

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU

FACULTE DES SCIENCES

DEPARTEMENT DE MATHEMATIQUES

THESE DE DOCTORAT

SPECIALITE : MATHEMATIQUES

OPTION : RECHERCHE OPERATIONNELLE ET OPTIMISATION

Présentée par :

M^{me} Kahina LOUADJ

Sujet :

Résolution de problèmes paramétrés de contrôle optimal

Devant le jury d'examen composé de :

Mr .Said Djennoune ;	Professeur ;	U.M.M.T.O ;	Président
Mr.Mohamed Aidene ;	Professeur ;	U.M.M.T.O ;	Rapporteur
Mr.Pierre Spiteri ;	Professeur ;	Université de Toulouse ;	Co-Directeur
Mr.Meziane Aider ;	Professeur ;	U.S.T .H.B ;	Examineur
Mr.Mohand Said Radjef ;	Professeur ;	Université de Béjaia ;	Examineur
Mr.Brahim Oukacha ;	M.Conférence Classe A ;	U.M.M.T.O ;	Examineur
Mr.Frédéric Messine ;	M. Conférence HDR ;	Université de Toulouse ;	Invité

Soutenue : le 09/05/2012

Remerciements

Je tiens à exprimer toute ma reconnaissance à *M^r* Aidene Mohamed mon directeur de thèse, professeur à l'université Mouloud Mammeri de Tizi-Ouzou, qui a proposé et accepté de diriger ce travail. Il a apporté une grande contribution à son aboutissement, et qui m'a fait découvrir ce domaine si riche du contrôle optimal, à la frontière de l'optimisation et du calcul différentiel. Son engagement personnel m'a permis de surmonter les démarches administratives. Il m'a aussi apporté toute facilité pour favoriser mon déplacement à Minsk et à Toulouse. Qu'il soit chaleureusement et sincèrement remercié pour le soutien qu'il m'a prodigué.

Je tiens à exprimer ma gratitude à *M^r* Pierre Spiteri, mon co-directeur de thèse à L'université de Toulouse, je le remercie pour m'avoir accordé sa confiance, pour m'avoir fait profiter de son expérience dans le domaine dense de l'optimisation, ainsi que pour son investissement indéniable et conséquent dans les travaux présentés ici. Ses grandes qualités scientifique ont sans nul doute contribué à améliorer et parfaite mes travaux de thèse. Ses recommandations et remarques furent à chaque fois riches de sens.

Je tiens à remercier *M^r* Frédéric Messine de l'université de Toulouse d'avoir accepter mon invitation et être parmi les membres de jury, comme rôle d'examineur. Je le remercie malgré ses occupations, pour sa contribution et son aide scientifique surtout dans la partie programmation.

Je remercie également *M^r* Said Djennoune qui m'a fait l'honneur de présider mon jury ainsi que *M^r* Méziane Aider, *M^r* Mohand Said Radjef, *M^r* Brahim Oukacha pour avoir accepté le rôle d'examineur.

J'adresse également mes remerciements à toute l'équipe du laboratoire de l'IRIT-ENSEEIH, qui m'ont entouré durant mes stages à Toulouse. Une pensée particulière pour Zehor et Meriem, sans oublier la secrétaire du laboratoire Sylvie.

J'adresse mes remerciements à tous les membres du laboratoire de L2CSP de l'université de Tizi-Ouzou, qui m'ont entouré durant ma thèse. Une pensée particulière pour Nacima et Aldjia, et pour Said Guermah, Rachid Mansouri et Kara Redouane.

Je remercie tous mes amis qui m'ont aidé de près ou de loin, surtout lynda, fadila, Cherifa Belkacem, Cherifa Bennani, Amina. Et sans notre chère secrétaire de département de Mathématiques Karima.

Je voudrais remercier du fond du cœur toute ma famille, mes parents, mes soeurs Aldjia, Lamia et Amel, mon cher frère Yacine de m'avoir soutenu dès le début. Et sans oublier, une pensée pour mes beaux frères Farouk et Hocine.

Table des matières

Table des matières	1
Introduction générale	4
Liste des travaux scientifique	9
1 Méthodes numériques de résolution des systèmes différentiels ordinaires	11
1.1 Introduction	11
1.2 Problème de Cauchy	12
1.3 Théorèmes d'existence et d'unicité	12
1.3.1 Théorème d'existence	12
1.3.2 Théorème d'unicité	13
1.4 Méthode d'Euler	13
1.4.1 Majoration de l'erreur dans la méthode d'Euler	14
1.5 Méthode numérique à un pas	16
1.5.1 Définition	16
1.5.2 Méthode de Taylor d'ordre p	16
1.6 Méthode du point milieu	17
1.6.1 Notions de consistance, de stabilité et de convergence	18
1.7 Runge-Kutta	19
1.7.1 Fonctions Matlab ode45	24
2 Introduction à la commande optimale	25
2.1 Introduction	25
2.2 Théorie du contrôle optimal et des systèmes de contrôle	25
2.2.1 Objet de la commande	26
2.2.2 Condition initiale du système	26
2.2.3 Le but de la commande	26
2.2.4 Classe des commandes admissibles	26
2.2.5 Critère de qualité	27
2.3 Contrôlabilité	28
2.3.1 Contrôlabilité des systèmes linéaires	29
2.3.2 Contrôlabilité des systèmes linéaires autonomes	30

2.3.3	Contrôlabilité des systèmes non-linéaire	30
2.4	Problème de contrôle optimal	31
2.4.1	Principe du maximum de Pontriaguine	31
2.4.2	Méthode directe de résolution d'un problème terminal d'un système dynamique linéaire :	33
2.5	Support-Contrôle	34
2.6	Formule d'accroissement de la fonctionnelle	36
2.7	Problème dual	37
2.8	Calcul de la valeur de suboptimalité	38
2.9	Critère d'optimalité	39
2.10	Critère d' ε -Optimalité	42
2.11	Détermination d'un support-contrôle de départ	44
2.12	Méthode de résolution	45
2.12.1	Changement de commande	45
2.12.2	Changement de support	48
2.12.3	Procédure finale	50
3	Optimisation d'un problème de contrôle optimale avec une entrée libre	52
3.1	Introduction	52
3.2	Position du problème	52
3.3	Définitions Essentielles	53
3.4	Support contrôle	54
3.5	Problème dual	55
3.6	Calcul de la valeur de suboptimalité	55
3.7	Critère d'Optimalité et $d'\varepsilon$ -Optimalité	57
3.8	Algorithme de résolution	58
3.8.1	Changement de commande	58
3.8.2	Changement de support	62
3.8.3	Procédure finale	65
3.9	Exemple numérique	67
4	Nouvelle approche de résolution d'un problème de contrôle optimal	69
4.1	Introduction	69
4.2	Position du problème	69
4.3	Définitions Essentielles	70
4.4	Problème discret du problème initial	71
4.5	La valeur de suboptimalité	72
4.6	Critère d'optimalité et $d'\varepsilon$ -optimalité	73
4.7	L'algorithme numérique pour la résolution du problème discret	73
4.7.1	Changement de commande	74
4.7.2	Changement de support	75
4.8	Procédure finale	77

5	Problème de contrôle optimal avec une entrée libre	79
5.1	Introduction	79
5.2	Position du problème	79
5.3	Définitions Essentielles :	80
5.4	Problème discret du problème initial	81
5.5	La valeur de suboptimalité	83
5.6	Critère d'optimalité et ε -optimalité	84
5.7	L'algorithme numérique pour la résolution du problème discret	84
5.7.1	Changement de commande	84
5.7.2	Changement de support	86
5.8	Procédure finale	89
5.9	Exemple numérique	91
6	Méthode de tir et méthode de Newton discrète	95
6.1	Introduction	95
6.2	La méthode de tir	95
6.3	Discrétisation totale : méthode directe	96
6.4	Méthode de Newton discrète	98
6.4.1	Différenciation approchée	98
6.4.2	Convergence de la méthode Newton discrète	104
7	Résolution par la méthode de relaxation d'un problème de	
	contrôle optimal avec une entrée libre	107
7.1	Introduction	107
7.2	Position du problème	108
7.2.1	La méthode de tir simple	109
7.3	Résolution numérique	111
7.4	Exemple numérique	112
	Conclusion générale	126
	Annexe 1	127
	Annexe 2	135
	Bibliographie	136

Introduction générale

La théorie du contrôle permet d'amener un système d'un état initial donné à un certain état final en respectant certains critères : c'est l'étape de réalisation de la commande. Par exemple, tout le monde sait maintenir en équilibre un pendule sur son doigt. En revanche il est beaucoup plus difficile de maintenir en équilibre sur son doigt un double pendule inversé. La théorie du contrôle permet pourtant de le faire. Mais pour réaliser effectivement cet équilibre, mieux vaut disposer d'un bon modèle mathématique et savoir résoudre les équations. Une voiture sur laquelle on agit avec les pédales d'accélérateur et de frein, et que l'on guide avec le volant est un exemple de système de contrôle, de système dynamique sur lequel on peut agir au moyen d'une commande représenté par le frein à main.

Un système de contrôle est un système dynamique sur lequel on peut agir au moyen d'une commande. Pour définir précisément le concept de système de contrôle, il faut utiliser le langage mathématique. Chaque système a une structure, des propriétés et des finalités spécifiques. Notons que ce concept peut aussi bien décrire des transformations discrètes que continues. Cela permet donc de modéliser le fonctionnement de robots, de systèmes adaptatifs à structure variable, etc. En considérant tous ces objets comme des systèmes de contrôle, on s'intéresse à leur comportement et à leurs caractéristiques fonctionnelles, sans forcément attacher d'importance à leurs propriétés internes ou intrinsèques. Par conséquent, deux systèmes de contrôle ayant, en un certain sens, même comportement et des caractéristiques similaires, sont considérés comme identiques. De nos jours, les systèmes automatisés font complètement partie de notre quotidien ; le but est d'améliorer notre qualité de vie et de faciliter certaines tâches.

L'objectif peut être aussi de stabiliser le système pour le rendre insensible à certaines perturbation, ou encore de déterminer des solutions optimales pour un certain critère d'optimisation (contrôle optimal). Mathématiquement, on modélise les phénomènes abordés par des systèmes d'équations différentielles ordinaires, aux dérivées partielles, etc.

Historiquement, la théorie du contrôle est liée d'une part avec le calcul des variations ([82],[90]) et d'autre part avec la résolution des équations différentielles ordinaires. Pour la première fois, Johann Bernouilli a soumis le problème de brachistochrone correspondant au problème de la trajectoire

la plus rapide entre deux points, en 1696, Leibniz, Newton, de l'Hôpital ainsi que son frère Jacques Bernouilli trouvèrent la solution. La méthode classique pour résoudre le problème est le calcul des variations. Cependant, la méthode utilisée par Jean Bernouilli est basée sur une analogie de l'optique [28]. Ceci est considéré comme un résultat pionnier dans le domaine du contrôle optimal. Cette théorie qui est une extension du calcul des variations, traite de la façon de trouver une loi de commande pour un système, modélisé par un ensemble d'équations différentielles décrivant les trajectoires d'état et de contrôle, de telle sorte qu'un certain critère d'optimalité soit atteint. La résolution de problème de contrôle optimal a pu démarré avec le célèbre principe du maximum de Pontriaguine (PMP) [73], qui fournit **une condition nécessaire** d'optimalité, ainsi que la programmation dynamique de Hamilton-Jacobi-Bellman (HJB) [5], [6] qui fournit **une condition suffisante** d'optimalité.

Plusieurs méthodes ont été inspiré du principe du maximum de Pontriaguine ; on les appelle des méthodes indirectes telle que la méthode de tir [88]. Dans ces méthodes, le calcul des variations est utilisé pour obtenir les conditions d'optimalité du premier ordre. Une autre approche est celle des méthodes directes. Ces dernières ont pris beaucoup d'importance au cours des trois dernières décennies. Dans une telle méthode, l'état et le contrôle sont estimés à l'aide d'une approximation d'une fonction appropriée, par exemple, par une approximation polynomiale, ou par un paramétrage constant par morceaux. Les coefficients des fonctions approximées sont traités comme des variables à déterminer et le problème est transcrit comme un problème d'optimisation non linéaire. L'éventail des problèmes qui peuvent être résolus par des méthodes directes est sensiblement supérieur à la gamme des problèmes qui peuvent être résolus par des méthodes indirectes.

Comme la résolution d'un problème de contrôle optimal est aujourd'hui souvent mis en oeuvre numériquement, la théorie du contrôle contemporaine concerne principalement les systèmes discontinues en temps. La théorie des approximations uniformes prévoit les conditions dans lesquelles des solutions d'une série de problèmes de contrôle optimal discrétisée, convergent vers la solution du problème original en temps continu. Malheureusement, toutes les méthodes de discrétisation ne possèdent pas cette propriété.

Une des méthodes pour résoudre des problèmes de contrôle, est basée sur la programmation linéaire ; c'est, sans doute, l'un des plus beaux succès de la recherche opérationnelle, en particulier, et des mathématiques appliquées, en général. C'est une technique qui s'intéresse aux méthodes de résolution du problème et qui consiste à minimiser ou à maximiser une fonction linéaire, appelée fonction objectif, sur un domaine délimité par un ensemble d'équations ou d'inéquations linéaires, appelées contraintes. Son origine remonte au dix-septième siècle lorsque Lagrange résolut en

1762, un problème d'optimisation avec un ensemble de contraintes linéaires. La première application des techniques de la programmation linéaire à l'économie fût réalisée par le mathématicien et économiste russe L.V. Kantorovich en 1939, et concerne un problème de planification de la production [44]. En raison de son importance tant sur le plan pratique que sur le plan théorique, la programmation linéaire a suscité, et suscite encore, l'intérêt d'un grand nombre de chercheurs à travers le monde. Par conséquent, plusieurs algorithmes intéressants de résolution des problèmes de programmation linéaire ont été introduits. Quelques méthodes ont été étudiées avant 1947 [24]; elles sont restreintes à la résolution certaines formes de problèmes de programmation linéaire. Elles étaient inspiré par le travail de Fourier sur les inégalités linéaire, et celui de G.B.Dantzig en 1947, qui a développé la méthode de simplexe [17]. La méthode de simplexe est efficace pour résoudre les problèmes pratiques de programmation linéaire. Cependant en 1972, Klee et Minty ont trouvé un exemple où la méthode de simplexe est de complexité exponentiel [47]. En 1979, Khachian développe un premier algorithme de point intérieur en temps polynomial pour résoudre un problème de programmation linéaire [46]. Par la suite en 1984, Karmarkar présente, l'algorithme de point intérieur en temps polynomial concurrentiel à la méthode de simplexe pour résoudre les problèmes de grande dimension [25].

En plus, des méthodes de points extrêmes (méthodes de simplexe et ses variantes) et les méthodes de points intérieurs, il existe d'autres méthodes pour résoudre les problèmes de programmation linéaire. On peut citer en quelques unes : les méthodes actives, la méthode de Gill et al. en 1973 [38],[39], la méthode de Gondzio en 1996 [40], la méthode de Santo-Palomo en 2004 [84], les méthodes "criss-cross", la méthode de Zions en 1969 [84], la méthode de Terlaky en 1985 [87], les méthodes de support, la méthode de Gabasov et Kirillova en 1977 [32]. Les méthodes de recherches à l'intérieur de la méthode de simplexe qui sont la méthode de Mitra et al. en 1988 [61], la méthode de simplexe généralisée, la méthode de Cardoso et Climaco en 1992 [20], l'algorithme "deficient-basis" de simplexe, la méthode de Pan en 1998 [69]. Les méthodes de points extérieurs introduites par Paparrizos et al. en 2003 [70], etc.

Dans [32], Les auteurs ont développé la méthode du support qui est une généralisation de la méthode de simplexe. Cette méthode de points intérieurs appelée aussi méthode adaptée utilisa certains métriques du simplexe. Elle a été utilisée pour la résolution de certains types de problèmes linéaires de contrôle optimal [13], [17], [27], [36], [48], [74], [78]. Une expérimentation sur les problèmes de programmation linéaire a montré son efficacité [48].

Le but de notre travail est la résolution de différents problèmes de contrôle optimal, principalement le problème avec une entrée libre. Chercher la condition initiale optimale, et une commande qui nous permet de ramener le système de l'état initial $x_0 \in X_0$, où X_0 est un ensemble fermé, vers l'état

final vérifiant une certaine contrainte $Hx(t^*) = g$. Trouver la meilleure condition initiale a beaucoup d'importance; nous renvoyons à [41] pour l'exposé de l'un des travaux effectués pour montrer son intérêt. Dans la pratique, on trouve quelques applications dans [42], [63], [64], [66], [83], [85] intervenant dans le domaine de l'aérospatiale, les techniques d'aérocaptures : les problèmes de guidage, transferts d'orbites aéroassistés, développement de lanceurs de satellites récupérables (l'enjeu financier est très important).

Par ailleurs, nous avons mis en œuvre la méthode de relaxation [50], [51], [52] couplée à la méthode de tir [88], pour résoudre un problème de contrôle optimal où l'état initial est libre. Par rapport aux travaux de ([59], [60]) où il n'y a pas de contraintes sur l'état final, la situation est plus complexe. En général, les conditions de la méthode de tir se traduisent par la formulation d'un problème aux deux bouts qui possède une structure particulière, car elles découlent de la dérivation du Hamiltonien. Le tir simple consiste à trouver un zéro de la fonction de tir associée au problème original. Il n'y a pas ici de discrétisation explicite, même si la méthode requiert l'intégration numérique du système et par conséquent la discrétisation temporelle explicite ou implicite; généralement on utilise des schémas explicite en temps. Le choix de ces méthodes se justifie par leurs avantages bien connus, à savoir une grande précision; cependant la convergence n'est pas garantie à coup sûr. Toutefois, l'inconvénient de cette méthode est la nécessité de disposer d'une donnée initiale de la commande suffisamment proche de la valeur exacte et permettant la convergence de l'algorithme bien qu'en général la convergence ne soit pas forcément assurée. Une des démarches classiques consiste à appliquer un algorithme de quasi-Newton à la fonction de tir; suivant la régularité du problème, le rayon de convergence peut être très faible. Ceci est particulièrement vrai pour des problèmes à contrôle Bang-Bang.

Le plan de notre manuscrit est le suivant :

Le premier chapitre sera consacré aux notions sur la résolution numérique d'équations différentielles qui sont la base fondamentale du contrôle optimal.

Au second chapitre, on s'intéresse à donner les fondements théoriques du contrôle, ainsi que quelques méthodes de résolution qui sont les méthodes indirectes découlant du principe du maximum de Pontriaguine et la méthode directe appelée la méthode adaptée, elle même constituée de trois procédures : le changement de commande, le changement de support et la procédure finale.

Le troisième chapitre est consacré à l'application des méthodes directes du 2^{ème} chapitre, à un problème terminal de contrôle optimal avec une entrée libre .

Au quatrième chapitre, on donne une nouvelle approche de résolution d'un problème de contrôle optimal qui consiste à un couplage de deux mé-

thodes : la méthode de discrétisation qui est constitué de deux procédures : le changement de commande et le changement de support, la solution du problème discret nous permet de trouver le support du problème de départ, en utilisant la procédure finale basée sur la méthode de Newton.

Au cinquième chapitre, on résout le problème comportant une entrée libre par la nouvelle approche décrite au quatrième chapitre.

Au sixième chapitre, on définit la méthode de tir et on étudie la convergence de la méthode de Newton discrète, intervenant dans la méthode de tir.

Au dernier chapitre, on résout numériquement un problème simple couplant la méthode de relaxation et la méthode de tir.

Le manuscrit se termine par une conclusion générale et quelques perspectives.

Liste des travaux scientifiques

Ce travail a été présenté dans des revues internationales et dans plusieurs manifestations scientifiques.

Présentation dans une revue Internationale

1. Kahina Louadj, Mohamed Aidene, "Optimization of an optimal control problem with free initial state", Applied Mathematical Sciences, Vol. 4, no. 5, 201 - 216, 2010.
2. Kahina Louadj, Mohamed Aidene, " Direct Method for Resolution of Optimal Control Problem with Free Initial Condition", International Journal of Differential Equations, Volume 2012 (2012), Article ID 173634, 18 pages.

Présentation dans des manifestations internationale

3. Kahina Louadj, Bachir Sadir, Brahim Oukacha, "Génération de clique maximale d'un graphe", Colloque International sur l'optimisation et les systèmes d'information COSI'05, organisé par l'université de Béjaia du 12 au 14 juin 2005.
4. Kahina Louadj, Mohamed Aidene, "A problem of optimal control with free initial condition", Colloque sur l'optimisation et les systèmes d'information COSI'09 organisé par l'université de Annaba du 25 au 27 mai 2009.
5. kahina Louadj, Mohamed Aidene, "Optimization d'un système dynamique avec une entrée libre", The first International Symposium on Operational Research, ISOR'08, organisé par l'université USTHB du 02 au 06 Novembre 2008.
6. Kahina Louadj, Mohamed Aidene, "Résolution d'un problème de contrôle optimal à variables mixtes", Colloque sur l'optimisation et les systèmes d'information COSI'08 organisé par l'université Tizi-Ouzou du 08 au 10 juin 2008.
7. Kahina Louadj, Nathalia Balashevich, "Optimization of hybrid systems with varying initial state", International conference " Dynamical systems : Stability, control, optimization", qui s'est tenue à Minsk, Belarus du 29 septembre au 04 octobre 2008.
8. Kahina Louadj, Mohamed Aidene, Pierre Spiteri, Frédéric Messine, " Analytical and Numerical solutions of a bi-criteria Optimal Control Problem", journées scientifiques du Laboratoire de Mathématiques de

la décision LAROMAD organisé par l'université de Tizi-Ouzou du 28 au 30 Novembre 2011.

9. Kahina Louadj, Pierre Spiteri, Mohamed Aidene, Frédéric Messine, "Solution of an optimal control problem by relaxation method", International Symposium on Operational Research ISOR'2011, organisé par l'USTHB du 30 mai au 02 juin 2011.
10. Kahina Louadj, Mohamed Aidene, " A problem of Optimal control with free initial condition under disturbance", conférence international en recherche Opérationnelle CIRO'10 s'est tenue à Marrakech du 24 au 27 mai 2010.
11. Kahina Louadj, Mohamed Aidene, "A problem of optimal control with free initial state", Proceeding du Congrès National de Mathématiques Appliquées et Industrielles, SMAI'11 qui s'est tenu à Orléans du 23 mai au 27 mai 2011, pages 184 - 190.
12. Kahina Louadj, Pierre Spiteri, Mohamed Aidene, Frédéric Messine, "Résolution par la méthode de relaxation d'un problème de contrôle optimal avec une entrée libre", Colloque sur l'optimisation et les systèmes d'information COSI'2012 organisé par l'université de Tlemcen du 12 mai au 15 mai 2012.

Présentation dans des manifestations nationales

13. Kahina Louadj, Mohamed Aidene, "Résolution d'un problème de contrôle optimal à variables mixtes", Rencontre sur l'Analyse Mathématiques et ses applications (RAMA VI) organisé par l'université de Tizi-Ouzou du 26 au 28 avril 2008.

Chapitre 1

Méthodes numériques de résolution des systèmes différentiels ordinaires

1.1 Introduction

Les équations différentielles ordinaires jouent un rôle important dans l'étude des applications dans le domaine de la physique, de la chimie, de l'économie, ...etc. Elles interviennent également pour la détermination de loi de commande optimale de systèmes gouvernés par des équations différentielles ordinaires.

Cependant, la résolution analytique des équations différentielles ordinaires n'est pas toujours possible. Depuis de nombreuses années on assiste à l'explosion de l'utilisation de méthodes numériques pour la résolution d'équations et de systèmes différentielles ordinaires.

Il est à noter que ces méthodes numériques bien qu'efficaces lorsque les méthodes analytiques sont défaillantes, permettent toutefois d'obtenir des solutions approchées. On est donc amené à déterminer une majoration d'erreur entre la solution approchée et la solution exacte. De plus, l'utilisation des méthodes numériques nécessite l'étude :

- De la consistance des schémas numériques, c'est à dire la pertinence du procédé d'approximation de la solution du problème d'évolution,
- De la stabilité des schémas numériques, c'est à dire de l'effet de la propagation des erreurs systématiques découlant du procédé d'approximation ainsi que de la propagation des erreurs de chute liées à la mauvaise représentation des nombres réels en machine,
- De la convergence de la solution numérique, vers la solution exacte, lorsque le pas de discrétisation en temps tend vers zéro.

Ces trois notions sont liées par un résultat théorique qui spécifie que la consistance et la stabilité d'un schéma entraîne la convergence de ce schéma.

Les schémas numériques de résolution d'équations différentielles ordinaires sont nombreux. Il paraît donc peu opportun de les présenter tous et c'est pourquoi dans ce chapitre nous nous limitons aux schémas à un pas : la méthode d'Euler, la méthode de Taylor et la méthode de Runge-Kutta.

1.2 Problème de Cauchy

Soit I_0 désigne un intervalle de \mathbb{R} non réduit à un point et t_0 un point fixé dans I_0 ; on se donne une fonction f définie et continue sur $I_0 \times \mathbb{R}^m$ à valeurs dans \mathbb{R}^m , ainsi qu'un élément y_0 de \mathbb{R}^m , et on cherche à trouver une fonction y continue et dérivable sur l'intervalle I_0 , à valeurs dans \mathbb{R}^m , telle que

$$\forall t \in I_0, y'(t) = f(t, y(t)), \quad (1.1)$$

et

$$y(t_0) = y_0. \quad (1.2)$$

Ce problème s'appelle un *problème de Cauchy* pour le système différentiel (1.1) ; la condition (1.2) s'appelle une *condition initiale*. Une fonction y qui vérifie les équations (1.1) est appelée une intégrale du système différentiel (1.1), (1.2). Nous nous intéresserons plus particulièrement au cas où I_0 est de la forme $[t_0, t^*]$; les cas où I_0 est de la forme $[t_0, t^*[$ ou $[t_0, +\infty[$ se traiteraient de même.

Dans de nombreux exemples physiques, la variable t représente le temps ; l'instant t_0 est alors appelé *instant initial*.

1.3 Théorèmes d'existence et d'unicité

1.3.1 Théorème d'existence

Théorème 1 (Cauchy-Péano) [23] *On suppose que la fonction f est continue dans un voisinage du point (t_0, y_0) dans $I_0 \times \mathbb{R}^m$; alors il existe un intervalle $J_0 \subset I_0$, au voisinage de t_0 et une fonction $y \in C^1(J_0)$ tels que*

$$\forall t \in J_0, y'(t) = f(t, y(t)), y(t_0) = y_0.$$

Définition 1.1 *On appelle solution locale du problème (1.1), (1.2) la donnée d'un couple (I, y) où I est un intervalle de \mathbb{R} qui est un voisinage de t_0 inclus dans I_0 et où y est une fonction appartenant à $C^1(I)$ telle que*

$$y(t_0) = y_0 \quad \text{et} \quad \forall t \in I, y'(t) = f(t, y(t))$$

Définition 1.2 *On dit que la solution locale (J, l) prolonge la solution locale (I, y) si on a $I \subset J$, et $\forall t \in I, y(t) = l(t)$; si de plus $I \neq J$, on dit que (J, l) prolonge strictement (I, y) .*

Définition 1.3 On dit que la solution locale (J, z) prolonge la solution locale (I, y) si on a $I \subset J$, et $\forall t \in I; y(t) = z(t)$; si de plus $I \neq J$, on dit que (J, z) prolonge strictement (I, y) .

Définition 1.4 On dit que la solution locale (I, y) est une solution maximale du problème (1.1), (1.2) s'il n'existe pas de solution locale de ce problème qui la prolonge strictement.

Définition 1.5 On dit que (I, y) est une solution globale du problème (1.1), (1.2) dans I_0 , (ou encore que y est solution du problème (1.1), (1.2)), si (I, y) est une solution locale de ce problème, et si $I = I_0$.

1.3.2 Théorème d'unicité

Définition 1.6 On dira que le problème (1.1), (1.2) admet une solution et une seule, s'il admet une solution globale et si toute solution locale est la restriction de cette solution globale.

Théorème 2 [23] On suppose que I_0 est de la forme $[t_0, t^*]$ ou $[t_0, t^*[$ ou $[t_0, +\infty[$, que f est continue sur $I_0 \times \mathbb{R}^m$ et qu'il existe une fonction $l \in \mathfrak{L}(I_0)$ telle que

$$\forall t \in I_0, \forall y, z \in \mathbb{R}^m, (f(t, y) - f(t, z), y - z) \leq l(t)|y - z|^2, \quad (1.3)$$

alors le problème (1.1), (1.2) admet une solution et une seule.

Avec $\mathfrak{L}(I_0)$ est un espace vectoriel normé des fonctions réelles mesurables sur I_0 telle que :

$$\|f\| = \int_{I_0} |f(x)| d\mu < +\infty.$$

Corollaire 1.1 [Cauchy-Lipschitz][23] On suppose que la fonction f est continue sur $I_0 \times \mathbb{R}^m$ et qu'il existe un réel L tel que

$$\forall (t, y) \text{ et } (t, z) \in I_0 \times \mathbb{R}^m, |f(t, y) - f(t, z)| \leq L|y - z|;$$

alors le problème (1.1), (1.2) admet une solution et une seule.

1.4 Méthode d'Euler

Nous allons maintenant nous consacrer à la résolution du problème de *Cauchy* décrit précédemment en nous restreignant, ce qui n'est pas une perte de généralité, au cas où $I_0 = [t_0, t^*]$. Tout d'abord nous allons étudier la plus simple des méthodes, la méthode d'Euler connue aussi sous le noms

de méthode d'Euler progressive. Considérons donc le problème différentiel suivant :

$$\forall t \in [t_0, t^*], y'(t) = f(t, y(t)), \quad (1.4)$$

$$y(0) = \eta \quad \text{donné dans } \mathbb{R}^m. \quad (1.5)$$

Nous nous donnons une subdivision $t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_N = t^*$ de l'intervalle $[t_0, t^*]$, et nous posons

$$h_n = t_{n+1} - t_n \quad \text{et } h = \max_{0 \leq n \leq N} h_n.$$

Les solutions de (1.4) vérifient, pour $0 \leq n < N$,

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(s, y(s)) ds;$$

soit $y_0 = \eta_h$ une approximation de $y(t_0) = \eta$; nous construisons par récurrence une approximation y_n de $y(t_n)$ en remplaçant la relation précédente par

$$y_{n+1} = y_n + h_n f(t_n, y_n), \quad n = 0, 1, \dots, N-1, \quad (1.6)$$

Ce qui revient à approcher, pour $s \in]t_n, t_{n+1}[$, $f(s, y(s))$ par $f(t_n, y_n)$. Le schéma défini par (1.6) s'appelle le schéma d'*Euler*.

Nous cherchons à obtenir une estimation de l'erreur

$$e_n = y(t_n) - y_n$$

entre la solution exacte de (1.4), (1.5) et la solution approchée donnée par (1.6); pour cela nous supposons f continue sur $[t_0, t^*] \times \mathbb{R}^m$ et nous ferons l'hypothèse de Lipschitz du corollaire 1.1.

1.4.1 Majoration de l'erreur dans la méthode d'Euler

Posons

$$\varepsilon_n = y(t_{n+1}) - y(t_n) - h_n f(t_n, y(t_n)), \quad (1.7)$$

où $y(\cdot)$ désigne la solution de (1.4), (1.5); cette quantité mesure avec quelle précision la solution exacte vérifie le schéma (1.6); elle s'appelle l'*erreur de consistance* à l'instant t_n de la méthode d'Euler. Nous avons

$$\varepsilon_n = \int_{t_n}^{t_{n+1}} (y'(t) - y'(t_n)) dt,$$

d'où

$$|\varepsilon_n| \leq \omega(y'; h) h_n \quad (1.8)$$

où $\omega(y'; h)h_n$ désigne le module de continuité de la fonction y' ; c'est-à-dire la quantité définie par

$$\omega(f, h) = \max_{t, t' \in I, |t-t'| \leq h} |f(t) - f(t')|.$$

Puisque $y' \in C^0[t_0, t^*]$, $\lim_{h \rightarrow 0} \omega(y'; h) = 0$. Si on suppose de plus que $y \in C^2[t_0, t^*]$, on obtient par utilisation de la formule de Taylor dans (1.7)

$$\varepsilon_n = \int_{t_n}^{t_{n+1}} (t_{n+1} - s)y''(s) ds,$$

d'où

$$|\varepsilon_n| \leq h \int_{t_n}^{t_{n+1}} |y''(s)| ds. \quad (1.9)$$

On déduit de (1.6) et (1.7) la relation

$$e_{n+1} = e_n + h_n[f(t_n, y(t_n)) - f(t_n, y_n)] + \varepsilon_n$$

d'où en utilisant la condition de Lipchitz

$$|e_{n+1}| \leq (1 + h_n L)|e_n| + |\varepsilon_n|. \quad (1.10)$$

Pour en déduire une majoration de $|e_n|$, nous allons utiliser le lemme technique fondamentale appelé lemme de GRONWALL suivant :

Lemme 1.1 [23] Soient θ_n et α_n deux suites de réels positifs ou nuls vérifiant

$$\forall n \geq 0, \quad \theta_{n+1} \leq (1 + h_n L)\theta_n + \alpha_n,$$

alors on a

$$\forall n \geq 0, \quad \theta_n \leq e^{L(t_n - t_0)}\theta_0 + \sum_{i=0}^{n-1} e^{L(t_n - t_{i+1})}\alpha_i.$$

Si on suppose de plus que $y \in C^2[t_0, t^*]$, on déduit de même de (1.9) et (1.10)

$$|y(t_n) - y_n| \leq e^{L(t_n - t_0)}|\eta - \eta_h| + h \int_{t_0}^{t_n} e^{L(t_n - s)}|y''(s)| ds. \quad (1.11)$$

1.5 Méthode numérique à un pas

On appelle *méthode à un pas*, une méthode permettant de calculer une valeur approchée y_{n+1} de $y(t_{n+1})$ à partir de t_n , de h_n et de la valeur approchée y_n de $y(t_n)$ obtenue au pas précédent. Une méthode à r pas est au contraire une méthode qui utilise les r valeurs antérieurs y_n, \dots, y_{n-r+1} (valeurs qui doivent donc être mémorisées) afin de faire le calcul de y_{n+1} .

1.5.1 Définition

Les méthodes à un pas sont les méthodes de résolution numérique qui peuvent s'écrire sous la forme

$$y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n), \quad n \geq 0, \quad (1.12)$$

$$y_0 = \eta_h \quad \text{donné dans } \mathbb{R}, \quad (1.13)$$

et nous supposons que Φ est une fonction continue de $[t_0, t^*] \times \mathbb{R} \times \mathbb{R} \times [0, h^*]$, $h^* > 0$, à valeurs réelles, qui ne dépend que de la fonction f . Notons que la méthode d'*Euler* est la méthode à un pas qui correspond à

$$\Phi(t, y, h) = f(t, y);$$

dans ce cas Φ est indépendant de h .

1.5.2 Méthode de Taylor d'ordre p

Supposons que f soit de classe C^p , alors toute solution exacte $y(\cdot)$ est de classe C^{p+1} , et on a défini des fonctions $f^{(k)}$, construite par récurrence à partir de f et de ses dérivées partielles telles que $y^{(k)}(t) = f^{(k-1)}(t, y(t))$; pour $k = 1, \dots, p+1$. La formule de Taylor à l'ordre $p+1$ s'écrit alors :

$$y(t_n + h_n) = y(t_n) + \sum_{k=1}^p \frac{1}{k!} h_n^k f^{(k-1)}(t, y(t_n)) + \frac{1}{(p+1)!} f^{(p)}(t_n, y(t_n)) h_n^{p+1}.$$

Lorsque h_n est assez petit, l'approximation est d'autant meilleure que p est plus grand. On est donc amené à considérer l'algorithme suivant, appelé méthode de Taylor d'ordre p :

$$\begin{cases} y_{n+1} = y_n + \sum_{k=1}^p \frac{1}{k!} h_n^k f^{(k-1)}(t_n, y_n), \\ t_{n+1} = t_n + h_n, \end{cases}$$

D'après la définition générale des méthodes à un pas, cet algorithme correspond au choix $\Phi(t, y, h) = \sum_{k=1}^p \frac{1}{k!} h^{k-1} f^{(k-1)}(t, y)$. Calculons l'erreur de

consistance ε_n . En supposant $y_n = y(t_n)$, la formule de Taylor d'ordre $p + 1$ donne

$$\begin{aligned}\varepsilon_n &= y(t_{n+1}) - y_{n+1} \\ &= y(t_n + h_n) - \sum_{k=0}^p \frac{1}{k!} h_n^k y^{(k)}(t_n) \\ &= \frac{1}{(p+1)!} h_n^{p+1} f^{(p)}(t_n, y_n) + O(h_n^{p+1}).\end{aligned}$$

L'erreur est donc de l'ordre de h_n^{p+1} . On dira d'une manière générale qu'une méthode est d'ordre p si l'erreur de consistance est en h_n^{p+1} chaque fois que f est de classe C^p au moins. La méthode d'Euler est le cas particulier où $p = 1$ pour la méthode de Taylor.

1.6 Méthode du point milieu

Notons M_n le point coordonnées $(t_n, y(t_n))$ du graphe de y . Le segment $[M_n, M_{n+1}]$ a une pente plus proche en général de $y'(t_n + \frac{h_n}{2})$ pente de la tangente au "point milieu" que de $y'(t_n)$, pente de la tangente de M_n .

On peut considérer qu'une approximation de $y(t_{n+1})$ à partir de $y(t_n)$ meilleur que l'expression $y(t_n + h_n)$ de la méthode d'Euler est :

$$y(t_n + h_n) \simeq y(t_n) + h y'(t_n + \frac{h_n}{2}) \quad (1.14)$$

Si y est de classe C^3 , avec $h_n = h$ il vient alors

$$\begin{aligned}y(t+h) &= y(t) + h y'(t) + \frac{1}{2} h^2 y''(t) + \frac{1}{6} h^3 y'''(t) + O(h^3); \\ y'(t + \frac{h}{2}) &= y'(t) + \frac{1}{2} h y''(t) + \frac{1}{6} h^2 y'''(t) + O(h^2)\end{aligned}$$

L'erreur comise est donc :

$$e = y(t+h) - y(t) - h y'(t + \frac{h}{2}) = \frac{1}{6} h^3 y'''(t) + O(h^3);$$

Soit une erreur en h^3 au lieu de h^2 dans la méthode d'Euler. On a par ailleurs

$$y'(t + \frac{h}{2}) = f(t + \frac{h}{2}, y(t + \frac{h}{2}))$$

Comme la valeur de $y(t + \frac{h}{2})$ n'est pas connue, on l'approxime par

$$y(t + \frac{h}{2}) \simeq y(t) + \frac{h}{2} f(t, y(t)), \quad (1.15)$$

d'où en définitive

$$y(t+h) \simeq y(t) + hf\left(t + \frac{h}{2}, y(t) + \frac{h}{2}f(t, y(t))\right).$$

L'algorithme du point milieu est associé au choix

$$\Phi(t, y, h) = f\left(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)\right)$$

et donne lieu au schéma numérique

$$\begin{cases} y_{n+\frac{1}{2}} = y_n + \frac{h_n}{2}f(t_n, y_n), \\ p_n = f\left(t_n + \frac{h_n}{2}, y_{n+\frac{1}{2}}\right), \\ y_{n+1} = y_n + h_n p_n, \\ t_{n+1} = t_n + h_n, \end{cases}$$

En utilisant les approximations faites en (1.14) et (1.15), on calcule l'erreur de consistance, et on en déduit que la méthode du point milieu est d'ordre 2.

1.6.1 Notions de consistance, de stabilité et de convergence

Définition 1.7 On dit qu'une méthode est consistante si pour toute solution exacte $y(\cdot)$ la somme des erreurs de consistances relatives à $y(\cdot)$, soit $\sum_{0 \leq n \leq N} |\varepsilon_n|$, tend vers 0 quand h_{max} tend vers 0.

Une autre notion importante est la notion de stabilité. Dans la pratique, le calcul récurrent des points y_n est en effet entaché d'erreurs d'arrondi e_n . Pour que les calculs soient significatifs, il est indispensable que la propagation de ces erreurs reste contrôlable. On est amené à la définition suivante :

Définition 1.8 On dit que la méthode est stable s'il existe une constante $S \geq 0$, telle que pour toutes suites $(y_n), (\tilde{y}_n)$ définies par

$$\begin{aligned} y_{n+1} &= y_n + h_n \Phi(t_n, y_n, h_n), & 0 \leq n \leq N \\ \tilde{y}_{n+1} &= \tilde{y}_n + h_n \Phi(t_n, \tilde{y}_n, h_n) + e_n, & 0 \leq n \leq N \end{aligned}$$

on ait

$$\max_{0 \leq n \leq N} |\tilde{y}_n - y_n| \leq S(|\tilde{y}_0 - y_0| + \sum_{0 \leq n \leq N} |e_n|).$$

Autrement dit, une petite erreur sur la donnée initiale $|\tilde{y}_0 - y_0|$ et de petites erreurs d'arrondis dans le calcul de \tilde{y}_n provoquent une erreur finale $\max |\tilde{y}_n - y_n|$ contrôlable. Une dernière notion importante est la suivante :

Définition 1.9 On dit que la méthode est convergente si pour toute solution exacte $y(\cdot)$, pour $t_n = nh_n$ fixé, la suite y_n telle que

$$y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n) \quad \text{vérifie} \quad \max_{0 \leq n \leq N} |y_n - y(t_n)| \rightarrow 0$$

$$\text{quand } y_0 \rightarrow y(t_0) \quad \text{et quand } h_{\max} \rightarrow 0$$

Théorème 3 Si la méthode à un pas est stable et consistante, elle est convergente.

