Ann. Sc. Math. Québec, 1977, vol. I, No. 2, pp. 211-225

APPROXIMATION DES EQUATIONS DIFFERENTIELLES ET PROBLEMES DE COMMANDE OPTIMALE*

Michel Delfour et François Trochu

1. Introduction

Dans cet article nous présentons quelques résultats sur l'approximation des solutions des équations différentielles par des éléments finis (EF) discontinus. On choisit un type d'approximation tel que dans le problème de commande optimale les solutions x et p du système d'optimalité soient toutes deux approchées par des EF discontinus. En faisant cela on retombe sur les méthodes de P. LESAINT [1] et LESAINT-RAVIART [1]. On récupère aussi plusieurs méthodes aux différences finies qui sont immédiatement généralisables au cas non-linéaires et entrent dans la famille des Runge-Kutta (cf. M. CROUZEIX [1]).

Mais l'utilisation des méthodes discontinues trouve aussi des applications intéressantes pour la commande optimale des systèmes à retard. En effet le noyau de la solution de l'équation différentielle opérationnelle de Riccati présente des propagatus de discontinuités. Comme l'approximation utilisée pour l'équation différentielle à retard se retrouve de façon naturelle dans celle du noyau de Riccati, ce genre d'approximation est tout à fait adéquat et donne de bons résultats (cf. M.C. DELFOUR [1] et DELFOUR-TROCHU [1]).

Pour en revenir au problème de commande, une fois le système approché discrétisé, il est possible d'utiliser la méthode de J.C. NEDELEC [1] pour construire des schémas d'approximation à Riccati qui seront d'ordre supérieur lorsque l'on augmente le degré des polynômes utilisés dans l'approximation de l'équation différentielle originale.

* Ces travaux ont été subventionnés par le programme F.C.A.C. du ministère de l'Education du Québec.

On trouvera des résultats plus complets sur le problème d'approximation dans TROCHU-DELFOUR [1].

Des essais numériques des méthodes introduites ici ont été présentés par A. CHALIFOUR [1].

Notations

Soit R le corps des nombres réels. Soient X et Y deux espaces de Hilbert sur R. Soit $\mathcal{L}(X,Y)$ l'espace de Banach des applications linéaires continues $T:X \to Y$, doté de la norme naturelle. Lorsque $X = Y$, on utilisera la notation $\mathcal{L}(X)$. L'application duale de T ($\in \mathcal{L}(X,Y)$) est un élément T^* de $\mathcal{L}(Y',X')$, où X' et Y' sont les duaux topologiques de X et Y, respectivement. Etant donné $a < b$, $L^p(a,b;X)$ sera l'espace de Banach des classes d'équivalence des fonctions $[a,b] \to X$ mesurables au sens de Lebesgue qui sont p-intégrables ($1 \le p < \infty$) ou essentiellement bornées ($p=\infty$). Lorsque $p=2$, le produit scalaire et la norme s'écriront $(\ ,\ )_2$ et $|\ |_2$. $H^k(a,b;X)$ sera l'espace de Sobolev des fonctions x dans $L^2(a,b;X)$ avec des dérivées distribution $D^i x$ dans $L^2(a,b;X)$, $1 \le i \le k$. On écrira le produit scalaire et la norme $(\ ,\ )_{H^k}$ et $|\ |_{H^k}$. Le produit de dualité d'un espace de Banach X et de son dual topologique X' s'écrira $\langle x,x' \rangle_X$.

2. Description du système et formulation du problème

2.1. Description du système

Soient $X = \mathbb{R}^n$ et $Y = \mathbb{R}^m$, où n et m sont des entiers positifs non nuls. On écrira $(\cdot,\cdot)$ (resp. $(\cdot,\cdot)_Y$) pour le produit scalaire et $|\cdot|$ (resp. $|\cdot|_Y$) pour la norme dans X (resp. Y). Soit $T > 0$ un nombre réel positif. On considère l'équation différentielle suivante dans $[0,T]$:

$$(2.1) \qquad \begin{cases} \dot{x} = A(t)x + B(t)v + f & \text{dans} \quad]0,T[\\ x(0) = x^0, \end{cases}$$

où $A:[0,T] \to \mathcal{L}(X)$ et $B:[0,T] \to \mathcal{L}(Y,X)$ sont des matrices mesurables et bornées. On suppose aussi que f appartient à $L^2(0,T;X)$ et que v est la fonction de commande dans $L^2(0,T;Y)$. Lorsqu'il n'y aura pas de confusion, on écrira L^2 ou $L^2(0,T)$ pour $L^2(0,T;Y)$.

