
4.6	Exercices	76

Dans ce chapitre, nous énonçons le résultat fondateur du contrôle optimal, le principe du maximum de Pontryagin [64]. La principale différence avec le principe faible du Chapitre 3, dans sa formulation, est que les conditions de Euler-Hamilton et de Legendre-Clebsch, cf. Section 3.4, sont remplacées par une condition de maximisation globale du pseudo-hamiltonien. D'un point de vue pratique, ce résultat nous permet de traiter des problèmes avec des contraintes sur le contrôle et demande moins de régularité sur les données du problème. Les conditions nécessaires d'optimalité que fournit le principe du maximum doivent être vérifiées par toute solution locale en un sens plus fort que pour le principe faible. En effet, toute paire trajectoire-contrôle localement optimale vis-à-vis de trajectoires-contrôles admissibles proches en norme C^0 pour la trajectoire et L^∞ pour le contrôle vérifie les conditions nécessaires d'optimalité du principe faible mais pas nécessairement celles du principe du maximum. On parle de solution locale faible. Cependant, si on relâche la notion de solution locale, en remplaçant la norme L^∞ par la norme L^1 pour le contrôle, alors toute paire trajectoire-contrôle localement optimale vis-à-vis de trajectoires-contrôles admissibles proches en norme C^0 pour la trajectoire et L^1 pour le contrôle vérifie les conditions nécessaires d'optimalité du principe du maximum, et bien entendu aussi celle du principe faible. On parle de solution locale forte ou de solution locale de Pontryagin. Ces notions locales faible et forte sont analogues à celles présentées Section B.2. Une fois le principe énoncé, nous donnons quelques propriétés et applications du principe du maximum.

4.1 Le principe du maximum

Dans [56, Section 3.4.5], on retrouve quelques critiques justifiées de l'approche variationnelle utilisée pour le principe faible. Pour le principe faible, l'ensemble des contrôles admissibles est trop restrictif puisque l'on se limite à des contrôles intérieurs aux contraintes U , ou alors à des problèmes sans contraintes sur le contrôle en prenant $U = \Pi$ un ouvert de \mathbb{R}^m comme dans le Chapitre 3. De plus, dans ce chapitre, nous avons supposé la dynamique et l'intégrande de la partie Lagrange du coût de classe \mathcal{C}^1 , ce qui ne permet pas de traiter

des problèmes de type minimisation de la consommation pour lesquels $f^0(x, u) = \|u\|$. Enfin, les perturbations du contrôle que l'on considère pour le principe faible sont de nature L^∞ , ce que l'on voit bien dans la preuve de la Proposition 3.1.1 où $\delta u \in L^\infty$. Ce procédé restreint l'optimalité si elle est locale à des petits voisinages de la trajectoire optimale, en comparaison avec le principe du maximum qui utilise des perturbations particulières de nature L^1 , appelées perturbations en aiguille. Autrement dit, en appliquant le principe du maximum on peut espérer obtenir moins de candidats à l'optimalité que pour le principe faible, car il peut exister des solutions locales faibles non fortes, cf. Exemple B.2.1, et ainsi découvrir plus facilement l'optimum global s'il existe.

On considère maintenant le problème de contrôle optimal dans sa formulation la plus générale, c'est-à-dire le Problème (OCP) présenté Section 2.2. Rappelons les hypothèses sur des données du problème.

Hypothèse H5. Soit un système contrôlé non autonome $\dot{x}(t) = f(t, x(t), u(t))$ où f est une application lisse de $\mathcal{I} \times \Omega \times \Pi$ dans \mathbb{R}^n , \mathcal{I} un intervalle ouvert de \mathbb{R} , Ω un ouvert de \mathbb{R}^n et Π un ouvert de \mathbb{R}^m . Considérons de plus une application f^0 lisse sur $\mathcal{I} \times \Omega \times \Pi$ et à valeurs dans \mathbb{R} et deux applications g et c lisses de $\mathcal{I} \times \Omega \times \mathcal{I} \times \Omega$ dans, respectivement, \mathbb{R} et \mathbb{R}^p , $p \leq 2(n+1)$. On notera (t_0, x_0, t_f, x_f) l'argument des applications g et c . On suppose enfin que c est une submersion sur $c^{-1}(\{0_{\mathbb{R}^p}\})$, c'est-à-dire $c'(t_0, x_0, t_f, x_f)$ est surjective pour tout point (t_0, x_0, t_f, x_f) vérifiant $c(t_0, x_0, t_f, x_f) = 0$. Enfin, on introduit l'ensemble $U \subset \Pi$ **quelconque**.

Remarque 4.1.1. Le qualificatif lisse veut dire suffisamment régulier. On peut supposer que les applications sont au moins de classe \mathcal{C}^1 . Cependant, dans le cadre du principe du maximum de Pontryagin, nous pouvons affaiblir cette hypothèse et supposer :

$$f(t, \cdot, u), f^0(t, \cdot, u) \text{ de classe } \mathcal{C}^1 \text{ pour tout } t \in \mathcal{I}, u \in U \text{ fixés}$$

et

$$f, f^0, \frac{\partial f}{\partial x}, \frac{\partial f^0}{\partial x} \text{ continues sur } \mathcal{I} \times \Omega \times U.$$

Cela permet de considérer des problèmes de type minimisation de la consommation pour lesquels $f^0(x, u) = \|u\|$. Il n'y a pas de problème d'existence de solutions maximales pour le système contrôlé dans ce cas, voir [67, Lemme 2.6.2].

Rappelons maintenant le problème de contrôle optimal dans sa formulation la plus générale, c'est-à-dire sous la forme de Bolza avec les temps initial et final libres :

$$(OCP) \quad \begin{cases} \min J(x, u, t_0, t_f) := g(t_0, x(t_0), t_f, x(t_f)) + \int_{t_0}^{t_f} f^0(t, x(t), u(t)) \, dt \\ \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U, \quad t \in [t_0, t_f] \text{ p.p.}, \\ c(t_0, x(t_0), t_f, x(t_f)) = 0_{\mathbb{R}^p}. \end{cases}$$

Nous avons alors les conditions nécessaires d'optimalité suivantes qui généralisent les propositions 3.1.1, 3.2.1 et 3.3.1. Le théorème suivant, le Principe du Maximum de Pontryagin, est fondamental et nous renvoyons aux références [1, 19, 26, 27, 47, 50, 54, 56, 64, 65, 70, 72, 74, 76] pour plus de détails sur ce principe et la théorie du contrôle optimal en général.

Théorème 4.1.1 – Principe du Maximum de Pontryagin (PMP)

Si (x, u, t_0, t_f) est solution du Problème (OCP) sous les Hypothèses (H5), alors il existe un vecteur adjoint $p(\cdot) \in AC([t_0, t_f], (\mathbb{R}^n)^*)$ un réel $p^0 \in \{-1, 0\}$, et une forme linéaire $\lambda \in (\mathbb{R}^p)^*$, t.q. $(p(\cdot), p^0) \neq (0, 0)$ et tels que les équations suivantes sont vérifiées pour $t \in [t_0, t_f]$ p.p. :

$$\begin{aligned}\dot{x}(t) &= \frac{\partial H}{\partial p}(t, x(t), p(t), p^0, u(t)), \\ \dot{p}(t) &= -\frac{\partial H}{\partial x}(t, x(t), p(t), p^0, u(t)),\end{aligned}\tag{4.1}$$

$$H(t, x(t), p(t), p^0, u(t)) = \max_{w \in U} H(t, x(t), p(t), p^0, w),$$

où $H(t, x, p, p^0, u) := p \cdot f(t, x, u) + p^0 f^0(t, x, u)$ est le pseudo-hamiltonien associé à (OCP). Les conditions aux limites $c(t_0, x(t_0), t_f, x(t_f)) = 0$ sont vérifiées. Soit

$$\xi(t_0, x_0, t_f, x_f) := p^0 g(t_0, x_0, t_f, x_f) + \sum_{i=1}^p \lambda_i c_i(t_0, x_0, t_f, x_f).$$

On a les conditions de transversalité suivantes :

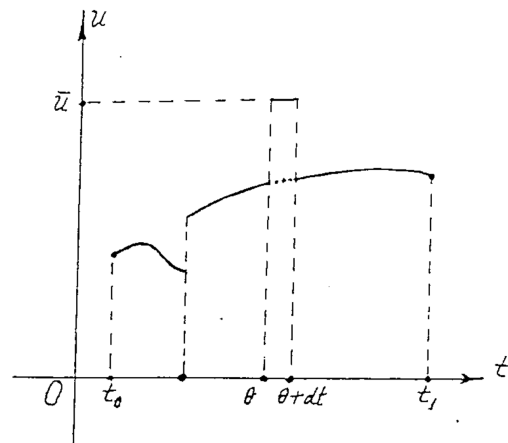
$$p(t_0) = -\frac{\partial \xi}{\partial x_0}(t_0, x(t_0), t_f, x(t_f)), \quad p(t_f) = \frac{\partial \xi}{\partial x_f}(t_0, x(t_0), t_f, x(t_f)).\tag{4.2}$$

Si $u(\cdot)$ est continu aux temps t_0 , respectivement t_f , alors on a :

$$H[t_0] = \frac{\partial \xi}{\partial t_0}(t_0, x(t_0), t_f, x(t_f)), \quad H[t_f] = -\frac{\partial \xi}{\partial t_f}(t_0, x(t_0), t_f, x(t_f)),\tag{4.3}$$

où $[t] := (t, x(t), p(t), p^0, u(t))$.

► La preuve s'appuie sur la construction de variations en aiguille introduites par V. G. Boltyanski [15]. Comme nous pouvons le voir sur la figure ci-contre, extraite de [15], les variations s'appliquent sur le contrôle. Elles ne sont pas petites en norme L^∞ mais en norme L^1 . La perturbation vient de la constante \bar{u} qui remplace le contrôle de référence sur un petit intervalle de temps de longueur dt . ■



Remarque 4.1.2. Pour des raisons similaires à celles évoquées dans la remarque 1.2.3, le vecteur adjoint $p(\cdot)$ est Lipschitz sous nos hypothèses.

Remarque 4.1.3. On rappelle que l'on peut considérer p et λ comme des vecteurs colonnes. Dans ce cas, nous écrivons

$$\dot{x}(t) = \nabla_p H(t, x(t), p(t), p^0, u(t)), \quad \dot{p}(t) = -\nabla_x H(t, x(t), p(t), p^0, u(t)),$$

les conditions de transversalité deviennent

$$p(t_0) = -\nabla_{x_0} \xi(t_0, x(t_0), t_f, x(t_f)), \quad p(t_f) = \nabla_{x_f} \xi(t_0, x(t_0), t_f, x(t_f))$$

et les conditions sur le hamiltonien s'écrivent

$$H[t_0] = \nabla_{t_0} \xi(t_0, x(t_0), t_f, x(t_f)), \quad H[t_f] = -\nabla_{t_f} \xi(t_0, x(t_0), t_f, x(t_f)).$$

La définition suivante fait écho à la Définition 3.3.2.

Définition 4.1.2 – Définitions importantes

- Une **extrémale de Pontryagin-Boltyanski** associée au Problème (OCP) est un quadruplet $(x(\cdot), p(\cdot), p^0, u(\cdot))$ solution des **équations pseudo-hamiltoniennes contraintes** (4.1) sur $[t_0, t_f]$.
- On parle de **BC-extrémale** (BC vient de “Boundary Conditions”) si l'extrémale vérifie en plus les **conditions aux limites** $c(t_0, x(t_0), t_f, x(t_f)) = 0$, les **conditions de transversalité** (4.2) et les **conditions sur le hamiltonien** (4.3).
- Une extrémale $(x(\cdot), p(\cdot), p^0, u(\cdot))$ est dite **anormale** si $p^0 = 0$ et **normale** dans le cas contraire, *i.e.* $p^0 = -1$.

Proposition 4.1.3

Soit $(\bar{x}(\cdot), \bar{p}(\cdot), p^0, \bar{u}(\cdot))$ une extrémale de Pontryagin-Boltyanski du Problème (OCP). Alors, l'application

$$t \mapsto H(t, \bar{x}(t), \bar{p}(t), p^0, \bar{u}(t))$$

est Lipschitz sur $[t_0, t_f]$. De plus, pour $t \in [t_0, t_f]$ p.p. on a :

$$\frac{d}{dt} H(t, \bar{x}(t), \bar{p}(t), p^0, \bar{u}(t)) = \frac{\partial}{\partial t} H(t, \bar{x}(t), \bar{p}(t), p^0, \bar{u}(t)),$$

si $\partial_t H(t, \bar{x}(t), \bar{p}(t), p^0, \bar{u}(t))$ existe sur presque tout $[t_0, t_f]$.

Corollaire 4.1.4

Le long d'une extrémale de Pontryagin-Boltyanski $(\bar{x}(\cdot), \bar{p}(\cdot), p^0, \bar{u}(\cdot))$ solution d'un problème **autonome**, le pseudo-hamiltonien $t \mapsto H(\bar{x}(t), \bar{p}(t), p^0, \bar{u}(t))$ est **constant**.

4.2 Exemples de maximisation de hamiltoniens

Dans chacun des exemples suivants, on cherche à maximiser sur U l'application partielle $u \mapsto M(u) := H(\bar{t}, \bar{x}, \bar{p}, p^0, u)$ où \bar{t} , \bar{x} , \bar{p} et p^0 sont fixés. Le pseudo-hamiltonien est supposé

suffisamment régulier et on ne note pas p^0 qui est fixé à -1 . On n'oublie pas que d'après le PMP, cette maximisation se fait à chaque instant t presque partout dans $[t_0, t_f]$.

Exemple 4.2.1. Si U est un ensemble **ouvert** (pas de contraintes sur le contrôle) et si $M(u)$ est strictement concave sur U alors M possède au plus un unique maximum sur U noté \bar{u} donné par la CN1 :

$$\bar{u} = \arg \max_{u \in U} M(u) \iff M'(\bar{u}) = \nabla_u H(\bar{t}, \bar{x}, \bar{p}, \bar{u}) = 0.$$

Remarque 4.2.1. Si de plus, $\nabla_{uu}^2 H(\bar{t}, \bar{x}, \bar{p}, \bar{u})$ est définie négative, alors d'après le théorème des fonctions implicites, il existe une fonction implicite $\varphi(t, x, p)$ telle que $\bar{u} = \varphi(\bar{t}, \bar{x}, \bar{p})$ et telle que localement $\nabla_u H(t, x, p, \varphi(t, x, p)) = 0$. En définissant localement $\mathbf{H}(t, z) := H(t, z, \varphi(t, z))$, $z := (x, p)$, on obtient :

$$\frac{\partial \mathbf{H}}{\partial z}(t, z) = \frac{\partial H}{\partial z}(t, z, \varphi(t, z)) + \frac{\partial H}{\partial u}(t, z, \varphi(t, z)) \frac{\partial \varphi}{\partial z}(t, z) = \frac{\partial H}{\partial z}(t, z, \varphi(t, z))$$

puisque $\partial_u H(t, z, \varphi(t, z)) = 0$. Par la suite, on notera u la fonction φ .

Voici deux exemples.

- Considérons un pseudo-hamiltonien de la forme

$$H(x, p, u) := H_0(x, p) + up - 0.5u^2,$$

avec x, p et u dans \mathbb{R} , et où H_0 est une fonction lisse de (x, p) . On a alors $\nabla_u H(x, p, u) = 0$ ssi $p - u = 0$ donc le contrôle maximisant est de la forme

$$u(p) := p.$$

- Considérons un hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + \sum_{i=1}^m u_i H_i(x, p) - 0.5\|u\|^2,$$

avec $u \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ et où les H_0, H_1, \dots, H_m sont des fonctions lisses. Posons

$$\Phi := (H_1, \dots, H_m).$$

On a alors $\nabla_u H(x, p, u) = 0$ ssi $\Phi - u = 0$ donc le contrôle maximisant est de la forme

$$u(x, p) := \Phi(x, p).$$

□

Considérons le cas où U est donné par la **contrainte sur le contrôle** $\|u\| \leq 1$.

Exemple 4.2.2. Considérons un hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + u H_1(x, p),$$

avec $u \in \mathbb{R}$, $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ et où H_0 et H_1 sont des fonctions lisses. La contrainte sur le contrôle nous donne $u(t) \in [-1, 1]$ pour presque tout temps t . Si au temps t , $H_1(z(t)) \neq 0$, $z = (x, p)$, alors le contrôle optimal est de la forme

$$u(z(t)) := \text{sign}(H_1(z(t))),$$

i.e. $\bar{u} = -1$ si $H_1 < 0$ où $\bar{u} = +1$ si $H_1 > 0$. Le contrôle le long de ces extrémales sature la contrainte, on parle d'arc **bang**.

La question qui reste est : comment calculer le contrôle si $H_1(z(t)) = 0$ sur $I \subset [t_0, t_f]$ non réduit à un point ?

Si $H_1(z(t)) = 0$ sur I alors pour tout temps $t \in I$, toutes les dérivées de $t \mapsto H_1(z(t))$ sont égales à 0. En dérivant suffisamment de fois, nous allons faire apparaître le contrôle, ce qui nous donnera sa valeur. Ce contrôle sera celui qui permet à l'extrémale de rester sur l'hypersurface $H_1 = 0$, voir Exemples 1.7.3 et 1.7.2. □

Exemple 4.2.3. Considérons maintenant un hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + \sum_{i=1}^m u_i H_i(x, p),$$

avec $u \in \mathbb{R}^m$, $\|u\| \leq 1$, $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ et où H_0, H_1, \dots, H_m sont des fonctions lisses. Posons

$$\Phi := (H_1, \dots, H_m).$$

Ainsi $H(z, u) = H_0(z) + (u | \Phi(z))$, $z = (x, p)$. D'après l'inégalité de Cauchy-Schwartz,

$$(u | \Phi(z)) \leq \|u\| \|\Phi(z)\|$$

et si $\Phi(z) \neq 0$, ce maximum est atteint pour

$$u(z) := \frac{\Phi(z)}{\|\Phi(z)\|}.$$

Il reste à calculer $u(z)$ dans le cas où $\Phi \equiv 0$ sur un intervalle $I \subset [t_0, t_f]$ non réduit à un point. Dans ce cas, c'est un peu plus compliqué que dans le cas précédent mais l'idée est toujours de dériver tous les H_i le long de l'extrémale afin d'obtenir des équations faisant intervenir le contrôle. Voir par exemple [28, page 5] pour le cas général et [15] pour une application concrète. □

Exemple 4.2.4. Considérons dans cet exemple, un hamiltonien de la forme

$$H(x, p, u) = H_0(x, p) + \sum_{i=1}^m u_i H_i(x, p) - \|u\|,$$

avec $u \in \mathbb{R}^m$, $\|u\| \leq 1$, $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ et où H_0, H_1, \dots, H_m sont des fonctions lisses. Posons

$$\Phi := (H_1, \dots, H_m).$$

On a alors $H(z, u) = H_0(z) + (u | \Phi(z)) - \|u\|$. D'après l'inégalité de Cauchy-Schwartz,

$$(u | \Phi(z)) - \|u\| \leq \|u\| \|\Phi(z)\| - \|u\| = \|u\|(\|\Phi(z)\| - 1)$$

et si $\Phi(z) \neq 0$, ce maximum est atteint pour un contrôle de la forme

$$u(z) := \rho(z) \frac{\Phi(z)}{\|\Phi(z)\|},$$

$\rho(z) \in [0, 1]$. Si $\Phi(z) = 0$ alors $u(z) = 0$. On a donc

$$H(z, u(z)) = H_0(z) + \rho(z) \|\Phi(z)\| - \rho(z) = H_0(z) + \rho(z) (\|\Phi(z)\| - 1).$$

Ainsi,

$$\begin{cases} \rho(z) = 1 & \text{si } \|\Phi(z)\| > 1, \\ \rho(z) = 0 & \text{si } \|\Phi(z)\| < 1, \\ \rho(z) \in [0, 1] & \text{si } \|\Phi(z)\| = 1. \end{cases}$$

Il reste à traiter le cas $\|\Phi(z)\| \equiv 1$ sur un intervalle d'intérieur non vide. Posons

$$\tilde{H}_1 := \|\Phi\| - 1,$$

et considérons ρ comme nouveau contrôle. Le nouveau pseudo-hamiltonien s'écrit

$$H(z, \rho) = H_0(z) + \rho \tilde{H}_1(z).$$

Les extrémales sont caractérisées par $\nabla_\rho H = \|\Phi\| - 1 = \tilde{H}_1 = 0$, et on se retrouve dans le cas de l'exemple précédent. Il faudra donc dériver autant de fois que nécessaire pour retrouver le contrôle.

□

4.3 Exemples de conditions de transversalité

Pour se simplifier la vie, on écrit parfois le multiplicateur de Lagrange λ et le vecteur adjoint p sous la forme d'un vecteur colonne.

Exemple 4.3.1. Considérons un problème de Lagrange (*i.e.* $g \equiv 0$) où t_0 , t_f , x_0 et x_f sont fixés mais tout de même inclus dans les conditions aux limites :

$$c(t_0, x_0, t_f, x_f) = (t_0 - \bar{t}_0, x_0 - \bar{x}_0, t_f - \bar{t}_f, x_f - \bar{x}_f).$$

Les contraintes sont donc

$$c(t_0, x(t_0), t_f, x(t_f)) = 0_{\mathbb{R}^p}.$$

On a alors $p = 2n + 2$ contraintes et un multiplicateur

$$\lambda := (\lambda_{t_0}, \lambda_{x_0}, \lambda_{t_f}, \lambda_{x_f}) \in \mathbb{R}^p.$$

Les conditions de transversalité et sur le hamiltonien nous donne

$$p(t_0) = -\lambda_{x_0}, \quad p(t_f) = \lambda_{x_f}, \quad H[t_0] = \lambda_{t_0}, \quad H[t_f] = -\lambda_{t_f}.$$

Ainsi, ces conditions n'apportent aucune informations supplémentaires.

□

Exemple 4.3.2. Prenons un problème de Lagrange 2D avec des contraintes $c(x(t_f)) = 0$ sous la forme $c(x) := 0.5(x_1^2 + x_2^2) - 1$, $x := (x_1, x_2)$, et où t_0 , x_0 et t_f sont fixés. D'après le PMP, on a la condition de transversalité

$$p(t_f) = \lambda \nabla c(x(t_f)) = \lambda x(t_f).$$

On a donc la condition liant $x(\cdot)$ et $p(\cdot)$ en t_f : $p_1(t_f)x_2(t_f) - p_2(t_f)x_1(t_f) = 0$. Remarquons que l'on a $n = 3$ inconnues $p_1(0)$, $p_2(0)$ et λ et 3 équations :

$$0 = c(x(t_f)), \quad 0 = p_1(t_f) - \lambda x_1(t_f), \quad 0 = p_2(t_f) - \lambda x_2(t_f),$$

que l'on peut tout aussi bien écrire sous la forme :

$$\begin{aligned} 0 &= c(x(t_f)), \\ 0 &= p_1(t_f)x_2(t_f) - p_2(t_f)x_1(t_f), \\ 0 &= \lambda - \frac{(p(t_f) | x(t_f))}{\|x(t_f)\|^2} = \lambda - \frac{(p(t_f) | x(t_f))}{2}. \end{aligned}$$

On remarque alors que l'on peut se retreindre à chercher dans un premier temps $p_1(0)$, $p_2(0)$, puis on trouve λ dans un second temps. \square

Exemple 4.3.3. Dans le cas d'un problème de Lagrange où l'on a des conditions terminales simples :

$$c(x_f) := x_f - \bar{x}_f = 0,$$

alors la condition de transversalité

$$p(t_f) = \lambda$$

ne donne pas d'informations utiles puisque l'on a ici une équation qui mélange seulement le vecteur adjoint et λ , que l'on cherche à déterminer. \square

Exemple 4.3.4. Dans le cas contraire où l'on n'a pas de contraintes terminales, *i.e.* $c(x_f) := 0$, toujours pour un problème de Lagrange ($g \equiv 0$), alors on a la condition de transversalité $p(t_f) = 0$, qui détermine ici totalement le vecteur adjoint en l'instant final. \square

Exemple 4.3.5. Toujours pour un problème de Lagrange, si

$$c(x_0, x_f) := x_f - x_0 = 0$$

qui est une condition pour une obtenir une trajectoire périodique, alors on obtient la condition

$$p(t_0) = p(t_f),$$

puisque les conditions de transversalité sont ici

$$p(t_0) = \lambda, \quad p(t_f) = \lambda.$$

On remarque encore une fois que si on détermine $p(t_0)$ ou $p(t_f)$ alors on peut retrouver λ facilement. \square

Exemple 4.3.6 (Voir aussi Exemple 3.3.1). Considérons un problème de Lagrange avec une contrainte terminale de la forme $c(x_f) = 0_{\mathbb{R}^p}$, $p \leq n$, et c une submersion. Alors, la condition de transversalité

$$p(t_f) = J_c(x(t_f))^T \lambda = \sum_{i=1}^p \lambda_i \nabla c_i(x(t_f))$$

signifie que

$$p(t_f) \perp T_{x(t_f)}M,$$

c'est-à-dire que $p(t_f)$ est orthogonal à l'espace tangent à M au point $x(t_f)$ où la cible terminale $M := c^{-1}(\{0_{\mathbb{R}^p}\})$ est une sous-variété de \mathbb{R}^n de dimension $n - p$. \square

Considérons le problème de contrôle optimal général (OCP) et notons $M := c^{-1}(\{0_{\mathbb{R}^p}\})$. L'ensemble M est une sous-variété différentielle de \mathbb{R}^N de dimension $N - p$, avec $p \leq N = 2n + 2$.¹ Pour tout $y \in M$, l'espace tangent à M en y est donné par $T_y M = \text{Ker } J_c(y)$, avec $J_c(y) \simeq c'(y)$ la matrice jacobienne de c en y . Soit $(x(\cdot), p(\cdot), p^0, u(\cdot))$ une BC-extrémale du Problème (OCP) sur $[t_0, t_f]$. Notons

$$y := (t_0, x(t_0), t_f, x(t_f)) \in M.$$

Les équations (4.2) et (4.3) du PMP se réécrivent alors

$$\begin{bmatrix} H[t_0] \\ -p(t_0) \\ -H[t_f] \\ p(t_f) \end{bmatrix} - p^0 \nabla g(y) \perp T_y M.$$

En effet, introduisons le point $q := (H[t_0], -p(t_0), -H[t_f], p(t_f)) - p^0 \nabla g(y) \in \mathbb{R}^N$ et la matrice $A := J_c(y) \in \mathbb{R}^{p \times N}$. Alors, la condition précédente devient $q \perp \text{Ker } A = (\text{Im } A^T)^\perp$, i.e. $q \in \text{Im } A^T$. Ainsi, il existe $\lambda \in \mathbb{R}^p$ t.q.

$$q = A^T \lambda$$

ce qui correspond aux conditions de transversalité (4.2) et sur le hamiltonien (4.3) du PMP.

Voyons maintenant comment se débarrasser de λ dans la formulation du PMP. Puisque c est une submersion sur M , alors A est de rang plein p et donc $\text{Ker } A$ est de dimension $N - p$. Soit (v_1, \dots, v_{N-p}) une base de $\text{Ker } A$. On note

$$B := \begin{bmatrix} v_1^T \\ \vdots \\ v_{N-p}^T \end{bmatrix} \in \mathbb{R}^{(N-p) \times N}.$$

La condition devient alors

$$q \perp \text{Ker } A \Leftrightarrow Bq = 0_{\mathbb{R}^{N-p}}.$$

et cette équation ne contient pas λ .

Voyons pour finir comment retrouver λ connaissant q . Attention, dire que l'on connaît q c'est dire que l'on a résolu le PMP et que l'on a trouvé une BC-extrémale. En effet, pour trouver q , il faut connaître B , donc A , or A dépend de y et y dépend de la BC-extrémale. Puisque

$$\mathbb{R}^N = \text{Ker } A \oplus \text{Im } A^T = \text{Im } B^T \oplus \text{Im } A^T,$$

alors

$$C := \begin{bmatrix} B^T & A^T \end{bmatrix} \in \mathbb{R}^{N \times N}$$

est inversible. C'est la matrice de passage de la base canonique de \mathbb{R}^N à la base

$$\{v_1, \dots, v_{N-p}, \nabla c_1(y), \dots, \nabla c_p(y)\}.$$

Pour la solution $q = A^T \lambda = \sum_{i=1}^p \lambda_i \nabla c_i(y)$, on a alors

$$\begin{pmatrix} 0_{\mathbb{R}^{N-p}} \\ \lambda \end{pmatrix} = C^{-1} q.$$

On retrouve donc λ en résolvant le système linéaire $C\mu = q$.

1. Attention, p désigne le vecteur adjoint et la dimension des contraintes $c = 0$.

4.4 Cadre hamiltonien et origine symplectique

Pour alléger les notations nous introduisons les définitions suivantes. Soit $H(t, x, p, p^0, u)$ un pseudo-hamiltonien, on appelle **système pseudo-hamiltonien**, le système suivant :

$$\vec{H}(t, x, p, p^0, u) := \left(\frac{\partial H}{\partial p}(t, x, p, p^0, u), -\frac{\partial H}{\partial x}(t, x, p, p^0, u) \right).$$

De même, soit $\mathbf{H}(t, x, p)$ un véritable hamiltonien (non autonome). On rappelle que le système hamiltonien associé est donné par :

$$\vec{\mathbf{H}}(t, x, p) := \left(\frac{\partial \mathbf{H}}{\partial p}(t, x, p), -\frac{\partial \mathbf{H}}{\partial x}(t, x, p) \right).$$

Remarque 4.4.1. Nous avons déjà introduit ces définitions dans la remarque 1.6.3 et dans la Définition 1.7.1.

Proposition 4.4.1

Soit $(\bar{x}(\cdot), \bar{p}(\cdot), p^0, \bar{u}(\cdot))$ une extrémale de Pontryagin-Boltyanski du Problème (OCP). On note $\bar{z}(\cdot) := (\bar{x}(\cdot), \bar{p}(\cdot))$.

Si pour presque tout $t \in [0, t_f]$, au voisinage de $\bar{z}(t)$, le pseudo-hamiltonien maximisé définit un hamiltonien

$$z \mapsto \mathbf{H}(t, z) := \max_{u \in U} H(t, z, p^0, u)$$

lisse, alors pour presque tout $t \in [0, t_f]$, on a $\dot{\bar{z}}(t) = \vec{\mathbf{H}}(t, \bar{z}(t))$.

► Puisque pour presque tout $t \in [0, t_f]$, d'après le PMP,

$$\dot{\bar{z}}(t) = \vec{H}(t, \bar{z}(t), p^0, \bar{u}(t)),$$

il suffit de montrer que, pour $t \in [0, t_f]$ p.p. : $\partial_z \mathbf{H}(t, \bar{z}(t)) = \partial_z H(t, \bar{z}(t), p^0, \bar{u}(t))$.

Soit $t \in [0, t_f]$ pour lequel (4.1) est vérifiée et pour lequel \mathbf{H} est défini et lisse, sur un voisinage ouvert de $\bar{z}(t)$, sur lequel on pose $F(z) := \mathbf{H}(t, z) - H(t, z, p^0, \bar{u}(t))$.

On a alors $F(z) \geq 0$ et $F(\bar{z}(t)) = 0$. Donc F est minimisée sur ce voisinage ouvert en $z = \bar{z}(t)$, ce qui implique $F'(\bar{z}(t)) = 0$. ■

Définition 4.4.2 – Extrémale de Hamilton

Une extrémale de Pontryagin-Boltyanski vérifiant les hypothèses de la Proposition 4.4.1 est appelée **extrémale de Hamilton**.

Supposons le pseudo-hamiltonien suffisamment régulier et p^0 fixé. Supposons que U soit un ouvert de \mathbb{R}^m et que $u \mapsto H(t, z, u)$, $z = (x, p)$, admette un unique maximum sur U en $u = u_m(t, z)$, pour tout (t, z) . On suppose de plus que u_m est lisse. Le pseudo-hamiltonien maximisé nous fournit un hamiltonien lisse (en z) donné par :

$$\mathbf{H}(t, z) = H(t, z, u_m(t, z)).$$

D'après le PMP, le long d'une extrémale de Pontryagin-Boltyanski est vérifiée

$$\dot{z}(t) = \vec{H}(t, z(t), u_m(t, z(t))).$$

D'après la définition et la proposition précédentes, cette extrémale est une extrémale de Hamilton qui vérifie donc aussi

$$\dot{z}(t) = \vec{\mathbf{H}}(t, z(t)),$$

ce qui se vérifie ici aisément dans ce cadre puisque

$$\frac{\partial \mathbf{H}}{\partial z}(t, z(t)) = \frac{\partial H}{\partial z}[t] + \frac{\partial H}{\partial u}[t] \cdot \frac{\partial u_m}{\partial z}(t, z(t)) = \frac{\partial H}{\partial z}[t]$$

car $\partial_u H[t] = 0$, où $[t] := (t, z(t), p^0, u_m(t, z(t)))$. Retrouvons pour finir les résultats de la Proposition 4.1.3 et du Corollaire 4.1.4 dans ce cadre. Notons $\mathbf{H}_t(z) := \mathbf{H}(t, z)$ pour avoir une fonction seulement de la variable z . Alors,

$$\begin{aligned} \frac{d}{dt} \mathbf{H}(t, z(t)) &= \frac{\partial \mathbf{H}}{\partial t}(t, z(t)) + \frac{\partial \mathbf{H}}{\partial z}(t, z(t)) \cdot \dot{z}(t) \\ &= \frac{\partial \mathbf{H}}{\partial t}(t, z(t)) + \mathbf{H}'_t(z(t)) \cdot \vec{\mathbf{H}}_t(z(t)) = \frac{\partial \mathbf{H}}{\partial t}(t, z(t)) + \{\mathbf{H}_t, \mathbf{H}_t\}(z(t)) \\ &= \frac{\partial \mathbf{H}}{\partial t}(t, z(t)) = \frac{\partial H}{\partial t}[t] + \frac{\partial H}{\partial u}[t] \cdot \frac{\partial u_m}{\partial t}(t, z(t)) = \frac{\partial H}{\partial t}[t] \end{aligned}$$

par antisymétrie du crochet de Poisson et puisque $\partial_u H[t] = 0$. Le hamiltonien est donc constant le long des extrémales si le problème est autonome.

Exemple 4.4.1. Supposons $U = \mathbb{R}^m$. On suppose aussi que le pseudo-hamiltonien est suffisamment régulier. On fixe t, z et p^0 . Si la fonction $u \mapsto H(t, z, p^0, u)$ est strictement concave et si la fonction est (négativement) coercive,² alors il existe un unique maximum à cette fonction. Si de plus

$$u \mapsto \frac{\partial^2 H}{\partial u^2}(t, z, p^0, u)$$

est définie négative en l'unique contrôle maximisant noté $u_{\max}(t, z, p^0)$, alors d'après le théorème des fonctions implicites, pour t et p^0 fixés, l'application partielle $z \mapsto u_{\max}(t, z, p^0)$ est lisse et donc les hypothèses de la Proposition 4.4.1 sont vérifiées le long de toute extrémale de Pontryagin-Boltyanski, c'est-à-dire le pseudo-hamiltonien maximisé défini bien un hamiltonien lisse. Dans ce cas, nous avons de plus que la condition forte de Legendre est vérifiée le long de toute extrémale. Tout ceci est vrai par exemple pour le pseudo-hamiltonien ($p^0 = -1$) :

$$H(x, p, u) = H_0(x, p) + \sum_{i=1}^m u_i H_i(x, p) - \frac{1}{2} \|u\|^2,$$

cf. Exemple 4.2.1. □

Pour compléter le Théorème 3.6.6, nous avons dans ce cadre, une condition suffisante d'optimalité locale forte à comparer avec le Théorème B.6.1. Le résultat à venir ne concerne

2. Voir <https://tinyurl.com/coercivite-wikipedia>.

que les problèmes autonomes (P_L) de Lagrange avec conditions aux limites simples. Rappelons la formulation du problème :

$$(P_L) \quad \begin{cases} \min J(x, u) := \int_0^{t_f} f^0(x(t), u(t)) dt \\ \dot{x}(t) = f(x(t), u(t)), \quad u(t) \in U, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

Nous avons alors le résultat suivant.

Théorème 4.4.3 – [1, Chapitre 21]

Pour une BC-extrémale normale de Hamilton du Problème (P_L) sous les Hypothèses (H5), la condition forte de Jacobi est une condition suffisante d'optimalité locale stricte pour la topologie \mathcal{C}^0 (sur x).

► Une preuve est donnée dans [23, Annexe A]. ■

En résumé, en combinant les Théorèmes 3.6.6 et 4.4.3, avant le premier temps conjugué, l'optimalité locale est vérifiée pour un grand voisinage dans \mathcal{C}^0 de la trajectoire, tandis qu'après le premier temps conjugué, l'optimalité locale est perdue, même pour des petits voisinages dans L^∞ sur le contrôle. Voir [28, Section 1.2] pour des algorithmes de calculs.

Origine symplectique. Les extrémales de Hamilton nous ramène aux théories de la dynamique hamiltonienne et de la géométrie symplectique [1, 6, 14, 55]. Rapidement dit, il y a deux objets centraux dans ce cadre qui sont fondamentaux ici, le hamiltonien \mathbf{H} et la forme symplectique ω . Le hamiltonien est une fonction sur l'espace des phases, tandis que la forme symplectique est une 2-forme différentielle sur l'espace des phases. C'est elle qui permet de définir la notion de champ de vecteurs hamiltonien donnant les équations de Hamilton. Ce hamiltonien avec la forme symplectique nous indique donc comment les extrémales de Hamilton se comportent dans l'espace des phases, qui lui-même définit donc la structure géométrique (symplectique) dans laquelle évolue les solutions du problème de contrôle optimal.

Nous avons déjà vu d'où vient le hamiltonien. Voyons comment obtenir la forme symplectique. Tout d'abord, à partir de l'espace d'état Ω , une variété différentielle, nous définissons l'espace des phases, ou fibré cotangent, comme l'espace

$$T^*\Omega := \{(x, p) \in \Omega \times T_x^*\Omega\},$$

où $T_x^*\Omega$ est l'espace dual de $T_x\Omega$, c'est-dire l'ensemble des formes linéaires sur l'espace tangent à Ω en x . Dans notre cas Ω est un ouvert de \mathbb{R}^n , donc

$$T_x\Omega \simeq \mathbb{R}^n, \quad T_x^*\Omega \simeq \mathbb{R}^n \quad \text{et} \quad T^*\Omega \simeq \Omega \times \mathbb{R}^n.$$

Nous pouvons de manière canonique définir une 1-forme différentielle sur $T^*\Omega$ donnée par

$$\alpha_{(x,p)} := \sum_{i=1}^n p_i dx_i.$$

Cette 1-forme est appelée la **forme de Liouville**. La structure symplectique de $T^*\Omega$ est donnée par la dérivée extérieure de la forme de Liouville, c'est-à-dire la 2-forme différentielle ω définie par

$$\omega_{(x,p)} := -d\alpha_{(x,p)} = \sum_{i=1}^n dx_i \wedge dp_i.$$

La forme symplectique est donc une forme différentielle non dégénérée. La structure symplectique de $T^*\Omega$ nous permet de définir la notion de champ de vecteurs hamiltonien. Soit H une fonction sur $T^*\Omega$. Le champ de vecteurs hamiltonien associé à H est le champ de vecteurs \vec{H} défini par

$$\omega(\vec{H}, \cdot) = dH.$$

Enfin, la structure symplectique nous permet de définir la notion de flot hamiltonien. Soit \vec{H} le champ de vecteurs hamiltonien associé à H . Le flot hamiltonien est la famille de difféomorphismes ϕ_t définie par

$$\frac{d}{dt}\phi_t(z_0) = \vec{H}(\phi_t(z_0)), \quad \phi_0(z_0) = z_0.$$

Le flot hamiltonien est donc la solution du problème de Cauchy

$$\dot{z}(t) = \vec{H}(z(t)), \quad z(0) = z_0.$$

Remarque 4.4.2. L'un des théorèmes centraux de la géométrie symplectique est le théorème de Darboux qui nous dit que, localement, toute forme symplectique est équivalente à la forme symplectique canonique ω_0 donnée par

$$\omega_0 = \sum_{i=1}^n dx_i \wedge dp_i.$$

Cela signifie que, localement, nous pouvons toujours trouver un changement de coordonnées qui nous permet de passer de ω à ω_0 . Un changement de coordonnées sur une structure symplectique est appelé un **symplectomorphisme**, c'est un difféomorphisme qui préserve la 2-forme symplectique. Dans notre cadre, à partir d'un changement de coordonnées sur Ω , nous pouvons définir de manière canonique un symplectomorphisme sur $T^*\Omega$. En effet, si φ est un changement de coordonnées sur Ω , alors le changement de coordonnées canonique sur $T^*\Omega$ est donné par

$$\varphi^* : T^*\Omega \rightarrow T^*\Omega, \quad (x, p) \mapsto (\varphi(x), p \varphi'(x)^{-1}).$$

On peut vérifier que φ^* est un symplectomorphisme. On rappelle aussi que

$$\varphi'(x)^{-1} = (\varphi^{-1})'(\varphi(x)).$$

Enfin, un autre théorème fondamental de la dynamique hamiltonienne est le théorème d'Arnold-Liouville et la notion de variables d'action-angles définies sur le tore dit de Liouville. Ce théorème permet de classer les extrémales de Hamilton en fonction de leurs propriétés topologiques et ainsi d'étudier comment les extrémales décrivent le tore. Dans le cas particulier de la mécanique classique présentée en Section 3.5.3, la théorie de Morse [61] appliquée à la fonction hauteur donnée par l'énergie mécanique permet de retrouver la description du tore de Liouville. C'est le cas aussi en géométrie riemannienne

de révolution en dimension 2 par exemple. Le point clé étant la présence d'un potentiel non dégénéré. Il existe également le théorème de Noether, fondamental, qui fait le lien entre les symétries et les lois de conservation.