1.7 Runge-Kutta

Principe Général On considère un problème de *Cauchy*

$$\begin{cases} y' = f(t, y), & t \in [t_0, t^*], \\ y(t_0) = y_0, \end{cases}$$

et on cherche à discrétiser ce problème par rapport à une subdivision $t_0 < t_1 < \dots < t_N = t^*$. L'idée est de calculer par récurrence les points (t_n, y_n) en utilisant des points intermédiaires $(t_{n,i}, y_{n,i})$ avec

$$t_{n,i} = t_n + c_i h_n, \quad 1 \leq i \leq q, \quad c_i \in [0, 1].$$

A chacun de ces points, on associe la pente correspondante

$$p_{n,i} = f(t_{n,i}, y_{n,i})$$

Soit y une solution exacte de l'équation. On a

$$\begin{aligned} y(t_{n,i}) &= y(t_n) + \int_{t_n}^{t_{n,i}} f(t, y(t)) dt \\ &= y(t_n) + h_n \int_0^{c_i} f(t_n + u h_n, y(t_n + u h_n)) du \end{aligned}$$

grâce au changement de variable $t = t_n + u h_n$. De même

$$y(t_{n+1}) = y(t_n) + h_n \int_0^1 f(t_n + u h_n, y(t_n + u h_n)) du$$

On se donne alors pour chaque $i = 1, 2, \dots, q$ une méthode d'intégration approchée

$$\int_0^{c_i} g(t) dt \simeq \sum_{1 \leq j \leq i} a_{ij} g(c_j), \quad (1.16)$$

ces méthodes pouvant être *a priori* différentes. On se donne également une méthode d'intégration approchée sur $[0, 1]$:

$$\int_0^1 g(t) dt \simeq \sum_{1 \leq j < q} b_j g(c_j). \tag{1.17}$$

En appliquant ces méthodes à $g(u) = f(t_n + uh_n, y(t_n + uh_n))$, il vient

$$y(t_{n,i}) \simeq y(t_n) + h_n \sum_{1 \leq j \leq i} a_{ij} f(t_{n,j}, y(t_{n,j})),$$

$$y(t_{n+1}) \simeq y(t_n) + h_n \sum_{1 \leq j < q} b_j f(t_{n,j}, y(t_{n,j}))$$

La méthode de Runge-Kutta correspondante est définie par l'algorithme

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} t_{n,i} = t_n + c_i h_n, \\ y_{n,i} = y_n + h_n \sum_{1 \leq j < i} a_{ij} p_{n,j}, \\ p_{n,i} = f(t_{n,i}, y_{n,i}) \end{array} \right. \quad 1 \leq i \leq q, \\ \\ t_{n+1} = t_n + h_n, \\ y_{n+1} = y_n + h_n \sum_{1 \leq j \leq q} b_j p_{n,j}, \end{array} \right.$$

On la représente traditionnellement par le tableau

(M_1)	c_1	0	0	...	0	0
(M_2)	c_2	a_{21}	0	...	0	0
	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
	\vdots	\vdots	\vdots		0	0
(M_q)	c_q	a_{q1}	a_{q2}	\cdots	a_{qq-1}	0
(M)		b_1	b_2	\cdots	b_{q-1}	b_q

dans lequel les méthodes d'intégration approchées correspondent aux lignes. On pose par convention $a_{ij} = 0$ pour $j \geq i$.

Exemples

Exemple 1.1 Pour $q = 1$, le seul choix possible est

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

On a ici

$c_1 = 0, a_{11} = 0, b_1 = 1$. L'algorithme est donné par

$$\begin{cases} p_{n,1} = f(t_n, y_n), \\ t_{n+1} = t_n + h_n, \\ y_{n+1} = y_n + h_n p_{n,1}, \end{cases}$$

Il s'agit de la méthode d'Euler.

Exemple 1.2 Pour $q = 2$, on considère les tableaux de la forme

$$\begin{array}{c|cc} 0 & 0 & 0 \\ a & a & 0 \\ \hline & 1 - \frac{1}{2a} & \frac{1}{2a} \end{array}$$

L'algorithme s'écrit ici

$$\begin{cases} p_{n,1} = f(t_n, y_n), \\ t_{n,2} = t_n + \alpha h_n, \\ y_{n,2} = y_n + \alpha h_n p_{n,1}, \\ p_{n,2} = f(t_{n,2}, y_{n,2}), \\ t_{n+1} = t_n + h_n, \\ y_{n+1} = y_n + h_n \left(\left(1 - \frac{1}{2\alpha}\right) p_{n,1} + \frac{1}{2\alpha} p_{n,2} \right), \end{cases}$$

ou encore, sous forme condensée :

$$y_{n+1} = y_n + h_n \left(\left(1 - \frac{1}{2\alpha}\right) f(t_n, y_n) + \frac{1}{2\alpha} f(t_n + \alpha h_n, y_n + \alpha h_n f(t_n, y_n)) \right),$$

- Pour $\alpha = \frac{1}{2}$, on retrouve la méthode du point milieu

$$y_{n+1} = y_n + h_n f\left(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, y_n)\right),$$

qui est basée sur la formule d'intégration du point milieu :

$$(M) \quad \int_0^1 g(t) dt \simeq g\left(\frac{1}{2}\right).$$

- Pour $\alpha = 1$, on obtient la méthode de Heun :

$$y_{n+1} = y_n + h_n \left(\frac{1}{2} f(t_n, y_n) + \frac{1}{2} f(t_n + h_n, y_n + h_n f(t_n, y_n)) \right),$$

qui repose sur la formule d'intégration des trapèzes :

$$(M) \quad \int_0^1 g(t) dt \simeq \frac{1}{2} (g(0) + g(1)).$$

Exemple 1.3 Méthode de Runge-Kutta "classique" :
 Il s'agit de la méthode définie par le tableau

$q = 4$,

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

L'algorithme correspondant s'écrit

$$\left\{ \begin{array}{l} p_{n,1} = f(t_n, y_n), \\ t_{n,2} = t_n + \frac{1}{2}h_n, \\ y_{n,2} = y_n + \frac{1}{2}h_n p_{n,1}, \\ p_{n,2} = f(t_{n,2}, y_{n,2}), \\ y_{n,3} = y_n + \frac{1}{2}h_n p_{n,2}, \\ p_{n,3} = f(t_{n,2}, y_{n,3}), \\ t_{n+1} = t_n + h_n, \\ y_{n,4} = y_n + h_n p_{n,3}, \\ p_{n,4} = f(t_{n+1}, y_{n,4}), \\ y_{n+1} = y_n + h_n \left(\frac{1}{6}p_{n,1} + \frac{2}{6}p_{n,2} + \frac{2}{6}p_{n,3} + \frac{1}{6}p_{n,4} \right), \end{array} \right.$$

Cette méthode est d'ordre 4. Dans ce cas les méthodes d'intégration (1.16) et (1.17) utilisées sont respectivement :

$$(M_2) \quad \int_0^{\frac{1}{2}} g(t) dt \simeq \frac{1}{2} g(0) : \quad \text{formule des rectangles à gauche,}$$

$$(M_3) \quad \int_0^{\frac{1}{2}} g(t) dt \simeq \frac{1}{2} g\left(\frac{1}{2}\right) : \quad \text{formule des rectangles à droite,}$$

$$(M_3) \quad \int_0^1 g(t) dt \simeq g\left(\frac{1}{2}\right) : \quad \text{formule du point milieu,}$$

$$(M) \quad \int_0^1 g(t) dt \simeq \frac{1}{6} g(0) + \frac{2}{6} g\left(\frac{1}{2}\right) + \frac{2}{6} g\left(\frac{1}{2}\right) + \frac{1}{6} g(1) : \quad \text{formule de Simpson.}$$

Ordre de la méthode de Runge-Kutta

Théorème 4 [26] La méthode de Runge-Kutta définie par le tableau des coefficients c_i, a_{ij}, b_j est

- d'ordre ≥ 2 ssi $\sum_j b_j c_j = \frac{1}{2}$.
- d'ordre ≥ 3 ssi $\sum_j b_j c_j = \frac{1}{2}$; $\sum_j b_j c_j^2 = \frac{1}{3}$; $\sum_{i,j} b_i a_{ij} c_j = \frac{1}{6}$.
- d'ordre ≥ 4 ssi

$$\begin{aligned} \sum_j b_j c_j &= \frac{1}{2}; & \sum_j b_j c_j^2 &= \frac{1}{3}; & \sum_j b_j c_j^3 &= \frac{1}{4} \\ \sum_{i,j} b_i a_{ij} c_j &= \frac{1}{6}; & \sum_{i,j} b_i a_{ij} c_j^2 &= \frac{1}{12}; & \sum_{i,j} b_i c_i a_{ij} c_j &= \frac{1}{8}; \\ & & \sum_{i,j,k} b_i a_{ij} a_{jk} c_k &= \frac{1}{12}. \end{aligned}$$

On voit ainsi que la méthode d'Euler est d'ordre 1, et que les méthodes de l'exemple 1.2 sont d'ordre 2. De plus, dans une méthode avec $q = 2$, il y a *a priori* un seul coefficient a_{ij} non nul, à savoir $\alpha = a_{21}$. On a alors $c_2 = \sum_{j < 2} a_{2j} = \alpha$ et la méthode est d'ordre 2 au moins si et seulement si $\sum_j b_j c_j = b_2 \alpha = \frac{1}{2}$, soit $b_2 = \frac{1}{2\alpha}$ et $b_1 = 1 - b_2 = 1 - \frac{1}{2\alpha}$. On voit donc qu'il n'y avait pas d'autres choix possibles pour une méthode d'ordre 2 avec $q = 2$. Enfin, la méthode Runge-Kutta "classique" présentée dans l'exemple 1.3 est d'ordre 4. La méthode la plus utilisée car elle a un ordre élevé et une grande stabilité. Il existe des méthodes d'ordre encore plus élevé, mais leur plus grande complexité les rend peut-être un peu moins praticables.

Contrôle du pas

La manière la plus simple d'utiliser une méthode de résolution numérique consiste à utiliser un pas constant $h_n = h$. La principale difficulté est alors de déterminer h_{max} de façon que l'erreur globale ne dépasse pas une certaine tolérance fixée à l'avance ; on ne sait pas en effet quelle sera l'évolution de la solution étudiée, de sorte qu'il est difficile de prévoir *a priori* les erreurs de consistance. L'utilisation d'algorithmes à pas variables présente de ce point de vue deux avantages majeurs :

- l'adaptation du pas à chaque étape permet d'optimiser l'erreur commise en fonction de la tolérance prescrite, sous réserve qu'on dispose d'une estimation "instantanée" de l'erreur de consistance ε_n .
- l'approche d'une discontinuité ou d'une singularité de l'équation différentielle ne peut se faire généralement qu'avec une réduction importante du pas. Dans ce cas, il convient d'arrêter l'algorithme avant de traverser cette discontinuité, faute de quoi les erreurs deviennent imprévisibles. Le calcul du pas sert alors de test d'arrêt.

1.7.1 Fonctions Matlab `ode45`

`ode45` résout l'équation différentielle ordinaire par une méthode d'ordre 4 similaire à la méthode de Runge-Kutta classique d'ordre 4 ; en plus, `ode45` utilise un pas de temps variable et choisit à chaque instant le pas le plus convenable de façon à satisfaire une tolérance fixée.

Chapitre 2

Introduction à la commande optimale

2.1 Introduction

Les processus physique consomment tous une certaine quantité d'énergie lors de leur fonctionnement. Il apparaît souhaitable, lorsque cela est possible, de chercher à minimiser cette dépense en s'écartant le moins possible de l'objectif à atteindre.

Ainsi, lorsque l'on s'intéresse à la fabrication d'un produit, il est naturel de rechercher une procédure conduisant à un compromis entre, d'une part une qualité et une quantité de produit fini maximale et d'autre part une dépense minimale en énergie et en matière première, sans oublier d'atténuer au maximum les nuisances éventuelles de tous ordres occasionnées par la fabrication du produit.

Si on possède un modèle mathématique décrivant le processus physique et si on a fait le choix d'un objectif à atteindre par la donnée d'une fonctionnelle à minimiser, il est alors possible de poser le problème mathématiquement et de rechercher les solutions optimales éventuelles.

Ce chapitre est composé de deux parties : la première partie est consacrée à la présentation des concepts de bases d'un problème de contrôle. On retrouve dans la seconde partie, en utilisant une méthode directe issue de la programmation linéaire, qui est la méthode adaptée introduite par R. Gabasov et F.M.Kirillova au début des années 1980.

2.2 Théorie du contrôle optimal et des systèmes de contrôle

Le problème général de contrôle optimal est constitué des données suivantes :

2.2.1 Objet de la commande

Le système peut comporter beaucoup de variables ou paramètres. On suppose que n variables sont nécessaires pour décrire son comportement. L'identification de ces variables et la description du système dépendant de celles-ci est une tâche très importante : c'est l'étape de modélisation mathématique.

Les variables, nommées variables d'état seront notées x_i , $i = 1, \dots, n$. Le système évolue dans le temps, donc les x_i sont des fonctions de t : $x_i(t)$. On ne le notera pas explicitement, mais cela sera sous entendu. Les n variables d'états vont être gouvernées par n équations différentielles du premier ordre sur un intervalle de temps $T = [t_0, t^*]$; ce sont des équations de la forme générale :

$$\dot{x}^i = f^i(t, x_1, \dots, x_n, u_1, \dots, u_m), \quad i = 1, \dots, n.$$

Les variables de contrôle seront notées $u_j(t)$, $j = 1, \dots, m$. Elles sont soumises à l'hypothèse d'intégrabilité par rapport à t . Cela simplifie beaucoup les traitements si les u_j sont connues; cependant cette hypothèse est bien souvent trop restrictive car ces fonctions peuvent être continues par morceaux ou de type Bang-Bang.

2.2.2 Condition initiale du système

La condition initiale du système, $x_0 = x(t_0)$ est un vecteur donné dans un plan de phase. En réalité, les composantes de $x(t)$ et de x_0 peuvent représenter physiquement : la position, la vitesse, la température et d'autres paramètres mesurables.

2.2.3 Le but de la commande

Dans un problème de contrôle, le but de la commande consiste à ramener l'objet de la position initiale $x_0 = x(t_0)$, ($x_0 \in M_0$) à une autre position $x^* = x(t^*)$, ($x^* \in M_1$) où M_0 est l'ensemble de départ, et M_1 l'ensemble d'arrivé.

2.2.4 Classe des commandes admissibles

U est l'ensemble des contrôles admissibles qui peut être non bornés, borné ou du type Bang-Bang.

Commande bornée

Dans beaucoup de problèmes de contrôle, on peut minorer et majorer les $u_j(t)$ par des constantes. Dans la suite, nous considérons ce type de problème avec $a_j \leq u_j \leq b_j$. Notons que l'on peut remplacer u_j par v_j en posant $u_j = \frac{1}{2}(a_j + b_j) + \frac{1}{2}(a_j - b_j)v_j$ et ainsi v_j est aussi intégrable et l'on a

$-1 \leq v_j \leq 1$. Donc lorsque U est borné, il est toujours pratique de se ramener à des commandes entre -1 et 1 .

Commande Bang-Bang

On suppose que U est un polyèdre (cube) $[-1, 1]^m$ dans \mathbb{R}^m . Un contrôle $u \in U$ est appelé contrôle Bang-Bang si pour chaque instant t et chaque indice $j = 1, \dots, m$, on a $|u_j(t)| = 1$. En d'autres termes, une commande Bang-Bang est une commande qui possède au moins un switch.

2.2.5 Critère de qualité

L'objectif, lors de la formulation d'un problème de contrôle, est de fournir la motivation physique pour la sélection d'une mesure de qualité pour le système. Le problème revient à définir une expression mathématique qui, lorsqu'elle est optimisée, indique que le système atteint un état désirable. Donc, choisir une mesure de qualité, est une traduction en termes mathématiques des exigences physiques du système.

Le critère de qualité, appelé aussi coût ou fonction objectif, est généralement décrit par la formule

$$J(x, u) = g(t^*, x^*) + \int_{t_0}^{t^*} f_0(t, x, u) dt.$$

Cette fonctionnelle comporte deux parties : $g(t^*, x^*)$ est le coût terminal, c'est une sorte de pénalité liée à la fin de l'évolution du système au temps final t^* ; il a son importance lorsque t^* est libre, sinon il est constant. Le second terme intervenant dans la fonction objectif $\int_{t_0}^{t^*} f_0(t, x, u) dt$ dépend de l'état du système tout au long de la trajectoire de la solution, définie par les variables d'état. Cette trajectoire dépend aussi du temps t mais surtout des variables de contrôle u . C'est une fonction d'efficacité de chaque commande sur l'intervalle T .

On peut classer les fonctions objectifs en deux critères physiques de performance :

Temps optimal

On parle d'un problème en temps optimal lorsque $f_0(t, x, u) = 1$, $g(t^*, x^*) = 0$ et le temps final t^* est libre dans l'expression de $\min_u \int_{t_0}^{t^*} 1 dt$.

Coût optimal

On parle d'un problème en coût optimal lorsque le temps final t^* est fixé dans l'expression

$$\min_u g(t^*, x^*) + \int_{t_0}^{t^*} f_0(t, x, u) dt.$$

Évidemment, il existe des problèmes qui combine les deux critères physiques de qualité, et on parlera dans ce cas d'un problème de contrôle en temps et en coût optimal.

Si dans l'expression de J , f_0 est proportionnelle à u^2 , on parle alors d'un coût quadratique.

Si u est un contrôle scalaire et f_0 est proportionnelle à u , on parle de problème de contrôle à coût d'approvisionnement. Lorsque les équations d'état $\dot{x} = f(x, u)$ ne dépendent pas explicitement de la variable t , on parle dans ce cas de problème autonome. Si t est présent dans les équations d'état on parle de problème non-autonome.

Problème de Mayer-Lagrange

Le problème de Mayer-Lagrange est donné sous la forme d'un système $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = x_0$, $x(t^*) = x^*$, $u \in U$, $t \in [t_0, t^*]$, l'objectif étant de minimiser le coût

$$J(t^*, u) = g(t^*, x^*) + \int_{t_0}^{t^*} f_0(t, x, u) dt.$$

Lorsque $g = 0$ dans l'expression de la fonctionnelle J , on parlera d'un problème de Lagrange; lorsque $f_0 = 0$, on parlera d'un problème de Mayer.

2.3 Contrôlabilité

Pour déterminer une trajectoire optimal joignant un ensemble initial à une cible, il faut d'abord savoir si cette cible est atteignable; c'est le problème; de contrôlabilité. La contrôlabilité est l'un des concepts centraux de la théorie du contrôle. C'est la possibilité d'influencer l'état du système (sortie) en manipulant les entrées (commandes). Existe-t-il un contrôle u tel que la trajectoire associée x conduit le système de $x_0 \in M_0$ à $x^* \in M_1$ en un temps fini?

La notion de contrôlabilité a été introduite en 1960 par Kalman [43] pour des systèmes linaires de la forme $\dot{x} = Ax + Bu$. L'état x évolue dans un espace vectoriel réel E , de dimension n . On dit que $\dot{x} = Ax + Bu$ est contrôlable, si l'on peut joindre deux points de l'espace d'état, c'est à dire si, et seulement

si, étant donnés deux points $x_0, x^* \in E$ et deux instants t_0, t^* avec $t_0 < t^*$, il existe une commande u , définie sur $[t_0, t^*]$, telle que $x(t_0) = x_0, x(t^*) = x^*$.

2.3.1 Contrôlabilité des systèmes linéaires

La formulation mathématique d'un système de contrôle linéaire est la suivante :

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t), \quad x(0) = x_0, \quad t \in I,$$

où I est un intervalle de \mathbb{R} , A, B , et r sont trois applications localement intégrables sur I à valeurs respectivement dans $M_n(\mathbb{R})$, $M_{n,m}(\mathbb{R})$ et \mathbb{R}^m .

Où $M_n(\mathbb{R})$ est l'ensemble des matrices réelles de dimension n , et $M_{n,m}(\mathbb{R})$ est l'ensemble des matrices de n lignes et de m colonnes.

L'ensemble des contrôles u considérés est l'ensemble des applications mesurables localement bornées sur I à valeurs dans un sous ensemble $U \subset \mathbb{R}^m$.

Soit $F(\cdot) : I \rightarrow M_n(\mathbb{R})$ la résolvante du système linéaire homogène $\dot{x}(t) = A(t)x(t)$ définie par

$$\begin{cases} \dot{F} = A(t)F(t), \\ F(t_0) = Id, \end{cases}$$

où Id : désigne la matrice identité.

Pour tout contrôle u le système $\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t)$, $x(0) = x_0$ admet une unique solution $x(\cdot) : I \rightarrow \mathbb{R}^n$ absolument continue donnée par

$$x(t) = F(t)x_0 + \int_{t_0}^{t^*} F(t)F(s)^{-1}(B(s)u(s) + r(s))ds,$$

pour tout $t \in I$.

Si $r = 0$ et $x_0 = 0$, la solution du système s'écrit

$$x(t) = F(t) \int_{t_0}^{t^*} F(s)^{-1}B(s)u(s)ds.$$

Elle est linéaire en u .

Le théorème suivant donne une condition générale pour la contrôlabilité des systèmes linéaires :

Théorème 5 *Le système $\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t)$ est contrôlable en temps t^* si et seulement si la matrice*

$$C(t^*) = \int_{t_0}^{t^*} F(t)^{-1}B(t)B(t)'F(t)^{-1}dt,$$

dite matrice de contrôlabilité, est inversible.

Cette condition ne dépend pas de x_0 , c'est-à-dire que si un système linéaire est contrôlable en temps t^* depuis x_0 , alors il est contrôlable en temps t^* depuis tout point.

2.3.2 Contrôlabilité des systèmes linéaires autonomes

Le système $\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t)$ est dit autonome lorsque les matrices A et B ne dépendent pas de t . Dans ce cas, la matrice $F(t) = e^{At}$, et la solution du système associée au contrôle u s'écrit, pour tout $t \in I$:

$$x(t) = e^{At} \left(x_0 + \int_{t_0}^{t^*} e^{-As} (B(s)u(s) + r(s)) ds \right).$$

Le théorème suivant donne une condition nécessaire et suffisante de contrôlabilité dans le cas sans contrainte sur le contrôle.

Théorème 6 *On suppose que $U = \mathbb{R}^n$. Le système $\dot{x}(t) = A(t)x(t) + B(t)u(t) + r(t)$ est contrôlable en temps t^* si et seulement si la matrice*

$$C = [B, AB, A^2B, \dots, A^{n-1}B]$$

est de rang n . La matrice C est appelée matrice de Kalman, et la condition rang $C = n$ est appelée condition de Kalman.

Dans le cas où le contrôle u est contraint, c'est-à-dire s'il appartient à un sous ensemble $U \subset \mathbb{R}^m$, les propriétés de contrôlabilité globale sont reliées aux propriétés de stabilité de la matrice A . En clair, si $r = 0$ et $0 \in U$, si de plus la condition de Kalman est vérifiée et si la matrice A est stable (toutes les valeurs propres de A sont de parties réelles strictement négatives), alors tout point de \mathbb{R}^n peut être conduit à l'origine en temps fini.

Dans le cas mono-entrée $m = 1$ (u est un contrôle scalaire), on a le théorème suivant :

Théorème 7 *On considère le système $\dot{x}(t) = Ax(t) + bu(t)$, $b \in \mathbb{R}^n$, $u(t) \in U$ où U est un intervalle de \mathbb{R} avec $0 \in U$. Alors tout point de \mathbb{R}^n peut être conduit à l'origine en temps fini si et seulement si la matrice $C = [b, Ab, A^2b, \dots, A^{n-1}b]$ est de rang n et la partie réelle de chaque valeur propre de A est inférieure ou égale à 0.*

2.3.3 Contrôlabilité des systèmes non-linéaire

La contrôlabilité est un concept clé pour la compréhension des propriétés structurelles et qualitatives, comme la stabilisation. L'extension de la contrôlabilité au cas non-linéaire de dimension finie et de dimension infinie a suscité depuis près de cinquante ans une littérature considérable, qui n'a en

rien épuisé ce sujet riche et varié. Les auteurs, dans leur quasi-totalité, ont considéré des généralisations naturelles de $\dot{x} = Ax + Bu$. Le résultat suivant donne une condition sur la contrôlabilité locale des systèmes non-linéaires.

Proposition 2.1 *Considérons le système $\dot{x}(t) = f(t, x(t), u(t))$ $x(0) = x_0$ avec $f(x_0, u^0) = 0$. On note $A = \frac{\partial f}{\partial x}(x_0, u^0)$ et $B = \frac{\partial f}{\partial u}(x_0, u^0)$. Si*

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

alors le système est localement contrôlable en x_0 .

En général, le problème de contrôlabilité est difficile. Cependant, il existe des techniques qui permettent de déduire la contrôlabilité locale dans le cas des systèmes linéarisés.

2.4 Problème de contrôle optimal

Une fois le problème de contrôlabilité est résolu, alors parmi toutes les trajectoires qui transfèrent l'objet de la position initiale $x(t_0)$ à la position finale $x(t^*) = x_1$, on cherche alors celle qui le fait en temps final, ou celle qui le fait en minimisant un certain critère, c'est le problème de contrôle optimal. La résolution du problème de contrôle optimal se fait par différentes méthodes. Parmi ces méthodes, on peut citer :

1. Principe du maximum de Pontriaguine,
2. méthode directe de résolution d'un problème terminal d'un système dynamique linéaire.

Dans les sous-paragraphes suivants, nous allons détaillé ces méthodes.

2.4.1 Principe du maximum de Pontriaguine

Ce principe formulé par le mathématicien russe donne une condition nécessaire d'optimalité [73]. Dans notre travail, on ne considérera seulement la condition que pour un problème terminal :

$$J(x, u) = \int_{t_0}^{t^*} f_0(t, x, u) dt \rightarrow \min, \quad (2.1)$$

$$\dot{x} = f(t, x, u), \quad x(t_0) = x_0 \in M_0, \quad (2.2)$$

$$u(t) \in U, \quad t \in [t_0, t^*], \quad (2.3)$$

$$x(t^*) \in M_1, \quad (2.4)$$

S'il n'existe pas de contrôle $u \in U$ satisfaisant le système $\dot{x} = f(t, x, u)$, $x(0) = x_0$ et $x(t^*) \in M_1$, on dit que le système n'est pas contrôlable de l'état initial aux états terminaux de M_1 . Dans ce cas, le problème

n'admet pas de solution. Si le système est contrôlable, il existe en général beaucoup de contrôles possibles et pour chacun de ces contrôles correspond une valeur pour J . Le problème est de déterminer un contrôle optimal $u^* \in U$ associé à des trajectoires optimales x^* qui optimise la valeur de J .

Dans la suite, on utilise l'Hamiltonien défini comme suit :

$$H(t, x(t), \psi(t), v(t)) = \sum_{i=0}^n \psi_i \cdot f_i(t, x(t), u(t)) = \psi'(t) \cdot (Ax(t) + bu(t)),$$

Finalement, On dit qu'une commande admissible $u(t)$, $t \in [t_0, t^*]$ vérifie le principe du maximum pour le problème (2.1)-(2.4) si la condition :

$$H(t, x(t), \psi(t), u(t)) = \max_{v \in U} H(t, x(t), \psi(t), v(t)),$$

est vérifiée. Classiquement, la trajectoire $x(t)$ du système vérifie :

$$\dot{x} = \frac{\partial H(x, \psi, u, t)}{\partial \psi}, \quad x(t_0) = x_0, \quad (2.5)$$

et si de plus, on utilise aussi la co-trajectoire $\psi(t)$ du système vérifie :

$$\dot{\psi} = -\frac{\partial H(x, \psi, u, t)}{\partial x}, \quad (2.6)$$

$\psi(t)$: co-trajectoire solution du système (2.6) est appelé système adjoint.

Remarque 2.1 *La convention $\psi_0 \leq 0$ conduit au principe du maximum, tandis que $\psi_0 \geq 0$ conduit au principe du minimum.*

Dans le cas où il n'y a pas de contrainte sur le contrôle ($U = \mathbb{R}^m$), un contrôle optimal u vérifie les conditions suivantes :

- $\psi_0 = -1$.
- u est une fonction telle que $H(x(t), \psi(t), v(t), t)$ atteint son maximum en u , $\forall v \in \mathbb{R}^m$. La condition du maximum devient $\frac{\partial H}{\partial v} = 0$.
- Les co-équations d'état adjoint ont une solution ψ , et les équations d'état ont une solution x qui prend les valeurs x_0 en $t_0 = 0$ et x^* au temps t^* . Le système vérifie les conditions de transversalité : à l'instant initial, $\psi(0)$ est perpendiculaire à l'espace tangent de M_0 en $x(0)$ et à l'instant final, $x(t^*)$ est perpendiculaire à l'espace tangent de M_1 en $x(t^*)$.
- L'Hamiltonien est constant le long de la trajectoire optimale, et cette constante vaut 0 si le temps terminal t^* est libre.

Si une solution existe, le principe du maximum de Pontriaguine produit des conditions nécessaires. On va donc chercher des solutions qui satisfont ces conditions nécessaires du PMP et l'on prendra celle qui minimise J . Il n'y a pas de garantie en toute généralité sur l'unicité de la solution optimale. Si

l'on ne trouve pas de solutions satisfaisantes toutes les conditions du principe du maximum de Pontriaguine, alors il n'existe pas de solution au problème de contrôle optimal.

Théorème 8 *Toute commande optimale d'un problème terminal de contrôle optimal vérifie le principe du maximum.*

2.4.2 Méthode directe de résolution d'un problème terminal d'un système dynamique linéaire :

Considérons le problème terminal de commande optimale suivant :

$$J(u) = c'x(t^*) \rightarrow \max_u, \quad (2.7)$$

$$\dot{x} = \frac{dx}{dt} = Ax + bu, \quad x(t_0) = x_0, \quad (2.8)$$

$$Hx(t^*) = g, \quad (2.9)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T = [t_0, t^*], \quad (2.10)$$

où $J(u)$ est le critère de qualité, $A(I, J)$ est une $n \times n$ matrice qui caractérise le système, $x(t) \in \mathbb{R}^n$ est l'état du système, $b(J)$ est un n -vecteur de commande donné, $H(I, J)$ est une $m \times n$ matrice de rang $m \leq n$, g est un m -vecteur, $u(t)$ est une commande constante par morceaux, bornée par f_* , $f^* \in \mathbb{R}^n$, $c = c(J)$ est le n -vecteur des coûts, $I = \{1, \dots, n\}$, $J = \{1, \dots, m\}$ sont des ensembles d'indices.

Notre problème consiste à trouver la commande admissible u^0 qui avec la trajectoire optimale x^0 , maximise le critère de qualité :

$$J(u^0) = c'x^0(t^*) = \max c'x(t^*) = \max J(u).$$

Par la formule de Cauchy, la solution du système dynamique

$$\dot{x} = Ax + bu, \quad x(t_0) = x_0,$$

est égale à

$$x(t) = F(t)[x_0 + \int_{t_0}^{t^*} F^{-1}(\tau)bu(\tau)d\tau], \quad (2.11)$$

où $F(t)$, $t \in T$, est la solution (résolvante) du système

$$\begin{cases} \dot{F} = AF, \\ F(t_0) = Id, \quad t \in T. \end{cases}$$

En utilisant la solution (2.11), le problème (2.7)-(2.10) devient un problème de la seule variable $u(t)$, $t \in T$:

$$J(u) = c'x(t^*) = c'F(t^*)x_0 + \int_{t_0}^{t^*} c(t)u(t)dt \rightarrow \max_u, \quad (2.12)$$

$$\int_{t_0}^{t^*} \phi(t)u(t)dt = g - HF(t^*)x_0, \quad (2.13)$$

$$f_* \leq u(t) \leq f^*, \quad (2.14)$$

où

$$\begin{cases} c(t) = c'F(t^*)F^{-1}(t)b, \\ \phi(t) = HF(t^*)F^{-1}(t)b, \quad t \in T. \end{cases} \quad (2.15)$$

Définition 2.1 Toute commande $u(t)$ vérifiant les contraintes

$$Hx(t^*) = g, \quad f_* \leq u(t) \leq f^*, \quad t \in [t_0, t^*],$$

est dite commande admissible du problème (2.7)-(2.10).

Une commande admissible $u^0 = u^0(t)$ est alors dite optimale si et seulement si :

$$J(u^0) = \max_{f_* \leq u \leq f^*} J(u), \quad (2.16)$$

et elle est ε -optimale si :

$$J(u^0) - J(u^\varepsilon) \leq \varepsilon \quad (\varepsilon > 0). \quad (2.17)$$

Définition 2.2 La valeur $u(\cdot)$ au point $t \in \text{int } T$, sera la valeur $u(t) = (u(t+0) + u(t-0))/2$, où $u(t+0)$ et $u(t-0)$ sont les limites à droite et à gauche de u au point t .

2.5 Support-Contrôle

Définition 2.3 L'ensemble $T_B = \{T_i = [\tau_i, \tau^i], T_i \subset T, i = \overline{1, m}/[\tau_i, \tau^i] \cap [\tau_j, \tau^j] = \emptyset, \forall i \neq j\}$ est appelé support généralisé du problème (2.7)-(2.10) si le système :

$$\begin{cases} \dot{x} = Ax + bu, \quad x(t_0) = x_0 \\ Hx(t^*) = 0, \quad t \in T \end{cases} \quad (2.18)$$

n'admet pour $u(t) \equiv 0, t \in T_H = T/T_B$ que la solution triviale $u(t) \equiv 0, t \in T_B$, mais pour tout intervalle $T^* = [\tau_*, \tau^*], T^* \subset T_H, \tau_* \neq \tau^*$ et

$$u(t) = \begin{cases} 0, & t \in T_H/T^* \\ u^*, & t \in T^*. \end{cases}$$

le système (2.18) admet une solution non triviale, i.e.,

$$u(t) \neq 0, t \in T_B \cup T^*$$

dans la classe des commandes constantes sur $T_i, i = \overline{1, m}$.

Lemme 2.1 L'ensemble T_B est un support généralisé si et seulement si la matrice du support généralisée

$$\phi(T_B) = \int_{\tau_i}^{\tau_i+h} \phi(t) dt, \quad i = \overline{1, m}. \quad (2.19)$$

est non dégénérée

Prenons une autre notion de support du problème. Choisissons les points isolés $\tau_i, i = \overline{1, m}$ qu'on appelle des moments de support et l'ensemble $\tau_B = \{\tau_i, i = \overline{1, m}\}$ est dit support si la matrice correspondante ϕ_B est non dégénérée

$$\phi_B = \phi(T_B) = \int_{\tau_i}^{\tau_i+h} \phi(t) dt, \quad i = \overline{1, m}. \quad (2.20)$$

Remarque 2.2 A chaque moment t_i , on fait correspondre un intervalle T_i tel que : $t_i = \tau_i$ ou τ^i .

L'existence du support est très lié à la notion de commandabilité du système (2.7)-(2.10).

Définition 2.4 Le couple $\{u, \tau_B\}$ formé de la commande admissible u et du support τ_B est appelé support-contrôle.

Il est dit dégénéré si la valeur de la fonctionnelle n'augmente pas après l'itération, et non dégénéré s'il existe de tels nombres $\lambda_0 > 0, \mu_0 > 0, u_i^\lambda, i = \overline{1, m}$ tel que, $\forall \lambda, 0 < \lambda < \lambda_0$, les relations suivantes sont vérifiées.

$$\begin{cases} \sum_{i=1}^m \int_{t_i-\lambda}^{t_i+\lambda} \phi(t) u(t) dt = \sum_{i=1}^m u_i^\lambda \int_{t_i-\lambda}^{t_i+\lambda} \phi(t) dt, \\ f_* + \mu_0 \leq u_i^\lambda \leq f^* - \mu_0, \quad i = \overline{1, m} \end{cases} \quad (2.21)$$

Par la suite, on supposera que la commande $u(t)$ prend des valeurs constantes dans les intervalles T_i , $i = \overline{1, m}$ i.e,

$$u(t) = u_i, \quad t \in T_i = [\tau_i, \tau^i], \quad d_* \leq u_i \leq d^*, \quad i = \overline{1, m}.$$

Comme u est admissible, alors

$$\int_{t_0}^{t^*} \phi(t)u(t)dt = g - HF(t^*)x_0,$$

et en utilisant le support, on obtient

$$\int_{T_B} \phi(t)u(t)dt = g - HF(t^*)x_0 - \int_{T_H} \phi(t)u(t)dt.$$

Comme $u(t)$ est constante par morceaux et de la définition du support, on obtient

$$u(T_B) = \phi^{-1}(T_B)\{g - HF(t^*)x_0 - \int_{T_H} \phi(t)u(t)dt\}, \quad (2.22)$$

où $T_B = \{T_i, \quad i = \overline{1, m}\}$.

2.6 Formule d'accroissement de la fonctionnelle

Soit $\{u, \tau_B\}$ un support-contrôle de départ, $x(t)$ sa trajectoire correspondante. En utilisant le support τ_B , on construit le vecteur potentiel :

$$\nu' = c'_B \phi_B^{-1}, \quad (2.23)$$

et la co-commande (vecteur des estimations) :

$$\Delta(t) = -\psi'(t)b, \quad t \in T, \quad (2.24)$$

où $\psi(t)$ est la solution du système conjuguée.

$$\dot{\psi}(t) = -A'\psi, \quad \psi(t^*) = c - H'\nu. \quad (2.25)$$

$\Delta(t)$ peut être écrite d'une autre manière :

$$\begin{aligned} \Delta(t) &= -\psi'(t)b = -((c - H'\nu)F(t^*)F^{-1}(t))'b \\ &= \nu'HF(t^*)F^{-1}(t)b - c'F(t^*)F^{-1}(t)b \\ &= \nu'\phi(t) - c(t) \end{aligned}$$

où

$$\phi(t) = HF(t^*)F^{-1}(t)b,$$

et

$$c(t) = c'F(t^*)F^{-1}(t)b.$$

d'où

$$\Delta(t) = \nu'\phi(t) - c(t), \quad t \in T \quad (2.26)$$

Soit $\bar{u}(t) = u(t) + \Delta u(t)$ une autre commande admissible et $\bar{x}(t) = x(t) + \Delta x(t)$, $t \in T$, la trajectoire correspondante associée.

Déterminons l'accroissement de la fonctionnelle

$$\begin{aligned} \Delta J(u) &= J(\bar{u}) - J(u) = \int_{t_0}^{t^*} c(t)u(t)dt \\ &= \int_{t_0}^{t^*} (-\Delta(t) + \nu'\phi(t))\Delta u(t)dt \\ &= -\int_{t_0}^{t^*} \Delta(t)\Delta u(t)dt + \int_{t_0}^{t^*} \nu\phi(t)\Delta u(t)dt \end{aligned}$$

De l'admissibilité de u et \bar{u} , on a :

$$\int_{t_0}^{t^*} \nu'\phi(t)\Delta u(t)dt = 0.$$

Donc,

$$\Delta J(u) = -\int_{t_0}^{t^*} \Delta(t)\Delta u(t)dt \quad (2.27)$$

2.7 Problème dual

Définissons le problème dual du problème (2.7)-(2.10) :

$$\left\{ \begin{array}{l} L(\varsigma, v, w) = \bar{b}'\varsigma - \int_{t_0}^{t^*} v(t)f_*dt + \int_{t_0}^{t^*} w(t)f^*dt \rightarrow \min, \\ \varsigma\phi(t) - v(t) + w(t) = c(t), \\ v(t) \geq 0, w(t) \geq 0, t \in T, \end{array} \right.$$

où

$$\bar{b} = g - HF(t^*)x_0.$$

Ici les fonctions $v(t)$, $w(t)$, $t \in T$ pour $\nu(I) = \zeta(I)$ sont définies de la manière suivante :

$$\begin{cases} v(t) = \Delta(t), w(t) = 0, & \text{si } \Delta(t) \geq 0 \\ v(t) = 0, w(t) = -\Delta(t), & \text{si } \Delta(t) \leq 0. \end{cases} \quad (2.28)$$

et forment un plan dual du problème (solution admissible du dual). Désignons par $(\zeta^0, v^0(t), w^0(t), t \in T)$ la solution optimale.

2.8 Calcul de la valeur de suboptimalité

La nouvelle commande $\bar{u}(t)$ est admissible donc :

$$f_* - u(t) \leq \Delta u(t) \leq f^* - u(t). \quad (2.29)$$

Le maximum de la fonctionnelle (2.27) sous les contraintes (2.29) est atteint pour :

$$\begin{cases} \Delta u(t) = f_* - u(t), & \text{si } \Delta(t) > 0 \\ \Delta u(t) = f^* - u(t), & \text{si } \Delta(t) < 0 \\ f_* - u(t) \leq \Delta u(t) \leq f^* - u(t), & \text{si } \Delta(t) = 0, t \in T. \end{cases} \quad (2.30)$$

et est égal à :

$$\beta = \beta(u, \tau_B) = \int_{T^+} \Delta(t)(u(t) - f_*)dt + \int_{T^-} \Delta(t)(u(t) - f^*)dt, \quad (2.31)$$

appelé valeur de suboptimalité du support-contrôle $\{u, \tau_B\}$, où

$$T^+ = \{t \in T : \Delta(t) > 0\}, \quad T^- = \{t \in T : \Delta(t) < 0\}.$$

De ce qui précède, on a toujours :

$$\Delta J(u) = J(\bar{u}) - J(u) \leq \beta(u, \tau_B),$$

pour tout \bar{u} , et pour tout $\bar{u} = u^0$, on obtient :

$$J(u^0) - J(u) \leq \beta(u, \tau_B).$$

De cette dernière inégalité, on déduit les critères suivants :

2.9 Critère d'optimalité

Théorème 9 [34] Pour l'optimalité du support-contrôle $\{u, \tau_B\}$, les relations :

$$\begin{cases} \Delta(t) > 0, & \text{si } u(t) = f_* \\ \Delta(t) < 0, & \text{si } u(t) = f^* \\ \Delta(t) = 0, & \text{si } f_* \leq u(t) \leq f^*, t \in T, \end{cases} \quad (2.32)$$

sont suffisantes et dans le cas de la non dégénérescence, elles sont nécessaires.