2.2. Problème de commande optimale

On associe à la commande v et à la donnée initiale x^0 la fonction coût

$$(2.2)\quad J(v,x) = (Q_1x(T)+2q_1,x(T)) + \int_0^T [(Q_0(t)x(t)+2q_0(t),x(t))+(N(t)v(t),v(t))_U]dt,$$

où $q = (q_0,q_1) \in L^2 \times X$, Q_1 et $Q_0(t)$ (resp. $N(t)$) sont des matrices semi-définies (resp. définies) positives de $\mathcal{L}(X)$ (resp. $\mathcal{L}(Y)$). Q_0 et N sont mesurables et bornées par rapport à t dans [0,T]. De plus

$$(2.3)\qquad \exists\ c > 0,\quad \forall\ v \in Y,\quad \forall\ t \in [0,T],\qquad (N(t)v,v)_U \geq c|v|_U^2.$$

Etant donné x^0, le problème de commande optimale est le suivant:

$$(2.4)\qquad \text{Inf}\{J(v,x)\colon\ v \in L^2(0,T;Y),\ x \text{ solution de (2.1) pour } x^0\}.$$

Ce problème possède une solution unique u dans $L^2(0,T;Y)$ qui est complètement caractérisée par le système d'optimalité suivant

$$(2.5)\qquad \begin{cases} \dot{x} = A(t)x + B(t)u + f, & x(0) = x^0 \\ \dot{p} + A^*(t)p + Q_0(t)x + q_0 = 0, & p(T) = Q_1x(T) + q_1 \\ u = -N^{-1}B^*p. & \end{cases}$$

Pour "découpler" le système (2.5) on considère (2.1) et le problème (2.4) dans l'intervalle [s,T] pour un s quelconque dans [0,T[:

$$(2.6)\qquad \dot{x} = A(t)x + B(t)v + f \text{ dans } [s,T],\quad x(s) = x^0$$

$$(2.7)\quad J_s(v,x) = (Q_1x(T)+2q_1,x(T)) + \int_s^T [(Q_0(t)x(t)+2q_0(t),x(t))+(N(t)v(t),v(t))_U]dt.$$

De nouveau la solution optimale u_s est complètement caractérisée par le système d'optimalité

$$(2.8)\qquad \begin{cases} \dot{x} = A(t)x + B(t)u_s + f, & x(s) = x^0 \\ \dot{p} + A^*(t)p + Q_0(t)x + q_0 = 0, & p(T) = Q_1x(T) + q_1 \\ u_s = -N^{-1}B^*p. & \end{cases}$$

On sait qu'il existe une famille de matrices symétriques $\{\Pi(t)\colon 0 \leq t \leq T\}$ dans $\mathcal{L}(X)$ et une famille de vecteurs $\{r(t)\colon 0 \leq t \leq T\}$ dans X telles que

$$(2.9)\qquad p(t) = \Pi(t)x(t) + r(t) \text{ dans } [0,T] \text{ (resp. } [s,T]).$$

Ces familles de matrices et de vecteurs sont solutions des équations:

(2.10) $$\frac{d\Pi}{dt} + A^*\Pi + \Pi A + Q_0 - \Pi R\Pi = 0, \quad \Pi(T) = Q_1, \quad R = BN^{-1}B$$

(2.11) $$\frac{dr}{dt} + [A-R\Pi]^*r + \Pi f + q_0 = 0, \quad r(T) = q_1.$$

On montre aussi que pour f=0, $q_0=0$ et $q_1=0$ dans (2.8)

(2.12) $$J_s(u_s,x) = (\Pi(s)x^0,x^0) = (p(s),x^0).$$

3. Approximation par éléments finis discontinus

On veut maintenant approcher les solutions du système d'optimalité (2.8) et la solution de l'équation différentielle matricielle de Riccati (2.10) à la façon de J.C. NEDELEC [1]. On utilisera pour nos approximations des fonctions polynomiales par morceaux. On fera tout cela de telle façon que les fonctions constantes par morceaux puissent être à la fois utilisées pour les deux équations du système d'optimalité. On utilisera comme outil principal un théorème fondamental sur les approximations du type Galerkin. La partie délicate du travail réside donc dans le choix des bons espaces et des bonnes fonctions.

3.1. Le théorème d'approximation du type Galerkin

Soit $\Gamma: \mathcal{U} \to \mathcal{V}'$ une bijection linéaire continue d'un espace de Hilbert $\mathcal{U}$ dans le dual topologique $\mathcal{V}'$ d'un autre espace de Hilbert $\mathcal{V}$. Alors il existe des constantes $m > 0$ et $M > 0$ telles que

$$\forall\ u \in \mathcal{U}, \quad m\|u\|_{\mathcal{U}} \le \|\Gamma u\|_{\mathcal{V}'} \le M\|u\|_{\mathcal{U}}.$$

Soient $\mathcal{U}_h$ et $\mathcal{V}_h$ deux sous espaces fermés de $\mathcal{U}$ et $\mathcal{V}$, respectivement. On associe à un élément ℓ de $\mathcal{V}'$ les problèmes suivants:

(3.1) trouver $u \in \mathcal{U}$ tel que $\forall\ v \in \mathcal{V}$, $\langle \Gamma u,v \rangle_{\mathcal{V}} = \langle \ell,v \rangle$,

(3.2) trouver $u_h \in \mathcal{U}_h$ tel que $\forall\ v_h \in \mathcal{V}_h$, $\langle \Gamma u_h,v_h \rangle = \langle \ell,v_h \rangle$.