4.5 Problèmes affines en la commande

On appelle **système contrôlé affine** tout système de la forme

$$\dot{x}(t) = F_0(x(t)) + \sum_{i=1}^m u_i(t) F_i(x(t)).$$

Le champ de vecteurs F_0 est appelé la **dérive** et $D = \{F_1, \dots, F_m\}$ est appelée la **distribution contrôlée**. La condition de Legendre n'est pas significative dans le cas affine où $\partial_{uu}^2 H \equiv 0$. On introduit donc la définition suivante. On dit qu'une extrémale de Pontryagin-Boltyanski $(x(\cdot), p(\cdot), p^0, u(\cdot))$ définie sur $[t_0, t_f]$ vérifie la **condition dégénérée de Legendre** si pour presque tout $t \in [t_0, t_f]$:

$$\frac{\partial^2 H}{\partial u^2}(t, x(t), p(t), p^0, u(t)) = 0.$$

Dans ce cas, pour un système affine par exemple, on a alors la condition nécessaire d'optimalité suivante. Si le contrôle est intérieur à la contrainte, alors la **condition de Goh** [38] :

$$\left\{ \frac{\partial H_t}{\partial u_i}, \frac{\partial H_t}{\partial u_j} \right\} (z(t)) = 0, \quad i, j = 1, \dots, m, \quad t \in [t_0, t_f] \text{ p.p.}, \quad (4.4)$$

est nécessaire pour que l'extrémale soit optimale. On rappelle que $z = (x, p)$. Nous avons introduit la notation $H_t(z) := H(t, z, p^0, u(t))$ et

$$\frac{\partial H_t}{\partial u_i}(z) := \frac{\partial H}{\partial u_i}(t, z, p^0, u(t)), \quad i = 1, \dots, m.$$

En plus de la condition de Goh, dans le cas où l'extrémale vérifie la condition dégénérée de Legendre, sous certaines hypothèses supplémentaires, on a la condition de Legendre-Clebsch généralisée [1] à vérifier. Dans le cas particulier affine mono-entrée autonome :

$$H(x, p, u) = H_0(x, p) + u H_1(x, p), \quad u \in \mathbb{R}, \quad H_i(x, p) := p \cdot F_i(x), \quad i = 0, 1,$$

la condition de Legendre-Clebsch généralisée est donnée par :

$$H_{101}(z(t)) := \{H_1, \{H_0, H_1\}\}(z(t)) \geq 0, \quad t \in [t_0, t_f] \text{ p.p..}$$

Classification des extrémales : système affine mono-entrée. On considère un système affine mono-entrée de la forme $\dot{x} = F_0(x) + u F_1(x)$, où $u \in \mathbb{R}$ et $x \in \Omega$, avec Ω un ouvert de \mathbb{R}^n . Soient H_0, H_1 les relèvements hamiltoniens de F_0, F_1 . Soit $(x(\cdot), p(\cdot), p^0, u(\cdot))$ une extrémale de Pontryagin-Boltyanski sur $I \subset \mathbb{R}$, I un ouvert d'intérieur non vide. Le long de cette extrémale, on a $\partial_u H(z(t), u(t)) = H_1(z(t)) = 0$, $z = (x, p)$. Nous sommes dans le cadre du Chapitre 1 et nous appelons cette extrémale, une **extrémale singulière**. La valeur

du contrôle se retrouve en dérivant au moins deux fois par rapport au temps l'application H_1 le long de l'extrémale, voir l'Exemple 1.7.2 pour les notations et le détail des calculs :

$$\begin{aligned} \frac{d}{dt} H_1(z(t)) &= H_{01}(z(t)) = 0 \\ \frac{d}{dt} H_{01}(z(t)) &= H_{001}(z(t)) + u(t) H_{101}(z(t)) = 0 \end{aligned} \quad (4.5)$$

et (4.5) permet la définition du contrôle singulier sous la forme fermée

$$u_s(z) := -\frac{H_{001}(z)}{H_{101}(z)}$$

en dehors de $\Sigma_{101}^0 := \{z \in T^*\Omega \mid H_{101}(z) = 0\}$. Une extrémale singulière définie en dehors de Σ_{101}^0 est appelée une **extrémale singulière d'ordre minimal**. Notons

$$\Sigma_1^0 := \{z \in T^*\Omega \mid H_1(z) = 0\}, \quad \Sigma_{01}^0 := \{z \in T^*\Omega \mid H_{01}(z) = 0\},$$

et définissons $\Sigma_s := \Sigma_1^0 \cap \Sigma_{01}^0$. En injectant u_s dans le pseudo-hamiltonien, on définit le hamiltonien

$$\mathbf{H}_s(z) := H(z, u_s(z)).$$

On a alors la proposition suivante.

Proposition 4.5.1

Soit $\bar{z} \in \Sigma_s \setminus \Sigma_{101}^0$.

Alors, il passe exactement une extrémale singulière d'ordre minimal par \bar{z} (à p^0 près). Elle est contenue dans Σ_s et elle est solution du système hamiltonien $\dot{z}(t) = \overrightarrow{\mathbf{H}}_s(z(t))$, le contrôle étant donné par $u_s(z(t))$.

► Soit $z(\cdot)$ la courbe intégrale de $\overrightarrow{\mathbf{H}}_s$ passant par \bar{z} à $t = 0$. Soit $\varphi(t) := \xi(z(t))$, avec $\xi(z) := (H_1(z), H_{01}(z))$. Alors $\varphi(0) = 0$ et $\varphi'(t) = \xi'(z(t)) \overrightarrow{\mathbf{H}}_s(z(t))$.

* Montrer que Σ_s est invariant par le flot hamiltonien associé à \mathbf{H}_s revient à montrer que $\varphi(t) = 0$ pour tout $t \in \mathbb{R}$. Or,

$$\begin{aligned} \varphi'_1(t) &= \{\mathbf{H}_s, H_1\}(z(t)) = \{H_0 + u_s H_1, H_1\}(z(t)) \\ &= \left(\{H_0, H_1\} + u_s \{H_1, H_1\} + H_1 \{u_s, H_1\} \right)(z(t)) \\ &= \{H_0, H_1\}(z(t)) + H_1(z(t)) \{u_s, H_1\}(z(t)) \end{aligned}$$

et

$$\begin{aligned} \varphi'_2(t) &= \{\mathbf{H}_s, H_{01}\}(z(t)) = \{H_0 + u_s H_1, H_{01}\}(z(t)) \\ &= \left(\underbrace{\{H_0, H_{01}\} + u_s \{H_1, H_{01}\}}_{=0 \text{ par définition de } u_s} + H_1 \{u_s, H_{01}\} \right)(z(t)) \\ &= H_1(z(t)) \{u_s, H_{01}\}(z(t)). \end{aligned}$$

Ainsi, $\varphi'(t) = A(t)\varphi(t)$, avec

$$A(t) := \begin{pmatrix} \{u_s, H_1\}(z(t)) & 1 \\ \{u_s, H_{01}\}(z(t)) & 0 \end{pmatrix}$$

et $\varphi(0) = 0$. On a donc $\varphi(t) = 0$ pour tout $t \in \mathbb{R}$.

* Maintenant, puisque

$$\mathbf{H}_s'(z) = \frac{\partial H}{\partial z}(z, u_s(z)) + \frac{\partial H}{\partial u}(z, u_s(z)) u_s'(z) = \frac{\partial H}{\partial z}(z, u_s(z)) + H_1(z) u_s'(z),$$

on a $\mathbf{H}_s'(z(t)) = \partial_z H(z(t), u_s(z(t)))$ le long de la courbe intégrale $z(\cdot)$ de $\overrightarrow{\mathbf{H}_s}$. Ainsi, $(z(\cdot), u_s(z(\cdot)))$ est une extrémale singulière (d'ordre minimal). Réciproquement, soit une extrémale d'ordre minimal passant par \bar{z} . Alors, automatiquement, elle est contenue dans Σ_s , le contrôle est donnée par u_s et elle est donc aussi solution du système hamiltonien $\dot{z}(t) = \overrightarrow{\mathbf{H}_s}(z(t))$ puisque $\mathbf{H}_s'(z(t)) = \partial_z H(z(t), u_s(z(t)))$ dès que $H_1(z(t)) = 0$. ■

Remarque 4.5.1. D'après la proposition précédente, $\Sigma_s = \xi^{-1}(\{0_{\mathbb{R}^2}\})$ est invariant par le flot hamiltonien associé à \mathbf{H}_s . Supposons que ξ soit une submersion sur $\Sigma_s \neq \emptyset$. Alors, Σ_s est une sous-variété de $T^*\Omega$ de codimension 2, et les ensembles Σ_1^0 et Σ_{01}^0 sont transverses en tout point de Σ_s .

Supposons que le contrôle soit borné entre -1 et 1 . La classification des extrémales bangs (de contrôle de norme maximale) et singulières se fait en fonction de l'ordre du contact des extrémales bangs avec la variété de commutation Σ_1^0 et des signes des crochets de Poisson H_{001} et H_{101} aux points de contact. L'optimalité locale des extrémales dépend de cette classification et on peut donner des conditions nécessaires et suffisantes d'optimalité locale sous des conditions d'absence de points conjugués comme dans les Sections 3.6 et 4.4. Voir les références [19, 18, 20] pour plus de détails. Voir aussi [28, Section 1.2] pour des algorithmes de calculs.

4.6 Exercices

Exercice 4.6.1: Soit le problème de contrôle optimal (non autonome) suivant :

$$\begin{cases} \min J(x, u) := \frac{1}{2} \int_{t_0}^{t_f} (t u(t) + x(t))^2 dt, \\ \dot{x}(t) = u(t), \quad u(t) \in \mathbb{R}, \quad t \in [t_0, t_f] \text{ p.p.}, \\ x(t_f) = 0, \end{cases}$$

avec $0 < t_0 < t_f$. Calculer les BC-extrémales du problème.

Exercice 4.6.2: [Problème de contrôle optimal à temps final libre] Soit le problème

de contrôle optimal suivant :

$$\begin{cases} \min J(x, u, t_f) := t_f, \\ \dot{x}(t) = (x_2(t), u(t)), \quad |u(t)| \leq 1, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = (a, 0), \quad x(t_f) = (0, 0), \end{cases}$$

avec $a < 0$ donné. L'objectif est de donner le 5-uplet $(t_f, u(\cdot), x(\cdot), p(\cdot), p^0)$ solution des conditions nécessaires issues de la proposition 3.6.1.

1. Application directe du PMP.

1.1. Donner le pseudo-hamiltonien $H(x, p, u)$.

1.2. Montrer que pour toute extrémale, $p(t) = (\alpha, -\alpha t + \beta)$ où α et β sont deux constantes.

1.3. Donner le contrôle maximisant $u[p_2]$.

2. Analyse du problème, partie 1.

2.1. Montrer que $(\alpha, \beta) \neq (0, 0)$ et en déduire que le contrôle commute au plus 1 fois.

2.2. Montrer qu'il existe $t_1 \in]0, t_f[$ tel que $p_2(t_1) = 0$, c'est-à-dire que le contrôle commute exactement 1 fois (utiliser le fait que $x(t_f) = (0, 0)$ et $x_2(0) = 0$). En déduire que $\alpha\beta > 0$.

2.3. Montrer que si $\alpha, \beta < 0$ alors on ne peut avoir $x_1(t_f) = 0$. En déduire que $\alpha, \beta > 0$.

3. Analyse du problème, partie 2.

3.1. Sachant que $H = -p^0$ le long de toute extrémale, montrer que $\beta = 1$ et que $p^0 = -1$.

3.2. Montrer que $t_1 = 1/\alpha$.

3.3. Calculer $x(t)$ pour $t \in [0, t_1]$, puis pour $t \in [t_1, t_f]$, en fonction de a et t_1 .

3.4. En utilisant le fait que $x(t_f) = (0, 0)$, montrer que $t_1 = \sqrt{-a}$ et $t_f = 2t_1$.

4. Conclusion.

4.1. Montrer que sur $[0, t_1]$, on a $x_1(t) = \frac{1}{2}x_2^2(t) + a$ et sur $[t_1, t_f]$, on a $x_1(t) = -\frac{1}{2}x_2^2(t)$.

4.2. Faire un dessin de la trajectoire solution dans le plan (x_1, x_2) .

Exercice 4.6.3: [Problème de turnpike avec singulière] Soit le problème suivant :

$$\begin{cases} \min J(x, u) := \int_0^2 x^2(t) dt, \\ \dot{x}(t) = u(t), \quad |u(t)| \leq 1, \quad t \in [0, 2] \text{ p.p.}, \\ x(0) = 1, \quad x(2) = 1/2. \end{cases}$$

L'objectif est de donner le 4-uplet $(u(\cdot), x(\cdot), p(\cdot), p^0)$ solution du PMP, c'est-à-dire calculer la BC-extrémale. On supposera que $p^0 = -1$.

1. Application directe du PMP.

1.1. Donner le pseudo-hamiltonien $H(x, p, u)$.

1.2. Donner le contrôle maximisant $u[p]$.

1.3. Donner l'équation adjointe.

1.4. Donner le problème aux deux bouts à résoudre. Est-il lisse ?

2. Analyse du problème, partie 1. Soit une trajectoire $x(\cdot)$ qui vérifie $x(t) < 0$ sur $]t_1, t_2[$, $0 < t_1 < t_2 < 2$, et $x(t_1) = x(t_2) = 0$. Alors, il est facile de voir que

$$\tilde{x}(t) = \begin{cases} x(t) & \text{sur } [0, t_1] \cup [t_2, 2], \\ 0 & \text{sur } [t_1, t_2] \end{cases}$$

est meilleure que $x(\cdot)$. Ainsi, la trajectoire optimale vérifie $x(\cdot) \geq 0$ sur $[0, 2]$ et $\dot{p} \geq 0$. On peut de même montrer que nécessairement $p(0) < 0$, ce que l'on supposera par la suite.

2.1. Montrer qu'il existe $t_1 > 0$ tel que $p(t_1) = 0$.

2.2. Montrer que si $x(t_1) > 0$ alors p ne s'annule plus ensuite.

2.3. En déduire que $x(t_1) = 0$.

3. Analyse du problème, partie 2.

3.1. Soit $t_2 \in [t_1, 2]$ tel que $x(t) = 0$ pour tout $t \in [t_1, t_2]$. On sait que $t_2 < 2$ car $x(t) > 0$ au voisinage de $t = 2$. Montrer que p ne s'annule plus sur $]t_2, 2]$. On utilisera le fait que $u(t_2) = 1$ pour rejoindre le demi-plan ouvert $x > 0$.

3.2. Calculer le contrôle permettant d'avoir $x(t) = 0$ pour tout $t \in]t_1, t_2[$, en supposant $t_1 < t_2$. On notera u_0 ce contrôle.

3.3. On sait que la solution est constituée de trois arcs et vérifie

$$\begin{cases} u(t) = -1 & \text{sur } [0, t_1], \\ u(t) = u_0 & \text{sur }]t_1, t_2[, \\ u(t) = 1 & \text{sur } [t_2, 2]. \end{cases}$$

Calculer la trajectoire solution $x(\cdot)$ avec le covecteur $p(\cdot)$, et donner les temps t_1 et t_2 .

Méthode indirecte de tir

5.1	Introduction par l'exemple	80
5.2	La démarche générale : condition aux limites simples	81
5.3	Exemple avec conditions aux limites sous-dimensionnées	83
5.4	Tir multiple de stabilité	85
5.5	Tir simple avec détections des commutations	86
5.6	Difficultés numériques des méthodes de tir simple	89
5.7	Tir multiple de structure	91

En contrôle optimal, il existe deux grandes familles de méthodes locales, les méthodes dites directes et les méthodes indirectes. Ce ne sont pas les seules, puisqu'il existe aussi des méthodes de modèles prédictifs pour résoudre des systèmes en temps réel à partir d'observations de l'état, ou encore des méthodes globales basées sur les équations de Hamilton-Jacobi-Bellman par exemple. Les méthodes directes consistent à transformer le problème de contrôle optimal en un problème d'optimisation de grande taille en dimension finie. Voir [10, 11] pour une présentation générale et [13] pour une présentation du tir direct. Dans les méthodes directes, l'état et/ou le contrôle sont discrétisés, sur une grille de temps choisie en amont, et les valeurs de l'état et/ou le contrôle en les instants de la grille deviennent les inconnues du problème d'optimisation. Ceci implique qu'il n'est pas nécessaire de connaître *a priori* la structure du contrôle optimal contrairement aux méthodes indirectes, telles que les méthodes de tir que l'on présente par la suite. Pour plus de détails sur l'implémentation des méthodes de tir indirect voir [8, 69] par exemple. Nous devons de plus faire la distinction entre plusieurs cas, car suivant le type de structure rencontrée, la méthode de tir choisie est différente. Supposons tout d'abord que le contrôle maximisant soit lisse par rapport au temps et s'écrive sous une forme fermée fonction de l'état et du co-état. Dans ce cas, nous utiliserons la méthode indirecte de tir simple pour résoudre le problème de contrôle optimal. La méthode de tir multiple peut quant à elle être utilisées pour plusieurs raisons. Initialement, le tir multiple a été introduit dans l'idée de renforcer la stabilité et la convergence de la méthode de tir simple, en discrétisant l'intervalle de temps [13]. On discrétise pour intégrer sur des intervalles de temps plus courts et ainsi réduire les erreurs numériques et la sensibilité. La subdivision est donc connue alors que les points intermédiaires à chaque instant de la grille sont des inconnues de la méthode de tir multiple. Maintenant, si le domaine du contrôle est fermé, autrement dit s'il y a des contraintes sur le contrôle, il est alors possible que le contrôle optimal sature la contrainte sur certains intervalles de temps et soit intérieur à la contrainte sur d'autres. Dans ce cas, pour résoudre des problèmes de type bang-bang ou bang-singulier par exemple, il est nécessaire

d'appliquer une méthode de tir multiple pour calculer précisément les instants de commutations entre les différents arcs. Nous verrons enfin que dans le cas particulier de problèmes bang-bang, nous pouvons aussi utiliser une méthode de tir simple avec détections des commutations [58]. Dans ce cas, seul le point initial est une inconnue (comme pour le tir simple), cependant il faut prendre garde à bien calculer la jacobienne de la fonction de tir en prenant en compte les sauts aux instants de commutations sur les équations variationnelles. Cette méthode de tir simple avec détections des commutations devient très utile lorsqu'il y a un nombre important de commutations pour réduire le nombre d'inconnues de la fonction de tir.

5.1 Introduction par l'exemple

Considérons le problème suivant :

$$(P_6) \quad \begin{cases} \min J(x, u) := \frac{1}{2} \int_0^{t_f} u(t)^2 dt \\ \dot{x}(t) = -x(t) + u(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

avec $t_f = 1$, $x_0 = -1$, $x_f = 0$ et $\forall t \in [0, t_f]$, $x(t) \in \mathbb{R}$. Le pseudo-hamiltonien du problème est

$$H(x, p, u) := p(-x + u) + p^0 u^2/2,$$

où $p^0 = -1$ car nous sommes dans le cas normal. La condition de maximisation du PMP, cf. Théorème 4.1.1, donne la loi de contrôle que l'on note $u(p)$:

$$u(p) := -p/p^0 = p$$

puisque $\partial_{uu}^2 H = p^0 = -1 < 0$. En introduisant cette loi de contrôle dans le système pseudo-hamiltonien de (4.1) et en considérant les conditions aux limites, on obtient le *problème aux deux bouts* (BVP) :

$$(BVP_1) \quad \begin{cases} \dot{x}(t) = \nabla_p H[t] = -x(t) + u(p(t)) = -x(t) + p(t), \\ \dot{p}(t) = -\nabla_x H[t] = p(t), \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

où $[t] := (x(t), p(t), u(p(t)))$. On s'intéresse maintenant à la résolution de (BVP₁). On introduit pour cela la notation connue

$$\vec{H}(z, u) := (\nabla_p H(z, u), -\nabla_x H(z, u)), \quad z := (x, p).$$

On note ensuite $z(\cdot, x_0, p_0)$ la solution de $\dot{z}(t) = \vec{H}(z(t), u(p(t)))$, $z(0) = (x_0, p_0)$. On définit enfin la **fonction de tir** suivante :

$$\begin{aligned} S: \mathbb{R} &\longrightarrow \mathbb{R} \\ p_0 &\longmapsto S(p_0) := \pi(z(t_f, x_0, p_0)) - x_f, \end{aligned}$$

où $\pi(x, p) := x$. Au final, résoudre (BVP₁) revient à résoudre $S(p_0) = 0$. C'est ce que l'on appelle la *méthode de tir simple indirect*. Les origines des méthodes de tir pour la résolution de problèmes aux deux bouts peuvent se retrouver dans les commentaires pages 139–140 de [9].

Calculons à la main la solution de (BVP₁). Nous pouvons tout d'abord calculer par la méthode de la variation de la constante la solution du système hamiltonien, ce qui nous donne :

$$p(t, x_0, p_0) = e^t p_0, \quad x(t, x_0, p_0) = (0.5 p_0 (e^{2t} - 1) + x_0) e^{-t}.$$

Et puisque $x_0 = -1$, $t_f = 1$, et puisque la cible est $x(t_f, x_0, p_0) = x_f = 0$, on obtient comme vecteur adjoint initial solution

$$\bar{p}_0 = \frac{2(x_f e^{t_f} - x_0)}{e^{2t_f} - 1} = \frac{2}{e^2 - 1} \approx 0.313.$$

Remarque 5.1.1. La fonction de tir est ici

$$S(p_0) = x(t_f, x_0, p_0) - x_f = p_0 \operatorname{sh}(t_f) + x_0 e^{-t_f} - x_f.$$

Elle est linéaire en p_0 et dire que \bar{p}_0 est solution c'est dire que $S(\bar{p}_0) = 0$.

5.2 La démarche générale : condition aux limites simples

Considérons le Problème (P_L) sous les Hypothèses (H2) et supposons que pour toute extrémale de Pontryagin-Boltyanski ($z(\cdot), p^0, u(\cdot)$), on peut écrire $u(t) = u(z(t))$, avec $z \mapsto u(z)$ de classe au moins \mathcal{C}^1 . En injectant $u(z)$ dans \vec{H} , trouver une BC-extrémale revient à résoudre le problème aux deux bouts suivant :

$$(BVP_2) \quad \begin{cases} \dot{z}(t) &= \vec{H}(z(t), u(z(t))), \\ 0_{\mathbb{R}^{2n}} &= b(z(0), z(t_f)) := (x(0) - x_0, x(t_f) - x_f). \end{cases}$$

On peut alors écrire le problème aux deux bouts (BVP₂) sous la forme d'un système d'équations non linéaires :

$$\begin{aligned} S: \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ p_0 &\longmapsto S(p_0) := \pi(z(t_f, x_0, p_0)) - x_f, \end{aligned}$$

où $\pi(x, p) = x$ et où $z(\cdot, x_0, p_0)$ est la solution de $\dot{z}(t) = \vec{H}(z(t), u(z(t)))$, $z(0) = (x_0, p_0)$. Résoudre le problème aux deux bouts (BVP₂) revient à résoudre $S(p_0) = 0$. C'est ce que l'on appelle la *méthode de tir simple indirect*.

Remarque 5.2.1. Si \bar{p}_0 vérifie $S(\bar{p}_0) = 0$, alors la courbe intégrale $\bar{z}(\cdot) := z(\cdot, x_0, \bar{p}_0)$, avec le contrôle $\bar{u}(\cdot) := u(\bar{z}(\cdot))$, est une BC-extrémale du Problème (P_L), i.e. cette extrémale satisfait les conditions nécessaires d'optimalité données par le PMP.

Remarque 5.2.2. D'après la Proposition 4.4.1 sur le hamiltonien vrai, la courbe intégrale $z(\cdot, x_0, p_0)$ est aussi solution de $\dot{z}(t) = \vec{H}(z(t))$, $z(0) = (x_0, p_0)$, où $\mathbf{H}(z) := H(z, u(z))$ et $\vec{H} = (\nabla_p \mathbf{H}, -\nabla_x \mathbf{H})$.

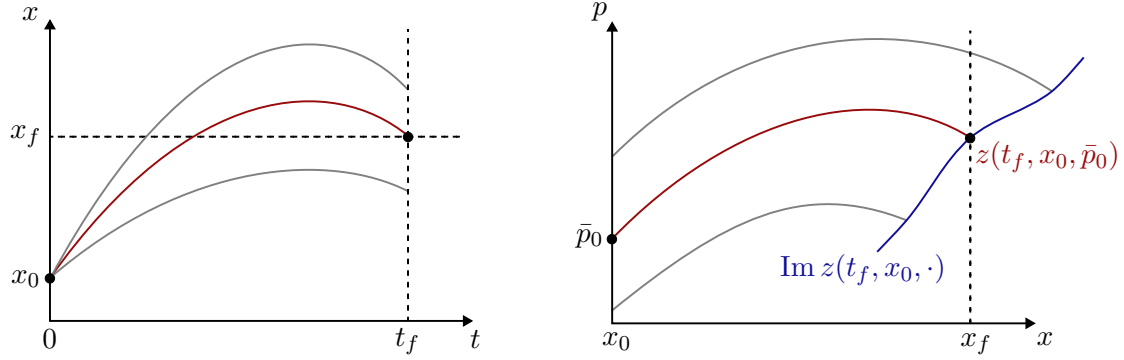


Illustration de la méthode de tir simple indirect.

Pour résoudre les équations de tir, nous avons besoin de calculer $z(t_f, x_0, p_0)$ qui est classiquement calculé numériquement par une méthode de type Runge-Kutta. Ensuite, pour résoudre le système d'équations $S(p_0) = 0$, nous pouvons utiliser une méthode de type Newton. Les méthodes de Newton sont connues pour être sensibles vis-à-vis de la condition initiale. Une difficulté est donc de donner un p_0 initial qui fasse converger la méthode. On rappelle l'itération de Newton :

$$p_0^{(k+1)} = p_0^{(k)} + d^{(k)},$$

avec $d^{(k)}$ solution de

$$S'(p_0^{(k)}) \cdot d = -S(p_0^{(k)}).$$

La Jacobienne de la fonction de tir est donnée par :

$$S'(p_0) \cdot d = \pi \left(\frac{\partial z}{\partial p_0}(t_f, x_0, p_0) \cdot d \right) = \pi \left(\frac{\partial z}{\partial z_0}(t_f, x_0, p_0) \cdot (0_{\mathbb{R}^n}, d) \right), \quad z_0 = (x_0, p_0),$$

et on a donc besoin de $\partial_{z_0} z(\cdot, x_0, p_0) \cdot \delta z_0$, $\delta z_0 = (0_{\mathbb{R}^n}, d)$, solution des **équations variationnelles** :

$$\dot{\delta z}(t) = \vec{\mathbf{H}}'(z(t, x_0, p_0)) \cdot \delta z(t), \quad \delta z(0) = \delta z_0.$$

Remarque 5.2.3. Si l'on écrit

$$z(t, x_0, p_0) = z(t, z_0) = z_0 + \int_0^t \vec{\mathbf{H}}(z(s, z_0)) ds,$$

et si $z(\cdot, z_0)$ est dérivable par rapport à z_0 , alors on retrouve le résultat en dérivant :

$$\frac{\partial z}{\partial z_0}(t, z_0) \cdot \delta z_0 = \delta z_0 + \int_0^t \vec{\mathbf{H}}'(z(s, z_0)) \cdot \left(\frac{\partial z}{\partial z_0}(s, z_0) \cdot \delta z_0 \right) ds.$$

Pour calculer la dérivée directionnelle $\partial_{z_0} z(t_f, z_0) \cdot \delta z_0$, la première option est d'utiliser les différences finies. Cependant, il est crucial d'utiliser en pratique des intégrateurs à pas variable qui calculent la discrétisation adaptée grâce au contrôle de l'erreur locale. Les différences finies sont connues ici pour donner de mauvais résultats car les deux grilles, évaluées dynamiquement, intervenant par exemple dans le calcul de $z(t_f, z_0)$ et $z(t_f, z_0 + \delta z_0)$, peuvent

être différentes et engendrées une non-différentiabilité artificielle. L'important est donc de forcer la grille à rester la même dans les deux cas. Ceci est connu sous le nom de **Internal Numerical Derivative** (IND [12]). Une alternative à cela est d'utiliser la **différentiation automatique** (DA) sur le code d'intégration. En raison du pas variable, le code définit seulement une fonction différentiable par morceaux et DA calcule effectivement la dérivée par morceaux, fournissant le même ordre de précision que IND pour des intégrateurs à pas variable à un pas [41]. La dernière option est d'assembler explicitement le système variationnel pour en calculer ses solutions, c'est-à-dire des champs de Jacobi. Puisque \vec{H}' est évalué le long de la solution courante, en pratique on doit intégrer simultanément sur $[t_0, t_f]$ les systèmes en z et δz ($2n + 4n^2$ dimensions pour obtenir la dérivée complète) :

$$(\dot{z}(t), \dot{\delta z}(t)) = (\vec{H}(z(t)), \vec{H}'(z(t)) \cdot \delta z(t)), \quad (z(0), \delta z(0)) = (z_0, \delta z_0).$$

Dans le cas d'un schéma explicite à un pas et pas variable, DA sur le code d'intégration et VAR (intégration du système variationnel augmenté) sont équivalents si le contrôle du pas est réalisé seulement sur la composante z (et non sur $(z, \delta z)$) et ainsi le schéma suivant commute.

$$\begin{array}{ccc} \text{(IVP)} & \xrightarrow{\text{Intégration numérique}} & z(t, z_0) \\ \text{Dérivation} \downarrow & & \downarrow \text{Dérivation} \\ \text{(VAR)} & \xrightarrow{\text{Intégration numérique}} & \frac{\partial z}{\partial z_0}(t, z_0) \end{array}$$

Exemple 5.2.1. Reprenons le Problème (P₆). Le système hamiltonien est :

$$\vec{H}(z) = (-x + p, p).$$

La fonction de tir est donnée par : $S(p_0) = p_0 \text{sh}(t_f) + x_0 e^{-t_f} - x_f$. On a donc $S'(p_0) = \text{sh}(t_f)$. Retrouvons ce résultat par les équations variationnelles. Tout d'abord, $\partial_{z_0} z(\cdot, x_0, p_0) \cdot \delta z_0$ est solution de

$$\dot{\delta z}(t) = \vec{H}'(z(t, x_0, p_0)) \cdot \delta z(t) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \delta z(t) =: A \delta z(t), \quad \delta z(0) = \delta z_0.$$

La solution en t_f de ce système linéaire est simplement

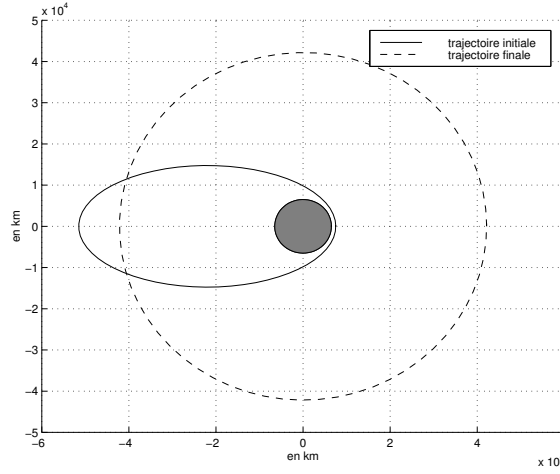
$$\frac{\partial z}{\partial z_0}(t_f, x_0, p_0) \cdot \delta z_0 = \exp(t_f A) \delta z_0 = \begin{pmatrix} e^{-t_f} & \text{sh}(t_f) \\ 0 & e^{t_f} \end{pmatrix} \delta z_0,$$

ainsi

$$S'(p_0) = \pi \left(\frac{\partial z}{\partial p_0}(t_f, x_0, p_0) \right) = \pi \left(\frac{\partial z}{\partial z_0}(t_f, x_0, p_0) \cdot (0, 1) \right) = \pi(\text{sh}(t_f), e^{t_f}) = \text{sh}(t_f). \quad \square$$

5.3 Exemple avec conditions aux limites sous-dimensionnées

Le système considéré est un satellite de masse fixée m libéré par une fusée dans le plan de l'équateur ; l'orbite initiale du satellite est une ellipse de forte excentricité. L'objectif est de réaliser le transfert en temps minimal de cette orbite elliptique à une orbite circulaire géostationnaire.



Le problème de transfert orbital à temps minimal s'écrit :

$$(P_7) \quad \left\{ \begin{array}{l} \min J(x, u, t_f) = g(t_f) := t_f \\ \dot{x}_1(t) = x_3(t), \\ \dot{x}_2(t) = x_4(t), \\ \dot{x}_3(t) = -\frac{\mu x_1(t)}{r^3(x(t))} + u_1(t), \\ \dot{x}_4(t) = -\frac{\mu x_2(t)}{r^3(x(t))} + u_2(t), \\ \|u(t)\| \leq \gamma, \quad t \in [0, t_f] \text{ p.p.}, \quad u(t) := (u_1(t), u_2(t)), \\ x_1(0) = x_{0,1}, \quad x_2(0) = x_{0,2}, \quad x_3(0) = x_{0,3}, \quad x_4(0) = x_{0,4}, \\ r(x(t_f)) = r_f, \quad x_3(t_f) = -\sqrt{\frac{\mu}{r_f^3}} x_2(t_f), \quad x_4(t_f) = \sqrt{\frac{\mu}{r_f^3}} x_1(t_f), \end{array} \right.$$

avec $r(x) := \sqrt{x_1^2 + x_2^2}$. On note $\alpha := \sqrt{\mu/r_f^3}$. La condition terminale peut se mettre sous la forme $c(x(t_f)) = 0$, avec $c: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ donné par :

$$c(x) := (r^2(x) - r_f^2, \quad x_3 + \alpha x_2, \quad x_4 - \alpha x_1).$$

Le temps final étant libre, on a la condition au temps final :

$$H(x(t_f), p(t_f), u(t_f)) = -p^0 g'(t_f) = -p^0 = 1. \quad (\text{cas normal})$$

On a de plus la condition de transversalité ($\lambda \in \mathbb{R}^3$) :

$$\begin{aligned} p(t_f) = c'(x(t_f))^T \lambda &= \begin{pmatrix} 2x_1(t_f) & 2x_2(t_f) & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ -\alpha & 0 & 0 & 1 \end{pmatrix}^T \lambda \\ &= (2x_1(t_f)\lambda_1 - \alpha\lambda_3, \quad 2x_2(t_f)\lambda_1 + \alpha\lambda_2, \quad \lambda_2, \quad \lambda_3). \end{aligned}$$

La condition de transversalité est équivalente à :

$$\begin{cases} \lambda_1 &= \frac{1}{2r(t_f)^2} \left(x_1(t_f)(p_1(t_f) + \alpha p_4(t_f)) + x_2(t_f)(p_2(t_f) - \alpha p_3(t_f)) \right), \\ \lambda_2 &= p_3(t_f), \\ \lambda_3 &= p_4(t_f), \\ 0 &= \Phi(x(t_f), p(t_f)) := x_2(t_f)(p_1(t_f) + \alpha p_4(t_f)) - x_1(t_f)(p_2(t_f) - \alpha p_3(t_f)). \end{cases}$$

En considérant les conditions aux limites, la condition finale sur le pseudo-hamiltonien et la condition de transversalité, la fonction de tir simple est donnée par :

$$S: \mathbb{R}^5 \longrightarrow \mathbb{R}^5$$

$$(p_0, t_f) \longmapsto S(p_0, t_f) := \begin{pmatrix} c(x(t_f, x_0, p_0)) \\ \Phi(z(t_f, x_0, p_0)) \\ H(z(t_f, x_0, p_0), u(z(t_f, x_0, p_0))) + p^0 \end{pmatrix}.$$

Remarque 5.3.1. Le contrôle maximisant $z \mapsto u(z)$ est donné par la condition de maximisation.

On peut ne pas éliminer λ et l'inclure dans la fonction de tir simple :

$$S: \mathbb{R}^8 \longrightarrow \mathbb{R}^8$$

$$(p_0, t_f, \lambda) \longmapsto S(p_0, t_f, \lambda) := \begin{pmatrix} c(x(t_f, x_0, p_0)) \\ p(t_f, x_0, p_0) - c'(x(t_f, x_0, p_0))^T \lambda \\ H(z(t_f, x_0, p_0), u(z(t_f, x_0, p_0))) + p^0 \end{pmatrix}.$$

Nous renvoyons à la discussion sur le multiplicateur de Lagrange λ page 69 pour comprendre comment de manière systématique on peut l'éliminer de la définition de la fonction de tir.

5.4 Tir multiple de stabilité

La méthode de tir multiple peut être utilisée afin de renforcer la stabilité des équations de tir [13]. Pour l'exemple, considérons le Problème (P_L) sous les Hypothèses (H2) et supposons que pour toute extrémale de Pontryagin-Boltyanski $(z(\cdot), p^0, u(\cdot))$, on peut écrire $u(t) = u(z(t))$, avec $z \mapsto u(z)$ de classe au moins \mathcal{C}^1 . Pour améliorer la stabilité, on discrétise l'intervalle $[0, t_f]$ en N sous-intervalles $[t_i, t_{i+1}]$. Les valeurs $z(t_i)$, à l'entrée de chaque intervalle, sont alors les inconnues du système à résoudre. On ajoute ensuite aux conditions aux limites et/ou de transversalité (dans cet exemple il n'y en a pas), des conditions de raccordement aux instants t_i assurant la continuité de la fonction d'état et du vecteur adjoint en ces instants.

