Examen du cours MOPSI

14 février 2013, 08h30-12h00.
Corrigé.

Exercice 1 : une équation aux dérivées partielles non linéaire.

1 Ce problème admet une unique solution par une application standard du théorème de Lax Milgram, en remarquant que si $u_0 \in H_0^1(\Omega)$, on a bien $f(u_0)$ dans $L^2(\Omega)$. En effet, puisque f est lipschitzienne,

$$|f(u)| \leq |f(u) - f(0)| + |f(0)| \leq L|u| + |f(0)|$$

et donc, si u est dans $L^2(\Omega)$, $f(u)$ est aussi dans $L^2(\Omega)$.

2 Soit $u_1 = \mathcal{T}(u_0)$ et $v_1 = \mathcal{T}(v_0)$. On a

$$\begin{cases} -\Delta u_1 = f(u_0) \text{ sur } \Omega, \\ u_1 = 0 \text{ sur } \partial\Omega, \end{cases}$$

et

$$\begin{cases} -\Delta v_1 = f(v_0) \text{ sur } \Omega, \\ v_1 = 0 \text{ sur } \partial\Omega. \end{cases}$$

Par conséquent

$$\begin{cases} -\Delta(u_1 - v_1) = f(u_0) - f(v_0) \text{ sur } \Omega, \\ v_1 = 0 \text{ sur } \partial\Omega. \end{cases}$$

En utilisant les formulations variationnelles associées à ces problèmes, on a donc : $\forall w \in H_0^1(\Omega)$,

$$\int_\Omega \nabla(u_1 - v_1) \cdot \nabla w = \int_\Omega (f(u_0) - f(v_0))w.$$

En choisissant $w = u_1 - v_1$ et en notant C_Ω la constante de Poincaré du domaine, on obtient

$$\begin{aligned} \|u_1 - v_1\|_{H_0^1}^2 &= \int_\Omega (f(u_0) - f(v_0))(u_1 - v_1) \\ &\leq L \int_\Omega |u_0 - v_0||u_1 - v_1| \\ &\leq L\|u_0 - v_0\|_{L^2}\|u_1 - v_1\|_{L^2} \\ &\leq LC_\Omega^2\|u_0 - v_0\|_{H_0^1}\|u_1 - v_1\|_{H_0^1} \end{aligned}$$

et donc

$$\|u_1 - v_1\|_{H_0^1} \leq LC_\Omega^2\|u_0 - v_0\|_{H_0^1}.$$

On en déduit que l'application $\mathcal{T}$ est contractante si $LC_\Omega^2 < 1$.

Dans ce cas, le théorème de point fixe de Picard assure l'existence et unicité d'un point fixe de $\mathcal{T}$, et donc l'existence et unicité d'une solution au problème (1) dans $H_0^1(\Omega)$. On a aussi vu en cours qu'une manière simple de construire une approximation de u solution de (1) est de considérer la suite récurrente $u_{n+1} = \mathcal{T}(u_n)$, c'est-à-dire :

$$\begin{cases} -\Delta u_{n+1} = f(u_n) \text{ sur } \Omega, \\ u_{n+1} = 0 \text{ sur } \partial\Omega. \end{cases}$$

La suite (u_n) converge vers u dans $H_0^1(\Omega)$. On peut résoudre les problèmes successifs sur u_n par éléments finis par exemple.

Exercice 2 : processus gaussien.

1 Soient $0 \leq t_1 < \cdots < t_n < T$. On approche les intégrales par des sommes de Riemann :

$$\left(W_{t_1} + \frac{t_1}{N} \sum_{k=0}^{N-1} \frac{W_{kt_1/N} - W_T}{T - kt_1/N}, \dots, W_{t_n} + \frac{t_n}{N} \sum_{k=0}^{N-1} \frac{W_{kt_n/N} - W_T}{T - kt_n/N} \right)$$

converge presque sûrement vers $(X_{t_1}, \dots, X_{t_n})$ lorsque $N \to +\infty$, car la fonction intégrée est p.s. continue. Ainsi, $(X_{t_1}, \dots, X_{t_n})$ est un vecteur gaussien comme limite en loi d'un vecteur gaussien. On a clairement $\mathbb{E}[X_t] = \mathbb{E}[W_t] + \int_0^t \frac{\mathbb{E}[W_s - W_T]}{T-s} ds = 0$.