Par hypothèse sur Γ, le problème (3.1) possède une solution unique $u = \Gamma^{-1}\ell$ dans $\mathcal{U}$. Si pour tout ℓ dans $\mathcal{V}'$, le problème (3.2) possède une solution unique $u_h \in \mathcal{U}_h$, alors il existe des constantes $m_h > 0$ et $M_h > 0$ telles que

(3.3) $$\forall\ u_h \in \mathcal{U}_h,\quad \forall\ v_h \in \mathcal{V}_h,\quad |\langle \Gamma u_h, v_h \rangle_{\mathcal{V}}| \le M_h \|u_h\|_{\mathcal{U}} \|v_h\|_{\mathcal{V}},$$

(3.4) $$\operatorname{Inf}\{\operatorname{Sup}[|\langle \Gamma u_h, v_h \rangle_{\mathcal{V}}|: v_h \in \mathcal{V}_h,\ \|v_h\|_{\mathcal{V}} \le 1]: u_h \in \mathcal{U}_h,\ \|u_h\|_{\mathcal{U}} = 1\} \ge m_h,$$

(3.5) $$\forall\ 0 \ne v_h \in \mathcal{V}_h,\quad \operatorname{Sup}\{|\langle \Gamma u_h, v_h \rangle_{\mathcal{V}}|: u_h \in \mathcal{U}_h\} > 0.$$

Théorème 3.1. Avec les notations et hypothèses précédentes, on a

$$\|u-u_h\|_{\mathcal{U}} \le (1+M/m_h)\ \|u-w\|_{\mathcal{U}} \quad \forall\ w \in \mathcal{U}_h.$$

Démonstration. Cf. BABUSKA-AZIZ [1]. □

3.2. Construction de l'application Γ

Nous associons au système (2.1) l'application

(3.7) $$x \to \Gamma x = (-\dot{x}+Ax, -x(0)): \mathcal{X} = H^1(0,T;X) \to \mathcal{Y} = L^2(0,T;X)\times X,$$

où $(Ax)(t) = A(t)x(t)$. L'équation (2.1) est complètement équivalente à

(3.8) $$\Gamma x + (Bv,0) + (f,x^0) = 0,$$

où $(Bv)(t) = B(t)v(t)$. Par construction nous avons une bijection linéaire et continue. Si l'on applique le Théorème 3.1 à Γ on s'aperçoit que l'on devra choisir nos approximations dans $H^1(0,T;X)$; ceci élimine donc d'office les approximations discontinues. Pour contourner cette difficulté il est nécessaire de faire un relèvement de la bijection Γ en une bijection $\bar{\Gamma}$ définie sur un plus gros espace U ($L^2(0,T;X)\times X$ par exemple). On se donne donc un espace de Hilbert U (plus gros que $\mathcal{X}$) et une injection continue

(3.9) $$i: \mathcal{X} \to U$$

dont l'image est dense dans U. On utilise ensuite le lemme suivant:

Lemme 3.2. Soient $\mathcal{X}$, $\mathcal{Y}$ et U trois espaces de Hilbert sur R, une injection $i: \mathcal{X} \to U$ dont l'image est dense dans U et une bijection linéaire continue $\Gamma: \mathcal{X} \to \mathcal{Y}$. Alors il existe un espace de Hilbert V sur R et une injection $j: V \to \mathcal{Y}'$ dont l'image est dense dans $\mathcal{Y}'$ (le dual topoligique de $\mathcal{Y}$) et une bijection linéaire continue $\bar{\Gamma}: U \to V'$ (le dual topologique de V) telle que

(3.10) $$j^*\Gamma = \bar{\Gamma} i.$$

Démonstration. Cf. TROCHU-DELFOUR [1] et DELFOUR-TROCHU [1]. □

Prenons par exemple l'application Γ de (3.7) avec

$$\mathcal{Y} = U = L^2(0,T;X)\times X \quad x \to i(x) = (x,x(T)). \tag{3.11}$$

On peut alors montrer que $V = H^1(0,T;X)$,

$$j(v) = (v,v(0)), \quad \bar{\Gamma}^*v = (\dot{v}+A^*v,-v(T)), \tag{3.12}$$

où $(A^*v)(t) = A(t)^*v(t)$. On peut alors réécrire l'équation (3.8) sous la forme

$$\bar{\Gamma}ix = j^*\Gamma x = -j^*((Bv+f,x^0)), \quad ix \in U. \tag{3.13}$$

On voit maintenant que dans ce cadre on peut utiliser des approximations discontinues $U_h \subset U = L^2(0,T;X)\times X$ pour approcher $ix \in U$ par le Théorème 3.1.

Il est évident que d'autres choix de $\mathcal{X}$, $\mathcal{Y}$, Γ, i et U sont possibles, mais nous allons nous limiter ici au cas discontinu.

Cependant l'objectif ultime est d'utiliser tout cela pour le problème de commande optimale (2.5)-(2.10)-(2.11). On verra que cela nous contraindra à modifier encore le problème (2.1) pour que l'équation en x dans (2.5) soit approchée par le même type d'approximation que l'équation en p dans (2.5).

On retourne donc au problème de commande optimale (2.1)-(2.4). On peut le réécrire sous la forme plus compacte

$$\Gamma x + (Bv+f,x^0) = 0, \quad J(v,x) = [Qix+2q,ix] + (Nv,v)_2, \tag{3.14}$$

où [,] indique le produit scalaire dans $U = L^2(0,T;X)\times X$, $Q \in \mathcal{L}(U)$ et $q \in U$

$$Q(u_0,u_1) = (Q_0u_0,Q_1u_1), \quad q = (q_0,q_1), \quad (Nv)(t) = N(t)v(t). \tag{3.15}$$