Dans le cas de la méthode de tir simple, sur cette exemple, nous sommes amenés à résoudre le problème aux deux bouts (BVP_2) que l'on peut réécrire sous la forme suivante :

$$\exp(t_f \vec{H})(z_0) = z_f, \quad b(z_0, z_f) = 0,$$

où $\exp(tF)(x_0)$ est la solution au temps t du problème de Cauchy $\dot{x}(t) = F(x(t))$, $x(0) = x_0$. On appelle ce genre de problème un (TPBVP) pour “Two-Points Boundary Value Problem”. Après discrétisation de l’intervalle de temps, on obtient le problème aux points multiples suivant :

$$\exp((t_{i+1} - t_i)\vec{H})(z_i) = z_{i+1}, \quad i \in \llbracket 0, N-1 \rrbracket, \quad b(z_0, z_f) = 0,$$

appelé (MPBVP) pour “Multiple-Points Boundary Value Problem”. Pour plus de détails sur la résolution de ce genre de problème, nous renvoyons à [69].

5.5 Tir simple avec détections des commutations

On s’intéresse ici au cas particulier de problèmes présentant des commutations, de type bang-bang par exemple. Supposons que la condition de maximisation du pseudo-hamiltonien nous donne

$$\begin{aligned} u(t) &= u_1(t, z(t)), \quad \text{si } \varphi(t, z(t)) < 0, \\ u(t) &= u_2(t, z(t)), \quad \text{si } \varphi(t, z(t)) > 0, \end{aligned}$$

où φ est la fonction de commutation dont les zéros correspondent aux commutations du contrôle optimal. Supposons que u_1 et u_2 soient lisses. Considérons une extrémale de Pontryagin-Boltyanski de référence $z(\cdot)$ et supposons qu’il n’existe qu’un nombre fini d’instant de commutations t tels que $\varphi(t, z(t)) = 0$ le long de celle-ci. On note $H_1(t, z) := H(t, z, u_1(t, z))$ et $H_2(t, z) := H(t, z, u_2(t, z))$. Supposons qu’au voisinage de l’extrémale de référence, toute autre extrémale de Pontryagin-Boltyanski soit solution du problème à valeur initiale discontinu suivant

$$(IVP)_{\text{disc}} \begin{cases} \dot{z}(t) = \vec{H}_1(t, z(t)), & \text{si } \varphi(t, z(t)) < 0 \\ \dot{z}(t) = \vec{H}_2(t, z(t)), & \text{si } \varphi(t, z(t)) > 0 \\ z(t_0) = z_0. \end{cases}$$

Dans ce cas il est possible de définir une fonction de tir ne dépendant que du point initial z_0 . Pour cela, il faut détecter durant l’intégration, les instants de commutations. L’idée est donc durant l’intégration, de détecter les commutations et de calculer l’instant et le point de commutation d’où repart l’intégration. Ces étapes sont primordiales, il ne faut pas laisser le contrôle du pas à l’intégrateur [41, 58] mais bien détecter pour améliorer à la fois le temps de calcul et la précision des calculs. La détection permet en effet de réduire considérablement le nombre de pas acceptés et rejetés. Une bonne détection est de plus importante afin de conserver l’ordre lors de l’intégration numérique [57].

Une fois le pas accepté et une commutation détectée, il est nécessaire de la localiser. Pour cela, nous présentons deux possibilités. Nous pouvons tout d’abord utiliser la sortie dense fournit par l’intégrateur pour calculer le lieu de commutation par dichotomie sur le signe de la fonction de commutation. Cependant, pour garder la précision du schéma utilisé, il est préférable d’utiliser une méthode de dichotomie (ou une méthode de Newton) toujours sur la fonction de commutation mais en prenant comme inconnue la longueur d’un pas. La fonction de commutation est donc calculé en un vrai pas d’intégration, voir schéma Figure 5.1. Supposons que l’on ait une commutation passant de \vec{H}_1 à \vec{H}_2 . On note alors $z_1(t)$ la partie de l’extrémale solution de $\dot{z}(t) = \vec{H}_1(t, z(t))$ et $z_2(t)$, celle solution de $\dot{z}(t) = \vec{H}_2(t, z(t))$, après la commutation. Ainsi, en l’instant initial on a $z_1(t_0, t_0, z_0) = z_0$, en l’instant de commutation τ ,

on a $z_1 := z_1(\tau, t_0, z_0)$ et $z_2(\tau, \tau, z_1) = z_1$, et en l'instant final, noté t_f , après la commutation, on a $z_2 := z_2(t_f, \tau, z_1)$.

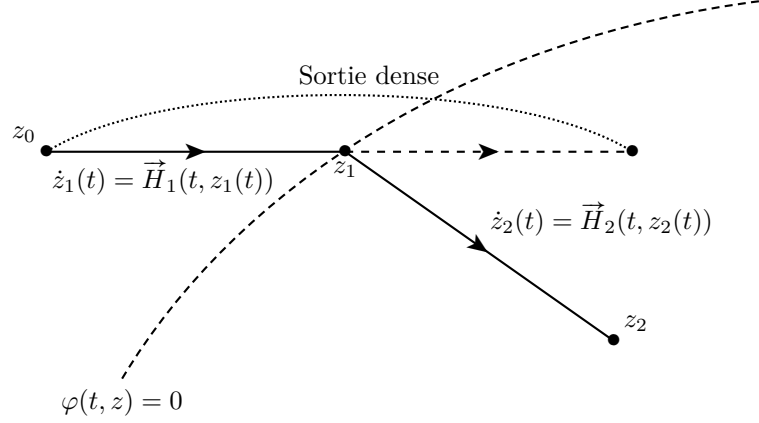


FIGURE 5.1 – Traversée d'une commutation.

Comme pour les méthodes de tir simple et multiple, nous avons besoin de la jacobienne de la fonction de tir. Celle-ci est décrite simplement comme pour la méthode de tir simple et fait donc intervenir la variation de l'extrémale en t_f par rapport à la condition initiale z_0 . Reprenons le calcul de [58] dans le cas particulier d'une commutation passant de \vec{H}_1 à \vec{H}_2 avec les notations définies précédemment. On pose $g(t, z) := \varphi(t, z_1(t, t_0, z))$, alors si

$$\frac{\partial g}{\partial t}(\tau, z_0) = \frac{\partial \varphi}{\partial t}(\tau, z_1) + \frac{\partial \varphi}{\partial z}(\tau, z_1) \vec{H}_1(\tau, z_1) \neq 0,$$

alors il existe localement $z_0 \mapsto \tau(z_0)$ tel que $\varphi(\tau(z_0), z_1(t, t_0, z_0)) = 0$ et

$$\tau'(z_0) = -\frac{\frac{\partial g}{\partial z}(\tau(z_0), z_0)}{\frac{\partial g}{\partial t}(\tau(z_0), z_0)} = -\frac{\frac{\partial \varphi}{\partial z}(\tau, z_1) \frac{\partial z_1}{\partial z_0}(\tau, t_0, z_0)}{\frac{\partial \varphi}{\partial t}(\tau, z_1) + \frac{\partial \varphi}{\partial z}(\tau, z_1) \vec{H}_1(\tau, z_1)}.$$

On note

$$\delta z_2(t) := \frac{d}{dz_0} z_2(t, \tau(z_0), z_1(\tau(z_0), t_0, z_0)) \cdot \delta z_0$$

et

$$\delta z_1(t) := \frac{\partial z_1}{\partial z_0}(t, t_0, z_0) \cdot \delta z_0.$$

La variation de z_2 en t_f par rapport à la condition initiale z_0 est alors donnée par

$$\delta z_2(t_f) = \frac{\partial z_2}{\partial t_0}(t_f, \tau, z_1) \tau'(z_0) \delta z_0 + \frac{\partial z_2}{\partial z_0}(t_f, \tau, z_1) \left(\frac{\partial z_1}{\partial t_f}(\tau, t_0, z_0) \tau'(z_0) \delta z_0 + \delta z_1(\tau) \right),$$

où τ et z_1 sont des fonctions de z_0 . On se place maintenant en $t_f = \tau^+$ juste après la commutation. En passant à la limite, puisque

$$\frac{\partial z_2}{\partial t_0}(\tau^+, \tau, z_1) = -\vec{H}_2(\tau^+, z_1), \quad \frac{\partial z_2}{\partial z_0}(\tau^+, \tau, z_1) = I_{2n},$$

et

$$\frac{\partial z_1}{\partial t_f}(\tau, t_0, z_0) = \vec{H}_1(\tau, z_1)$$

alors

$$\delta z_2(\tau^+) = \delta z_1(\tau) + \left(\vec{H}_1(\tau, z_1) - \vec{H}_2(\tau^+, z_1) \right) \tau'(z_0) \delta z_0.$$

En remplaçant $\tau'(z_0)$ par sa valeur et puisque \vec{H}_2 est continue au temps τ , on obtient finalement,

$$\delta z_2(\tau^+) = \left(I_{2n} + \frac{\left(\vec{H}_2(\tau, z_1) - \vec{H}_1(\tau, z_1) \right) \frac{\partial \varphi}{\partial z}(\tau, z_1)}{\frac{\partial \varphi}{\partial t}(\tau, z_1) + \frac{\partial \varphi}{\partial z}(\tau, z_1) \vec{H}_1(\tau, z_1)} \right) \delta z_1(\tau) =: (I_{2n} + \alpha) \delta z_1(\tau). \quad (5.1)$$

Ainsi si les champs de vecteurs hamiltoniens associés à H_1 et H_2 sont égaux en τ , alors le terme de saut α est nul, et les équations variationnelles sont continues en ce point. En principe, en une commutation on a $\alpha \neq 0$ et on observe une discontinuité dans le calcul des champs de Jacobi. Dans la pratique, on intègre le système en z et δz et lorsqu'une commutation est détectée et localisée, on met à jour δz puis on continue l'intégration. Ce calcul est utilisé à la fois pour le calcul de la jacobienne de la fonction de tir et pour les champs de Jacobi, la condition initiale δz_0 étant arbitraire. Dans le cas particulier où $\varphi = H_2 - H_1$ alors $\partial_z \varphi = \partial_z H_2 - \partial_z H_1$ et $\partial_z \varphi \vec{H}_1 = \{H_1, H_2\}$.

Il est possible que l'extrémale en z_2 dépende aussi de l'instant initial si par exemple on utilise la détection de commutation à l'intérieur d'un arc dont l'instant initial est une inconnue d'un problème résolu par des méthodes de tir multiple. Dans ce cas la formule de la dérivée par rapport à l'instant initial reste identique à (5.1), simplement la condition initiale δz_0 est remplacée par $\delta t_0 = -H_1(t_0, z_0)$.

Un des intérêts de la méthode de tir simple avec détections des commutations et mise à jour des équations variationnelles, est qu'il est possible à la fois de résoudre des problèmes de type bang-bang via les conditions nécessaire du premier ordre, mais aussi de vérifier l'optimalité locale via les conditions du deuxième ordre, dans un cadre particulier, comme défini dans [66] par exemple. Considérons un système affine mono-entrée dont les solutions optimales sont bang-bang. Dans [66] les auteurs définissent le premier instant conjugué d'une extrémale de manière analogue à la définition classique, seulement la forme quadratique impliquée dans la définition provient d'une application différente de l'application entrée-sortie dans le cadre lisse. De nombreuses références sur les conditions du deuxième ordre pour des problèmes bang-bang sont données dans [66]. Les auteurs utilisent une procédure de régularisation qui consiste à introduire de nouveaux contrôles artificiels sous certaines contraintes, pour rendre le problème de contrôle optimal lisse pour $\lambda \in [0, 1[$ par exemple. L'apport principal dans cet article est alors de prouver la convergence, lorsque $\lambda \rightarrow 1$, du premier instant conjugué défini dans le cadre lisse, vers le premier instant conjugué pour le problème de départ, qui lui est bang-bang. Calculer le premier instant conjugué revient à trouver le premier zéro d'un déterminant, défini à partir des champs de Jacobi, le long de l'extrémale de référence. Nous reprenons Figure 5.2 l'exemple numérique du problème de Rayleigh dont le premier instant conjugué se trouve être le deuxième instants de commutation. Nous pouvons voir Figure 5.2, que le calcul précis des équations variationnelles permet de retrouver ce résultat.

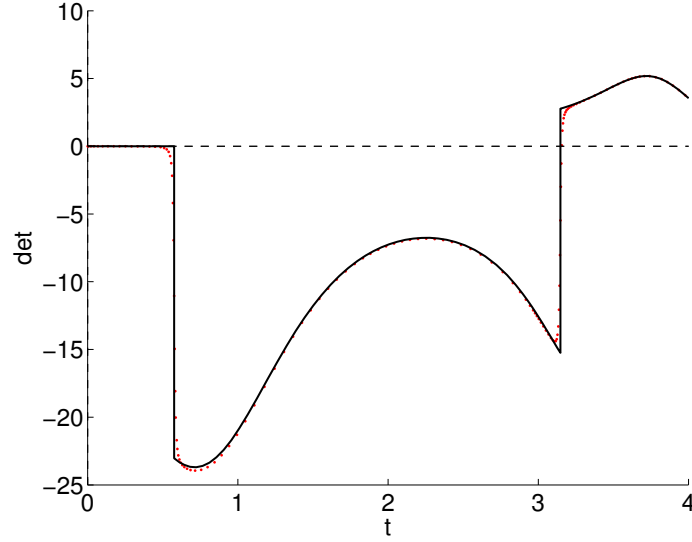


FIGURE 5.2 – Déterminant s’annulant au premier instant conjugué, en points rouges pour le problème régularisé avec $\varepsilon = 0.01$, cf. [66, Figure 10] et en trait plein noir pour le problème de départ bang-bang.

5.6 Difficultés numériques des méthodes de tir simple

Nous reprenons dans cette section, l’illustration des difficultés numériques des méthodes de tir simple présentées dans [36]. Cette présentation est une motivation à l’introduction des méthodes de tir multiple et des méthodes homotopiques dans la résolution des problèmes de contrôle optimal. Nous renvoyons à [36] pour plus de détails et pour une présentation des méthodes homotopiques dans le cadre de la théorie du degré topologique.

Pour illustrer les difficultés numériques, nous considérons le problème du double intégrateur avec la minimisation L^1 du contrôle :

$$\begin{cases} \min \int_0^{t_f} |u(t)| dt, \\ \dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = u(t), \quad |u(t)| \leq \gamma, \\ x(0) = x_0, \quad x(t_f) = x_f, \end{cases}$$

avec x_0, x_f dans \mathbb{R}^2 , $t_f \geq 0$ et $\gamma > 0$ fixés. D’après le PMP, nous devons considérer le pseudo-hamiltonien suivant (sous la forme normale) :

$$H(x, p, u) = p_1 x_2 + p_2 u - |u|.$$

Le contrôle maximisant est donné par $u = 0$ si $|p_2| < 1$, $u = \gamma \operatorname{sign} p_2$ si $|p_2| > 1$ et $u \in [-\gamma, \gamma]$ si $|p_2| = 1$. On note $u(p_2)$ le contrôle maximisant défini ci-avant, voir Figure 5.3. On est alors amené à résoudre le problème aux deux bouts non lisse (même mal défini en $p_2 = \pm 1$) suivant :

$$\begin{cases} \dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = u(p_2(t)), \quad \dot{p}_1(t) = 0, \quad \dot{p}_2(t) = -p_1(t), \\ x(0) = x_0, \quad x(t_f) = x_f. \end{cases}$$

La fonction de tir est donnée Figure 5.4 avec le covecteur initial $p_0 =: (\alpha, \beta)$.

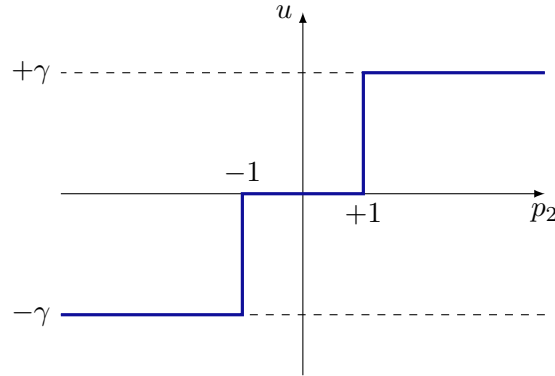


FIGURE 5.3 – Le contrôle maximisant n'est pas continu. Puisque $p_2(t) = -\alpha t + \beta$, alors pour une paire (α, β) donnée, il y aura 0, 1 ou 2 commutations (c'est-à-dire passage par $|p_2| = 1$).

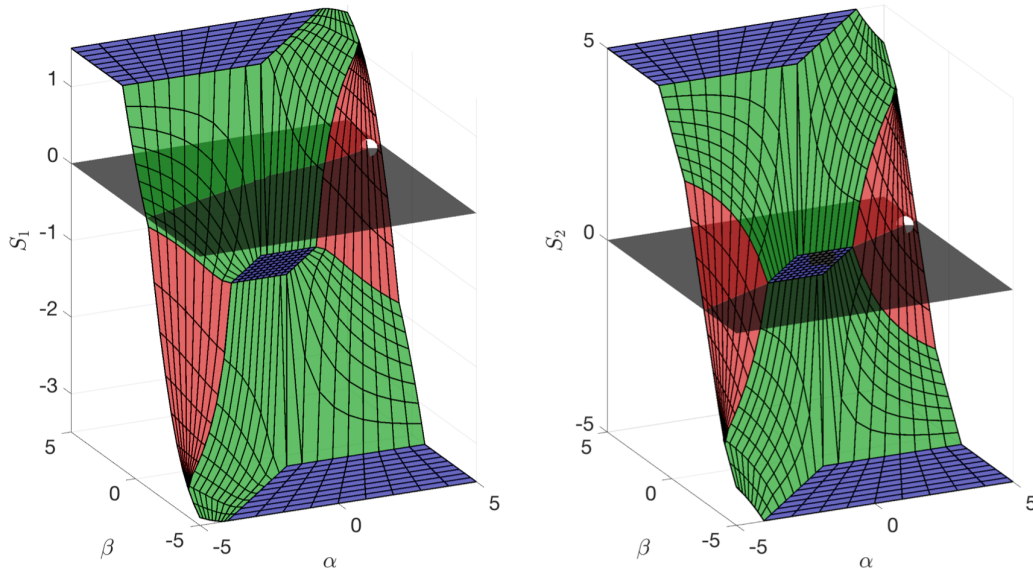


FIGURE 5.4 – Fonction de tir du problème du double intégrateur avec minimisation L^1 . Sur cet exemple, $\gamma = 5$, $T = 1$, $x_0 = (-1, 0)$ et $x_f = (0, 0)$. La boule blanche est la solution à l'intersection de la fonction de tir et des plans $S_1 = 0$ et $S_2 = 0$.

On peut remarquer que la fonction de tir, cf. Figure 5.4, n'est pas continue en les points $(0, \pm 1)$, elle est non différentiable aux interfaces des différentes régions, et elle est de plus constante sur les régions bleues. Les neuf régions sont caractérisées par des structures de contrôle différentes, cf. Figure 5.5. Une conséquence des propriétés de la fonction de tir concernant l'utilisation de la méthode de Newton pour résoudre l'équation de tir simple $S(p_0) = 0$ est que si le point initial p_0 n'est pas choisi dans la bonne région, c'est-à-dire celle qui correspond à la bonne structure de la commande optimale, alors l'algorithme risque de ne pas converger. Sur ce petit exemple en effet, l'algorithme peut générer un point dans une région où la fonction de tir est constante. Or dans la pratique nous ne connaissons pas

la structure de la commande optimale. L'introduction des méthodes homotopiques dans le chapitre suivant a pour objectif de palier à cette difficulté et fournit à la fois la structure optimale et un point initial dans la bonne région. Une fois la structure connue et un bon point initial, nous pouvons utiliser ces informations et données pour résoudre le problème avec une méthode de tir multiple présentée dans la section suivante.

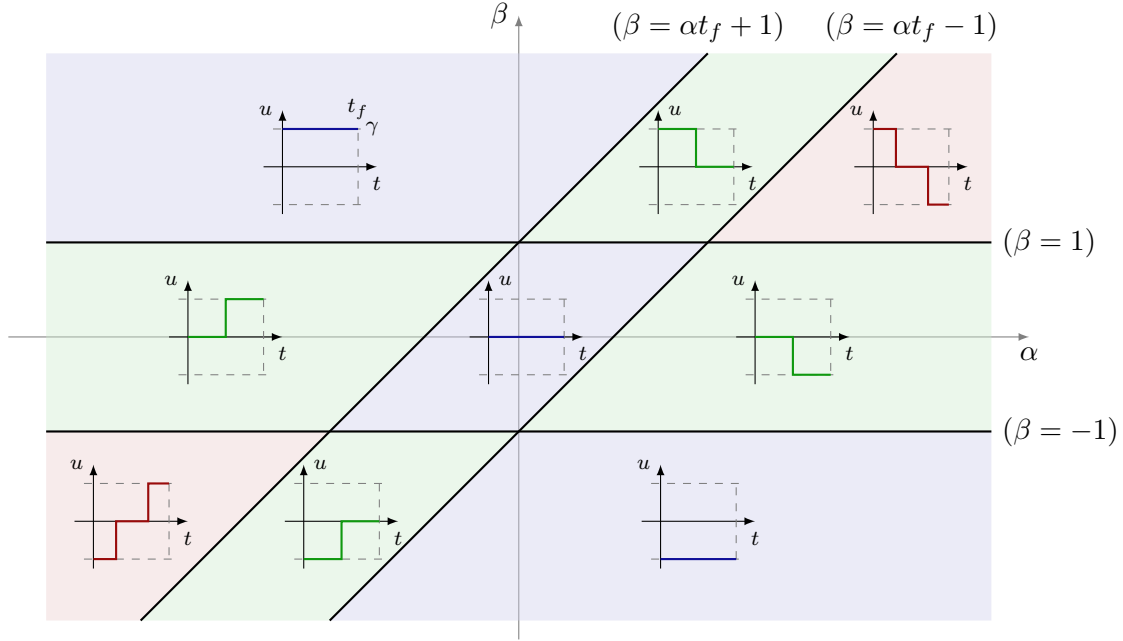


FIGURE 5.5 – Diagramme de bifurcation des structures des lois de commandes (sur $[0, t_f]$) dans le plan (α, β) . On rappelle que $p_2(t) = -\alpha t + \beta$ et que l'on a $u(p_2) = 0$ si $p_2 \in]-1, 1[$, $u(p_2) = \gamma \text{sign}(p_2)$ si $|p_2| > 1$ et $u(p_2) \in [-\gamma, \gamma]$ sinon.

Remarque 5.6.1. Une alternative à l'utilisation des méthodes homotopiques et l'utilisation des méthodes directes afin de déterminer ma structure optimale et obtenir un point initial dans la bonne région, suffisamment proche de la solution pour faire converger la méthode de Newton utilisée dans la méthode de tir multiple.

5.7 Tir multiple de structure

Considérons un système affine en le contrôle scalaire de la forme $H = H_0 + u H_1$ avec la contrainte $\|u\| \leq 1$. Si l'extrémale solution du problème de contrôle optimal est la concaténation d'arcs bangs (où le contrôle vérifie $\|u\| = 1$) et d'arcs singuliers (où $H_1 = 0$) alors il devient nécessaire d'utiliser une méthode de tir multiple que nous qualifierons de **tir multiple de structure**. Imaginons que la solution soit composée de trois arcs Bang-Singulier-Bang, alors la fonction de tir en plus de l'adjoint initial, aura comme inconnues les deux instants de commutations. Puisque nous avons deux inconnues supplémentaires, on s'attend à avoir deux conditions supplémentaires. Celles-ci sont données par la Proposition 4.5.1 : puisque la surface singulière Σ_s est invariante par le flot singulier, il suffit d'imposer que le premier arc bang rejoigne Σ_s au premier instant de commutation. En imposant le contrôle singulier sur

le deuxième arc, on s'assure que celui-ci reste sur Σ_s . Enfin, pour le troisième arc, il suffit d'imposer le contrôle bang pour quitter la surface singulière.

Remarque 5.7.1. Il est possible de combiner tir multiple de structure et tir multiple de stabilité. Dans ce cas, les inconnues sont les instants de commutations et les points de raccordement. Il est aussi possible le long d'un arc quelconque d'ajouter des noeuds de raccordement.

Exemple 5.7.1. Considérons le problème du double intégrateur suivant :

$$\begin{cases} \min \frac{1}{2} \int_0^{t_f} (x_1^2(t) + x_2^2(t)) dt, & t_f = 5, \\ \dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = u(t), \quad u(t) \in [-1, 1] \\ x(0) = (0, 1). \end{cases}$$

Supposons que l'on ait déterminé la structure optimale par n'importe quelle méthode, par exemple une méthode directe ou par homotopie. On sait donc que la structure optimale est composée de deux arcs : un premier arc bang avec $u = -1$ puis un arc singulier, voir Figure 5.6.

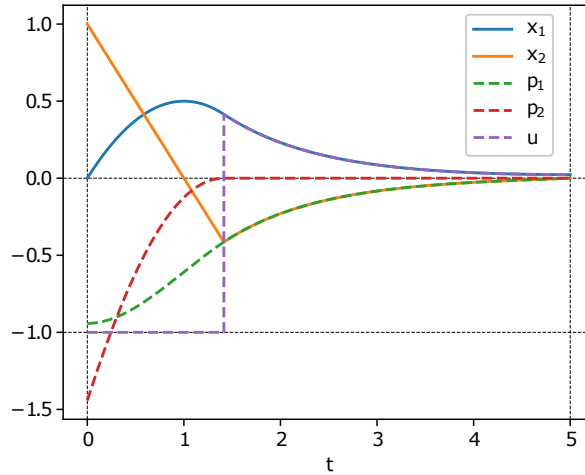


FIGURE 5.6 – Etats, adjoints et contrôles pour le problème du double intégrateur.

D'après le PMP, nous devons considérer le pseudo-hamiltonien suivant (sous la forme normale) :

$$H(x_1, x_2, p_1, p_2, u) = -(x_1^2 + x_2^2)/2 + p_1 x_2 + p_2 u.$$

Les arcs singuliers sont d'ordre minimal, donnés par la condition $p_2 = 0$. La surface singulière est donnée par l'ensemble $\Sigma_s = \{p_2 = x_2 - p_1 = 0\}$ du fibré cotangent $T^*\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$, et le contrôle singulier en forme fermée est $u = x_1$. Nous avons donc une compétition entre trois flots hamiltoniens, respectivement associés à

$$H_{\pm}(x, p) := H(x, p, \pm 1), \quad H_s(x, p) := H(x, p, x_1).$$

Sachant la structure bang-singulière de la solution optimale, nous pouvons définir une fonction de tir multiple de structure ayant comme inconnues le vecteur adjoint initial $p_0 \in \mathbb{R}^2$ et

l'instant de commutation $\tau \in \mathbb{R}$. Les conditions sont données par les conditions de transversalité du PMP et par la Proposition 4.5.1 :

$$p_1(t_f) = 0, \quad p_2(t_f) = 0, \quad p_2(\tau) = 0, \quad x_2(\tau) - p_1(\tau) = 0.$$

Nous pouvons remarquer que la condition $p_2(t_f) = 0$ est redondante puisque c'est automatiquement le cas pour une extrémale singulière. Nous avons donc trois conditions pour trois inconnues, ce qui définit notre fonction de tir. \square

Exemple 5.7.2. On considère le problème de contrôle optimal suivant :

$$\begin{cases} \min \int_0^{t_f} x^2(t) dt \\ \dot{x}(t) = u(t), \quad |u(t)| \leq 1, \quad t \in [0, t_f] \text{ p.p.}, \\ x(0) = 1, \quad x(t_f) = 1/2. \end{cases}$$

où $t_f = 2$. On suppose que l'on connaît la structure optimale de la solution, composée d'un arc bang avec $u = -1$, suivi d'un arc singulier avec $u = 0$, puis d'un arc bang avec $u = 1$, voir Figure 5.7.

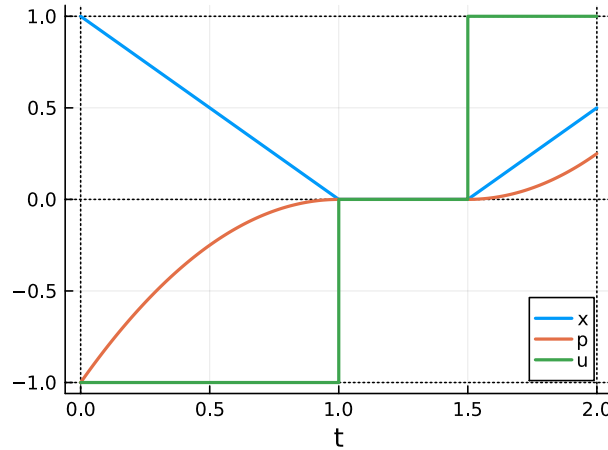


FIGURE 5.7 – Etat, adjoints et contrôle pour le problème bang-singulier-bang.

D'après le PMP, nous devons considérer le pseudo-hamiltonien suivant (sous la forme normale) : $H(x, p, u) = -x^2 + pu$. Les arcs singuliers sont d'ordre minimal, donnés par la condition $p = 0$. La surface singulière est donnée par l'ensemble $\Sigma_s = \{p = x = 0\}$ du fibré cotangent $T^*\mathbb{R} \simeq \mathbb{R} \times \mathbb{R}$, et le contrôle singulier en forme fermée est $u = 0$. Nous avons donc une compétition entre trois flots hamiltoniens, respectivement associés à $H_{\pm}(x, p) := H(x, p, \pm 1)$, $H_s(x, p) := H(x, p, 0)$. Sachant la structure bang-singulière-bang de la solution optimale, nous pouvons définir une fonction de tir multiple de structure ayant comme inconnues le vecteur adjoint initial $p_0 \in \mathbb{R}$ et les instants de commutations $\tau_1, \tau_2 \in \mathbb{R}$. Les conditions sont données par la condition terminale $x(t_f) = 1/2$ et par la Proposition 4.5.1 :

$$x(t_f) = 1/2, \quad p(\tau_1) = 0, \quad x(\tau_1) = 0.$$

Nous avons donc trois conditions pour trois inconnues, ce qui définit notre fonction de tir. \square

Corrections des exercices

Solution de l'Exercice 1.7.1 p. 19 :

1. Soit u un contrôle singulier sur $[0, T]$. On note $x(\cdot)$ la trajectoire associée et $p(\cdot)$ l'application adjointe associée issue de la caractérisation hamiltonienne des contrôles singuliers, cf. Théorème 1.6.5. Le contrôle étant singulier, il vérifie $\partial_u H(x(t), p(t), u(t)) = 0$ pour tout $t \in [0, T]$. Or,

$$\frac{\partial H}{\partial u}(x, p, u) = 0 \iff -p_1 x_1 + p_2 u = 0.$$

Donc, en supposant $p_2 \neq 0$, le contrôle singulier vérifie

$$u(t) = u_s(x(t), p(t)) = \frac{p_1(t) x_1(t)}{p_2(t)}$$

avec u_s le contrôle singulier en fonction de l'état $x \in \mathbb{R}^2$ et de l'adjoint $p \in \mathbb{R}^2$ donné par

$$u_s(x, p) := \frac{p_1 x_1}{p_2}.$$

Solution de l'Exercice 3.5.1 p. 46 :

1. On a

$$\frac{\partial L}{\partial v}(t, x, v) = \frac{v}{\sqrt{1+v^2}}.$$

Soit x une extrémale, *i.e.* une trajectoire vérifiant l'équation de Euler-Lagrange. Alors (on omet l'argument t dans x et sa dérivée) :

$$\begin{aligned} \frac{\partial L}{\partial x}(t, x, \dot{x}) - \frac{d}{dt} \frac{\partial L}{\partial v}(t, x, \dot{x}) &= -\frac{d}{dt} \frac{\dot{x}}{\sqrt{1+\dot{x}^2}} = -\left(\frac{\ddot{x}\sqrt{1+\dot{x}^2} - \dot{x} \frac{\dot{x}\ddot{x}}{\sqrt{1+\dot{x}^2}}}{1+\dot{x}^2} \right) \\ &= \frac{\ddot{x}}{(1+\dot{x}^2)^{\frac{3}{2}}} (-1 - \dot{x}^2 + \dot{x}^2) = 0 \iff \ddot{x} = 0. \end{aligned}$$

2. L'équation de Euler-Lagrange nous donne

$$\frac{\partial L}{\partial x}(t, x, \dot{x}) - \frac{d}{dt} \frac{\partial L}{\partial v}(t, x, \dot{x}) = 2 \left(x - \frac{d}{dt} \dot{x} \right) = 2(x - \ddot{x}) = 0 \iff \ddot{x} - x = 0.$$

3. Le lagrangien associé au problème est donné par $L(t, x, v) := x^2 + v^2$. D'après la question précédente les extrémales vérifient $\ddot{x}(t) - x(t) = 0$ pour tout temps $t \in [0, 1]$. Elles sont donc de la forme

$$x(t) = \lambda \cosh t + \mu \sinh t.$$

L'unique extrémale vérifiant les conditions aux limites $x(0) = 0$ et $x(1) = 1$ est donnée par $\lambda = 0$ et $\mu = 1/\sinh 1$, c'est-à-dire

$$x(t) = \frac{\sinh t}{\sinh 1}.$$

Solution de l'Exercice 3.5.2 p. 46 :

1. L'équation de Euler-Lagrange nous donne

$$\frac{\partial L}{\partial x}(t, x, \dot{x}) - \frac{d}{dt} \frac{\partial L}{\partial v}(t, x, \dot{x}) = -\frac{d}{dt} \dot{x} = -\ddot{x} = 0.$$

Ainsi les extrémales vérifient $x(t) = \lambda t + \mu$ où λ et μ sont deux constantes. Les conditions aux limites nous donne $x(0) = \mu = A$ et $x(T) = \lambda T + A = B$, c'est-à-dire

$$x(t) = \frac{T-t}{T}A + \frac{t}{T}B.$$

Remarque. La trajectoire est le segment $[AB]$.

2. On a

$$\dot{x}(t) = \frac{1}{T}(B - A).$$

Ainsi,

$$l(x) = \int_0^T \frac{\|B - A\|}{T} dt = \|B - A\|.$$

3. C'est la longueur du segment $[AB]$, c'est-à-dire la distance entre les points A et B !

Solution de l'Exercice 3.5.4 p. 50 :

1. Le hamiltonien est donné par $H(\alpha, p) = p\varphi(\alpha, p) - L(\alpha, \varphi(\alpha, p))$ où $\varphi(\alpha, p)$ est la solution de $p - \partial_{\dot{\alpha}} L(\alpha, \dot{\alpha}) = 0$, l'inconnue étant $\dot{\alpha}$. On a donc

$$p - ml^2 \dot{\alpha} = 0 \iff \varphi(\alpha, p) = \dot{\alpha} = \frac{p}{ml^2} = \frac{p}{I}$$

et le hamiltonien est donné par

$$H(\alpha, p) = \frac{p^2}{I} - \left(\frac{p^2}{2I} + \omega^2 I \cos \alpha \right) = \frac{p^2}{2I} - \omega^2 I \cos \alpha.$$

Remarque. On aurait pu utiliser directement le fait que $H = T + V$.

2. H est constant le long de toute extrémale car le système est autonome, *i.e.* Il ne dépend pas du temps.
3. Puisque le hamiltonien est constant le long de l'extrémale et puisque $I > 0$, on a pour tout temps t :

$$\frac{p(t)^2}{2I} = h + \omega^2 I \cos \alpha(t) \geq 0,$$

c'est-à-dire $-\omega^2 I \cos \alpha(t) \leq h$.

4.

$$\begin{aligned} H = \omega^2 I = \frac{p_0^2}{2I} - \omega^2 I \cos \alpha_0 &\iff p_0^2 = 2\omega^2 I^2 (1 + \cos \alpha_0) \\ &\iff \dot{\alpha}_0^2 = 2\omega^2 (1 + \cos \alpha_0). \end{aligned}$$

Solution de l'Exercice 3.7.1 p. 55 :

1. On pose $H(q, p, u) = p \cdot u + p^0 \frac{1}{2} \|u\|^2$. La condition de maximisation implique $p^0 = -1$ car sinon $u \mapsto p \cdot u$ n'a de max que si $p = 0$, or $(p, p^0) \neq (0, 0)$. L'application partielle $u \mapsto H(x, p, u)$ est alors concave (strictement) donc la condition de maximisation est équivalente à la CN1 : $\frac{\partial H}{\partial u}(x, p, u) = p - u = 0$. Ainsi le contrôle maximisant est donné par $u[p] = p$. Les équations hamiltoniennes sont donc données par

$$\dot{q}(t) = p(t), \quad \dot{p}(t) = 0.$$

Ainsi $p(t) = \alpha$ et $q(t) = \alpha t + \beta$ où α et β sont deux constantes à déterminer. La condition initiale nous donne $q(0) = b = q_0$ et la condition terminale $q(t_f) = at_f + q_0 = q_1$, c-a-d $a = (q_1 - q_0)/t_f$. Au final, on a

$$p(t) = \frac{q_1 - q_0}{t_f}, \quad u(t) = p(t), \quad q(t) = (q_1 - q_0) \frac{t}{t_f} + q_0, \quad p^0 = -1.$$

2.

$$\int_0^{t_f} \|u(t)\| dt = \int_0^{t_f} \frac{\|q_1 - q_0\|}{t_f} dt = \|q_1 - q_0\|.$$

3. On vient de calculer la distance entre les points q_0 et q_1 .

Solution de l'Exercice 3.7.2 p. 55 :

1. Par les mêmes calculs qu'à l'exercice précédent, le système hamiltonien est donné par

$$\dot{q}(t) = p(t), \quad \dot{p}(t) = 0.$$

Ainsi $p(t) = \alpha$ et $q(t) = \alpha t + \beta$ où α et β sont deux constantes à déterminer. La condition initiale nous donne $q(0) = \beta = 0_{\mathbb{R}^2}$. La condition terminale peut s'écrire sous la forme $c(q) = x - 1$, où $q = (x, y)$. Si on note $\alpha = (\alpha_1, \alpha_2)$, alors la condition terminale nous donne $x(t_f) = \alpha_1 t_f = 1$, c-a-d $\alpha_1 = 1/t_f$. Enfin, la

condition de transversalité nous donne $p(t_f) = \lambda \nabla c(q(t_f)) = (\lambda, 0) = \alpha$, c-a-d $\lambda = 1/t_f$ et $\alpha_2 = 0$. Au final, on a

$$p(t) = (\frac{1}{t_f}, 0), \quad u(t) = p(t), \quad \lambda = \frac{1}{t_f}, \quad q(t) = (\frac{t}{t_f}, 0), \quad p^0 = -1.$$

Deuxième partie

Annexes

Théorèmes d'analyse et optimisation

A.1	Théorèmes des accroissements finis et applications	101
A.2	Dérivabilité dans le cas fonctionnel	105
A.3	Théorème du point fixe	109
A.4	Application inversion	110
A.5	Théorème d'inversion locale	113
A.6	Théorème des fonctions implicites	117
A.7	Conditions nécessaires d'optimalité	118

A.1 Théorèmes des accroissements finis et applications

Le théorème des accroissements finis donne une majoration de $f(b) - f(a)$ sous des hypothèses appropriées portant sur la dérivée de f . Il s'agit essentiellement d'un résultat portant sur les fonctions de la variable réelle. Rappelons que le théorème de Rolle¹ pour des fonctions à valeurs réelles assure l'existence d'un point $c \in]a, b[$, $a < b$ réels, tel que

$$f(b) - f(a) = f'(c)(b - a)$$

sous les hypothèses suivantes : f continue sur $[a, b]$ et dérivable sur $]a, b[$. Ce résultat ne subsiste pas pour les fonctions à valeurs vectorielles, comme le montre l'exemple de la fonction à valeurs dans \mathbb{R}^2 , $f(x) = (\cos(2\pi x), \sin(2\pi x))$ sur $[0, 1]$.