2 En utilisant le même argument qu'à la question précédente, $(X_{t_1}, \dots, X_{t_n}, W_T)$ est la limite de $\left(W_{t_1} + \frac{t_1}{N} \sum_{k=0}^{N-1} \frac{W_{kt_1/N} - W_T}{T - kt_1/N}, \dots, W_{t_n} + \frac{t_n}{N} \sum_{k=0}^{N-1} \frac{W_{kt_n/N} - W_T}{T - kt_n/N}, W_T\right)$ et est donc un vecteur gaussien centré. Pour $t \in [0, T[$, on a $\mathbb{E}[X_t W_T] = t + \int_0^t \frac{u-T}{T-u} du = 0$ ce qui prouve l'indépendance de W_T avec $(X_{t_1}, \dots, X_{t_n})$, quels que soient $n \in \mathbb{N}^*$ et $0 \leq t_1 < \cdots < t_n < T$.

3 Soient $s, t \in [0, T[$. Comme $\mathbb{E}[W_s W_t] = s \wedge t$, on a :

$$\mathbb{E}[X_s X_t] = s \wedge t + \int_0^s \frac{u \wedge t - t}{T - u} du + \int_0^t \frac{u \wedge s - s}{T - u} du + \int_0^s \int_0^t \frac{u \wedge v - (u + v) + T}{(T - u)(T - v)} du dv.$$

On suppose sans perte de généralité que $s \leq t$. On a :

$$\begin{aligned} \mathbb{E}[X_s X_t] =& s + \int_0^s \frac{u - t}{T - u} du + \int_0^s \frac{u - s}{T - u} du \\ &+ \int_0^s \int_0^s \frac{u \wedge v - (u + v) + T}{(T - u)(T - v)} du dv + \int_0^s \underbrace{\left(\int_s^t \frac{u - (u + v) + T}{(T - u)(T - v)} dv \right)}_{= \frac{t-s}{T-u}} du. \end{aligned}$$

On calcule l'intégrale double. Par symétrie, elle est égale à

$$2 \int_{0 \leq u \leq v \leq s} \frac{u - (u + v) + T}{(T - u)(T - v)} du dv = 2 \int_0^s \frac{1}{T - u} \left(\int_u^s dv \right) du = 2 \int_0^s \frac{s - u}{T - u} du.$$

Ainsi, on obtient

$$\mathbb{E}[X_s X_t] = s + \int_0^s \frac{u - t + u - s + 2(s - u) + t - s}{T - u} du = s.$$

Ainsi, on a $\mathbb{E}[X_t] = 0$ et $\mathbb{E}[X_s X_t] = s \wedge t$ pour $s, t \in [0, T[$. Comme $(X_t, t \in [0, T[)$ est un processus gaussien, on en déduit que c'est un mouvement brownien, indépendant de W_T.

Exercice 3 : calcul de projection orthogonale.

1 Comme les variables sont indépendantes, on a $\mathbb{E}[f(X_1, \ldots, X_n)|X_i] = f_i(X_i)$.

2 Soit $l \in \{1, \ldots, n\}$. Comme, $f_l(X_l) - h_l(X_l)$ et $f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)$ sont de carré intégrables, leur produit est une variable aléatoire intégrable. On commence par calculer l'espérance conditionnelle $\mathbb{E}\left[f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)|X_l\right] = f_l(X_l) - f_l(X_l) - \sum_{i \neq l} \mathbb{E}[f_i(X_i)] = 0$, grâce à l'indépendance de X_i et de X_l et au fait que $\mathbb{E}[f_i(X_i)] = \mathbb{E}[f(X_1, \ldots, X_n)] = 0$. Alors,

$$\begin{aligned}
&\mathbb{E}\left[(f_l(X_l) - h_l(X_l))\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)\right)\right] \\
&= \mathbb{E}\left[(f_l(X_l) - h_l(X_l))\mathbb{E}\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)\middle| X_l\right)\right] = 0.
\end{aligned}$$

3 On écrit

$$\begin{aligned}
\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n h_i(X_i)\right)^2 &= \left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i) + \sum_{i=1}^n (f_i(X_i) - h_i(X_i))\right)^2 \\
&= \left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)\right)^2 \\
&\quad + 2\sum_{l=1}^n (f_l(X_l) - h_l(X_l))\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)\right) \\
&\quad + \left(\sum_{i=1}^n (f_i(X_i) - h_i(X_i))\right)^2,
\end{aligned}$$

et obtient grâce à la question précédente

$$\begin{aligned}
\mathbb{E}\left[\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n h_i(X_i)\right)^2\right] &= \mathbb{E}\left[\left(f(X_1, \ldots, X_n) - \sum_{i=1}^n f_i(X_i)\right)^2\right] \\
&\quad + \mathbb{E}\left[\left(\sum_{i=1}^n (f_i(X_i) - h_i(X_i))\right)^2\right],
\end{aligned}$$

ce qui donne le résultat voulu.