La commande optimale v^* correspondant au problème de minimisation (3.14) est caractérisée par l'équation d'Euler:

$$[Qix+q,i(y-x)] + (Nv^*,v-v^*)_2 = 0, \quad \forall\, v \in L^2(0,T;Y), \tag{3.16}$$

où y (resp. x) est la solution de la première équation (3.14) associée à v (resp. v^*) et x^0. On simplifie l'équation (3.16) en introduisant le système adjoint

$$\bar{\Gamma}^*p + Qix + q = 0. \tag{3.17}$$

Si l'on écrit $[\ ,\ ]_Y$ le produit scalaire dans Y, on obtient

$$(3.18)\qquad 0 = -[\bar{\Gamma}^*p, i(y-x)] + (Nv^*, v-v^*)_2 = -[jp, \Gamma(y-x)]_Y + (Nv^*, v-v^*)_2$$
$$= (B^*p+Nv^*, v-v^*)_2.$$

Enfin la commande optimale v^* est complètement caractérisée par

$$(3.19)\qquad \Gamma x + (Bv^*+f, x^0) = 0, \quad \bar{\Gamma}^*p + Qix + q = 0, \quad v^* = -N^{-1}B^*p$$

ou les systèmes relevés

$$(3.20)\qquad \bar{\Gamma}ix + j^*((Bv^*+f, x^0)) = 0, \quad \Gamma^*jp + i^*(Qix+q) = 0, \quad v^* = -N^{-1}B^*p.$$

<u>Les remarques suivantes sont d'une importance capitale et motivent le développement subséquent de ce travail</u>.

On considère le système approché correspondant à deux sous-espaces $U_h \subset U$ et $V_h \subset V$ et le système relevé

$$(3.21)\qquad \bar{\Gamma}ix + j^*((Bv+f, x^0)) = 0, \quad J(v, ix) = [Qix+2q, ix] + (Nv, v)_2.$$

On obtient le système approché

$$(3.22)\qquad u_h \in U_h, \quad \langle \bar{\Gamma}u_h + j^*((Bv+f, 0)+F\phi), v_h \rangle_V = 0, \quad \forall\, v_h \in V_h$$

et la fonction coût approchée

$$(3.23)\qquad J(v, u_h) = [Qu_h+2q, u_h] + (Nv, v)_2.$$

La commande minimisante v_h^* est complètement caractérisée par l'équation d'Euler

$$(3.24)\qquad [Qu_h+q, w_h-u_h] + (Nv_h^*, v-v_h^*)_2 = 0 \quad \forall\, v \in L^2(0,T;Y),$$

où w_h (resp. u_h) est la solution de (3.22) correspondant à v (resp. v_h^*). On est alors obligé de choisir le système adjoint approché suivant

$$(3.25)\qquad p_h \in V_h, \quad [\bar{\Gamma}^*p_h + Qu_h + q, w_h] = 0, \quad \forall\, w_h \in U_h.$$

Le premier terme de l'équation (3.24) devient

$$(3.26)\qquad [-\bar{\Gamma}^*p_h, w_h-u_h] = -\langle p_h, \bar{\Gamma}w_h - \bar{\Gamma}u_h \rangle_V = -[jp_h, (B(v-v^*), 0)]_Y$$

et nécessairement

$$v_h^* = -N^{-1}B^*p_h. \tag{3.27}$$

Le système d'optimalité est donné par les équations (3.22) avec $v = v_h^*$ et (3.25)-(3.27). Il est fondamental de comprendre qu'en choisissant le sous-espace $U_h \subset U$, nous sommes contraints de prendre un sous-espace $V_h \subset V$ pour l'approximation de la variable adjointe p. Il en résulte que p devra être approché par un sous-espace de $H^1(0,T;X)$; ceci élimine les approximations discontinues pour p. On ne pourrait pas obtenir (3.26) si l'on cherchait $y_h \in \mathcal{Y}_h \subset \mathcal{Y}$ tel que

$$\langle \Gamma^* y_h + i^*(Qu_h+q), z_h \rangle_{\mathcal{X}} = 0, \quad \forall\, z_h \in \mathcal{X}_h, \tag{3.28}$$

pour un sous-espace approprié $\mathcal{X}_h$ de $\mathcal{X}$ de dimension finie, à moins que l'on choisisse $U_h = i\mathcal{X}_h$. Dans ce dernier cas, on doit prendre dans l'équation (3.24)

$$w_h = iz_h \quad \text{et} \quad u_h = ix_h$$

pour z_h et x_h dans $\mathcal{X}_h$ et le premier terme de cette équation devient

$$\begin{aligned}[Qu_h+q, i(z_h-x_h)] = \langle i^*(Qu_h+q), z_h-x_h \rangle_{\mathcal{X}} &= -\langle \Gamma^* y_h, z_h-x_h \rangle_{\mathcal{X}} \\ &= [y_h, (B(v-v^*),0)]_{\mathcal{Y}}\end{aligned}$$

et l'on obtient l'analogue de (3.27). En conclusion si l'on approche j_p dans le sous-espace $Y_h \subset Y = L^2(0,T;X)\times X$, nous sommes contraints à approcher x dans le sous-espace $\mathcal{X}_h \subset \mathcal{X} = H^1(0,T;X)$; ceci élimine les approximations discontinues pour x.

Pour éliminer ces difficultés, nous allons introduire un nouvel espace $\mathcal{X}$ et une nouvelle application Γ. Ce sera la formulation discontinue.