Exemple A.1.1. La fonction $f: \mathbb{R} \rightarrow \mathbb{R}^2$ définie par $f(x) = (\cos(2\pi x), \sin(2\pi x))$, a, en chaque point une dérivée de norme 2π , et il n'existe aucun point $c \in]0, 1[$ tel que $f(1) - f(0) = f'(c)$. En effet, $f'(x) = 2\pi(-\sin(2\pi x), \cos(2\pi x))$, donc pour tout $x \in \mathbb{R}$, $\|f'(x)\| = 2\pi$. Et puisque $f(1) = f(0)$, on ne peut avoir $\|f(1) - f(0)\| = 0 = \|f'(c)\| = 2\pi$. \square

Rappelons qu'il existe aussi une version intégrale du théorème des accroissements finis pour les fonctions à valeurs dans \mathbb{R} . Pour une fonction f à valeurs réelles, définie et continue sur $[a, b]$, il existe $c \in]a, b[$ tel que

$$f(c) = \frac{1}{b-a} \int_a^b f(x) \, dx.$$

Pour les fonctions de la variable réelle et à valeurs dans un espace vectoriel normés quelconque, nous avons le résultat suivant.

1. cf. <https://tinyurl.com/Rolle-wikipedia>.

Théorème A.1.1 – 1^{re} forme des accroissements finis

Soient $a < b$ deux réels, $(F, \|\cdot\|_F)$ un espace vectoriel normé, $f \in \mathcal{C}^0([a, b], F)$ et $g \in \mathcal{C}^0([a, b], \mathbb{R})$. Supposons f et g dérivables dans $]a, b[$. Si $\forall x \in]a, b[, \|f'(x)\|_F \leq g'(x)$, alors :

$$\|f(b) - f(a)\|_F \leq g(b) - g(a).$$

► Soit $\varepsilon > 0$. Notons

$$I_\varepsilon := \{y \in [a, b] \mid \forall x \in [a, y], \|f(x) - f(a)\|_F \leq g(x) - g(a) + \varepsilon(x - a) + \varepsilon\}.$$

* Nous allons montrer que $I_\varepsilon = [a, b]$. Nous aurons alors l'inégalité pour $x = b$, et ce, quelque soit $\varepsilon > 0$. La conclusion viendra en fixant $x = b$ et en faisant tendre ε vers 0.

* Notons $c_\varepsilon := \sup I_\varepsilon$. Montrons que $c_\varepsilon = b$. Tout d'abord, l'application

$$x \mapsto \varphi(x) := \|f(x) - f(a)\|_F - (g(x) - g(a) + \varepsilon(x - a)), \quad x \in [a, b],$$

est continue en a , puisque f et g le sont, et $\varphi(a) = 0$, donc pour tout $x > a$ suffisamment proche de a , $\varphi(x) \leq \varepsilon$, donc $c_\varepsilon > a$. Supposons maintenant que $a < c_\varepsilon < b$ et raisonnons par l'absurde. Puisque $c_\varepsilon \in]a, b[$, alors, par hypothèse, f et g sont dérivables en c_ε . Ainsi, il existe $y \in]c_\varepsilon, b[$, suffisamment proche de c_ε , tel que pour tout $x \in]c_\varepsilon, y]$, on a

$$\left\| \frac{f(x) - f(c_\varepsilon)}{x - c_\varepsilon} - f'(c_\varepsilon) \right\|_F \leq \frac{\varepsilon}{2} \quad \text{et} \quad \left| \frac{g(x) - g(c_\varepsilon)}{x - c_\varepsilon} - g'(c_\varepsilon) \right| \leq \frac{\varepsilon}{2},$$

d'où, puisque $x - c_\varepsilon > 0$:

$$\begin{aligned} \|f(x) - f(c_\varepsilon)\|_F &= \|f(x) - f(c_\varepsilon) - f'(c_\varepsilon)(x - c_\varepsilon) + f'(c_\varepsilon)(x - c_\varepsilon)\|_F \\ &\leq \|f(x) - f(c_\varepsilon) - f'(c_\varepsilon)(x - c_\varepsilon)\|_F + \|f'(c_\varepsilon)(x - c_\varepsilon)\|_F \\ &\leq \left(\frac{\varepsilon}{2} + \|f'(c_\varepsilon)\|_F\right)(x - c_\varepsilon). \end{aligned}$$

Ensuite, puisque nous avons $g'(c_\varepsilon) \geq 0$, alors, le théorème de Rolle nous donne $g(x) - g(c_\varepsilon) \geq 0$ et il vient :

$$\begin{aligned} g'(c_\varepsilon)(x - c_\varepsilon) &= |g(x) - g(c_\varepsilon) - g'(c_\varepsilon)(x - c_\varepsilon) - (g(x) - g(c_\varepsilon))| \\ &\leq \frac{\varepsilon}{2}(x - c_\varepsilon) + |g(x) - g(c_\varepsilon)| \\ &= \frac{\varepsilon}{2}(x - c_\varepsilon) + g(x) - g(c_\varepsilon). \end{aligned}$$

Par hypothèse, puisque $a < c_\varepsilon < b$, on a $\|f'(c_\varepsilon)\|_F \leq g'(c_\varepsilon)$. Ainsi, pour tout $x \in]c_\varepsilon, y]$:

$$\|f(x) - f(c_\varepsilon)\|_F \leq \varepsilon(x - c_\varepsilon) + g(x) - g(c_\varepsilon).$$

En outre, la continuité de f et g nous assure que $c_\varepsilon \in I_\varepsilon$. Et finalement, on obtient

$$\begin{aligned} \|f(x) - f(a)\|_F &= \|f(x) - f(c_\varepsilon) + f(c_\varepsilon) - f(a)\|_F \leq \|f(x) - f(c_\varepsilon)\|_F + \|f(c_\varepsilon) - f(a)\|_F \\ &\leq \varepsilon(x - c_\varepsilon) + g(x) - g(c_\varepsilon) + g(c_\varepsilon) - g(a) + \varepsilon(c_\varepsilon - a) + \varepsilon \\ &= g(x) - g(a) + \varepsilon(x - a) + \varepsilon. \end{aligned}$$

En prenant $x = y$ il vient que $y \in I_\varepsilon$, ce qui est impossible puisque $y > c_\varepsilon = \sup I_\varepsilon$. On a donc $c_\varepsilon = b$. En fixant $x = b$ et en faisant tendre ε vers 0, on obtient le résultat attendu. ■

Définition A.1.2

Soient (X, d_X) et (Y, d_Y) deux espaces métriques. Une application $f: X \rightarrow Y$ est dite ***k-lipschitzienne***, $k \geq 0$ dans \mathbb{R} , si pour tout x_1, x_2 dans X ,

$$d_Y(f(x_1), f(x_2)) \leq k d_X(x_1, x_2).$$

$f: X \rightarrow X$ est dite ***contractante*** si elle est k -lipschitzienne pour un $k \in [0, 1[$.

En prenant $g(x) = kx$, $k \in \mathbb{R}_+$, on obtient le corollaire suivant.

Corollaire A.1.3

Soit $f \in \mathcal{C}^0([a, b], F)$, f dérivable dans $]a, b[$ et telle que $\|f'(x)\|_F \leq k$ pour tout $x \in]a, b[$, alors f est k -lipschitzienne :

$$\|f(y) - f(x)\|_F \leq k |y - x|, \quad \forall x, y \in [a, b].$$

Lorsque $k = 0$, on obtient en particulier :

Corollaire A.1.4

Soit $f \in \mathcal{C}^0([a, b], F)$, admettant dans $]a, b[$ une dérivée nulle, alors f est constante.

Enfin, en prenant $f \equiv 0$ dans le Théorème A.1.1, on obtient :

Corollaire A.1.5

Soit $g \in \mathcal{C}^0([a, b], \mathbb{R})$, admettant dans $]a, b[$ une dérivée ≥ 0 , alors g est croissante.

Revenons au cas des fonctions définies sur un ouvert d'un espace vectoriel normé E . Comme pour les fonctions dérivables d'une variable réelle, le résultat fondamental du calcul différentiel est celui qui, à partir de la connaissance des différentielles d'une fonction en chaque point de son domaine, permet d'estimer l'accroissement de cette fonction entre deux points fixés à l'avance. On se ramène pour cela au cas des fonctions d'une variable, en considérant la restriction de la fonction au segment qui joint les deux points. Rappelons donc que pour une paire $(a, b) \in E^2$, on note

$$[a, b] := \nu([0, 1]) \quad \text{et} \quad]a, b[:= \nu(]0, 1]),$$

où

$$\begin{aligned} \nu: [0, 1] \subset \mathbb{R} &\longrightarrow E \\ t &\longmapsto \nu(t) := (1 - t)a + tb = a + t(b - a). \end{aligned}$$

Nous avons alors le résultat suivant.

Théorème A.1.6 – 2^e forme des accroissements finis

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soient U un ouvert de E , $(a, b) \in E^2$ tel que $[a, b] \subset U$ et $f: U \rightarrow F$. Si f est continue dans $[a, b]$ et différentiable dans $]a, b[$, alors :

$$\|f(b) - f(a)\|_F \leq \left(\sup_{x \in]a, b[} \|f'(x)\|_{\mathcal{L}(E, F)} \right) \|b - a\|_E \in \overline{\mathbb{R}}_+.$$

► Posons $M := \sup_{x \in]a, b[} \|f'(x)\|_{\mathcal{L}(E, F)}$. Si $M = +\infty$, alors le résultat trivialement vrai. Supposons donc $M < +\infty$. Posons sur l'intervalle $[0, 1]$, l'application $\varphi := f \circ \nu$, $\nu(t) = a + t(b - a)$, de telle sorte que $\varphi(0) = f(a)$ et $\varphi(1) = f(b)$. Introduisons de même sur $[0, 1]$, l'application $g(t) := tM\|b - a\|_E$. Ainsi définies, $\varphi \in \mathcal{C}^0([0, 1], F)$ par composition d'applications continues, $g \in \mathcal{C}^0([0, 1], \mathbb{R})$ et pour tout $t \in]0, 1[$:

$$\begin{aligned} \|\varphi'(t)\|_F &= \|f'(\nu(t)) \cdot \nu'(t)\|_F = \|f'(\nu(t)) \cdot (b - a)\|_F \\ &\leq \|f'(\nu(t))\|_{\mathcal{L}(E, F)} \|b - a\|_E \leq M\|b - a\|_E = g'(t). \end{aligned}$$

D'après le Théorème A.1.1, on obtient finalement

$$\|\varphi(1) - \varphi(0)\|_F = \|f(b) - f(a)\|_F \leq g(1) - g(0) = M\|b - a\|_E,$$

ce qui donne le résultat attendu. ■

Corollaire A.1.7

Soit $f: U \rightarrow F$ une fonction différentiable définie sur un ouvert **convexe** telle que $\|f'(x)\|_{\mathcal{L}(E, F)} \leq k$ pour tout $x \in U$. Alors, f est k -lipschitzienne :

$$\|f(y) - f(x)\|_F \leq k \|y - x\|_E, \quad \forall x, y \in U.$$

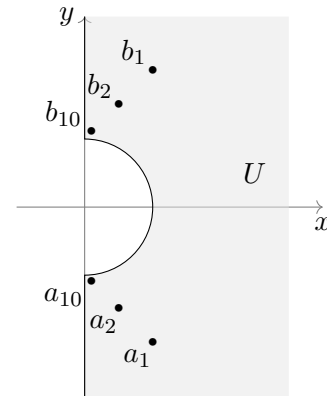
► Puisque U convexe, alors pour tout $(x, y) \in U^2$, le segment $[x, y] \subset U$ et on peut appliquer le Théorème A.1.6. ■

L'hypothèse de convexité de U dans ce dernier résultat ne peut être remplacé par la connexité de U . On a en effet l'exemple suivant.

Exemple A.1.2. Soient U l'ouvert de \mathbb{R}^2 défini par

$$U := \{(x, y) \mid x > 0 \text{ et } x^2 + y^2 > 1\}$$

et $\psi: U \rightarrow \mathbb{R}$ telle que $\psi(x, y) = \arctan(y/x)$. Alors U est un ouvert connexe, la différentielle de ψ vérifie $\|\psi'(x, y)\| < 1$ pour tout (x, y) dans U et ψ n'est pas k -lipschitzienne pour $k < \pi/2$. □



▷ **Correction.** Il est clair que U est connexe mais pas convexe. Sur U , l'application ψ est dérivable et on a :

$$\psi'(x, y) = \frac{1}{1 + \left(\frac{y}{x}\right)^2} \left(-\frac{y}{x^2}, \frac{1}{x}\right),$$

donc

$$\|\psi'(x, y)\| = \frac{x^2}{x^2 + y^2} \left(\frac{y^2}{x^4} + \frac{1}{x^2}\right)^{\frac{1}{2}} = \frac{x^2}{x^2 + y^2} \left(\frac{y^2 + x^2}{x^4}\right)^{\frac{1}{2}} = (x^2 + y^2)^{-\frac{1}{2}} < 1.$$

Montrons que ψ n'est pas k -lipschitzienne pour $k < \pi/2$. On définit dans U , les points :

$$a_n := \left(\frac{1}{n}, -\frac{1+n}{n}\right) \quad \text{et} \quad b_n := \left(\frac{1}{n}, \frac{1+n}{n}\right),$$

pour $n \in \mathbb{N}^*$. Alors, $\|b_n - a_n\| \rightarrow 2$ et

$$|\psi(b_n) - \psi(a_n)| = 2 \arctan(n+1) \rightarrow \pi,$$

donc ψ ne peut pas être k -lipschitzienne sur U pour $k < \pi/2$. □

Sur un ouvert connexe, nous avons le résultat suivant.

Corollaire A.1.8

*Soit $f: U \rightarrow F$ une fonction différentiable définie sur un ouvert **connexe**, admettant une dérivée nulle. Alors, f est constante.*

► Soit $x_0 \in U$. Posons

$$A := \{x \in U \mid f(x) = f(x_0)\}.$$

L'application f étant différentiable sur U , elle y est continue. L'ensemble A est donc fermé comme préimage d'un fermé par une application continue. Pour tout $x \in A$, le Corollaire A.1.7 montre que toute boule ouverte $B(x, r)$, $r > 0$, contenue dans U est aussi contenue dans A , puisque f est 0-lipschitzienne sur la boule $B(x, r)$ qui est convexe. Ainsi, l'ensemble A est ouvert dans U . Finalement, A est ouvert et fermé dans U , U connexe, donc $A = U$, et on obtient le résultat attendu. ■

Remarque A.1.1. Sur un ouvert formé de plusieurs composantes connexes, si la dérivée est nulle, alors l'application est constante sur chaque composante connexe, mais rien oblige à ce que les constantes soient égales.

A.2 Dérivabilité dans le cas fonctionnel

On s'intéresse ici à la dérivabilité de la composition à droite d'une application \mathcal{C}^1 avec une application continue définie sur un intervalle compact de \mathbb{R} . Concrètement, on considère deux espaces vectoriels normés $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$, et une application $f \in \mathcal{C}^1(E, F)$. On définit l'application

$$\begin{aligned} \Phi_f: \mathcal{C}^0([0, 1], E) &\longrightarrow \mathcal{C}^0([0, 1], F) \\ g &\longmapsto \Phi_f(g) := f \circ g, \end{aligned}$$

où $\mathcal{C}^0([0, 1], E)$ et $\mathcal{C}^0([0, 1], F)$ sont munis de la norme de la convergence uniforme $\|\cdot\|_\infty$,² et on s'intéresse à la dérivabilité de Φ_f . Si f est linéaire, i.e. $f \in \mathcal{L}(E, F)$, alors la réponse est plutôt simple : l'application Φ_f est linéaire (évident) et continue puisque

$$\begin{aligned} \forall t \in [0, 1], \quad \|\Phi_f(g)(t)\|_F &= \|f(g(t))\|_F \leq \|f\|_{\mathcal{L}(E, F)} \|g(t)\|_E \\ &\leq \|f\|_{\mathcal{L}(E, F)} \|g\|_\infty \end{aligned}$$

donc

$$\|\Phi_f(g)\|_\infty \leq \|f\|_{\mathcal{L}(E, F)} \|g\|_\infty.$$

L'application Φ_f est donc de classe \mathcal{C}^1 (toute application linéaire continue est \mathcal{C}^∞) et sa différentielle est donnée pour tout g et v dans $\mathcal{C}^0([0, 1], E)$, par :

$$\Phi'_f(g) \cdot v = \Phi_f(v) = f \circ v.$$

Dans le cas non linéaire, nous avons le résultat suivant.

Proposition A.2.1

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Soit $f \in \mathcal{C}^1(E, F)$. On définit l'application

$$\begin{aligned} \Phi_f: \mathcal{C}^0([0, 1], E) &\longrightarrow \mathcal{C}^0([0, 1], F) \\ g &\longmapsto \Phi_f(g) := f \circ g, \end{aligned}$$

où $\mathcal{C}^0([0, 1], E)$ et $\mathcal{C}^0([0, 1], F)$ sont munis de la norme de la convergence uniforme $\|\cdot\|_\infty$. Alors, l'application Φ_f est de classe \mathcal{C}^1 et pour tout g, v dans $\mathcal{C}^0([0, 1], E)$, la différentielle est donnée par :

$$\begin{aligned} \Phi'_f(g) \cdot v: [0, 1] &\longrightarrow F \\ t &\longmapsto (\Phi'_f(g) \cdot v)(t) = f'(g(t)) \cdot v(t). \end{aligned}$$

Remarque A.2.1. On a donc $\Phi_f(g)(t) = f(g(t))$ et sa différentielle est donnée par $(\Phi'_f(g) \cdot v)(t) = f'(g(t)) \cdot v(t) = (f' \circ g)(t) \cdot v(t)$ que l'on peut donc noter

$$\Phi'_f(g) \cdot v = (f' \circ g) \cdot v.$$

Φ_f est un opérateur de Nemytskii (ou de superposition), voir [4] pour plus détails.

Nous aurons besoin du lemme suivant.

Lemme A.2.1. Soient $F: (X, d_X) \rightarrow (Y, d_Y)$ une application continue entre deux espaces métriques et $K \subset X$ un sous-ensemble compact. On pose pour $\delta \in \mathbb{R}_+$:

$$M(\delta) := \sup \{d_Y(F(x_1), F(x_2)) \mid x_1 \in K, x_2 \in X, d_X(x_1, x_2) \leq \delta\}.$$

Alors,

$$M(\delta) = o(1) \text{ quand } \delta \rightarrow 0.$$

2. Il y a ici un léger abus de notation (sans conséquence) car dans $\mathcal{C}^0([0, 1], E)$, la norme sup est donnée par $\|g\|_\infty = \sup_{t \in [0, 1]} (\|g(t)\|_E)$, tandis que dans $\mathcal{C}^0([0, 1], F)$, on a $\|h\|_\infty = \sup_{t \in [0, 1]} (\|h(t)\|_F)$.

► Soit $\varepsilon > 0$. Si la conclusion est fausse, alors il existe deux suites $(x_{1,n})$ dans K et $(x_{2,n})$ dans X telles que $d_X(x_{1,n}, x_{2,n}) \rightarrow 0$ et $d_Y(F(x_{1,n}), F(x_{2,n})) > \varepsilon$ pour tout n . Quitte à extraire une sous-suite, on peut supposer que $x_{1,n} \rightarrow x \in K$. Alors, $x_{2,n} \rightarrow x$ et puisque F est continue sur X , $F(x_{1,n})$ et $F(x_{2,n})$ converge vers $F(x)$ d'où la contradiction. ■

Nous pouvons maintenant démontrer le théorème.

► Soit $g \in \mathcal{C}^0([0, 1], E)$. On introduit dans $\mathcal{C}^0([0, 1], E)$, l'application

$$h(v) := \Phi_f(g + v) - \Phi_f(g) - (f' \circ g) \cdot v = f \circ (g + v) - f \circ g - (f' \circ g) \cdot v.$$

* L'idée est donc de montrer dans un premier temps que $h(v) = o(v)$. Soit $v \in \mathcal{C}^0([0, 1], E)$. On introduit l'application

$$\begin{aligned} \varphi: [0, 1]^2 &\longrightarrow F \\ (t, s) &\longmapsto \varphi(t, s) := h(sv)(t) = f(g(t) + sv(t)) - f(g(t)) - s f'(g(t)) \cdot v(t) \end{aligned}$$

de telle sorte que $\|h(v)\|_\infty = \sup_{t \in [0, 1]} (\|\varphi(t, 1)\|_F)$. Pour $t \in [0, 1]$ fixé, l'application partielle $s \mapsto \varphi(t, s)$ est continue sur $[0, 1]$, dérivable sur $]0, 1[$, de dérivée

$$\frac{\partial \varphi}{\partial s}(t, s) = (f'(g(t) + sv(t)) - f'(g(t))) \cdot v(t),$$

donc d'après le Théorème A.1.6 des accroissements finis :

$$\begin{aligned} \|\varphi(t, 1) - \varphi(t, 0)\|_F &= \|\varphi(t, 1)\|_F \leq \sup_{s \in [0, 1]} (\|(f'(g(t) + sv(t)) - f'(g(t))) \cdot v(t)\|_F) \\ &\leq \sup_{s \in [0, 1]} (\|(f'(g(t) + sv(t)) - f'(g(t)))\|_{\mathcal{L}(E, F)}) \|v\|_\infty. \end{aligned}$$

Ainsi,

$$\|h(v)\|_\infty \leq \sup_{(t, s) \in [0, 1]^2} (\|(f'(g(t) + sv(t)) - f'(g(t)))\|_{\mathcal{L}(E, F)}) \|v\|_\infty \leq M(\|v\|_\infty) \|v\|_\infty,$$

où l'on a introduit pour $\delta \in \mathbb{R}_+$:

$$M(\delta) := \sup \{ \|f'(y) - f'(x)\|_{\mathcal{L}(E, F)} \mid x \in g([0, 1]), y \in E, \|y - x\|_E \leq \delta \}. \quad (\text{A.1})$$

On applique le lemme précédent avec $K := g([0, 1])$ compact de E (car g continue) et $F := f'$ continue (car f de classe \mathcal{C}^1). Ainsi, $M(\|v\|_\infty) = o(1)$ quand $\|v\|_\infty \rightarrow 0$, donc $M(\|v\|_\infty) \|v\|_\infty = o(\|v\|_\infty)$ et puisque $\|h(v)\|_\infty \leq M(\|v\|_\infty) \|v\|_\infty$, finalement on obtient bien que $h(v) = o(v)$.

* Montrons que $v \mapsto (f' \circ g) \cdot v$ est une application linéaire continue, car alors, on aura que Φ_f est dérivable en g de différentielle donnée par $\Phi'_f(g) \cdot v = (f' \circ g) \cdot v$. Il est clair que $v \mapsto (f' \circ g) \cdot v$ est linéaire. Elle est continue car, pour tout $v \in \mathcal{C}^0([0, 1], E)$, on a

$$\forall t \in [0, 1], \|(f' \circ g) \cdot v(t)\|_F \leq \|f'(g(t))\|_{\mathcal{L}(E, F)} \|v(t)\|_E \leq \|f' \circ g\|_\infty \|v\|_\infty$$

donc $\|(f' \circ g) \cdot v\|_\infty \leq \|f' \circ g\|_\infty \|v\|_\infty$. On a utilisé le fait que $f' \circ g$ est continue sur le compact $[0, 1]$ donc $\|f' \circ g\|_\infty < \infty$. Finalement, on a bien $\Phi'_f(g) \cdot v = (f' \circ g) \cdot v$.

* Montrons enfin que Φ_f est de classe \mathcal{C}^1 , i.e. que Φ'_f est continue sur $\mathcal{C}_E^0 := \mathcal{C}^0([0, 1], E)$. Soit $g \in \mathcal{C}_E^0$. Montrons que $\|\Phi'_f(g + v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = o(1)$ quand $\|v\|_\infty \rightarrow 0$, où l'on a noté $\mathcal{C}_F^0 := \mathcal{C}^0([0, 1], F)$. Puisque

$$\|\Phi'_f(g + v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = \sup_{\|u\|_\infty=1} \|(\Phi'_f(g + v) - \Phi'_f(g)) \cdot u\|_\infty,$$

et puisque

$$\begin{aligned} \|(\Phi'_f(g + v) - \Phi'_f(g)) \cdot u\|_\infty &= \sup_{t \in [0, 1]} \|(f'(g(t) + v(t)) - f'(g(t))) \cdot u(t)\|_F \\ &\leq \sup_{t \in [0, 1]} \|f'(g(t) + v(t)) - f'(g(t))\|_{\mathcal{L}(E, F)}, \end{aligned}$$

on obtient donc

$$\|\Phi'_f(g + v) - \Phi'_f(g)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \sup_{t \in [0, 1]} \|f'(g(t) + v(t)) - f'(g(t))\|_{\mathcal{L}(E, F)} \leq M(\|v\|_\infty),$$

avec M définie comme en (A.1) et puisque $M(\|v\|_\infty) = o(1)$ quand $\|v\|_\infty \rightarrow 0$ d'après le lemme précédent, on obtient le résultat voulu. ■

Remarque A.2.2. Montrons d'une autre manière que Φ_f est de classe \mathcal{C}^1 , i.e. que Φ'_f est continue sur $\mathcal{C}_E^0 := \mathcal{C}^0([0, 1], E)$. Soient $a \in \mathcal{C}_E^0$ et $\varepsilon > 0$. Il faut montrer qu'il existe $\eta_{\varepsilon, a} > 0$ tel que pour tout $g \in \mathcal{C}_E^0$:

$$\|g - a\|_\infty \leq \eta_{\varepsilon, a} \implies \|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \varepsilon,$$

où $\mathcal{C}_F^0 := \mathcal{C}^0([0, 1], F)$. Or,

$$\|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} = \sup_{\|v\|_\infty=1} \|(\Phi'_f(g) - \Phi'_f(a)) \cdot v\|_\infty,$$

et

$$\begin{aligned} \|(\Phi'_f(g) - \Phi'_f(a)) \cdot v\|_\infty &= \sup_{t \in [0, 1]} \|(f'(g(t)) - f'(a(t))) \cdot v(t)\|_F \\ &\leq \sup_{t \in [0, 1]} \|f'(g(t)) - f'(a(t))\|_{\mathcal{L}(E, F)}, \end{aligned}$$

donc

$$\|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \sup_{t \in [0, 1]} \|f'(g(t)) - f'(a(t))\|_{\mathcal{L}(E, F)}.$$

Mais par hypothèse, $f \in \mathcal{C}^1(E, F)$ donc $f' \in \mathcal{C}^0(E, \mathcal{L}(E, F))$, donc f' est uniformément continue sur le compact $K_a := a([0, 1]) \subset E$ et ainsi pour le $\varepsilon > 0$ choisi : $\exists \eta_{\varepsilon, K_a} > 0$, tel que $\forall (x, y) \in K_a \times E$:

$$\|y - x\|_E \leq \eta_{\varepsilon, K_a} \implies \|f'(y) - f'(x)\|_F \leq \varepsilon.$$

Alors en prenant $\eta_{\varepsilon, a} := \eta_{\varepsilon, K_a}$, on voit que si $\|g - a\|_\infty \leq \eta_{\varepsilon, a}$, alors pour tout $t \in [0, 1]$, $\|g(t) - a(t)\|_E \leq \eta_{\varepsilon, K_a}$ donc $\|f'(g(t)) - f'(a(t))\|_F \leq \varepsilon$. D'où :

$$\|g - a\|_\infty \leq \eta_{\varepsilon, a} \implies \|\Phi'_f(g) - \Phi'_f(a)\|_{\mathcal{L}(\mathcal{C}_E^0, \mathcal{C}_F^0)} \leq \varepsilon.$$

A.3 Théorème du point fixe

Théorème A.3.1 – du point fixe de Picard

Soit $f: X \rightarrow X$ une application contractante sur un espace métrique complet³ (X, d) non vide. Il existe alors un unique point fixe $x \in X$ de f , c'est-à-dire tel que $f(x) = x$.

► On suppose f contractante de constante de Lipschitz $k \in [0, 1[$.

* (Unicité). Soit $(x_1, x_2) \in X^2$ t.q. $f(x_1) = x_1$ et $f(x_2) = x_2$. Alors

$$d(f(x_1), f(x_2)) = d(x_1, x_2) \leq k d(x_1, x_2) < d(x_1, x_2),$$

ce qui n'est possible que pour $x_1 = x_2$. On a donc l'unicité.

* (Existence). On fixe $x_0 \in X$ (X est non vide). On définit pour $n \in \mathbb{N}$, $x_{n+1} := f(x_n)$. Montrons que la suite (x_n) est de Cauchy. Nous avons pour $n \in \mathbb{N}$:

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq k d(x_n, x_{n-1}) \leq k^n d(x_1, x_0) =: C k^n.$$

Ainsi, pour $(n, p) \in \mathbb{N}^2$:

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{i=0}^{p-1} d(x_{n+i+1}, x_{n+i}) && \text{(d'après l'inégalité triangulaire)} \\ &\leq C \sum_{i=0}^{p-1} k^{n+i} = C k^n \frac{1 - k^p}{1 - k} \leq \frac{C}{1 - k} k^n. && \text{(car } k \in [0, 1[) \end{aligned}$$

Soit $\varepsilon > 0$. On choisit $N_\varepsilon \in \mathbb{N}$ t.q.

$$\frac{C}{1 - k} k^{N_\varepsilon} \leq \varepsilon,$$

alors pour tous $n \geq N_\varepsilon$ et $p \in \mathbb{N}$, on a

$$d(x_{n+p}, x_n) \leq \frac{C}{1 - k} k^n \leq \frac{C}{1 - k} k^{N_\varepsilon} \leq \varepsilon,$$

c'est-à-dire (x_n) est une suite de Cauchy. Puisque X est complet, la suite (x_n) converge vers un élément de X noté x . Posons $u_n := f(x_n) - x_n$. La suite (u_n) converge vers $f(x) - x$ (car f continue). Mais puisque $u_n = f(x_n) - x_n = f(x_n) - f(x_{n-1})$ (pour $n \geq 1$), la suite (u_n) converge vers $f(x) - f(x) = 0$. Par unicité de la limite, on obtient $f(x) = x$, c'est-à-dire l'existence d'un point fixe. ■

3. Soit (X, d) un espace métrique. Il est complet si toute suite de Cauchy converge. Une suite (x_n) de X vérifiant $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n, m \geq N, d(x_n, x_m) \leq \varepsilon$, est dite de Cauchy.

Théorème A.3.2 – du point fixe de Picard à paramètre

Soient (X, d) un espace métrique complet, Λ un espace topologique, $\varphi: X \times \Lambda \rightarrow X$ une application continue et $0 < k < 1$. Supposons que pour tout $\lambda \in \Lambda$, l'application $\varphi(\cdot, \lambda)$ soit k -contractante. Alors, pour tout $\lambda \in \Lambda$, il existe un unique point fixe $x(\lambda) \in X$ de $\varphi(\cdot, \lambda)$ et l'application $x(\cdot): \Lambda \rightarrow X$ est continue.

► On définit $(\mathcal{C}_b^0(\Lambda, X), d_\infty)$, l'ensemble des applications continues et bornées de Λ dans X muni de la distance de la convergence uniforme : $d_\infty(f, g) = \sup_{\lambda \in \Lambda} d(f(\lambda), g(\lambda))$. Puisque (X, d) est complet, $(\mathcal{C}_b^0(\Lambda, X), d_\infty)$ l'est aussi. On définit

$$\begin{aligned} \Phi: (\mathcal{C}_b^0(\Lambda, X), d_\infty) &\longrightarrow (\mathcal{C}_b^0(\Lambda, X), d_\infty) \\ f &\longmapsto \Phi(f) := \varphi(f(\cdot), \cdot). \end{aligned}$$

Ainsi,

$$d_\infty(\Phi(f), \Phi(g)) = \sup_{\lambda \in \Lambda} d(\Phi(f)(\lambda), \Phi(g)(\lambda)) = \sup_{\lambda \in \Lambda} d(\varphi(f(\lambda), \lambda), \varphi(g(\lambda), \lambda)),$$

or

$$d(\varphi(f(\lambda), \lambda), \varphi(g(\lambda), \lambda)) \leq k d(f(\lambda), g(\lambda)) \leq k d_\infty(f, g),$$

donc

$$d_\infty(\Phi(f), \Phi(g)) \leq k d_\infty(f, g),$$

c'est-à-dire, Φ est k -contractante. D'après le théorème du point fixe de Picard A.3.1, Φ admet un unique point fixe noté x . Pour tout $\lambda \in \Lambda$, $\Phi(x)(\lambda) = \varphi(x(\lambda), \lambda) = x(\lambda)$. Puisque pour tout $\lambda \in \Lambda$, $\varphi(\cdot, \lambda)$ est k -contractante, elle admet un unique point fixe $x_\lambda \in X$, qui vérifie donc $x_\lambda = \varphi(x_\lambda, \lambda) = x(\lambda)$. Le théorème est démontré. ■

A.4 Application inversion

Rappelons quelques définitions.

Définition A.4.1 – Espace de Banach

Un **espace de Banach** est un espace vectoriel normé $(E, \|\cdot\|_E)$ qui est complet pour la distance $d_E(x, y) := \|x - y\|_E$.

Définition A.4.2

Soient E, F deux espaces vectoriels topologiques (par exemple, normés).

- Un **morphisme** (topologique) entre E et F est une application linéaire continue de E dans F . On note $\mathcal{L}(E, F)$ ou $L_c(E, F)$ l'ensemble des applications linéaires continues de E dans F .
- Un **endomorphisme** (topologique) de E est une application linéaire continue de E dans E . On note $\mathcal{L}(E) := \mathcal{L}(E, E)$ ou $L_c(E)$ l'ensemble des endomorphismes de E .

- Un **isomorphisme** (topologique) entre E et F est une application de E dans F , linéaire, bijective, continue, dont l'application réciproque (ou inverse) est continue. C'est donc un **homéomorphisme** (*i.e.* une application bijective continue d'inverse continue) linéaire. On note $\text{Isom}_c(E, F)$ l'ensemble des isomorphismes de E sur F .
- Un **automorphisme** (topologique) de E est une application dans $\text{GL}_c(E) := \text{Isom}_c(E, E)$. L'ensemble des automorphismes peut être aussi noté $\mathcal{GL}(E)$.
- Si $F = \mathbb{R}$, on parle de **forme linéaire continue**. On note $E' := \mathcal{L}(E, \mathbb{R})$, l'ensemble des formes linéaires continue de E sur \mathbb{R} . Cet espace est appelé **l'espace dual topologique** de E .

On considère deux espaces de Banach E et F et on suppose que $\text{Isom}_c(E, F)$ est non vide. D'après [75, Théorème 3.19.8], l'ensemble $\text{Isom}_c(E, F)$ est ouvert dans $\mathcal{L}(E, F)$. Introduisons l'*application inversion* :

$$\begin{aligned} \text{inv} : \text{Isom}_c(E, F) &\longrightarrow \text{Isom}_c(F, E) \\ T &\longmapsto \text{inv}(T) := T^{-1}. \end{aligned}$$

Toujours d'après le même résultat, l'application inv est un homéomorphisme de $\text{Isom}_c(E, F)$ dans $\text{Isom}_c(F, E)$. Montrons que inv est de classe \mathcal{C}^1 .

* Nous nous ramenons dans un premier temps à $\text{Isom}_c(E, E) = \text{GL}_c(E)$. Pour cela, considérons $B \in \text{Isom}_c(F, E)$ et introduisons l'isomorphisme (à vérifier) :

$$\begin{aligned} \varphi : \text{Isom}_c(E, F) &\longrightarrow \text{GL}_c(E) \\ T &\longmapsto \varphi(T) := B \circ T. \end{aligned}$$

Ajoutons la notation $\text{inv}_{E,F}$ pour représenter l'application inversion sur $\text{Isom}_c(E, F)$. Avec cette notation, $\text{inv} = \text{inv}_{E,F}$ et l'application inversion sur $\text{GL}_c(E)$ s'écrit $\text{inv}_E := \text{inv}_{E,E}$. Il apparaît alors que :

$$\text{inv}(T) = \varphi(T)^{-1} \circ B = \text{inv}_E(\varphi(T)) \circ B. \quad (\text{A.2})$$

Ainsi, par le théorème de dérivation des applications composées, il est clair que inv est de classe \mathcal{C}^1 si et seulement si inv_E l'est. En effet, φ est linéaire continue donc \mathcal{C}^1 tout comme l'application de composition à gauche $A \mapsto A \circ B$, avec $A \in \text{GL}_c(E)$.

* Montrons donc que inv_E est de classe \mathcal{C}^1 . Considérons T, H dans $\text{GL}_c(E)$ et remarquons tout d'abord que

$$T + H = T \circ (\text{Id}_E + \text{inv}_E(T) \circ H) =: T(I + T^{-1}H),$$

où dans le dernier terme de ces égalités, nous utilisons les notations T^{-1} et I qui sont plus concises que $\text{inv}_E(T)$ et Id_E et où nous ne notons pas le symbole de composition pour avoir une notation semblable à l'écriture matricielle. D'après [75, Proposition 3.19.6], si $\|T^{-1}H\|_{\mathcal{L}(E)} < 1$, alors $I - T^{-1}H$ est inversible et

$$K := (I - T^{-1}H)^{-1} = I - T^{-1}H + (T^{-1}H)^2 + \cdots + (-1)^n (T^{-1}H)^n + \cdots.$$

Ainsi, avec ces notations : $T + H = TK^{-1}$. En combinant cela avec le fait que

$$K = I - T^{-1}H + (T^{-1}H)^2K,$$

nous obtenons, pour H suffisamment petit,

$$\begin{aligned} (T + H)^{-1} &= KT^{-1} = (I - T^{-1}H + (T^{-1}H)^2K)T^{-1} \\ &= T - T^{-1}HT^{-1} + (T^{-1}H)^2KT^{-1} = T - T^{-1}HT^{-1} + o(H). \end{aligned}$$

Autrement dit,

$$\text{inv}_E(T + H) = \text{inv}_E(T) - \text{inv}_E(T) \circ H \circ \text{inv}_E(T) + o(H).$$

Puisque $H \mapsto -\text{inv}_E(T) \circ H \circ \text{inv}_E(T)$ est linéaire continue (de norme inférieure à $2\|T^{-1}\|_{\mathcal{L}(E)}$), on a inv_E dérivable en T , donc sur tout $\text{GL}_c(E)$ (car T quelconque). Sa dérivée est donnée par

$$\text{inv}'_E(T) \cdot H = -\text{inv}_E(T) \circ H \circ \text{inv}_E(T).$$

* Montrons maintenant que $\text{inv}'_E: \text{GL}_c(E) \rightarrow \mathcal{L}(\text{GL}_c(E))$ est continue. Définissons pour cela $B: \text{GL}_c(E) \times \text{GL}_c(E) \rightarrow \mathcal{L}(\text{GL}_c(E))$ par $B(M, N) \cdot H := -M \circ H \circ N$. Ainsi définie, l'application B est une application bilinéaire sur $\text{GL}_c(E) \times \text{GL}_c(E)$. On munit cet espace de la norme

$$\|B\| := \sup \{ \|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} \mid \|M\|_{\mathcal{L}(E)} \leq 1 \text{ et } \|N\|_{\mathcal{L}(E)} \leq 1 \}.$$

À l'aide de B , on peut écrire $\text{inv}'_E(T) = B(\text{inv}_E(T), \text{inv}_E(T))$ et il est alors clair que inv'_E est continue si B l'est. Puisque

$$\|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} = \sup_{\|H\|_{\mathcal{L}(E)} \leq 1} \|B(M, N) \cdot H\|_{\mathcal{L}(E)} = \sup_{\|H\|_{\mathcal{L}(E)} \leq 1} \|M \circ H \circ N\|_{\mathcal{L}(E)},$$

et puisque

$$\|M \circ H \circ N\|_{\mathcal{L}(E)} \leq \|M\|_{\mathcal{L}(E)} \|H\|_{\mathcal{L}(E)} \|N\|_{\mathcal{L}(E)},$$

on obtient $\|B(M, N)\|_{\mathcal{L}(\text{GL}_c(E))} \leq \|M\|_{\mathcal{L}(E)} \|N\|_{\mathcal{L}(E)}$ et donc finalement $\|B\| \leq 1$. Nous avons donc montré que B est continue donc inv'_E aussi, donc inv_E est de classe \mathcal{C}^1 . Nous pouvons alors conclure que $\text{inv} = \text{inv}_{E,F}$ est elle-même de classe \mathcal{C}^1 .