Preuve. voir l'annexe 2 **Condition suffisante :**

Si les relations (2.32) sont vérifiées alors de (2.31), on a $\beta(u, \tau_B) = 0$. Comme $J(\bar{u}) - J(u) \leq \beta(u, \tau_B) = 0$, pour tout \bar{u} , alors $\{u, \tau_B\}$ est un support-contrôle optimale.

Condition nécessaire :

Soit $\{u, \tau_B\}$ un support contrôle optimal non dégénéré et supposons que les relations (2.31) ne sont pas vérifiées, c'est-à-dire il existe $t_0 \in T$ tel que $\Delta(t_0) > 0$ et $u(t_0) > f_0$. De la continuité de $\Delta(t)$, $t \in T$, il existe un tel nombre petit η_0 , tel que pour tout $0 < \eta < \eta_0$, on a :

1. t_0 n'est pas un moment du support.
2. $t_0 \in T_\kappa^\eta$, $T_\kappa^\eta = [\tau_\kappa - \eta, \tau_\kappa + \eta]$, $\kappa = \overline{1, m}$.
3. $\Delta(t) > 0$, pour $u(t) > f_0$, $t \in T_*^\eta$, $T_*^\eta = [t_0 - \eta, t_0 + \eta]$.
4. $T_*^\eta \cap T_\kappa^\eta = \emptyset$, $\kappa = \overline{1, m}$.
5. le moment t_0 peut être considéré comme un point de continuité de $u(t)$, $t \in T_*^\eta$.

Choisissons η_0 , suffisamment petit pour tout η , $0 < \eta < \eta_0$.

Construisons les variations de la commande

$$\Delta u(t) = \Delta u(t, \eta), \quad t \in T,$$

en posant :

$\Delta u(t) = -\gamma$, $t \in T_*^\eta$ où $0 < \gamma = \gamma(\eta) = \kappa\eta \leq 1$, $\kappa > 0$, et Δu ne dépend pas de η ,

$$\Delta u(t) = 0, \quad t \in T / (\cup_{\kappa=1}^m T_\kappa^\eta \cup T_*^\eta).$$

Au voisinage du moment du support τ_j , qui permet de changer la valeur $u(t)$ à la nouvelle valeur $\bar{u}(t)$.

$$\bar{u}(t) = \{[u(\tau_\kappa - 0) + u(\tau_\kappa)]/2 = u(\tau_\kappa), \quad t \in T_\kappa^\eta, \quad \kappa = \overline{1, m}\}.$$

D'après de la non dégénérescence de $u(t)$, on a : $f_* < \bar{u}(\tau_\kappa) < f^*$, $\kappa = \overline{1, m}$.
 Pour toute fonction continue :

$$\begin{aligned} \int_{t_0}^{t^*} \phi(t)u(t)dt &= \int_{\tau_\kappa - \eta}^{\tau_\kappa} \phi(t)dt u(\tau_\kappa - 0) + \int_{\tau_\kappa}^{\tau_\kappa + \eta} \phi(t)dt u(\tau_\kappa + 0) \\ &= \eta\phi(\tau_\kappa)u(\tau_\kappa - 0) + \mathcal{O}_1(\eta) + \eta\phi(\tau_\kappa)u(\tau_\kappa + 0) + \mathcal{O}_2(\eta) \\ &= 2\eta\phi(\tau_\kappa)u(\tau_\kappa) + \mathcal{O}(\eta). \end{aligned}$$

En changeant $u(t)$, $t \in T$, en $\bar{u}(t)$, $t \in T$, la contrainte (2.12) prend la forme :

$$\begin{aligned} H F(t^*)x_0 + \int_{t_0}^{t^*} \phi(t)\bar{u}(t)dt - HF(t^*)x_0 - \int_{t_0}^{t^*} \phi(t)u(t)dt \\ = \int_{t_0}^{t^*} \phi(t)(\bar{u}(t) - u(t))dt = \sum_{\kappa=1}^m \int_{T_\kappa^\eta} (\bar{u}(t) - u(t))dt \\ = 2\eta \sum_{\kappa=1}^m \phi(\tau_\kappa)u(\tau_\kappa) + \mathcal{O}_1(\eta) - 2\eta \sum_{\kappa=1}^m \phi(\tau_\kappa) + \mathcal{O}_2(\eta) = \mathcal{O}(\eta). \end{aligned}$$

D'après le théorème des fonctions implicites (voir annexe 2), la commande $\bar{u}(t)$, $t \in T$, peut être transformée de telle sorte que $\bar{u}(t)$, $t \in T$, vérifie la contrainte (2.12).

$$\bar{u}(t) = \begin{cases} u(t), & t \in T_\kappa^\eta \\ u(\tau_\kappa) + \mathcal{O}(\eta), & t \in T_\kappa^\eta, \kappa = \overline{1, m} \end{cases}$$

L'accroissement de la fonctionnelle devient :

$$\begin{aligned} J(\bar{u}) - J(u) &= \int_{t_0}^{t^*} c(t)(\bar{u}(t) - u(t))dt \\ &= \sum_{\kappa=1}^m \int_{T_\kappa^\eta} c(t)(\bar{u}(t) - u(t))dt = 2\eta \sum_{\kappa=1}^m c(\tau_\kappa)u(\tau_\kappa) + \mathcal{O}_1(\eta) \\ &\quad - 2\eta \sum_{\kappa=1}^m c(\tau_\kappa)u(\tau_\kappa) + \mathcal{O}_2(\eta) = \mathcal{O}(\eta). \end{aligned}$$

La variation de la commande sur les intervalles T_κ^η , $\kappa = \overline{1, m}$, sera construite de la manière suivante :

$$\Delta u(t) = \Delta u(\tau_\kappa), \quad t \in T_\kappa^\eta, \quad \kappa = \overline{1, m}.$$

De l'inégalité, on a :

$$HF(t^*)x_0 + \sum_{\kappa=1}^m \int_{T_\kappa^\eta} \phi(t) \Delta u(t) dt = \gamma \int_{T_\kappa^\eta} \phi(t) dt.$$

$$HF(t^*)x_0 + \sum_{\kappa=1}^m 2\eta^\kappa \phi(\tau_\kappa) \Delta u(\tau_\kappa) + \mathcal{O}(\eta) = \gamma \int_{T_\kappa^\eta} \phi(t) dt.$$

La matrice de coefficients de ce système :

$$P_B = (\phi(t), t \in T_B).$$

est confondue avec la matrice du support ϕ_B .

Introduisons le vecteur :

$$\Delta u(\tau_\kappa), \kappa = \overline{1, m}.$$

$$\Delta u_B = 2mQ\phi(t_0) + \mathcal{O}(\eta).$$

Pour η suffisamment petit, la norme Δu_B est aussi suffisamment petite. Par conséquent, les inégalités suivantes sont vérifiées :

$$f_* \leq u^*(\tau_\kappa) + \Delta u(\tau_\kappa) \leq f^*, \kappa = \overline{1, m}.$$

Par conséquent, la commande généralisée $\bar{u}(t)$ de composantes : $\tilde{u}(t) = u(t)$, $t \in T / (\cup_{\kappa=1}^m T_\kappa^\eta \cup T_*^\eta)$, $\tilde{u}(t) = \tilde{u}(t) + \Delta u(\tau_\kappa)$, $t \in T_\kappa^\eta$, $\kappa = \overline{1, m}$, $\tilde{u}(t) = u(t) - \gamma$, $t \in T_*^\eta$ est admissible.

Calculons l'accroissement de la fonctionnelle pour les deux commandes u et \bar{u} :

$$\begin{aligned} \Delta J(u) = J(\tilde{u}) - J(u) &= - \sum_{\kappa=1}^m \int_{T_\kappa^\eta} \Delta(t)(u(\tau_\kappa) + \Delta u(\tau_\kappa) - u(t)) dt \\ &+ \gamma \int_{T_*^\eta} \Delta(t) dt = \kappa \eta \int_{T_*^\eta} \Delta(t) dt + \mathcal{O}(\eta) > 0. \end{aligned}$$

Pour η suffisamment petit, $0 < \eta < \eta_0$, l'inégalité obtenue contredit l'optimalité de u ; ce qui démontre la condition nécessaire.

Remarque 2.3 *Le critère d'optimalité peut être écrit sous la forme traditionnelle du principe du maximum de Pontriaguine. Pour cela en utilisant la fonction Hamiltonienne, on obtient :*

$$H(x, \psi, u) = \psi'(Ax + bu)$$

La condition du maximum :

$$H(x(t), \psi(t), u(t)) = \max_{f_* \leq u \leq f^*} H(x(t), \psi(t), u(t)), t \in T. \quad (2.33)$$

est suffisante, et dans le cas de la non dégénérescence, elle est nécessaire pour l'optimalité du support-contrôle $\{u, \tau_B\}$.

2.10 Critère d' ε -Optimalité

Théorème 10 *Pour tout $\varepsilon \geq 0$, la commande admissible $u(t)$, $t \in T$ est ε -optimale si et seulement s'il existe un tel support τ_B sur lequel le long de $u(t)$ et des solutions $x(t)$, $\psi(t)$, $t \in T$ des systèmes direct et conjugué, la condition ε -maximum est vérifiée :*

$$H(x(t), \psi(t), u(t)) = \max_{f_* \leq u \leq f^*} H(x(t), \psi(t), u(t)) - \varepsilon(t), \quad t \in T. \quad (2.34)$$

$$\text{avec } \int_{t_0}^{t^*} \varepsilon(t) dt \leq \varepsilon.$$

Preuve.

Condition suffisante :

Supposons que la condition (2.34) soit vérifiée ; de la valeur de suboptimalité (2.31) du support-contrôle $\{u, \tau_B\}$, et de la formule (2.24) on déduit :

$$\begin{aligned} \beta = \beta(u, \tau_B) &= \int_{T^+} \psi'(t) b(f_* - u(t)) dt + \int_{T^-} \psi'(t) b(f^* - u(t)) dt \\ &= \int_{T^+} (\psi'(t) A x(t) + \psi'(t) b f_* - \psi'(t) A x(t) - \psi'(t) b u(t)) dt \\ &\quad + \int_{T^-} (\psi'(t) A x(t) + \psi'(t) b f^* - \psi'(t) A x(t) - \psi'(t) b u(t)) dt \\ &= \int_{t_0}^{t^*} [\max_{f_* \leq u \leq f^*} H(x(t), \psi(t), u(t)) - H(x(t), \psi(t), u(t))] dt \\ &= \int_{t_0}^{t^*} \varepsilon(t) dt \leq \varepsilon. \end{aligned}$$

De l'accroissement de la fonctionnelle, on déduit :

$$J(u^0) - J(u) \leq \beta < \varepsilon,$$

ce qui implique que la commande $u(t)$, $t \in T$ est ε -optimale.

Condition nécessaire :

Soit $u(t)$, $t \in T$ une commande ε -optimale. De la formule (2.31) pour un certain support τ_B , on calcule la valeur de suboptimalité $\beta(u, \tau_B)$ écrite sous la forme :

$$\beta = \beta(u, \tau_B) = \int_{t_0}^{t^*} \Delta(t) u(t) dt + \int_{T^+} \Delta(t) f_* dt - \int_{T^-} \Delta(t) f^* dt. \quad (2.35)$$

De la valeur de suboptimalité (2.34) et en vertu des relation (2.27) et (2.35), on obtient :

$$\beta = \beta(u, \tau_B) = \int_{t_0}^{t^*} \nu' \phi(t) u(t) dt - \int_{t_0}^{t^*} c(t) u(t) dt - \int_{t_0}^{t^*} v(t) f_* dt + \int_{t_0}^{t^*} w(t) f^* dt.$$

Or, de l'admissibilité de u , on déduit :

$$\int_{t_0}^{t^*} \nu' \phi(t) u(t) dt = \bar{b}' \nu,$$

et de la relation de dualité, on obtient :

$$J(u^0) = L(\nu^0, v^0, w^0).$$

ce qui nous donne :

$$\int_{t_0}^{t^*} c(t) u^0(t) dt = \bar{b}' \nu^0 + \int_{t_0}^{t^*} v^0(t) f_* dt - \int_{t_0}^{t^*} w^0(t) f^* dt = 0.$$

De cette dernière égalité, la valeur de suboptimalité prend la forme :

$$\begin{aligned} \beta &= \int_{t_0}^{t^*} c(t) u^0(t) dt - \int_{t_0}^{t^*} c(t) u(t) dt \\ &+ \bar{b}' \nu - \int_{t_0}^{t^*} v(t) f_* dt + \int_{t_0}^{t^*} w(t) f^* dt + \bar{b}' \nu^0 \\ &+ \int_{t_0}^{t^*} v^0(t) f_* dt - \int_{t_0}^{t^*} w^0(t) f^* dt. \end{aligned}$$

$$\beta = \beta(u, \tau_B) = \beta_u + \beta_B,$$

où

$$\beta_u = \int_{t_0}^{t^*} c(t) u^0(t) dt - \int_{t_0}^{t^*} c(t) u(t) dt = J(u^0(t)) - J(u(t)),$$

est la mesure de la non optimalité de la commande $u(t)$, $t \in T$, et

$$\begin{aligned} \beta_B &= \bar{b}' \nu - \int_{t_0}^{t^*} v(t) f_* dt + \int_{t_0}^{t^*} w(t) f^* dt + \bar{b}' \nu^0 + \int_{t_0}^{t^*} v^0(t) f_* dt - \int_{t_0}^{t^*} w^0(t) f^* dt \\ &= L(\varsigma, v, w) - L(\varsigma^0, v^0, w^0), \end{aligned}$$

est la mesure de la non optimalité du support τ_B .

Si à la commande $u(t)$, $t \in T$, on associe un support τ_B^0 optimal, c'est-à-dire $\beta(\tau_B^0) = \beta_B = 0$, alors :

$$\beta(u, \tau_B^0) = \beta_u \leq \varepsilon, \quad (2.36)$$

car u est ε -optimale. Posons :

$$\varepsilon(t) = \begin{cases} \Delta(t)(u(t) - f_*), & \text{si } t \in T^+ \\ \Delta(t)(u(t) - f^*), & \text{si } t \in T^- \\ 0, & \text{si } \Delta(t) = 0, t \in T. \end{cases}$$

En utilisant la co-commande (2.24), on obtient :

$$\varepsilon(t) = -\psi'(t)b(u(t) - f_*),$$

$$\varepsilon(t) = \psi'(t)Ax(t) + \psi'(t)bf_* - \psi'(t)Ax(t) - \psi'(t)bu(t),$$

$$\varepsilon(t) = \psi'(t)(Ax(t) + bf_*) - \psi'(t)(Ax(t) + bu(t)) \quad \psi'(t)b < 0,$$

$$\varepsilon(t) = \psi'(t)(Ax(t) + bf^*) - \psi'(t)(Ax(t) + bu(t)) \quad \psi'(t)b > 0,$$

$$\varepsilon(t) = 0 \quad \psi'(t)b = 0.$$

En utilisant la fonction Hamiltonienne, $\varepsilon(t)$ sera égale à :

$$\varepsilon(t) = \max_{f_* \leq u \leq f^*} H(x(t), \psi(t), u(t)) - H(x(t), \psi(t), u(t)), \quad t \in T \quad (2.37)$$

Des conditions (2.36) et (2.37) découle alors la condition ε - maximum.

Conclusion

Si $\Delta J(u) = J(u^0) - J(u) = \beta = \beta(u, \tau_B) \leq \varepsilon$, alors u est ε - optimale.

Pour $\varepsilon = 0$, le principe ε - maximum devient le principe traditionnel du maximum.

2.11 Détermination d'un support-contrôle de départ

Pour déterminer un support contrôle de départ admissible, on prend une commande $u(t)$, $t \in T$ vérifiant la contrainte

$$f_* \leq u(t) \leq f^*, \quad t \in T.$$

Par la suite, on détermine sa trajectoire correspondante $x(t)$, $t \in T$, solution de

$$\dot{x} = Ax + bu, \quad x(0) = x_0.$$

Si la condition $Hx(t^*) = g$ est vérifiée, alors $u(t)$, $t \in T$ est une commande admissible. Sinon, on ajoute une variable artificielle pour avoir un support de départ en utilisant la méthode de deux phases (phase 1) ou M -méthode [17] permet de déterminer la condition initiale du problème. Quant au support τ_B , on le choisit de manière à avoir $|\phi_B| \neq 0$, et si on ne peut pas déterminer un tel support, on démarre l'algorithme à partir de $\tau_B = \emptyset$ avec $\det \phi_B(t) \neq 0$.

2.12 Méthode de résolution

Soit $\{u, \tau_B\}$ un support contrôle de départ, pour lequel la condition ε -maximum n'est pas vérifiée.

Une itération de l'algorithme $\{u, \tau_B\} \rightarrow \{\bar{u}, \bar{\tau}_B\}$ se fait en trois étapes :

1. Changement de commande : $u \rightarrow \bar{u}$.
2. Changement de support : $\tau_B \rightarrow \bar{\tau}_B$.
3. Procédure finale.

2.12.1 Changement de commande

Soit $\bar{u}(t) = u(t) + \theta \Delta u(t)$, $t \in T$ une autre commande admissible, où $\Delta u(t)$, $t \in T$ est la direction, et θ le pas maximal le long de cette direction, qui seront cherchés comme solution du problème :

$$\left\{ \begin{array}{l} \Delta J(u) = -\theta \int_{t_0}^{t^*} \Delta(t) \Delta u(t) dt \rightarrow \max_{\theta, \Delta u(t)}, \\ \int_{T^*} \phi(t) \theta \Delta u(t) dt = 0, \\ f_* - u(t) \leq \theta \Delta u(t) \leq f^* - u(t), \quad t \in T. \end{array} \right. \quad (2.38)$$

Choisissons les nombres : $\alpha > 0$, $h > 0$ (paramètres de la méthode), et construisons les ensembles suivants :

$$T_0 = \{t \in T : |\Delta(t)| < \alpha\}, \quad T_1 = \{t \in T : |\Delta(t)| \geq \alpha\},$$

$T_1 = T/T_0$. Comme :

$$\Delta(t) = 0, \quad t \in T_B \text{ alors } T_B \subset T_0.$$

Subdivisons l'ensemble T_0 en sous ensembles :

$$T_0 = \bigcup_{j=1}^N [\tau_j, \tau^j[, \quad j = \overline{1, N},$$

tel que :

$$[\tau_j, \tau^j] \cap [\tau_i, \tau^i] = \emptyset, \quad \forall i \neq j \text{ et } \tau^j - \tau_j = h.$$

Posons :

$$u(t) \equiv u_j = \text{constante}, \quad t \in [\tau_j, \tau^j], \quad j = \overline{1, N}.$$

$$\Delta u(t) = \begin{cases} f^* - u(t), & \text{si } \Delta(t) \leq -\alpha, \\ f_* - u(t), & \text{si } \Delta(t) \geq \alpha, \quad t \in T_1. \end{cases} \quad (2.39)$$

Faisons un changement de variables :

$$l_j = \begin{cases} \theta \Delta u(t), & t \in [\tau_j, \tau^j], \quad j = \overline{1, N}, \\ \theta, & \text{pour } j = N + 1, \end{cases} \quad (2.40)$$

où $j = N + 1$ est un indice supplémentaire correspondant à l'ensemble T_1 .
Calculons les quantités suivantes :

$$g_j = \begin{cases} - \int_{\tau_i}^{\tau^i} \Delta(t) dt, & j = \overline{1, N} \\ - \int_{T_1} \Delta(t) \Delta u(t) dt, & j = N + 1. \end{cases} \quad (2.41)$$

$$q_j = \begin{cases} - \int_{\tau_i}^{\tau^i} \phi(t) dt, & j = \overline{1, N} \\ - \int_{T_1} \phi(t) \Delta u(t) dt, & j = N + 1. \end{cases} \quad (2.42)$$

et prenons :

$$f_{*j} = f_* - u_j, \quad d_j^* = d^* - u_j, \quad j = \overline{1, N}$$

avec :

$$0 \leq \theta \leq 1.$$

Soit alors :

$$f_{*N+1} = 0 \text{ et } f_{N+1}^* = 1.$$

De ce qui précède, le problème (2.38) devient le problème suivant appelé problème du support :

$$\left\{ \begin{array}{l} \sum_{j=1}^{N+1} g_j l_j \rightarrow \max_{l_j}, \\ \sum_{j=1}^{N+1} q_j l_j = 0, \\ f_{*j} \leq l_j \leq f_j^*, \quad j = \overline{1, N+1}. \end{array} \right. \quad (2.43)$$

Ce problème est un problème de programmation linéaire qu'on va résoudre par "la méthode adaptée" avec comme plan de départ ($l = 0, J_B$) où :

$$\phi_B = \phi(J_B) = (q_j, j \in J_B), \quad J_B = \{j \in \{1, \dots, N+1\} / T_j \subset T_B\}$$

où

$$q_j = \int_{\tau_j}^{\tau_j+h} \phi(t) dt, \quad j = \overline{1, N+1}.$$

En utilisant $\{\bar{l}, \bar{J}_B\}$ la solution optimale du problème (2.43) trouvée par la méthode adaptée, la nouvelle commande sera :

$$\bar{u}(t) = \begin{cases} u(t) + l_{N+1} \Delta u(t), & t \in T_1 \\ u(t) + l, & t \in [\tau_j, \tau^j[, \quad j = \overline{1, N}. \end{cases} \quad (2.44)$$

La commande \bar{u} ainsi construite vérifie l'inégalité :

$$J(\bar{u}) \geq J(u).$$

- Si l'indice $(N+1) \notin \bar{J}_B$, alors on pose $\tilde{J} = \bar{J}$.
- Sinon on l'exclut du support \bar{J}_B en utilisant la méthode duale. Pour cela, on cherche la nouvelle co-commande :

$$\bar{\Delta}(t) = \Delta(t) + \sigma \delta(t),$$

où $\delta(t)$ est la direction de changement de support et σ le pas le long de cette direction.

Déterminons j_* tel que :

$$\sigma_{j_*} = \min \sigma_j, \quad j \in J/J_B.$$

$$\sigma_j = \begin{cases} -\bar{\Delta}_j / \delta_j, & \text{si } \delta_j \bar{\Delta}_j \leq 0, \delta_j \neq 0, \\ 0, & \text{si } \Delta_j = 0, u(t) \neq f_*, \delta_j > 0 \text{ ou } \Delta_j = 0, u(t) \neq f^*, \delta_j < 0, j \in J_H, \\ \infty, & \text{sinon.} \end{cases}$$

la direction sera calculée de la manière suivante :

$$\delta(t) = \begin{cases} 0, & \text{sur } T_B/t_{N+1} \\ 1, & \text{si } \bar{u}(t) = f_* \\ -1, & \text{si } \bar{u}(t) = f^*. \end{cases}$$

$$\delta_j = (g_j, j \in \bar{J}_B)' \phi^{-1}(\bar{J}_B) q_j - g_j.$$

et le pas :

Le nouveau support du problème (2.43) sera alors :

$$\tilde{J}_B = (\bar{J}_B/N + 1) \cup \{j^*\}$$

En utilisant ce dernier support, on calcule le support du problème initial (2.38) en posant :

$$\tilde{\tau}_B = \{\tau_j, j \in \tilde{J}_B\}$$

et construisons la matrice

$$\phi_B = (\phi(t_j), j \in \tilde{J}_B) = \phi(\tilde{\tau}_B),$$

on vérifie facilement que :

$$\det \phi_B \neq 0,$$

par suite, on a la nouvelle valeur de suboptimalité correspondant à $\tilde{\tau}_B$, $\beta(\bar{u}, \tilde{\tau}_B)$:

- Si $\beta(\bar{u}, \tilde{\tau}_B) = 0$ alors $\bar{u}(t)$, $t \in T$ est une commande optimale.
- Si $\beta(\bar{u}, \tilde{\tau}_B) \leq \varepsilon$ alors $\bar{u}(t)$, $t \in T$ est une commande ε -optimale.
- Sinon on passe soit à une nouvelle itération en démarrant avec $\{\bar{u}, \tilde{\tau}_B\}$, $\bar{\alpha} < \alpha$, $\bar{h} < h$, ou à la procédure de changement de support.

2.12.2 Changement de support

Soit $\{\bar{u}, \tilde{\tau}_B\}$ le support contrôle trouvé précédemment et $\beta(\bar{u}, \tilde{\tau}_B) > \varepsilon$. En utilisant le support $\tilde{\tau}_B$; on construit la solution $\tilde{\psi}(t)$, $t \in T$ du système conjugué (2.26). Posons

$$\tilde{\Delta}(t) = -\tilde{\psi}'(t)b, t \in T.$$

et calculons la quasi-commande suivante :

$$\tilde{w}(t) = \begin{cases} f^*, & \text{si } \tilde{\Delta}(t) < 0, \\ f_*, & \text{si } \tilde{\Delta}(t) > 0, \\ \in [f_*, f^*], & \text{si } \tilde{\Delta}(t) = 0, t \in T \end{cases} \quad (2.45)$$

et la trajectoire correspondante $\chi = (\chi(t), t \in T)$ solution de :

$$\dot{\chi} = A\chi + b\tilde{w}, \quad \chi(0) = x_0, \quad (2.46)$$

Si $H\chi(t^*) = g$, alors $\tilde{w}(t)$, $t \in T$ est optimale pour le problème (2.7)-(2.10).
Si $H\chi(t^*) \neq g$, construisons alors le vecteur :

$$\lambda(\tilde{T}_B) = \phi^{-1}(\tilde{T}_B)(g - H\chi(t^*)). \quad (2.47)$$

- Si $\lambda(\tilde{T}_B) = 0$, alors la quasi-commande \tilde{w} construite par le support $\tilde{\tau}_B$ est optimale pour le problème (2.7)-(2.10).
- Sinon calculons $\|\lambda(\tilde{T}_B)\|$; pour $\mu > 0$, paramètre de la méthode, deux cas peuvent se présenter :
 1. Si $\|\lambda(\tilde{\tau}_B)\| > \mu$, alors on change $\tilde{\tau}_B$ en $\bar{\tau}_B$ par la méthode duale.
 2. Si $\|\lambda(\tilde{\tau}_B)\| < \mu$, on passe à la procédure finale.

Méthode duale.

Soit $t_1 \in \tau_B$,

$$|\lambda(t_1)| = \max |\lambda(t)|, \quad t \in \tilde{T}_B$$

On a :

$$\begin{aligned} \bar{\Delta}(t) &= \bar{\nu}'\varphi(t) - c(t), \quad \Delta(t) = \nu'\varphi(t) - c(t), \\ \bar{\Delta}(t) &= \Delta(t) + \sigma^0\delta(t) = \bar{\nu}'\varphi(t) - c(t) \\ &= \nu'\varphi(t) - c(t) + \sigma^0\delta(t) \\ &\Rightarrow \Delta\nu'\varphi(t) = \sigma^0\delta'(t). \end{aligned}$$

Ici, $\delta(t)$ est la direction, et σ^0 le pas dual. Par suite :

$$\Delta\nu' = \sigma^0\delta'(T_B)\varphi_B^{-1}.$$

Posons :

$$\delta(t) = \begin{cases} 0, & \text{sur } T_B/t_1, \\ +1, & \text{si } u(t) = f_* \text{ pour } t = t_1; \\ -1, & \text{si } u(t) = f^* \text{ pour } t = t_1; \end{cases}$$

Déterminons le pas dual :

$$\sigma(t) = \begin{cases} -\frac{\tilde{\Delta}(t)}{\delta(t)}, & \text{si } \tilde{\Delta}(t)\delta(t) < 0 \\ 0, & \text{si } \tilde{\Delta}(t) = 0, \delta(t) > 0, \tilde{w}(t) \neq f_* \text{ ou } \tilde{\Delta}(t) = 0, \delta(t) < 0, \tilde{w}(t) \neq f^* \\ +\infty, & \text{sinon, } t \in T. \end{cases} \quad (2.48)$$

Construisons l'ensemble :

$$T(\sigma) = \{t \in T : \sigma(t) < \sigma\}.$$

et désignons par :

$$\alpha(\sigma) = -|\lambda(t_1)| + 2 \int_{T(\sigma)} |\delta(t)| dt.$$

la vitesse de décroissance de la fonctionnelle duale. Par construction :

$$\alpha(0) < 0, \text{ et } \alpha(\sigma) \leq \alpha(\bar{\sigma}), \text{ pour } \sigma < \bar{\sigma}.$$

Si $\alpha(\sigma) < 0$ pour $\sigma > 0$, alors le problème (2.7)-(2.10) ne possède pas de commande admissible, dans le cas contraire cherchons un tel $\sigma^0 \geq 0$ tel que :

$$\alpha(\sigma^0 - y) < 0, \alpha(\sigma^0 + 0) \geq 0, \forall y/0 < y \leq \sigma^0.$$

Soit $t_* \in T/\tilde{T}_B$ un tel moment vérifiant :

$$\tilde{\Delta}(t_*) + \sigma_0 \delta(t_*) = 0, \delta(t_*) \neq 0. \quad (2.49)$$

On obtient alors le nouveau support

$$\bar{T}_B = (\tilde{T}_B/\{t_1\}) \cup \{t_*\}.$$

- Si $\lambda(\bar{T}_B) = 0$, alors la quasi-commande \tilde{w} est optimale pour le problème (2.7)-(2.10).
- Sinon dans le cas où $\|\lambda(\tilde{T}_B)\| > \mu$, on refait l'itération jusqu'à ce que $\|\lambda(\tilde{T}_B)\| < \mu$, puis on applique la procédure finale.

Remarque 2.4 en faisant une itération, $\bar{T}_B \rightarrow \tilde{T}_B$ la fonctionnelle duale diminue de la quantité :

$$\int_0^{\sigma_0} \alpha(\sigma) d\sigma.$$

2.12.3 Procédure finale

Supposons que pour la quasi commande \tilde{w} et la quasi-trajectoire correspondante χ construite par le support \tilde{T}_B on ait la condition :

$$\|\lambda(\tilde{T}_B)\| \leq \mu.$$

Désignons par :

$$T^0 = \{t \in T : \tilde{\Delta}(t) = 0\}, T^0 = \{t_j, j = \overline{1, m}\}$$

l'ensemble des points isolés t_j , $j = \overline{1, m}$ et supposons que :

$$\dot{\Delta}(t_j) \neq 0, \quad j = \overline{1, m}.$$

La procédure finale consiste à déterminer $\tau_B^0 = \{\tau_j^0, j = \overline{1, m}\}$ à partir des équations :

$$(f^* - f_*) \sum_{j=1}^m \text{sign} \dot{\Delta}(t_j) \int_{\tau_j}^{\tau_j} \varphi(t) dt = g - H\chi(t^*), \quad (2.50)$$

obtenues à partir de la contrainte :

$$\begin{aligned} g - H\chi(t^*) &= g - HF(t^*)x_0 - \int_T \varphi(t)u(t)dt \\ &= g - HF(t^*)x_0 - \int_{T_H} \varphi(t)u(t)dt - \int_{T_0} \varphi(t)u(t)dt; \end{aligned}$$

en identifiant la partie hors base à zéro et la partie basique à μ .

La résolution de (2.50) se fera par la méthode de Newton (c'est pourquoi on cherche à avoir $\|\lambda(\tilde{T}_B)\| \leq \mu$, μ petit) pour assurer la convergence de cette méthode. Pour la résolution des équations (2.50), on prend comme approximation initiale de

$$\tau_B^0 = \tilde{\tau}_B = \{\tau_j^{(k)}, j = \overline{1, m}\},$$

On a alors la relation de récurrence suivante :

$$\tau_B^{(k+1)} = \tau_B^{(k)} + \frac{1}{f^* - f_*} \{\text{sign} \tilde{\Delta}(t_j) \lambda_j(\tau_B^{(k)}), j = \overline{1, m}\} \quad (2.51)$$

où $\lambda(\tau_B^{(k)})$ est un vecteur calculé par la relation (2.47). La fonction $w^0(t) = \tilde{w}(t)$, $t \in T$ calculée par le support τ_B^0 solution de (2.50), est une commande optimale du problème (2.7)-(2.10).

Chapitre 3

Optimisation d'un problème de contrôle optimale avec une entrée libre

3.1 Introduction

La théorie du contrôle analyse les propriétés des systèmes commandés c'est à dire des systèmes dynamiques sur lequel on peut agir. Les problèmes de contrôle optimale occupent une large place dans la littérature depuis plus de 40 ans. Durant cette période, des résultats fondamentaux ont été retrouvés comme le principe du maximum de Pontriaguine, ou le principe de programmation dynamique de Bellman. A cet effet, beaucoup de problèmes de contrôle optimal ont été résolus, et appliqués dans différents domaines de la science, mécanique, électricité, électronique, économie,...etc. Dans ce chapitre, on s'intéresse à un problème de contrôle optimal avec une entrée libre c'est-à-dire que la condition initiale n'est pas fixé. On cherche la meilleur condition initiale $x_0 \in X_0$, et une commande qui nous permet de ramener le système de l'état initial vers l'état final soumis à la contrainte $Hx(t^*) = g$. Pour sa résolution, on applique la méthode directe vue au chapitre 2.

3.2 Position du problème

Considérons le problème de contrôle optimal suivant :

$$J(z, u) = c'x(t^*) \rightarrow \max \quad (3.1)$$

$$\begin{aligned} \dot{x} &= Ax + bu, x(0) = z \in X_0, \\ X_0 &= \{z \in R^n, Gz = \gamma, d_* \leq z \leq d^*\}, \end{aligned} \quad (3.2)$$

$$Hx(t^*) = g, \quad (3.3)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T = [0, t^*]. \quad (3.4)$$

où $x = x(t) \in \mathbb{R}^n$ est le vecteur d'état du système dynamique (3.2) à l'instant t ; $u(\cdot) = (u(t), t \in T = [0, t^*])$, est une fonction continue par morceaux; $A \in \mathbb{R}^{n \times n}$; $b, c \in \mathbb{R}^n$; $H \in \mathbb{R}^{m \times n}$, $\text{rang } H = m \leq n$; $g \in \mathbb{R}^m$, f_* , f^* sont des scalaires; $d_* = (d_{*j}, j \in J)$, $d^* = d^*(J) = (d_j^*, j \in J)$ sont des n -vecteurs; $G \in \mathbb{R}^{l \times n}$, $\text{rang } G = l \leq n$, $\gamma \in \mathbb{R}^l$, $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$, $L = \{1, \dots, l\}$ sont des ensembles d'indices.

En utilisant la formule de Cauchy, on trouve la solution du système (3.2) :

$$x(t) = F(t)\left(z + \int_0^t F^{-1}(\vartheta)bu(\vartheta)d\vartheta\right), t \in T, \quad (3.5)$$

où $F(t) = e^{At}$, $t \in T = [0, t^*]$ est la solution (résolvante) du système suivant :

$$\begin{cases} \dot{F}(t) = AF(t) & , \\ F(0) = Id & . \end{cases}$$

En remplaçant (3.5) dans (3.1)-(3.4), on obtient le problème des seules variables z et de la commande $u(t)$, $t \in T$:

$$J(z, u) = \tilde{c}'z + \int_0^{t^*} c(t)u(t)dt \longrightarrow \max_{z, u}, \quad (3.6)$$

$$D(I, J)z + \int_0^{t^*} \varphi(t)u(t)dt = g, \quad (3.7)$$

$$G(L, J)z = \gamma, \quad d_* \leq z \leq d^*, \quad (3.8)$$

$$f_* \leq u(t) \leq f^*, t \in T, \quad (3.9)$$

où $\tilde{c}' = c'F(t^*)$, $c(t) = c'F(t^*)F^{-1}(t)b$, $D(I, J) = HF(t^*)$, $\varphi(t) = HF(t^*)F^{-1}(t)b$, c' est le transposé de c .

3.3 Définitions Essentielles

Définition 3.1 La paire $v = (z, u(\cdot))$ formé d'un n -vecteur z et de la fonction continue par morceaux $u(\cdot)$ est appelée commande généralisée.

Définition 3.2 La commande généralisée $v = (z, u(\cdot))$ est dite admissible si elle vérifie les contraintes (3.2)-(3.4).

Définition 3.3 La commande admissible $v^0 = (z^0, u^0(\cdot))$ est dite optimale si elle réalise le maximum du critère de qualité c'est-à-dire :

$$J(v^0) = \max_v J(v).$$

Définition 3.4 Soit $\varepsilon \geq 0$ donné. La commande $v^\varepsilon = (z^\varepsilon, u^\varepsilon(\cdot))$ est dite ε -optimale si

$$J(v^0) - J(v^\varepsilon) \leq \varepsilon.$$

3.4 Support contrôle

De l'ensemble T , choisissons un sous-ensemble T_B constitué de k éléments et un sous-ensemble $J_B \subset J$ de $m+l-k$ éléments, avec $k \leq m$. En utilisant ces deux ensembles, formons la matrice :

$$P_B = \begin{pmatrix} D(I, J_B) & \varphi(t), t \in T_B \\ G(L, J_B) & 0 \end{pmatrix} \quad (3.10)$$

L'ensemble $S_B = \{T_B, J_B\}$ est appelé support du problème (3.1)-(3.4) si $\det P_B \neq 0$.

La paire $\{v, S_B\}$ formé d'une commande admissible $v = (z, u(\cdot))$ et du support S_B est appelé support contrôle. Le support contrôle $\{v, S_B\}$ est dit non dégénéré si $d_{*j} < z_j < d_j^*, j \in J_B, f_* < u(t) < f^*, t \in T_B$.

Considérons une commande admissible $\bar{v} = (\bar{z}, \bar{u}(\cdot)) = v + \Delta v$, où $\bar{z} = z + \Delta z, \bar{u}(t) = u(t) + \Delta u(t), t \in T$, et calculons l'accroissement de la fonctionnelle :

$$\Delta J(v) = J(\bar{v}) - J(v) = \bar{c}' \Delta z + \int_{t \in T} c(t) \Delta u(t) dt.$$

De l'admissibilité de v et \bar{v} , on a :

$$D(I, J) \Delta z + \int_{t \in T} \varphi(t) \Delta u(t) dt = 0,$$

et

$$G(L, J) \Delta z = 0.$$

En remplaçant \bar{c} et $c(t)$ par leurs valeurs, l'accroissement de la fonctionnelle prend la forme :

$$\Delta J(v) = (\bar{c}' - \nu' \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix}) \Delta z + \int_{t \in T} (\varphi(t) - \nu' c(t)) \Delta u(t) dt,$$

où $\nu = \begin{pmatrix} \nu_u \\ \nu_z \end{pmatrix} \in R^{m+l}$, $\nu_u \in R^m$, $\nu_z \in R^l$ est la fonction des multiplicateurs de Lagrange appelée vecteur des potentiels, calculé comme suit : $\nu' = q'_B Q$, où $Q = P_B^{-1}$, $q_B = (\tilde{c}_j, j \in J_B, c(t), t \in T_B)$. Introduisons le n -vecteur des estimations $\Delta' = \nu' \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix} - \tilde{c}'$, et la fonction de la co-commande $\Delta(\cdot) = (\Delta(t) = \nu'_u \varphi(t) - c(t), t \in T)$. En utilisant ces notions, la valeur de l'accroissement de la fonctionnelle prend la forme suivante :

$$\Delta J(v) = \Delta' \Delta z - \int_{t \in T} \Delta(t) \Delta u(t). \quad (3.11)$$

Le support contrôle $\{v, S_B\}$ est non dégénérée pour le problème dual si $\Delta(t) \neq 0, t \in T_H, \Delta_j \neq 0, j \in J_H$, où $T_H = T/T_B, J_H = J/J_B$.

3.5 Problème dual

Définissons le problème dual du problème (3.6)-(3.9) :

$$\left\{ \begin{array}{l} L(\varsigma, v, w) = g\varsigma - \int_{t_0}^{t^*} v(t) f_* dt + \int_{t_0}^{t^*} w(t) f^* dt \rightarrow \min, \\ \varsigma(\phi(t) + HF(t^*)) - v(t) + w(t) = c(t), \\ v(t) \geq 0, w(t) \geq 0, t \in T, \end{array} \right.$$

Ici les fonctions $v(t), w(t), t \in T$ pour $\nu(I) = \varsigma(I)$ sont définies de la manière suivante :

$$\left\{ \begin{array}{l} v(t) = \Delta(t), w(t) = 0, \quad \text{si } \Delta(t) \geq 0 \\ v(t) = 0, w(t) = -\Delta(t), \quad \text{si } \Delta(t) \leq 0. \end{array} \right. \quad (3.12)$$

et forment un plan dual du problème (solution admissible du dual).

Désignons par $(\varsigma^0, v^0(t), w^0(t), t \in T)$ sa solution optimale.