3.3. Construction de Γ à partir d'une formulation discontinue

Jusqu'à présent nous n'avons pas pu approcher de façon simultanée le système d'optimalité en x et p par des éléments finis discontinus. Nous remédions à cette difficulté en construisant une nouvelle application Γ de façon à ce que les espaces associés $\mathcal{X}$ et $\mathcal{Y}$ contiennent déjà des fonctions discontinues.

On subdivise l'intervalle $[0,T]$ en E sous-intervalles $\{J_e: e=1,\ldots,E\}$ et l'on décompose le problème original en E problèmes. Soit $J_e = [t_{e-1},t_e]$, $e=1,\ldots,E$, pour un ensemble ordonné de points $0 = t_0 < t_1 <\ldots< t_E = T$. Nous considérons l'équation (2.1) dans chaque intervalle J_e. On obtient alors E problèmes:

$$\left.\begin{cases} \dot{x}_e = Ax_e + Bu + f \text{ dans } J_e \\ x_e(t_{e-1}) - x_{e-1}(t_{e-1}) = 0 \quad (x_1(0) = x^0 \text{ pour } e=1) \end{cases}\right\} e=1,\dots,E. \tag{3.29}$$

L'application

$$\Gamma_0: \mathcal{X}_0 = \sum_{e=1}^{E} H^1(J_e;X) \to \mathcal{Y}_0 = L^2(0,T;X) \times X^E$$

est obtenue en mettant les E problèmes bout à bout

$$\Gamma_0(x_1,\dots,x_E) = (\sum_{e=1}^{E} [-\dot{x}_e+Ax_e]\chi_e, -x_1(0), -\{[\![x]\!]_e\}_{e=1}^{E-1}), \tag{3.30}$$

où χ_e est la fonction caractéristique de l'intervalle J_e et

$$[\![x]\!]_e = x_{e+1}(t_e) - x_e(t_e) \tag{3.31}$$

est le "saut au point t_e". Ce problème est évidemment plus général que (2.1) puisque l'on peut introduire des données qui créent des sauts en x aux instants t_e, $e=1,\dots,E-1$. On peut obtenir quelque chose d'encore plus général en ajoutant des sauts à t_0 et t_E:

$$\begin{cases} \Gamma_1: \mathcal{X}_1 = [\sum_{e=1}^{E} H^1(J_e;X)] \times X \to \mathcal{Y}_1 = L^2(0,T;X) \times X^{E+1} \\ \Gamma_1(x_1,\dots,x_E,x_{E+1}) = (\sum_{e=1}^{E} [-\dot{x}_e+Ax_e]\chi_e, -x_1(0), -\{[x]_e\}_{e=1}^{E-1}, x_E(t_E)-x_{E+1}) \end{cases} \tag{3.32}$$

$$\begin{cases} \Gamma_2: \mathcal{X}_2 = X \times \prod_{e=1}^{E} H^1(J_e;X) \to \mathcal{Y}_2 = L^2(0,T;X) \times X^{E+1} \\ \Gamma_2(x_0,x_1,\dots,x_E) = (\sum_{e=1}^{E} [-\dot{x}_e+Ax_e]\chi_e, -x_0, x_0-x_1(0), -\{[x]_e\}_{e=1}^{E-1}). \end{cases} \tag{3.33}$$

On associe à chaque Γ_k un espace U_k et un injection i_k: $\mathcal{X}_k \to U_k$ ce qui déterminera un espace V_k et une injection j_k: $V_k \to Y_k$. Evidemment le choix doit être fait de façon à ce que les éléments de V_k soient discontinus aux points t_e, $e=1,\dots,E-1$. La façon la plus simple (et la plus symétrique) de faire cela est de choisir $U_k = \mathcal{Y}_k'$, k=0,1,2 et les i_k correspondantes. On obtient alors les applications Γ_0, Γ_1 et Γ_2 suivantes:

$$(3.34)\quad \begin{cases} i_0(x_1,\dots,x_E) = \left(\sum_{e=1}^{E} x_e\chi_e, \{\overline{x(t_e)}^{\alpha}\}_{e=1}^{E-1}, x_E(t_E)\right) : X_0 \to U_0 = L^2(0,T;X)\times X^E \\ j_0(v_1,\dots,v_E) = \left(\sum_{e=1}^{E} v_e\chi_e, v_1(0), \{\overline{v(t_e)}^{(1-\alpha)}\}_{e=1}^{E-1}\right) : V_0 = \prod_{e=1}^{E} H^1(J_e;X) \to Y_0 \\ \bar{\Gamma}_0^*(v_1,\dots,v_E) = \left(\sum_{e=1}^{E} [\dot{v}_e + A^*v_e]\chi_e, \{[\![v]\!]_e\}_{e=1}^{E-1}, -v_E(t_E)\right) : V_0 \to U_0, \end{cases}$$

où

$$\overline{x(t_e)}^{\alpha} = \alpha x_{e-1}(t_e) + (1-\alpha)x_e(t_e) \text{ et } \overline{v(t_e)}^{(1-\alpha)} = (1-\alpha)v_{e-1}(t_e) + \alpha v_e(t_e),\ e=1,\dots,E-1,$$