4 On applique ce qui précède à $\tilde{f}(X_1, \ldots, X_n) = f(X_1, \ldots, X_n) - m$ avec $m = \mathbb{E}[f(X_1, \ldots, X_n)]$. On a $\tilde{f}_i(x) = f_i(x) - m$, et on a grâce à la question précédente :

$$\mathbb{E}\left[\left(\tilde{f}(X_1, \ldots, X_n) - \sum_{i=1}^n \tilde{f}_i(X_i)\right)^2\right] \leq \mathbb{E}\left[\left(\tilde{f}(X_1, \ldots, X_n) - \sum_{i=1}^n h_i(X_i)\right)^2\right].$$

Comme $\tilde{f}(X_1,\ldots,X_n)-\sum_{i=1}^n \tilde{f}_i(X_i) = f(X_1,\ldots,X_n)+(n-1)m-\sum_{i=1}^n f_i(X_i)$, on voit que $Y=\sum_{i=1}^n f_i(X_i)-(n-1)m$ satisfait l'inégalité voulue.

5 $f_1(X_1)=\mathbb{E}[(\sum_{i=1}^n X_i)^2|X_1] = X_1^2+2X_1\mathbb{E}[\sum_{i=2}^n X_i|X_1]+\mathbb{E}[(\sum_{i=2}^n X_i)^2|X_1] = X_1^2+(n-1)$. Par symétrie, on obtient $f_i(X_i)=X_i^2+(n-1)$ puis $Y=\sum_{i=1}^n X_i^2$.

Problème 1 : contrôle optimal.

1 Il s'agit simplement des équations du mouvement de Newton : le premier terme est la masse m_0 fois l'accélération $\ddot{y}$, le second correspond à la force de rappel du ressort linéaire, qui a une longueur d'équilibre l_0 et une constante de raideur k_0, et la force extérieure exercée sur le système est au second membre.

On se ramène à un problème du premier ordre de la forme $\dot{x}=Ax+Bu$ en posant $x=\begin{pmatrix} y-l_0 \\ \dot{y} \end{pmatrix}$, $A=\begin{bmatrix} 0 & 1 \\ -k_0/m_0 & 0 \end{bmatrix}$ et $B=\begin{pmatrix} 0 \\ 1/m_0 \end{pmatrix}$.

2 C'est une simple formule de Duhamel, qui se vérifie en différentiant par rapport au temps t.

3 Soit $x^1(T)$ et $x^2(T)$ deux points de $\mathrm{Acc}(x_0,T)$, associés respectivement aux trajectoires $x^1(t)$ et $x^2(t)$ et aux contrôles $u^1(t)$ et $u^2(t)$. Soit $\lambda\in(0,1)$. Par linéarité, on a évidemment : $\lambda x^1(T)+(1-\lambda)x^2(T)=\exp(At)x_0+\int_0^t \exp((t-s)A)[B(\lambda u^1(s)+(1-\lambda)u^2(s))+r(s)]\,ds$. Ceci montre que le point $\lambda x^1(T)+(1-\lambda)x^2(T)$ est atteint pour le contrôle $\lambda u^1+(1-\lambda)u^2$. L'ensemble $\mathrm{Acc}(x_0,T)$ est donc convexe.

4 On suppose que $\mathrm{rg}(K)<n$. Il existe donc un vecteur ligne non nul $\psi\in\mathbb{R}^n$ tel que $\psi K=0$: il suffit de choisir ψ dans l'orthogonal de l'image de K. On a donc $\psi B=\psi AB=\ldots=\psi A^{n-1}B=0$. Par ailleurs, on sait par le théorème de Cayley-Hamilton qu'il existe des réels $a_{n-1},\ldots,a_0$ tels que

$$A^n+a_{n-1}A^{n-1}+\ldots+a_1A+a_0=0.$$

On en déduit que $\psi A^nB=-\psi\left(a_{n-1}A^{n-1}+\ldots+a_1A+a_0\right)=0$, et par récurrence immédiate, $\psi A^kB=0$ pour tous les $k\geq 0$. Ceci implique en particulier que

$$\psi\exp(tA)B=\sum_{k\geq 0}\frac{t^k}{k!}\psi A^kB=0,$$

et donc, quelque soit $u\in L^\infty([0,T],\mathbb{R}^m)$,

$$\psi\Phi(u)=\int_0^T \psi\exp((T-t)A)Bu(t)\,dt=0.$$

Ceci montre que $\Phi(u)$ n'est pas surjective, puisque $\psi\in\mathrm{Im}(\Phi)^\perp$.

5 C'est essentiellement le même raisonnement que ci-dessus. Si Φ n'est pas surjective, il existe un vecteur ligne non nul ψ (dans l'orthogonal de l'image de Φ) tel que, pour tout $u\in L^\infty([0,T],\mathbb{R}^m)$,

$$0=\psi\Phi(u)=\int_0^T \psi\exp((T-t)A)Bu(t)\,dt.$$

En choisissant $u(t) = \psi \exp((T-t)A)B$, on a donc, pour tout $t \in [0,T]$,

$$\psi \exp((T-t)A)B = 0.$$

Cette relation en $t = T$ donne $\psi B = 0$. En dérivant cette relation k fois par rapport au temps t, puis en choisissant $t = T$, on obtient plus généralement $\psi A^k B = 0$. En particulier, $\psi K = 0$ et donc $\mathrm{rg}(K) < n$, puisque l'orthogonal de l'image de K contient un vecteur non nul.