3.6 Calcul de la valeur de suboptimalité

De l'admissibilité de $\bar{v}(t)$, on a les relations suivantes : $\bar{v}(t)$ est admissible, si elle satisfait les contraintes :

$$d_* - z \leq \Delta z \leq d^* - z; f_* - u(t) \leq \Delta u(t) \leq f^* - u(t), t \in T. \quad (3.13)$$

Le maximum de l'accroissement de la fonctionnelle (3.11) sous les contraintes (3.13) est atteint pour :

$$\begin{cases} \Delta z_j = d_{*j} - z_j & \text{si } \Delta_j > 0 \\ \Delta z_j = d_j^* - z_j & \text{si } \Delta_j < 0 \\ d_{*j} - z_j \leq \Delta z_j \leq d_j^* - z_j, & \text{si } \Delta_j = 0, j \in J, \end{cases}$$

$$\begin{cases} \Delta u(t) = f_* - u(t) & \text{si } \Delta(t) > 0 \\ \Delta u(t) = f^* - u(t) & \text{si } \Delta(t) < 0 \\ f_* \leq \Delta u(t) \leq f^*, & \text{si } \Delta(t) = 0, t \in T. \end{cases}$$

et égal à :

$$\begin{aligned} \beta &= \beta(v, S_B) = \sum_{j \in J_H^+} \Delta_j (z_j - d_{*j}) + \sum_{j \in J_H^-} \Delta_j (z_j - d_j^*) \\ &+ \int_{t \in T^+} \Delta(t) (u(t) - f_*) dt + \int_{t \in T^-} \Delta(t) (u(t) - f^*) dt, \end{aligned} \quad (3.14)$$

où

$$\begin{aligned} T^+ &= \{t \in T_H, \Delta(t) > 0\}, \quad T^- = \{t \in T_H, \Delta(t) < 0\}, \\ J_H^+ &= \{j \in J_H, \Delta_j > 0\}, \quad J_H^- = \{j \in J_H, \Delta_j < 0\}. \end{aligned}$$

La valeur $\beta(v, S_B)$ est appelée valeur de suboptimalité du support contrôle $\{v, S_B\}$.

On en déduit que

$$J(\bar{v}) - J(v) \leq \beta(v, S_B),$$

pour tout \bar{v} ; et pour $\bar{v} = v^0$, on aura :

$$J(v^0) - J(v) \leq \beta(v, S_B).$$

De cette dernière inégalité, on déduit un critère d'Optimalité et d' ε -optimalité :

3.7 Critère d'Optimalité et $d'\varepsilon$ -Optimalité

Théorème 11 [34] *Pour l'optimalité du support-contrôle $\{v, S_B\}$, les relations suivantes :*

$$\left\{ \begin{array}{ll} u(t) = f_*, & \text{si } \Delta(t) > 0, \\ u(t) = f^*, & \text{si } \Delta(t) < 0, \\ f_* \leq u(t) \leq f^*, & \text{si } \Delta(t) = 0, t \in T, \\ z_j = d_{*j}, & \text{si } \Delta_j > 0, \\ z_j = d_j^*, & \text{si } \Delta_j < 0, \\ d_{*j} \leq z_j \leq d_j^*, & \text{si } \Delta_j = 0, j \in J, \end{array} \right. \quad (3.15)$$

sont suffisantes, et dans le cas de la non dégénérescence, elles sont nécessaires.

Preuve.

Condition suffisante Si les relations (3.15) sont satisfaites, alors $\beta(v, S_B) = 0$ et de là, on a :

$$J(v^0) - J(v) = \Delta J(v) \leq \beta(v, S_B) = 0,$$

donc $\{v, S_B\}$ est un support contrôle optimale.

Condition Nécessaire Procédons par l'absurde :

Soit $\{v, S_B\}$ un support contrôle optimal non dégénéré et supposons que les relations (3.15) ne sont pas satisfaites, c'est-à-dire :

$\exists \bar{t} \in T$ tel que :

$$\Delta(\bar{t}) > 0 \quad \text{et} \quad u(\bar{t}) > f_* \quad \text{ou} \quad \Delta(\bar{t}) < 0 \quad \text{et} \quad u(\bar{t}) < f^*;$$

où $\exists j_0 \in J$ tel que :

$$\Delta_{j_0} > 0 \quad \text{et} \quad z_{j_0} > d_{*j_0} \quad \text{ou} \quad \Delta_{j_0} < 0 \quad \text{et} \quad z_{j_0} < d_{j_0}^*.$$

Il est facile de construire les variations admissibles de Δv , telles que la valeur de l'accroissement de la fonctionnelle soit :

$$\Delta J(v) = \Delta' \Delta z - \int_{t \in T} \Delta(t) \Delta u(t) > 0,$$

$$J(\bar{v}) - J(v) > 0,$$

et ceci contredit l'optimalité de $\{v, S_B\}$.

Théorème 12 *Pour $\varepsilon \geq 0$, la commande admissible v est ε -optimale si et seulement s'il existe un tel support S_B tel que $\beta(v, S_B) \leq \varepsilon$.*

La preuve se fait de manière analogue que celle du critère 4 du deuxième chapitre.

3.8 Algorithme de résolution

Pour $\varepsilon > 0$ donné, supposons que le support contrôle $\{v, S_B\}$ ne vérifie pas le critère d' ε -optimalité. De là on passe à l'itération de l'algorithme $\{v, S_B\} \rightarrow \{\bar{v}, \bar{S}_B\}$ tel que $\beta(v, S_B) \geq \beta(\bar{v}, \bar{S}_B)$ constitué de trois procédures :

1. Changement de commande admissible $v \rightarrow \bar{v}$,
2. Changement de support $S_B \rightarrow \bar{S}_B$,
3. Procédure finale.

3.8.1 Changement de commande

Soient $\alpha_1 > 0, \alpha_2 > 0, h > 0, \mu > 0$ les paramètres de la méthode, et construisons les ensembles suivants :

$$J_0 = \{j \in J : |\Delta_j| \leq \alpha_2\}, \quad J_* = \{j \in J : |\Delta_j| > \alpha_2\}, \quad J_* = J/J_0,$$

et

$$T_0 = \{t \in T : |\Delta(t)| \leq \alpha_1\}, \quad T_* = \{t \in T : |\Delta(t)| > \alpha_1\}, \quad T_* = T/T_0.$$

Subdivisons T_0 en sous intervalles $[\tau_i, \tau^i[$, $i = \overline{1, N}$; $\tau_i < \tau^i$, $T_0 = \bigcup_{i=1}^N [\tau_i, \tau^i[$, $\tau^i - \tau_i \leq h$, $T_B \subset T_0$, $u(t) = u_i = \text{constante}$, $t \in [\tau_i, \tau^i[$, $i = \overline{1, N}$.

La nouvelle commande admissible $\bar{v} = (\bar{z}, \bar{u}(t), t \in T)$ est telle que :

$$\begin{cases} \bar{z}_j = z_j + \kappa \Delta z_j, & j \in J \\ \bar{u}(t) = u(t) + \theta \Delta u(t), & t \in T. \end{cases} \quad (3.16)$$

Posons :

$$\Delta z_j = \begin{cases} d_j^* - z_j, & \text{si } \Delta_j < -\alpha_2, \\ d_{*j} - z_j, & \text{si } \Delta_j > \alpha_2, \quad j \in J_*, \\ 0, & \text{si } \Delta_j = 0, \quad j \in J_0, \end{cases}$$

$$\Delta u(t) = \begin{cases} f^* - u(t), & \text{si } \Delta(t) < -\alpha_1, \\ f_* - u(t), & \text{si } \Delta(t) > \alpha_1, \quad t \in T_*, \\ u_i = \text{constante}, & \text{si } t \in [\tau_i, \tau^i], \quad i = \overline{1, N}, t \in T_0. \end{cases}$$

On introduit les vecteurs suivants :

$l_i = \theta u_i, i = \overline{1, N}, h_j = \kappa \Delta z_j, j \in J_0, h_{K+1} = \kappa$, et on calcule les quantités suivantes :

$$g_i = - \int_{\tau_i}^{\tau^i} \Delta(t) dt, i = \overline{1, N}, g_{N+1} = - \int_{T_*} \Delta(t) \Delta u(t) dt,$$

$$\phi_i = - \int_{\tau_i}^{\tau^i} \varphi(t) dt, i = \overline{1, N}, \phi_{N+1} = - \int_{T_*} \varphi(t) \Delta u(t),$$

$$q_j = -\Delta_j, j \in J_0, q_{K+1} = \sum_{j \in J_*} -\Delta_j \Delta z_j, j \in J_*,$$

$$D_j = D(I, j), j \in J_0, D_{K+1} = \sum_{j \in J_*} D(I, j) \Delta z_j,$$

$$f_{*i} = f_* - u_i, f_i^* = f^* - u_i, i = \overline{1, N}, f_{*N+1} = 0, f_{N+1}^* = 1,$$

$$d_{*j} = d_* - z_j, d_j^* = d^* - z_j, j = \overline{1, K}, d_{*K+1} = 0, d_{K+1}^* = 1.$$

Pour trouver la nouvelle commande, on cherche la solution $(h_j, l_i), j = \overline{1, K+1}, i = \overline{1, N+1}$ du problème (3.6)-(3.9) qui devient un problème de programmation linéaire suivant obtenu en utilisant l'accroissement de la fonctionnelle :

$$\left\{ \begin{array}{l} \Delta J(v) = \sum_{j \in J_0 \cup \{K+1\}} q_j h_j + \sum_{i=1}^{N+1} g_i l_i \rightarrow \max_{h_j, l_i}, \\ \sum_{j \in J_0 \cup \{K+1\}} D(I, j) h_j + \sum_{i=1}^{N+1} \phi_i l_i = 0, \\ \sum_{j \in J_0 \cup \{K+1\}} G(l, j) h_j = 0, \\ f_{*i} \leq l_i \leq f_i^*, \quad i = \overline{1, N+1} \\ d_{*j} \geq h_j \geq d_j^*, \quad j = \overline{1, K+1}. \end{array} \right.$$

le problème (3.17) est appelé problème du support qu'on résoud par la méthode adaptée dont la solution de départ admissible avec le support plan $(h_j = 0, l_i = 0, J_B, T_B)$. On obtient un ε -support plan optimal $(h_j^\varepsilon, l_i^\varepsilon, \bar{J}_B, \bar{T}_B)$. La nouvelle commande $\bar{v} = (\bar{z}, \bar{u}(t), t \in T)$ sera alors :

$$\bar{z}_j = \begin{cases} z_j + h_{K+1} \Delta z_j, & j \in J_* \\ z_j + h_j, & j \in J_0. \end{cases} \quad (3.17)$$

$$\bar{u}(t) = \begin{cases} u(t) + l_{N+1} \Delta u(t), & t \in T_* \\ u(t) + l_i, & t \in [\tau_i, \tau^i], i = \overline{1, N}. \end{cases} \quad (3.18)$$

il est clair que : $J(\bar{v}) \geq J(v)$.

– Cas $K + 1 \notin \bar{J}_B$ et $t_{N+1} \notin \bar{T}_B$, alors on pose :

$$\tilde{S}_B = \{\tilde{J}_B = \bar{J}_B, \tilde{T}_B = \bar{T}_B\}.$$

– Dans le cas contraire, On aura les cas suivants :

1. Cas $K + 1 \notin \bar{J}_B$ et $t_{N+1} \in \bar{T}_B$, alors on exclut l'indice $N + 1$ du support de la manière suivante : soit :

$$\bar{\Delta}(t) = \Delta(t) + \sigma \delta(t),$$

où σ est le pas dual maximal le long de la direction $\delta(t)$. Déterminons i_* tel que :

$$\sigma(t_{i_*}) = \min \sigma(t_i), t_i \in T_H,$$

avec

$$\sigma(t_i) = \begin{cases} -\Delta(t_i)/\delta(t_i), & \text{si } \Delta(t_i) \times \delta(t_i) \leq 0, \delta(t_i) \neq 0 \\ +\infty, & \text{sinon.} \end{cases}$$

$$\delta(t) = \begin{cases} 0, & \text{sur } T_B/\{t_{N+1}\}; \\ 1, & \text{si } \bar{u}(t) = f_*; \\ -1, & \text{si } \bar{u}(t) = f^*. \end{cases}$$

$$\delta(t) = \delta'_B P_B^{-1} \phi(t), t \in T.$$

alors le nouveau support sera :

$$\tilde{J}_B = \bar{J}_B; \tilde{T}_B = (\bar{T}_B/\{t_{N+1}\}) \cup \{t_{i_*}\}.$$

2. Cas $K + 1 \in \bar{J}_B$ et $t_{N+1} \notin \bar{T}_B$, On exclut l'indice $K + 1$ du support de la manière suivante :

$$\bar{\Delta}_j = \Delta_j + \sigma_j \delta_j,$$

où σ_j est le pas dual maximal le long de la direction δ_j .
Déterminons j_* tel que :

$$\sigma_{j_*} = \min \sigma_j, j \in J_H,$$

avec

$$\sigma_j = \begin{cases} -\Delta_j/\delta_j, & \text{si } \Delta_j \times \delta_j \leq 0, \delta_j \neq 0 \\ +\infty, & \text{sinon.} \end{cases}$$

$$\delta_j = \begin{cases} 0, & \text{sur } J_B/\{K + 1\}; \\ 1, & \text{si } \bar{z}_j = d_*; \\ -1, & \text{si } \bar{z}_j = d^*. \end{cases}$$

$$\delta_j = \delta'_B P_B^{-1} \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix}, j \in J.$$

Alors le nouveau support sera :

$$\tilde{J}_B = (\bar{J}_B/\{K + 1\}) \cup \{j_*\}; \tilde{T}_B = \bar{T}_B.$$

3. Cas $K + 1 \in \bar{J}_B, t_{N+1} \in \bar{T}_B$, le nouveau support sera :

$$\tilde{J}_B = (\bar{J}_B/\{K + 1\}) \cup \{j_*\}; \tilde{T}_B = (\bar{T}_B/\{t_{N+1}\}) \cup \{t_{i_*}\}.$$

A ce stade, dénotons le nouveau support par \tilde{S}_B et construisons la matrice du support $P(\tilde{S}_B)$ avec laquelle on calcule la nouvelle valeur de suboptimalité $\beta(\tilde{v}, \tilde{S}_B)$.

1. Si $\beta(\tilde{v}, \tilde{S}_B) = 0$, alors \bar{v} est une commande optimale.
2. Si $\beta(\tilde{v}, \tilde{S}_B) \leq \varepsilon$, alors \bar{v} est une commande ε -optimale.
3. Sinon, on passe soit à la nouvelle itération avec $\{\bar{v}, \tilde{S}_B\}, \bar{\alpha}_1 < \alpha_1, \bar{\alpha}_2 < \alpha_2, \bar{h} < h$, soit à la procédure du changement de support.

3.8.2 Changement de support

Supposons que pour la nouvelle commande \bar{v} , on a $\beta(\bar{v}, \tilde{S}_B) > \varepsilon$, alors on passe au changement de support. En utilisant le support \tilde{S}_B , on construit la quasi-commande $\tilde{v} = (\tilde{z}, \tilde{u}(t), t \in T)$:

$$\tilde{z}_j = \begin{cases} d_{j*} & \text{si } \tilde{\Delta}_j > 0 \\ d_j^* & \text{si } \tilde{\Delta}_j < 0 \\ \in [d_{j*}, d_j^*] & \text{si } \tilde{\Delta}_j = 0, j \in J, \end{cases}$$

$$\tilde{u}(t) = \begin{cases} f_*, & \text{si } \tilde{\Delta}(t) < 0 \\ f^*, & \text{si } \tilde{\Delta}(t) > 0, \\ \in [f_* f^*], & \text{si } \tilde{\Delta}(t) = 0, t \in T, \end{cases}$$

où : $\tilde{\Delta}(t) = -\tilde{\psi}'(t)b, t \in T, \tilde{\Delta}' = (\tilde{\Delta}_j, j \in J)' = \nu' \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix} - \tilde{c}'$.

Ici, $\tilde{\psi}(t), t \in T$, est la solution du système adjoint correspondant à \tilde{S}_B , et la quasi-trajectoire correspondante $\chi = (\chi(t), t \in T), \chi(0) = z \in X_0$ est solution du système $\dot{\chi} = A\chi + b\tilde{u}, \chi(0) = z \in X_0$.

Si

$$D(I, J)\tilde{z} + \int_0^{t^*} \varphi(t)\tilde{u}(t)dt = g,$$

$$G(L, J)\tilde{z} = \gamma,$$

alors \bar{v} est une commande optimale, sinon on construit le vecteur $\lambda(\tilde{J}_B, \tilde{T}_B)$ comme suit :

$$P(\tilde{S}_B) \cdot \lambda(\tilde{J}_B, \tilde{T}_B) = \begin{pmatrix} D(I, J)\tilde{z} + \int_0^{t^*} \tilde{u}(t)dt - g \\ G(L, J)\tilde{z} - \gamma \end{pmatrix},$$

c'est-à-dire

$$\lambda(\tilde{J}_B, \tilde{T}_B) = P_B^{-1}(\tilde{S}_B) \begin{pmatrix} D(I, J)\tilde{z} + \int_0^{t^*} \tilde{u}(t)dt - g \\ G(L, J)\tilde{z} - \gamma \end{pmatrix}.$$

Maintenant, on va étudier les cas suivants :

- Si $\|\lambda(\tilde{J}_B, \tilde{T}_B)\| = 0$, alors la quasi-commande \tilde{v} est optimale pour le problème (3.1) – (3.4).
- Si $\|\lambda(\tilde{J}_B, \tilde{T}_B)\| > \mu$, alors on change le support de \tilde{S}_B en \bar{S}_B par la méthode duale.
- Si $\|\lambda(\tilde{J}_B, \tilde{T}_B)\| < \mu$, alors on passe à la procédure finale.

Méthode duale.

Ici notre but est de diminuer la valeur de $\lambda(\tilde{J}_B, \tilde{T}_B)$ pour pouvoir aller en procédure finale et utiliser la méthode de Newton pour cela soit :

$$\lambda_0 = \max_{\{j \in J_B, t \in T_B\}} |\lambda(t), \lambda_j|.$$

On aura à considérer deux cas :

1. Si $\lambda_0 = |\lambda_{j_0}|, j_0 \in \tilde{J}_B$,
on exclut l'indice j_0 de \tilde{J}_B . Pour cela on calcul le pas dual de la manière suivante :

$$\sigma_j = \begin{cases} -\tilde{\Delta}_j / \delta_j, & \text{si } \tilde{\Delta}_j \delta_j < 0, \delta_j \neq 0 \\ 0, & \text{si } \tilde{\Delta}_j = 0, \delta_j > 0, \bar{z} \neq d_* \text{ ou } \tilde{\Delta}_j = 0, \delta_j < 0, \bar{z} \neq d^*, \\ +\infty, & \text{sinon, } j \in J. \end{cases}$$

Construisons l'ensemble :

$$J(\sigma) = \{j \in J : \sigma_j < \sigma\},$$

et la vitesse de décroissance de la fonctionnelle duale est égale à :

$$\alpha(\sigma) = -|\lambda_{j_0}| + 2 \sum_{J(\sigma)} |\delta_j|.$$

Par construction :

$$\alpha(0) = -|\lambda_{j_0}| < 0 \text{ et } \alpha(\sigma) < \alpha(\bar{\sigma}) \text{ si } \sigma < \bar{\sigma}, \text{ si } \alpha(\sigma) < 0 \text{ pour } \sigma > 0,$$

alors le problème (3.1) – (3.4) ne possède pas de commande admissible. Dans le cas contraire, cherchons un tel $\sigma_0 \geq 0$ tel que :

$$\alpha(\sigma_0 - y) < 0, \alpha(\sigma_0 + 0) \geq 0, \forall 0 \leq y \leq \sigma_0.$$

Soit $j_* \in J/\tilde{J}_B$ un tel moment vérifiant :

$$\tilde{\Delta}(j_*) + \sigma^0 \delta_{j_*} = 0, \delta_{j_*} \neq 0,$$

alors le nouveau support \tilde{S}_B change en \hat{S}_B .

$$\hat{S}_B = \{\hat{J}_B = (\tilde{J}_B / \{j_0\}) \cup \{j_*\}, \hat{T}_B = \tilde{T}_B\}.$$

2. Si $\lambda_0 = |\lambda(t_0)|$, $t_0 \in \tilde{T}_B$. Calculons les quantités suivantes (directions de changement du support) :

Le pas dual sera :

$$\sigma(t) = \begin{cases} -\tilde{\Delta}(t)/\delta(t), \text{ si } \tilde{\Delta}(t) \cdot \delta(t) < 0, \delta(t) \neq 0, \\ 0, \text{ si } \tilde{\Delta}(t) = 0, \delta(t) > 0, \bar{u}(t) \neq f_* \text{ ou } \tilde{\Delta}(t) = 0, \delta(t) < 0, \bar{u}(t) \neq f^*, \\ +\infty, \text{ sinon; } t \in T. \end{cases}$$

Construisons l'ensemble suivant :

$$T(\sigma) = \{t \in T : \sigma(t) < \sigma\}.$$

La vitesse de décroissance de la fonctionnelle duale est égale à :

$$\alpha(\sigma) = -|\lambda(t_0)| + 2 \int_{T(\sigma)} |\delta(t)| dt.$$

Par construction :

$$\alpha(0) = -|\lambda(t_0)| < 0 \text{ et } \alpha(\sigma) < \alpha(\bar{\sigma}) \text{ si } \sigma < \bar{\sigma}, \text{ si } \alpha(\sigma) < 0 \text{ pour } \sigma > 0,$$

alors le problème (3.1) – (3.4) ne possède pas de commande admissible.

Dans le cas contraire, cherchons un tel $\sigma^0 \geq 0$ tel que :

$$\alpha(\sigma^0 - y) < 0, \alpha(\sigma^0 + 0) \geq 0, \forall 0 \leq y \leq \sigma^0.$$

Soit $t_* \in T/\tilde{T}_B$ un tel moment vérifiant :

$$\tilde{\Delta}(t_*) + \sigma^0 \delta(t_*) = 0, \delta(t_*) \neq 0,$$

alors le nouveau support \tilde{S}_B change en \hat{S}_B .

$$\hat{S}_B = \{\hat{J}_B = \tilde{J}_B, \hat{T}_B = (\tilde{T}_B/\{t_0\}) \cup \{t_*\}\}.$$

Calculons la nouvelle valeur de suboptimalité $\beta(\bar{v}, \hat{S}_B)$ et vérifiant les cas suivants :

1. Si $\beta(\bar{v}, \hat{S}_B) = 0$, alors la commande \bar{v} est optimale pour le problème (3.1)-(3.4) .
2. Si $\beta(\bar{v}, \hat{S}_B) < \varepsilon$, alors la commande \bar{v} est ε -optimale pour le problème (3.1)-(3.4) .
3. Si $\beta(\bar{v}, \hat{S}_B) > \varepsilon$, alors on passe à une nouvelle itération avec $\{\bar{v}, \hat{S}_B\}$.

3.8.3 Procédure finale

Soit S_B un tel support tel que $\|\lambda(\tilde{J}_B, \tilde{T}_B)\| < \mu$. A partir du support S_B , on construit la quasi-commande $w = (z, u(t))$, $t \in T$:

$$z_j = \begin{cases} d_{j*} & \text{si } \Delta_j > 0 \\ d_j^* & \text{si } \Delta_j < 0 \\ \in [d_{j*}, d_j^*], & \text{si } \Delta_j = 0, j \in J \end{cases}$$

$$u(t) = \begin{cases} f_*, & \text{si } \Delta(t) < 0 \\ f^*, & \text{si } \Delta(t) > 0, t \in T \\ \in [f_*, f^*], & \text{si } \Delta(t) = 0, t \in T \end{cases}$$

si

$$D(I, J)z + \int_0^{t^*} \varphi(t)u(t)dt = g,$$

$$G(L, J)z = \gamma,$$

alors w est optimale, et si

$$D(I, J)z + \int_0^{t^*} \varphi(t)u(t)dt \neq g,$$

$$G(L, J)z \neq \gamma,$$

alors désignons par : $T^0 = \{t_i, i = \overline{1, s}\}$, $s = |T_B|$. Ici, $t_i, i = \overline{1, s}$ l'ensemble des points isolés $\Delta(t) = 0, t \in T$; $t_0 = 0, t_{s+1} = t^*$.

Supposons que

$$\dot{\Delta}(t_i) \neq 0, i = \overline{1, s}.$$

L'annulation du vecteur λ est équivalent à chercher les racines de la fonction :

$$f(\Theta) = \begin{pmatrix} D(I, J_B)z(J_B) + D(I, J_H)z(J_H) + \sum_{i=0}^s \left(\frac{f^* + f_*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i) \right) \int_{t_i}^{t_{i+1}} \varphi(t)dt - g \\ G(L, J_B)z(J_B) + G(L, J_H)z(J_H) - \gamma \end{pmatrix}$$

où

$$z_j = \frac{d_j^* + d_{j*}}{2} - \frac{d_j^* - d_{j*}}{2} \text{sign} \Delta_j, j \in J_H.$$

$$\Theta = (t_i, i = \overline{1, s}; z_j, j \in J_B).$$

La procédure finale consiste à trouver la solution

$$\Theta^0 = (t_i^0, i = \overline{1, s}; z_j^0, j \in J_B)$$

du $(m + l)$ équations non linéaire :

$$f(\Theta) = 0. \quad (3.19)$$

On résout ce système par la méthode de Newton, en utilisant l'approximation initiale :

$$\Theta^{(0)} = (\bar{t}_i, i = \overline{1, s}; \bar{z}_j, j \in J_B).$$

La $(k + 1)^{th}$ approximation $\Theta^{(k+1)}$, sera déterminée par la formule :

$$\Theta^{(k+1)} = \Theta^{(k)} + \Delta\Theta^{(k)}, \quad \text{où } \Delta\Theta^{(k)} = -\frac{\partial f^{-1}(\Theta^{(k)})}{\partial \Theta^{(k)}} \cdot f(\Theta^{(k)}).$$

La Jacobienne de l'équation (3.19) est égale à :

$$\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}} = \begin{pmatrix} D(I, J_B) & (f_* - f^*) \text{sign} \dot{\Delta}(t_i^{(k)}) \varphi(t_i^{(k)}), i = \overline{1, s} \\ G(L, J_B) & 0 \end{pmatrix}.$$

Comme $\det P_B \neq 0$, alors on peut montrer facilement que :

$$\det \frac{\partial f(\Theta^{(0)})}{\partial \Theta^{(0)}} \neq 0. \quad (3.20)$$

Pour l'instant $t \in T_B$, il existe $\mu > 0$ petit, tel que pour chaque $\tilde{t}_i \in [t_i - \mu, t_i + \mu]$, $i = \overline{1, s}$, la matrice $(\varphi(\tilde{t}_i), i = \overline{1, s})$ est non dégénérée et la matrice $\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}}$ est aussi non dégénérée, si les éléments $t_i^{(k)}$, $i = \overline{1, s}$, $k = 1, 2, \dots$ sont au voisinage de μ . Le vecteur $\Theta^{(k^*)}$ est pris comme solution de l'équation (3.19) si

$$\| f(\Theta^{(k^*)}) \| \leq \eta, \quad \eta > 0, \quad \text{donné.} \quad (3.21)$$

Donc posons $\theta^0 = \theta^{(k^*)}$.

La commande suboptimale du problème (3.1)-(3.4) est calculée comme, suit

$$z_j^0 = \begin{cases} z_j^0, & j \in J_B \\ \hat{z}_j, & j \in J_H; \end{cases}$$

$$u^0(t) = \frac{f^* + f_*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i^0), \quad t \in [t_i^0, t_{i+1}^0[, \quad i = \overline{1, s}.$$

Si la méthode de Newton ne converge pas, on diminue les paramètres $\alpha_1 > 0, \alpha_2 > 0, h > 0$ et on refait l'itération, pour faire diminuer $\|\lambda\|$.

3.9 Exemple numérique

Nous illustrons les résultats obtenus par un exemple numérique :

$$c'x(2) \rightarrow \max \quad (3.22)$$

$$\dot{x}_1 = x_2, \dot{x}_2 = u, z \in X_0 = \{z \in \mathbb{R}^2 : Gz = \gamma, -2 \leq z_i \leq 2, i = 1, 2\} \quad (3.23)$$

$$Hx(2) = g, \quad (3.24)$$

$$|u(t)| \leq 1, t \in [0, 2], \quad (3.25)$$

où

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, H = (1 \quad -2), c' = (0 \quad 1),$$

$$g = 2, G = (1 \quad 2), \gamma = 3, f_* = -1, f^* = 1, d_* = (-2, -2), d^* = (2, 2),$$

$$n = 2, m = 1, l = 1, t^* = 2.$$

Soit la commande de départ :

$$u(t) = \begin{cases} 1/2, & t \in [0, 1[\\ -1/2, & t \in]1, 2]. \end{cases}, z = (z_1, z_2) = (3/2, 3/4) \in X_0.$$

$$c(t) = 1, \varphi(t) = -t, \bar{c} = (0 \quad 1), D(I, J) = (1 \quad 0).$$

De la trajectoire $x(t), t \in [0, 2]$ correspondante à (3.23), on obtient :

$$Hx(t^*) = 2, J(v) = 3/4.$$

Pour la première étape de l'algorithme posons $h = 0.25, \alpha_1 = 0.25, \alpha_2 = 0.25, \mu = 0.4, \varepsilon = 0.2$. Déterminons le vecteur (h_j, l_i) solution du problème du support (3.17) en utilisant la méthode adaptée.

Soit le support initial S_B tel que $T_B =]0.85, 1]$ correspondant à $\tau_B = 1, J = \{1, 2\}, J_B = \emptyset$.

La solution du problème (3.17) est : $h_1 = 1, l_1 = 0.5, l_2 = 0.06, \bar{T}_B = [5/4, 2[$ correspond $\tau_B = 5/4, J_B = \emptyset$,

Après trois itérations, Les relations (3.21) sont satisfaites pour l'ensemble des paramètres suivants :

$$\lambda(\bar{T}_B, J_B) = (-0.35, 0), \bar{T}_B = [5/4, 2[\text{ correspondant à } \tau_B = 5/4, J_B = \emptyset,$$

$$\chi(t^*) = \begin{pmatrix} 7.4375 \\ 2.5 \end{pmatrix}.$$

Notons par t_1 le zéro de la fonction $\bar{\Delta}(t)$, $t \in [0, 2]$. On passe à la procédure finale.

Les données obtenues à l'étape précédente sont utilisées pour former le vecteur des paramètres θ et son approximations initiale $\theta^{(0)} = (\bar{z}_j, t_1 = 1)$ avec $\bar{z} = (3/2, 3/4)$, le vecteur $\theta^{(1)}$ est trouvé avec une précision de $\eta = 0.6$ en résolvant le système (3.19), et en utilisant la méthode de Newton. Notons que la forme du système (3.19) est déterminée uniquement par l'ensemble des paramètres $\bar{T}_B = [5/4, 2[$ correspondant à $\tau_B = 5/4, J_B = \emptyset$. A la troisième itération de la méthode de Newton, la condition (3.21) est satisfaite pour le vecteur paramètre

$$\theta^{(*)} = (\bar{z}_j, \tilde{\tau}_B = 1.23),$$

lequel est pris comme une solution approximative du système (3.19). Ce vecteur est utilisé pour améliorer la commande \hat{v} , et la trajectoire $\chi(t^*)$ et vérifier la satisfaction des contraintes et la condition ε -optimalité. On obtient,

$$\chi(t^*) = \begin{pmatrix} 7.4071 \\ 2.46 \end{pmatrix}, H\chi(t^*) = 2.4871.$$

Par conséquent :

$$J(v^0) = 2.46.$$

Avec la condition initiale :

$$x_0 = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

et la commande

$$u^0(t) = \begin{cases} 1, & \text{si } t \in [0, 1.23[, \\ -1, & \text{si } t \in [1.23, 2]. \end{cases}$$

La commande $v^0 = (z^0, u^0(t))$ est admissible, et elle est ε -optimale, le principe ε -optimalité est satisfait pour $\varepsilon = 0.2$.

Chapitre 4

Nouvelle approche de résolution d'un problème de contrôle optimal

4.1 Introduction

Dans ce chapitre, on résout le problème de contrôle optimal par une autre approche qui est un couplage de la méthode de discrétisation et de la procédure finale. La solution du problème discret permet de trouver un support qu'on utilisera dans la résolution du problème initial.

4.2 Position du problème

Dans l'intervalle de temps $T = [0, t^*]$, considérons un problème terminal de contrôle optimal :

$$J(u) = c'x(t^*) \rightarrow \max_u, \quad (4.1)$$

$$\dot{x} = Ax + bu, \quad x(0) = x_0, \quad (4.2)$$

$$Hx(t^*) = g, \quad (4.3)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T = [0, t^*]. \quad (4.4)$$

où $x = x(t) \in \mathbb{R}^n$, $t \in T$, est l'état du système (4.2) à l'instant t ; $u(\cdot) = (u(t), t \in T)$, $T = [0, t^*]$, la fonction continue par morceaux; $A \in \mathbb{R}^{n \times n}$; $b, c \in \mathbb{R}^n$; $H \in \mathbb{R}^{m \times n}$, $\text{rang}H = m \leq n$; $g \in \mathbb{R}^m$, f_*, f^* sont des scalaires; $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$, sont des ensembles d'indices, c' est le transposé

de c .

En utilisant la formule de Cauchy, On obtient la solution du système (4.2) :

$$x(t) = F(t)\left(x_0 + \int_0^t F^{-1}(\vartheta)bu(\vartheta)d\vartheta\right), \quad t \in T, \quad (4.5)$$

où $F(t) = e^{At}$, $t \in T = [0, t^*]$ est la solution (résolvante) du système :

$$\begin{cases} \dot{F}(t) = AF(t), \\ F(0) = Id. \end{cases}$$

En utilisant la formule (4.5), Le problème (4.1) – (4.4) devient un problème de variable $u(t)$, $t \in T$:

$$J(u) = \tilde{c}'x_0 + \int_0^{t^*} c(t)u(t)dt \longrightarrow \max_u, \quad (4.6)$$

$$D(I, J)x_0 + \int_0^{t^*} \varphi(t)u(t)dt = g, \quad (4.7)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T, \quad (4.8)$$

où $\tilde{c}' = c'F(t^*)$, $c(t) = c'F(t^*)F^{-1}(t)b$, $D(I, J) = HF(t^*)$, $\varphi(t) = HF(t^*)F^{-1}(t)b$.

4.3 Définitions Essentielles

Définition 4.1 La commande $u(t)$ est admissible si elle satisfait les contraintes (4.2)-(4.4).

Définition 4.2 Une commande admissible $u^0(\cdot)$ est optimale si elle réalise le maximum de la fonctionnelle :

$$J(u^0) = \max_u J(u).$$

Définition 4.3 Pour $\varepsilon \geq 0$, la commande $u^\varepsilon(\cdot)$ est dite ε -optimale (solution approchée) si

$$J(u^0) - J(u^\varepsilon) \leq \varepsilon.$$

4.4 Problème discret du problème initial

Dans l'intervalle T , choisissons un sous-ensemble $T_h = \{0, h, \dots, t^* - h\}$ formé de valeurs discrètes au temps, où $h = t^*/N$, N est un entier. Faisons une discrétisation de la fonction $u(t)$, $t \in T$:

$$u(t) \equiv u(\tau), \quad t \in [\tau, \tau + h), \tau \in T_h.$$

En utilisant cette discrétisation, le problème (4.6)-(4.8) devient :

$$\tilde{c}'x_0 + \sum_{t \in T_h} q(t)u(t) \rightarrow \max_u, \quad (4.9)$$

$$D(I, J)x_0 + \sum_{t \in T_h} d(t)u(t) = g, \quad (4.10)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T. \quad (4.11)$$

$d(t)$ est définie par l'expression suivante :

$$d(t) = \int_t^{t+h} \varphi(\vartheta) d\vartheta = \int_t^{t+h} \psi'(\vartheta) b(\vartheta) d\vartheta,$$

et $q(t)$ est égale à

$$q(t) = \int_t^{t+h} c(\vartheta) d\vartheta = \int_t^{t+h} \varrho(\vartheta) d\vartheta, \quad t \in T_h.$$

Ici $\psi(t)$, $t \in T$, est la solution du système adjoint (conjugué) :

$$\dot{\psi} = -A' \psi, \quad (4.12)$$

avec la condition initiale

$$\psi(t^*) = c;$$

et $\varrho(t)$, $t \in T$ est une matrice $m \times n$, solution de l'équation suivante :

$$\dot{\varrho} = -\varrho A, \quad (4.13)$$

avec la condition initiale

$$\varrho(t^*) = H.$$

Par suite on résout le problème linéaire (4.9)-(4.11) par la méthode adaptée. Choisissons un sous-ensemble arbitraire $T_B \subset T_h$ de m éléments et formons la matrice

$$P_B = (d(t), t \in T_B) \quad (4.14)$$

L'ensemble T_B est dit support du problème (4.1) – (4.4) si $\det P_B \neq 0$. La paire $\{v, S_B\}$ formée d'une commande admissible $u(\cdot)$ et du support T_B est dite support contrôle. Le support contrôle $\{u, T_B\}$ est dit non dégénérée si $f_* < u(t) < f^*, t \in T_B$.

Considérons une autre commande admissible $\bar{u}(t) = u(t) + \Delta u(t), t \in T$, et calculons l'accroissement de la fonctionnelle :

$$\Delta J(u) = J(\bar{u}) - J(u) = \sum_{t \in T_h} q(t) \Delta u(t).$$

De l'admissibilité de u et \bar{u} on a :

$$\sum_{t \in T_h} d(t) \Delta u(t) = 0,$$

et par conséquent l'accroissement de la fonctionnelle devient :

$$\Delta J(u) = \sum_{t \in T_h} (q(t) - \nu' d(t)) \Delta u(t),$$

où $\nu \in \mathbb{R}^m$ est la fonction des multiplicateurs de Lagrange appelée vecteur des potentiels, calculé comme solution de l'équation : $\nu' = q'_B Q$, où $Q = P_B^{-1}, q_B = (q(t), t \in T_B)$. Introduisons le n -vecteur des estimations $\Delta(t) = \nu' d(t) - q(t), t \in T_h$.

En utilisant ce vecteur, l'accroissement de la fonctionnelle prend la forme suivante :

$$\Delta J(u) = - \sum_{t \in T_h} \Delta(t) \Delta u(t). \tag{4.15}$$

Le support contrôle $\{u, T_B\}$ est non dégénérée pour le dual si $\Delta(t) \neq 0, t \in T_H$, où $T_H = T_h / T_B$.

4.5 La valeur de suboptimalité

Comme u et \bar{u} sont admissibles, on a alors :

$$f_* - u(t) \leq \Delta u(t) \leq f^* - u(t), t \in T. \tag{4.16}$$

Le maximum de la fonctionnelle (4.15) sous les contraintes (4.16) est atteint pour :

$$\left\{ \begin{array}{l} \Delta u(t) = f_* - u(t) \quad \text{si } \Delta(t) > 0 \\ \Delta u(t) = f^* - u(t) \quad \text{si } \Delta(t) < 0 \\ f_* \leq \Delta u(t) \leq f^*, \quad \text{si } \Delta(t) = 0, t \in T, \end{array} \right.$$

et est égal à :

$$\beta = \beta(u, T_B) = \sum_{t \in T_h^+} \Delta(t)(u(t) - f_*) + \sum_{t \in T_h^-} \Delta(t)(u(t) - f^*)$$

où

$$T_h^+ = \{t \in T_H, \Delta(t) > 0\}, \quad T_h^- = \{t \in T_H, \Delta(t) < 0\}.$$

Le nombre $\beta(v, S_B)$ est appelé valeur de suboptimalité du support contrôle $\{u, S_B\}$.

On en déduit que l'inégalité,

$$J(\bar{u}) - J(u) \leq \beta(u, T_B)$$

est toujours vérifiée, et pour $\bar{u} = u^0$, on a :

$$J(u^0) - J(u) \leq \beta(u, T_B).$$

4.6 Critère d'optimalité et $d'\varepsilon$ -optimalité

Théorème 13 [34] *Pour l'optimalité du support-contrôle $\{u, T_B\}$, les relations suivantes :*

$$\begin{cases} u(t) = f_*, & \text{si } \Delta(t) > 0, \\ u(t) = f^*, & \text{si } \Delta(t) < 0, \\ f_* \leq u(t) \leq f^*, & \text{si } \Delta(t) = 0, t \in T_H, \end{cases}$$

sont suffisantes, et dans le cas de la non dégénérescence, elles sont nécessaires.

Théorème 14 *Pour $\varepsilon \geq 0$ donné, la commande admissible u est ε -optimale si et seulement si il existe un tel support T_B tel que $\beta(u, T_B) \leq \varepsilon$.*

4.7 L'algorithme numérique pour la résolution du problème discret

Soit $\varepsilon > 0$ donné. Supposons que le support contrôle $\{u, T_B\}$ de départ ne vérifie pas l'optimalité et l' ε -optimalité. Pour cela on passe au changement du support contrôle : $\{u, T_B\} \rightarrow \{\bar{u}, \bar{T}_B\}$ pour lequel $\beta(\bar{u}, \bar{T}_B) \leq \beta(u, T_B)$. Cette itération est constitué de deux procédures :

1. Changement de commande $u \rightarrow \bar{u}$.
2. Changement de support $T_B \rightarrow \bar{T}_B$.

4.7.1 Changement de commande

Soit $\{u, T_B\}$ le support contrôle de départ et soit \bar{u} la nouvelle commande :

$$\bar{u}(t) = u(t) + \theta^0 l(t), \quad t \in T_h, \quad (4.17)$$

où $l = (l(t), t \in T_h)$ est une direction admissible de changement de commande u ; θ^0 est le pas maximum le long de cette direction.