$$(3.35)\quad \begin{cases} i_1(x_1,\dots,x_{E+1}) = \left(\sum_{e=1}^{E} x_e\chi_e, x_1(0), \{\overline{x(t_e)}^{\alpha}\}_{e=1}^{E-1}, \alpha x_E(t_E) + (1-\alpha)x_{E+1}\right) \\ \qquad : X_1 \to U_1 = L^2(0,T;X)\times X^{E+1} \\ j_1(v_0,\dots,v_E) = \left(\sum_{e=1}^{E} v_e\chi_e, v_0, \{\overline{v(t_e)}^{(1-\alpha)}\}_{e=1}^{E-1}, (1-\alpha)v_E(t_E)\right) \\ \qquad : V_1 = X\times\prod_{e=1}^{E} H^1(J_e;X) \to Y_1 \\ \bar{\Gamma}_1^*(v_0,\dots,v_E) = \left(\sum_{e=1}^{E} [\dot{v}_e + A^*v_e]\chi_e, v_1(0)-v_0, \{[\![v]\!]_e\}_{e=1}^{E-1}, -v_E(t_E)\right) \\ \qquad : V_1 \to U_1 \end{cases}$$

$$(3.36)\quad \begin{cases} i_2(x_0,\dots,x_E) = \left(\sum_{e=1}^{E} x_e\chi_e, \alpha x_0 + (1-\alpha)x_1(0), \{\overline{x(t_e)}^{\alpha}\}_{e=1}^{E-1}, x_E(t_E)\right) \\ \qquad : X_2 \to U_2 = L^2\times X^{E+1} \\ j_2(v_0,\dots,v_E) = \left(\sum_{e=1}^{E} v_e\chi_e, v_0, (1-\alpha)v_0 + \alpha v_1(0), \{\overline{v(t_e)}^{(1-\alpha)}\}_{e=1}^{E-1}\right) \\ \qquad : V_2 = X\times\prod_{e=1}^{E} H^1(J_e;X) \to Y_2 \\ \bar{\Gamma}_2^*(v_0,\dots,v_E) = \left(\sum_{e=1}^{E} [\dot{v}_e + A^*v_e]\chi_e, v_1(0)-v_0, \{[\![v]\!]_e\}_{e=1}^{E-1}, -v_E(t_E)\right) : V_2 \to U_2. \end{cases}$$

On peut évidemment faire d'autres choix qui mèneront à différents $\bar{\Gamma}$ et à d'autres schémas.

Les trois cas considérés ont la même structure et nous pouvons oublier les indices 1,2,3 dans ce qui suit. Pour spécifier le problème approché, on choisit des sous-espaces $X_h \subset X$ et $V_h \subset V$. On applique le Théorème 3.1 à $\bar{\Gamma}: U \to V'$ avec

$$U_h = i(X_h) \subset U \quad \text{et} \quad V_h \subset V.$$

Ensuite on applique le même théorème à $\Gamma^*: \mathcal{Y} \to \mathcal{X}'$ avec cette fois

$$\mathcal{Y}_h = j(V_h) \subset \mathcal{Y} \quad \text{et} \quad \mathcal{X}_h \subset \mathcal{X}:$$

primal	dual (ou adjoint)
$\Gamma: \mathcal{X} \to \mathcal{Y} \equiv \mathcal{Y}'$, $i: \mathcal{X} \to U$	$\bar{\Gamma}^*: V \to U' \equiv U$, $j: V \to \mathcal{Y}$
$\bar{\Gamma}: U \to V'$	$\Gamma^*: \mathcal{Y} \to \mathcal{X}'$
$U_h = i(\mathcal{X}_h) \subset U$, $V_h \subset V$	$\mathcal{Y}_h = j(V_h) \subset \mathcal{Y}$, $\mathcal{X}_h \subset \mathcal{X}$
$\Gamma x = \ell$ dans $\mathcal{Y}$	$\bar{\Gamma}^* p = \ell$ dans U
$[\Gamma x_h, jv_h]_{\mathcal{Y}} = [\ell, jv_h]_{\mathcal{Y}}, \quad \forall\, v_h \in V_h$	$[\bar{\Gamma}^* p_h, iy_h]_U = [\ell, iy_h]_U, \quad \forall\, y_h \in \mathcal{X}_h$.

On remarquera que les matrices associées aux discrétisations de Γ et $\bar{\Gamma}^*$ sont adjointes l'une de l'autre. Ce cadre permet donc l'approximation simultanée des solutions des problèmes primal et dual dans les espaces choisis U et $\mathcal{Y}$ (c'est-à-dire globalement dans la norme L^2 et localement aux noeuds du maillage).

3.4. Approximation du problème de commande optimale

Ici on se borne au cas 1 (équations (3.32) et (3.35)) et l'on laisse tomber l'indice 1 pour Γ, i, U, j, V, $\mathcal{X}$ et $\mathcal{Y}$. Le problème de commande optimale s'écrit

$$J(v,x) = [Qix+2q, ix] + (Nv,v)_2, \tag{3.37}$$

où $[\cdot,\cdot]$ est le produit scalaire dans $U = L^2(0,T;X) \times X^{E+1}$, $Q \in \mathcal{L}(U)$ et $q \in U$

$$Q(u,u_0,\ldots,u_E) = (Q_0 u, 0, \ldots, 0, Q_1 u_E), \quad q = (q_0, 0, \ldots, 0, q_1). \tag{3.38}$$

Le problème de commande optimale peut maintenant être reformulé comme suit:

$$\operatorname{Inf}\{J(v,x): v \in L^2(0,T;Y),\ \bar{\Gamma} ix + j^*((Bv+f, x^0, 0, \ldots, 0))\}. \tag{3.39}$$