6 On a la suite d'équivalence : $\mathrm{Acc}(x_0, T) = \mathbb{R}^n$ ssi $\forall y \in \mathbb{R}^n$, $\exists u \in L^\infty([0,T], \mathbb{R}^m)$, $\exp(TA)x_0 + \int_0^T \exp((T-s)A)(Bu(s) + r(s))\, ds = y$ ssi $\forall y \in \mathbb{R}^n$, $\exists u \in L^\infty([0,T], \mathbb{R}^m)$, $\Phi(u) = y - \exp(TA)x_0 - \int_0^T \exp((T-s)A)r(s)\, ds$ ssi $\forall z \in \mathbb{R}^n$, $\exists u \in L^\infty([0,T], \mathbb{R}^m)$, $\Phi(u) = z$ ssi Φ est surjective ssi $\mathrm{rg}(K) = n$. La dernière équivalence est une conséquence des questions 4 et 5.

7 Dans le cas de l'oscillateur décrit à la question 1, on a $n = 2$, $m = 1$, $r = 0$, $B = \begin{pmatrix} 0 \\ 1/m_0 \end{pmatrix}$ et $AB = \begin{pmatrix} 1/m_0 \\ 0 \end{pmatrix}$. On en déduit que $K = [A, AB]$ est bien de rang 2, et donc $\mathrm{Acc}(x_0, T) = \mathbb{R}^2$.

8 On note tout d'abord que la fonction $C(u)$ est bien définie pour $u \in L^2([0,T], \mathbb{R}^m)$, puisque dans ce cas,

$$x(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)Bu(s)\, ds$$

est bien également une fonction de carré intégrable. De plus elle est minorée (car positive) et on peut donc considérer son infimum et une suite minimisante associée : $u^n \in L^2([0,T], \mathbb{R}^m)$ telle que

$$\lim_{n\to\infty} C(u^n) = \inf_{u \in L^2([0,T], \mathbb{R}^m)} C(u).$$

On note que $C(u) \geq \|u\|^2_{L^2([0,T])}$ et donc, la suite minimisante u^n est bornée dans $L^2([0,T], \mathbb{R}^m)$. On peut par conséquent supposer (quitte à extraire une sous-suite) que u^n converge faiblement dans $L^2([0,T], \mathbb{R}^m)$ vers une fonction $u^* \in L^2([0,T], \mathbb{R}^m)$. Les trajectoires associées

$$x^n(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)Bu^n(s)\, ds$$

convergent simplement (i.e. pour tout $t \in [0,T]$) vers

$$x^*(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)Bu^*(s)\, ds.$$

Puisque la suite de trajectoires $(x^n(t))_{0 \leq t \leq T}$ est bornée dans $L^2([0,T], \mathbb{R}^n)$, on en déduit par convergence dominée que x^n converge vers x^* dans $L^2([0,T], \mathbb{R}^n)$. On a par conséquent

$$\lim_{n\to\infty} \left(\|x^n(T)\|^2 + \int_0^T \|x^n(t)\|^2\, dt \right) = \|x^*(T)\|^2 + \int_0^T \|x^*(t)\|^2\, dt.$$

Par ailleurs, on a vu en cours que la convergence faible implique que

$$\|u^*\|_{L^2([0,T])} \leq \liminf_{n\to\infty} \|u^n\|_{L^2([0,T])}.$$

On en déduit que

$$C(u^*) \leq \lim\inf_{n\to\infty} C(u_n) = \inf_{u\in L^2([0,T],\mathbb{R}^m)} C(u)$$

et donc

$$C(u^*) = \inf_{u\in L^2([0,T],\mathbb{R}^m)} C(u) = \min_{u\in L^2([0,T],\mathbb{R}^m)} C(u).$$

9 On a

$$x(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)Bu(s)\,ds$$

et

$$(x+\delta x)(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)B(u+\delta u)(s)\,ds$$

d'où

$$\delta x(t) = \int_0^t \exp((t-s)A)B\delta u(s)\,ds.$$

Par ailleurs,

$$\begin{aligned} C(u+\delta u) &= \|x(T)+\delta x(T)\|^2 + \int_0^T \left(\|x(t)+\delta x(t)\|^2 + \|u(t)+\delta u(t)\|^2\right)\,dt, \\ &= C(u) + 2\left((x(T))^T\delta x(T) + \int_0^T \left(x(t)^T\delta x(t) + u(t)^T\delta u(t)\right)\,dt\right) \\ &\quad + \|\delta x(T)\|^2 + \int_0^T \left(\|\delta x(t)\|^2 + \|\delta u(t)\|^2\right)\,dt. \end{aligned}$$