* Maintenant que nous savons que inv est dérivable (même \mathcal{C}^1), nous pouvons retrouver facilement sa dérivée. En effet, sachant que l'on a pour tout $T \in \text{Isom}_c(E, F)$, la relation $\Psi(T) := \text{inv}(T) \circ T = \text{Id}_E$, on obtient

$$\Psi'(T) \cdot H = \text{inv}'(T) \cdot H \circ T + \text{inv}(T) \circ H = 0_{\mathcal{L}(E)},$$

que l'on peut transformer en : $\text{inv}'(T) \cdot H = -\text{inv}(T) \circ H \circ \text{inv}(T)$. Il est à noter que l'on retrouve la même expression à partir de la relation (A.2).

Remarque A.4.1. Si $E = F = \mathbb{R}^n$, alors inv est définie sur $\text{GL}_n(\mathbb{R})$, l'ensemble des matrices inversibles. On a la relation plus familière, pour $A \in \text{GL}_n(\mathbb{R})$,

$$\text{inv}'(A) \cdot H = -A^{-1}HA^{-1}.$$

* Enfin, montrons que inv est \mathcal{C}^∞ . Nous pouvons écrire,

$$\text{inv}'_E = B \circ \varphi \circ \text{inv}_E$$

avec $\varphi(A, B) := (A, B)$ linéaire continue donc \mathcal{C}^∞ et B bilinéaire continue donc \mathcal{C}^∞ . Ainsi, par induction, inv_E est \mathcal{C}^∞ tout comme inv .

A.5 Théorème d'inversion locale

Nous aurons besoin du résultat suivant, voir [75, Corollaire 3.11.3].

Théorème A.5.1 – de Banach

Soient E et F deux espaces de Banach, alors, toute bijection linéaire et continue de E sur F est un isomorphisme (topologique).

Définition A.5.2 – Difféomorphisme, \mathcal{C}^k -difféomorphisme

Soient E, F deux espaces de Banach et U, V deux ouverts respectivement de E, F . Soit une application $f: U \rightarrow V$.

- i) On dit que f est un difféomorphisme de U dans V si f est une bijection de U dans V , dérivable dans U et dont l'inverse $f^{-1}: V \rightarrow U$ est dérivable dans V .
- ii) On dit que f est un \mathcal{C}^k -difféomorphisme, $k \in \overline{\mathbb{N}}^*$, de U dans V si f est une bijection de U dans V , de classe \mathcal{C}^k et dont l'inverse $f^{-1}: V \rightarrow U$ est de classe \mathcal{C}^k .

Un difféomorphisme et *a fortiori* un \mathcal{C}^k -difféomorphisme, sont des homéomorphismes.

Remarque A.5.1. En fait, on peut voir un homéomorphisme, un difféomorphisme ou un \mathcal{C}^k -difféomorphisme comme une application inversible, respectivement, dans l'ensemble des applications continues, dérivables ou de classe \mathcal{C}^k .

Proposition A.5.3

Un homéomorphisme $f: U \rightarrow V$ est un \mathcal{C}^k -difféomorphisme, $k \in \overline{\mathbb{N}}^*$, de U dans V si et seulement si f est de classe \mathcal{C}^k et, pour tout $x \in U$, $f'(x) \in \mathcal{L}(E, F)$ est bijective.

► Soit un homéomorphisme $f: U \rightarrow V$.

1. Supposons que f soit un \mathcal{C}^k -difféomorphisme de U dans V . En particulier, f est \mathcal{C}^k . En particulier aussi, f est dérivable sur U et son inverse $f^{-1}: V \rightarrow U$ est elle dérivable sur V . Montrons que $f'(x)$ est bijective. Par hypothèse,

$$f^{-1} \circ f = \text{Id}_U \quad \text{et} \quad f \circ f^{-1} = \text{Id}_V,$$

donc pour tout $x \in U$, en définissant $y := f(x) \in V$, on obtient par dérivation :

$$(f^{-1} \circ f)'(x) = (f^{-1})'(y) \circ f'(x) = \text{Id}_E \quad \text{et} \quad (f \circ f^{-1})'(y) = f'(x) \circ (f^{-1})'(y) = \text{Id}_F,$$

autrement dit $f'(x)$ est bijective avec $(f'(x))^{-1} = (f^{-1})'(y)$.

2. Supposons maintenant que f soit de classe \mathcal{C}^k et que pour tout $x \in U$, $f'(x)$ soit bijective. Nous voulons montrer que f est un \mathcal{C}^k -difféomorphisme, autrement dit que f est \mathcal{C}^k (ceci est vrai par hypothèse) et que f est une bijection (vraie par hypothèse) dont l'inverse $f^{-1}: V \rightarrow U$ est de classe \mathcal{C}^k .

2.a. Montrons dans un premier temps que $g := f^{-1}$ est dérivable sur V . D'après ce qui précède, si g est dérivable en $y \in V$, alors nécessairement $g'(y)$ serait donnée par $(f'(x))^{-1}$ où l'on a noté $x := f^{-1}(y) \in U$. Remarquons tout de suite que, d'après le Théorème A.5.1 de Banach, $f'(x) \in \text{Isom}_c(E, F)$, autrement dit $(f'(x))^{-1} \in \mathcal{L}(F, E)$. Il ne nous reste plus qu'à montrer que pour $y \in V$ donné :

$$\varphi(h) := g(y + h) - g(y) - T_y(h) = o(h),$$

où l'on a introduit la notation $T_y := (f'(x))^{-1} = (f'(f^{-1}(y)))^{-1}$ et où φ est définie pour h suffisamment petit.

* Posons $x_h := g(y + h) = f^{-1}(f(x) + h)$. Alors, $h = f(x_h) - f(x)$. Ainsi,

$$h = f(x_h) - f(x) = f'(x) \cdot (x_h - x) + o(x_h - x).$$

Introduisons la notation $\psi(x_h - x) := f(x_h) - f(x) - f'(x) \cdot (x_h - x)$. Puisque T_y est l'inverse de $f'(x)$, on obtient

$$T_y(h) = x_h - x + T_y(\psi(x_h - x)) = g(y + h) - g(y) + T_y(\psi(x_h - x)). \quad (\text{A.3})$$

Supposons que l'on ait montré que $\psi(x_h - x)$, comme fonction de h , est un $o(h)$, alors on peut écrire $\psi(x_h - x) =: \|h\|_F \varepsilon(h)$ avec $\|\varepsilon(h)\|_F \rightarrow 0$ quand $\|h\|_F \rightarrow 0$. On obtiendrait finalement que

$$\|\varphi(h)\|_E = \|T_y(\psi(x_h - x))\|_E \leq \|T_y\|_{\mathcal{L}(F, E)} \|\varepsilon(h)\|_F \|h\|_F = o(h),$$

autrement dit que $\varphi(h) = o(h)$, c'est-à-dire ce que l'on voulait démontrer dans un premier temps, au point 2.a.

* Montrons donc que $\psi(x_h - x) = o(h)$. On sait déjà que $\psi(x_h - x) = o(x_h - x)$. On peut donc écrire $\psi(x_h - x) =: \|x_h - x\|_E \varepsilon(x_h - x)$, avec $\|\varepsilon(x_h - x)\|_F \rightarrow 0$ quand $\|x_h - x\|_E \rightarrow 0$. Puisque $g = f^{-1}$ est continue, on a $x_h - x = g(y + h) - g(y) = o(1)$ quand $\|h\|_F \rightarrow 0$. Supposons que l'on ait montré que $x_h - x = O(h)$,⁴ alors il existe $\eta > 0$ et $C > 0$, tels que pour tout $\|h\|_F \leq \eta$:

$$\|\psi(x_h - x)\|_F = \|x_h - x\|_E \|\varepsilon(x_h - x)\|_F \leq C \|\varepsilon(x_h - x)\|_F \|h\|_F = o(h),$$

car $\|\varepsilon(x_h - x)\|_F \rightarrow 0$ quand $\|h\|_F \rightarrow 0$ (car $\|x_h - x\|_E \rightarrow 0$ quand $\|h\|_F \rightarrow 0$). On aurait donc $\psi(x_h - x) = o(h)$.

* Pour finir de prouver que $\varphi(h) = o(h)$, il nous reste à montrer que $x_h - x = O(h)$. Puisque $0 < \|T_y\|_{\mathcal{L}(F, E)}^{-1} < +\infty$ ($T_y \in \mathcal{L}(F, E)$ et $T_y \neq 0$) et puisque $\psi(x_h - x) = o(x_h - x)$, on a pour h suffisamment petit :

$$\|\psi(x_h - x)\|_F \leq \frac{1}{2} \|T_y\|_{\mathcal{L}(F, E)}^{-1} \|x_h - x\|_E.$$

Ainsi, d'après (A.3), on a :

$$\begin{aligned}\|x_h - x\|_E &\leq \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F + \|T_y\|_{\mathcal{L}(F,E)} \|\psi(x_h - x)\|_F \\ &\leq \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F + \frac{1}{2} \|x_h - x\|_E,\end{aligned}$$

ce qui nous donne $\|x_h - x\|_E \leq 2 \|T_y\|_{\mathcal{L}(F,E)} \|h\|_F$, autrement dit $x_h - x = O(h)$. En conclusion, on a bien $\psi(x_h - x) = o(h)$ ce qui entraîne $\varphi(h) = o(h)$. Le point 2.a est démontré : $g = f^{-1}$ est dérivable sur V .

2.b. L'application réciproque $g = f^{-1}$ admet donc une application dérivée (cf. 2.a) donnée par $g' : y \mapsto g'(y) = (f'(g(y)))^{-1}$. Introduisons

$$\begin{aligned}\text{inv} : \text{Isom}_c(E, F) &\longrightarrow \text{Isom}_c(F, E) \\ T &\longmapsto \text{inv}(T) := T^{-1},\end{aligned}$$

de telle sorte que $g' = \text{inv} \circ f' \circ g$. Sachant que inv est \mathcal{C}^∞ , cf. A.4, que f' est \mathcal{C}^{k-1} et que g est continue, on obtient que g' est continue donc g de classe \mathcal{C}^1 , donc finalement g' de classe \mathcal{C}^1 , donc g de classe \mathcal{C}^2 et ainsi de suite jusqu'à obtenir g de classe \mathcal{C}^k . Nous avons finalement démontré que f est un \mathcal{C}^k -difféomorphisme. Le théorème est démontré. ■

Remarque A.5.2. Rappelons que l'ensemble $\text{Isom}_c(E, F)$ est ouvert dans $\mathcal{L}(E, F)$ d'après [75, Théorème 3.19.8 page 400]. L'application inv est même un homéomorphisme de $\text{Isom}_c(E, F)$ dans $\text{Isom}_c(F, E)$. Et puisque pour tous T et H dans $\text{Isom}_c(E, F)$, on a $\text{inv}'(T) \cdot H = -\text{inv}(T) \circ H \circ \text{inv}(T)$, on peut montrer que $\text{inv}'(T) \in \mathcal{L}(\text{Isom}_c(E, F), \text{Isom}_c(F, E))$ est bijective. Donc, d'après la Proposition A.5.3, l'application inv est un \mathcal{C}^∞ -difféomorphisme.

Théorème A.5.4 – d'inversion locale

Soient E, F deux espaces de Banach et $f : E \rightarrow F$ une application de classe \mathcal{C}^k , $k \in \mathbb{N}^*$, sur un voisinage de $x \in E$. On suppose que $f'(x) \in \mathcal{L}(E, F)$ est bijective.⁵ Il existe alors un ouvert $U \subset E$ contenant x et un ouvert $V \subset F$ contenant $f(x)$, tels que f est un \mathcal{C}^k -difféomorphisme de U dans $V = f(U)$.

► L'idée est de montrer que f est un homéomorphisme de classe \mathcal{C}^k de U sur V t.q. $\forall x \in U$, $f'(x) \in \mathcal{L}(E, F)$ est bijective. On appliquera ensuite la proposition précédente.

* Notons $y := f(x) \in F$. Introduisons

$$f_0(u) := f'(x)^{-1} \cdot (f(x + u) - y),$$

de telle sorte que f est un \mathcal{C}^k -difféomorphisme d'un voisinage de x sur un voisinage de y si et seulement si f_0 est un \mathcal{C}^k -difféomorphisme d'un voisinage de 0_E sur un voisinage de 0_E . On se ramène donc à l'étude de f_0 . Remarquons que f_0 est de classe \mathcal{C}^k , que $f_0(0_E) = 0_E$

4. Soit $g : E \rightarrow F$. On dit que $g(v) = O(v)$ si $\exists \eta > 0$ et $C > 0$ t.q. $\|v\|_E \leq \eta \Rightarrow \|g(v)\|_F \leq C\|v\|_E$.

5. D'après le Théorème A.5.1, dans ce cas $f'(x) \in \text{Isom}_c(E, F)$.

et que $f'(0_E) = \text{Id}_E$. Puisque f'_0 est continue sur E , alors $\forall \varepsilon > 0, \exists \eta_\varepsilon > 0$ t.q. $\forall u \in E$ vérifiant $\|u\|_E \leq \eta_\varepsilon$ on a $f'_0(u) \in B(\text{Id}_E, \varepsilon)$. Enfin, puisque $\text{GL}_c(E) = \text{Isom}_c(E, E)$ est ouvert dans $\mathcal{L}(E)$, il existe $\varepsilon > 0$ t.q. $B(\text{Id}_E, \varepsilon) \subset \text{GL}_c(E)$.

* Il reste donc à montrer que f_0 est un homéomorphisme d'un voisinage de 0_E sur un autre voisinage de 0_E . Puisque f_0 est continue, il reste à montrer que f_0 est bijective d'un voisinage U_0 de 0_E sur un autre voisinage V_0 de 0_E et que sur le voisinage V_0 , l'application inverse est continue. Introduisons pour cela, pour u, v dans E

$$\varphi_v(u) := v + u - f_0(u),$$

de telle sorte que $\varphi_v(u) = u$ si et seulement si $f_0(u) = v$, c'est-à-dire si et seulement si φ_v admet un point fixe. Vérifions tout d'abord que φ_v est contractante dans un voisinage de 0_E . D'après le Théorème A.1.6 des accroissements finis (et son corollaire), pour tous $r > 0$ et u_1, u_2 dans $B(0, r)$, on a :

$$\|\varphi_v(u_1) - \varphi_v(u_2)\|_E \leq \left(\sup_{u \in B(0, r)} \|\text{Id}_E - f'_0(u)\|_{\mathcal{L}(E)} \right) \|u_1 - u_2\|_E.$$

Ainsi, pour $\rho := \min(\varepsilon, 1/2) \in [0, 1[$ avec $\varepsilon > 0$ t.q. $B(\text{Id}_E, \varepsilon) \subset \text{GL}_c(E)$, et pour $\eta := \eta_\rho$, alors pour tous u_1, u_2 dans $B(0, \eta)$, on a :

$$\|\varphi_v(u_1) - \varphi_v(u_2)\|_E \leq \rho \|u_1 - u_2\|_E,$$

c'est-à-dire, φ_v est ρ -contractante. On a de plus en particulier (pour $u_2 = 0_E$)

$$\|\varphi_v(u_1) - v\|_E \leq \rho \|u_1\|_E,$$

et donc pour tous $u \in B_f(0, \eta/2)$ et $v \in B_f(0, (1 - \rho)\eta/2)$, on a

$$\|\varphi_v(u)\|_E \leq \|v\|_E + \rho \|u\|_E \leq (1 - \rho) \frac{\eta}{2} + \rho \frac{\eta}{2} = \frac{\eta}{2},$$

autrement dit $\varphi_v(B_f(0, \eta/2)) \subset B_f(0, \eta/2)$ si $v \in B_f(0, (1 - \rho)\eta/2)$. Posons

$$V_0 := B(0, (1 - \rho)\eta/2)$$

et introduisons

$$\begin{aligned} \varphi: B_f(0, \eta/2) \times V_0 &\longrightarrow B_f(0, \eta/2) \\ (u, v) &\longmapsto \varphi(u, v) := \varphi_v(u), \end{aligned}$$

de telle sorte que φ vérifie les hypothèses du Théorème A.3.2 du point fixe de Picard à paramètre; φ est continue et uniformément ρ -contractante. Ainsi, $\forall v \in V_0, \exists! u_v \in B_f(0, \eta/2)$ t.q. $\varphi(u_v, v) = \varphi_v(u_v) = u_v$, et $v \mapsto u_v$ est continue sur V_0 . Finalement, f_0 est bijective de $U_0 := f_0^{-1}(V_0)$ dans V_0 et l'application réciproque $f_0^{-1}(v) = u_v$ est continue.

* Nous avons donc démontré que f_0 est un homéomorphisme de U_0 dans V_0 de classe \mathcal{C}^k t.q. $\forall u \in U_0, f'_0(u) \in \mathcal{L}(E)$ est bijective, donc d'après la Proposition A.5.3, f_0 est un \mathcal{C}^k -difféomorphisme de U_0 dans V_0 et finalement le théorème est démontré. ■

A.6 Théorème des fonctions implicites

Théorème A.6.1 – des fonctions implicites

Soient E , F et G trois espaces de Banach, Ω un ouvert de $E \times F$, $f: \Omega \subset E \times F \rightarrow G$, une application de classe \mathcal{C}^k , $k \in \mathbb{N}^*$, et $(\bar{x}, \bar{y}) \in \Omega$, tels que $\partial_y f(\bar{x}, \bar{y}) \in \mathcal{L}(F, G)$ est bijective.

Il existe alors $U \times V \subset \Omega$ un voisinage ouvert du point (\bar{x}, \bar{y}) tel que :

- i) $\forall x \in U$, $\exists! y_x \in V$, tels que $f(x, y_x) = f(\bar{x}, \bar{y})$;
- ii) L'application (unique) $\varphi: U \rightarrow V$, $x \mapsto \varphi(x) := y_x$ est de classe \mathcal{C}^k et $\forall x \in U$:

$$\varphi'(x) = -(\partial_y f(x, \varphi(x)))^{-1} \circ \partial_x f(x, \varphi(x)) ;$$

► On pose

$$\begin{aligned} \Phi: \quad \Omega &\longrightarrow E \times G \\ (x, y) &\longmapsto \Phi(x, y) := (x, f(x, y)). \end{aligned}$$

* L'application Φ est \mathcal{C}^k sur Ω et sa différentielle en (\bar{x}, \bar{y}) est donnée par

$$\begin{aligned} \Phi'(\bar{x}, \bar{y}) \cdot (u, v) &= \begin{pmatrix} \text{Id}_E & 0_{\mathcal{L}(F, E)} \\ \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) & \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \end{pmatrix} \cdot (u, v) \\ &= \left(u, \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) \cdot u + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \cdot v \right) \in E \times G, \end{aligned}$$

pour tout $(u, v) \in E \times F$. On rappelle que $\Phi'(\bar{x}, \bar{y}) \in \mathcal{L}(E \times F, E \times G)$. De plus, pour tout $(a, b) \in E \times G$, il existe un unique couple $(u, v) \in E \times F$ tel que $\Phi'(\bar{x}, \bar{y}) \cdot (u, v) = (a, b)$. En effet, $u = a$ et $v = -(\partial_y f(\bar{x}, \bar{y}))^{-1} \cdot (\partial_x f(\bar{x}, \bar{y}) \cdot a)$, car $\partial_y f(\bar{x}, \bar{y})$ inversible par hypothèse, donc $\Phi'(\bar{x}, \bar{y}) \in \mathcal{L}(E \times F, E \times G)$ est bijective et puisque $E \times F$ et $E \times G$ sont deux espaces de Banach, nous pouvons appliquer le Théorème A.5.4 d'inversion locale à Φ .

* D'après le théorème d'inversion locale, il existe $\Omega' := U' \times V \subset \Omega$, U' et V deux ouverts contenant respectivement \bar{x} et \bar{y} , tel que Φ est un \mathcal{C}^k -difféomorphisme de Ω' dans $\Phi(\Omega')$. On note $\Phi^{-1} =: (g_1, g_2)$ l'inverse de Φ défini sur $\Phi(\Omega') \subset E \times G$ et à valeurs dans $\Omega' \subset E \times F$. Les applications g_1 et g_2 sont de classe \mathcal{C}^k . Alors, pour tout $(x, z) \in \Phi(\Omega')$,

$$(x, z) = (\Phi \circ \Phi^{-1})(x, z) = \Phi(\Phi^{-1}(x, z)) = (g_1(x, z), f(g_1(x, z), g_2(x, z))).$$

Ceci nous donne $g_1(x, z) = x$ et $f(x, g_2(x, z)) = z$ pour tout $(x, z) \in \Phi(\Omega')$. En particulier, pour $\bar{z} := f(\bar{x}, \bar{y})$, en introduisant $\varphi := g_2(\cdot, \bar{z})$ et $U := \{x \in U' \mid (x, \bar{z}) \in \Phi(\Omega')\}$, nous avons :

$$\forall x \in U, \quad f(x, \varphi(x)) = \bar{z}.$$

Défini ainsi, U est bien un ouvert de E (car préimage de l'ouvert $\Phi(\Omega')$ par l'application continue $x \mapsto (x, \bar{z})$ définie sur U') contenant \bar{x} et $\varphi(U) \subset V$ (par définition).

* L'application φ est \mathcal{C}^k sur U donc par composition $x \mapsto f(x, \varphi(x))$ aussi, cette application est constante sur U donc de dérivée nulle et sa dérivée est donnée par :

$$\forall x \in U, \quad \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x)) \circ \varphi'(x) = 0_{\mathcal{L}(E, G)},$$

ce qui conclut la démonstration. ■

Remarque A.6.1. On peut réécrire les conclusions *i)* et *ii)* du théorème des fonctions implicites sous la forme suivante. L'application $\varphi: U \rightarrow V$ est unique, de classe \mathcal{C}^k , on a :

$$x \in U, y \in V \text{ et } f(x, y) = f(\bar{x}, \bar{y}) \iff x \in U \text{ et } y = \varphi(x)$$

et la différentielle de φ est donnée pour tout $x \in U$ par :

$$\varphi'(x) = -(\partial_y f(x, \varphi(x)))^{-1} \circ \partial_x f(x, \varphi(x)).$$

A.7 Conditions nécessaires d'optimalité

Théorème A.7.1 – de l'application ouverte [75, Théorème 3.11.1]

Soient E et F deux espaces de Banach et $f \in \mathcal{L}(E, F)$. Si f est surjective, alors il existe une constante $\alpha > 0$ telle que $f(B_E(0, 1)) \supset B_F(0, \alpha)$, ce qui est équivalent à dire que f est ouverte (par linéarité).

que l'on va étendre au cas non linéaire sous l'hypothèse F de dimension finie. Nous avons besoin du lemme suivant.

Lemme A.7.1. Soit E un Banach tel que E est somme directe algébrique⁶ de deux sous-espaces vectoriels G et H , tous deux fermés, alors, E est somme directe topologique de G et H .

► On pose

$$\begin{aligned} \varphi: G \times H &\longrightarrow E \\ (x, y) &\longmapsto \varphi(x, y) := x + y. \end{aligned}$$

Par hypothèse, φ est un isomorphisme algébrique. Montrons que c'est un isomorphisme topologique. Les sous-espaces G et H sont munis de la norme induite par E , notée $\|\cdot\|$. On munit $G \times H$ de la norme produit $\|(x, y)\|_{G \times H} := \|x\| + \|y\|$. L'application φ est continue car $\|\varphi(x, y)\| = \|x + y\| \leq \|x\| + \|y\| = \|(x, y)\|_{G \times H}$. Le sous-espace G est fermé dans un espace complet donc G complet (et fermé) donc G Banach, de même H Banach. Un produit fini de Banach étant un Banach, $G \times H$ Banach. On a donc que l'application $\varphi \in \text{Isom}(G \times H, E)$ est continue, $G \times H$ et E Banach, donc d'après le Théorème A.5.1 de Banach, φ^{-1} est continue, et $\varphi \in \text{Isom}_c(G \times H, E)$. ■

6. C'est à dire $E = G \oplus H$, i.e. $E = \text{Vect}(G \cup H)$, et $G \cap H = \{0\}$.

Théorème A.7.2 – de l'application ouverte non linéaire

Soit $f: U \subset E \rightarrow F := \mathbb{R}^n$ une application de classe \mathcal{C}^1 dans U , définie sur un ouvert U d'un espace de Banach E . Soit $\bar{x} \in U$ un point régulier de f (i.e. $f'(\bar{x})$ surjective), alors f est localement ouverte⁷ en \bar{x} .

► Puisque $f'(\bar{x}) \in \mathcal{L}(E, F)$ surjective, $E/\text{Ker } f'(\bar{x}) \simeq F$ (isomorphisme algébrique). Donc en notant $G := \text{Ker } f'(\bar{x})$, il existe un supplémentaire algébrique $H \simeq F$, de G dans E , i.e. $E = G \oplus H$ (algébrique). Le sous-espace G est fermé car $f'(\bar{x})$ est continue et H est fermé car c'est un sous-espace vectoriel de dimension finie (comme F). D'après le Lemme A.7.1, $E = G \oplus H$ (topologique) et donc en particulier, la projection $p: E \rightarrow G$ est continue. Introduisons l'application $\psi: U \subset E \rightarrow F \times G$, définie par $\psi := (f, p)$, de classe \mathcal{C}^1 et montrons que $\psi'(\bar{x}) = (f'(\bar{x}), p) \in \mathcal{L}(E, F \times G)$ est bijective. L'application $f'(\bar{x})$ est surjective par hypothèse, donc $\psi'(\bar{x})$ est surjective. Maintenant, si $\psi'(\bar{x}) \cdot v = (0, 0)$, alors $v \in \text{Ker } f'(\bar{x}) \cap \text{Ker } p = G \cap H = \{0\}$, donc $\psi'(\bar{x})$ est injective. Finalement, $\psi'(\bar{x})$ est bijective, donc d'après le Théorème A.5.4 d'inversion locale, il existe un ouvert $\Omega \in \mathcal{V}_U(\bar{x})$ tel que $\psi|_\Omega$ est un \mathcal{C}^1 -difféomorphisme de Ω sur $\psi(\Omega)$. L'application $\psi|_\Omega$ est donc ouverte, tout comme $f|_\Omega = q \circ \psi|_\Omega$, $q: F \times G \rightarrow F$ projection canonique, car q ouverte. Pour conclure, pour tout $V \in \mathcal{V}_U(\bar{x})$, $V \cap \Omega \in \mathcal{V}_\Omega(\bar{x})$ donc $f|_\Omega(V \cap \Omega) \in \mathcal{V}_F(f(\bar{x}))$ (car $f|_\Omega$ ouverte), et puisque $f|_\Omega(V \cap \Omega) \subset f(V)$, on a $f(V) \in \mathcal{V}_F(f(\bar{x}))$. ■

Théorème A.7.3 – CN1 d'optimalité

Soit E un espace de Banach. Soient $f \in \mathcal{C}^1(U, \mathbb{R})$ et $h \in \mathcal{C}^1(U, \mathbb{R}^m)$, avec $m \in \mathbb{N}^*$ et U un ouvert de E . Soit $\bar{x} \in U$ une solution (locale ou globale) au problème

$$\min \{f(x) \mid x \in U, h(x) = 0_{\mathbb{R}^m}\}.$$

Il existe alors $(\lambda, \lambda^0) \in (\mathbb{R}^{m+1})^*$, $(\lambda, \lambda^0) \neq 0_{(\mathbb{R}^{m+1})^*}$, tel que $\lambda^0 f'(\bar{x}) + \lambda \circ h'(\bar{x}) = 0_{E'}$.

► Soit $\bar{x} \in U \subset E$ un minimum de f sous la contrainte $h = 0_{\mathbb{R}^m}$. Il existe alors $V \in \mathcal{V}_U(\bar{x})$ (l'ensemble des voisinages de \bar{x} dans U) tel que $\forall x \in V : h(x) = 0_{\mathbb{R}^m} \Rightarrow f(x) \geq f(\bar{x})$. Posons $\varphi: U \subset E \rightarrow \mathbb{R}^m \times \mathbb{R}$, $\varphi(x) := (h(x), f(x))$ de sorte que $\varphi(\bar{x}) = (h(\bar{x}), f(\bar{x})) = (0_{\mathbb{R}^m}, f(\bar{x})) \in \mathbb{R}^{m+1}$. L'application φ ne peut être localement ouverte en \bar{x} car sinon $\varphi(V) \in \mathcal{V}_{\mathbb{R}^{m+1}}(\varphi(\bar{x}))$ et il existerait alors $\varepsilon > 0$ tel que $(0_{\mathbb{R}^m}, f(\bar{x}) - \varepsilon) \in \varphi(V)$, ce qui contredirait l'optimalité de \bar{x} . Puisque φ n'est pas localement ouverte en \bar{x} , d'après le Théorème A.7.2, par contraposée, $\varphi'(\bar{x}) \in \mathcal{L}(E, \mathbb{R}^{m+1})$ n'est pas surjective; ainsi $\text{Im } \varphi'(\bar{x}) \subsetneq \mathbb{R}^{m+1}$ et il existe donc $\tilde{\lambda} := (\lambda, \lambda^0) \in (\mathbb{R}^{m+1})^*$, $\tilde{\lambda} \neq 0_{(\mathbb{R}^{m+1})^*}$ tel que $\text{Im } \varphi'(\bar{x}) \subset \text{Ker } \tilde{\lambda}$ (puisque $\text{Im } \varphi'(\bar{x})$ est inclus dans un hyperplan, et tout hyperplan est le noyau d'une forme linéaire), i.e. $\forall v \in E$, $\tilde{\lambda}(\varphi'(\bar{x}) \cdot v) = 0$, c-à-d $\tilde{\lambda} \circ \varphi'(\bar{x}) = 0_{E'}$. ■

Remarque A.7.1. Dans le théorème précédent, on ne fait pas d'hypothèses de qualification des contraintes, le paramètre λ^0 peut être nul.

7. i.e. envoie les voisinages de \bar{x} dans U sur les voisinages de $f(\bar{x})$ dans \mathbb{R}^n : $\forall V \in \mathcal{V}_U(\bar{x}) \Rightarrow f(V) \in \mathcal{V}_{\mathbb{R}^n}(f(\bar{x}))$.

Calcul des variations

B.1	Introduction	121
B.2	Solutions des problèmes variationnels	124
B.3	Équation de Euler-Lagrange	128
B.4	Condition de Legendre	135
B.5	Extrémum libre et contraintes affines	139
B.5.1	Extrémum libre	139
B.5.2	Extremum lié : contraintes affines	144
B.5.3	“Two-norm discrepancy”	145
B.6	Condition de Jacobi	150

B.1 Introduction

Le *calcul des variations* est une discipline ancienne qui pris naissance suite à l’introduction du problème de la courbe *brachistochrone*¹ [71] vers la fin du XVII^e siècle par le mathématicien Johann Bernoulli. La courbe brachistochrone est par définition une courbe dans un plan vertical de coordonnées (x, y) sur laquelle un point matériel pesant placé dans un champ de pesanteur uniforme, glissant sans frottement et sans vitesse initiale, présente un temps de parcours minimal parmi toutes les courbes joignant deux extrémités fixées. Notons $A := (x_a, y_a)$ et $B := (x_b, y_b)$ les deux extrémités de la courbe. Sans perte de généralité, nous pouvons supposer $x_a = y_a = 0$. En choisissant un repère dirigé vers le bas, la condition $y_b > y_a$ doit être vérifiée pour assurer l’existence (au sens physique) d’une solution au problème. En revanche, par symétrie de celui-ci nous pouvons fixer $x_b > x_a$. On élimine le cas de la chute libre, c’est-à-dire $x_a = x_b$. Voir la Figure B.1 pour une illustration du problème.

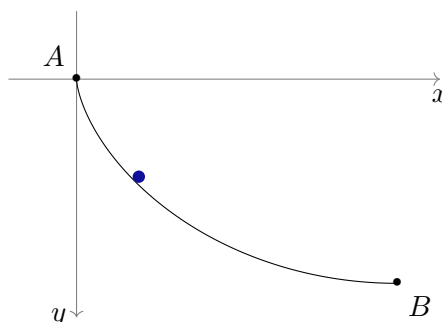


FIGURE B.1 – Courbe brachistochrone

1. Voir https://fr.wikipedia.org/wiki/Courbe_brachistochrone pour un historique.

Pour fixer les idées, nous recherchons une courbe Γ dans le plan (x, y) donnée par le graphe d'une fonction f de classe \mathcal{C}^1 , *i.e.* $\Gamma := \{(x, f(x)) \mid x \in [x_a, x_b]\}$. Le point matériel est déterminé par sa position à l'instant t que nous notons $p(t) := (x(t), y(t))$. Pour tout temps t ce point appartient à Γ , ainsi son vecteur vitesse est donné par :

$$\dot{p}(t) := \frac{d}{dt}p(t) = (\dot{x}(t), \dot{y}(t)) = (\dot{x}(t), f'(x(t)) \dot{x}(t)) = (1, f'(x(t))) \dot{x}(t) \in T_{p(t)}\Gamma,$$

où $T_{p(t)}\Gamma$ est l'espace tangent à la courbe Γ au point $p(t)$. D'un autre côté, la conservation de l'énergie mécanique au cours du temps nous permet d'exprimer l'énergie cinétique en fonction de l'énergie potentielle de pesanteur par la relation :

$$\frac{1}{2}m\|\dot{p}(t)\|^2 = mgy(t) = mgf(x(t)),$$

où m est la masse du point matériel (qui ne joue aucun rôle) et g est l'accélération de pesanteur. En combinant les deux relations précédentes (on suppose $\dot{x}(t) > 0$), nous obtenons l'équation différentielle

$$\dot{x}(t) = \sqrt{\frac{2gf(x(t))}{1 + f'(x(t))^2}} =: \varphi(x(t)).$$

Remarquons qu'à l'instant initial, nous avons $x(0) = x_a$ et $\varphi(x_a) = 0$ car $f(x_a) = y_a = 0$ (la vitesse initiale est nulle). Quelle que soit la fonctionnelle f vérifiant $f(x_a) = y_a = 0$, la trajectoire triviale $x(t) = x_a = 0$ est solution de l'équation $\dot{x}(t) = \varphi(x(t))$. Cependant, cette solution ne nous intéresse pas car nous voulons que le point matériel ait comme abscisse $x(T) = x_b > x_a$ à l'instant terminal T . Fort heureusement, en introduisant

$$F(x) := \int_{x_a}^x \frac{dz}{\varphi(z)} = \int_0^x \frac{dz}{\varphi(z)},$$

et en supposant que $|F(x_a + \varepsilon)| = F(\varepsilon) < +\infty$ pour un certain $\varepsilon > 0$, alors $x(t) := F^{-1}(t)$, $x(0) := x_a$, est une autre solution différente de la solution triviale. Notons pour ce convaincre de cela que si $x(t)$ est solution de l'équation $\dot{x}(t) = \varphi(x(t))$, alors pour tout temps t nous avons $F(x(t)) = t$, cf. :

$$\frac{d}{dt}(F \circ x)(t) = F'(x(t)) \dot{x}(t) = \frac{\dot{x}(t)}{\varphi(x(t))} = 1.$$

En conclusion, le temps de parcours du point matériel est donné par

$$T(f) := F(x_b) = \int_{x_a}^{x_b} \sqrt{\frac{1 + f'(x)^2}{2gf(x)}} dx,$$

où nous avons introduit la notation $T(f)$ pour mettre en évidence la dépendance à f , car f est pour nous l'inconnue du problème. Nous cherchons donc la fonctionnelle f qui minimise le temps de parcours $T(f)$ et qui vérifie les conditions aux extrémités $f(x_a) = y_a$ et $f(x_b) = y_b$. Ce problème peut s'écrire sous la forme

$$\inf_f \{T(f) \mid f(x_a) = y_a, f(x_b) = y_b\}. \quad (\text{B.1})$$

La solution à ce problème, toujours notée f , est connue depuis longtemps : plus précisément, la courbe solution $\Gamma = \{(x, f(x)) \mid x \in [x_a, x_b]\}$ est ce que l'on appelle une **cycloïde**, qui sous une forme paramétrée est donnée par

$$\begin{cases} x(\theta) = \frac{D}{2}(\theta - \sin(\theta)), \\ y(\theta) = \frac{D}{2}(1 - \cos(\theta)), \end{cases}$$

pour $\theta \in [\theta_a, \theta_b]$. Nous pouvons remarquer que $f(x_a) = f(x(\theta_a)) = y_a = y(\theta_a) = 0$ est bien vérifiée (prendre $\theta_a = 0$) quels que soient D et θ_b , qui eux, sont déterminés par la relation $f(x_b) = y_b$ pour $\theta = \theta_b$, *i.e.* par $x(\theta_b) = x_b$ et $y(\theta_b) = y_b$.

Bien que ce problème soit ancien, il n'en reste pas moins difficile. Tout d'abord, nous avons fixé les idées en supposant la fonctionnelle f de classe \mathcal{C}^1 , sous-entendu sur $[x_a, x_b]$. Pourtant, puisque $x'(\theta) = y(\theta)$ et $y'(\theta) = D \sin(\theta)/2$, un calcul rapide montre que

$$\frac{y'(\theta)}{x'(\theta)} \rightarrow +\infty \quad \text{quand} \quad \theta \rightarrow 0,$$

autrement dit, la fonctionnelle f a une pente verticale en $x = x_a = 0$. Elle n'y est donc pas \mathcal{C}^1 ! Ceci nous amène à une première difficulté que nous avons cachée lorsque nous avons défini le Problème (B.1) : comment bien choisir le domaine sur lequel on minimise T ? C'est-à-dire :

Q1. À quelle classe de fonctions appartient f ?

Une fois la classe choisie, que l'on note \mathcal{F} , nous pouvons nous demander s'il existe bien une solution au problème, c'est-à-dire si l'infimum est bien atteint, *i.e.* si l'infimum est un minimum :

Q2. Existe-t-il une fonction $\bar{f} \in \mathcal{F}$ telle que

$$T(\bar{f}) = \inf_{f \in \mathcal{F}} \{T(f) \mid f(x_a) = y_a, f(x_b) = y_b\} \quad ?$$

Enfin, connaissant l'existence d'un tel minimum, nous pouvons nous demander comment le calculer d'un point de vue pratique, autrement dit, nous nous demandons :

Q3. Quelles conditions nécessaires et/ou suffisantes d'optimalité sont associées à notre problème d'optimisation ?

Une possibilité, pour résoudre notre problème d'optimisation, est de montrer dans un premier temps qu'une solution au problème existe, et ce, à l'aide de théorèmes généraux. On montre que :

$$\exists \bar{f} \in \mathcal{F} \quad \text{t.q.} \quad \forall f \in \mathcal{F} : T(\bar{f}) \leq T(f).$$

L'idée ensuite est de trouver cet (ou un) optimum (il n'est pas nécessairement unique) à l'aide de conditions nécessaires d'optimalité. Pour faire cela, on peut comparer les différents candidats obtenus à l'aide des conditions nécessaires et ne retenir que le meilleur. Nous pourrions choisir une autre approche qui reviendrait plutôt à appliquer tout d'abord les conditions nécessaires d'optimalité pour réduire l'ensemble des candidats à l'optimalité. On

ne garderait ensuite que le (ou les) meilleur candidat, que l'on note \tilde{f} (si on le suppose unique), pour montrer, en utilisant des propriétés spécifiques au problème (de convexité par exemple), que ce candidat est bien meilleur que toute autre possibilité. Autrement dit, on montrerait que :

$$\forall f \in \mathcal{F} : T(\tilde{f}) \leq T(f),$$

ce qui nous permettrait de conclure qu'il existe un minimum \bar{f} au problème et que $\bar{f} = \tilde{f}$.

Dans ce chapitre, nous nous intéressons seulement aux conditions nécessaires d'optimalité. Pour plus de détails sur l'existence de solutions et sur les conditions suffisantes d'optimalité, nous renvoyons le lecteur aux références [25, 27, 35, 65, 71].