La commande optimale v^* dans $[0,T]$ est complètement caractérisée par le système d'optimalité

$$\Gamma x + (Bv^*+f, x^0, \ldots, 0) = 0, \quad \bar{\Gamma}^* p + Qix + q = 0, \quad v^* = -N^{-1}B^* p. \tag{3.40}$$

En utilisant l'identité $\bar{\Gamma} i = j^*\Gamma$ on peut aussi écrire

$$\bar{\Gamma} ix + j^*((Bv^*+f, x^0, \ldots, 0)) = 0, \quad \Gamma^* jp + i^*[Qix+q] = 0, \quad v^* = -N^{-1}B^* p. \tag{3.41}$$

Des expressions analogues peuvent être obtenues pour la commande optimale v_s dans $[s,T]$:

(3.42) $\bar{\Gamma}_s ix + j^*((Bv_s+f,x^0,\ldots,0)) = 0, \quad \Gamma_s^* jp + i^*[Qix+q] = 0, \quad v_s = -N^{-1}B^*p.$

Finalement on doit faire l'approximation du problème

$$\begin{cases} \bar{\Gamma}_s ix - j^*Rjp = -j^*(f,x^0,\ldots,0) \\ i^*Qix+\Gamma_s^* jp = -i^*q, \end{cases} \tag{3.43}$$

où $R \in \mathcal{L}(V)$ est défini par

$$R(y,y_0,\ldots,y_E) = (BN^{-1}B^*y,0,\ldots,0). \tag{3.44}$$

4. Le système approché et le problème de commande optimale approché

4.1. Espaces d'approximation et erreurs d'interpolation

On introduit un nombre réel $h > 0$

$$h = \max\{\text{mesure } (J_e): e=1,\ldots,E\}$$

et les sous-espaces $X_i^h \subset X_i$, $V_i^h \subset V_i$, $i=0,1,2$:

$$\begin{cases} X_0^h = V_0^h = \{(x_1,\ldots,x_E): x_e \in P^k,\ e=1,\ldots,E\} \\ X_1^h = X_0^h \times X,\ V_1^h = X \times V_0^h \\ X_2^h = V_2^h = X \times X_0^h, \end{cases} \tag{4.1}$$

où P^k est l'espace des polynômes de degré inférieur ou égal à k. On définit pour chaque x dans X_j, l'erreur d'interpolation

$$I_j(x) = \text{Inf}\{\|i_j(x-x^h)\|_{Y_j} : x^h \in X_j^h\}. \tag{4.2}$$

Théorème 4.1. (i) Pour $k \geq 1$, $x_0 \in X$, $x_{E+1} \in X$ et $x_e \in H^{k+1}(J_e)$, $e=1,\ldots,E$, il existe une constante $c>0$ (indépendante de h, lorsque h tend vers zéro) telle que

$$(4.3)\qquad \begin{aligned} I_0(x) &\le c\ h^{k+1}\|x\|_{k+1} \\ I_1(x) &\le c\ h^{k+1}[\|x\|^2_{k+1}+|x_{E+1}|^2]^{\frac{1}{2}} \\ I_2(x) &\le c\ h^{k+1}[\|x\|^2_{k+1}+|x_0|^2]^{\frac{1}{2}}, \end{aligned}$$

où

$$(4.4)\qquad \|x\|^2_{k+1} = \sum_{e=1}^{E} \|x_e\|^2_{H^{k+1}(J_e)}.$$

(ii) Soit $k=0$ et $x_e \in H^1(J_e)$, $e=1,\dots,E$. Alors

$$(4.5)\qquad \begin{cases} I_j(x) \le c\ h\|x\|_{X_j}, & \frac{1}{2} \le \alpha \le 1, \quad j=0 \text{ ou } 2 \\ I_1(x) \le c\ h\|x\|_{X_1}, & 0 \le \alpha \le \frac{1}{2}. \quad \square \end{cases}$$

4.2. Solution des problèmes approchés.

Dans ce qui précède nous avons toujours imposé que les espaces d'approximation $X_i^h \subset X_i$ et $V_i^h \subset V_i$ soient tels que

$$(4.6)\qquad \dim X_i^h = \dim V_i^h.$$

On peut montrer que lorsque $k \ge 1$ et h est petit les problèmes approchés admettent une solution unique. Le cas $k=0$ n'est pas aussi immédiat et il est nécessaire de regarder chaque cas particulier.

Cas 0. (équations (3.30) et (3.34)). On obtient

$$(4.8)\qquad \begin{cases} (\alpha I-hA_1)x_1 + (1-\alpha)x_2 = x^0+hf_1 \\ -\alpha x_{e-1} + [(2\alpha-1)I-hA_2]x_e + (1-\alpha)x_{e+1} = hf_e, \quad e=2,\dots,E-1 \\ -\alpha x_{E-1} + (\alpha I-hA_E)x_E = hf_E, \end{cases}$$

où I est l'identité dans $\mathcal{L}(X)$ et

$$(4.9)\qquad A_e = h^{-1}\int_{J_e} A(t)dt, \qquad f_e = h^{-1}\int_{J_e} f(t)dt.$$

Lorsque $A=0$, le déterminant de la matrice définie en (4.8) est égal à α^{E-1}. Lorsque