On a ensuite, en utilisant la formule ci-dessus pour $\delta x(t)$ et Fubini :

$$\begin{aligned} &(x(T))^T\delta x(T) + \int_0^T \left(x(t)^T\delta x(t) + u(t)^T\delta u(t)\right)\,dt \\ &= (x(T))^T\int_0^T \exp((T-s)A)B\delta u(s)\,ds + \int_0^T x(t)^T\int_0^t \exp((t-s)A)B\delta u(s)\,ds\,dt + \int_0^T u(t)^T\delta u(t)\,dt \\ &= (x(T))^T\int_0^T \exp((T-s)A)B\delta u(s)\,ds + \int_0^T\int_s^T x(t)^T\exp((t-s)A)B\,dt\delta u(s)\,ds + \int_0^T u(s)^T\delta u(s)\,ds \\ &= \int_0^T \left((x(T))^T\exp((T-s)A)B + \int_s^T x(t)^T\exp((t-s)A)B\,dt + u(s)^T\right)\delta u(s)\,ds \\ &= \int_0^T D(u)(s)\delta u(s)\,ds. \end{aligned}$$

On en déduit la relation demandée.

10 D'après la question précédente, on a

$$C(u+\delta u) > C(u) + 2\int_0^T D(u)(s)\delta u(s)\,ds$$

dès que $\delta u \neq 0$. En appliquant cette inégalité avec $u+\delta u = u_0$ et $u = \lambda u_0 + (1-\lambda)u_1$, on en déduit immédiatement que

$$C(u_0) > C(\lambda u_0 + (1-\lambda)u_1) + 2(1-\lambda)\int_0^T D(\lambda u_0 + (1-\lambda)u_1)(s)(u_0-u_1)(s)\,ds.$$

De même, on a

$$C(u_1) > C(\lambda u_0 + (1-\lambda)u_1) - 2\lambda \int_0^T D(\lambda u_0 + (1-\lambda)u_1)(s)(u_0 - u_1)(s)\, ds.$$

On en déduit facilement que $C(\lambda u_0 + (1-\lambda)u_1) < \lambda C(u_0) + (1-\lambda)C(u_1)$. En particulier, C admet un unique minimiseur, noté u^* dans la suite.

11 Pour que les résultats des questions 8 et 10 soient encore valables, il suffit de montrer que la fonction C prend des valeurs finies pour certains contrôles u, dans le cas $T = \infty$. Pour cela, il faut en particulier que x tendent vers 0 à l'infini et soit dans $L^2(\mathbb{R}_+)$ pour certains contrôles u. Or, comme $\mathrm{rg}(K) = n$, on sait (cf. la question 6) que pour un temps $\tilde{T}$ fixé, il existe un $u \in L^2([0,\tilde{T}], \mathbb{R}^m)$ tel que $x(\tilde{T}) = 0$. En prolongeant u par zéro au-delà de $\tilde{T}$, on a $x(t) = 0$ pour tout $t > \tilde{T}$, et donc $C(u) < \infty$. Tous les raisonnements précédents s'appliquent donc, et on a bien un unique minimiseur u^* à la fonction C dans ce cas.

12 On sait que $C(u) \geq C(u^*)$, et donc, d'après la question 9, en prenant $u = u^*$, on a

$$2\int_0^T D(u^*)(s)\delta u(s)\, ds + \|\delta x(T)\|^2 + \int_0^T \left(\|\delta x(t)\|^2 + \|\delta u(t)\|^2\right)\, dt \geq 0.$$

Comme cette relation est valable pour tout δu, elle est en particulier valable pour $\epsilon\delta u$ et en faisant tendre ϵ vers zéro, on obtient

$$\int_0^T D(u^*)(s)\delta u(s)\, ds \geq 0.$$

En changeant ensuite le signe de δu, on a finalement

$$\int_0^T D(u^*)(s)\delta u(s)\, ds = 0.$$

Ceci étant valable pour tout δu, on en déduit que

$$D(u^*) = 0$$

et donc

$$(x^*(T))^T \exp((T-s)A)B + \int_s^T (x^*(t))^T \exp((t-s)A)B\, dt + (u^*(s))^T = 0.$$

Ceci donne la formule suivante pour le contrôle optimal :

$$u^*(s) = -\left[\left((x^*(T))^T \exp((T-s)A) + \int_s^T (x^*(t))^T \exp((t-s)A)\, dt\right) B\right]^T.$$