B.2 Solutions des problèmes variationnels

Soit \mathcal{I} un intervalle ouvert de \mathbb{R} , Ω un ouvert de \mathbb{R}^n et une application continue

$$\begin{aligned} L : \mathcal{I} \times \Omega \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (t, x, u) &\longmapsto L(t, x, u). \end{aligned}$$

Nous appellerons une telle application, un **lagrangien**. Nous pouvons remarquer que notre point de vue est local car nous considérons la variable x dans Ω , un ouvert de \mathbb{R}^n , au lieu par exemple, de considérer x dans une variété différentielle. Si nous avons fait le choix d'une variété différentielle, alors le vecteur u appartiendrait à l'espace tangent à la variété en x , et nous aurions défini le lagrangien légèrement différemment. Ce point de vue local est donc une simplification.

Soient a et b deux points de \mathcal{I} tels que $a < b$. Soit $x(\cdot) \in \mathcal{C}^1([a, b], \Omega)$, c'est-à-dire une application continue sur $[a, b]$, à valeurs dans Ω , dérivable sur $]a, b[$, dont la dérivée $\dot{x}(\cdot)$ est continue sur $]a, b[$ et se prolonge de manière continue sur $[a, b]$. Pour faire plus simple, on peut imaginer $x(\cdot) \in \mathcal{C}^1(\mathcal{I}, \Omega)$ ou $x(\cdot)$ de classe \mathcal{C}^1 sur un ouvert $I \subset \mathcal{I}$ contenant $[a, b]$.

Remarque B.2.1. La paramètre t jouera le rôle du “temps” et la dérivée de $t \mapsto x(t)$ par rapport au temps sera notée $\dot{x}(\cdot)$ au lieu de $x'(\cdot)$. Nous avons aussi mis en évidence la dépendance de x par rapport au temps en écrivant $x(\cdot)$ au lieu de x . En effet, ici $x(\cdot)$ est une fonction du temps t contrairement à la variable x dans le lagrangien L qui est simplement un point de \mathbb{R}^n . Cependant, nous allons vite abandonner ce réflexe et nous écrirons x aussi bien pour le point de \mathbb{R}^n que la fonction du temps. Le contexte permettra de s'y retrouver.

On définit le **coût** associé à x par la fonctionnelle

$$J(x) := \int_a^b L(t, x(t), \dot{x}(t)) dt.$$

Étant donnés deux points A et B dans Ω , on s'intéresse au problème d'optimisation

$$\text{minimiser } J(x) : x \in \mathcal{C}^1([a, b], \Omega), x(a) = A, x(b) = B, \quad (\text{PCV})$$

que l'on peut aussi écrire sous la forme

$$\inf_{x \in \mathcal{C}^1([a, b], \Omega)} \{J(x) \mid x(a) = A, x(b) = B\},$$

ou encore

$$\inf_x \{J(x) \mid x \in \mathcal{K}\},$$

avec $\mathcal{K} := \{x \in \mathcal{C}^1([a, b], \Omega) \mid x(a) = A, x(b) = B\}$. Ce **problème variationnel** est le plus classique en **calcul des variations**. Dans ce problème, les instants initiaux et finaux sont fixés, tout comme les extrémités : on parle de **conditions aux limites** simples.

Remarque B.2.2. En introduction, nous avons parlé du problème de la courbe brachistochrone. Dans cet exemple, les notations étaient différentes pour être plus explicites. La courbe était définie dans le plan (x, y) , où x jouait le rôle de t dans la formulation du Problème (PCV). Nous recherchions une fonction $f(x)$, ainsi f jouait le rôle de la variable x . Les coordonnées x_a et x_b sont ici les temps a et b et les coordonnées y_a et y_b sont ici les points A et B . Enfin, le temps de parcours $T(f)$ est remplacé par la fonction de coût $J(x)$. Il est à noter que le problème de la courbe brachistochrone ne s'écrit pas sous la forme (PCV). En effet, le lagrangien

$$L(t, x, u) = \sqrt{\frac{1 + u^2}{2gx}}$$

est défini sur $\mathbb{R} \times \mathbb{R}^* \times \mathbb{R}$, c'est-à-dire que l'on a dans ce cas $\Omega := \mathbb{R}^*$. Jusqu'ici tout va bien, mais la condition initiale s'écrit $x(a) = A$ avec $A := 0 \notin \Omega$. Ici $A \in \partial\Omega$, c'est-à-dire à la frontière de Ω . C'est ce qui explique que la solution \bar{x} , i.e. la courbe brachistochrone, peut avoir une pente verticale en l'instant initial $a := 0$. Dans le problème de la courbe brachistochrone, la solution \bar{x} appartient à $\mathcal{C}^0([a, b], \bar{\Omega}) \cap \mathcal{C}^1(]a, b[, \bar{\Omega})$ et non à $\mathcal{C}^1([a, b], \Omega)$. Nous faisons donc une simplification dans cette présentation : la fonctionnelle $J(x)$ est pour nous une intégrale de Riemann au sens usuel, alors que dans le cas du brachistochrone c'est une intégrale impropre. Pour une présentation du calcul des variations dans un cadre moins régulier, nous renvoyons à [27].

Définition B.2.1 – Fonctions admissibles et solutions (globales)

Une fonction $x: [a, b] \rightarrow \Omega$ est dite **admissible** pour (PCV) si $x \in \mathcal{C}^1([a, b], \Omega)$ et si les conditions aux limites sont satisfaites : $x(a) = A$ et $x(b) = B$, c'est-à-dire si $x \in \mathcal{K}$.

On appelle **solution globale** (ou minimum global) du Problème (PCV), une fonction \bar{x} admissible telle que $J(\bar{x}) \leq J(x)$ pour toute autre fonction admissible x .

En optimisation en dimension finie, la propriété d'optimalité locale ne dépend pas du choix de la norme car toutes les normes sont équivalentes. En dimension infinie, cet aspect est différent. L'espace vectoriel des fonctions continues $\mathcal{C}^0([a, b], \mathbb{R}^n)$ est généralement muni de la norme de la convergence uniforme :

$$\|x\|_{\mathcal{C}^0} := \|x\|_{\infty} = \max \{\|x(t)\| \mid t \in [a, b]\},$$

où $\|\cdot\|$ est ici la norme euclidienne dans \mathbb{R}^n . Rappelons que l'ensemble $\mathcal{C}^0([a, b], \mathbb{R}^n)$ muni de la norme $\|\cdot\|_{\mathcal{C}^0}$ est un espace de Banach. L'ensemble $\mathcal{C}^1([a, b], \mathbb{R}^n)$ depuis fonctions de classe \mathcal{C}^1 est quant à lui généralement muni de la norme

$$\|x\|_{\mathcal{C}^1} := \|x\|_{\infty} + \|\dot{x}\|_{\infty}.$$

Nous rappelons de même que $\mathcal{C}^1([a, b], \mathbb{R}^n)$ muni de la norme $\|\cdot\|_{\mathcal{C}^1}$ est un espace de Banach. Notons que

$$\mathcal{C}^1([a, b], \mathbb{R}^n) \subset \mathcal{C}^0([a, b], \mathbb{R}^n)$$

et il serait donc tout à fait possible de munir $\mathcal{C}^1([a, b], \mathbb{R}^n)$ de la norme $\|\cdot\|_{\mathcal{C}^0}$. Ceci nous amène, pour $\eta > 0$, à l'introduction des notations suivantes :

$$\begin{aligned} B_w(x, \eta) &:= \{y \in \mathcal{C}^1([a, b], \mathbb{R}^n) \mid \|x - y\|_{\mathcal{C}^1} < \eta\}, \\ B_s(x, \eta) &:= \{y \in \mathcal{C}^1([a, b], \mathbb{R}^n) \mid \|x - y\|_{\mathcal{C}^0} < \eta\}. \end{aligned}$$

L'indice w signifie “weak” (*i.e.* faible) tandis que le s est pour “strong” (*i.e.* fort). Notons que le η n'est pas nécessairement petit. Introduisons la définition suivante.

Définition B.2.2 – Solutions locales faibles et fortes

Soit \bar{x} une fonction admissible, *i.e.* $\bar{x} \in \mathcal{K}$.

On dit que \bar{x} est une **solution locale faible** si :

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B_w(\bar{x}, \eta) \cap \mathcal{K} : J(\bar{x}) \leq J(x).$$

On dit que \bar{x} est une **solution locale forte** si :

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B_s(\bar{x}, \eta) \cap \mathcal{K} : J(\bar{x}) \leq J(x).$$

Remarque B.2.3. On peut remplacer dans la définition précédente les boules ouvertes $B_w(\bar{x}, \eta)$ et $B_s(\bar{x}, \eta)$ par les boules fermées

$$\begin{aligned} \bar{B}_w(\bar{x}, \eta) &:= \{x \in \mathcal{C}^1([a, b], \mathbb{R}^n) \mid \|\bar{x} - x\|_{\mathcal{C}^1} \leq \eta\}, \\ \bar{B}_s(\bar{x}, \eta) &:= \{x \in \mathcal{C}^1([a, b], \mathbb{R}^n) \mid \|\bar{x} - x\|_{\mathcal{C}^0} \leq \eta\}. \end{aligned}$$

Puisque pour tout $x \in \mathcal{C}^1([a, b], \mathbb{R}^n)$, $\|x\|_{\mathcal{C}^0} \leq \|x\|_{\mathcal{C}^0} + \|\dot{x}\|_{\mathcal{C}^0} = \|x\|_{\mathcal{C}^1}$, il est clair que

$$B_w(x, \eta) \subset B_s(x, \eta).$$

Ainsi, un minimum local faible est un minimum local relatif à un ensemble plus petit qu'un minimum local fort, et on a :

Proposition B.2.3

Soit $\bar{x} \in \mathcal{K}$.

$$\bar{x} \text{ solution globale} \Rightarrow \bar{x} \text{ solution locale forte} \Rightarrow \bar{x} \text{ solution locale faible}.$$

Voici un exemple de solution locale faible qui n'est pas une solution locale forte.

Exemple B.2.1 ([65]). Considérons le problème suivant :

$$\inf_x \left\{ J(x) := \int_0^\pi x^2(t)(1 - \dot{x}^2(t)) dt \mid x \in \mathcal{K} \right\},$$

avec $\mathcal{K} := \{x \in \mathcal{C}^1([0, \pi], \mathbb{R}) \mid x(0) = 0, x(\pi) = 0\}$. La fonction $\bar{x} \equiv 0$ est une solution locale faible mais non forte. En effet, il est clair que :

$$\forall x \in B_w(\bar{x}, 1) \cap \mathcal{K} : J(\bar{x}) = 0 \leq J(x),$$

car si $x \in B_w(\bar{x}, 1)$ alors $\|x - \bar{x}\|_{\mathcal{C}^1} = \|x\|_{\mathcal{C}^1} = \|x\|_{\infty} + \|\dot{x}\|_{\infty} < 1$ donc en particulier $\|\dot{x}\|_{\infty} < 1$ et ainsi pour tout temps t , on a $\dot{x}^2(t) < 1$, et donc $J(x) \geq 0$. Cependant, \bar{x} n'est pas une solution locale forte, et donc a fortiori ce n'est pas non plus une solution globale. Pour montrer cela, il suffit de trouver pour tout $\varepsilon > 0$, une perturbation h de faible amplitude (*i.e.* vérifiant $\|h\|_{\infty} < \varepsilon$) telle que $J(\bar{x} + h) < J(\bar{x}) = 0$ et $\bar{x} + h$ admissible. Pour arriver à nos fins, il est nécessaire de choisir une fonction h ayant de fortes oscillations car il faut $\dot{x}^2(t) > 1$ souvent pour obtenir un coût strictement négatif. Posons pour $\varepsilon > 0$ et $n \in \mathbb{N}^*$:

$$h_n(t) := \varepsilon \sin(nt).$$

Ainsi, $x_n := \bar{x} + h_n = h_n \in \mathcal{K}$ puisque $h_n(0) = h_n(\pi) = 0$ et $h_n \in \mathcal{C}^1([0, \pi], \mathbb{R})$. Et puisque $\sin(nt) = 1$ pour $t = \pi/2n \in [0, \pi]$, on a $\|h_n\|_{\infty} = \varepsilon$, autrement dit :

$$x_n \in \bar{B}_s(\bar{x}, \varepsilon) \cap \mathcal{K}.$$

Par contre

$$\begin{aligned} J(x_n) &= \int_0^{\pi} \varepsilon^2 \sin^2(nt) (1 - \varepsilon^2 n^2 \cos^2(nt)) \, dt \\ &= \varepsilon^2 \int_0^{\pi} \sin^2(nt) (1 - \mu \cos^2(nt)) \, dt, \quad \mu := (\varepsilon n)^2, \\ &= \varepsilon^2 \left(\int_0^{\pi} \sin^2(nt) \, dt - \frac{\mu}{4} \int_0^{\pi} \sin^2(2nt) \, dt \right) \\ &= \varepsilon^2 \left(\frac{1}{n} \int_0^{n\pi} \sin^2(\tau) \, d\tau - \frac{\mu}{8n} \int_0^{2n\pi} \sin^2(\tau) \, d\tau \right) \\ &= \varepsilon^2 \left(\int_0^{\pi} \sin^2(\tau) \, d\tau - \frac{\mu}{4} \int_0^{\pi} \sin^2(\tau) \, d\tau \right) \\ &= \varepsilon^2 \left(1 - \left(\frac{\varepsilon n}{2} \right)^2 \right) \int_0^{\pi} \sin^2(\tau) \, d\tau, \end{aligned}$$

donc en choisissant $n > 2/\varepsilon$, on obtient $J(x_n) < J(\bar{x}) = 0$, ce qui prouve que \bar{x} n'est pas une solution locale forte. Remarquons de plus que pour $\varepsilon > 0$ fixé :

$$\lim_{n \rightarrow +\infty} J(x_n) = -\infty,$$

et donc le problème n'admet pas de solution globale ! □

Exemple B.2.2. Insistons sur la différence entre solutions locales faibles et fortes. Reprenons la solution locale faible $\bar{x} \equiv 0$ de l'exemple précédent et les perturbations au sens fort $h_n(t) = \varepsilon \sin(nt)$. Pour fixer les idées, on considère

$$\varepsilon \in \{0.02, 0.05, 0.1, 0.25\} \quad \text{et} \quad n_{\varepsilon} := \frac{1}{\varepsilon} - 1 \in \{49, 19, 9, 3\}.$$

On pose alors

$$h_{s,\varepsilon}(t) := h_{n_{\varepsilon}}(t) = \varepsilon \sin \left(\left(\frac{1}{\varepsilon} - 1 \right) t \right),$$

de telle sorte que $h_{s,\varepsilon}(0) = h_{s,\varepsilon}(\pi) = 0$, $\|h_{s,\varepsilon}\|_{\mathcal{C}^0} = \varepsilon$ et $\|h_{s,\varepsilon}\|_{\mathcal{C}^1} = 1$, *i.e.* $\bar{x} + h_{s,\varepsilon} \in \bar{B}_s(\bar{x}, \varepsilon) \cap \mathcal{K}$. Cette perturbation au sens fort a une faible amplitude mais de fortes oscillations. Nous pouvons comparer cette perturbation avec la perturbation au sens faible suivante :

$$h_{w,\varepsilon}(t) := r\varepsilon \sin\left(\left(\frac{1}{r} - 1\right)t\right),$$

où $r := 1/4$ est utile seulement pour fixer la norme $\|h_{w,\varepsilon}\|_{\mathcal{C}^1} = \varepsilon$, de telle sorte que $\bar{x} + h_{w,\varepsilon} \in \bar{B}_w(\bar{x}, \varepsilon) \cap \mathcal{K}$. Cette perturbation a une faible amplitude mais aussi de faibles oscillations comme le montre la figure B.2. \square

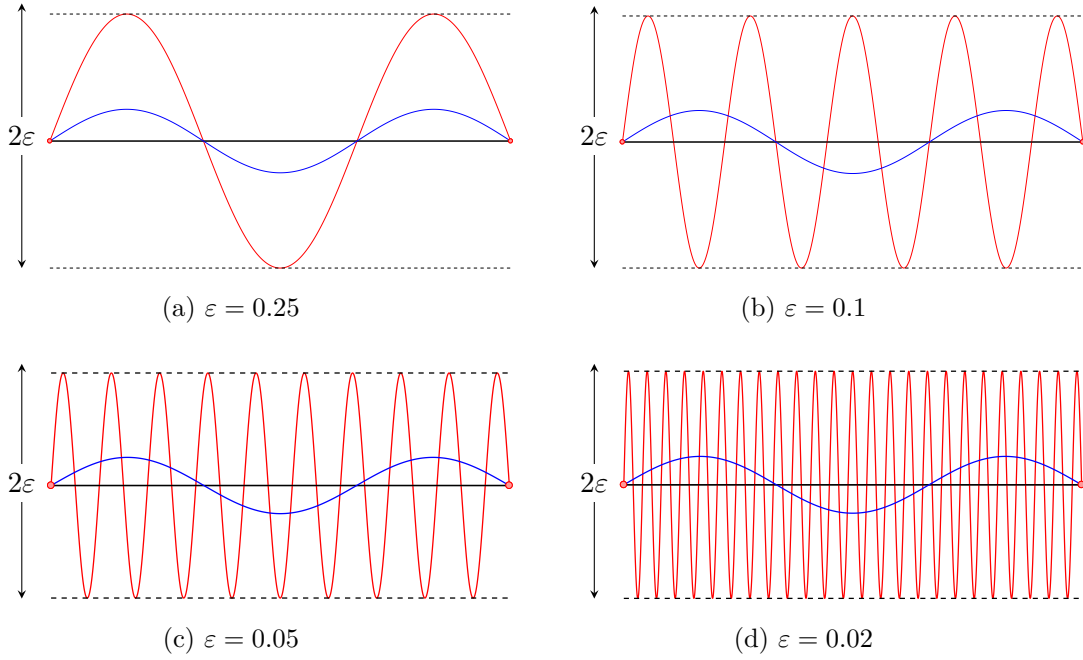


FIGURE B.2 – Perturbations faibles vs fortes. La trajectoire de référence $\bar{x} \equiv 0$ est représentée par un trait plein noir, la perturbation forte $\bar{x} + h_{s,\varepsilon}$ est tracée en rouge tandis que la perturbation faible $\bar{x} + h_{w,\varepsilon}$ est en bleu. Les extrémités (a, A) et (b, B) sont représentées par de petits disques rouges. On ajoute l'enveloppe $\bar{x} \pm \|h_{s,\varepsilon}\|_{\mathcal{C}^0}$ qui est représentée par des tirets noirs. Rappelons que $\bar{x} + h_{s,\varepsilon} \in \bar{B}_s(\bar{x}, \varepsilon) \cap \mathcal{K}$ et $\bar{x} + h_{w,\varepsilon} \in \bar{B}_w(\bar{x}, \varepsilon) \cap \mathcal{K}$.

B.3 Équation de Euler-Lagrange

Dans l'Exemple B.2.1 précédent, nous avons montré que la fonction $\bar{x} \equiv 0$ est une solution locale faible, elle vérifie donc certaines conditions d'optimalité que nous allons détailler dans cette section. Nous avons besoin dans un premier temps du résultat suivant : la fonctionnelle J est de classe \mathcal{C}^1 si L est aussi \mathcal{C}^1 .

Proposition B.3.1

Supposons le lagrangien L de classe \mathcal{C}^1 . Alors, l'application

$$\begin{aligned} J: (\mathcal{C}^1([a, b], \Omega), \|\cdot\|_{\mathcal{C}^1}) &\longrightarrow (\mathbb{R}, |\cdot|) \\ x &\longmapsto J(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt \end{aligned}$$

est de classe \mathcal{C}^1 .

► Définissons quelques applications qui vont nous être utiles.

1. Introduisons l'application

$$\begin{aligned} A: (\mathcal{C}^1([a, b], \Omega), \|\cdot\|_{\mathcal{C}^1}) &\longrightarrow (\mathcal{C}^0([a, b], \Omega) \times \mathcal{C}^0([a, b], \mathbb{R}^n), \|\cdot\|_{\mathcal{C}^0 \times \mathcal{C}^0}) \\ x &\longmapsto A(x) := (x, \dot{x}), \end{aligned}$$

où l'on définit $\|(x, v)\|_{\mathcal{C}^0 \times \mathcal{C}^0} := \|x\|_{\mathcal{C}^0} + \|v\|_{\mathcal{C}^0}$. Cette application est linéaire (évident) et continue puisque :

$$\|A(x)\|_{\mathcal{C}^0 \times \mathcal{C}^0} = \|x\|_{\mathcal{C}^1}.$$

2. Introduisons l'application

$$\begin{aligned} \mathcal{L}: (\mathcal{C}^0([a, b], \Omega) \times \mathcal{C}^0([a, b], \mathbb{R}^n), \|\cdot\|_{\mathcal{C}^0 \times \mathcal{C}^0}) &\longrightarrow (\mathcal{C}^0([a, b], \mathbb{R}), \|\cdot\|_{\mathcal{C}^0}) \\ (x, u) &\longmapsto \mathcal{L}(x, u) \end{aligned}$$

telle que pour tout $t \in [a, b]$, $\mathcal{L}(x, u)(t) := L(t, x(t), u(t))$. Puisque L est \mathcal{C}^1 , alors \mathcal{L} aussi, cf. Proposition A.2.1.

3. Introduisons enfin l'application

$$\begin{aligned} \varphi: (\mathcal{C}^0([a, b], \mathbb{R}), \|\cdot\|_{\mathcal{C}^0}) &\longrightarrow (\mathbb{R}, |\cdot|) \\ f &\longmapsto \varphi(f) := \int_a^b f(t) dt. \end{aligned}$$

Cette application est linéaire (évident) et continue puisque

$$|\varphi(f)| \leq \int_a^b |f(t)| dt \leq (b-a)\|f\|_{\mathcal{C}^0}.$$

À l'aide de ces applications, nous pouvons écrire $J = \varphi \circ \mathcal{L} \circ A$. Puisque toute application linéaire continue est \mathcal{C}^1 et puisque la composée d'applications \mathcal{C}^1 et elle-même de classe \mathcal{C}^1 , on peut alors conclure que J est de classe \mathcal{C}^1 . ■

Remarque B.3.1. Il est à noter que $\mathcal{L}(x, u)$ appartient à $(\mathcal{C}^0([a, b], \mathbb{R}), \|\cdot\|_{\mathcal{C}^0})$. Nous avons donc utilisé la même notation pour la norme, que ce soit pour des applications à valeurs dans \mathbb{R}^n ou dans \mathbb{R} . Dans ce dernier cas, on a :

$$\|\mathcal{L}(x, u)\|_{\mathcal{C}^0} = \|\mathcal{L}(x, u)\|_{\infty} = \max_{t \in [a, b]} |\mathcal{L}(x, u)(t)|.$$

La différentielle de J (au sens de Fréchet) en un point $x \in \mathcal{C}^1([a, b], \Omega)$ appliquée à un vecteur $h \in \mathcal{C}^1([a, b], \mathbb{R}^n)$ est donnée par² :

$$J'(x) \cdot h = (\varphi \circ \mathcal{L}'(A(x)) \circ A) \cdot h.$$

On rappelle, pour obtenir cela, que la différentielle d'une application linéaire T au point x appliquée en le vecteur h est donnée par $T'(x) \cdot h = T(h)$, c'est-à-dire $T'(x) = T$. Ensuite, puisque :

$$\mathcal{L}'(A(x)) \cdot A(h) = \frac{\partial \mathcal{L}}{\partial x}(x, \dot{x}) \cdot h + \frac{\partial \mathcal{L}}{\partial u}(x, \dot{x}) \cdot \dot{h},$$

nous obtenons finalement :

$$J'(x) \cdot h = \int_a^b \left(\frac{\partial L}{\partial x}(t, x(t), \dot{x}(t)) \cdot h(t) + \frac{\partial L}{\partial u}(t, x(t), \dot{x}(t)) \cdot \dot{h}(t) \right) dt.$$

Introduisons alors les formes linéaires $\alpha(t) := \partial_x L(t, x(t), \dot{x}(t))$ et $\beta(t) := \partial_u L(t, x(t), \dot{x}(t))$, de telle sorte que

$$J'(x) \cdot h = \int_a^b (\alpha(t) \cdot h(t) + \beta(t) \cdot \dot{h}(t)) dt, \quad (\text{B.2})$$

et supposons un instant que L et x soient de classe \mathcal{C}^2 . Alors dans ce cas, β est de classe \mathcal{C}^1 et à l'aide d'une simple intégration par parties, on peut écrire :

$$J'(x) \cdot h = [\beta \cdot h]_a^b + \int_a^b (\alpha(t) - \dot{\beta}(t)) \cdot h(t) dt.$$

Nous allons montrer dans ce qui suit que même si x et L ne sont que de classe \mathcal{C}^1 , et si x est solution du Problème (PCV), alors $\alpha(t) - \dot{\beta}(t) = 0$ pour tout $t \in [a, b]$. Nous aurons pour cela besoin du lemme suivant fondamental en calcul des variations :

Lemme B.3.1 (de Du Bois-Reymond, voir [65]). *Soient α et β dans $\mathcal{C}^0([a, b], (\mathbb{R}^n)^*)$. Supposons que*

$$\int_a^b (\alpha(t) \cdot h(t) + \beta(t) \cdot \dot{h}(t)) dt = 0$$

pour tout $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n) := \{h \in \mathcal{C}^1([a, b], \mathbb{R}^n) \mid h(a) = h(b) = 0_{\mathbb{R}^n}\}$. Alors, il vient que $\beta \in \mathcal{C}^1([a, b], (\mathbb{R}^n)^)$ et pour tout $t \in [a, b]$, on a $\dot{\beta}(t) = \alpha(t)$.*

► Soient $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$ et $c \in (\mathbb{R}^n)^*$. On peut montrer à l'aide d'une simple intégration par parties que :

$$0 = \int_a^b (\alpha(t) \cdot h(t) + \beta(t) \cdot \dot{h}(t)) dt = \int_a^b (\beta(t) - A(t) - c) \cdot \dot{h}(t) dt,$$

où $A(t) := \int_a^t \alpha(s) ds$. Ainsi, pour le choix particulier

$$c := \frac{1}{b-a} \int_a^b (\beta(t) - A(t)) dt,$$

et pour h défini par

$$h(t) := \int_a^t (\beta(s) - A(s) - c)^T ds,$$

2. $\mathcal{C}^1([a, b], \Omega)$ est un ouvert de l'espace vectoriel normé $\mathcal{C}^1([a, b], \mathbb{R}^n)$ car Ω est un ouvert de \mathbb{R}^n .

il vient que $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$ et que

$$0 = \int_a^b (\alpha(t) \cdot h(t) + \beta(t) \cdot \dot{h}(t)) dt = \int_a^b (\beta(t) - A(t) - c) \cdot \dot{h}(t) dt = \int_a^b \| \dot{h}(t) \|^2 dt.$$

Puisque \dot{h} est continue sur $[a, b]$ alors c'est la fonction nulle et ainsi pour tout $t \in [a, b]$, $\beta(t) = A(t) + c$, c'est-à-dire β appartient à $\mathcal{C}^1([a, b], (\mathbb{R}^n)^*)$ et $\dot{\beta}(t) = \alpha(t)$. ■

Nous pouvons maintenant donner nos premières conditions nécessaires d'optimalité.

Théorème B.3.2 – Euler-Lagrange (1744)

Supposons le lagrangien L de classe \mathcal{C}^1 . Si \bar{x} est une solution globale, locale forte ou locale faible du Problème (PCV), alors \bar{x} satisfait l'équation de Euler-Lagrange :

$$\forall t \in [a, b] : \frac{\partial L}{\partial x}(t, \bar{x}(t), \dot{\bar{x}}(t)) - \frac{d}{dt} \left(\frac{\partial L}{\partial u}(t, \bar{x}(t), \dot{\bar{x}}(t)) \right) = 0_{(\mathbb{R}^n)^*}. \quad (\text{B.3})$$

► Supposons que \bar{x} soit une solution faible pour un $\eta > 0$. Soit $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$. Alors, il existe $\rho_{\eta, h} > 0$ t.q. $\forall s \in]-\rho_{\eta, h}, \rho_{\eta, h}[:$

$$\bar{x} + sh \in B_w(\bar{x}, \eta) \cap \mathcal{K}.$$

Puisque \bar{x} est une solution faible alors $\forall s \in]-\rho_{\eta, h}, \rho_{\eta, h}[:$

$$J(\bar{x}) \leq J(\bar{x} + sh),$$

autrement dit, $f(s) := J(\bar{x} + sh)$ est une fonction \mathcal{C}^1 qui atteint son minimum en le point intérieur $\bar{s} := 0$. Ainsi,

$$f'(\bar{s}) = J'(\bar{x}) \cdot h = 0.$$

Puisque ceci est vrai pour tout $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$, alors le lemme B.3.1 avec l'expression (B.2) de la différentielle de J nous donne l'équation de Euler-Lagrange (B.3). Enfin, puisque ceci est vrai pour toute solution locale faible, alors c'est aussi vrai pour toute solution globale et locale forte, cf. Proposition B.2.3. ■

Remarque B.3.2. L'équation de Euler-Lagrange est présentée comme une égalité entre formes linéaires. On peut préférer la notation

$$\nabla_x L(t, x(t), \dot{x}(t)) - \frac{d}{dt} (\nabla_u L(t, x(t), \dot{x}(t))) = 0_{\mathbb{R}^n}.$$

Remarque B.3.3. Il est important de noter que l'équation de Euler-Lagrange n'est qu'une condition nécessaire comme le montre l'Exemple B.2.1. On peut noter de même que nous n'avons pas vraiment besoin de la différentiabilité au sens de Fréchet de la fonctionnelle J mais seulement au sens de Gâteaux, car nous n'avons réellement besoin que des dérivées directionnelles.

Exemple B.3.1. Prenons le temps d'un exemple des notations plus habituelles en théorie des équations aux dérivées partielles. On note u l'inconnue (qui remplace x) et x la dimension

d'espace (qui remplace t). Définissons le lagrangien suivant :

$$L(x, u, v) := \frac{1}{2}v^2 + \frac{1}{2}c(x)u^2 - f(x)u,$$

où c et f sont des fonctions suffisamment régulières. Si $u(\cdot)$ est solution du problème variationnel associé à L alors il vérifie les équations de Euler-Lagrange :

$$\frac{\partial}{\partial u}L(x, u(x), u'(x)) - \frac{d}{dx} \left(\frac{\partial}{\partial v}L(x, u(x), u'(x)) \right) = c(x)u(x) - f(x) - \frac{d}{dx}u'(x) = 0,$$

ainsi, si on impose les conditions aux limites (de Dirichlet) $u(0) = \alpha$ et $u(1) = \beta$, on trouve l'équation aux dérivées partielles stationnaire de dimension 1 suivante :

$$\begin{cases} -u''(x) + c(x)u(x) = f(x), \\ u(0) = \alpha, \quad u(1) = \beta \end{cases}$$

qui n'est rien d'autre qu'un problème aux deux bouts. \square

Définition B.3.3 – Extrémale et BC-extrémale

Une fonction $x \in \mathcal{C}^1([a, b], \Omega)$ qui vérifie Euler-Lagrange (B.3) est une **extrémale**.

Une extrémale qui vérifie les conditions aux limites, $x \in \mathcal{K}$, est une **BC-extrémale**.³

Exercice B.3.2 (cf. [27]).

- i) Montrer que les extrémales de $L(t, x, u) := \sqrt{1 + u^2}$ sont affine (*i.e.* $\ddot{x} = 0$).
- ii) Montrer que l'équation de Euler-Lagrange pour le lagrangien $L(t, x, u) := x^2 + u^2$ est donnée par $\ddot{x} - x = 0$.
- iii) Trouver l'unique extrémale admissible (*i.e.* BC-extrémale) pour le problème :

$$\inf_x \left\{ J(x) := \int_0^1 (\dot{x}^2(t) + x^2(t)) dt \mid x \in \mathcal{C}^1([0, 1], \mathbb{R}), x(0) = 0, x(1) = 1 \right\}. \quad \square$$

Nous travaillions jusqu'ici avec des lagrangiens de classe \mathcal{C}^1 et des extrémales associées elles-mêmes de classe \mathcal{C}^1 . Voyons maintenant comment la régularité du lagrangien influence celle des extrémales [27, 65].

Théorème B.3.4 – Hilbert-Weierstrass (1875)

Soient L un lagrangien de classe \mathcal{C}^k , $k \geq 2$, et x une extrémale t.q. :

$$\frac{\partial^2 L}{\partial u^2}(t, x(t), \dot{x}(t)) \text{ inversible,}$$

pour tout $t \in [a, b]$. Alors, x est de classe \mathcal{C}^k sur $]a, b[$.

3. "BC" vient de "Boundary Conditions" qui signifie conditions aux limites.

► Par hypothèse, x satisfait l'équation de Euler-Lagrange (B.3). Ainsi, il existe une constante $\lambda \in (\mathbb{R}^n)^*$ t.q. :

$$\int_a^t \frac{\partial L}{\partial x}(\tau, x(\tau), \dot{x}(\tau)) \, d\tau - \frac{\partial L}{\partial u}(t, x(t), \dot{x}(t)) = \lambda$$

pour tout $t \in [a, b]$. On pose

$$\begin{aligned} F:]a, b[\times \mathbb{R}^n &\longrightarrow (\mathbb{R}^n)^* \\ (t, u) &\longmapsto F(t, u) := \int_a^t \frac{\partial L}{\partial x}(\tau, x(\tau), \dot{x}(\tau)) \, d\tau - \frac{\partial L}{\partial u}(t, x(t), u), \end{aligned}$$

de telle sorte que pour $t \in]a, b[$, on a $F(t, \dot{x}(t)) = \lambda$ et

$$\frac{\partial F}{\partial u}(t, \dot{x}(t)) = -\frac{\partial^2 L}{\partial u^2}(t, x(t), \dot{x}(t))$$

qui est inversible. Ainsi, d'après le théorème des fonctions implicites, la solution $\dot{x}(t)$ est unique (localement) et \dot{x} a la même régularité que F . Si $k = 2$, il faut montrer que F est de classe \mathcal{C}^1 et pour cela, il suffit de montrer que $\partial_t F$ et $\partial_u F$ sont continues sur un voisinage de $(t, \dot{x}(t))$. Or,

$$\frac{\partial F}{\partial u}(t, u) = -\frac{\partial^2 L}{\partial u^2}(t, x(t), u)$$

est bien continue car L est de classe \mathcal{C}^2 et x est continue (car x est \mathcal{C}^1), et

$$\frac{\partial F}{\partial t}(t, u) = \frac{\partial L}{\partial x}(t, x(t), \dot{x}(t)) - \frac{\partial^2 L}{\partial t \partial u}(t, x(t), u) - \frac{\partial^2 L}{\partial x \partial u}(t, x(t), u) \cdot \dot{x}(t)$$

est de même continue. En conclusion, si $k = 2$ alors x est de classe \mathcal{C}^2 sur $]a, b[$. Maintenant si $k = 3$, on sait que x est \mathcal{C}^2 et par la même mécanique, on peut montrer que F est de classe \mathcal{C}^2 . En conclusion, par récurrence sur k , on peut montrer pour tout $k \geq 2$, que F est \mathcal{C}^{k-1} et donc x de classe \mathcal{C}^k sur $]a, b[$. ■

On déduit des équations de Euler-Lagrange la propriété suivante, utile d'un point de vue pratique pour calculer des extrémales.

Proposition B.3.5 – condition de Erdmann [27] ou identité de Beltrami

Soient un lagrangien L et une extrémale x tous deux de classe \mathcal{C}^2 . On pose

$$p(t) := \partial_u L[t] \quad \text{et} \quad E(t) := p(t) \cdot \dot{x}(t) - L[t],$$

où $[t] := (t, x(t), \dot{x}(t))$. Alors E est de classe \mathcal{C}^1 et pour tout $t \in [a, b]$, on a :

$$\dot{E}(t) = -\partial_t L[t],$$

ainsi donc, si L ne dépend pas explicitement de t (on dit que L est **autonome**) :

$$E(t) = p(t) \cdot \dot{x}(t) - L[t] = \text{const.}$$

► Cette preuve est laissée au lecteur. ■

Exemple B.3.3 (Problème de surface de révolution d'aire minimale⁴). Une caténoïde (du latin *Catena*, chaîne) se forme lorsque deux anneaux circulaires parallèles sont séparés lentement après avoir été plongés dans une solution savonneuse. C'est la surface minimale entre ces deux cercles. Elle fut découverte par Leonhard Euler en 1744 et l'équation de sa surface est la solution d'un problème variationnel. Modélisons ce problème. On va supposer que la solution est une surface dans \mathbb{R}^3 de révolution autour de l'axe des temps t , engendrée par une courbe donnée comme le graphe d'une fonction $x(t)$.

Le problème variationnel correspondant (on omet le facteur 2π dans le critère) s'écrit :

$$\inf_x \left\{ J(x) := \int_a^b x(t) \sqrt{1 + \dot{x}^2(t)} dt \mid x \in \mathcal{C}^1([a, b], \mathbb{R}), x(a) = A, x(b) = B \right\},$$

où a, b, A et B sont donnés, et où le lagrangien associé est :

$$L(t, x, u) := x \sqrt{1 + u^2}.$$

Le sous-graphe (a) de la figure B.3 est celui de la courbe solution $x(t)$ tandis qu'à droite, en (b), on peut voir la surface de révolution engendrée.

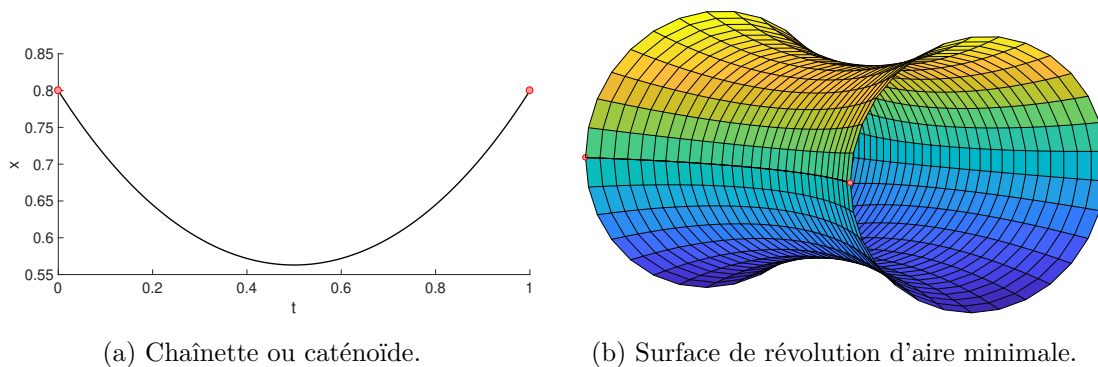


FIGURE B.3 – Illustration du problème de surface de révolution d'aire minimale. Dans cet exemple, $a = 0$, $b = 1$ et $A = B = 0.8$.