$\alpha=1$, on obtient le schéma aux différences implicite d'Euler

$$(4.10) \qquad [I-hA_1]x_1 = x^0+hf_1, \quad [I-hA_e]x_e = x_{e-1}+hf_e, \quad e=2,\dots,E.$$

Cas 1. (équations (3.32)-(3.35)). Pour $\alpha \neq 1$ on obtient le schéma explicite suivant

$$(4.11) \qquad \begin{cases} x_1 = x^0, \; x_2 = (1-\alpha)^{-1}\{[(1-\alpha)I+hA_1]x_1+hf_1\} \\ x_e = (1-\alpha)^{-1}\{[(1-2\alpha)I+hA_{e-1}]x_{e-1}+\alpha x_{e-2}+hf_{e-1}\}, \; e=3,\dots,E-1, \end{cases}$$

où A_e et f_e sont comme dans le cas 0. Pour $\alpha=0$, (4.11) coincide avec le schéma d'Euler explicite et pour $\alpha = \frac{1}{2}$ avec le schéma d'Euler amélioré.

Cas 2. (équations (3.30)-(3.33)). Ce schéma est semblable au cas 0 sauf pour les deux premières équations

$$(4.12) \qquad \begin{cases} \alpha x_0 + (1-\alpha)x_1 = x^0 \\ -\alpha x_0 + [(2\alpha-1)]-hA_1]x_1 + (1-\alpha)x_2 = hf_1 \\ -\alpha x_{e-1} + [(2\alpha-1)I-hA_e]x_e + (1-\alpha)x_{e+1} = hf_e, \; e=2,\dots,E-1, \\ -\alpha x_{E-1} + [\alpha I-hA_E]x_E = hf_E, \end{cases}$$

où A_e et f_e sont définis comme au cas 0. Pour $A=0$, le déterminant de la matrice associée à (4.12) est α^{E+1}. Lorsque $\alpha=1$ on récupère (4.10) avec $x_0 = x^0$.

4.3. Le problème de commande optimale approché

On utilise les résultats précédents avec ceux de la section 3.3. Le cas 1 pour k=0 est un cas particulier de M.C. DELFOUR [4].

Références

I. BABUSKA et A.K. AZIZ [1], Survey lecture on the mathematical foundations of the finite element method, in The mathematical foundations of the finite element method with applications to partial differential equations, éd. A.K. Aziz, Academic Press, New York 1972.

A. CHALIFOUR [1], Analyse d'une méthode d'éléments finis discontinus pour les équations différentielles linéaires, mémoire de maîtrise, Université de Montréal, 1976.

M. CROUZEIX [1], Approximation des équations hyperboliques du second ordre par des méthodes de Runge-Kutta, Séminaire Ciarlet-Glowinski-Raviart 1974-1975.

[2], Sur l'approximation des équations différentielles opérationnelles linéaires par des méthodes de RUNGE-KUTTA. Thèse de doctorat d'état ès-sciences mathématiques, Université de Paris VI, mars 1975.

M.C. DELFOUR [1], Solution numérique de l'équation différentielle de Riccati rencontrée en théorie de la commande optimale de systèmes héréditaires linéaires, dans Control theory, numerical methods and computer systems modelling, éds. A. BENSOUSSAN and J.L. LIONS, Springer-Verlag, New York 1975, 362-383.
[2], Numerical solution of the operational Riccati differential equation in the optimal control theory of linear hereditary differential systems with a linear-quadratic cost function, Proc. 1974 IEEE Conference on Decision and Control, Phoenix, 784-790.
[3], Numerical solution of the operator Riccati equation for the filtering of linear stochastic hereditary differential systems, dans "Proceedings of the 7th IFIP Conference on Optimization Techniques", à être publié par Springer-Verlag.
[4], The linear quadratic optimal control problem for hereditary differential systems: theory and numerical solution, à paraître dans Journal of Applied Mathematics and Optimization.

M.C. DELFOUR et F. TROCHU [1], Discontinuous finite element methods for the approximation of optimal control problems governed by hereditary differential systems, Proc. IFIP Working Conference on Distributed Parameter Systems, Rome (Italie), 21-24 juin 1976.

P. LESAINT et P.A. RAVIART [1], On a finite element method for solving the neutron transport equation, Proc. Symposium on Mathematical Aspects of Finite Elements in Partial Differential Equations, Mathematics Research Center, Univ. of Wisconsin, Madison, 1-3 avril 1974.

J.L. LIONS [1], Some aspects of the optimal control of distributed parameter systems, SIAM, Philadelphie, 1972.

J.C. NEDELEC [1], Schéma d'approximation pour des équations intégro-différentielles de Riccati. Thèse de doctorat d'état ès-sciences mathématiques, 7 octobre 1970, Paris, France.
[2], Approximation par éléments finis des équations de Riccati, Rapport interne, Université de Rennes.

J. REVERDY [1], Discrétisation d'une équation aux différences-différentielles, comparaison entre différents schémas, Instituto Lombardo (Rend. Sc.) A 107 (1973), 511-527.

F. TROCHU et M. DELFOUR [1], Approximation des systèmes différentiels par éléments finis discontinus, Rapport CRM, juin 1976 (Centre de Recherches Mathématiques, Université de Montréal).

M.D. et F.T.
Centre de Recherches Mathématiques
Université de Montréal
Montréal, Québec, Canada H3C 3J7