Si on pose

$$p(s) = -(x^*(T))^T \exp((T-s)A) - \int_s^T (x^*(t))^T \exp((t-s)A)\, dt$$

on a donc bien $u^*(s) = (p(s)B)^T$, et on vérifie par dérivation de la formule définissant p que p est la solution de l'EDO linéaire :

$$\begin{cases} \dot{p}(s) = -p(s)A + (x^*(s))^T, \\ p(T) = -(x^*(T))^T. \end{cases}$$

13 Soit $(x(t), p(t))_{0 \le t \le T}$ solution du problème

$$\begin{cases} \dot{x}(t) = Ax(t) + BB^T p(t)^T, \\ x(0) = x_0, \\ \dot{p}(t) = -p(t)A + (x(t))^T, \\ p(T) = -(x(T))^T. \end{cases}$$

On pose $u(t) = (p(t)B)^T$. On a, par intégration explicite de l'EDO sur p,

$$\begin{aligned} D(u)(s) &= (x(T))^T \exp((T-s)A)B + \int_s^T (x(t))^T \exp((t-s)A)B\, dt + (u(s))^T \\ &= -p(s)B + p(s)B = 0. \end{aligned}$$

En utilisant la question 9, on a donc, pour toute fonction δu non nulle,

$$\begin{aligned} C(u + \delta u) &= C(u) + 2\int_0^T D(u)(s)\delta u(s)\, ds + \|\delta x(T)\|^2 + \int_0^T \left(\|\delta x(t)\|^2 + \|\delta u(t)\|^2\right) dt \\ &> C(u) \end{aligned}$$

ce qui montre bien que u réalise le minimum global de C. On note que le système d'EDO sur le couple (x, p) n'est pas standard, car on a une condition initiale sur x et finale sur p : la solution ne peut pas se calculer par des méthodes d'approximation habituelles pour les EDOs.

14 Le théorème de Cauchy-Lipschitz assure l'existence d'une unique solution maximale sur un intervalle de temps (t_0, T) avec $t_0 < T$. Si $t_0 > -\infty$, on sait que $\|E(t_0)\| = \infty$. On vérifie facilement que si E est solution de l'EDO, alors E^T est solution de la même EDO, et donc, par l'unicité dans le théorème de Cauchy-Lipschitz, $E = E^T$.

15 Par les raisonnements précédents, en remplaçant le temps initial 0 par t_0 dans tous les réultats, on sait que $u \mapsto C(u; t_0, x_0)$ admet un unique minimiseur u associée à la trajectoire optimale x. D'après la question 13, on sait que ce minimiseur est $u(t) = (p(t)B)^T$, avec $(x(t), p(t))_{t_0 \le t \le T}$ solution de :

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ x(t_0) = x_0, \\ \dot{p}(t) = -p(t)A + (x(t))^T, \\ p(T) = -(x(T))^T. \end{cases}$$

On pose donc $p(t) = (E(t)x(t))^T$, de sorte que $u(t) = (p(t)B)^T$. On a bien $p(T) = -(x(T))^T$, et

$$\begin{aligned} \dot{p}(t) &= \dot{x}(t)^T E(t) + (x(t))^T \dot{E}(t) \\ &= (Ax(t) + Bu(t))^T E(t) + (x(t))^T (\mathrm{Id} - A^T E(t) - E(t)A - E(t)BB^T E(t)) \\ &= (x(t))^T A^T E(t) + p(t)BB^T E(t) + (x(t))^T - (x(t))^T A^T E(t) - (x(t))^T E(t)A - (x(t))^T E(t)BB^T E(t) \\ &= (x(t))^T - (x(t))^T E(t)A \\ &= (x(t))^T - p(t)A \end{aligned}$$

puisque $p(t) = (E(t)x(t))^T$. On en déduit, d'après la question 13, que u est le contrôle optimal, et x la trajectoire optimale associée, pour le coût $C(u; t_0, x_0)$.

16 On calcule (en omettant la dépendance en t pour plus de clarté) :

$$\begin{aligned}
&\frac{d}{dt}\left(x(t)^T E(t)x(t)\right) = \dot{x}(t)^T E(t)x(t) + x(t)^T \dot{E}(t)x(t) + x(t)^T E(t)\dot{x}(t) \\
&= (Ax(t) + Bu(t))^T E(t)x(t) + x(t)^T \left(\mathrm{Id} - A^T E(t) - E(t)A - E(t)BB^T E(t)\right) x(t) \\
&\quad + x(t)^T E(t)(Ax(t) + Bu(t)) \\
&= x^T A^T Ex + u^T B^T Ex + \|x\|^2 - x^T A^T Ex - x^T EAx - x^T EBB^T Ex + x^T EAx + x^T EBu \\
&= x^T EBB^T Ex + \|x\|^2 - x^T EBB^T Ex + x^T EBB^T Ex \\
&= x^T EBB^T Ex + \|x\|^2 \\
&= \|u\|^2 + \|x\|^2.
\end{aligned}$$