Construction de la direction admissible

Sur les ensembles hors base on pose $\theta = 1$ et la direction sera construite de telle sorte que la nouvelle commande vérifie le critère d'optimalité.

$$l(t) = \begin{cases} f_* - u(t), & \Delta(t) \geq 0; \\ f^* - u(t), & \text{si } \Delta(t) \leq 0, \quad t \in T_H; \\ 0, & \text{si } \Delta(t) = 0. \end{cases}$$

Les valeurs de la direction correspondantes aux ensembles du support seront calculées à partir de l'admissibilité de u et \bar{u} :

$$(l(t), t \in T_B) = P_B^{-1} \cdot P_H \cdot (l(T_H)).$$

Construction du pas maximal

Le pas sera calculé de telle sorte que \bar{u} vérifie les contraintes directes sur les ensembles de base :

$$f_* \leq \bar{u}(t) \leq f^*, \quad t \in T_B,$$

c'est-à-dire

$$f_* \leq u(t) + \theta(t^0)l(t) \leq f^*, \quad t \in T_B.$$

où $\theta(t^0) = \min_{t \in T_B} \theta(t)$:

$$\theta(t) = \begin{cases} \frac{f^* - u(t)}{l(t)}, & \text{si } l(t) > 0 \\ \frac{f_* - u(t)}{l(t)}, & \text{si } l(t) < 0 \\ +\infty, & \text{si } l(t) = 0, \quad t \in T_B. \end{cases}$$

$\theta^0 = \min\{1; \theta(t_0)\}$. Calculons la valeur de suboptimalité du nouveau support contrôle $\{\bar{u}, T_B\}$, avec \bar{u} calculé en utilisant la relation (4.17) :

$$\beta(\bar{u}, T_B) = (1 - \theta^0)\beta(u, T_B).$$

Par conséquent

1. Si $\theta^0 = 1$, alors \bar{u} est une commande optimale.
2. Si $\beta(\bar{u}, T_B) \leq \varepsilon$, alors \bar{u} est une commande ε -optimale.
3. Si $\beta(\bar{u}, T_B) > \varepsilon$, alors on passe au changement du support.

4.7.2 Changement de support

De la procédure issue du changement de commande, nous avons $\theta^0 = \theta(t^0)$, $t^0 \in T_B$, $\beta(\bar{u}, T_B) > \varepsilon$. Alors le moment t^0 doit sortir de la base et être remplacé par un autre moment qui va permettre de diminuer la mesure de la non-optimalité de support β_B ; cela se fait par le changement support, $T_B \rightarrow \bar{T}_B$.

Construisons la pseudo-commande $(\tilde{u}(t), t \in T)$ comme suit :

$$\tilde{u}(t) = \begin{cases} f^* & \text{si } \Delta(t) \leq 0, \\ f_* & \text{si } \Delta(t) \geq 0 \\ \in [f_*, f^*], & \text{si } \Delta(t) = 0, t \in T_H. \end{cases}$$

La première étape consiste à déterminer les variations du vecteur des potentiels $\Delta\nu$, par la résolution d'un système d'équations linéaires :

$$-d(t)\Delta\nu = \delta(t), t \in T_B,$$

où $\delta(t) = \text{sign } \bar{u}(t^0)$ et $\delta(t) = 0$, $t \in T_B/t^0$.

Notons que $\delta(t)$, $t \in T_h$ est la variation de la co-commande engendrée par les variations $\Delta\nu$; elle est définie par :

$$\delta(t) = -\Delta\nu'd(t), t \in T_H.$$

Le nouveau vecteur des estimations est donné par :

$$\bar{\Delta}(t, \sigma) = \Delta(t) + \sigma\delta(t), t \in T_h, \sigma \geq 0.$$

Ainsi, le problème dual du problème (4.6) – (4.8) est donné comme suit : e

$$\begin{cases} L(\zeta, v, w) = g\zeta - \sum_{t \in T_h} v(t)f_*dt + \sum_{t \in T_h} w(t)f^*dt \rightarrow \min_{v, w}, \\ \zeta'd(t) - v(t) + w(t) = c(t), \\ v(t) \geq 0, w(t) \geq 0, t \in T, \end{cases}$$

L'ensemble $\varpi(\sigma) = (\nu(\sigma), v(t, \sigma), w(t, \sigma), t \in T_h)$ est donné par :

$$\begin{cases} \nu(\sigma) = \nu + \sigma\Delta\nu, \\ v(t, \sigma) = \Delta(t, \sigma), w(t, \sigma) = 0, & \text{si } \Delta(t, \sigma) \geq 0 \\ v(t, \sigma) = 0, w(t, \sigma) = -\Delta(t, \sigma), & \text{si } \Delta(t, \sigma) \leq 0, t \in T_h, \\ v(t^0, \sigma) = \Delta(t^0, \sigma) = \sigma, w(t^0, \sigma) = 0, & \text{pour } \sigma(t^0) = 1, (\text{i.e } l(t^0) < 0), \\ v(t^0, \sigma) = 0, w(t^0, \sigma) = -\Delta(t^0, \sigma) = -\sigma, & \text{pour } \sigma(t^0) = -1, (\text{i.e } l(t^0) > 0). \end{cases}$$

$$\begin{aligned}
L(\varpi(\sigma)) &= b'\nu(\sigma) - \sum_{t \in T_h} v(t, \sigma) f_* + \sum_{t \in T_h} w(t, \sigma) f^* \\
&= b'\nu(\sigma) - \sigma b' \Delta \nu - u(T_H) \Delta(T_H, \sigma) - \sigma u(T_H) \delta(T_H) \sigma \bar{u}(t^0) \delta(t^0) \\
&= L(\varpi) + \sigma (b' \Delta \nu - u(T_H) \delta(T_H) - \bar{u}(t^0) \delta(t^0)) \\
&= L(\varpi) + \sigma (b' \Delta \nu - u'(T_H) P'_H \Delta \nu - \bar{u}(t^0) \delta(t^0)) \\
&= L(\varpi) + \sigma ((b' - u'(T_H) P'_H) Q \delta(T_H) - \bar{u}(t^0) \delta(t^0)) \\
&= L(\varpi) + \sigma (u'(T_B) \delta(T_H) - \bar{u}(t^0) \delta(t^0)) \\
&= L(\varpi) + \sigma |\tilde{u}(t^0) - \bar{u}(t^0)|
\end{aligned} \tag{4.18}$$

On pose :

$$\alpha = -|\tilde{u}(t^0) - \bar{u}(t^0)| < 0 \tag{4.19}$$

Selon (4.18), (4.19), la fonction dual $L(\varpi(\sigma))$ diminue, jusqu'à ce que l'une de ses composantes de $\delta(t)$, $t \in T_H$ s'annule, et la valeur du pas maximal est égal à :

$$\sigma^0 = \min(\sigma_t^0),$$

où

$$\sigma_t^0 = \sigma(t_1) = \min \sigma(t), \quad t \in T_H,$$

où

$$\sigma(t) = \begin{cases} -\Delta(t)/\delta(t) & \text{si } \Delta(t)\delta(t) < 0, \\ +\infty & \text{si } \Delta(t)\delta(t) \geq 0, \quad t \in T_H. \end{cases}$$

Construction du nouveau support

Le nouveau support sera construit de la manière suivante :

$$\theta^0 = \theta(t^0), \quad \sigma^0 = \sigma(t_1).$$

Le nouveau support est

$$\bar{T}_B = T_B / \{t^0\} \cup \{t_1\}.$$

La valeur de suboptimalité $\beta(\bar{u}, \bar{T}_B)$ du nouveau support contrôle est égale à :

$$\beta(\bar{u}, \bar{T}_B) = (1 - \theta^0) \beta(u, T_B) - \alpha \sigma^0,$$

où

$$\alpha = |\tilde{u}(t^0) - \bar{u}(t^0)|, \quad \text{si } \theta^0 = \theta(t_0).$$

1. Si $\beta(\bar{u}, \bar{T}_B) = 0$, alors la commande \bar{u} est optimale pour le problème (4.9)-(4.11) .
2. Si $\beta(\bar{u}, \bar{T}_B) < \varepsilon$, alors la commande \bar{u} est ε -optimale pour le problème (4.9)-(4.11) .
3. Si $\beta(\bar{u}, \bar{T}_B) > \varepsilon$, alors on passe à une nouvelle itération avec le support contrôle $\{\bar{u}, \bar{T}_B\}$ ou à la procédure finale.

4.8 Procédure finale

En utilisant le support \overline{T}_B , on construit la quasi-commande $\widehat{u}(t), t \in T$, en utilisant

$$\widehat{u}(t) = \begin{cases} f_*, & \text{si } \Delta(t) < 0 \\ f^*, & \text{si } \Delta(t) > 0, t \in T_H \end{cases}$$

Si

$$\int_0^{t^*} \varphi(t) \widehat{u}(t) dt = g,$$

alors \widehat{u} est une commande optimale, et si

$$\int_0^{t^*} \varphi(t) \widehat{u}(t) dt \neq g, \quad (4.20)$$

alors notons par $T^0 = \{t_i, i = \overline{1, s}\}, s = |T_B|$.

Ici, $t_i, i = \overline{1, s}$ sont les points isolés optimaux de la co-commande $\Delta(t) = 0, t \in T; t_0 = 0, t_{s+1} = t^*$. Supposons que

$$\dot{\Delta}(t_i) \neq 0, i = \overline{1, s}.$$

De (4.20), on déduit la construction de la fonction suivante :

$$f(\Theta) = \sum_{i=0}^s \left(\frac{f^* + f_*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i) \right) \int_{t_i}^{t_{i+1}} \varphi(t) dt - g,$$

où

$$\Theta = (t_i, i = \overline{1, s}).$$

La procédure finale consiste à trouver la solution

$$\Theta^0 = (t_i^0, i = \overline{1, s})$$

du système de m équations non linéaire

$$f(\Theta) = 0. \quad (4.21)$$

On résout ce système par le méthode de Newton en utilisant l'approximation initiale :

$$\Theta^{(0)} = (\bar{t}_i, i = \overline{1, s}).$$

la $(k+1)^{\text{ème}}$ approximation $\Theta^{(k+1)}$, à l'étape $k+1 \geq 1$, est égale :

$$\Theta^{(k+1)} = \Theta^{(k)} + \Delta\Theta^{(k)}, \quad \Delta\Theta^{(k)} = -\frac{\partial f^{-1}(\Theta^{(k)})}{\partial \Theta^{(k)}} \cdot f(\Theta^{(k)}),$$

où

$$\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}} = (f_* - f^*) \text{sign} \dot{\Delta}(t_i^{(k)}) \varphi(t_i^{(k)}), \quad i = \overline{1, s}.$$

Comme $\det P_B \neq 0$, on peut montrer facilement que

$$\det \frac{\partial f(\Theta^{(0)})}{\partial \Theta^{(0)}} \neq 0. \quad (4.22)$$

Pour chaque instant $t_i \in T_B$, il existe un $\mu > 0$ petit tel que $\tilde{t}_i \in [t_i - \mu, t_i + \mu]$, $i = \overline{1, s}$, les matrices $(\varphi(\tilde{t}_i), i = \overline{1, s})$ et $\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}}$ sont non dégénérées. Si les éléments $t_i^{(k)}$, $i = \overline{1, s}$, $k = 1, 2, \dots$ sont au voisinage de μ , $i = \overline{1, s}$, le vecteur $\Theta^{(k^*)}$ est solution de l'équation (4.21) et assure que :

$$\| f(\Theta^{(k^*)}) \| \leq \eta,$$

pour $\eta > 0$ donné. Donc, posons $\theta^0 = \theta^{(k^*)}$. La commande suboptimale du problème (4.1)-(4.4) est calculé comme suit :

$$u^0(t) = \frac{f^* + f_*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i^0), \quad t \in [t_i^0, t_{i+1}^0[, \quad i = \overline{1, s}.$$

Si la méthode de Newton ne converge pas, on fait décroître le paramètre $h > 0$ et on passe à une autre itération.

Chapitre 5

Problème de contrôle optimal avec une entrée libre

5.1 Introduction

Dans ce chapitre, on résout le problème du chapitre 3 par une autre approche qui est un couplage de la méthode de discrétisation et de la procédure finale. La solution du problème discret nous permet de trouver un support qu'on utilisera dans la résolution du problème initial.

5.2 Position du problème

Dans l'intervalle de temps $T = [0, t^*]$, considérons un problème terminal de contrôle optimal avec une condition initiale libre :

$$J(z, u(t)) = c'x(t^*) \rightarrow \max_{z, u}, \quad (5.1)$$

$$\dot{x} = Ax + bu, x(0) = z \in X_0 = \{z \in \mathbb{R}^n, \quad Gz = \gamma, \quad d_* \leq z \leq d^*\}, \quad (5.2)$$

$$Hx(t^*) = g, \quad (5.3)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T = [0, t^*]. \quad (5.4)$$

où $x = x(t) \in \mathbb{R}^n$, $t \in T$, est l'état du système (5.2) à l'instant t ; $u(\cdot) = (u(t), t \in T)$, $T = [0, t^*]$, la fonction continue par morceaux; $A \in \mathbb{R}^{n \times n}$; $b, c \in \mathbb{R}^n$; $g \in \mathbb{R}^{m \times n}$, $\text{rang } H = m \leq n$; f_*, f^* sont des scalaires; $d_* = (d_{*j}, j \in J)$, $d^* = d^*(J) = (d_j^*, j \in J)$ sont des n -vecteurs; $G \in \mathbb{R}^{l \times n}$, $\text{rang } G = l \leq n$, $\gamma \in \mathbb{R}^l$, $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$, $L = \{1, \dots, l\}$ sont des ensembles

Chapitre 5. Problème de contrôle optimale avec une entrée libre80

d'indices, c' est le transposé de c .

En utilisant la formule de Cauchy, on obtient la solution du système (5.2) :

$$x(t) = F(t)\left(z + \int_0^t F^{-1}(\vartheta)bu(\vartheta)d\vartheta\right), t \in T, \quad (5.5)$$

où $F(t) = e^{At}$, $t \in T = [0, t^*]$ est la solution (résolvante) du système :

$$\begin{cases} \dot{F}(t) = AF(t), \\ F(0) = Id. \end{cases}$$

En utilisant la formule (5.5), Le problème (5.1) – (5.4) devient un problème des seules variables z et $u(t)$, $t \in T$:

$$J(z, u(t)) = \tilde{c}'z + \int_0^{t^*} c(t)u(t)dt \longrightarrow \max_{z, u}, \quad (5.6)$$

$$D(I, J)z + \int_0^{t^*} \varphi(t)u(t)dt = g, \quad (5.7)$$

$$G(L, J)z = \gamma, \quad d_* \leq z \leq d^*, \quad (5.8)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T, \quad (5.9)$$

où $\tilde{c}' = c'F(t^*)$, $c(t) = c'F(t^*)F^{-1}(t)b$, $D(I, J) = HF(t^*)$, $\varphi(t) = HF(t^*)F^{-1}(t)b$.

5.3 Définitions Essentielles :

Définition 5.1 La paire $v = (z, u(\cdot))$ formée d'un n -vecteur z et de la fonction continue par morceaux $u(\cdot)$ est appelé commande généralisée .

Définition 5.2 La commande généralisée $v = (z, u(\cdot))$ est admissible si elle satisfait les contraintes (5.2)-(5.4).

Définition 5.3 Une commande admissible $v^0 = (z^0, u^0(\cdot))$ est optimale si elle réalise le maximum de la fonctionnelle :

$$J(v^0) = \max_v J(v).$$

Définition 5.4 Pour $\varepsilon \geq 0$, la commande $v^\varepsilon = (z^\varepsilon, u^\varepsilon(\cdot))$ est dite ε -optimale (solution approchée) si

$$J(v^0) - J(v^\varepsilon) \leq \varepsilon.$$

5.4 Problème discret du problème initial

Dans l'intervalle T , choisissons un sous-ensemble $T_h = \{0, h, \dots, t^* - h\}$ formé d'instants discrets, où $h = t^*/N$, N est un entier. Faisons une discrétisation de la fonction $u(t)$, $t \in T$:

$$u(t) \equiv u(\tau), \quad t \in [\tau, \tau + h), \tau \in T_h.$$

En utilisant cette discrétisation, le problème (5.6)-(5.9) devient :

$$\tilde{c}'z + \sum_{t \in T_h} q(t)u(t) \rightarrow \max_{z, u}, \quad (5.10)$$

$$D(I, J)z + \sum_{t \in T_h} d(t)u(t) = g, \quad (5.11)$$

$$G(L, J)z = \gamma, \quad d_* \leq z \leq d^*, \quad (5.12)$$

$$f_* \leq u(t) \leq f^*, \quad t \in T. \quad (5.13)$$

$d(t)$ est définie par l'expression suivante :

$$d(t) = \int_t^{t+h} \varphi(\vartheta) d\vartheta = \int_t^{t+h} \psi'(\vartheta) b(\vartheta) d\vartheta,$$

et $q(t)$ est égale à

$$q(t) = \int_t^{t+h} c(\vartheta) d\vartheta = \int_t^{t+h} \varrho(\vartheta) d\vartheta, \quad t \in T_h.$$

Ici $\psi(t)$, $t \in T$, est la solution du système adjoint (conjugué) :

$$\dot{\psi} = -A' \psi, \quad (5.14)$$

avec la condition initiale

$$\psi(t^*) = c;$$

et $\varrho(t)$, $t \in T$ est une matrice $m \times n$, solution de l'équation suivante :

$$\dot{\varrho} = -\varrho A, \quad (5.15)$$

avec la condition initiale

$$\varrho(t^*) = H.$$

Par suite on résout le problème linéaire (5.10)-(5.13) par la méthode adaptée. Choisissons un sous-ensemble arbitraire $T_B \subset T_h$ de $k \leq m$ éléments

Chapitre 5. Problème de contrôle optimale avec une entrée libre 82

et un sous-ensemble arbitraire $J_B \subset J$ de $m + l - k$ éléments et formons la matrice

$$P_B = \begin{pmatrix} D(I, J_B) & d(t), t \in T_B \\ G(L, J_B) & 0 \end{pmatrix} \quad (5.16)$$

L'ensemble $S_B = \{T_B, J_B\}$ est dit support du problème (5.1) – (5.4) si $\det P_B \neq 0$.

La paire $\{v, S_B\}$ formée d'une commande admissible $v = (z, u(\cdot))$ et du support S_B est dite support contrôle. Le support contrôle $\{v, S_B\}$ est dit non dégénéré si $d_{*j} < z_j < d_j^*, j \in J_B, f_* < u(t) < f^*, t \in T_B$.

Considérons une autre commande admissible $\bar{v} = (\bar{z}, \bar{u}(\cdot)) = v + \Delta v$, où $\bar{z} = z + \Delta z, \bar{u}(t) = u(t) + \Delta u(t), t \in T$, et calculons l'accroissement de la fonctionnelle :

$$\Delta J(v) = J(\bar{v}) - J(v) = \bar{c}' \Delta z + \sum_{t \in T_h} q(t) \Delta u(t).$$

De l'admissibilité de v et \bar{v} on a :

$$D(I, J) \Delta z + \sum_{t \in T_h} d(t) \Delta u(t) = 0,$$

et

$$G(L, J) \Delta z = 0,$$

et par conséquent l'accroissement de la fonctionnelle devient :

$$\Delta J(v) = (\bar{c}' - \nu' \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix}) \Delta z + \sum_{t \in T_h} (q(t) - \nu' d(t)) \Delta u(t),$$

où $\nu = \begin{pmatrix} \nu_u \\ \nu_z \end{pmatrix} \in \mathbb{R}^{m+l}$, $\nu_u \in \mathbb{R}^m$, $\nu_z \in \mathbb{R}^l$ sont les multiplicateurs de Lagrange et le vecteur ν est le vecteur des potentiels, calculé comme solution de l'équation : $\nu' = q'_B Q$, où $Q = P_B^{-1}, q_B = (\tilde{c}_j, j \in J_B, q(t), t \in T_B)$.

Introduisons le n -vecteur des estimations $\Delta' = \nu' \begin{pmatrix} D(I, J) \\ G(L, J) \end{pmatrix} - \bar{c}'$, et la fonction co-commande $\Delta(\cdot) = (\Delta(t) = \nu'_u d(t) - q(t), t \in T_h)$.

En utilisant ce vecteur, l'accroissement de la fonctionnelle prend la forme suivante :

$$\Delta J(v) = \Delta' \Delta z - \sum_{t \in T_h} \Delta(t) \Delta u(t). \quad (5.17)$$

Le support contrôle $\{v, S_B\}$ est non dégénérée pour le duale si $\Delta(t) \neq 0, t \in T_H, \Delta_j \neq 0, j \in J_H$, où $T_H = T_h / T_B, J_H = J / J_B$.

5.5 La valeur de suboptimalité

Comme v et \bar{v} sont admissibles, alors on a :

$$d_* - z \leq \Delta z \leq d^* - z; \quad f_* - u(t) \leq \Delta u(t) \leq f^* - u(t), \quad t \in T. \quad (5.18)$$

Le maximum de la fonctionnelle (5.17) sous les contraintes (5.18) est atteint pour :

$$\left\{ \begin{array}{ll} \Delta z_j = d_{*j} - z_j & \text{si } \Delta_j > 0 \\ \Delta z_j = d_j^* - z_j & \text{si } \Delta_j < 0 \\ d_{*j} - z_j \leq \Delta z_j \leq d_j^* - z_j, & \text{si } \Delta_j = 0, \quad j \in J. \end{array} \right.$$

$$\left\{ \begin{array}{ll} \Delta u(t) = f_* - u(t) & \text{si } \Delta(t) > 0 \\ \Delta u(t) = f^* - u(t) & \text{si } \Delta(t) < 0 \\ f_* \leq \Delta u(t) \leq f^*, & \text{si } \Delta(t) = 0, \quad t \in T, \end{array} \right.$$

et est égal à :

$$\begin{aligned} \beta = \beta(v, S_B) &= \sum_{j \in J_H^+} \Delta_j (z_j - d_{*j}) + \sum_{j \in J_H^-} \Delta_j (z_j - d_j^*) \\ &+ \sum_{t \in T_h^+} \Delta(t) (u(t) - f_*) + \sum_{t \in T_h^-} \Delta(t) (u(t) - f^*) \end{aligned}$$

où

$$\begin{aligned} T_h^+ &= \{t \in T_H, \Delta(t) > 0\}, \quad T_h^- = \{t \in T_H, \Delta(t) < 0\}, \\ J_H^+ &= \{j \in J_H, \Delta_j > 0\}, \quad J_H^- = \{j \in J_H, \Delta_j < 0\}. \end{aligned}$$

Le nombre $\beta(v, S_B)$ est appelé valeur de suboptimalité du support contrôle $\{v, S_B\}$.

On en déduit que l'inégalité,

$$J(\bar{v}) - J(v) \leq \beta(v, S_B)$$

est toujours vérifiée, et pour $\bar{v} = v^0$, on a :

$$J(v^0) - J(v) \leq \beta(v, S_B).$$

5.6 Critère d'optimalité et ε -optimalité

Théorème 15 [34] *Pour l'optimalité du support-contrôle $\{v, S_B\}$, les relations suivantes :*

$$\left\{ \begin{array}{ll} u(t) = f_*, & \text{si } \Delta(t) > 0, \\ u(t) = f^*, & \text{si } \Delta(t) < 0, \\ f_* \leq u(t) \leq f^*, & \text{si } \Delta(t) = 0, t \in T_H, \\ z_j = d_{*j}, & \text{si } \Delta_j > 0, \\ z_j = d_j^*, & \text{si } \Delta_j < 0, \\ d_{*j} \leq z_j \leq d_j^*, & \text{si } \Delta_j = 0, j \in J_H, \end{array} \right.$$

sont suffisantes, et dans le cas de la non dégénérescence, elles sont nécessaires.

Théorème 16 *Pour $\varepsilon \geq 0$ donné, la commande admissible v est ε -optimale si et seulement si il existe un tel support S_B tel que $\beta(v, S_B) \leq \varepsilon$.*

5.7 L'algorithme numérique pour la résolution du problème discret

Soit $\varepsilon > 0$ donné. Supposons que le support contrôle $\{v, S_B\}$ de départ ne vérifie pas l'optimalité et l' ε -optimalité. Pour cela on passe au changement du support contrôle : $\{v, S_B\} \rightarrow \{\bar{v}, \bar{S}_B\}$ pour lequel $\beta(\bar{v}, \bar{S}_B) \leq \beta(v, S_B)$. Cette itération est constituée de deux procédures :

1. Changement de commande $v \rightarrow \bar{v}$.
2. Changement de support $S_B \rightarrow \bar{S}_B$.

5.7.1 Changement de commande

Soit $\{v, S_B\}$ le support contrôle de départ et soit $\bar{v} = (\bar{z}, \bar{u})$ la nouvelle commande :

$$\left\{ \begin{array}{ll} \bar{z}_j = z_j + \theta^0 l_j, & j \in J, \\ \bar{u}(t) = u(t) + \theta^0 l(t), & t \in T_h, \end{array} \right. \quad (5.19)$$

où $l = (l_j, j \in J, l(t), t \in T_h)$ est une direction admissible de changement de commande v ; θ^0 est le pas maximum le long de cette direction.

Chapitre 5. Problème de contrôle optimale avec une entrée libre 85

Construction de la direction admissible

Sur les ensembles hors base on pose $\theta = 1$ et la direction sera construite de telle sorte que la nouvelle commande vérifie le critère d'optimalité.

$$l_j = \begin{cases} d_{*j} - z_j, & \Delta_j > 0; \\ d_j^* - z_j, & \text{si } \Delta_j < 0, j \in J_H; \\ 0, & \text{si } \Delta_j = 0. \end{cases}$$

$$l(t) = \begin{cases} f_* - u(t), & \Delta(t) > 0; \\ f^* - u(t), & \text{si } \Delta(t) < 0, t \in T_H; \\ 0, & \text{si } \Delta(t) = 0. \end{cases}$$

Les valeurs de la direction correspondantes aux ensembles du support seront calculées à partir de l'admissibilité de v et \bar{v} :

$$(l_j, j \in J_B, l(t), t \in T_B) = P_B^{-1} \cdot P_H \cdot (l(J_H), l(T_H))'.$$

Construction du pas maximal

Le pas sera calculé de telle sorte que \bar{v} vérifie les contraintes directes sur les ensembles de base :

$$d_{j*} \leq \bar{z} \leq d_j^*, j \in J_B; \quad f_* \leq \bar{u}(t) \leq f^*, t \in T_B,$$

c'est-à-dire

$$d_* \leq z_j + \theta_{j0} l_j \leq d^*, j \in J_B;$$

$$f_* \leq u(t) + \theta(t_0) l(t) \leq f^*, t \in T_B.$$

où

$$\theta_{j0} = \min \theta_j :$$

$$\theta_j = \begin{cases} \frac{d_j^* - z_j}{l_j}, & \text{si } l_j > 0 \\ \frac{d_{*j} - z_j}{l_j}, & \text{si } l_j < 0 \\ +\infty, & \text{si } l_j = 0, j \in J_B. \end{cases}$$

$$\theta(t_0) = \min_{t \in T_B} \theta(t) :$$

$$\theta(t) = \begin{cases} \frac{f^* - u(t)}{l(t)}, & \text{si } l(t) > 0 \\ \frac{f_* - u(t)}{l(t)}, & \text{si } l(t) < 0 \\ +\infty, & \text{si } l(t) = 0, t \in T_B. \end{cases}$$

$\theta^0 = \min\{1; \theta(t^0); \theta_{j_0}\}$. Calculons la valeur de suboptimalité du nouveau support contrôle $\{\bar{v}, S_B\}$, avec \bar{v} calculé en utilisant la relation (5.19) :

$$\beta(\bar{v}, S_B) = (1 - \theta^0)\beta(v, S_B).$$

Par conséquent

1. Si $\theta^0 = 1$, alors \bar{v} est une commande optimale.
2. Si $\beta(\bar{v}, S_B) \leq \varepsilon$, alors \bar{v} est une commande ε -optimale.
3. Si $\beta(\bar{v}, S_B) > \varepsilon$, alors on passe au changement du support.

5.7.2 Changement de support

Le changement de support $S_B \rightarrow \bar{S}_B$ se fera de sorte que $\beta(\bar{v}, \bar{S}_B) < \beta(\bar{v}, S_B)$. Introduisons la pseudo-commande $\tilde{v} = (\tilde{z}, \tilde{u}(t), t \in T)$.

$$\tilde{z}_j = \begin{cases} d_{j*}, & \text{si } \Delta_j \geq 0 \\ d_j^*, & \text{si } \Delta_j \leq 0, \\ \in [d_{j*}, d_j^*], & j \in J_H; \end{cases}$$

$$\tilde{u}(t) = \begin{cases} f^*, & \text{si } \Delta(t) \leq 0, \\ f_*, & \text{si } \Delta(t) \geq 0, \\ \in [f_*, f^*], & t \in T_H. \end{cases}$$

Ici, on a $\theta^0 = \min(\theta(t^0), t^0 \in T_B; \theta_{j_0}, j_0 \in J_B)$.

Par suite, on distingue deux cas possibles :

- (a) $\theta^0 = \theta_{j_0}, j_0 \in J_B$.
- (b) $\theta^0 = \theta(t^0), t^0 \in T_B$.

Le changement du support entraîne le changement des vecteurs des potentiels et le vecteur des estimations :

$$\nu' = \nu + \sigma^0 \Delta \nu; \bar{\Delta}_j = \Delta_j + \sigma^0 \delta_j, j \in J; \bar{\Delta}(t) = \Delta(t) + \sigma^0 \delta(t), t \in T_h, \quad (5.20)$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 87

où $(\delta_j, j \in J, \delta(t), t \in T_h)$ est une direction admissible du changement du support, σ^0 est la pas maximal le long de cette direction et $\Delta\nu$ est la variation du vecteur des potentiels.

Ceci permet de diminuer la fonctionnelle du dual. Définissons le problème dual du problème (5.6) – (5.9) est le suivant :

$$\left\{ \begin{array}{l} L(\varsigma, \tilde{\eta}, \tilde{\varrho}, \iota, v, w) = \varsigma'g + \gamma'\tilde{\eta} - \tilde{\varrho}'d_* + \iota'd^* - \sum_{t \in T_h} v(t)f_*dt + \sum_{t \in T_h} w(t)f^*dt \rightarrow \min, \\ \varsigma'D(I, J) + \tilde{\eta}'G(L, J) - \tilde{\varrho} + \iota = \tilde{c}, \\ \varsigma'\varphi(t) - v(t) + w(t) = c(t), \\ \tilde{\eta} \geq 0, \tilde{\varrho} \geq 0, \iota \geq 0, v(t) \geq 0, w(t) \geq 0, t \in T, \end{array} \right.$$

Ici les fonctions $\tilde{\eta} \geq 0, \tilde{\varrho} \geq 0, \iota \geq 0$ et $v(t), w(t), t \in T$ pour $\nu(I) = \varsigma(I)$ sont définies de la manière suivante :

$$\left\{ \begin{array}{ll} \varrho_j = \Delta_j, \iota = 0, & \text{si } \Delta_j \geq 0 \\ \varrho_j = 0, \iota = -\Delta_j, & \Delta_j \leq 0, j \in J \\ v(t) = \Delta(t), w(t) = 0, & \text{si } \Delta(t) \geq 0 \\ v(t) = 0, w(t) = -\Delta(t), & \text{si } \Delta(t) \leq 0, t \in T \end{array} \right. \quad (5.21)$$

forment un plan dual du problème (solution admissible du dual).

Désignons par $(\varsigma^0, \tilde{\eta}^0, \tilde{\varrho}^0, \iota^0, v^0(t), w^0(t))$ sa solution optimale.

Construction de la direction admissible $(\delta_j, j \in J, \delta(t), t \in T_h)$

On construit premièrement les composantes de la direction correspondant au support S_B , suivant les deux cas :

Cas a) $\theta^0 = \theta_{j_0}$. Posons :

$$\left\{ \begin{array}{ll} \delta(t) = 0 & \text{si } t \in T_B \\ \delta_j = 0 & \text{si } j \neq j_0, j \in J_B \\ \delta_{j_0} = 1 & \text{si } \bar{z}_{j_0} = d_{*j_0} \\ \delta_{j_0} = -1 & \text{si } \bar{z}_{j_0} = d_{j_0}^*. \end{array} \right.$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 88

Cas b) $\theta^0 = \theta(t_0)$. Posons :

$$\left\{ \begin{array}{ll} \delta_j = 0 & \text{si } j \in J_B \\ \delta(t) = 0 & \text{si } t \in T_B/t^0 \\ \delta(t^0) = 1 & \text{si } \bar{u}(t^0) = f_* \\ \delta(t^0) = -1 & \text{si } \bar{u}(t^0) = f^*. \end{array} \right.$$

En utilisant les valeurs de δ_B , on calcule les valeurs composantes de la direction correspondants aux indices hors base

$$(\delta_j, j \in J_H) = \Delta \nu' \left(\begin{array}{c} D(I, j) \\ G(L, j) \end{array} \right),$$

$$(\delta(t), t \in T_H) = \Delta \nu'_u(d(t), t \in T_H).$$

Construction du pas maximal σ^0

Le pas maximal est égal à : $\sigma^0 = \min(\sigma_j^0, \sigma_t^0)$, où

$$\sigma_j^0 = \sigma_{j_1} = \min \sigma_j, j \in J_H; \sigma_t^0 = \sigma(t_1) = \min \sigma(t), t \in T_H,$$

où

$$\sigma_j = \begin{cases} -\Delta_j/\delta_j & \text{si } \Delta_j \delta_j < 0, \\ +\infty & \text{si } \Delta_j \delta_j \geq 0, \quad j \in J_H, \end{cases}$$

et

$$\sigma(t) = \begin{cases} -\Delta(t)/\delta(t) & \text{si } \Delta(t)\delta(t) < 0, \\ +\infty & \text{si } \Delta(t)\delta(t) \geq 0, \quad t \in T_H. \end{cases}$$

Construction du nouveau support

Le nouveau support sera construit suivant les différents cas possibles :

1. $\theta^0 = \theta(t^0)$, $\sigma^0 = \sigma(t_1)$.

La nouveau support est $\bar{S}_B = \{\bar{T}_B, \bar{J}_B\}$, où :

$$\bar{T}_B = T_B/\{t^0\} \cup \{t_1\}, \bar{J}_B = J_B.$$

2. $\theta^0 = \theta(t^0)$, $\sigma^0 = \sigma_{j_1}$.

Le nouveau support est $\bar{S}_B = \{\bar{T}_B, \bar{J}_B\}$, où :

$$\bar{T}_B = T_B/\{t^0\}, \bar{J}_B = J_B \cup \{j_1\}.$$

3. $\theta^0 = \theta_{j_0}$, $\sigma^0 = \sigma_{j_1}$.

Le nouveau support est $\bar{S}_B = \{\bar{T}_B, \bar{J}_B\}$, où :

$$\bar{T}_B = T_B, \bar{J}_B = J_B/\{j_0\} \cup \{j_1\}.$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 89

4. $\theta^0 = \theta_{j_0}$, $\sigma^0 = \sigma(t_1)$.

Le nouveau support est $\bar{S}_B = \{\bar{T}_B, \bar{J}_B\}$ où :

$$\bar{T}_B = T_B \cup \{t_1\}, \bar{J}_B = J_B / \{j_0\}.$$

La valeur de suboptimalité $\beta(\bar{v}, \bar{S}_B)$ du nouveau support contrôle est égale à :

$$\beta(\bar{v}, \bar{S}_B) = (1 - \theta^0)\beta(v, S_B) - \alpha\sigma^0,$$

où

$$\alpha = \begin{cases} |\tilde{z}_{j_0} - \bar{z}_{j_0}|, & \text{si } \theta^0 = \theta_{j_0}, \\ |\tilde{u}(t^0) - \bar{u}(t^0)|, & \text{si } \theta^0 = \theta(t^0). \end{cases}$$

1. Si $\beta(\bar{v}, \bar{S}_B) = 0$, alors la commande \bar{v} est optimale pour le problème (5.10)-(5.13) .
2. Si $\beta(\bar{v}, \bar{S}_B) < \varepsilon$, alors la commande \bar{v} est ε -optimale pour le problème (5.10)-(5.13) .
3. Si $\beta(\bar{v}, \bar{S}_B) > \varepsilon$, alors on passe à une nouvelle itération avec le support contrôle $\{\bar{v}, \bar{S}_B\}$ ou à la procédure finale.

5.8 Procédure finale

En utilisant le support \bar{S}_B , on construit la quasi-commande $\hat{v} = (\hat{z}, \hat{u}(t), t \in T)$, en utilisant

$$\hat{z}_j = \begin{cases} d_{j^*} & \text{si } \Delta_j > 0 \\ d_j^* & \text{si } \Delta_j < 0, j \in J_H, \end{cases}$$

$$\hat{u}(t) = \begin{cases} f_*, & \text{si } \Delta(t) < 0 \\ f^*, & \text{si } \Delta(t) > 0, t \in T_H. \end{cases}$$

Si

$$D(I, J)\hat{z} + \int_0^{t^*} \varphi(t)\hat{u}(t)dt = g,$$

$$G(L, J)\hat{z} = \gamma,$$

alors \hat{v} est une commande optimale, et si

$$D(I, J)\hat{z} + \int_0^{t^*} \varphi(t)\hat{u}(t)dt \neq g, \quad G(L, J)\hat{z} \neq \gamma, \quad (5.22)$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 90

alors notons par $T^0 = \{t_i, i = \overline{1, s}\}$, $s = |T_B|$.

Ici, $t_i, i = \overline{1, s}$ sont les points isolés de la co-commande $\Delta(t) = 0, t \in T$; $t_0 = 0, t_{s+1} = t^*$. Supposons que

$$\dot{\Delta}(t_i) \neq 0, i = \overline{1, s}.$$

De (7.33), on déduit la construction de la fonction suivante :

$$f(\Theta) = \begin{pmatrix} D(I, J_B)z(J_B) + D(I, J_H)z(J_H) + \sum_{i=0}^s \left(\frac{f_*^* + f^*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i) \right) \int_{t_i}^{t_{i+1}} \varphi(t) dt - g \\ G(L, J_B)z(J_B) + G(L, J_H)z(J_H) - \gamma \end{pmatrix}$$

où

$$z_j = \frac{d_j^* + d_{j^*}}{2} - \frac{d_j^* - d_{j^*}}{2} \text{sign} \Delta_j, j \in J_H.$$

$$\Theta = (t_i, i = \overline{1, s}; z_j, j \in J_B).$$

La procédure finale consiste à trouver la solution

$$\Theta^0 = (t_i^0, i = \overline{1, s}; z_j^0, j \in J_B)$$

du système de $m + l$ équations non linéaire

$$f(\Theta) = 0. \quad (5.23)$$

On résout ce système par le méthode de Newton en utilisant l'approximation initiale :

$$\Theta^{(0)} = (\bar{t}_i, i = \overline{1, s}; \bar{z}_j, j \in J_B).$$

la $(k + 1)^{\text{ème}}$ approximation $\Theta^{(k+1)}$, à l'étape $k + 1 \geq 1$, est égale :

$$\Theta^{(k+1)} = \Theta^{(k)} + \Delta\Theta^{(k)} \quad \Delta\Theta^{(k)} = -\frac{\partial f^{-1}(\Theta^{(k)})}{\partial \Theta^{(k)}} \cdot f(\Theta^{(k)}),$$

où

$$\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}} = \begin{pmatrix} D(I, J_B) & (f_* - f^*) \text{sign} \dot{\Delta}(t_i^{(k)}) \varphi(t_i^{(k)}), i = \overline{1, s} \\ G(L, J_B) & 0 \end{pmatrix}.$$

Comme $\det P_B \neq 0$, on peut montrer facilement que

$$\det \frac{\partial f(\Theta^{(0)})}{\partial \Theta^{(0)}} \neq 0. \quad (5.24)$$

Pour chaque instant $t_i \in T_B$, il existe un $\mu > 0$ petit tel que $\tilde{t}_i \in [t_i - \mu, t_i + \mu], i = \overline{1, s}$, les matrices $(\varphi(\tilde{t}_i), i = \overline{1, s})$ et $\frac{\partial f(\Theta^{(k)})}{\partial \Theta^{(k)}}$ sont non dégénérées. Si les

Chapitre 5. Problème de contrôle optimale avec une entrée libre 91

éléments $t_i^{(k)}$, $i = \overline{1, s}$, $k = 1, 2, \dots$ sont au voisinage de μ , $i = \overline{1, s}$, le vecteur $\Theta^{(k^*)}$ est solution de l'équation (5.23) et assure que :

$$\|f(\Theta^{(k^*)})\| \leq \eta,$$

pour $\eta > 0$ donné.

Donc, posons $\theta^0 = \theta^{(k^*)}$. La commande suboptimale du problème (5.1)-(5.4) est calculé comme suit :

$$z_j^0 = \begin{cases} z_j^0, & j \in J_B \\ \widehat{z}_j, & j \in J_H; \end{cases}$$

$$u^0(t) = \frac{f^* + f_*}{2} - \frac{f^* - f_*}{2} \text{sign} \dot{\Delta}(t_i^0), \quad t \in [t_i^0, t_{i+1}^0[, \quad i = \overline{1, s}.$$

Si la méthode de Newton ne converge pas, on décroît le paramètre $h > 0$ et on passe à une autre itération.