Supposons qu'il existe une solution x de classe \mathcal{C}^2 au problème. Attention, il n'existe pas de solutions pour tous les jeux de données (a, b, A, B) , il faut donc bien les choisir. De plus, puisque $x(t)$ correspond au rayon du cercle de notre surface de révolution au temps t , il est raisonnable de supposer que la solution vérifie $x(t) > 0$ pour tout temps t . Dans ces conditions, l'équation de Euler-Lagrange est donnée par

$$\ddot{x}(t) = \frac{1 + \dot{x}^2(t)}{x(t)},$$

et puisque le lagrangien est de classe \mathcal{C}^2 et autonome, d'après la Proposition B.3.5, il existe une constante $K > 0$ telle que

$$-E(t) = \frac{x(t)}{\sqrt{1 + \dot{x}^2(t)}} = K.$$

4. Voir <https://fr.wikipedia.org/wiki/Chaînette> et https://fr.wikipedia.org/wiki/Surface_minimale.

Utilisons cette dernière relation pour déterminer $x(t)$. Pour cela, changeons tout d'abord la paramétrisation du temps $t = \varphi(\tau)$ et définissons $\tilde{x}(\tau) := x(\varphi(\tau))$ de telle sorte que :

$$\frac{d\tilde{x}}{d\tau}(\tau) = \dot{x}(\varphi(\tau)) \varphi'(\tau), \quad (\text{B.4})$$

puis imposons (comme proposé dans [65]) :

$$\dot{x}(\varphi(\tau)) = \sinh \tau. \quad (\text{B.5})$$

Alors, $1 + \dot{x}^2(\varphi(\tau)) = 1 + \sinh^2 \tau = \cosh^2 \tau$ et donc :

$$\tilde{x}(\tau) = x(\varphi(\tau)) = K \sqrt{1 + \dot{x}^2(\varphi(\tau))} = K \cosh \tau.$$

En dérivant, on trouve

$$\frac{d\tilde{x}}{d\tau}(\tau) = K \sinh \tau$$

et à l'aide des équations (B.4) et (B.5), on obtient finalement $\varphi'(\tau) = K$ ou encore

$$\varphi(\tau) = K\tau + \beta,$$

pour une certaine constante β . En conclusion

$$x(t) = x(\varphi(\tau)) = \tilde{x}(\varphi^{-1}(t)) = K \cosh \left(\frac{t - \beta}{K} \right), \quad \beta \in \mathbb{R}, \quad K > 0,$$

ce qui est l'expression d'une caténoïde. □

B.4 Condition de Legendre

Supposons un instant que L soit un lagrangien de classe \mathcal{C}^2 . Toute extrémale x telle que

$$\frac{\partial^2 L}{\partial u^2}(t, x(t), \dot{x}(t)) \text{ inversible}$$

pour tout $t \in [a, b]$, vérifie alors l'équation différentielle d'ordre 2 :

$$0 = \frac{\partial L}{\partial x}[t] - \frac{d}{dt} \left(\frac{\partial L}{\partial u}[t] \right) = \frac{\partial L}{\partial x}[t] - \left(\frac{\partial^2 L}{\partial t \partial u}[t] + \frac{\partial^2 L}{\partial x \partial u}[t] \cdot \dot{x}(t) + \frac{\partial^2 L}{\partial u^2}[t] \cdot \ddot{x}(t) \right),$$

où $[t] := (t, x(t), \dot{x}(t))$, que l'on peut réécrire (avec une notation plus succincte mais claire) :

$$\ddot{x} = (L_{uu})^{-1} (L_x - L_{tu} - L_{xu} \dot{x})$$

car $\partial_{uu}^2 L[t]$ est inversible. Dans cette situation, \ddot{x} se prolonge de manière continue aux extrémités a et b et ainsi $x \in \mathcal{C}^2([a, b], \Omega)$. Nous voyons ici que $\partial_{uu}^2 L[t]$ joue un rôle important et nous allons montrer ci-après que $\partial_{uu}^2 L[t] \geq 0$ le long de toute solution faible. Avant cela, nous présentons le résultat suivant, similaire à la Proposition B.3.1.

Proposition B.4.1

Supposons le lagrangien L de classe \mathcal{C}^k , pour $k \in \bar{\mathbb{N}}^*$. Alors, l'application

$$\begin{aligned} J: (\mathcal{C}^1([a, b], \Omega), \|\cdot\|_{\mathcal{C}^1}) &\longrightarrow (\mathbb{R}, |\cdot|) \\ x &\longmapsto J(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt \end{aligned}$$

est de classe \mathcal{C}^k .

► La preuve est similaire à celle de la Proposition B.3.1 : le point clé est que l'application \mathcal{L} est de classe \mathcal{C}^k car L est de classe \mathcal{C}^k . Ensuite, toute application linéaire continue est \mathcal{C}^∞ ce qui permet de conclure via le théorème de dérivation des applications composées. ■

Nous avons déjà vu que la différentielle de J en $x \in \mathcal{C}^1([a, b], \Omega)$ appliquée à $h \in \mathcal{C}^1([a, b], \mathbb{R}^n)$ est donnée par :

$$J'(x) \cdot h = (\varphi \circ \mathcal{L}'(A(x)) \circ A) \cdot h = \varphi(\mathcal{L}'(A(x)) \cdot A(h)),$$

où φ , A et \mathcal{L} sont définies dans la démonstration de la Proposition B.3.1. Introduisons

$$g(x) := J'(x) \cdot h$$

et l'application linéaire continue $T(f) := f \cdot A(h)$ pour $f \in \mathcal{L}(\mathcal{C}^0([a, b], \mathbb{R}^n)^2, \mathcal{C}^0([a, b], \mathbb{R}))$, de telle sorte que

$$g = \varphi \circ T \circ \mathcal{L}' \circ A.$$

Nous voulons calculer la dérivée seconde de J en x appliquée en $(h, k) \in \mathcal{C}^1([a, b], \mathbb{R}^n)^2$, i.e.

$$J''(x) \cdot (h, k),$$

or,

$$\begin{aligned} J''(x) \cdot (h, k) &= (J'(\cdot) \cdot h)'(x) \cdot k = g'(x) \cdot k = (\varphi \circ T \circ \mathcal{L}''(A(x)) \circ A) \cdot k \\ &= \varphi(\mathcal{L}''(A(x)) \cdot (A(h), A(k))) \\ &= \varphi(\mathcal{L}''(x, \dot{x}) \cdot ((h, \dot{h}), (k, \dot{k}))) \\ &= \int_a^b (\mathcal{L}''(x, \dot{x}) \cdot ((h, \dot{h}), (k, \dot{k}))) (t) dt, \end{aligned}$$

et

$$\begin{aligned} \mathcal{L}''(x, \dot{x}) \cdot ((h, \dot{h}), (k, \dot{k})) &= \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \dot{x}) \cdot (h, k) + \frac{\partial^2 \mathcal{L}}{\partial u \partial x}(x, \dot{x}) \cdot (k, \dot{h}) \\ &\quad + \frac{\partial^2 \mathcal{L}}{\partial x \partial u}(x, \dot{x}) \cdot (\dot{k}, h) + \frac{\partial^2 \mathcal{L}}{\partial u^2}(x, \dot{x}) \cdot (\dot{h}, \dot{k}). \end{aligned}$$

Finalement, on obtient

$$\begin{aligned} J''(x) \cdot (h, k) &= \int_a^b \left(\frac{\partial^2 L}{\partial x^2}[t] \cdot (h(t), k(t)) + \frac{\partial^2 L}{\partial u \partial x}[t] \cdot (k(t), \dot{h}(t)) \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial x \partial u}[t] \cdot (\dot{k}(t), h(t)) + \frac{\partial^2 L}{\partial u^2}[t] \cdot (\dot{h}(t), \dot{k}(t)) \right) dt, \end{aligned}$$

avec $[t] := (t, x(t), \dot{x}(t))$, que l'on peut écrire sous la forme

$$J''(x) \cdot (h, k) = \int_a^b \left(\frac{\partial^2 L}{\partial(x, u)^2} [t] \cdot ((h(t), \dot{h}(t)), (k(t), \dot{k}(t))) \right) dt.$$

Nous pouvons maintenant donner une nouvelle condition nécessaire dans le cas d'un lagrangien de classe \mathcal{C}^2 . Avant cela, introduisons quelques notations. Une matrice $A \in M_n(\mathbb{R})$ vérifie $A \succeq 0$ si A est semi-définie positive. On note $A \succ 0$ si elle est définie positive.

Théorème B.4.2 – Condition de Legendre (1786)

Supposons le lagrangien L de classe \mathcal{C}^2 . Si \bar{x} est une solution globale, locale forte ou locale faible du Problème (PCV), alors \bar{x} satisfait la condition de Legendre :

$$\forall t \in [a, b] : \frac{\partial^2 L}{\partial u^2}(t, \bar{x}(t), \dot{\bar{x}}(t)) \succeq 0. \quad (\text{B.6})$$

► Supposons que \bar{x} soit une solution faible pour un $\eta > 0$. Soit $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$. Alors, il existe $\rho_{\eta, h} > 0$ t.q. $\forall s \in]-\rho_{\eta, h}, \rho_{\eta, h}[$:

$$\bar{x} + sh \in B_w(\bar{x}, \eta) \cap \mathcal{K}.$$

Puisque \bar{x} est une solution faible alors $\forall s \in]-\rho_{\eta, h}, \rho_{\eta, h}[$:

$$J(\bar{x}) \leq J(\bar{x} + sh),$$

autrement dit, $f(s) := J(\bar{x} + sh)$ est une fonction \mathcal{C}^2 qui atteint son minimum en le point intérieur $\bar{s} := 0$. Ainsi,

$$f'(\bar{s}) = J'(\bar{x}) \cdot h = 0 \quad \text{et} \quad f''(\bar{s}) = J''(\bar{x}) \cdot (h, h) \geq 0,$$

avec

$$J''(\bar{x}) \cdot (h, h) = \int_a^b \left(\frac{\partial^2 L}{\partial x^2} [t] \cdot (h(t))^2 + 2 \frac{\partial^2 L}{\partial u \partial x} [t] \cdot (h(t), \dot{h}(t)) + \frac{\partial^2 L}{\partial u^2} [t] \cdot (\dot{h}(t))^2 \right) dt,$$

où $[t] := (t, \bar{x}(t), \dot{\bar{x}}(t))$. Puisque $J''(\bar{x}) \cdot (h, h)$ doit être positif quel que soit $h \in \mathcal{C}_0^1([a, b], \mathbb{R}^n)$, il suffit de construire une fonction h telle que $\|h\|_\infty$ soit négligeable devant $\|\dot{h}\|_\infty$, car alors le terme $\partial^2 L_{uu} [t] \cdot (\dot{h}(t), \dot{h}(t))$ serait dominant et devrait donc être positif.

* **Cas $n = 1$.** Montrons le résultat par l'absurde et suivons [65]. Supposons qu'il existe un temps $\tau \in [a, b]$ tel que $\partial_{uu}^2 L[\tau] = -2\beta < 0$. Alors par continuité de $t \mapsto \partial_{uu}^2 L[t]$, on peut supposer que $\tau \in]a, b[$. Choisissons un $\varepsilon > 0$ suffisamment petit pour que $\partial^2 L_{uu} [\tau] < -\beta < 0$ pour tout $t \in [\tau - \varepsilon, \tau + \varepsilon]$ et définissons la fonction

$$h_\varepsilon(t) := \begin{cases} \sin^2\left(\frac{\pi}{\varepsilon}(t - \tau)\right) & \text{pour } |t - \tau| < \varepsilon, \\ 0 & \text{sinon.} \end{cases}$$

Puisque pour tout t tel que $|t - \tau| < \varepsilon$:

$$\dot{h}_\varepsilon(t) = \frac{2\pi}{\varepsilon} \sin\left(\frac{\pi}{\varepsilon}(t - \tau)\right) \cos\left(\frac{\pi}{\varepsilon}(t - \tau)\right) = \frac{\pi}{\varepsilon} \sin\left(\frac{2\pi}{\varepsilon}(t - \tau)\right),$$

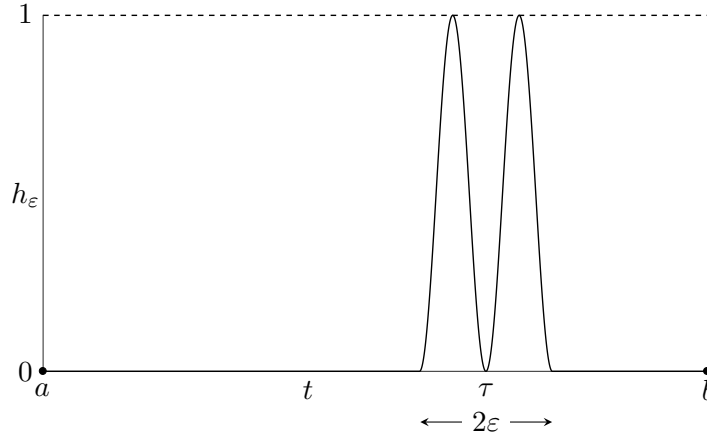
alors $h_\varepsilon \in \mathcal{C}_0^1([a, b], \mathbb{R})$ et

$$\begin{aligned} J''(\bar{x}) \cdot (h_\varepsilon, h_\varepsilon) &= \int_{\tau-\varepsilon}^{\tau+\varepsilon} \partial_{xx}^2 L[t] \sin^4 \left(\frac{\pi}{\varepsilon} (t - \tau) \right) dt \\ &\quad + \frac{2\pi}{\varepsilon} \int_{\tau-\varepsilon}^{\tau+\varepsilon} \partial_{ux}^2 L[t] \sin^2 \left(\frac{\pi}{\varepsilon} (t - \tau) \right) \sin \left(\frac{2\pi}{\varepsilon} (t - \tau) \right) dt \\ &\quad + \left(\frac{\pi}{\varepsilon} \right)^2 \int_{\tau-\varepsilon}^{\tau+\varepsilon} \partial_{uu}^2 L[t] \sin^2 \left(\frac{2\pi}{\varepsilon} (t - \tau) \right) dt. \end{aligned}$$

Introduisons M et N respectivement les bornes supérieures des valeurs absolues des fonctions continues $t \mapsto \partial_{xx}^2 L[t]$ et $t \mapsto \partial_{ux}^2 L[t]$ sur $[a, b]$, de telle sorte que

$$J''(\bar{x}) \cdot (h_\varepsilon, h_\varepsilon) \leq 2\varepsilon M + 4\pi N - 2\beta \frac{\pi^2}{\varepsilon} \rightarrow -\infty \quad \text{quand } \varepsilon \rightarrow 0^+.$$

Il est donc possible de choisir un $\varepsilon > 0$ tel que $J''(\bar{x}) \cdot (h_\varepsilon, h_\varepsilon) < 0$, d'où la contradiction et donc $\partial_{uu}^2 L[t]$ est positif sur $[a, b]$. Voici la visualisation de h_ε :



* **Cas général.** On suppose qu'il existe un temps $\tau \in]a, b[$ et un vecteur propre $v_\tau \in \mathbb{R}^n$ tel que $\partial_{uu}^2 L[\tau] \cdot (v_\tau, v_\tau) = -2\beta < 0$. Ensuite, on choisit un $\varepsilon > 0$ suffisamment petit pour que $\partial_{uu}^2 L[t] \cdot (v_\tau, v_\tau) < -\beta < 0$ pour tout $t \in [\tau - \varepsilon, \tau + \varepsilon]$ et on définit la fonction $h_\varepsilon(t) := \alpha_\varepsilon(t)v_\tau$ avec

$$\alpha_\varepsilon(t) := \begin{cases} \sin^2 \left(\frac{\pi}{\varepsilon} (t - \tau) \right) & \text{pour } |t - \tau| < \varepsilon, \\ 0 & \text{sinon.} \end{cases}$$

Le reste de la démonstration se généralise bien et au final on montre que l'on peut choisir $\varepsilon > 0$ tel que $J''(\bar{x}) \cdot (h_\varepsilon, h_\varepsilon) < 0$ ce qui est en contradiction avec les hypothèses et donc toutes les valeurs propres de $\partial_{uu}^2 L[t]$ sont positives sur $[a, b]$.

* Puisque ceci est vrai pour toute solution locale faible, alors c'est aussi vrai pour toute solution globale et locale forte, cf. Proposition B.2.3. ■

Remarque B.4.1. Attention, la *condition forte de Legendre* (ou stricte)

$$\forall t \in [a, b] : \frac{\partial^2 L}{\partial u^2}(t, \bar{x}(t), \dot{\bar{x}}(t)) \succ 0,$$

n'est pas une condition suffisante d'optimalité, cf. [27, Exercice 14.10]. Il faut une condition supplémentaire, la condition de Jacobi. Cependant, si la longueur de l'intervalle $[a, b]$ est suffisamment petite, alors la condition de Legendre forte est suffisante, car la condition de Jacobi est automatiquement vérifiée sur des temps courts.

B.5 Extrémum libre et contraintes affines

Nous avons dans les preuves des Théorèmes B.3.2 (Euler-Lagrange) et B.4.2 (Legendre) utilisé des notions d'optimisation que nous allons formaliser dans cette section.

B.5.1 Extrémum libre

Définition B.5.1

- i) Soient X un ensemble quelconque, $J: X \rightarrow \mathbb{R}$ et $\bar{x} \in X$. On dit que \bar{x} est un **minimum global** de J si $J(\bar{x}) \leq J(x)$ pour tout $x \in X$.
- ii) Si X est un espace topologique, on dit que \bar{x} est un **minimum local** de J si $\exists V \in \mathcal{V}_X(\bar{x})$ (i.e. un voisinage ouvert de \bar{x} dans X) t.q. $\forall x \in V : J(\bar{x}) \leq J(x)$.
- iii) Un minimum local est dit **strict** si : $\forall x \in V \setminus \{\bar{x}\} : J(\bar{x}) < J(x)$. On a de même les notions de maximum global, local et local strict.
- iv) On dit que \bar{x} est un **extremum** de J (respectivement global, local, local strict) si \bar{x} est un minimum ou maximum de J (respectivement global, local, local strict).
- v) Soient E un espace vectoriel normé (evn) et $J: U \subset E \rightarrow \mathbb{R}$, U un ouvert de E , J dérivable en $\tilde{x} \in U$. On dit que \tilde{x} est un **point critique** de J si $J'(\tilde{x}) = 0_{E'}$, et alors $J(\tilde{x})$ est appelée une **valeur critique**.

Remarque B.5.1.

- On rappelle que E' est l'ensemble des formes linéaires continues sur E et que la notion de dérivabilité est ici sous-entendue au sens de Fréchet. On notera $\|\cdot\|$ la norme de E .
- Un point critique n'est pas nécessairement un extrémum.
- En un point critique, on dit aussi que J est **stationnaire**.

Proposition B.5.2 – Condition du premier ordre (CN1)

Soient $J: U \subset E \rightarrow \mathbb{R}$ dérivable sur U ouvert de E evn et $\bar{x} \in U$. Si \bar{x} est un extrémum local de J alors l'équation de Euler est vérifiée :

$$J'(\bar{x}) = 0_{E'}.$$

► Si \bar{x} est par exemple un minimum (c'est similaire pour un max.) local alors :

$$\exists \eta > 0, \text{ t.q. } B(\bar{x}, \eta) \subset U \text{ et } \forall x \in B(\bar{x}, \eta) : J(\bar{x}) \leq J(x),$$

puisque U est un ouvert de E . Soit maintenant $h \in E$. Alors, il existe $\rho_{\eta, h} > 0$ t.q. :

$$\begin{aligned} f: V_{\eta, h} &:=]-\rho_{\eta, h}, \rho_{\eta, h}[\longrightarrow \mathbb{R} \\ s &\longmapsto f(s) := J(\bar{x} + sh) \end{aligned}$$

soit bien définie et t.q. $\bar{x} + sh \in B(\bar{x}, \eta)$, $\forall s \in V_{\eta, h}$. Si $h \neq 0_E$, prendre $\rho_{\eta, h} := \eta/\|h\|$. Ainsi définie, f est dérivable en 0 et $f'(0) = J'(\bar{x}) \cdot h$. Mais puisque \bar{x} est un minimum local de J , alors 0 est un minimum de f sur $V_{\eta, h}$. Ainsi, puisque $V_{\eta, h}$ est un ouvert de \mathbb{R} , il vient $f'(0) = J'(\bar{x}) \cdot h = 0$, car

$$0 \geq \lim_{s \rightarrow 0^-} \frac{f(s) - f(0)}{s} = f'(0) = \lim_{s \rightarrow 0^+} \frac{f(s) - f(0)}{s} \geq 0.$$

■

Proposition B.5.3 – Cas convexe (CNS)

Soient $J: U \subset E \rightarrow \mathbb{R}$ convexe et dérivable sur U un ouvert convexe de E , et $\bar{x} \in U$. Si \bar{x} est un point critique de J alors \bar{x} est un minimum global de J (sous-entendu sur U). Nous avons donc

$$\bar{x} \in U \text{ minimum global} \Leftrightarrow \bar{x} \in U \text{ point critique.}$$

► Puisque J est dérivable et convexe sur U alors $\forall (x, y) \in U^2 : J(x) - J(y) \geq J'(y) \cdot (x - y)$. En particulier, au point $y := \bar{x}$, on a $\forall x \in U : J(x) - J(\bar{x}) \geq J'(\bar{x}) \cdot (x - \bar{x}) = 0$, puisque \bar{x} est un point critique de J . Il vient donc que \bar{x} est un minimum global de J sur U . ■

Proposition B.5.4 – Condition nécessaire d'ordre supérieur (CNk)

Soit $J: U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $\bar{x} \in U$ un point critique de J t.q. $\forall i \in \llbracket 1, k-1 \rrbracket : J^{(i)}(\bar{x}) = 0$ et $J^{(k)}(\bar{x}) \neq 0$. On a :

$$\bar{x} \text{ minimum local de } J \Rightarrow k \text{ pair et } \forall h \in E : J^{(k)}(\bar{x}) \cdot h^k \geq 0.$$

► Supposons donc que \bar{x} soit un minimum local de J et considérons $h \in E$. On définit $f(s)$ comme dans la preuve de la Proposition B.5.2 et on garde les mêmes notations. Alors, $\forall s \in V_{\eta, h}$:

$$f(s) - f(0) = J(\bar{x} + sh) - J(\bar{x}) \geq 0.$$

Puisque J est k fois dérivable sur U alors f l'est sur $V_{\eta, h}$ donc en particulier en 0. D'après la formule de Taylor-Young :

$$f(s) - f(0) = \sum_{i=1}^k \frac{s^i}{i!} f^{(i)}(0) + s^k \varepsilon(s) \quad \text{avec} \quad \lim_{s \rightarrow 0} \varepsilon(s) = 0.$$

Par hypothèse :

$$f(s) - f(0) = \frac{s^k}{k!} f^{(k)}(0) + s^k \varepsilon(s) = s^k \left(\frac{1}{k!} J^{(k)}(\bar{x}) \cdot h^k + \varepsilon(s) \right) \geq 0,$$

En divisant par s^k et en passant à la limite quand $s \rightarrow 0^+$ dans le terme de droite on en déduit que $J^{(k)}(\bar{x}) \cdot h^k \geq 0$. Montrons pour finir que k est pair. Puisque $J^{(k)}(\bar{x}) \neq 0$, alors $\exists h_0 \in E$ t.q. $J^{(k)}(\bar{x}) \cdot h_0^k \neq 0$ ⁵ et donc $J^{(k)}(\bar{x}) \cdot h_0^k > 0$. Posons

$$\Omega := \left\{ s \in \mathbb{R}^* \mid |\varepsilon(s)| < \frac{1}{k!} J^{(k)}(\bar{x}) \cdot h_0^k \right\} \cap V_{\eta, h_0},$$

de telle sorte que $\forall s \in \Omega$:

$$\text{sign}(f(s) - f(0)) = \text{sign}(s^k) \times \text{sign} \left(J^{(k)}(\bar{x}) \cdot h_0^k \right) = \text{sign}(s^k).$$

Puisque $f(s) - f(0) \geq 0$, $\Omega \cap \mathbb{R}_- \neq \emptyset$ et $\Omega \cap \mathbb{R}_+ \neq \emptyset$, on en déduit que k est pair. ■

Proposition B.5.5 – Conditions suffisantes d'ordre supérieur (CSk)

Soit $J: U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $\bar{x} \in U$ un point critique de J t.q. $\forall i \in \llbracket 1, k-1 \rrbracket : J^{(i)}(\bar{x}) = 0$ et $J^{(k)}(\bar{x}) \neq 0$. On a :

i) Si

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E : J^{(k)}(\bar{x}) \cdot h^k \geq \alpha \|h\|^k,$$

alors k est pair et \bar{x} est un minimum local **strict**.

ii) Si

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B(\bar{x}, \eta) \subset U \quad \text{et} \quad \forall h \in E : J^{(k)}(x) \cdot h^k \geq 0,$$

alors k est pair et \bar{x} est un minimum local.

► i) Pour tout $h \in E$, $J^{(k)}(\bar{x}) \cdot (-h)^k = (-1)^k J^{(k)}(\bar{x}) \cdot h^k \geq \alpha \|h\|^k \geq 0$ donc k doit être pair. La formule de Taylor appliquée à J nous donne pour $h \in E$ t.q. $\bar{x} + h \in U$:

$$J(\bar{x} + h) - J(\bar{x}) = \frac{1}{k!} J^{(k)}(\bar{x}) \cdot h^k + \|h\|^k \varepsilon(h) \quad \text{avec} \quad \lim_{\|h\| \rightarrow 0} |\varepsilon(h)| = 0.$$

Par hypothèse :

$$J(\bar{x} + h) - J(\bar{x}) \geq \left(\frac{\alpha}{k!} + \varepsilon(h) \right) \|h\|^k$$

et donc il existe $\eta > 0$ tel que pour tout $h \neq 0_E$ t.q. $\|h\| < \eta$ on ait $\bar{x} + h \in U$ et

$$J(\bar{x} + h) - J(\bar{x}) > 0.$$

Autrement dit, \bar{x} est un minimum local strict.

ii) Il est évident que k soit pair car $J^{(k)}(\bar{x}) \neq 0$ (cf. Proposition précédente). Par hypothèse, J est donc convexe sur la boule ouverte $B(\bar{x}, \eta)$ et puisque \bar{x} est un point critique,

5. Voir [40, Chapitre IV, § 3, Problème 3, page 92].

alors \bar{x} est un minimum global de J sur $B(\bar{x}, \eta)$, autrement dit un minimum local de J , cf. Proposition B.5.3. ■

Définition B.5.6

Soit $B \in \mathcal{L}_2(E^2, \mathbb{R})$ une forme bilinéaire symétrique définie sur $(E, \|\cdot\|)$ un espace vectoriel normé.

- B est dite **semi-définie positive**, notée $B \succeq 0$, si $\forall h \in E : B(h, h) \geq 0$.
- B est dite **définie positive**, notée $B \succ 0$, si $\forall h \in E \setminus \{0_E\} : B(h, h) > 0$.
- B est dite **elliptique** si $\exists \alpha > 0$ t.q. $\forall h \in E : B(h, h) \geq \alpha \|h\|^2$.

D'après la Proposition B.5.4, en un point critique \bar{x} , une condition nécessaire d'ordre deux est donc que $J''(\bar{x})$ soit semi-définie positive, et d'après la Proposition B.5.5, une condition suffisante d'ordre deux est que $J''(\bar{x})$ soit elliptique :

$$\begin{aligned} \text{(CN2)} \quad & \forall h \in E : J''(\bar{x}) \cdot (h, h) \geq 0, \\ \text{(CS2)} \quad & \exists \alpha > 0 \text{ t.q. } \forall h \in E : J''(\bar{x}) \cdot (h, h) \geq \alpha \|h\|^2. \end{aligned} \tag{B.7}$$

Une autre condition suffisante d'ordre deux est que $J''(x)$ soit semi-définie positive pour tout x contenu dans une boule ouverte centrée en \bar{x} . Attention, la définie positivité ne suffit pas en général en dimension infinie mais suffit en dimension finie comme le montre le résultat suivant.

Proposition B.5.7 – Cas particulier de la dimension finie

Soit $J : U \subset E \rightarrow \mathbb{R}$ une application $k \geq 2$ fois dérivable sur U un ouvert de E (evn) et soit $x \in U$. Si $\dim E \in \mathbb{N}$ alors :

$$\exists \alpha > 0 \text{ t.q. } \forall h \in E : J^{(k)}(x) \cdot h^k \geq \alpha \|h\|^k \Leftrightarrow \forall h \in E \setminus \{0_E\} : J^{(k)}(x) \cdot h^k > 0.$$

► La conditions nécessaire est évidente et vraie même si $\dim E = +\infty$. Supposons donc que $\forall h \in E \setminus \{0_E\} : J^{(k)}(x) \cdot h^k > 0$. Introduisons

$$\begin{aligned} \varphi : E &\longrightarrow \mathbb{R} & \psi : E &\longrightarrow E^k \\ h &\longmapsto \varphi(h) := f^{(k)}(x) \cdot h^k & \text{et} & & h &\longmapsto \psi(h) = (h, \dots, h) = h^k \end{aligned}$$

de telle sorte que $\varphi = f^{(k)}(x) \circ \psi$, avec ψ une application linéaire continue. Puisque $f^{(k)}(x) \in \mathcal{L}_k(E^k, \mathbb{R})$ est aussi continue, alors φ est continue. Or, $\dim E < +\infty$, donc sa sphère unité $S(0_E, 1)$ est compacte, donc φ est bornée sur $S(0_E, 1)$ et atteint sa borne inférieure, c'est-à-dire :

$$\exists h_0 \in S(0_E, 1) \text{ t.q. } \forall h \in S(0_E, 1) : \varphi(h_0) \leq \varphi(h).$$

Notons $\alpha := f^{(k)}(x) \cdot h_0^k > 0$ (car $h_0 \neq 0_E$). Alors, $\forall h \in S(0_E, 1)$ on a $f^{(k)}(x) \cdot h^k \geq \alpha$ et donc par homogénéité, on en déduit :

$$\forall h \in E \setminus \{0_E\} : \frac{h}{\|h\|} \in S(0_E, 1) \Rightarrow f^{(k)}(x) \cdot \left(\frac{h}{\|h\|} \right)^k \geq \alpha \Rightarrow f^{(k)}(x) \cdot h^k \geq \alpha \|h\|^k.$$

■

Remarque B.5.2. Dans les preuves des propositions B.5.2 (CN1) et B.5.4 (CNk), nous nous sommes ramené dans \mathbb{R} , et nous avons utilisé directement les conditions nécessaires pour une fonctionnelle dans \mathbb{R} pour obtenir les conditions nécessaires en dimension infinie. Ce procédé ne fonctionne pas pour obtenir les conditions suffisantes de la Proposition B.5.5 (CSk). En effet, même si \bar{x} est un minimum local le long de n'importe quelle direction h , c'est-à-dire un minimum local de $f(s) = J(\bar{x} + sh)$, alors il est toujours possible que \bar{x} ne soit pas un minimum local de J , et ce, même en dimension finie (supérieure ou égale à 2), cf. exercice B.5.1. De plus, comme nous l'avons vu, il y a une différence notable entre les dimensions finie et infinie concernant les conditions suffisantes. La définie positivité de la dérivée seconde n'est pas une condition suffisante d'optimalité locale en dimension infinie, cf. Exemple B.5.2. C'est pour ces raisons que nous avons reporté la formalisation des concepts d'optimisation. Dans les sections précédentes, pour obtenir les équations de Euler-Lagrange et la condition de Legendre, seules les conditions nécessaires étaient utiles. De plus, comme nous allons vite le voir, ces conditions suffisantes ne nous seront pas utiles, c'est pourquoi nous allons en donner d'autres dans la prochaine section.

Voici un exercice montrant que l'optimalité "directionnelle" n'implique pas l'optimalité.

Exercice B.5.1 ([46]). On considère la fonction

$$\begin{aligned} J: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ x &\longmapsto J(x) := 2x_1^4 - 4x_1^2x_2 + x_2^2. \end{aligned}$$

1. Montrer que l'unique point critique de J est $\tilde{x} := (0, 0)$.
2. La condition nécessaire de solution du deuxième ordre est-elle vérifiée ?
3. Les conditions suffisantes de solution du deuxième ordre sont-elles vérifiées ?
4. On fixe $h \in \mathbb{R}^2$ et on considère la fonction

$$\begin{aligned} f: \mathbb{R} &\longrightarrow \mathbb{R} \\ s &\longmapsto f(s) := J(\tilde{x} + sh). \end{aligned}$$

Montrer que $\bar{s} = 0$ est un minimum local de f .

5. Calculer $J(x_1, 2x_1^2)$. Conclusion (faire le lien avec la remarque B.5.2) ? □

L'exemple suivant montre que la définie positivité de la dérivée seconde n'est pas une condition suffisante d'optimalité locale.

Exemple B.5.2 ([17]). On considère le problème

$$\inf_x \left\{ J(x) := \sum_{n=1}^{\infty} \left(\frac{x_n^2}{n} - x_n^3 \right) \mid x \in l_2 \right\},$$

où $l_2 := \{(x_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ est l'espace de Hilbert des suites de carrés sommables. Alors, le point $\tilde{x} := 0$ est un point critique tel que pour tout $h := (h_n) \in l_2$ non nul :

$$J''(\tilde{x}) \cdot (h, h) = \sum_{n=1}^{\infty} \frac{2h_n^2}{n} > 0.$$

Cependant, considérons les suites $h_k := (h_{k,n})$ telles que toutes les coordonnées sont nulles sauf la k^{e} qui vérifie $h_{k,k} = 2k^{-1}$. Alors, $J(\tilde{x} + h_k) = J(h_k) = -4k^{-3} < 0 = J(\tilde{x})$ et donc \tilde{x} n'est pas localement optimal. Voici donc un exemple de dérivée seconde définie positive mais non elliptique. □

B.5.2 Extremum lié : contraintes affines

Considérons un espace vectoriel normé $(E, \|\cdot\|_E)$ et une fonctionnelle $J: E \rightarrow \mathbb{R}$. Considérons de plus un autre espace vectoriel normé $(F, \|\cdot\|_F)$, une application linéaire continue $T \in \mathcal{L}(E, F)$ et un point $y \in F$. On s'intéresse au problème aux contraintes affines suivant :

$$\inf_x \{J(x) \mid x \in \mathcal{K}\}, \quad (P_A)$$

où $\mathcal{K} := \{x \in E \mid \Phi(x) := T(x) - y = 0_F\} = \Phi^{-1}(\{0_F\})$ avec $\Phi: E \rightarrow F$ une application affine continue de E dans F . Supposons que $y \in \text{Im } T$, c'est-à-dire que :

$$\exists u \in E \text{ t.q. } T(u) = y, \text{ i.e. t.q. } \Phi(u) = 0_F.$$

Introduisons alors

$$E_0 := \text{Ker } T \subset E$$

de telle sorte que $\forall v \in E_0 : \Phi(u+v) = T(u+v) - y = T(u) + T(v) - y = \Phi(u) + T(v) = 0_F$, c'est-à-dire :

$$\Phi(u+v) = 0_F \Leftrightarrow v \in E_0.$$

On munit E_0 de la norme induite par la norme de E , i.e. $\|v\|_{E_0} = \|v\|_E$ pour tout $v \in E_0$. Considérons alors la fonctionnelle

$$\begin{aligned} F: E_0 &\longrightarrow \mathbb{R} \\ v &\longmapsto F(v) := J(u+v), \end{aligned}$$

de telle sorte que le Problème (P_A) est équivalent au problème sans contraintes suivant :

$$\inf_{v \in E_0} F(v). \quad (P)$$

Ici, "équivalent" signifie : $\bar{x} := u + \bar{v}$ solution de $(P_A) \Leftrightarrow \bar{v}$ solution de (P) .

Remarque B.5.3. Si on restreint la recherche de la solution de (P_A) à un ouvert $U \subset E$, c'est-à-dire si on s'intéresse au problème

$$\inf_x \{J(x) \mid x \in \mathcal{K}\}, \quad \mathcal{K} := \Phi^{-1}(\{0_F\}) \cap U,$$

alors ce problème est équivalent au problème sans contraintes :

$$\inf_v \{F(v) \mid v \in U_0 \subset E_0\},$$

où $U_0 := E_0 \cap U$ est un ouvert de E_0 . On est alors précisément dans le cadre de l'item v) de la définition B.5.1.

Écrit de manière schématique, les conditions nécessaires et suffisantes d'ordres 1 et 2 pour le problème avec contraintes affines (P_A) sont donc :

$$\begin{aligned} (\text{CN1}) \quad & F'(\bar{v}) = J'(\bar{x}) = 0_{E_0'}, \text{ i.e. } \forall h \in E_0 : J'(\bar{x}) \cdot h = 0 ; \\ (\text{CN2}) \quad & \forall h \in E_0 : F''(\bar{v}) \cdot (h, h) = J''(\bar{x}) \cdot (h, h) \geq 0 ; \\ (\text{CS2}) \quad & \exists \alpha > 0 \text{ t.q. } \forall h \in E_0 : F''(\bar{v}) \cdot (h, h) = J''(\bar{x}) \cdot (h, h) \geq \alpha \|h\|_{E_0}^2 ; \end{aligned} \quad (\text{B.8})$$

où bien sûr $\bar{x} := u + \bar{v}$ doit être un point admissible, c'est-à-dire $\bar{x} \in \mathcal{K}$.

Application au calcul des variations. Nous avons déjà utilisé sans le savoir les CN1 et CN2 pour le problème classique du calcul des variations, *i.e.* pour le Problème (PCV), que l'on peut écrire sous la forme :

$$\inf_x \{J(x) \mid x \in \mathcal{K}\},$$

avec $\mathcal{K} := \Phi^{-1}(\{0_{\mathbb{R}^{2n}}\}) \cap U$, où $\Phi: E \rightarrow \mathbb{R}^{2n}$ est définie par

$$\Phi(x) := (x(a) - A, x(b) - B) =: T(x) - y := (x(a), x(b)) - (A, B),$$

et où $U := \mathcal{C}^1([a, b], \Omega)$ est un ouvert de l'espace vectoriel normé $E := (\mathcal{C}^1([a, b], \mathbb{R}^n), \|\cdot\|_{\mathcal{C}^1})$. On rappelle que Ω est un ouvert de \mathbb{R}^n et donc U est un ouvert de E . Le Problème (PCV) est alors équivalent au problème sans contraintes :

$$\inf_v \{F(v) \mid v \in U_0 \subset E_0\},$$

où $E_0 := \mathcal{C}_0^1([a, b], \mathbb{R}^n) = \text{Ker } T$ et $U_0 := E_0 \cap U = \mathcal{C}_0^1([a, b], \Omega)$. Pour trouver l'équation de Euler-Lagrange, cf. Théorème B.3.2, nous avons utilisé la CN1 :

$$\forall h \in E_0 : J'(\bar{x}) \cdot h = 0$$

et pour trouver la condition de Legendre, cf. Théorème B.4.2, nous avons utilisé la CN2

$$\forall h \in E_0 : J''(\bar{x}) \cdot h^2 \geq 0.$$

B.5.3 “Two-norm discrepancy”

Dans la Section B.6 sur les conditions de Jacobi, nous allons utiliser des conditions suffisantes du deuxième ordre. Cependant, comme nous allons le voir dans l'exemple suivant, la condition suffisante (CS2) présentée ci-avant n'est pas toujours utile en pratique pour des problèmes en dimension infinie, notamment en calcul des variations.