On en déduit, par intégration entre t_0 et T,

$$x(T)^T E(T)x(T) - x(t_0)^T E(t_0)x(t_0) = \int_{t_0}^{T} \left(\|u(t)\|^2 + \|x(t)\|^2\right) dt$$

et donc, puisque $E(T) = -\mathrm{Id}$,

$$C(u; t_0, x_0) = -x_0^T E(t_0)x_0.$$

Puisque $C(u; t_0, x_0) < \infty$ pour tout x_0, on a $\|E(t_0)\| < \infty$ et donc, nécessairement $t_0 = -\infty$: la solution de l'équation de Riccati est définie pour tout temps $t \leq T$.

En pratique, on peut donc calculer numériquement la trajectoire optimale, partant de x_0 à l'instant 0 en (i) calculant la solution $E(t)$ de l'équation de Riccati puis (ii) en considérant la solution de l'EDO

$$\begin{cases} \dot{x}(t) = Ax(t) + BB^T E(t)x(t), \\ x(0) = x_0. \end{cases}$$

Problème 2 : une marche aléatoire sur $\mathbb{Z}$.

1 On pose pour $x \in \mathbb{Z}$, $P(x, x) = 1/2$, $P(x, x+1) = P(x, x-1) = 1/4$ et $P(x, y) = 0$ pour $y \notin \{-1, 0, 1\}$. On a $P(M_0 = 0, M_1 = x_1) = P(0, x_1)$ puis par récurrence sur $n \in \mathbb{N}$, $P(M_0 = 0, M_1 = x_1, \ldots, M_n = x_n) = P(0, x_1) \ldots P(x_{n-1}, x_n)$. Cette chaine de Markov est irréductible car pour tout $x \in \mathbb{Z}$, $n \in \mathbb{N}$, $P(x, x-1) \ldots P(x-(n-1), x-n) > 0$ et $P(x, x+1) \ldots P(x+(n-1), x+n) > 0$.

2 Clairement, on a $|M_n| \leq n$ pour tout n, et M_n est en particulier intégrable. De plus, Y_{n+1} est indépendante de $Y_1, \ldots, Y_n$ et donc de $\mathcal{F}_n$, ce qui donne $\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \mathbb{E}[Y_{n+1}] = 0$ puis $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[M_n|\mathcal{F}_n] + \mathbb{E}[Y_{n+1}|\mathcal{F}_n] = M_n$, puisque M_n est $\mathcal{F}_n$-mesurable.

3 M_n étant bornée, elle est en particulier de carré intégrable. D'après le cours, on a $[M]_n = \sum_{k=0}^{n-1} \mathbb{E}[(M_{k+1} - M_k)^2|\mathcal{F}_k] = \sum_{k=0}^{n-1} \mathbb{E}[(Y_{k+1})^2|\mathcal{F}_k] = \sum_{k=0}^{n-1} \mathbb{E}[(Y_{k+1})^2] = \frac{n}{2}$ en utilisant l'indépendance de Y_{k+1} avec $\mathcal{F}_k$ et $\mathbb{E}[(Y_1)^2] = 1/2$.

4 On écrit : $\{\tau_x \leq n\} = \cup_{i=1}^n \{M_i = x\}$. On a $\{M_i = x\} \in \mathcal{F}_i \subset \mathcal{F}_n$, et donc $\{\tau_x \leq n\} \in \mathcal{F}_n$. Enfin, τ est un temps d'arrêt puisque c'est le minimum de deux temps d'arrêt.

5 Pour $n \in \mathbb{N}$, $\tau \wedge n$ est un temps d'arrêt borné, et on a par le théorème d'arrêt appliqué à la martingale $M_n^2 - [M]_n$: $\mathbb{E}[M_{\tau\wedge n}^2 - \frac{1}{2}\tau \wedge n] = 0$. Comme $M_{\tau\wedge n}^2 \leq (a \vee b)^2$, le théorème de convergence monotone assure que $\mathbb{E}[\tau] = \lim_{n\to+\infty} \mathbb{E}[\tau \wedge n] \leq 2(a \vee b)^2$. Cela donne en particulier que $\mathbb{P}(\tau < \infty) = 1$. Par ailleurs, le théorème de convergence dominée donne $\mathbb{E}[M_\tau^2] = \lim_{n\to+\infty} \mathbb{E}[M_{\tau\wedge n}^2] = \frac{1}{2}\mathbb{E}[\tau]$.