5.9 Exemple numérique

Illustrons les résultats obtenus par l'exemple suivant :

$$\begin{aligned} & \int_0^{25} u(t) dt \rightarrow \min, \\ \dot{x}_1 &= x_3, \\ \dot{x}_2 &= x_4, \\ \dot{x}_3 &= -x_1 + x_2 + u, \\ \dot{x}_4 &= 0.1x_1 - 1.01x_2, \\ x_1(25) &= x_2(25) = x_3(25) = x_4(25) = 0, \\ 0 &\leq u(t) \leq 1, \quad t \in [0, 25]. \end{aligned} \tag{5.25}$$

On définit les matrices et les vecteurs suivants :

$$\begin{aligned} A &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0.1 & -1.01 & 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ g &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0.1 \\ 0.25 \\ 2 \\ 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad d_* = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad d^* = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}. \end{aligned}$$

Le problème (5.25) est déduit à la forme canonique de (5.1) – (5.4) en introduisant une nouvelle variable $\dot{x}_5 = u$, $x_5(0) = 0$. La fonctionnelle prend alors la forme

$$-x_5(t^*) \rightarrow \max.$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 92

Dans la classe des commandes discrètes avec une période $h = 25/1000 = 0.025$, le problème (5.25) est équivalent à un problème de programmation linéaire de dimension 4×1000 .

Soit $T_B = \{5, 10, 15, 20\}$ le support initial. Le problème est résolu en 18 itérations, il consiste à construire la commande optimale, la matrice du support 4×4 qui change 18 fois. La valeur optimale du critère de qualité est égale à 6.602499 et le temps d'exécution est de 2.30 secondes.

Le mouvement de la co-commande $\Delta(t)$ au cours des itérations est donné dans la figure (5.1).

A notre avis, le temps d'exécution pour construire une commande op-

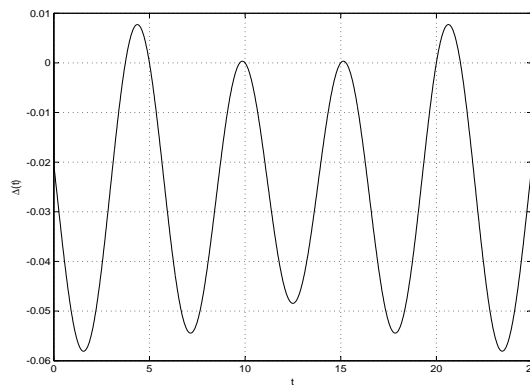


FIGURE 5.1 –

timale n'a pas beaucoup d'importance, il est seulement important que la méthode soit capable de construire la solution réalisable en un temps raisonnable.

Il est important de définir l'efficacité de la méthode en utilisant le nombre d'intégration du problème primal ou le système adjoint. A ce stade, la complexité du temps d'intégration du problème primal ou du système adjoint est pris le long de l'intervalle T . L'algorithme est implémenté par le logiciel Matlab.

Le tableau suivant contient la valeur du critère de qualité du problème (5.25) dans différentes périodes.

h	Le nombre d'itérations	la valeur du critère	le temps d'exécution
0.25	11	6.6243433	2.27
0.0025	26	6.602054	2.30
0.001	32	6.602050	2.69

En faites, le problème (5.25) peut être résolu par la méthode de programmation linéaire en transformant le problème (5.10)-(5.13). Donc, une seule intégration du système suffit pour former la matrice du problème de

Chapitre 5. Problème de contrôle optimale avec une entrée libre 93

programmation linéaire (5.1)-(5.4).

Dans la figure (5.2), on représente la commande $u^*(\tau), \tau \in T_h$. Dans la figure (5.3), on représente la trajectoire x_1 en fonction de x_3 du système (5.25). Dans la figure (5.4), on représente la trajectoire x_2 en fonction de x_4 du système (5.25).

La condition initiale optimale est donnée par

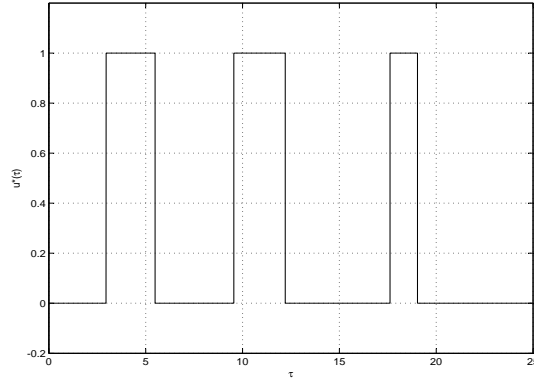


FIGURE 5.2 –

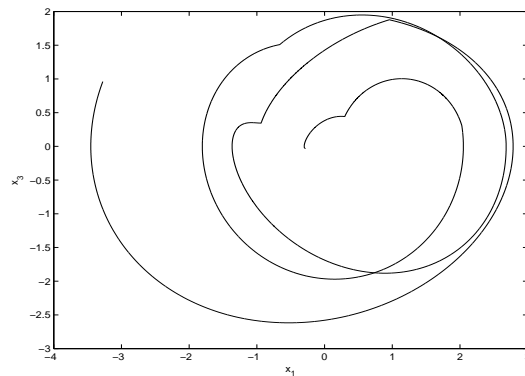


FIGURE 5.3 –

$$x_1(0) = 0.1009729, x_2(0) = 0.2502507, x_3(0) = 0.9933905, x_4(0) = 1.0010008.$$

Chapitre 5. Problème de contrôle optimale avec une entrée libre 94

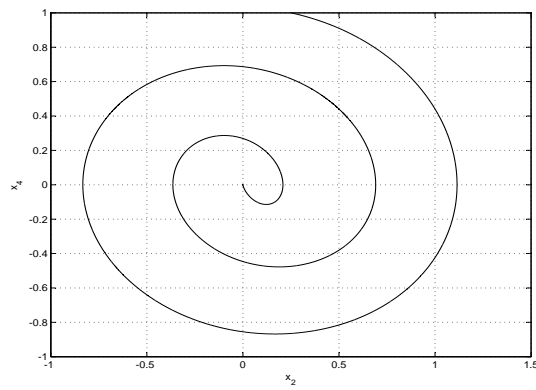


FIGURE 5.4 –

Chapitre 6

Méthode de tir et méthode de Newton discrète

6.1 Introduction

La résolution numérique de problèmes de commande optimale, nécessite dans certains cas, l'utilisation de la méthode de tir [88]. Cet algorithme, utilise la méthode de Newton et, on sait, par expérience que la méthode de tir nécessite, l'utilisation de la méthode de différenciation numérique; les résultats sont convenables à condition toutefois que le pas de discrétisation lors de la différenciation approchée soit petit.

Dans le paragraphe suivant, nous exposons une étude bibliographique concernant la convergence de la méthode de Newton liée à la seule condition que le pas de discrétisation lors de différenciation numérique soit petit.

6.2 La méthode de tir

Le principe de la méthode de tir est le suivant. Considérons le problème de contrôle optimal sous forme de Bolza, et supposons que le temps final t^* fixé. Le principe du maximum donne une condition nécessaire d'optimalité et affirme que toute trajectoire optimale est la projection d'une extrémale. Si l'on est capable, à partir de la condition de maximum, d'exprimer le contrôle extrémal en fonction de $(x(t), \psi(t))$, alors le système extrémal est un système différentiel de la forme $\dot{\vartheta}(t) = V(t, \vartheta(t))$, où $\vartheta(t) = (x(t), \psi(t))$, et les conditions initiales, finales, et les conditions de transversalité, se mettent sous la forme $R(\vartheta(0), \vartheta(t^*)) = 0$. Finalement, on obtient le problème aux valeurs limites

$$\begin{cases} \dot{\vartheta}(t) = V(t, \vartheta(t)), \\ R(\vartheta(0), \vartheta(t^*)) = 0, \end{cases} \quad (6.1)$$

Notons par $\vartheta(t, \vartheta_0)$ la solution du problème de Cauchy :

$$\dot{\vartheta}(t) = V(t, \vartheta(t)), \quad \vartheta(0) = \vartheta_0,$$

Le problème (6.1) aux valeurs limites est alors équivalent à

$$R(\vartheta_0, \vartheta(t^*, \vartheta_0)) = 0$$

il s'agit de *déterminer un zéro de cette équation* . Ceci peut se résoudre par la méthode de Newton.

Remarque 6.1 *Si le temps final t^* est libre, on peut se ramener à la formulation précédente, en considérant t^* comme une inconnue auxiliaire. On augmente alors la dimension de l'état en considérant l'équation supplémentaire $\frac{dt^*}{dt} = 0$. On peut utiliser le même si le contrôle est bang-bang, pour déterminer les temps de commutation. Il peut cependant s'avérer préférable, lorsque le temps final est libre, d'utiliser la condition de transversalité sur l'Hamiltonien.*

6.3 Discrétisation totale : méthode directe

C'est la méthode la plus simple lorsqu'on aborde un problème de contrôle optimal. En discrétisant l'état et le contrôle, on se ramène à un problème d'optimisation non linéaire en dimension finie de la forme

$$\min_{Z \in C} V(Z), \tag{6.2}$$

où $Z = (x_1, x_2, \dots, x_N, u_1, u_2, \dots, u_n)$, et

$$\tilde{C} = \{Z / \tilde{g}_i(Z) = 0, \quad i \in 1, \dots, r, \tilde{g}_j(Z) \leq 0, \quad j \in r + 1, \dots, m\}. \tag{6.3}$$

Plus précisément, la méthode consiste à choisir les contrôles dans un espace de dimension finie, et à utiliser une méthode d'intégration numérique des équations différentielles. Considérons donc une subdivision $0 = t_0 < t_1 < t_2 < \dots < t_N = t^*$ de l'intervalle $[0, t^*]$. Réduisons l'espace des contrôles en considérant (par exemple) les contrôles constants par morceaux selon cette subdivision. Par ailleurs, choisissons une discrétisation de l'équation différentielle, par exemple choisissons ici (pour simplifier) la méthode d'Euler. On obtient alors, en posant

$$h_i = t_{i+1} - t_i$$

,

$$x_{i+1} = x_i + h_i f(t_i, x_i, u_i).$$

Remarque 6.2 *Il existe une infinité de méthodes d'intégration numérique. D'une part, on peut discrétiser l'ensemble des contrôles admissibles par des contrôles constants par morceaux, ou affines par morceaux, ou par des fonctions splines, etc. D'autre part, il existe de nombreuses méthodes pour discrétiser une équation différentielle ordinaire : méthode d'Euler (explicite ou implicite), méthode du point milieu, méthode de Heun, méthode Runge-Kutta, méthode d'Adams Moulton, etc [23]. De plus l'introduction d'éventuelles contraintes sur l'état ne pose aucun problème.*

La discrétisation précédente conduit donc au problème de programmation linéaire

$$\begin{aligned} x_{i+1} &= x_i + h_i f(t_i, x_i, u_i), \quad i = 0, \dots, N, \\ \min \quad &\tilde{C}(x_0, x_1, \dots, x_N, u_0, u_1, \dots, u_N), \\ &u_i \in \Omega, \quad i = 0, \dots, N - 1, \end{aligned}$$

un problème du type (6.2).

D'un point de vue général, cela revient à choisir une discrétisation des contrôles, ainsi que de l'état, dans des espaces de dimension finie :

$$\begin{aligned} u \in \text{Vect}(U_1, \dots, U_N), \quad \text{i.e. } u(t) &= \sum_{i=1}^N u_i U_i(t), \quad u_i \in \mathbb{R}, \\ x \in \text{Vect}(X_1, \dots, X_N), \quad \text{i.e. } x(t) &= \sum_{i=1}^N x_i X_i(t), \quad X_i \in \mathbb{R}, \end{aligned}$$

où les $U_i(t)$ et $X_i(t)$ représentent une base. Typiquement, on peut choisir des approximations polynomiales par morceaux. L'équation différentielle, ainsi que les éventuelles contraintes sur l'état ou le contrôle, ne sont vérifiées que sur les points de discrétisation. On se ramène bien à un problème d'optimisation non linéaire en dimension finie de la forme (6.2).

La résolution numérique d'un problème de programmation non linéaire du type (6.2) est standard. Elle peut être effectuée, par exemple, par une méthode de pénalisation, ou par une méthode SQP (séquentiel quadratic programming).

Dans ces méthodes, le but est de se ramener à des sous-problèmes plus simples, sans contraintes, en utilisant des fonctions de pénalisation pour les contraintes, ou bien d'appliquer les conditions nécessaires de Kuhn-Tucker pour des problèmes d'optimisation avec contraintes. Pour le problème (6.2), (6.3), les conditions de Kuhn-Tucker s'écrivent

$$\nabla V(Z) + \sum_{i=1}^m \lambda_i \nabla \tilde{g}_i(Z) = 0,$$

où les multiplicateurs de Lagrange λ_i vérifient

$$\lambda_i \tilde{g}_i(Z) = 0, \quad i \in \{1, \dots, r\}, \quad \text{et } \lambda_i \geq 0, \quad i \in \{r+1, \dots, m\}.$$

Les méthodes SQP consistent à calculer de manière itérative ces multiplicateurs de Lagrange, en utilisant des méthodes de Newton ou quasi-Newton. A chaque itération, on utilise une méthode de quasi-Newton pour estimer le hessien du Lagrangien associé au problème de programmation non linéaire, et on résout un sous-problème de programmation quadratique du Lagrangien.

6.4 Méthode de Newton discrète

Dans tout ce qui suit, D_0 est un compact, D est un ouvert et $D_h = \{h \in \mathbb{R}^2 / h_1 \neq 1\}$.

6.4.1 Différenciation approchée

Soit $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application. On se propose de résoudre le système algébrique

$$F(\tilde{x}^*) = 0 \tag{6.4}$$

La méthode de Newton pour résoudre le système (6.4) se formule comme suit :

Soit x^0 donné, à chaque pas on doit résoudre le système linéaire suivant

$$x^{k+1} = x^k - [(F'(x^k))^{-1}F(x^k)], \tag{6.5}$$

à condition toutefois que F' soit une matrice inversible dans un voisinage S_0 de x^* ; dans la suite on considère une approximation de $F'(\tilde{x}^k)$ est une application linéaire $J : \mathbb{R}^n \times \mathbb{R}^m \rightarrow L(\mathbb{R}^n)$ tel qu'on ait à résoudre à la place de (6.5) le système suivant :

$$x^{k+1} = x^k - J(x^k, h^k)^{-1}F(x^k), \quad k = 0, 1, \dots \tag{6.6}$$

Définition 6.1 Soit $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application G -différentiable sur $D_0 \subset D$ et $J : D_0 \times D_h \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow L(\mathbb{R}^n)$. Alors J est une approximation consistante de F' sur $D_0 \subset D_J$, si $0 \in \mathbb{R}^m$ est un point limite de D_h et

$$\lim_{h \rightarrow 0, h \in D_h} J(x, h) = F'(x), \quad \text{uniformément pour } x \in D_0. \tag{6.7}$$

Si de plus, il existe c et $\tilde{r} > 0$ tel que

$$\|F'(x) - J(x, h)\| \leq c\|h\|, \quad \forall x \in D_0, \quad h \in D_h \cap S(0, \tilde{r}), \tag{6.8}$$

Alors J est une approximation fortement consistante de F' sur D_0 .

Lemme 6.1 Soit $G : D \times D_h \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$; supposons qu'il existe un ensemble $S = S(x^*, \tilde{\delta}) \subset D$ et $D'_h \subset D_h$, tel que $\tilde{\alpha} < 1$ vérifie

$$\|G(x, h) - \tilde{x}^*\| \leq \tilde{\alpha} \|x - \tilde{x}^*\|, \quad \forall x \in S, \quad \forall h \in D'_h. \quad (6.9)$$

Alors, pour tout $x^0 \in S$ et toute suite $\{h^k\} \subset D'_h$, les itérations générés par :

$$x^{k+1} = G(x^k, h^k), \quad k = 0, 1, \dots \quad (6.10)$$

sont contenues dans S et convergent vers \tilde{x}^* . De plus,

$$R_1\{x^k\} \leq Q_1\{x^k\} \leq \tilde{\alpha}. \quad (6.11)$$

Où $Q_1\{x^k\}$ et $R_1\{x^k\}$ sont données dans la définition 7.7 de l'annexe 1 avec $p = 1$.

Preuve. voir l'annexe 1.

Lemme 6.2 Supposons que $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ soit G -différentiable sur un voisinage ouvert $S_0 \subset D$ contenant $\tilde{x}^* \in D$ pour lequel $F(\tilde{x}^*) = 0$. Si de plus F' est continue au point \tilde{x}^* et $F'(\tilde{x}^*)$ non singulière, alors $J : D_0 \times D_h \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow L(\mathbb{R}^n)$ est une approximation consistante de F' sur S_0 . Il existe alors $\tilde{\delta} > 0$ et $r > 0$ tel que l'application

$$G(x, h) = x - J^{-1}(x, h)F(x). \quad (6.12)$$

soit bien définie pour tout $x \in S = S(x^*, \tilde{\delta})$, $h \in D'_h = D_h \cap S(0, \tilde{r})$, et satisfasse

$$\|\tilde{x}^* - G(x, h)\| \leq w(x, h)\|x - \tilde{x}^*\|, \quad \forall x \in S, \quad h \in D'_h, \quad (6.13)$$

où

$$w(x, h) \rightarrow 0 \text{ quand } x \rightarrow \tilde{x}^* \text{ et } h \rightarrow 0, \quad h \in D'_h. \quad (6.14)$$

De plus, si J est une approximation fortement consistante de F' sur S_0 et si

$$\|F'(x) - F'(\tilde{x}^*)\| \leq \tilde{\gamma}\|x - \tilde{x}^*\|, \quad \forall x \in S_0. \quad (6.15)$$

il existe $\tilde{\alpha}_1, \tilde{\alpha}_2$ tel que :

$$\|x^* - G(x, h)\| \leq \tilde{\alpha}_1 \|x - x^*\|^2 + \tilde{\alpha}_2 \|h\| \|x - x^*\|, \quad \forall x \in S, \quad h \in D'_h. \quad (6.16)$$

Preuve.

On pose $\tilde{\beta} = \|F'(x^*)^{-1}\|$ et soit $\varepsilon \in (0, \frac{1}{2}\tilde{\beta}^{-1})$. J étant une approximation consistante sur S_0 , il existe $\tilde{r} > 0$ tel que D'_h soit non vide et

$$\|F'(x) - J(x, h)\| \leq \frac{1}{2}\varepsilon, \quad \forall x \in S_0, \quad h \in D'_h.$$

De plus, d'après la continuité de F' au point \tilde{x}^* , il existe $\tilde{\delta} > 0$ tel que $S = S(x^*, \tilde{\delta}) \subset S_0$ et

$$\|F'(x) - F'(x^*)\| \leq \frac{1}{2}\varepsilon, \quad \forall x \in S.$$

D'où

$$\|F'(x) - J(x, h)\| \leq \varepsilon, \quad \forall x \in S, \quad h \in D'_h.$$

En effet

$$\|F'(x^*) - F'(x) + F'(x) - J(x, h)\| \leq \|F'(x) - F'(x^*)\| + \|F'(x) - J(x, h)\| \leq \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

D'après le lemme de Perturbation 7.7 de l'annexe 1, alors $J(x, h)^{-1}$ existe et vérifie :

$$\|J(x, h)^{-1}\| \leq \frac{\tilde{\beta}}{1 - \tilde{\beta}\varepsilon},$$

dans notre cas : $A = F'(x^*)$, $C = J(x, h)$, $\tilde{\alpha} = \tilde{\beta}$, $\tilde{\beta} = \varepsilon$.
et G est bien définie sur $D \times D'_h$ et

$$\begin{aligned} \|G(x, h) - x^*\| &= \|J(x, h)^{-1}[J(x, h)(x - x^*) - F(x)]\| \\ &= \|J(x, h)^{-1}[J(x, h)(x - x^*) - F'(x)(x - x^*) + F'(x)(x - x^*) \\ &\quad - F'(x^*)(x - x^*) + F'(x^*)(x - x^*) - F(x^*) - F(x)]\| \\ &\leq \bar{\eta}[\|J(x, h) - F'(x)\| + \|F'(x) - F'(x^*)\|]\|x - x^*\| \\ &\quad + \bar{\eta}[\|F(x) - F(x^*) - F'(x)(x - x^*)\|]. \end{aligned}$$

et la relation (6.13) est vérifiée avec

$$w(x, h) = \bar{\eta}[\|J(x, h) - F'(x)\| + \|F'(x) - F'(x^*)\|] + q(x), \quad (6.17)$$

où

$$q(x) = \frac{\|F(x) - F(\tilde{x}^*) - F'(x)(x - \tilde{x}^*)\|}{\|x - \tilde{x}^*\|}, \quad \text{pour } x \neq \tilde{x}^*, \quad q(\tilde{x}^*) = 0.$$

donc la relation (6.14) est vérifiée .

A présent, supposons que la relation (6.15) soit vérifiée. Alors, d'après le lemme (7.9) de l'annexe 1, on obtient :

$$\|q(x)\| \leq \tilde{\gamma} \|x - \tilde{x}^*\| \quad (6.18)$$

avec

$$\tilde{\gamma} = \sup_{0 \leq t \leq 1} \frac{\|F'(\tilde{x}^* + t(x - \tilde{x}^*)) - F'(x)\|}{\|x - \tilde{x}^*\|}$$

Donc la relation (6.15), et la relation (6.16) soit aussi vérifiée immédiatement grâce aux relations (6.13) et (6.17) avec $\tilde{\alpha}_1 = 2\tilde{\gamma}\bar{\eta}$ et $\tilde{\alpha}_2 = \bar{\eta}c$ où c est la constante intervenant dans la relation (6.8).

On déduit que la vitesse de la convergence de la suite (6.6) est super linéaire quand $\lim_{h \rightarrow \infty} \{h^k\} = 0$.

Remarque 6.3 On a besoin de la continuité de F' au point \tilde{x}^* pour assurer que l'application $q(x) \rightarrow 0$ et que $F'(x) - F'(\tilde{x}^*) \rightarrow 0$ quand $x \rightarrow \tilde{x}^*$.

La convergence uniforme implique $J(x, h) - F'(x) \rightarrow 0$ quand $h \rightarrow 0$ et $x \rightarrow \tilde{x}^*$.

Corollaire 6.1 Supposons que $F : D \subset \mathbb{R}^n$ est G -différentiable dans un voisinage ouvert S_0 de $\tilde{x}^* \in D$ tel que $F(\tilde{x}^*) = 0$, et que F' soit continue au point \tilde{x}^* où $F'(\tilde{x}^*)$ est non singulière. Soit $J : D_J \times D_h \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow L(\mathbb{R}^n)$ une approximation consistante de F' sur S_0 . Alors il existe une boule $S_1 = S(\tilde{x}^*, \delta_1) \subset S_0$ et un nombre réel $r_1 > 0$ tel que pour tout $x^0 \in S_1$ et pour toute suite $\{h^k\} \subset D_h \cap S(0, r_1)$, les itérations $\{x^k\}$ donnée par la relation (6.6) appartiennent toujours à S_1 et convergent vers \tilde{x}^* . De plus, si

$$\lim_{h \rightarrow \infty} h^k = 0,$$

alors

$$R_1\{x^k\} = Q_1\{x^k\} = 0.$$

Preuve.

Soit $\tilde{\delta} > 0$ et $\tilde{r} > 0$ des constantes obtenues du lemme 6.2 ; alors, pour $\tilde{\alpha} \in]0, 1[$ donné, la relation (6.14) assure qu'on peut choisir $\delta_1 \leq \tilde{\delta}$ et $r_1 \leq \tilde{r}$ tel que

$$w(x, h) \leq \tilde{\alpha}, \quad \forall x \in S(\tilde{x}^*, \delta_1), \quad h \in D_h \cap S(0, r_1)$$

D'où, l'existence et convergence d'une suite $\{x^k\}$ d'après le lemme 6.1.

De plus, si $\lim_{k \rightarrow \infty} h^k = 0$ alors grâce au lemme 6.1, et les relations (6.13) et (6.14), on déduit que :

$$R_1\{x^k\} \leq Q_1\{x^k\} \leq \lim_{k \rightarrow \infty} \sup w(x^k, h^k) = 0$$

Afin d'appliquer les itérations correctement au lemme 6.2 et au corollaire 6.1, il est nécessaire d'assurer que J est une approximation consistante.

Corollaire 6.2 Soit F et \tilde{x}^* satisfait les conditions du corollaire 6.1, alors il existe une constante $1 > c_1 > 0, c_2 > 0$ et une boule $S_1 = S(\tilde{x}^*, \delta_1) \subset S_0$ telles que, pour tout $\tilde{x}^* \in S_1$ et les suites $\{w_k\}, \{\lambda_k\}$ satisfassent :

$$1 - c_1 \leq w_k \leq 1 + c_1, \quad -c_2 \leq \lambda_k \leq c_2, \quad k = 0, 1, \dots$$

Les itérations

$$x^{k+1} = x^k - w_k [F'(x^k) + \lambda_k I]^{-1} F(x^k), \quad k = 0, 1, \dots$$

sont contenues dans S_1 et convergent vers \tilde{x}^* . De plus, si

$$\lim_{k \rightarrow +\infty} \lambda_k = 0,$$

alors

$$R_1\{x^k\} = Q_1\{x^k\} = 0.$$

Preuve.

On définit $J : S_0 \times D_h \subset \mathbb{R}^n \times \mathbb{R}^2 \rightarrow L(\mathbb{R}^n)$ par

$$J(x, h) = (1 - h_1)^{-1} [F'(x) + h_2 I],$$

où $D_h = \{h \in \mathbb{R}^2 / h_1 \neq 1\}$, comme F' est continue en \tilde{x}^* , il existe $\eta > 0$ donnée, $\delta > 0$ telle que $\|F'(x)\| \leq \|F'(\tilde{x}^*)\| + \eta = \eta_1$ pour tout $x \in S(\tilde{x}^*, \delta) \subset D_0$. On a

$$\begin{aligned} \|J(x, h) - F'(x)\| &= \left\| \frac{h_1}{1 - h_1} F'(x) + \frac{h_2}{1 - h_2} I \right\| \\ &\leq \frac{|h_1| \eta_1 + |h_2|}{|1 - h_1|}, \end{aligned}$$

en effet :

$$\begin{aligned} \|J(x, h) - F'(x)\| &= \|(1 - h_1)^{-1} (F'(x) + h_2 I) - F'(x)\| \\ &= \left\| \frac{1}{1 - h_1} F'(x) + \frac{h_2}{1 - h_1} I - F'(x) \right\| \\ &= \left\| \frac{h_1}{1 - h_1} F'(x) + \frac{h_2}{1 - h_1} I \right\| \\ &\leq \frac{|h_1| \eta_1 + |h_2|}{|1 - h_1|}. \end{aligned}$$

Ce qui montre que J est une approximation consistante de F' sur $S(\tilde{x}^*, \delta)$. Le résultat suivant s'obtient grâce au corollaire 6.1. Considérons, par exemple, que les composantes de la matrice $J(x, h) \in L(\mathbb{R}^n)$ définies par :

$$[J(x, h)]_{i,j} = \begin{cases} \frac{1}{h_{ij}} [f_i(x + \beta \sum_{i=1}^{j-1} h_{ik} e^k + h_{ij} e^j) - f_i(x + \tilde{\beta} \sum_{i=1}^{j-1} h_{ik} e^k + h_{ij} e^k)], & \text{si } h_{ij} \neq 0; \\ \partial_j f_i(x + \tilde{\beta} \sum_{i=1}^{j-1} h_{ik} e^k + h_{ij} e^k), & \text{si } h_{ij} = 0. \end{cases} \quad (6.19)$$

Où $\tilde{\beta} \in [0, 1]$ et e^1, \dots, e^n les vecteurs de la base canonique.

Si $\tilde{\beta} = 1$, alors (6.19) correspond à l'approximation de $\partial_j f_i(x)$

$$\partial_j f_i(x) = \frac{1}{h_{ij}} [f_i(x + \sum_{i=1}^j h_{ik} e^k) - f_i(x + \sum_{i=1}^{j-1} h_{ik} e^k)],$$

Si $\tilde{\beta} = 0$, alors la relation (6.19) correspond à

$$\partial_j f_i(x) = \frac{1}{h_{ij}} [f_i(x + h_{ij} e^j) - f_i(x)].$$

Si F est G -différentiable au voisinage de x , alors $J(x, h) \rightarrow F'(x)$ quand $h \rightarrow 0$,

Corollaire 6.3 *Supposons que $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ soit continûment différentiable sur un ensemble ouvert D . Alors, pour tout ensemble compact $D_0 \subset D$, il existe un $r > 0$ tel que l'application $J : D_0 \times D_h \subset \mathbb{R}^n \times \mathbb{R}^{n^2} \rightarrow L(\mathbb{R}^n)$, donnée par la relation (6.19) soit bien définie pour tout $\beta \in [0, 1]$, soit une approximation consistante de F' sur D_0 , avec $D_h = \{h \in \mathbb{R}^{n^2} / |h_{ij}| \leq r, i, j = \overline{1, n}\}$. Si de plus J*

$$\|F'(x) - F'(y)\| \leq \tilde{\gamma} \|x - y\|. \quad \forall x, y \in D. \quad (6.20)$$

Alors J est une approximation fortement consistante de F' sur D_0 .

Preuve.

Comme D_0 est un compact et D est un ouvert, il existe $\delta > 0$ tel que l'ensemble compact $D_1 = \{x / \|x - y\|_1 \leq \tilde{\delta}, \text{ pour un certain } y \in D_0\}$ soit dans D .

Clairement, F' est uniformément continue sur D_1 , et, par conséquence pour $\varepsilon > 0$ donnée, il existe $\delta_1 \in (0, \delta)$ tel que

$$|\partial_j f_i(x) - \partial_j f_i(y)| \leq \varepsilon, \quad i, j = \overline{1, n}, \quad \forall x, y \in D_1, \quad \|x - y\|_1 \leq \delta_1.$$

On pose $\tilde{r} = \frac{\delta_1}{n}$ et $\Delta_{ij}(h) = \tilde{\beta} \sum_{k=1}^{j-1} h_{ik} e^k$, alors, pour tout $h \in D_h$,

$$\|\Delta_{ij}(h) + h_{ij} e^j\|_1 \leq n\tilde{r} \leq \delta_1 < \tilde{\delta}, \quad i, j = \overline{1, n}.$$

ce qui montre que $x + \Delta_{ij}(h) + h_{ij} e^j \in D_1, \forall x \in D_0$.

Par conséquent, d'après le théorème 17 de l'annexe 1, on obtient :

$$\begin{aligned} & \left| \frac{1}{h_{ij}} [f_i(x + \Delta_{ij}(h) + h_{ij} e^j) - f_i(x + \Delta_{ij}(h))] - \partial_j f_i(x) \right| \\ & \leq \left| \frac{1}{h_{ij}} [f_i(x + \Delta_{ij}(h) + h_{ij} e^j) - f_i(x + \Delta_{ij}(h))] - \partial_j f_i(x + \Delta_{ij}(h)) \right| \\ & + |\partial_j f_i(x + \Delta_{ij}(h)) - \partial_j f_i(x)| \leq \varepsilon + \varepsilon = 2\varepsilon \end{aligned} \quad (6.21)$$

d'où

$$\|F'(x) - J(x, h)\|_1 \leq 2n\varepsilon, \quad \forall x \in D_0, \quad h \in D_h.$$

Comme ε est arbitraire, ce qui prouve que J est bien une approximation consistante de F' sur D_0 . Si la relation (6.20) est vérifiée, alors,

$$|\partial_j f_i(x) - \partial_j f_i(y)| \leq \tilde{\gamma}_1 \|x - y\|_1, \quad \forall x, y \in D.$$

D'après le théorème 17 de l'annexe 1, on en déduit que la partie droite de l'inégalité (6.21) peut être remplacé par :

$$\gamma_1 [|h_{ij}| + |\Delta_{ij}(h)|] \leq \gamma_1 \sum_{k=1}^n |h_{ik}|$$

D'où

$$\|F'(x) - J(x, h)\|_1 \leq \gamma_1 \sum_{i,j=1}^n h = \gamma_1 \|h\|_1, \quad \forall x \in D_0,$$

Ce qui prouve que J est une approximation fortement consistante sur D_0 . Pour assurer que J soit bien définie, il faut que h soit petit. Si F est bien définie sur tout \mathbb{R}^n , on prend $D_h = \mathbb{R}^{n^2}$.

D'après les corollaires 6.3 et 6.1, on obtient un résultat de convergence résumé dans le paragraphe suivant :

6.4.2 Convergence de la méthode Newton discrète

Supposons que $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ soit continûment différentiable, et qu'il existe une solution \tilde{x}^* de $F(x) = 0$ tel que $F'(\tilde{x}^*)$ est non singulière. Définissons $J : \mathbb{R}^n \times \mathbb{R}^{n^2} \rightarrow L(\mathbb{R}^n)$ par (6.19).

Alors il existe $r_1 > 0$ et $\delta_1 > 0$ tel que, pour tout $x^0 \in S(\tilde{x}^*, \delta_1)$ et toute suite $\{h^k\} \subset S(0, r_1) \subset \mathbb{R}^{n^2}$, les itérations $\{x^k\}$ donné par (6.6) sont bien définies et convergent vers \tilde{x}^* . En plus, si $\lim_{k \rightarrow \infty} h^k = 0$, alors $R_1\{x^k\} = Q_1\{x^k\} = 0$.

Pour obtenir une convergence rapide, il est généralement nécessaire d'introduire les trois conditions supplémentaires :

1. F est une fonction infiniment différentiable.
2. J est une approximation fortement consistante.
3. Le taux de décroissance de h^k est suffisamment rapide.

Si F' satisfait la condition de Lipschitz :

$$\|F'(x) - F'(\tilde{x}^*)\| \leq \tilde{\gamma} \|x - \tilde{x}^*\|, \quad \forall x \in S_0,$$

et J est une approximation fortement consistante, alors d'après (6.16) on a :

$$\|x^{k+1} - \tilde{x}^*\| \leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \|h\| \|x^k - \tilde{x}^*\|.$$

- Si $\tilde{\alpha}_2 = 0$, on a une convergence quadratique de la suite $\{x^k\}$.
- Si $\tilde{\alpha}_2 \neq 0$, alors le comportement de h^k quand $k \rightarrow \infty$, joue un rôle important pour estimer le taux de convergence ; le résultat suivant indique deux conditions sur h^k pour assurer la convergence rapide.

Proposition 6.1 *Supposons que $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ est G -différentiable dans un voisinage ouvert $S_0 \subset D$ de $\tilde{x}^* \in D$, où $F(\tilde{x}^*) = 0$, si de plus la condition de Lipschitz (6.15) est vérifiée et si $F'(\tilde{x}^*)$ non singulière alors : soit $J : D_J \times D_h \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow L(\mathbb{R}^n)$ une approximation fortement consistante de F' sur S_0 . Supposons que pour un certain $\{h^k\} \subset D_h$ les itérations $\{x^k\}$ donnée par la relation (6.6) sont bien définies et convergent vers \tilde{x}^* . Si, en plus, la condition*

$$\|h^k\| \leq \beta_1 \|F(x^k)\|, \forall k \geq k_0, \text{ vérifiée, alors, } O_R\{x^k\} \geq O_Q\{x^k\} \geq 2, \quad (6.22)$$

sinon, si

$$\|h^k\| \leq \beta_2 \|x^k - x^{k-1}\|, \forall k \geq k_0, \text{ vérifiée, alors, } O_R\{x^k\} \geq \frac{1}{2}(1 + \sqrt{5}). \quad (6.23)$$

Où $O_Q\{x^k\}$ est donné dans la définition 7.8 de l'annexe 1, et $O_R\{x^k\}$ est donné dans la définition 7.9 de l'annexe 1.

Preuve.

D'après le lemme 6.2, il existe $\tilde{\delta} > 0$, $\tilde{r} > 0$ tel que, pour tout $x \in S = S(\tilde{x}^*, \tilde{\delta}) \subset S_0$ et $h \in D'_h = D_h \cap S(0, \tilde{r})$, la relation (6.16) est vérifiée.

Supposons que la relation (6.22) soit aussi vérifiée, alors $\lim_{k \rightarrow \infty} x^k = \tilde{x}^*$ implique que $\lim_{k \rightarrow +\infty} F(x^k) = 0$, et $x^k \in S$, $h^k \in D'_h$, pour tout $k \geq k_1 \geq k_0$.

Par conséquent, d'après la relation (6.16) on a

$$\|x^{k+1} - \tilde{x}^*\| \leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_1 \|F(x^k)\| \|x^k - \tilde{x}^*\|, \forall k \geq k_1.$$

Mais

$$\begin{aligned} \|F(x^k)\| &\leq \|F(x^k) - F(\tilde{x}^*) - F'(\tilde{x}^*)(x^k - \tilde{x}^*)\| + \|F'(\tilde{x}^*)(\tilde{x}^* - x^k)\| \\ &\leq [\varepsilon_k + \|F'(\tilde{x}^*)\|] \|x^k - \tilde{x}^*\|, \end{aligned}$$

en effet :

$$\begin{aligned} \|F(x^k)\| &= \|F(x^k) - F(\tilde{x}^*) - F'(\tilde{x}^*)(x^k - \tilde{x}^*) + F'(\tilde{x}^*)(x^k - \tilde{x}^*)\| \\ &\leq \|F(x^k) - F(\tilde{x}^*) - F'(\tilde{x}^*)(x^k - \tilde{x}^*)\| + \|F'(\tilde{x}^*)(\tilde{x}^* - x^k)\| \\ &\leq \varepsilon_k \|x^k - \tilde{x}^*\| + \|F'(\tilde{x}^*)\| \|x^k - \tilde{x}^*\| \\ &= [\varepsilon_k + \|F'(\tilde{x}^*)\|] \|x^k - \tilde{x}^*\|. \end{aligned}$$

avec

$$\varepsilon_k = \|x^k - \tilde{x}^*\|,$$

on a

$$\|F(x^k) - F(\tilde{x}^*) - F'(\tilde{x}^*)(x^k - \tilde{x}^*)\| \leq \|x^k - \tilde{x}^*\|^2 \leq \varepsilon_k \|x^k - \tilde{x}^*\|,$$

où $\lim_{k \rightarrow \infty} \varepsilon_k = 0$, et le résultat demandé c'est une conséquence directe de la définition 7.1 de l'annexe 1 avec $p = 2$.

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_1 [\varepsilon_k + \|F'(\tilde{x}^*)\|] \|x^k - \tilde{x}^*\|^2 \\ &\leq c_2 \|x^k - \tilde{x}^*\|^2, \end{aligned}$$

avec

$$c_2 = \tilde{\alpha}_1 + \tilde{\alpha}_2 \beta_1 [\varepsilon_k + \|F'(\tilde{x}^*)\|].$$

De la même manière, si la relation (6.13) est vérifiée, alors

$$\begin{aligned} \|x^{k+1} - \tilde{x}^*\| &\leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^k - x^{k-1}\| \|x^k - \tilde{x}^*\| \\ &\leq (\tilde{\alpha}_1 + \tilde{\alpha}_2 \beta_2) \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^{k-1} - x^*\| \|x^k - \tilde{x}^*\|, \quad \forall k \geq k_1, \end{aligned}$$

en effet,

$$\begin{aligned} \|x^{k+1} - \tilde{x}^*\| &\leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^k - x^{k-1}\| \|x^k - \tilde{x}^*\| \\ &\leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^k - x^{k-1} - \tilde{x}^* + \tilde{x}^*\| \|x^k - \tilde{x}^*\| \\ &\leq \tilde{\alpha}_1 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^k - x^{k-1}\| \|x^k - \tilde{x}^*\| \\ &\leq (\tilde{\alpha}_1 + \tilde{\alpha}_2 \beta_2) \|x^k - \tilde{x}^*\|^2 + \tilde{\alpha}_2 \beta_2 \|x^{k-1} - x^*\| \|x^k - \tilde{x}^*\|. \end{aligned}$$

Il s'en suit que $O_R\{x^k\} \geq \tau$ où $\tau = \frac{1}{2}(1 + \sqrt{5})$ est la racine positive du polynôme $t^2 - t - 1 = 0$.

Chapitre 7

Résolution par la méthode de relaxation d'un problème de contrôle optimal avec une entrée libre

7.1 Introduction

Dans ce chapitre, nous présentons une méthode pour déterminer analytiquement et numériquement la solution d'un problème de contrôle optimal avec un temps terminal fixé ainsi qu'une contrainte sur l'état final et initial. Nous considérons ici un problème de contrôle optimal avec une entrée libre ; de plus, le critère à minimiser est une somme pondérée entre d'une part, la distance entre l'état du système et un état désiré et d'autre part, la distance entre la commande nominale et une commande conduisant asymptotiquement à un état stationnaire. Nous présentons préalablement les équations décrivant le modèle mathématique et, en utilisant le principe du minimum de Pontriaguine, nous donnons une caractérisation de la solution du problème à résoudre, permettant ainsi de déterminer la commande optimale en fonction de la variable d'état adjoint. Nous comparons ensuite, sur un exemple simple, la solution analytique à la solution numérique calculée en utilisant la méthode de relaxation couplée à la méthode de tir ; nous terminons en donnant quelques remarques sur les performances de l'algorithme numérique (vitesse de convergence et temps de calculs).