Exemple B.5.3. On s'intéresse au problème suivant :

$$\inf_x \left\{ J(x) := \int_0^1 \sin(\dot{x}(t)) dt \mid x \in \mathcal{C}^1([0, 1], \mathbb{R}), x(0) = 0, x(1) = -\frac{\pi}{2} \right\},$$

dont le lagrangien associé est donné par $L(t, x, u) := \sin(u)$. Il est facile de voir que la fonction définie par $\bar{x}(t) := -\frac{\pi}{2}t$ est la solution globale du problème et vérifie

$$J(\bar{x}) = -1.$$

La CN1 est bien vérifiée car :

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J'(\bar{x}) \cdot h = \int_0^1 \cos(\dot{\bar{x}}(t)) \dot{h}(t) dt = 0.$$

La CN2 est aussi vérifiée car :

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = - \int_0^1 \sin(\dot{\bar{x}}(t)) \dot{h}^2(t) dt = \int_0^1 \dot{h}^2(t) dt = \|\dot{h}\|_{L^2}^2 \geq 0,$$

où pour toute fonction mesurable (sous-entendu pour la mesure de Lebesgue) $f: [0, 1] \rightarrow \mathbb{R}$, on a défini :

$$\|f\|_{L^2} := \left(\int_0^1 |f(t)|^2 dt \right)^{\frac{1}{2}} \in \mathbb{R}_+,$$

de telle sorte que $\|\cdot\|_{L^2}$ est la norme usuelle associée à l'espace de Hilbert $L^2([0, 1], \mathbb{R})$ des (classes de) fonctions mesurables de $[0, 1]$ à valeurs dans \mathbb{R} de carrés intégrables, c'est-à-dire :

$$L^2([0, 1], \mathbb{R}) := \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ mesurable et } \|f\|_{L^2} < +\infty\}.$$

En revanche, la CS2 d'ellipticité de la dérivée seconde $J''(\bar{x})$ n'est pas vérifiée car on a dans notre exemple :

$$\|\dot{h}\|_{L^2} \leq \|\dot{h}\|_{\infty} \leq \|h\|_{\mathcal{C}^1} = \|h\|_{\infty} + \|\dot{h}\|_{\infty},$$

mais pas l'inverse, c'est-à-dire, il n'existe pas de $\alpha > 0$ tel que $\|\dot{h}\|_{L^2} \geq \alpha \|h\|_{\mathcal{C}^1}$. \square

Remarque B.5.4. Dans l'exemple précédent, nous ne pouvons pas montrer que \bar{x} est un minimum local strict en utilisant la condition suffisante d'ordre deux d'ellipticité de $J''(\bar{x})$. Cependant, il est aisé de montrer que \bar{x} est un minimum local à l'aide de la condition suffisante de semi-définie positivité de $J''(x)$ pour tout $x \in B(\bar{x}, \eta)$, pour un certain $\eta > 0$. Il suffit de prendre $\eta := \pi/2$ par exemple (la preuve est laissée en exercice).

Poursuivons avec l'exemple précédent. Nous avons donc

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2.$$

Ce qui est gênant ici, c'est qu'il y a dans le terme de gauche h tandis que dans celui de droite, on retrouve \dot{h} . On aimerait avoir quelque chose qui ressemble à :

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) \subset H_0 : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_{H_0}^2,$$

où $\|\cdot\|_{H_0}$ est une norme à définir sur un espace H_0 lui-même à définir. Introduisons pour cela l'espace de Sobolev $H^1([0, 1], \mathbb{R}) := W^{1,2}([0, 1], \mathbb{R})$ des fonctions f dans $L^2([0, 1], \mathbb{R})$ qui s'écrivent ⁶ sous la forme $f(t) = f(0) + \int_0^t g(s) ds$ avec $g \in L^2([0, 1], \mathbb{R})$, c'est-à-dire :

$$H^1([0, 1], \mathbb{R}) := \left\{ f \in L^2([0, 1], \mathbb{R}) \mid \exists g \in L^2([0, 1], \mathbb{R}) \text{ t.q. } f(t) = f(0) + \int_0^t g(s) ds \right\}.$$

Il est important de rappeler que l'application g est unique (modulo le fait que l'on identifie deux fonctions intégrables qui coïncident p.p. = presque partout = sauf sur un ensemble négligeable). Ainsi défini, nous avons

$$\mathcal{C}^1([0, 1], \mathbb{R}) \subset H^1([0, 1], \mathbb{R}) \subset \mathcal{C}^0([0, 1], \mathbb{R}),$$

autrement dit, toute fonction de H^1 (sous-entendu $H^1([0, 1], \mathbb{R})$) est continue mais possède moins de régularité que les fonctions de classe \mathcal{C}^1 . En effet, les fonctions de H^1 ne possède

6. On rappelle qu'en réalité, $W^{1,2}([0, 1], \mathbb{R})$ est un ensemble de classes de fonctions et non de fonctions. Nous ne donnons pas ici sa définition, mais plutôt une propriété. En effet, d'après [22, Théorème VIII.2], il existe un unique représentant continu pour toute classe dans $W^{1,2}([0, 1], \mathbb{R})$ que nous appelons ici f .

pas de dérivée continue puisqu'elles ne sont même pas dérivables en tout point $t \in [0, 1]$. Les fonctions de H^1 sont dérivables presque partout sur $[0, 1]$ et on a :

$$\forall f \in H^1 : f'(t) = g(t) \text{ p.p. sur } [0, 1].$$

On fera donc clairement l'abus de notation $f' = g$ et puisque l'on dérive par rapport à la variable t qui correspond au "temps" pour nous, on écrira plutôt

$$\dot{f} = g.$$

Remarque B.5.5. Nous pouvons voir g comme une dérivée en un sens plus faible que d'ordinaire. Cependant, si g est de classe \mathcal{C}^0 , alors la notion de dérivée faible coïncide avec celle de dérivée classique et f est alors de classe \mathcal{C}^1 . Ainsi, nous pouvons voir les fonctions de $H^1 = W^{1,2}$ comme des primitives de fonctions L^2 , de la même manière que nous pouvons voir les fonctions de \mathcal{C}^1 comme des primitives de fonctions \mathcal{C}^0 , et nous avons pour rappel :

$$\mathcal{C}^1([0, 1], \mathbb{R}) \subset \mathcal{C}^0([0, 1], \mathbb{R}) \quad \text{et} \quad H^1([0, 1], \mathbb{R}) \subset L^2([0, 1], \mathbb{R}).$$

Il nous faut maintenant une norme sur H^1 . Il existe deux possibilités équivalentes suivant si l'on considère la norme usuelle de $W^{1,2}$:

$$\|f\|_{W^{1,2}} := \|f\|_{L^2} + \|\dot{f}\|_{L^2},$$

ou si l'on considère la norme

$$\|f\|_{H^1} := (\|f\|_{L^2}^2 + \|\dot{f}\|_{L^2}^2)^{\frac{1}{2}}$$

induite par le produit scalaire usuel sur H^1 :

$$(f_1 | f_2)_{H^1} := (f_1 | f_2)_{L^2} + (\dot{f}_1 | \dot{f}_2)_{L^2},$$

où $(\cdot | \cdot)_{L^2}$ est le produit scalaire usuel sur L^2 . Les deux normes ainsi définies sont équivalentes et on peut noter que l'espace H^1 muni du produit scalaire $(\cdot | \cdot)_{H^1}$ forme un espace de Hilbert.

Exemple B.5.4 (suite). Revenons à l'exemple précédent. Nous avons donc

$$\forall h \in \mathcal{C}_0^1([0, 1], \mathbb{R}) : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2.$$

Rappelons que le problème classique du calcul des variations est un problème d'optimisation avec des contraintes affines et nous avons vu dans la Section B.5.2 comment définir un problème équivalent sans contraintes. Appliquons les notations du cadre général de la Section B.5.2 à cet exemple. Nous nous intéressons donc au problème :

$$\inf_x \{J(x) \mid x \in \mathcal{K}\},$$

avec $\mathcal{K} = \Phi^{-1}(\{0_{\mathbb{R}^2}\}) \subset E$, $E = \mathcal{C}^1([0, 1], \mathbb{R})$ et où $\Phi(x) = T(x) - y$ définit les contraintes affines. On a dans cet exemple $T(x) = (x(0), x(1))$ et $y = (0, -\pi/2)$. Pour obtenir de nouvelles conditions suffisantes d'optimalité locale, nous venons d'introduire ci-avant l'espace $H :=$

$H^1([0, 1], \mathbb{R})$ de telle sorte que $E \subset H$. Dans l'optique de se ramener à un problème sans contraintes, nous avons introduit

$$E_0 = E \cap \text{Ker } T = \mathcal{C}_0^1([0, 1], \mathbb{R}).$$

Toujours dans cette optique, nous introduisons :

$$H_0 := H \cap \text{Ker } T = H_0^1([0, 1], \mathbb{R}) := \{h \in H^1([0, 1], \mathbb{R}) \mid h(a) = h(b) = 0\},$$

de telle sorte que

$$E_0 \subset H_0.$$

L'idée générale est d'introduire un espace vectoriel normé plus grand que l'espace de travail. Ici, en raison des contraintes affines, l'espace de travail est E_0 et nous avons introduit H_0 (nous utilisons cette notation pour mettre un cadre général).

Remarque B.5.6. Pour être un peu plus précis, l'application linéaire continue⁷ T appartient à $\mathcal{L}(H, \mathbb{R}^2)$ au lieu de $\mathcal{L}(E, \mathbb{R}^2)$ comme dans la Section B.5.2. Nous devons tenir compte ici du rôle du nouvel espace H .

L'espace H_0 est muni de la norme induite par celle de H , *i.e.* $\|h\|_{H_0} = \|h\|_H$ et d'après le lemme B.5.1 (ci-après), on a :

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E_0 = \mathcal{C}_0^1([0, 1], \mathbb{R}) \subset H_0 : J''(\bar{x}) \cdot h^2 = \|\dot{h}\|_{L^2}^2 \geq \alpha \|h\|_{H_0}^2. \quad (\text{B.9})$$

Ceci est l'objectif voulu. Car, d'après la Proposition B.5.8 (ci-après), on peut conclure que \bar{x} est un minimum local (faible) strict, c'est-à-dire :

$$\exists \eta > 0 \quad \text{t.q.} \quad \forall x \in B_E(\bar{x}, \eta) \cap \mathcal{K}, \quad x \neq \bar{x} : J(\bar{x}) < J(x),$$

où $B_E(\bar{x}, \eta) := \{x \in E \mid \|x - \bar{x}\|_E < \eta\}$. □

Remarque B.5.7. C'est cette idée d'utiliser deux normes, $\|\cdot\|_E$ pour la différentiabilité et $\|\cdot\|_H$ pour la **condition de croissance** $J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_H^2$ (qui remplace la condition d'ellipticité), que l'on nomme “two-norm discrepancy”.

Lemme B.5.1. Soit $a < b$ dans \mathbb{R} . Alors,

$$\exists c > 0 \quad \text{t.q.} \quad \forall h \in H_0^1([a, b], \mathbb{R}) : \|h\|_{H_0^1} \leq c \|\dot{h}\|_{L^2}.$$

► Soit $h \in H_0^1([a, b], \mathbb{R})$. Alors, puisque $h(a) = 0$ et d'après l'inégalité de Cauchy-Schwarz [22] :

$$h^2(t) = \left(\int_a^t \dot{h}(s) \, ds \right)^2 \leq (t - a) \int_a^t \dot{h}^2(s) \, ds \leq (b - a) \int_a^b \dot{h}^2(s) \, ds = (b - a) \|\dot{h}\|_{L^2}^2.$$

Ainsi,

$$\|h\|_{L^2}^2 = \int_a^b h^2(t) \, dt \leq (b - a)^2 \|\dot{h}\|_{L^2}^2,$$

7. L'application T est bien continue car il existe $C > 0$ t.q. $\|T(x)\|_{\mathbb{R}^2} \leq \|x\|_\infty \leq C \|x\|_{H^1}$ puisque H_1 s'injecte de manière continue dans \mathcal{C}^0 , cf. [22]. On peut noter que puisque T est continue, alors H_0^1 est un fermé de l'espace de Hilbert H^1 , donc lui-même un Hilbert, si on le munit du produit scalaire de H^1 .

donc

$$\|h\|_{H_0^1}^2 = \|h\|_{H^1}^2 = \|h\|_{L^2}^2 + \|\dot{h}\|_{L^2}^2 \leq (1 + (b-a)^2) \|\dot{h}\|_{L^2}^2$$

et la conclusion suit avec $c := \sqrt{1 + (b-a)^2} > 0$. ■

Proposition B.5.8 – Condition suffisante d'ordre 2 généralisée (CS2G)

Soient $(E, \|\cdot\|_E)$ et $(H, \|\cdot\|_H)$ deux espaces vectoriels normés t.q. $E \subsetneq H$, i.e. $E \subset H$ et $\exists c > 0$ t.q. $\forall h \in E : \|h\|_H \leq c \|h\|_E$. Soit $J : U \subset E \rightarrow \mathbb{R}$ une application 2 fois dérivable sur U un ouvert de E et soit $\bar{x} \in U$ un point critique de J . Si

$$\exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_H^2, \quad (\text{B.10})$$

alors \bar{x} est un minimum local strict (sous-entendu pour la norme $\|\cdot\|_E$).

► La formule de Taylor appliquée à J nous donne pour $h \in E$ t.q. $\bar{x} + h \in U$:

$$J(\bar{x} + h) - J(\bar{x}) = \frac{1}{2!} J''(\bar{x}) \cdot h^2 + \|h\|_E^2 \varepsilon(h) \quad \text{avec} \quad \lim_{\|h\|_E \rightarrow 0} |\varepsilon(h)| = 0.$$

Par hypothèse :

$$J(\bar{x} + h) - J(\bar{x}) \geq \frac{\alpha}{2} \|h\|_H^2 + \varepsilon(h) \|h\|_E^2 \geq \left(\frac{\alpha}{2} + \frac{\varepsilon(h)}{c} \right) \|h\|_E^2$$

et donc il existe $\eta > 0$ tel que pour tout $h \neq 0_E$ t.q. $\|h\|_E < \eta$ on ait $\bar{x} + h \in U$ et

$$J(\bar{x} + h) - J(\bar{x}) > 0.$$

Autrement dit, \bar{x} est un minimum local strict (sous-entendu pour la norme $\|\cdot\|_E$). ■

Remarque B.5.8. Dans la proposition précédente, il suffit de faire l'hypothèse $E \subset H$. L'hypothèse d'injection continue n'est pas nécessaire car elle est automatiquement vérifiée dès lors que l'on a l'équation (B.10). En effet, $J''(\bar{x})$ est continue car J est deux fois dérivable en \bar{x} et donc, il existe une constante $K > 0$ telle que

$$\forall h \in E : |J''(\bar{x}) \cdot h^2| \leq K \|h\|_E^2$$

et finalement puisque (B.10) est vérifiée, on obtient automatiquement :

$$\forall h \in E : \alpha \|h\|_H^2 \leq J''(\bar{x}) \cdot h^2 \leq K \|h\|_E^2, \quad (\text{B.11})$$

et la constante $c := \sqrt{K/\alpha}$ fait l'affaire pour l'hypothèse de la proposition précédente.

Remarque B.5.9. Dans le cas particulier où $E = H$ et $\|\cdot\|_E = \|\cdot\|_H$, on retrouve la condition suffisante d'ordre deux (CS2) classique, cf. conditions (B.7). Introduisons dans ce cas la quantité $\|h\| := \sqrt{J''(\bar{x}) \cdot h^2}$. Il est alors intéressant de noter (ce que l'on aurait pu mentionner avant) que $\|\cdot\|$ définit une norme équivalente à la norme $\|\cdot\|_E$, cf. l'équa-

tion (B.11), mais que cette nouvelle norme est induite par le produit scalaire

$$(h | k) := J''(\bar{x}) \cdot (h, k).$$

L'ensemble E munit du produit scalaire $(\cdot | \cdot)$ est donc un espace préhilbertien (un espace de Hilbert si E est un espace vectoriel normé complet, *i.e.* si E est un espace de Banach), on dit alors que $(E, \|\cdot\|_E)$ est un espace vectoriel normé **homéomorphe** à un espace préhilbertien. On arrive alors à la conclusion remarquable que si E n'est pas homéomorphe à un espace préhilbertien, alors la condition d'ellipticité ne peut jamais être vérifiée, quel que soit le point critique ! Et donc dans ce cas, il est nécessaire d'utiliser une nouvelle norme pour obtenir la condition de croissance.

Remarque B.5.10. La Proposition B.5.8 donne une nouvelle condition suffisante d'ordre deux d'optimalité locale pour un minimum libre, *i.e.* dans le cadre d'un problème d'optimisation sans contraintes. Bien entendu, nous pouvons donner une version de cette proposition dans le cas de contraintes affines de manière similaire à ce qui a été présenté dans la Section B.5.2 (voir les conditions (B.8)) :

$$(CS2G) \quad \exists \alpha > 0 \quad \text{t.q.} \quad \forall h \in E_0 : J''(\bar{x}) \cdot h^2 \geq \alpha \|h\|_{H_0}^2,$$

où $E_0 := E \cap \text{Ker } T$, $H_0 := H \cap \text{Ker } T$ et où $\bar{x} \in \mathcal{K}$. Cette version ayant longuement été détaillée dans cette section, cf. notamment l'équation (B.9).

B.6 Condition de Jacobi

Supposons pour simplifier que $n = 1$, $\Omega = \mathbb{R}$, *i.e.* $x(t) \in \mathbb{R}$ et que le lagrangien L soit de classe \mathcal{C}^3 . On rappelle la dérivée seconde de la fonction coût :

$$J''(x) \cdot h^2 = \int_a^b \left(\frac{\partial^2 L}{\partial x^2}[t] h^2(t) + 2 \frac{\partial^2 L}{\partial u \partial x}[t] \dot{h}(t) h(t) + \frac{\partial^2 L}{\partial u^2}[t] \dot{h}^2(t) \right) dt,$$

où $[t] := (t, x(t), \dot{x}(t))$. Soit \bar{x} une BC-extrémale qui satisfait la condition forte de Legendre :

$$\forall t \in [a, b] : \frac{\partial^2 L}{\partial u^2}(t, \bar{x}(t), \dot{\bar{x}}(t)) > 0.$$

Alors, d'après le Théorème B.3.4, $\bar{x} \in \mathcal{C}^3([a, b], \mathbb{R})$, donc en particulier $\bar{x} \in \mathcal{C}^2([a, b], \mathbb{R})$ mais comme nous l'avons déjà vu (cf. le début de la Section B.4), nous pouvons prolonger \bar{x} de manière continue aux extrémités, c'est-à-dire $\bar{x} \in \mathcal{C}^2([a, b], \mathbb{R})$ (nous n'avons pas besoin de plus de régularité). Puisque L est \mathcal{C}^3 , alors

$$t \mapsto \frac{\partial^2 L}{\partial u \partial x}(t, \bar{x}(t), \dot{\bar{x}}(t))$$

est \mathcal{C}^1 , et à l'aide d'une simple intégration par partie, on peut écrire :

$$\int_a^b 2 \frac{\partial^2 L}{\partial u \partial x}[t] \dot{h}(t) h(t) dt = - \int_a^b \frac{d}{dt} \left(\frac{\partial^2 L}{\partial u \partial x}[t] \right) h^2(t) dt,$$

où maintenant $[t] := (t, \bar{x}(t), \dot{\bar{x}}(t))$. Ceci nous permet de réécrire la dérivée seconde sous la forme :

$$J''(\bar{x}) \cdot h^2 = \int_a^b (Q(t) h^2(t) + R(t) \dot{h}^2(t)) dt,$$

avec

$$Q(t) := \frac{\partial^2 L}{\partial x^2}[t] - \frac{d}{dt} \left(\frac{\partial^2 L}{\partial u \partial x}[t] \right) \quad \text{et} \quad R(t) := \frac{\partial^2 L}{\partial u^2}[t].$$

Sous l'hypothèse L de classe \mathcal{C}^3 et sous la condition de Legendre stricte, on a $Q \in \mathcal{C}^0([a, b], \mathbb{R})$ et $R \in \mathcal{C}^1([a, b], \mathbb{R})$.

On appelle **équation de Jacobi** l'équation différentielle linéaire du second ordre :

$$Q(t) h(t) - \frac{d}{dt} (R(t) \dot{h}(t)) = 0, \quad (\text{B.12})$$

Un temps $t_c \in [a, b]$ est dit **conjugué** à a (le point $x(t_c)$ est dit conjugué à $x(a)$) si l'équation de Jacobi admet une solution h telle que

$$h(t_c) = h(a) = 0 \quad \text{et} \quad R(t_c) \dot{h}(t_c) \neq 0.$$

Sous la condition de Legendre forte, la condition $R(t_c) \dot{h}(t_c) \neq 0$ se réduit à $\dot{h}(t_c) \neq 0$. On dit que la **condition faible de Jacobi** est satisfaite si l'intervalle ouvert $]a, b[$ ne contient aucun temps conjugué à a . On parle de **condition forte de Jacobi** si l'intervalle $]a, b]$ ne contient aucun temps conjugué à a . On a alors le résultat suivant :

Théorème B.6.1 – Jacobi-Weierstrass (1877)

Supposons le lagrangien L de classe \mathcal{C}^3 et la dimension $n = 1$.

Si \bar{x} est une solution globale, locale forte ou locale faible du Problème (PCV), alors \bar{x} vérifie :

- l'équation de Euler-Lagrange et les conditions aux limites ;
- la condition faible de Legendre ;
- la condition faible de Jacobi.

Une condition suffisante pour que \bar{x} donne un minimum local faible strict est que soient vérifiées :

- l'équation de Euler-Lagrange et les conditions aux limites ;
- la condition forte de Legendre ;
- la condition forte de Jacobi.

► Voir [65, Section 1.4]. On peut remarquer que dans la preuve du [65, Théorème 1.4.4], on utilise (implicitement) la condition suffisante d'ordre 2 généralisée (B.5.8) basée sur l'idée de la “two-norm discrepancy”. ■

Bibliographie

- [1] A. A. Agrachev & Y. L. Sachkov, *Control theory from the geometric viewpoint*, vol **87** of *Encyclopaedia of Mathematical Sciences*, Springer-Verlag, Berlin (2004), 412 pages. \hookleftarrow [13](#), [20](#), [41](#), [53](#), [54](#), [62](#), [72](#) et [74](#).
- [2] E. Allgower & K. Georg, *Introduction to numerical continuation methods*, vol **45** of *Classics in Applied Mathematics*, Soc. for Industrial and Applied Math., Philadelphia, PA, USA, (2003), xxvi+388.
- [3] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. D. Croz, A. Greenbaum, S. Hammarling, A. McKenney & D. Sorensen, *LAPACK Users' Guide*, Soc. for Industrial and Applied Math., Philadelphia, PA, USA, third edn (1999)
- [4] J. Appell & P. P. Zabrejko, *Nonlinear Superposition Operators*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1990. \hookleftarrow [8](#) et [106](#).
- [5] V. I. Arnold, *Équations différentielles ordinaires*, MIR, Moscou (1974). \hookleftarrow [3](#).
- [6] V. I. Arnold, *Mathematical methods of classical mechanics*. Translated from the Russian by K. Vogtmann and A. Weinstein. Second edition. Graduate Texts in Mathematics, **60**. Springer-Verlag, New York, 1989, 508 pages. \hookleftarrow [72](#).
- [7] V. I. Arnold, *Singularities of Caustics and Wave Fronts*, Kluwer Acad. Publ., Dordrecht-Boston-London, 1990. \hookleftarrow [28](#).
- [8] U. Ascher, R. Mattheij & R. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, vol **13** of *Classics in Applied Mathematics*, Soc. for Industrial and Applied Math., Philadelphia, PA, USA, (1995), xxvi+595. \hookleftarrow [79](#).
- [9] P. B. Bailey, L. F. Shampine & P. E. Waltman, *Nonlinear two point boundary value problems*, Vol. 44. New York : Academic press, 1968. \hookleftarrow [81](#).
- [10] J. T. Betts, *Survey of Numerical Methods for Trajectory Optimization*, AIAA J. of Guidance, Control and Dynamics, **21** (1998), 193–207. \hookleftarrow [79](#).
- [11] J. T. Betts, *Practical methods for optimal control using nonlinear programming*, vol 19 of *Advances in Design and Control*, Soc. for Industrial and Applied Math., Philadelphia, PA, USA, second edn (2010), xiv+434. \hookleftarrow [79](#).
- [12] H. G. Bock, *Numerical treatment of inverse problems in chemical reaction kinetics*, vol 18 of *Springer Series in Chemical Physics*, Eds. K. H. Ebert, P. Deuflhard & W. Jäger, in *Modelling of Chemical Reaction Systems*, Springer, Heidelberg, (1981), 102–125. \hookleftarrow [83](#).
- [13] H. Bock & K. Plitt, *A multiple shooting algorithm for direct solution of optimal control problems*, Pergamon Press, (1984), 243–247. Proceedings of the 9th IFAC World Congress, Budapest. \hookleftarrow [79](#) et [85](#).
- [14] A. V. Bolsinov & A. T. Fomenko, *Integrable Hamiltonian Systems, Geometry, Topology, Classification*. Chapman and Hall/CRC, London, 2004, 724 pages. \hookleftarrow [72](#).
- [15] V. G. Boltyanski, *The Maximum Principle — How it came to be?*, Mathematisches Institut, München, Germany, Report No. 526, 1994. \hookleftarrow [63](#).

- [16] J. F. Bonnans, D. Giorgi, V. Grelard, S. Maindrault & P. Martinon, *Bocop — A collection of examples*, Tech. report, 2015.
- [17] J. F. Bonnans & A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research and Financial Engineering, Springer-Verlag New York 2000, XVIII, 601. \leftrightarrow 143.
- [18] B. Bonnard, J.-B. Caillau & E. Trélat, *Second order optimality conditions in the smooth case and applications in optimal control*, ESAIM Control Optim. Calc. Var., **13** (2007), no. 2, 207–236. \leftrightarrow 54 et 76.
- [19] B. Bonnard & M. Chyba, *Singular trajectories and their role in control theory*. Vol 40 of *Mathematics & Applications*, Springer-Verlag, Berlin (2003), 357 pages. \leftrightarrow 12, 14, 62 et 76.
- [20] B. Bonnard & I. Kupka, *Théorie des singularités de l'application entrée/sortie et optimalité des trajectoires singulières dans le problème du temps minimal*. Forum Math., **5** (1993), no. 2, pp. 111–159. \leftrightarrow 12, 14 et 76.
- [21] B. Bonnard, N. Shcherbakova & D. Sugny, *The smooth continuation method in optimal control with an application to quantum systems*, ESAIM Control Optim. Calc. Var., **17** (2011), no 1, 267–292.
- [22] H. Brezis, *Analyse fonctionnelle, Théorie et applications*, Collection Mathématiques appliquées pour la maîtrise, Masson, 1987. \leftrightarrow 8, 146 et 148.
- [23] J.-B. Caillau, Z. Chen & Y. Chitour, *L^1 -minimization for mechanical systems*, SIAM J. Control Optim. **54** (2016), no. 3, 1245–1265. \leftrightarrow 72.
- [24] J.-B. Caillau & B. Daoud, *Minimum time control of the circular restricted three-body problem*, SIAM J. Control Optim., **50** (2012), no. 6, 3178–3202.
- [25] C. Carathéodory, *Calculus of Variations and Partial Differential Equations of the First Order, Part 1, Part 2*, Holden-Day, San Francisco, California, 1965–1967; Reprint : 2nd AMS printing, AMS Chelsea Publishing, Providence, RI, USA, 2001, 412 pages. \leftrightarrow 27 et 124.
- [26] L. Cesari, *Optimization-theory and applications: problems with ordinary differential equations*, vol 17 of *Applications of mathematics*, Springer-Verlag, New York, 1983, 542 pages. \leftrightarrow 62.
- [27] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, Graduate Texts in Mathematics 264, Springer-Verlag London 2013. \leftrightarrow 46, 62, 124, 125, 132, 133 et 139.
- [28] O. Cots, *Contrôle optimal géométrique : méthodes homotopiques et applications*. Phd thesis, Institut Mathématiques de Bourgogne, Dijon, France, 2012. \leftrightarrow 66, 72 et 76.
- [29] B. Daoud, *Contribution au contrôle optimal du problème circulaire restreint des trois corps*, Thèse de doctorat, Université de Bourgogne (2011).
- [30] J. Demailly, *Analyse numérique et équations différentielles*, Collection Grenoble Sciences. EDP Sciences (2006). \leftrightarrow 7.
- [31] M. P. do Carmo, *Riemannian geometry*. Birkhäuser, Mathematics : Theory & applications, 2nd edn 1988, 300 pages. \leftrightarrow 28.

- [32] J. J. Dongarra, J. Du Croz, S. Hammarling & I. S. Duff, *A set of level 3 basic linear algebra subprograms*, ACM Trans. Math. Softw., **16** (1990), no 1, 1–17.
- [33] J. J. Dongarra, J. Du Croz, S. Hammarling & R. J. Hanson, *An extended set of fortran basic linear algebra subprograms*, ACM Trans. Math. Softw., **14** (1988), no 1, 1–17.
- [34] C. B. García & W. I. Zangwill, *An approach to homotopy and degree theory*, Math. Oper. Res., **4** (1979), no 4, 390–405.
- [35] I. M. Gelfand & S. V. Fomin, *Calculus of Variations*, Dover Publications, New York, 2000. \hookleftarrow 27 et 124.
- [36] J. Gergaud, *Sur la résolution numérique de problèmes de contrôle optimal à solution bang-bang via les méthodes homotopiques*, HDR thesis, Université de Toulouse, 2008. \hookleftarrow 89.
- [37] R. H. Goddard, *A Method of Reaching Extreme Altitudes*, volume 71(2), Smithsonian Miscellaneous Collections. Smithsonian institution, City of Washington, (1919). \hookleftarrow 31.
- [38] B. S. Goh, *Necessary conditions for singular extremals involving multiple control variables*, SIAM Journal on Control, 1966, vol. 4, no. 4, p. 716–731. \hookleftarrow 74.
- [39] É.ourgoulhon, *Relativité restreinte : des particules à l’astrophysique*, EDP sciences, 2010. \hookleftarrow 48.
- [40] W. Greub, *Multilinear algebra*, Springer, New York (1967). \hookleftarrow 141.
- [41] E. Hairer, S. P. Nørsett & G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, vol 8 of *Springer Serie in Computational Mathematics*, Springer-Verlag, second edn (1993). \hookleftarrow 83 et 86.
- [42] E. Hairer & G. Wanner, *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems*, vol 14 of *Springer Serie in Computational Mathematics*, Springer-Verlag, second edn (1996).
- [43] L. Hascoët & V. Pascual, *The Tapenade Automatic Differentiation tool : principles, model, and specification*, Rapport de recherche RR-7957, INRIA (2012).
- [44] A. Herasimenka, *Contrôle optimal des voiles solaires*, Thèse de doctorat, Université Côte d’Azur (2023).
- [45] J.-B. Hiriart-Urruty, *Bases, outils et principes pour l’analyse variationnelle*, Vol. 70, Springer Science & Business Media, 2012.
- [46] J.-B. Hiriart-Urruty, *L’optimisation*, Que sais-je, Presses Universitaires de France, 1996. \hookleftarrow 143.
- [47] A. D. Ioffe & V. M. Tikhomirov, *Theory of extremal problems*, Elsevier, 2009. \hookleftarrow 62.
- [48] F. Jean, *Géométrie Différentielle et Application au Contrôle Géométrique*, Notes de cours, 2009/2010. \hookleftarrow 14 et 21.
- [49] F. Jean, *Stabilité et Commande des Systèmes Dynamiques. Cours et exercices corrigés*, Coll. Les Cours, Les Presses de l’ENSTA, 2011, 197 pages. \hookleftarrow 4 et 14.
- [50] V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1996. \hookleftarrow 14 et 62.
- [51] W. Klingenberg, *Riemannian geometry*, de Gruyter Studies in Mathematics, vol. 1, Walter de Gruyter & Co., Berlin, 1995, 409 pages. \hookleftarrow 28.

- [52] J. Lafontaine, *Introduction aux variétés différentielles*, Presses Universitaires De Grenoble, 1996, 299 pages. [↔ 21](#).
- [53] M. Lapert, *Développement de nouvelles techniques de contrôle optimal en dynamique quantique : de la RMN à la Physique moléculaire*, Thèse de doctorat, Université de Bourgogne (2011).
- [54] E. B. Lee & L. Markus, *Foundations of Optimal Control Theory*, Wiley, New York, 1967. [↔ 62](#).
- [55] A. Lesfari, *Géométrie symplectique, calcul des variations et dynamique hamiltonienne*, Published in Great Britain by ISTE Editions Ltd, London, 2021. [↔ 17](#), [19](#) et [72](#).
- [56] D. Liberzon, *Calculus of variations and optimal control theory : a concise introduction*, Princeton university press, 2011. [↔ 61](#) et [62](#).
- [57] R. Mannshardt, *One-step methods of any order for ordinary differential equations with discontinuous right-hand side*, Numer. Math., **31** (1978), no 2, 131–152. [↔ 86](#).
- [58] P. Martinon & J. Gergaud, *Using switching detection and variational equations for the shooting method*, Optimal Control Appl. Methods, **28** (2007), no 2, 95–116. [↔ 80](#), [86](#) et [87](#).
- [59] H. Maurer, *Numerical solution of singular control problems using multiple shooting techniques*, Journal of optimization theory and applications, Vol.18, No.2, (1976). [↔ 31](#).
- [60] H. Maurer & D. Augustin, *Second order sufficient conditions and sensitivity analysis for the controlled Rayleigh problem*, in Parametric Optimization and Related Topics, IV: 15–55, J. Guddat, Ed. et al. Frankfurt, Germany: Peter Lang Publishing, 1997. [↔ 30](#).
- [61] J. Milnor, *Morse Theory*, Princeton University Press, 1963. [↔ 73](#).
- [62] J. J. Moré, B. S. Garbow & K. E. Hillstom, *User Guide for MINPACK-1*, ANL-80-74, Argonne National Laboratory, (1980).
- [63] F. Paulin, *Géométrie différentielle élémentaire*, Notes de cours, 2006–2007. [↔ 21](#).
- [64] L. S. Pontryagin, V. G. Boltyanskiï, R. V. Gamkrelidze & E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Translated from the Russian by K. N. Trirogoff, edited by L. W. Neustadt, Interscience Publishers John Wiley & Sons, Inc., New York-London, 1962, 360 pages. [↔ 61](#) et [62](#).
- [65] H. Schättler & U. Ledzewicz, *Geometric optimal control: theory, methods and examples*, vol **38** of *Interdisciplinary applied mathematics*, Springer Science & Business Media, New York (2012), xiv+640. [↔ 14](#), [41](#), [62](#), [124](#), [126](#), [130](#), [132](#), [135](#), [137](#) et [151](#).
- [66] C. Silva & E. Trélat, *Smooth regularization of bang-bang optimal control problems*, IEEE Trans. Automat. Control, **55** (2010), no 11, 2488–2499. [↔ 30](#), [88](#) et [89](#).
- [67] E. D. Sontag, *Mathematical control theory: deterministic systems*, Springer-Verlag New York, Inc., New York, NY, USA, 1998. [↔ 7](#), [9](#), [14](#), [38](#) et [62](#).
- [68] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc., third edition, 1999, 493 pages. [↔ 21](#).
- [69] J. Stoer & R. Bulirsch, *Introduction to numerical analysis*, vol 12 of *Texts in Applied Mathematics*, Springer-Verlag, New York, second edn (1993). [↔ 79](#) et [86](#).

- [70] H. J. Sussmann, *Geometry and optimal control*, *Mathematical control theory*, Springer, New York (1999), 140–198. \leftrightarrow 62.
- [71] H. J. Sussmann & J. C. Willems, *300 years of optimal control: from the Brachistochrone to the maximum principle*, IEEE Control Systems, (1997), pp. 32–44. \leftrightarrow 121 et 124.
- [72] E. Trélat, *Contrôle optimal : théorie et applications*, Coll. Mathématiques concrètes, Vuibert, 2008, 250 pages. pdf. \leftrightarrow 12, 14 et 62.
- [73] E. Trélat, *Optimal control and applications to aerospace: some results and challenges*, J. Optim. Theory Appl., **154** (2012), no 3, 52.
- [74] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000. \leftrightarrow 62.
- [75] C. Wagschal, *Topologie et analyse fonctionnelle*, Hermann, 2012. \leftrightarrow 16, 111, 113, 115 et 118.
- [76] L. C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, 1969. \leftrightarrow 62.

Index

Application

- contractante, 103, 109, 116
- entrée/sortie, 12, 41
- entrée/sortie augmentée, 35, 51
- lipschitzienne, 7, 103

BC-extrémale, 39, 64, 72, 132, 150

Calcul des variations, 27, 45, 121, 125

Caténoïde, 134

Condition

- d'ellipticité, 142, 148, 150
- d'Euler-Hamilton, 41, 61
- de croissance, 148, 150
- de Erdmann, 133
- de Jacobi, 41, 54, 72, 151
- de Legendre, 41, 137, 139, 145
- nécessaire d'ordre deux, 142, 144
- nécessaire d'ordre supérieur, 140
- nécessaire du premier ordre, 139, 144
- nécessaire et suffisante, 140
- suffisante d'ordre deux, 142, 144
- suffisante d'ordre deux généralisée, 149
- suffisante d'ordre supérieur, 141

Conditions

- aux limites simples, 125

Conditions aux limites, 39, 64

Conditions de transversalité, 39, 64

Conservation du hamiltonien, 42, 49, 64

Contraintes

- affines, 144

Contrôlabilité, 14, 15

Contrôle

- régulier, 14
- singulier, 14

Critère de Kalman, 14

Crochet de Poisson, 19

Difféomorphisme, \mathcal{C}^k -difféomorphisme, 113

Double intégrateur

- énergie min, 29

temps min, 29

Ensemble

- accessible, 13
- accessible augmenté, 35
- des contrôles admissibles, 7, 11

Équation

- de Euler, 139
- de Euler-Lagrange, 131, 145
- linéarisée, 11
- variationnelle, 11, 82

équation

- de Euler-Lagrange, 46
- de Hamilton, 46, 72

Espace

- de Banach, 110
- métrique, 103

Etat augmenté, 24, 35

Extrémale, 39, 132

- anormale, 64
- bang, 64
- de Hamilton, 70
- de Pontryagin-Boltyanski, 64
- normale, 64

Extrémum

- libre, 139
- lié, 144

Flot d'un système contrôlé, 8

Fonction

- admissible, 125
- coût, 124

Fonction de tir, 44, 80

Forme bilinéaire

- définie positive, 142
- elliptique, 142
- semi-définie positive, 142

Gradient

- symplectique, 19

Hamiltonien, 18, 46, 49, 70

- Homéomorphisme, 110
- Identité
 - de Beltrami, 133
- Isomorphisme, 110
- Lagrangien, 28, 45, 124
- Loi de commande, 6
- Méthode
 - de Newton, 82
 - de Runge-Kutta, 82
- Métrique riemannienne, 28
- Minimum
 - global, 139
 - local, 139
 - local strict, 139
- Paradoxe des jumeaux de Langevin, 47
- Point critique, 139
- Principe
 - du Maximum de Pontryagin, 62
 - Fondamental de la Dynamique, 49
- Problème
 - à temps final libre, 23
 - à temps final fixé, 23
 - de Bolza, 23, 37
 - de calcul des variations, 27
 - de contrôle optimal, 22, 62
 - de géométrie riemannienne, 28
 - de Goddard, 31
 - de la courbe Brachistochrone, 121
 - de Lagrange, 23, 35, 72
 - de Mayer, 23, 33
 - de Rayleigh, 30
 - de surface d'aire minimale, 134
 - de transfert orbital, 24, 83
 - du temps minimal, 23
 - linéaire quadratique, 27
 - linéaire quadratique (LQR), 57
 - linéaire, 27
 - variationnel, 125
- Pseudo-hamiltonien, 16, 24
- Solution
 - d'un problème de contrôle optimal, 23
 - d'un système contrôlé, 6
 - globale, 125
 - locale faible, 126
 - locale forte, 126
- Somme directe algébrique, topologique, 118
- Submersion, 22
- Système
 - augmenté, 24, 35
 - contrôlé, 5
 - hamiltonien, 18, 19, 46, 49
- Théorème
 - CN1 d'optimalité, optimisation, 119
 - d'inversion locale, 115
 - de Banach, 113
 - de Cauchy-Lipschitz généralisé, 7
 - de l'application ouverte, 118
 - de l'application ouverte non linéaire, 119
 - de l'application ouverte non linéaire - ordre 2, 52
 - de régularité du flot, 8
 - des accroissements finis 1, 102
 - des accroissements finis 2, 104
 - des fonctions implicites, 117
 - du point fixe, 109
- Théorème
 - de Euler-Lagrange, 131
 - de Hilbert-Weierstrass, 132
 - de Jacobi-Weierstrass, 151
 - de Legendre, 137
- Tir multiple
 - de stabilité, 85
 - de structure, 91
- Tir simple avec détections des commutations, 86
- Tir simple indirect, 81
- Two-norm discrepancy, 148
- Valeur critique, 139
- Vecteur adjoint, covecteur, 16, 24