6 Pour $n \in \mathbb{N}$, $\tau \wedge n$ est un temps d'arrêt borné, et on a par le théorème d'arrêt $\mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_0] = 0$. Comme $\mathbb{P}(\tau < \infty) = 1$ et $|M_{\tau\wedge n}| \leq a \vee b$, le théorème de convergence dominée donne $\mathbb{E}[M_\tau] = \lim_{n\to+\infty} \mathbb{E}[M_{\tau\wedge n}] = 0$. Comme $M_\tau = -a\mathbf{1}_{\tau_{-a}<\tau_b} + b\mathbf{1}_{\tau_{-a}>\tau_b}$, on obtient $0 = -a(1 - \mathbb{P}(\tau_b < \tau_{-a})) + b\mathbb{P}(\tau_b < \tau_{-a})$ puis

$$\mathbb{P}(\tau_b < \tau_{-a}) = \frac{a}{a+b}.$$

Il vient ensuite $\mathbb{E}[\tau] = 2\left(a^2\frac{b}{a+b} + b^2\frac{a}{a+b}\right) = 2ab$.

7 Comme M_n est $\mathcal{F}_n$-mesurable, on a $\mathbb{E}[(M_n + Y_{n+1})^3|\mathcal{F}_n] = M_n^3 + 3M_n^2\mathbb{E}[Y_{n+1}|\mathcal{F}_n] + 3M_n\mathbb{E}[Y_{n+1}^2|\mathcal{F}_n] + \mathbb{E}[Y_{n+1}^3|\mathcal{F}_n] = M_n^3 + \frac{3}{2}M_n$ en utilisant l'indépendance de Y_{n+1} et $\mathcal{F}_n$ et $\mathbb{E}[Y_1] = \mathbb{E}[Y_1^3] = 0$ et $\mathbb{E}[Y_1^2] = 1$. Par conséquent $\mathbb{E}[M_{n+1}^3 - \frac{3}{2}(n+1)M_{n+1}|\mathcal{F}_n] = M_n^3 + \frac{3}{2}M_n - \frac{3}{2}(n+1)M_n = M_n^3 - \frac{3}{2}nM_n$. Le théorème d'arrêt donne $\mathbb{E}[M_{\tau\wedge n}^3] = \frac{3}{2}\mathbb{E}[(\tau\wedge n)M_{\tau\wedge n}]$, puis on obtient par convergence dominée $\mathbb{E}[M_\tau^3] = \frac{3}{2}\mathbb{E}[\tau M_\tau]$ (on utilise à nouveau que $|M_{\tau\wedge n}| \leq a \vee b$). Ainsi, $\mathbb{E}[\tau M_\tau] = \frac{2}{3}\left(-a^3\frac{b}{a+b} + b^3\frac{a}{a+b}\right) = \frac{2}{3}ab(b-a)$. On voit que pour $b \neq a$, $\mathbb{E}[\tau M_\tau] \neq 0 = \mathbb{E}[\tau]\mathbb{E}[M_\tau]$, si bien que τ et M_τ ne peuvent pas être indépendants. Supposons désormais $b = a$. Comme Y_1 à même loi que $-Y_1$, les processus $(M_n, n \in \mathbb{N})$ et $(-M_n, n \in \mathbb{N})$ ont même loi : ce sont tous deux des chaînes de Markov de matrice de transition P issues de 0. De plus $\tau = \inf\{n \in \mathbb{N}, |M_n| = a\} = \inf\{n \in \mathbb{N}, |-M_n| = a\}$. Par conséquent $\mathbb{P}(M_\tau = a, \tau = k) = \mathbb{P}(-M_\tau = a, \tau = k)$ et donc $\mathbb{P}(M_\tau = a, \tau = k) = \frac{1}{2}\mathbb{P}(\tau = k)$ ce qui prouve l'indépendance.

8 Soit $(Z_k, k \in \mathbb{N})$ une famille de variables aléatoires indépendantes telles que $\mathbb{P}(Z_k = 1) = \mathbb{P}(Z_k = -1) = 1/2$. Alors $\Sigma_0 = 0$ et $\Sigma_n = \sum_{k=1}^n Z_k$ est une marche aléatoire standard sur $\mathbb{Z}$. Il est facile de voir que $Z_1 + Z_2$ a même loi que $2Y_1$, et les variables $Z_{2n-1} + Z_{2n}$ sont i.i.d. Par conséquent, $(\Sigma_{2n}, n \in \mathbb{N})$ a même loi que $(2M_n, n \in \mathbb{N})$ ce qui prouve le résultat. Ainsi, τ_b a même loi que $\inf\{n \in \mathbb{N}, \Sigma_{2n} = 2b\} = \frac{1}{2}\inf\{n \in \mathbb{N}, \Sigma_n = 2b\}$ puisque Σ_{2n+1} est toujours impair. Ainsi, grâce au résultat du cours sur la transformée de Laplace du temps d'atteinte par la marche aléatoire standard, on obtient $\mathbb{E}[e^{-y\tau_b}] = \exp(-2b\mathrm{acosh}(e^{y/2}))$.