7.2 Position du problème

Soit le système dynamique suivant :

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ x(0) = z \in X_0 = \{z \in \mathbb{R}^n, Gz = \gamma, d_* \leq z \leq d^*\}, x(t^*) = x^*, t \in [0, t^*], \\ u(t) \in \mathcal{U}_{ad}. \end{cases} \quad (7.1)$$

où $x(t)$ est un n -vecteur représentant l'état du système à l'instant t , $x(0) = z$ est la condition initiale et $x(t^*) = x^*$ est l'état final. $u(t)$ est un r -vecteur représentant la commande agissant sur le système à l'instant $t \in [0, t^*]$, \mathcal{U}_{ad} est l'ensemble de commandes admissibles. A , B sont des $n \times n$ - et $n \times r$ -matrices données, G est une matrice de l lignes et de n colonnes de rang $G = l \leq n$, γ est un l -vecteur. De plus on suppose que A est l'opposé d'une M -matrice (voir la définition 7.2 et la proposition 7.1 de l'annexe 1).

On cherche une commande admissible \hat{u} qui transfère le système d'un état initial $x(0)$ vers un état final x^* fixé et minimisant la fonction coût J définie par :

$$J(u) = \frac{1}{2} \int_0^{t^*} (x^t Q x + u^t N u) dt,$$

les matrices Q et N sont symétriques, Q est définie non-négative et N est définie positive.

L'Hamiltonien est donné par :

$$H(x(t), \psi(t), u(t), t) = \frac{1}{2} [x^t Q x + u^t N u] + \psi^t \cdot [Ax + Bu].$$

Cherchons maintenant la commande qui minimise l'Hamiltonien ; cela revient à chercher $\hat{u}(t)$, tel que

$$H(\hat{x}(t), \hat{\psi}(t), \hat{u}(t)) \leq H(x(t), \psi(t), u(t)); \forall u(t) \in \mathcal{U}_{ad}, \forall t \in [0, t^*].$$

Les équations d'optimalité s'écrivent :

$$\begin{cases} \frac{dx}{dt} = \frac{\partial H}{\partial \psi} = Ax + Bu; x(0) = z, x(t^*) = x^*, \\ -\frac{d\psi}{dt} = \frac{\partial H}{\partial x} = A^t \psi(t) + Qx(t), \\ \frac{\partial H}{\partial u} = 0 = Nu(t) + B^t \psi(t). \end{cases} \quad (7.2)$$

Ces équations sont connues sous le nom d'équations d'Hamilton-Pontriaguine.

Condition de transversalité sur $\psi(t)$:

De manière générale, lorsque l'on prend en compte un coût terminal, le critère à minimiser s'écrit :

$$J = g(t^*, x(t^*)) + \int_0^{t^*} f(x(t), u(t), t) dt,$$

où g est le coût terminal, l'état final étant fixé, on a donc classiquement :

$$\begin{cases} \text{il existe } (k_0, k_1) \neq 0, \text{ tel que} \\ \tilde{\varphi}(0, z, t^*, x(t^*), k_0, k_1) = g(0, x_0, t^*, x(t^*)) + (p_0(0, z)|k_0) + (p_1(t^*, x(t^*))|k_1), \end{cases} \quad (7.3)$$

où $p_0(0, z) = 0$ représente des contraintes sur les conditions initiales, et $p_1(t^*, x(t^*)) = 0$ représente des contraintes sur les conditions finales avec p_0 et p_1 de classe C^1 par rapport à x .

$$\begin{cases} \psi(0) = \frac{\partial \tilde{\varphi}}{\partial z}(0, z, t^*, x(t^*), k_0, k_1), \\ \psi(t^*) = -\frac{\partial \tilde{\varphi}}{\partial x(t^*)}(0, z, t^*, x(t^*), k_0, k_1), \end{cases} \quad (7.4)$$

$k_i, i = 1, 2$ étant les multiplicateurs de Lagrange.

Dans notre problème, $g(t^*, x(t^*)) = 0$, alors les conditions de transversalité (7.3) et (7.4) sur le vecteur adjoint s'écrivent :

$$\begin{cases} \text{il existe } (k_0, k_1) \neq 0, \text{ tel que} \\ \tilde{\varphi}(0, z, t^*, x(t^*), k_0, k_1) = (p_0(0, z)|k_0) + (p_1(t^*, x(t^*))|k_1), \end{cases} \quad (7.5)$$

$$\begin{cases} \psi(0) = \frac{\partial \tilde{\varphi}}{\partial z}(0, z, t^*, x(t^*), k_0, k_1), \\ \psi(t^*) = -\frac{\partial \tilde{\varphi}}{\partial x(t^*)}(0, z, t^*, x(t^*), k_0, k_1). \end{cases} \quad (7.6)$$

Remarque 7.1 *On aboutit à la résolution d'un système algébro-différentiel. L'équation d'état, décrivant le système physique est munie d'une condition initiale $x(0) = z$ et d'une condition finale $x(t^*) = x^*$. Par contre, la seconde équation, correspondante à l'équation de l'état adjoint, n'est munie d'aucune condition initiale ou d'aucune condition terminale exploitable pratiquement. On va donc utiliser la méthode de tir pour en déduire la valeur de $\psi(0)$.*

7.2.1 La méthode de tir simple

La méthode de tir permet d'obtenir la valeur de $\psi(0)$ nécessaire à la résolution du problème aux deux bouts obtenus par application du principe de Pontriaguine. Si on est capable, à partir de la condition de minimisation de l'Hamiltonien d'exprimer le contrôle en fonction de $(x(t), \psi(t))$

alors on obtient un système différentiel de la forme $\dot{v}(t) = V(t, v(t))$ où $v(t) = (x(t), \psi(t))$ et, où les conditions initiales et finales s'écrivent sous la forme $R(v(0), v(t^*)) = 0$. Finalement, on obtient le problème aux valeurs limites suivant :

$$\dot{v}(t) = V(t, v(t)), R(v(0), v(t^*)) = 0. \quad (7.7)$$

La solution du problème de Cauchy

$$\dot{v}(t) = V(t, v(t)), v(0) = v_0.$$

La fonction de tir est définie par

$$\varphi(t^*, v_0) = R(v_0, v(t^*, v_0)).$$

Soit $v(t, v_0)$, le problème (7.7) est équivalent à :

$$\varphi(t^*, v_0) = 0. \quad (7.8)$$

Il s'agit donc de déterminer un zéro de la fonction de tir φ . Comme (7.8) représente un système algébrique non linéaire, si on connaît une approximation de v , il peut se résoudre par la méthode de Newton. Dans notre problème, φ s'écrit $\varphi = x(t^*) - x^*$. Rappelons la formulation de la méthode de Newton ; il s'agit à présent de résoudre numériquement $\varphi(v) = 0$, où φ est une fonction de classe C^1 ; le principe de la méthode de Newton est le suivant :

A une étape k donnée, soit v^k une approximation d'un zéro v de φ ; donc v s'écrit $v = v^k + \Delta v^k$, et on a alors :

$$0 = \varphi(v) = \varphi(v^k + \Delta v^k) = \varphi(v^k) + \frac{\partial \varphi}{\partial v}(v^k) \cdot (v - v^k) + o(v - v^k),$$

ce qui entraîne

$$\frac{\partial \varphi}{\partial v}(v^k) \cdot (v - v^k) = -\varphi(v^k),$$

où $\frac{\partial \varphi}{\partial v}(v^k)$ est la matrice Jacobienne de l'application $v \mapsto \varphi(v)$ calculée quand $v = v^k$; or on ne connaît la fonction $v \mapsto \varphi(z)$ que numériquement. On va donc utiliser un procédé de dérivation numérique basé sur la méthode des différences finies. Pour calculer $\frac{\partial \varphi}{\partial v}(v^k)$, nous retiendrons une approximation de $\frac{\partial \varphi}{\partial v}(v^k)$ (voir [67]), incluant pour des choix particuliers des paramètres de la méthode des différences finies symétriques et définie ci-dessous,

$$\frac{\partial \varphi_i}{\partial v_j}(v^k) \simeq \frac{1}{h_{ij}} [\varphi_i(v + \sum_{k=1}^j h_{ik} e^k) - \varphi_i(v + \sum_{k=1}^{j-1} h_{ik} e^k)],$$

où h_{ij} sont des paramètres de discrétisation correspondant à la $i^{\text{ème}}$ équation et à la $j^{\text{ème}}$ variable, et e^k est le $k^{\text{ème}}$ vecteur de la base canonique. Soit

$\Delta_{ij}(v, h)$ une approximation de $\frac{\partial \varphi_i}{\partial v_j}(v)$; si l'approximation par différences finies est consistante alors,

$$\lim_{h \rightarrow 0} \Delta_{ij}(v, h) = \frac{\partial \varphi_i}{\partial v_j}(v), i, j = 1, \dots, n.$$

On pose,

$$J(x, h) = (\Delta_{ij}(v, h)).$$

De manière générale, on a à considérer l'itération suivante :

$$v^{k+1} = v^k - J(v^k, h^k)^{-1} \cdot \varphi(v^k).$$

Le problème de la convergence de ce processus itératif est résolu grâce à un résultat du livre d'Ortega et Rheinboldt [67]; en effet si les pas de discrétisation h_{ij} sont petits et tendent vers zéro, on est assuré de la convergence de ce processus.

7.3 Résolution numérique

Pour résoudre le problème, nous considérons le couplage de la méthode de relaxation avec la méthode de tir, cette dernière étant destinée à calculer $\psi(0)$. Les étapes de la méthode sont :

1. Approximation initiale de la commande $u^0(t), t \in [0, t^*]$, et de l'état adjoint $\psi^0(0)$.
2. $r \leftarrow 0$
3. **Tant que** convergence $> \varepsilon$ **faire**
 - Détermination de l'état $x^r(t)$ et de l'état adjoint $\psi^r(t)$ composante par composante séquentiellement pour $t \in [0, t^*]$ par intégration numérique, pour le temps croissants.

$$\begin{cases} \frac{dx^r}{dt} = Ax^r + Bu^r(t), \\ x(0) = z. \end{cases} \quad (7.9)$$

$$\begin{cases} -\frac{d\psi^r}{dt} = A^t \psi^r + Qx^r, \\ \psi^r(0) \end{cases} \quad (7.10)$$

où $\psi^r(0)$ est calculé par la méthode de tir.

- Détermination de la commande $u^{r+1}(t)$

$$u^{r+1}(t) \leftarrow -N^{-1}B^t \psi^r(t). \quad (7.11)$$

- Convergence $\leftarrow |u^{r+1}(t) - u^r(t)|$.
- Détermination de la fonction de tir :

$$\varphi(\psi) = x^r(t^*) - x^*.$$

- Solution de l'équation de tir par la méthode de Newton et détermination de la nouvelle valeur de $\psi(0)$

$$\psi^{r+1}(0) \leftarrow \psi^r(0) + \text{correction}.$$

- $r \leftarrow r + 1$.

fin de tant que.

Remarque 7.2 *Les étapes (7.9) à (7.11) de la boucle correspondent à la méthode de relaxation alors que les étapes suivantes correspondent à la mise en oeuvre de la méthode de tir.*

7.4 Exemple numérique

On considère le problème suivant :

$$\left\{ \begin{array}{l} \text{Déterminer } u \in \mathcal{U}_{ad} \text{ tel que,} \\ J(u) \leq J(v), \forall v \in \mathcal{U}_{ad}, \end{array} \right.$$

où

$$J(u) = \frac{1}{2} \int_0^{t^*} ((x - x_d)^2 + ku^2) dt. \quad (7.12)$$

sous les contraintes suivantes :

$$\left\{ \begin{array}{l} \dot{x}_1 = -bx_1 + ax_2, \\ \dot{x}_2 = ax_1 - bx_2 + u, \\ x(0) = z \in X_0 = \{z \in \mathbb{R}^2 : Gz = \gamma, 0 \leq z_1 \leq 20, -10 \leq z_2 \leq 10\}, \\ \kappa x(t^*) = g, \end{array} \right. \quad (7.13)$$

où $G = (1, 2)$, $\gamma = 3$, $\kappa = (1, 0)$, $g = 2$, $t^* = 2$, $a > 0$, $b > 0$.

Dans le modèle mathématique précédent, x_d correspond à un état désiré, u est la commande, k permet de doser deux critères distincts à minimiser, un critère de précision et un autre de minimisation d'énergie et \mathcal{U}_{ad} est l'ensemble des commandes admissibles. On prend $N = k.Id$, avec évidemment

$k > 0$ pour vérifier l'hypothèse de définie positivité de N et $Q = Id$.
 On peut reformulé le problème (7.13) d'une autre manière :

$$\left\{ \begin{array}{l} \dot{x}_1 = -bx_1 + ax_2, \\ \dot{x}_2 = ax_1 - bx_2 + u, \\ x_1(0) = z_1, x_2(0) = z_2 \text{ tel que } z_1 + 2z_2 = 3, \quad 0 \leq z_1 \leq 20, \quad -10 \leq z_2 \leq 10, \\ x_1(t^*) = 2. \end{array} \right. \quad (7.14)$$

La forme matricielle du système d'état s'écrit :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -b & a \\ a & -b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u,$$

et le critère de coût est reformulé comme suit

$$J(u) = \frac{1}{2} \int_0^{t^*} [(x_1 - x_{1d}, x_2 - x_{2d})^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 - x_{1d} \\ x_2 - x_{2d} \end{pmatrix} + ku^2] dt.$$

L'Hamiltonien du système est donné par :

$$H(x(t), \psi(t), u(t), t) = \frac{1}{2} ((x_1 - x_{1d})^2 + (x_2 - x_{2d})^2 + ku^2) + \psi_1(t)(-bx_1 + ax_2) + \psi_2(t)(ax_1 - bx_2) + \psi_2(t)u.$$

Les équations d'optimalité sont données par :

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{dx_1}{dt} = \frac{\partial H}{\partial \psi_1} = -bx_1 + ax_2, \\ \frac{dx_2}{dt} = \frac{\partial H}{\partial \psi_2} = ax_1 - bx_2, \\ z_1 + 2z_2 = 3, \end{array} \right. \\ \left\{ \begin{array}{l} -\frac{d\psi_1}{dt} = \frac{\partial H}{\partial x_1} = -b\psi_1 + a\psi_2 + x_1 - x_{1d}, \quad \psi_1(0) = \alpha_1, \quad \psi_1(t^*) = \alpha_2, \\ -\frac{d\psi_2}{dt} = \frac{\partial H}{\partial x_2} = a\psi_1 - b\psi_2 + x_2 - x_{2d}, \quad \psi_2(0) = \alpha_3, \quad \psi_2(t^*) = \alpha_4, \end{array} \right. \\ \frac{\partial H}{\partial u} = 0 = ku + \psi_2(t). \end{array} \right.$$

où les coefficients α_i , $i = \{1, 2, 3, 4\}$ seront calculés à partir des conditions de transversalité :

$$\begin{cases} \text{il existe } (k_0, k_1) \neq 0, \text{ tel que} \\ \tilde{\varphi}(0, z, t^*, x(t^*), k_0, k_1) = (p_0(0, z|k_0) + (p_1(t^*, x(t^*))|k_1), \end{cases}$$

où

$$\begin{aligned} p_0(0, z) &= Gz - \gamma, \\ p_1(t^*, x(t^*)) &= \kappa x(t^*) - g, \\ \tilde{\varphi}(0, z, t^*, x(t^*), k_0, k_1) &= (Gz - \gamma|k_0) + (\kappa x(t^*) - g|k_1). \end{aligned}$$

$$\begin{cases} \psi(0) = \frac{\partial \tilde{\varphi}}{\partial z}(0, z, t^*, x(t^*), k_0, k_1) = G^t k_0, \\ \psi(t^*) = -\frac{\partial \tilde{\varphi}}{\partial x(t^*)}(0, z, t^*, x(t^*), k_0, k_1) = -\kappa^t k_1, \end{cases}$$

$$\begin{cases} (\psi_1(0), \psi_2(0)) = (1, 2)k_0 = (k_0, 2k_0), \\ (\psi_1(t^*), \psi_2(t^*)) = -(1, 0)k_1 = (-k_1, 0). \end{cases}$$

Les coefficients α_i , $i = \{1, 2, 3, 4\}$, dépendent des paramètres k_0 , k_1 , qui seront calculés en utilisant la méthode du tir. Cette méthode consiste à résoudre le système suivant en utilisant les conditions aux limites, $x_1(t^*) = 2$, $x_2(t^*) = 0$, $x_1(0) + 2x_2(0) = 3$:

$$\begin{cases} \dot{x}_1 = -bx_1 + ax_2, \\ \dot{x}_2 = ax_1 - bx_2 + u, \\ z_1 + 2z_2 = 3, \\ k_1 = -\psi_1(t^*), \\ k_0 = \psi_1(0), \\ u = -\frac{\psi_2(t)}{k}. \end{cases}$$

La résolution de ce problème sera faite de deux manières différentes, analytiquement et numériquement.

Solution numérique

En optimisation classique, minimiser une fonction H par rapport à un paramètre u implique dans un premier temps de chercher les points stationnaires du système, c'est-à-dire les valeurs u pour lesquelles $\frac{\partial H}{\partial u}(x, u, \psi) = 0$,

puis d'étudier en ces points la positivité de la matrice Hessienne $\frac{\partial^2 H}{\partial^2 u}(x, u, \psi)$.
On aura alors :

$$\frac{\partial^2 H}{\partial^2 u}(x, u, \psi) = k;$$

pour assurer la positivité de $\frac{\partial^2 H}{\partial^2 u}(x, u, \psi)$ on a nécessairement $k > 0$, c'est une condition naturelle pour le problème considéré.

On a à résoudre le système suivant :

$$\left\{ \begin{array}{l} \dot{v}_1 = -bv_1 + av_2 + u, \\ \dot{v}_2 = av_1 - bv_2, \\ \dot{v}_3 = bv_3 - av_4 - v_1 + x_{1d}, \\ \dot{v}_4 = -av_2 + bv_4 - v_2 + x_{2d}, \\ u = -\frac{v_4}{k} \\ v_1(0) \in \mathbb{R}, v_2(0) \in \mathbb{R}, \\ v_3(0) \in \mathbb{R}, v_4(0) \in \mathbb{R}. \end{array} \right.$$

Soit $v(t, 0, 0, \psi_1, \psi_2)$ une solution du système au temps t avec les conditions initiales $(v_1(0), v_2(0), v_3(0), v_4(0))$.

On construit une fonction de tir qui est une équation algébrique non linéaire de la variable ψ à l'instant $t = 0$; cette fonction de tir est calculée par une procédure d'intégration numérique d'équations différentielles ordinaires (Euler, Runge-Kutta, etc); la fonction de tir s'écrit :

$$\varphi(\psi) = \begin{pmatrix} v_1(2, 0, 0, \psi_1, \psi_2) - 2 \\ v_2(2, 0, 0, \psi_1, \psi_2) \end{pmatrix}.$$

Le problème à résoudre s'écrit alors : Déterminer $\psi(0)$ tel que $\varphi(\psi(0))$ soit nul ce qui revient à déterminer $x(t^*)$ désiré.

L'algorithme de résolution numérique de ce problème sera alors complètement défini si l'on se donne :

1. l'algorithme d'intégration d'un système différentiel à valeur initiale (par exemple une procédure d'Euler ou de Runge-Kutta), pour calculer la fonction de tir φ .
2. l'algorithme de résolution de $\varphi(\psi) = 0$.

Appliquons l'algorithme vu au paragraphe (7.3) à notre exemple :

- Approximation initiale de la l'état adjoint $\psi^{(0)}(0)$ et de la commande $u^{(0)}(t)$, $t \in [0, t^*]$.

- $r \leftarrow 0$.
- Tant que pas de convergence faire
- Détermination de l'état $x^{(r)}(t)$ et de l'état adjoint $\psi^{(r)}(t)$, successivement composante par composante, $t \in [0, t^*]$ par intégration numérique dans le sens direct de l'équation d'état et de l'équation d'état adjoint.

$$\begin{cases} \frac{dx_1^{(r)}}{dt} = -bx_1^{(r)} + ax_2^{(r)}, \\ \frac{dx_2^{(r)}}{dt} = ax_1^{(r)} - bx_2^{(r)} + u^{(r)}(t), \\ z_1 + 2z_2 = 3, \end{cases}$$

$$\begin{cases} -\frac{d\psi_1^{(r)}}{dt} = -b\psi_1^{(r)} + a\psi_2^{(r)} + x_1^{(r)} - x_{1d}, & \psi_1^{(r)}(0), \\ -\frac{d\psi_2^{(r)}}{dt} = a\psi_1^{(r)} - b\psi_2^{(r)} + x_2^{(r)} - x_{2d}, & \psi_2^{(r)}(0), \end{cases}$$

où $\psi_1^{(r)}(0)$, $\psi_2^{(r)}(0)$ sont calculés par la méthode de tir.

- Détermination de la commande $u^{(r+1)}(t)$:

$$u^{(r+1)}(t) \leftarrow -\frac{\psi_2^{(r)}(t)}{k}, \quad t \in [0, t^*].$$

- Convergence \leftarrow norme $(u^{(r+1)} - u^{(r)})$.
- Détermination de la fonction de tir.
- Solution de l'équation de tir par la méthode de Newton.
- Actualiser la valeur de $\psi^{(r+1)}(0)$

$$\psi^{(r+1)}(0) \leftarrow \psi^{(r)}(0) + \text{correction}.$$

- $r \leftarrow r + 1$.
- fin de tant que.

La convergence de cette méthode itérative est un problème ouvert. Nous conjecturons qu'elle peut se montrer par des techniques de contraction analogue à celle utilisées dans [67] ; elle découle certainement du fait que A est l'opposé d'une M -matrice.

Résolution analytique

Pour calculer de manière analytique la solution optimale $x(t)$, et la commande optimale correspondante $u(t)$ du problème (7.12)-(7.14), nous avons utilisé les équations d'optimalité ainsi que la condition de transversalité sur $\psi(t)$

Pour trouver la solution exacte du problème de contrôle optimal, on utilise la méthode de dérivation au niveau des équations. On a :

$$\dot{\psi}_1 = b\psi_1 - a\psi_2 - x_1 + x_{1d},$$

en dérivant par rapport à t , on obtient :

$$\ddot{\psi}_1 = b\dot{\psi}_1 - a\dot{\psi}_2 - \dot{x}_1,$$

soit,

$$\ddot{\psi}_1 = b(b\psi_1 - a\psi_2 - x_1 + x_{1d}) - a(-a\psi_1 + b\psi_2 - x_2 + x_{2d}) - (-bx_1 + ax_2),$$

$$\ddot{\psi}_1 = b^2\psi_1 - ab\psi_2 - bx_1 + bx_{1d} + a^2\psi_1 - ab\psi_2 + ax_2 - ax_{2d} + bx_1 - ax_2,$$

$$\ddot{\psi}_1 = (a^2 + b^2)\psi_1 - 2ab\psi_2 + bx_{1d} - ax_{2d},$$

d'où,

$$\ddot{\psi}_1 = (a^2 + b^2)\psi_1 - 2ab\psi_2 + bx_{1d} - ax_{2d}.$$

De la même manière, on obtient :

$$\dot{\psi}_2 = -a\psi_1 + b\psi_2 - x_2 + x_{2d},$$

soit

$$\ddot{\psi}_2 = -a\dot{\psi}_1 + b\dot{\psi}_2 - \dot{x}_2,$$

$$\ddot{\psi}_2 = -a(b\psi_1 - a\psi_2 - x_1 + x_{1d}) + b(-a\psi_1 + b\psi_2 - x_2 + x_{2d}) - (ax_1 - bx_2 - \frac{\psi_2}{k}),$$

$$\ddot{\psi}_2 = -ab\psi_1 + a^2\psi_2 + ax_1 - ax_{1d} - ab\psi_1 + b^2\psi_2 - bx_2 + bx_{2d} - ax_1 + bx_2 + \frac{\psi_2}{k},$$

$$\ddot{\psi}_2 = -2ab\psi_1 + (a^2 + b^2 + \frac{1}{k})\psi_2 - ax_{1d} + bx_{2d}.$$

Par conséquent,

$$\ddot{\psi}_1 = (a^2 + b^2 + \frac{1}{k})\psi_2 - 2ab\psi_1 - ax_{1d} + bx_{2d}. \quad (7.15)$$

Dérivons deux fois l'équation (7.15), on obtient :

$$\psi_1^{(4)} = (a^2 + b^2)\ddot{\psi}_1 - 2ab\ddot{\psi}_2,$$

$$\psi_1^{(4)} = (a^2 + b^2)\ddot{\psi}_1 - 2ab(-2ab\psi_1 + (a^2 + b^2 + \frac{1}{k})\psi_2 - ax_{1d} + bx_{2d}),$$

$$\psi_1^{(4)} = (a^2 + b^2)\ddot{\psi}_1 + 4a^2b^2\psi_1 - 2ab(a^2 + b^2 + \frac{1}{k})\psi_2 + 2a^2bx_{1d} - 2ab^2x_{2d},$$

d'où

$$\begin{aligned} \psi_1^{(4)} &= (a^2 + b^2)\ddot{\psi}_1 + 4a^2b^2\psi_1 - 2ab(a^2 + b^2 + \frac{1}{k}) \\ &\quad + 2a^2bx_{1d} - 2ab^2x_{2d}, \end{aligned} \quad (7.16)$$

(7.15) entraîne :

$$2ab\psi_2 = (a^2 + b^2)\psi_1 - \ddot{\psi}_1 + bx_{1d} - ax_{2d}. \quad (7.17)$$

En injectant (7.17) dans (7.16), on obtient :

$$\begin{aligned} \psi_1^{(4)} &= (a^2 + b^2)\ddot{\psi}_1 + 4a^2b^2\psi_1 - (a^2 + b^2 + \frac{1}{k})((a^2 + b^2)\psi_1 - \ddot{\psi}_1 + bx_{1d} - ax_{2d}) \\ &\quad + 2a^2bx_{1d} - 2ab^2x_{2d}, \end{aligned}$$

$$\begin{aligned} \psi_1^{(4)} &= (a^2 + b^2)\ddot{\psi}_1 + 4a^2b^2\psi_1 - ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2))\psi_1 + (a^2 + b^2 + \frac{1}{k})\ddot{\psi}_1 \\ &\quad - b(a^2 + b^2 + \frac{1}{k})x_{1d} + a(a^2 + b^2 + \frac{1}{k})x_{2d} + 2a^2bx_{1d} - 2ab^2x_{2d}, \end{aligned}$$

$$\begin{aligned} \psi_1^{(4)} &= (2a^2 + 2b^2 + \frac{1}{k})\ddot{\psi}_1 - ((a^2 + b^2)^2 + \frac{a^2 + b^2}{k} - 4a^2b^2)\psi_1 \\ &\quad + (-b^3 - \frac{b}{k} + a^2b)x_{1d} + (a^3 - ab^2 + \frac{a}{k})x_{2d}, \end{aligned}$$

d'où,

$$\begin{aligned} \psi_1^{(4)} &- (2a^2 + 2b^2 + \frac{1}{k})\ddot{\psi}_1 - ((a^2 + b^2)^2 + \frac{a^2 + b^2}{k} - 4a^2b^2)\psi_1 \\ &= (-b^3 - \frac{b}{k} + a^2b)x_{1d} + (a^3 - ab^2 + \frac{a}{k})x_{2d}, \end{aligned} \quad (7.18)$$

L'équation caractéristique correspondante à l'équation (7.18) d'écrivent comme suit :

$$C^4 - 2(a^2 + b^2 + \frac{1}{2k})C^2 + ((a^2 + b^2)^2 + \frac{a^2 + b^2}{k} - 4a^2b^2) = 0.$$

Le déterminant est donné par

$$\Delta = \frac{1}{4k^2} + 4a^2b^2,$$

Les racines de l'équation caractéristique sont données par :

$$C_1^2 = (a^2 + b^2 + \frac{1}{2k}) - \sqrt{\frac{1}{4k^2} + 4a^2b^2},$$

$$C_2^2 = (a^2 + b^2 + \frac{1}{2k}) + \sqrt{\frac{1}{4k^2} + 4a^2b^2}.$$

D'où, on obtient

$$\psi_1(t) = \lambda e^{C_1 t} + \beta e^{-C_1 t} + \mu e^{C_2 t} + \alpha e^{-C_2 t} + \nu. \quad (7.19)$$

Détermination de ν , on a :

$$\dot{\psi}_1(t) = \lambda C_1 e^{C_1 t} - \beta C_1 e^{-C_1 t} + \mu C_2 e^{C_2 t} - \alpha C_2 e^{-C_2 t},$$

$$\ddot{\psi}_1(t) = \lambda C_1^2 e^{C_1 t} + \beta C_1^2 e^{-C_1 t} + \mu C_2^2 e^{C_2 t} + \alpha C_2^2 e^{-C_2 t},$$

$$\psi_1^{(3)}(t) = \lambda C_1^3 e^{C_1 t} - \beta C_1^3 e^{-C_1 t} + \mu C_2^3 e^{C_2 t} - \alpha C_2^3 e^{-C_2 t},$$

$$\psi_1^{(4)}(t) = \lambda C_1^4 e^{C_1 t} + \beta C_1^4 e^{-C_1 t} + \mu C_2^4 e^{C_2 t} + \alpha C_2^4 e^{-C_2 t}.$$

On remplace dans (7.18), et on obtient :

$$\begin{aligned} & \lambda [C_1^4 - ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2) - 4a^2b^2) - 2(a^2 + b^2 + \frac{1}{2k})C_1^2]e^{C_1 t} \\ & + \beta [C_1^4 - ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2) - 4a^2b^2) - 2(a^2 + b^2 + \frac{1}{2k})C_1^2]e^{-C_1 t} \\ & + \mu [C_2^4 - ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2) - 4a^2b^2) - 2(a^2 + b^2 + \frac{1}{2k})C_2^2]e^{C_2 t} \\ & + \alpha [C_2^4 - ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2) - 4a^2b^2) - 2(a^2 + b^2 + \frac{1}{2k})C_2^2]e^{-C_2 t} \\ & + ((a^2 + b^2)^2 + \frac{1}{k}(a^2 + b^2) - 4a^2b^2)\nu \\ & = (-b^3 - \frac{b}{k} + a^2b)x_{1d} + (a^3 - ab^2 + \frac{a}{k})x_{2d}, \end{aligned}$$

Par identification, on obtient la valeur exacte de ν donnée par :

$$\nu = \frac{(-b^3 - \frac{b}{k} + a^2b)x_{1d} + (a^3 - ab^2 + \frac{a}{k})x_{2d}}{((a^2 + b^2)^2 + \frac{a^2+b^2}{k} - 4a^2b^2)}. \quad (7.20)$$

On en déduit aisément $\psi_2(t)$ de (7.17), et on obtient :

$$\psi_2(t) = \frac{1}{2ab}[(a^2 + b^2)\psi_1 - \ddot{\psi}_1 + bx_{1d} - ax_{2d}2ab],$$

$$\begin{aligned} \psi_2(t) &= \frac{a^2 + b^2}{2ab}[\lambda e^{C_1 t} + \beta e^{-C_1 t} + \mu e^{C_2 t} + \alpha e^{-C_2 t} + \nu] - \frac{1}{2ab}[\lambda C_1^2 e^{C_1 t} \\ &+ \beta C_1^2 e^{-C_1 t} + \mu C_2^2 e^{C_2 t} + \alpha C_2^2 e^{-C_2 t}] + \frac{1}{2a}x_{1d} - \frac{1}{2b}x_{2d}, \end{aligned}$$

alors

$$\begin{aligned} \psi_2(t) &= \lambda \left[\frac{a^2 + b^2 - C_1^2}{2ab} \right] e^{C_1 t} + \beta \left[\frac{a^2 + b^2 - C_1^2}{2ab} \right] e^{-C_1 t} + \mu \left[\frac{a^2 + b^2 - C_2^2}{2ab} \right] e^{C_2 t} \\ &+ \alpha \left[\frac{a^2 + b^2 - C_2^2}{2ab} \right] e^{-C_2 t} + \frac{a^2 + b^2}{2ab} \nu + \frac{1}{2a}x_{1d} - \frac{1}{2b}x_{2d}. \end{aligned} \quad (7.21)$$

De même compte tenu des équations :

$$x_1 = b\psi_1 - a\psi_2 - \dot{\psi}_1 + x_{1d}, \quad (7.22)$$

$$x_2 = -a\psi_1 + b\psi_2 - \dot{\psi}_2 + x_{2d}. \quad (7.23)$$

En remplaçant (7.19), (7.21) et $\dot{\psi}_1, \dot{\psi}_2$ dans (7.23), on obtient les résultats suivants :

$$\begin{aligned} x_1(t) &= b[\lambda e^{C_1 t} + \beta e^{-C_1 t} + \mu e^{C_2 t} + \alpha e^{-C_2 t} + \nu] - a \left[\lambda \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{C_1 t} \right. \\ &+ \beta \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{-C_1 t} + \mu \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{C_2 t} + \alpha \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{-C_2 t} \\ &+ \frac{a^2 + b^2}{2ab} \nu + \frac{1}{2a}x_{1d} - \frac{1}{2b}x_{2d}] - [\lambda C_1 e^{C_1 t} - \beta C_1 e^{-C_1 t} + \mu C_2 e^{C_2 t} \\ &- \alpha C_2 e^{-C_2 t}] + x_{1d}. \end{aligned}$$

D'où

$$\begin{aligned} x_1(t) &= \lambda \left[\frac{b^2 - a^2 + C_1^2}{2b} - C_1 \right] e^{C_1 t} + \beta \left[\frac{b^2 - a^2 + C_1^2}{2b} + C_1 \right] e^{-C_1 t} \\ &+ \mu \left[\frac{b^2 - a^2 + C_2^2}{2b} - C_2 \right] e^{C_2 t} + \alpha \left[\frac{b^2 - a^2 + C_2^2}{2b} + C_2 \right] e^{-C_2 t} \\ &+ \frac{b^2 - a^2}{2b} \nu + \frac{1}{2}x_{1d} + \frac{a}{2b}x_{2d}. \end{aligned} \quad (7.24)$$

et

$$\begin{aligned}
x_2(t) &= -a[\lambda e^{C_1 t} + \beta e^{-C_1 t} + \mu e^{C_2 t} + \alpha e^{-C_2 t} + \nu] + b[\lambda(\frac{a^2 + b^2 - C_1^2}{2ab})e^{C_1 t} \\
&+ \beta(\frac{a^2 + b^2 - C_1^2}{2ab})e^{-C_1 t} + \mu(\frac{a^2 + b^2 - C_2^2}{2ab})e^{C_2 t} + \alpha(\frac{a^2 + b^2 - C_2^2}{2ab})e^{-C_2 t} \\
&+ \frac{a^2 + b^2}{2ab}\nu + \frac{1}{2a}x_{1d} - \frac{1}{2d}] - [\lambda C_1(\frac{a^2 + b^2 - C_1^2}{2ab})e^{C_1 t} - \beta C_2(\frac{a^2 + b^2 - C_1^2}{2ab}) \\
&e^{-C_1 t} + \mu C_2(\frac{a^2 + b^2 - C_1^2}{2ab})e^{C_2 t} - \alpha C_2(\frac{a^2 + b^2 - C_1^2}{2ab})e^{-C_2 t}] + x_{2d},
\end{aligned}$$

d'où

$$\begin{aligned}
x_2(t) &= \lambda[\frac{b^2 - a^2 - C_1^2}{2a} - C_1(\frac{b^2 + a^2 - C_1^2}{2ab})]e^{C_1 t} + \beta[\frac{b^2 - a^2 - C_1^2}{2a} \\
&+ C_1(\frac{b^2 + a^2 - C_1^2}{2ab})]e^{-C_1 t} + \mu[\frac{b^2 - a^2 - C_2^2}{2a} - C_2(\frac{b^2 + a^2 - C_2^2}{2ab})]e^{C_2 t} \\
&+ \alpha[\frac{b^2 - a^2 - C_2^2}{2a} + C_2(\frac{b^2 + a^2 - C_2^2}{2ab})]e^{-C_2 t} + \frac{a^2 - b^2}{2a}\nu \\
&+ \frac{b}{2a}x_{1d} + \frac{1}{2}x_{2d}. \tag{7.25}
\end{aligned}$$

Les constantes étant déterminées par les conditions aux limites suivantes :

$$z_1 + 2z_2 = 3, \quad x_1(t^*) = 2, \quad \psi_1(0) = k_0, \quad \psi_2(0) = 2k_0, \quad \psi_1(t^*) = -k_1, \quad \psi_2(t^*) = 0.$$

On résoud le système linéaire numériquement :

$$\left\{ \begin{array}{l}
3 = \lambda \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 \right) + \beta \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 \right) + \mu \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 \right) + \alpha \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 \right) \\
+ \frac{b^2 - a^2}{2b} \nu + \frac{1}{2} x_{1d} + \frac{a}{2b} x_{2d} + 2 \left[\lambda \left[\frac{b^2 - a^2 - C_1^2}{2a} - C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] + \beta \left[\frac{b^2 - a^2 - C_1^2}{2a} \right. \right. \\
\left. \left. + C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] + \mu \left[\frac{b^2 - a^2 - C_2^2}{2a} - C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] + \alpha \left[\frac{b^2 - a^2 - C_2^2}{2a} + C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] \right. \\
\left. + \left(\frac{a^2 - b^2}{2a} \right) \nu + \frac{b}{2a} x_{1d} + \frac{1}{2} x_{2d} \right], \\
2 = \lambda \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 \right) e^{C_1 t^*} + \beta \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 \right) e^{-C_1 t^*} + \mu \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 \right) e^{C_2 t^*} \\
+ \alpha \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 \right) e^{-C_2 t^*} + \frac{b^2 - a^2}{2b} \nu + \frac{1}{2} x_{1d} + \frac{a}{2b} x_{2d}, \\
k_0 = \lambda + \beta + \mu + \alpha + \nu, \\
2k_0 = \lambda \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) + \beta \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) + \mu \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) + \alpha \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) + \left(\frac{a^2 + b^2}{2ab} \right) \nu \\
+ \frac{1}{2a} x_{1d} - \frac{1}{2b} x_{2d}, \\
-k_1 = \lambda e^{2C_1} + \beta e^{-2C_1} + \mu e^{2C_2} + \alpha e^{-2C_2} + \nu, \\
0 = \lambda \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{2C_1} + \beta \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{-2C_1} + \mu \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{2C_2} + \alpha \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{-2C_2} \\
+ \left(\frac{a^2 + b^2}{2ab} \right) \nu + \frac{1}{2a} x_{1d} - \frac{1}{2b} x_{2d},
\end{array} \right.$$

Et le système linéaire précédent est équivalent à :

$$\left\{ \begin{array}{l} \lambda \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 + 2 \left[\frac{b^2 - a^2 - C_1^2}{2a} - C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] \right) + \beta \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 + \right. \\ \left. 2 \left[\frac{b^2 - a^2 - C_1^2}{2a} + C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] \right) + \mu \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 + 2 \left[\frac{b^2 - a^2 - C_2^2}{2a} - C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] \right) \\ + \alpha \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 + 2 \left[\frac{b^2 - a^2 - C_2^2}{2a} + C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] \right) = -\frac{b^2 - a^2}{2b} \nu - \frac{1}{2} x_{1d} - \frac{a}{2b} x_{2d} \\ - 2 \left(\frac{a^2 - b^2}{2a} \right) \nu - \frac{b}{2a} x_{1d} - \frac{1}{2} x_{2d}, \\ \lambda \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 \right) e^{C_1 t^*} + \beta \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 \right) e^{-C_1 t^*} + \mu \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 \right) e^{C_2 t^*} \\ + \alpha \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 \right) e^{-C_2 t^*} = 2 - \frac{b^2 - a^2}{2b} \nu - \frac{1}{2} x_{1d} - \frac{a}{2b} x_{2d}, \\ \lambda + \beta + \mu + \alpha - k_0 = -\nu, \\ \lambda \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) + \beta \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) + \mu \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) + \alpha \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) - 2k_0 = -\left(\frac{a^2 + b^2}{2ab} \right) \nu \\ - \frac{1}{2a} x_{1d} + \frac{1}{2b} x_{2d}, \\ \lambda e^{2C_1} + \beta e^{-2C_1} + \mu e^{2C_2} + \alpha e^{-2C_2} + k_1 = -\nu, \\ \lambda \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{2C_1} + \beta \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{-2C_1} + \mu \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{2C_2} + \alpha \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{-2C_2} \\ = -\left(\frac{a^2 + b^2}{2ab} \right) \nu - \frac{1}{2a} x_{1d} + \frac{1}{2b} x_{2d}, \end{array} \right.$$

Soit encore sous la forme matricielle :

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & 0 & 0 \\ b_1 & b_2 & b_3 & b_4 & 0 & 0 \\ 1 & 1 & 1 & 1 & -1 & 0 \\ v_1 & v_2 & v_3 & v_4 & -2 & 0 \\ e^{2C_1} & e^{-2C_1} & e^{2C_2} & e^{-2C_2} & 0 & 1 \\ w_1 & w_2 & w_3 & w_4 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \beta \\ \mu \\ \alpha \\ k_0 \\ k_1 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ -\nu \\ f_3 \\ -\nu \\ f_3 \end{pmatrix},$$

avec

$$\begin{aligned} a_1 &= \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 + 2 \left[\frac{b^2 - a^2 - C_1^2}{2a} - C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] \right), \\ a_2 &= \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 + 2 \left[\frac{b^2 - a^2 - C_1^2}{2a} + C_1 \left(\frac{b^2 + a^2 - C_1^2}{2ab} \right) \right] \right), \\ a_3 &= \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 + 2 \left[\frac{b^2 - a^2 - C_2^2}{2a} - C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] \right), \\ a_4 &= \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 + 2 \left[\frac{b^2 - a^2 - C_2^2}{2a} + C_2 \left(\frac{b^2 + a^2 - C_2^2}{2ab} \right) \right] \right), \end{aligned}$$

$$\begin{aligned} b_1 &= \left(\frac{b^2 - a^2 + C_1^2}{2b} - C_1 \right) e^{C_1 t^*}, \quad b_2 = \left(\frac{b^2 - a^2 + C_1^2}{2b} + C_1 \right) e^{-C_1 t^*}, \\ b_3 &= \left(\frac{b^2 - a^2 + C_2^2}{2b} - C_2 \right) e^{C_2 t^*}, \quad b_4 = \left(\frac{b^2 - a^2 + C_2^2}{2b} + C_2 \right) e^{-C_2 t^*}, \end{aligned}$$

$$\begin{aligned} w_1 &= \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{C_1 t^*}, \quad w_2 = \left(\frac{a^2 + b^2 - C_1^2}{2ab} \right) e^{-C_1 t^*}, \\ w_3 &= \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{C_2 t^*}, \quad w_4 = \left(\frac{a^2 + b^2 - C_2^2}{2ab} \right) e^{-C_2 t^*}, \end{aligned}$$

$$\begin{aligned} f_1 &= 3 - \left[\frac{b^2 - a^2}{2b} \nu - \frac{1}{2} x_{1d} - \frac{a}{2b} x_{2d} + 2 \left(\left(\frac{a^2 - b^2}{2a} \right) \nu - \frac{b}{2a} x_{1d} - \frac{1}{2} x_{2d} \right) \right], \\ f_2 &= 2 - \frac{b^2 - a^2}{2b} \nu - \frac{1}{2} x_{1d} - \frac{a}{2b} x_{2d}, \quad f_3 = - \frac{a^2 + b^2}{2ab} \nu - \frac{1}{2a} x_{1d} + \frac{1}{2b} x_{2d}. \end{aligned}$$

Dans cette section, nous déterminons explicitement la trajectoire optimale $x(t)$ du problème (7.12)-(7.13) en résolvant le système linéaire numériquement.

Comparaison des deux approches

L'algorithme numérique est implémenté par le logiciel Matlab. En particulier, on a utilisé les fonctions *ode45* et *fsolve*.

Dans cet exemple, la méthode converge indépendamment du point de départ $\psi(0) = (-1.8706, 342.0562)$.

L'expérimentation numérique est déterminée pour les valeurs de $a = 1$, $b = 3$ et $k = 2$. On déduit que la solution analytique et la solution numérique correspondent parfaitement (voir figure 1 et figure 2). La performance de la procédure numérique est résumé dans le tableau suivant, pour différentes valeurs de k .

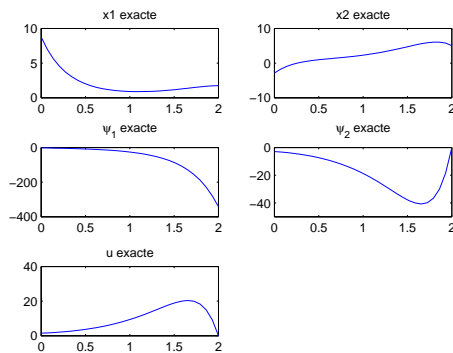


FIGURE 7.1 – solution exacte

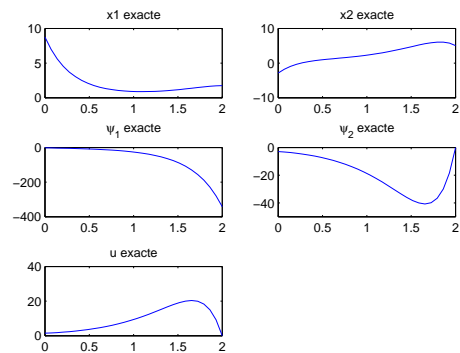


FIGURE 7.2 – solution approchée

k	temps d'exécution C.P.U	Nombre d'itération
0.5	0.2028	3
1	0.2340	4
1.5	0.2496	4
2	0.2808	5
2.5	0.3276	5

Notons que l'algorithme converge rapidement, pour un nombre d'itérations très petit nécessaire pour atteindre la convergence. De plus, le temps de calcul utilisé est très faible.

Conclusion générale

Cette étude a été consacrée essentiellement à la résolution d'un problème de contrôle optimal à entrée libre. Notre but était de trouver une meilleure condition initiale et une commande optimale qui pouvait nous ramener de l'état initial $x_0 \in X_0$ vers l'état final décrit par l'équation $Hx(t^*) = g$. Pour sa résolution, on a utilisé trois approches :

- La méthode adaptée.
- La méthode de discrétisation couplée à la procédure finale.
- La méthode de relaxation couplée à la méthode de tir.

La méthode adaptée est constituée de trois procédures : le changement de commande, le changement de support et la procédure finale.

En exploitant cette méthode, on a proposé une méthode originale qui est un couplage entre la méthode de discrétisation et la procédure finale. Le but de la méthode de discrétisation est d'approcher le support optimal. Par la suite en utilisant ce support et le critère d'optimalité découlant du principe du maximum, on a effectué une résolution appelée procédure finale basé sur la méthode de Newton.

En dernier lieu, on a effectué une approche appelée méthode de relaxation, basée sur le principe du maximum et couplée à la méthode de tir.

On a appliqué ces trois méthodes à différents types de problèmes de contrôle optimal.

Ce domaine de recherche offre de nombreuses perspectives qu'elles soient théoriques ou pratiques :

- * En théorie, il est intéressant d'appliquer ces algorithmes à des problèmes de contrôle optimal non linéaire, stochastique, feedback, multicritère, etc..
- * En pratique, différents problèmes que ce soit en économie, en agriculture, en automatique, en aéronautique, etc., peuvent être modélisés par des problèmes de contrôle optimal et résolus par les méthodes proposées.

Annexe 1

Rappels mathématiques sur l'Algèbre linéaire

On définit $L(\mathbb{R}^n)$ l'espace linéaire des opérateurs linéaires définies de \mathbb{R}^n vers \mathbb{R}^n . Et \mathbb{C}^n est l'espace des nombres complexes de dimension n .

Définition 7.1 Soit $A \in L(\mathbb{R}^n)$ une matrice inversible. Alors $A^T A$ est symétrique définie positive, où A^T est la matrice transposée.

Définition 7.2 Une matrice $A \in L(\mathbb{R}^n)$ soit une M -matrice si A est inversible, $A^{-1} \geq 0$, et $a_{ij} \leq 0$ pour tout $i, j = 1, \dots, n, i \neq j$.

On donne une des caractéristiques des M -matrices la plus utilisé :

Proposition 7.1 Soit $A \in L(\mathbb{R}^n)$, avec $a_{ij} \leq 0, i \neq j$. Alors A est une M -matrice si et seulement si :

1. Les éléments diagonales de A sont positives,
2. La matrice $B = I - D^{-1}A$, où $D = \text{diag}(a_{11}, \dots, a_{nn})$, satisfait $\rho(B) < 1$, où B est la matrice de Jacobi associée à A .

Définition 7.3 Une application $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est hémicontinue au point $x \in D$, si pour tout $h \in \mathbb{R}^n$ et $\varepsilon > 0$, il existe $\delta = \delta(\varepsilon, h)$ tel que $|t| < \infty$ et $x + th \in D$, alors

$$\|F(x + th) - F(x)\| < \varepsilon.$$

Définition 7.4 Si $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est G -différentiable en chaque point d'un ensemble convexe $D_0 \subset D$, et F' est hémicontinue sur D_0 , alors, pour tout $x, y \in D_0$, la relation suivante est vérifiée

$$F(y) - F(x) = \int_0^1 F'(x + t(y-x))(y-x)dt. \quad (7.26)$$

Définition 7.5 Si $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est G -différentiable au point $x \in D$, alors F est hémicontinue au point x .

Définition 7.6 Si $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est G -dérivable à chaque point d'un ensemble convexe $D_0 \subset D$ et F' est hémicontinue sur D_0 , alors, pour tout $x, y \in D_0$,

$$F(y) - F(x) = \int_0^1 F'(x + t(y - x))(y - x) dt,$$

est vérifiée.

Définition 7.7 Soit $\{x^k\} \subset \mathbb{R}^n$ une suite qui converge vers x^* . Pour $p \in [1, \infty[$, on définit les quantités suivantes :

$$Q_p\{x^k\} = \begin{cases} 0, & \text{si } x^k = x^* \quad k < \infty; \\ \lim_{k \rightarrow +\infty} \sup \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^p}, & \text{si } x^k \neq x^* \quad k < \infty; \\ +\infty, & \text{sinon;} \end{cases}$$

est appelée Q -facteurs (the quotient convergence factors).

Et

$$R_p\{x^k\} = \begin{cases} \lim_{k \rightarrow +\infty} \sup \|x^k - x^*\|^p, & \text{si } p = 1; \\ \lim_{k \rightarrow +\infty} \sup \|x^k - x^*\|^p, & \text{si } p > 1; \end{cases}$$

est appelée R -facteurs (the root-convergence factors).

Définition 7.8 Soit $Q_p(\mathfrak{F}, \{x^*\})$ un Q -facteurs du processus itératif \mathfrak{F} avec un point limite x^* dans \mathbb{R}^n , alors la quantité

$$O_Q(\mathfrak{F}, \{x^*\}) = \begin{cases} +\infty, & \text{si } Q_p(\mathfrak{F}, x^*) = 0, \quad \forall p \in [1, \infty[, \\ \inf\{p \in [1, \infty[, Q_p(\mathfrak{F}, x^*) = +\infty\}, & \text{sinon,} \end{cases}$$

est appelée l'ordre de Q (Q -order) du processus \mathfrak{F} au point x^* .

Définition 7.9 Soit un processus itératif \mathfrak{F} avec un point limite x^* , alors la quantité :

$$O_R(\mathfrak{F}, \{x^*\}) = \begin{cases} +\infty, & \text{si } R_p(\mathfrak{F}, x^*) = 0, \quad \forall p \in [1, \infty[, \\ \inf\{p \in [1, \infty[, R_p(\mathfrak{F}, x^*) = 1\}, & \text{sinon,} \end{cases}$$

est appelée l'ordre de R (R -order) du processus \mathfrak{F} au point x^* .

Lemme 7.1 [67] Soit \mathfrak{F} un processus itératif de point limite x^* , supposons qu'il existe $p \in [1, \infty[$ et une constante c_2 tel que, pour tout suite $\{x^k\}$ on a :

$$\|x^{k+1} - x^*\| \leq c_2 \|x^k - x^*\|^p, \quad \forall k \geq k_0$$

alors

$$O_R\{x^k\} \geq O_Q\{x^k\} \geq p.$$

Lemme 7.2 Soit $\{x^k\} \subset \mathbb{R}^n$ une suite convergente vers x^* , alors

$$R_1\{x^k\} \leq Q_1\{x^k\},$$

pour chaque norme. Par suite, si \mathfrak{F} est un processus itératif avec un point limite x^* , alors

$$R_1(\mathfrak{F}, x^*) \leq Q_1(\mathfrak{F}, x^*),$$

pour chaque norme.

Preuve.

Supposons que $Q_1\{x^k\} < \infty$, et on pose $\varepsilon_k = \|x^k - x^*\|$, alors, pour tout $\varepsilon > 0$ et $\gamma = Q_1\{x^k\} + \varepsilon$, il existe $k_0 \geq 0$ tel que

$$\varepsilon_k \leq \gamma \varepsilon_{k-1} \leq \dots \leq \gamma^{k-k_0} \varepsilon_{k_0}, \quad \forall k \geq k_0.$$

Par conséquent,

$$R_1\{x^k\} \leq \gamma \limsup_{k \rightarrow \infty} \left[\frac{\varepsilon_{k_0}}{\gamma^{k-k_0}} \right] = \gamma,$$

et comme ε est arbitraire, donc

$$R_1\{x^k\} \leq Q_1\{x^k\}.$$

Si $\mathbb{C}(\mathfrak{F}, x^*)$ est l'ensemble de toute les suites générées par \mathfrak{F} qui convergent vers x^* , on obtient

$$\begin{aligned} R_1(\mathfrak{F}, x^*) = \sup\{R_1\{x^k\}/\{x^k\} \in \mathbb{C}(\mathfrak{F}, x^*)\} &\leq \sup\{Q_1\{x^k\}/\{x^k\} \in \mathbb{C}(\mathfrak{F}, x^*)\} \\ &= Q_1(\mathfrak{F}, x^*). \end{aligned}$$

Preuve du lemme 6.1 du chapitre 6

La preuve est immédiate. Une simple induction montre que

$$\|x^{k+1} - x^*\| = \|G(x^k, h^k) - x^*\| \leq \tilde{\alpha} \|x^k - x^*\| \leq \dots \leq \tilde{\alpha}^{k+1} \|x^0 - x^*\|;$$

d'où, pour toute suite x^k contenue dans S et convergeant vers x . La première inégalité de la relation (6.11) est donnée par le lemme 7.2, la seconde découle de la relation (6.9) et de la définition de Q_1 .

Lemme 7.3 Soit $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ où \mathbb{R}^n et \mathbb{R}^m ont munis de la norme l_i -norme, $i = 1, 2, \infty$. Alors

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad (7.27)$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|, \quad (7.28)$$

et

$$\|A\|_2 = \sqrt{\lambda}, \quad (7.29)$$

où λ est le maximum des valeurs propres de $A^T A$.

Preuve.

Considérons la norme l_1 . Pour chaque $x \in \mathbb{R}^n$,

$$\begin{aligned} \|Ax\|_1 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| |x_j| \leq \sum_{j=1}^n |x_j| \sum_{i=1}^m |a_{ij}| \\ &\leq \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \|x\|_1. \end{aligned} \quad (7.30)$$

il est suffisant de montrer pour qu'un certain $x \in \mathbb{R}^n$ l'égalité (7.30) soit vraie. Soit k un indice pour lequel le maximum de (7.27) soit atteint, alors

$$\|Ae^k\|_1 = \sum_{i=1}^m |a_{ik}| = \max_{1 \leq i \leq m} |a_{ij}|.$$

$$\|A\| = \sup_{\|x\|=1} \|Ax\|'. \quad (7.31)$$

qui est le supremum de la relation (7.31), atteint pour la $k^{\text{ème}}$ composante du vecteur unitaire.

La preuve pour l_∞ -norme est similaire. On note seulement, dans ce cas, le supremum pour $x \in \mathbb{R}^n$ définie par :

$$x_i = \begin{cases} \frac{a_{ki}}{|a_{ki}|}, & a_{ki} \neq 0; \\ 1, & a_{ki} = 0, i = 1, \dots, n, \end{cases}$$

où k est un indice pour lequel le maximum de (7.30) est atteint.

Finalement, pour la l_2 -norme, on a $\|Ax\|_2 = (x^T A^T A x)^{\frac{1}{2}}$, et le résultat découle de la définition (7.1).

Lemme 7.4 Soit $\|\cdot\|$ une norme arbitraire sur \mathbb{R}^n (ou \mathbb{C}^n) et P une matrice réelle $n \times n$ (ou complexe) non singulière. Alors l'application est définie par

$$\|x\|' = \|Px\|, \text{ pour tout } x \in \mathbb{R}^n,$$

est une norme sur \mathbb{R}^n (ou \mathbb{C}^n). De plus, si $A \in L(\mathbb{R}^n)$, alors

$$\|A\|' = \|PAP^{-1}\|. \quad (7.32)$$

Preuve.

Il est facile de montrer que $\|\cdot\|'$ est une norme. Alors la relation (7.32) est déduite de

$$\|A\|' = \sup_{\|x\|'=1} \|Ax\|' = \sup_{\|Px\|=1} \|PAx\|' = \sup_{\|y\|=1} \|PAP^{-1}y\| = \|PAP^{-1}\|.$$

Corollaire 7.1 Soit $A \in L(\mathbb{C}^n)$, alors, pour tout $\epsilon > 0$, il existe une norme dans \mathbb{C}^n tel que

$$\|A\| \leq \rho(A) + \epsilon \quad (7.33)$$

Lemme 7.5 Soit $A \in L(\mathbb{C}^n)$. Alors $\lim_{k \rightarrow \infty} A^k = 0$ si et seulement si $\rho(A) < 1$.

Preuve.

Si $\rho(A) < 1$, alors, d'après le corollaire 7.1, il existe une norme telle que $\|A\| < 1$. Par conséquent, on a $\|A^k\| \leq \|A\|^k$, il s'en suit que $A^k \rightarrow 0$ quand $k \rightarrow \infty$. Par contre, supposons que A admet des valeurs propres λ tel que $|\lambda| \geq 1$ et des vecteurs propres correspondants $x \neq 0$. Alors $A^k x = \lambda^k x$ pour tout k , donc $A^k x$ ne tend pas vers zéro.

Lemme 7.6 (Lemme de Neumann)

Soit $B \in L(\mathbb{R}^n)$ et supposons que $\rho(B) < 1$ alors $(I - B)^{-1}$ existe et

$$(I - B)^{-1} = \lim_{k \rightarrow \infty} \sum_{i=1}^k B^i \quad (7.34)$$

Preuve.

Comme $\rho(B) < 1$, clairement $I - B$ n'admet pas des valeurs propres nulles d'où elle est inversible. Vérifions maintenant le relation (7.34), Notons que

$$(I - B)(I + \dots + B^k) = I - B^{k+1},$$

de plus

$$I + B + \dots + B^{k-1} = (I - B)^{-1} - (I - B)^{-1}B^k.$$

D'après le corollaire 7.6, la partie droite tend vers $(I - B)^{-1}$. Comme $(I - B)$ est inversible quand $\|B\| \leq 1$, d'après (7.34) on a

$$\|(I - B)^{-1}\| \leq \sum_{i=0}^{\infty} \|B\|^i = \frac{1}{1 - \|B\|}. \quad (7.35)$$

Lemme 7.7 (Lemme de Perturbation)

Soit $A, C \in L(\mathbb{R}^n)$ et supposons que A soit inversible avec $\|A^{-1}\| \leq \tilde{\alpha}$ si $\|A - C\| < \beta$; si de plus $\tilde{\alpha}\beta < 1$, alors C est aussi inversible et

$$\|C^{-1}\| \leq \frac{\tilde{\alpha}}{(1 - \tilde{\alpha}\beta)}.$$

Preuve.

On a $\|I - A^{-1}C\| = \|A^{-1}(A - C)\| \leq \tilde{\alpha}\beta < 1$ et $A^{-1}C = I - (I - A^{-1}C)$, d'après le lemme de (7.6) on déduit que $A^{-1}C$ est inversible. D'où, C est inversible. De plus, d'après la relation (7.35) on déduit que

$$\|C\| = \|[I - (I - A^{-1}C)]^{-1}A^{-1}\| \leq \tilde{\alpha} \sum_{i=0}^{\infty} (\tilde{\alpha}\beta)^i = \frac{\tilde{\alpha}}{(1 - \tilde{\alpha}\beta)}.$$

Lemme 7.8 Supposons que $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ soit G -différentiable sur un ensemble convexe $D_0 \subset D$. Alors, pour tout, $x, y \in D_0$,

$$\|F(y) - F(x)\| \leq \sup_{0 \leq t \leq 1} \|F'(x + t(y - x))\| \|x - y\|. \quad (7.36)$$

Preuve.

Supposons que $M = \sup_{0 \leq t \leq 1} \|F'(x + t(y - x))\| < \infty$ et, pour $\varepsilon > 0$, soit Γ est l'ensemble de $t \in [0, 1]$ pour lequel

$$\|F'(x + t(y - x)) - F'(x)\| \leq Mt\|y - x\| + \varepsilon t\|y - x\|, \quad (7.37)$$

est vérifiée. Clairement, $0 \in \Gamma$, et $\gamma = \sup_{t \in \Gamma} t$ est bien définie; comme la définition 7.5 implique que $F(x + t(y - x))$ est continue par rapport à t , on a

$$\|F(x + \gamma(y - x)) - F(x)\| \leq M\gamma\|y - x\| + \varepsilon\gamma\|y - x\|. \quad (7.38)$$

Comme ε est arbitraire, clairement le résultat est prouvé si $\gamma = 1$.

Supposons que $\gamma < 1$. Alors, comme F' existe au point $x + \gamma(y - x)$, il existe $\beta \in (\gamma, 1)$ tel que

$$\begin{aligned} \|F(x + \beta(y - x)) - F(x + \gamma(y - x)) - F'(x + \gamma(y - x))(\beta - \gamma)(y - x)\| \\ \leq \varepsilon(\beta - \gamma)\|y - x\|, \end{aligned}$$

d'où

$$\|F(x + \beta(y - x)) - F(x + \gamma(y - x))\| \leq M(\beta - \gamma)\|y - x\| + \varepsilon(\beta - \gamma)\|y - x\|.$$

Mais, d'après la relation (7.38), on a

$$\begin{aligned} \|F(x + \beta(y - x)) - F(x)\| &\leq (M\gamma + \varepsilon\gamma)\|y - x\| + (M + \varepsilon)(\beta - \gamma)\|y - x\| \\ &= (M + \varepsilon)\|y - x\|; \end{aligned}$$

qui est en contradiction avec la définition de γ . Donc la relation (7.37) est vérifiée pour $1 > \beta > \gamma$.

Lemme 7.9 Si $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est G -différentiable sur un ensemble convexe $D_0 \subset D$, alors, pour tout $x, y, z \in D_0$.

$$\|F(y) - F(z) - F'(x)(y - z)\| \leq \sup_{0 \leq t \leq 1} \|F'(z + t(y - z)) - F'(x)\| \|y - z\|. \quad (7.39)$$

Preuve .

Pour $x \in D_0$, on définit l'application

$$G(\omega) = F(\omega) - F'(x)\omega, \quad \omega \in D.$$

Les conditions de la définition 7.5 sont satisfaites pour G et comme

$$G'(\omega) = F'(\omega) - F'(x),$$

la relation (7.39) est précisément la relation

$$\|G(y) - G(z)\| \leq \sup_{0 \leq t \leq 1} \|G'(z + t(y - z))\| \|y - z\|.$$

Corollaire 7.2 Si $G : [a, b] \subset \mathbb{R}^1 \rightarrow \mathbb{R}^m$ est continue sur $[a, b]$, alors

$$\left\| \int_a^b G(t) dt \right\| \leq \int_a^b \|G(t)\| dt.$$

Preuve.

Comme la norme est une fonction continue, $\|G(\cdot)\|$ est intégrable au sens de Riemann, et, pour $\varepsilon > 0$ arbitraire, il existe une partition $a < t_0 < \dots < t_p < b$ tel que

$$\left\| \int_a^b G(t) dt - \sum_{i=1}^p G(t_i)(t_i - t_{i-1}) \right\| \leq \varepsilon,$$

et

$$\left| \int_a^b \|G(t)\| dt - \sum_{i=1}^p \|G(t_i)\|(t_i - t_{i-1}) \right| \leq \varepsilon.$$

D'où,

$$\left\| \int_a^b G(t) dt \right\| \leq \left\| \sum_{i=1}^p \|G(t_i)\|(t_i - t_{i-1}) \right\| + \varepsilon \leq \sum_{i=1}^p \|G(t_i)\|(t_i - t_{i-1}) + \varepsilon \leq \int_a^b \|G(t)\| dt + 2\varepsilon,$$

Théorème 17 Soit $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ continûment différentiable sur un ensemble convexe $D_0 \subset D$ et supposons que, pour des constantes $\tilde{\alpha} \geq 0$ et $p \geq 0$, F' satisfait

$$\|F'(u) - F'(v)\| \leq \tilde{\alpha} \|u - v\|^p, \quad u, v \in D_0.$$

Alors, pour tout $x, y \in D_0$

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq \left[\frac{\tilde{\alpha}}{(p+1)} \right] \|y - x\|^{p+1}.$$

Preuve.

D'après la définition 7.4 et le corollaire 7.2 on a

$$\begin{aligned}\|F(y) - F(x) - F'(x)(y - x)\| &= \left\| \int_0^1 [F'(x + t(y - x)) - F'(x)](y - x) dt \right\| \\ &\leq \int_0^1 \|F'(x + t(y - x)) - F'(x)\| \|y - x\| dt \\ &\leq \tilde{\alpha} \|y - x\|^{p+1} \int_0^1 t^p dt.\end{aligned}$$

Annexe 2

Théorème des fonctions implicites

En mathématiques, le théorème des fonctions implicites est un résultat de géométrie différentielle. Certaines courbes sont définies par une équation cartésienne, c'est-à-dire la forme $f(x, y) = 0$, où x et y décrivent les nombres réels. Le théorème indique que si la fonction f est suffisamment régulière au voisinage d'un point de la courbe, il existe une fonction φ de \mathbb{R} dans \mathbb{R} et au moins aussi régulière que f telle que localement, la courbe et le graphe de la fonction φ sont confondus.

Plus précisément, si (x_0, y_0) vérifie l'équation, si f est continûment différentiable et que sa dérivée partielle en (x_0, y_0) n'est pas nulle, alors il existe un voisinage de (x_0, y_0) sur lequel la zone s'identifie au graphe de φ .

On peut le traduire comme suit :

Soit $f(x, y)$ une fonction de classe C^p définie sur un ouvert U de \mathbb{R}^2 et à valeurs dans \mathbb{R} . Soit (x_0, y_0) un point de U tel que $f(x_0, y_0) = 0$ et tel que la dérivée partielle de f , par rapport à la deuxième variable, ne soit pas nulle en (x_0, y_0) . Il existe un ouvert V contenu dans U et contenant (x_0, y_0) et une fonction φ de classe C^p et de \mathbb{R} dans \mathbb{R} , tels que l'équivalence suivante soit vraie :

$$\forall (x, y) \in V \quad f(x, y) = 0 \Leftrightarrow \varphi(x) = y.$$

Bibliographie

- [1] M. Aidene, I.L.Vorob'ev, and B.Oukacha. Algorithm for solving a linear optimal control problem with minimax. *Performance Index, Computational Mathematics and mathematical Physics*, 45(10) :1691–1700, 2005.
- [2] J. Arnal, V. Migallon, and J. Penadés. Newton two-stage parallel iterative methods for nonlinear systems. *Spanish DGESIC*, 40(4) :001–004, 2000.
- [3] N.V. Balashevich, R. Gabasov, and F.M. Kirillova. Numerical methods of program and positional optimization of the linear control systems. *Zh. Vychisl. Mat. Mat. Fiz.*, 40(6) :838–859, 2000.
- [4] G.M. Baudet. Asynchronous iterative methods for multiprocesseurs. *Journal of ACM*, 25 :226–244, 1978.
- [5] R.E. Bellman. Dynamic programming. *Princeton University Press, Princeton, NJ*, 1963.
- [6] R.E. Bellman, I. Glicksberg, and O.A. Gross. Some aspects of the mathematical theory of control processes. *Report R-313, Rand Corporation, Santa Monica, CA*, 1958.
- [7] M. Bentobache. A new method for solving linear programming problems in canonical form and with bounded variables. *Master thesis, University of Bejaia, Algeria, (in French)*, 2005.
- [8] M. Bentobache. Nouvelle méthode pour la résolution des problèmes linéaire sous forme canonique et à variables bornées. *Thèse de magister, l'université A.Mira de Béjaia*, 2005.
- [9] M. Bentobache and M.O. Bibi. Two-phase support method for solving linear programming problems with nonneagitive variables : Numerical experiments. *Proceedings of COSI'08, University of Tizi-Ouzou (in French)*, pages 314–325, 2008.
- [10] M. Bentobache and M.O. Bibi. Two-phase support method for solving linear programming problems with bounded variables : Numerical experiments. *Proceedings of COSI'09, University of Annaba (in French)*, pages 109–120, 2009.
- [11] M. Bentobache and M.O. Bibi. Adaptive method with hybrid direction : theory and numerical experiments. *to appear in Proceedings of*

- Optimization 2011, Universidade nova de Lisboa, Portugal, 24-27 July 2011.*
- [12] D.P. Bertsekas and J. Tsitsiklis. Parallel and distributed computation. *Numerical Methods, Englewood cliffs : Prentice Hall, 1989.*
 - [13] M.O. Bibi. Support method for solving a linear-quadratic problem with polyedral constraints on control. *Optimization, 37(2) :139–147, 1996.*
 - [14] M.O. Bibi and M. Bentobache. The adaptive method with hybrid direction for solving linear programming problems with bounded variables. *Optimization, Prceedings of COSI'2011, University of Guelma, Algeria :80–91, 2011.*
 - [15] M.O. Bibi and M. Bentobache. An hybrid direction algorithm for solving linear programs. *Proceedings of DIMACOS'11, University of Mohamadia, Morocco, pages 28–30, 2011.*
 - [16] P. Borne, G. Dauphin-Tanguy, J.P. Richard, F. Rotella, and I. Zambettakis. Commande et optimisation des processus. *Editions Techniq, 1990.*
 - [17] B. Brahmi and M.O. Bibi. Dual support method for solving convex quadratic programs. *Optimisation, 59(6) :851–872, 2010.*
 - [18] R. Bru, V. Migallon, J. Penadés, and D. B. Szyld. Parralel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis, 3 :24–35, March 1995.*
 - [19] A.E. Bryson and Yu-Chi Ho. Applied optimal control. *Blaisdell, Toronto, Canada, 1969.*
 - [20] D.M. Cardoso and J.N. Climaco. The generalized simplex method. *Operations Research Letters 12, 3 :337–348, 1992.*
 - [21] D. Chazan and Miranker. Chaoric relaxation. *Linear Algebra Appl, 2 :199–222, 1969.*
 - [22] M. Chebbah. Résolution et implementation du problème min-max en contrôle optimale. *Thèse de magister, Tizi-Ouzou, 2006.*
 - [23] M. Crouzeix and A.L. Mignot. Analyse numérique des équations différentielles. *2ème Edition Masson, 1989.*
 - [24] G.B. Dantzig. Maximisation of a linear function of variables subject to linear inequalities. *in Koopmans RC (ed.) , Activity Analysis of Production and Allocation, Wiley, New-York, pages 339–347, 1951.*
 - [25] G.B. Dantzig. Linear programming and extensions. *Princeton University Press, Princeton, N.J., 1963.*
 - [26] J.P. Denailly. Analyse numérique et équations différentielles. *EDP Sciences, 2006.*
 - [27] E.Kostina. The long step rule in the bounded-variable dual simplex method : Numerical experiments. *Mathematical Methods of Operations Research 55, pages 413–429, 2002.*

- [28] Herman Erichson. Brachistochrone solution using fermat's principle of least time. *Eur. J. Phys.*, 20(5) :299–304, 1999.
- [29] F.Kara. Optimisation de problème de contrôle d'une commande polyédrale. *Thèse de magister, l'Université Mouloud Mammeri de Tizi-Ouzou*, 2008.
- [30] A. Frommer. Implementing brown's method for systems of nonlinear equations with dense banded jacobian. *Computing* 41, pages 123–129, 1989.
- [31] A. Frommer and D.B. Szyld. H-splitting and two-stage iterative methods. *Numerische Mathematik*, 63 :345–356, 1992.
- [32] R. Gabasov and F.M. Kirillova. Methods of linear programming. *in 3 parts, Edition of University Press, Minsk*, 63, (1977, 1978 and 1980).
- [33] R. Gabasov and F.M. Kirillova. Constructive methods of optimization part 2. control problems. *University Press, Minsk*, (*in Russian*), 39(4), 1984.
- [34] R. Gabasov and F.M. Kirillova. Adaptive method of solving linear programming problems. *Preprints Series of University of Karlsruhe, Institute for Statistics and Mathematics*, 1994.
- [35] R. Gabasov, F.M. Kirillova, and N.V. Balashevich. On the synthesis problem for optimal control systems. *SIAM J. Control OPTIM*, 39(4) :1008–1042, 2000.
- [36] R. Gabasov, F.M. Kirillova, and O.I. Kostyukova. Solution of linear quadratic extremal problems. *Soviet Mathematics Doklady*, 31 :99–103, 1985.
- [37] R. Gabasov, F.M. Kirillova, and N.S. Pavlenok. Optimization of dynamical systems by using dynamical controllers. *Autm. Telemekh.*, 5 :8–28, 2004.
- [38] P.E. Gill and W. Murray. A numerically stable form of the simplex algorithm. *Linear Algebra and its applications*, 7 :99–138, 1973.
- [39] P.E. Gill, W. Murray, and M.H. Wright. Numerical linear algebra and optimization. *Addison-Wesley Publishing Company, Redwood City, CA.94065*, 1, 1991.
- [40] J. Gondzio. Another simplex type method for large scale linear programming. *Control and Cybernetics*, 25(4) :739–760, 1996.
- [41] B. Grunn and Saeid Nahavandi. Why initial conditions are important. *Proceeding of the 2002 Winter Simulation Conference, School of Engineering and Technology, Deakin University, Geelong, Victoria 3217, Australia*, 2002.
- [42] E. Issamoto, F.T. Miki, J.I da Luz, J.D.da Silva, P.B. de Oliveira, and H.F. de Campos Velho. An inverse initial condition problem

- in heat conductions. *A Neural Network Approach, Braz. Cong. Mech. Eng. (COBEM), Proc. in CD-ROM paper code AAAGHA, Unicamp, Campinas (SP), Brazil*, 238, 1999.
- [43] R.E. Kalman. Mathematical description of linear dynamical systems. *J.SIAM Control*, 1 :152–192, 1963.
- [44] L.V. Kantorovich. Mathematical methods in the organisation and planning of production. *Publication House of the Leningrad State University, 1939*, Translated in management science 6 :366–422, 1960.
- [45] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica* 4, pages 373–395, 1984.
- [46] L.G. Khachiyan. A polynomial algorithm for linear programming. *Soviet Mathematics Doklady* 20, pages 191–194, 1979.
- [47] V. Klee and G.J. Minty. How good is the simplex algorithm. *In O. Shisha (Ed.), Inequalities III, Academic Press, New York*, pages 159–175, 1972.
- [48] E. Kostina and S.V. Prishchepova. A numerical experiments with respect to the solution of large sparse problems of linear programming by the adaptive method. *Izv. Akad. Nauk BSSR, Ser. Fiz-Mat. Nauk* 6, (in Russian), pages 3–5, 1990.
- [49] B. Lang, J.C. Miellou, and P. Spiteri. Asynchronous relaxation algorithms for optimal control. *Mathematics and computers in simulation, Elsevier Science Publishers B.V (North Halland)*, 1986.
- [50] B. Lang and P. Spiteri. Algorithmes de relaxation appliqués à la résolution de problèmes de commande optimale : Etude d'un système en anneau. 1983.
- [51] B. Lang and P. Spiteri. Application des méthodes de relaxation à la résolution de problèmes de commande optimale dans le cas des systèmes singulièrement perturbés. *Commande optimale de systèmes complexes*, 1983.
- [52] B. Lang and P. Spiteri. Application des méthodes de relaxation à la résolution de problèmes multicritères. *Commandes optimales de systèmes complexes*, 1983.
- [53] E.B. Lee and L. Markus. Foundations of optimal control theory. *Wiley and Sons, New York*, 1967.
- [54] I. Maros and G. Mitra. Strategies for creating advanced bases fr large-scale linear programming problems. *Inform Journal on Computing*, 10(2) :248–260, 1998.
- [55] P. Martinon. Résolution numérique de problèmes de contrôle optimal par une méthode homotopique simpliciale. Thèse, Chapitre 1 "Méthode de tir et homotopie". 2005.

- [56] H. Mauer, C. Buskens, J.H.R. Kim, and C.Y. Kaya. Optimization methods for the verification of second order sufficient conditions for bang-bang controls. *Journal optimal control, Applications and methods*, 26 :129–156, 2005.
- [57] A. Merakeb. Optimisation multicritères en contrôle optimal : Application au véhicule électrique. *Thèse de doctorat, l'université Mouloud Mammeri de Tizi-Ouzou*, 2011.
- [58] K. Messaoudene. Optimisation d'un système dynamique linéaire avec une trajectoire discontinue. *Thèse de magister, l'université Mouloud Mammeri de Tizi-Ouzou*, 2007.
- [59] J.C. Miellou and P. Spiteri. La notion de h-accrétivité : application à la résolution de problèmes aux deux bouts. *Commande Optimale de systèmes complexes*, 1982.
- [60] J.C. Miellou and P. Spiteri. A parallel asynchronous relaxation algorithm for optimal control problems. *Mathematical analysis and its applications*, Pergamon Press, 1985.
- [61] G. Mitra, M. Tamiz, and J. Yadegar. Experimental investigation of an interior search method within a simplex framework. *Communications of the ACM* 31, pages 1474–1482, 1988.
- [62] J.J. Moré and M.Y. Cosnard. Numerical solution of nonlinear equations. *ACM Transactions on Mathematical Software*, 5(1) :64–85, March 1979.
- [63] W.B. Muniz, F. Haroldo de Campos Velho, and Fernand Manuel Ramos. A comparison of some inverse methods for estimating the initial condition of the heat equation. *Journal of Computational and Applied Mathematics*, 103 :145–163, 1999.
- [64] W.B. Muniz, H.F. de Campos Velho, and F.M. Ramos. A comparison of some inverse methods for estimating the initial condition of the heat equation. *J.Comp.Appl.Math.*, 103, 145(1999).
- [65] A. Nait-Abdesselam. Résolution d'un problème de contrôle optimal des systèmes hybrides. *Thèse de magister, l'université Mouloud Mammeri de Tizi-Ouzou*, 2010.
- [66] A.J. Silva Neto and M.N. Özisik. An inverse heat conduction problem of estimating initial condition. *Proc. 12th Brazilian Congress of Mechanical Engineering, Brasilia*, 1993.
- [67] J.M. Ortega and W.C. Rheinboldt. Iterative solution of non linear equations in several variables. *Computer Science and Applied Mathematics*, 103, 1970.
- [68] B. Oukacha. Résolution de problème de contrôle optimal. *thèse de doctorat, l'université Mouloud Mammeri de Tizi-Ouzou*, 2006.
- [69] P.Q. Pan. A basis-deficiency-allowing variation of the simplex method. *Computers and Mathematics with Applications*, 36(3) :33–53, 1998.

- [70] K. Paparrizos, N. Samaras, and G. Stephanides. An efficient simplex type algorithm for sparse and dense linear programs. *European Journal of Operational Research*, 148 :323–334, 2003.
- [71] P.J.Lanzkron, D. J. Rose, and D.B. Szyld. Convergence of nested classical iterative methods for linear systems. *Numerische Mathematik*, 58, pages 685–702, 1991.
- [72] L. Poggialini. On local state optimality of bang-bang extremals in a free horizon bolza problem. *Rend.Sem.Mat.Univ.Pol.Torino, contr.th.and Stab.*, 63(4), 2005.
- [73] L.S. Pontriaguine, V.G. Boltyanski, R.V. Gamkrelidze, and E.F. Mischenko. The mathematical theory of optimal processes. *Interscience Publishers, New York*, 1962.
- [74] S. Radjef and M.O. Bibi. An effective generalization of the direct support method. *To appear in Mathematical Problems in Engineering*, 2011.
- [75] F.M. Ramos and A. Giovannini. Résolution d'un problème inverse multidimensionnel de diffusion de la chaleur par la méthode des éléments analytiques et par le principe de l'entropie maximale. *Internat.J.Heat Mass Transfer*, 38 :101–111, 1995.
- [76] F. Rezki. Résolution d'un problème terminal de contrôle optimal d'un système dynamique linéaire avec une trajectoire discontinue. *Thèse de magister, l'université Mouloud Mammeri de Tizi-Ouzou*, 2006.
- [77] R.Gabasov, F.M.Kirillova, and S.V.Prischepova. Optimal feedback control. *Springer-Verlag, London*, 1995.
- [78] R.Gabasov, F.M. Kirillova, and O.I. Kostyukova. A method for solving general linear programming problems. *Doklady AN BSSR (in Russian)*, 23(3) :179–200, 1979.
- [79] R.Gabasov, F.M. Kirillova, and N.S. Pavlenok. Optimal discrete-pulse of linear systems, automation and remote control. 69(3) :443–462, 2008.
- [80] F. Robert. Convergence locale d'itération chaotique non linéaire. *C.R.A.S série A*, 284 :679–682, 1977.
- [81] F. Robert, M. Charnay, and F. Musy. Itérations chaotiques série-parrallèle pour des équations non linéaires de point fixe. *Aplikace Matematik*, 20 :1–38, 1975.
- [82] R.Weinstock. Calculus of variations, chapitre 3. *Dover publications, Inc., New-York*, 1974.
- [83] P. Sabatier. Inverse problems-an introduction. *Inverse Problems 1*, pages 1–4, 1985.
- [84] A. Santos-Palomo. The sagitta method for solving linear programming. *European Journal of Operational Research*, 157(3) :572–539, 2004.

-
- [85] Elcio H. Shiguemori, José Demisio S. da Silva, and Haroldo F. de Campos Velho. Estimation of initial condition in heat conduction by neural network. *4th International Conference on Inverse Problems in Engineering, Rio de Janeiro, Brazil*, 2001.
- [86] B. Daniel Szyld and Jian-Jun Xu. Convergence of some asynchronous nonlinear multisplitting methods. Report99-6-30, June 1999.
- [87] T. Terlaky. A convergent criss-cross method. *A convergent criss-cross method*, pages 683–690, 1985.
- [88] E. Trelat. Contrôle optimal : Théorie et applications. *Université Paris-Sud, Laboratoire EN-EDP, Mathématiques, UMR 8628, Bat.425, 91405 Orsay Cedex*, 2005.
- [89] E. Trelat and J.M. Coron. Tout est sous contrôle. *Laboratoire de Mathématique, Equipe AN-EDP, Université Paris-Sud, MatAplic n°83*, pages 1–15, Juillet 2007.
- [90] Bruce van Brunt. The calculus of variations. *Springer-Verlag, Inc., New-York*, 2004.
- [91] S. Zionts. The criss-cross method for solving linear programming problems. *Management Science 15*, pages 426–445, 1969.