1. Un problème de mise en jambes

1.1. Énoncé : commande optimale d'un système gouverné par une équation différentielle linéaire —

L'objet de ce problème est l'étude d'un système gouverné par une équation différentielle linéaire dont le second membre dépend d'une fonction $u(\cdot)$ appelée «commande» (ou «contrôle»), et la recherche d'une «commande optimale» minimisant un critère donné.

1.1.1. Rappels — Une fonction $\varphi(\cdot) : [0, T] \to \mathbb{R}^d$ est dite *continue par morceaux* sur $[0, T]$ si elle est continue sur $[0, T]$ ou s'il existe un nombre fini de points $t_0 = 0 < t_1 < \ldots < t_{k-1} < t_k = T$ tels que :

- $\varphi(\cdot)$ est continue sur chaque intervalle $]t_{i-1}, t_i[$, $i = 1, \ldots, k$;
- en tout t_i ($i = 1, \ldots, k - 1$), $\varphi(\cdot)$ admet une limite à gauche et une limite à droite ;
- $\varphi(\cdot)$ admet une limite à droite en t_0 et une limite à gauche en T.

Pour une telle fonction, on convient d'adopter (sans perte de généralité) sa version continue à gauche sur $[0, T]$ et continue à droite en 0, c'est-à-dire que :

- $\varphi(t_i) = \varphi(t_i^-)$ (limite à gauche de $\varphi(\cdot)$ en t_i) pour tout $i = 1, \ldots, k$;
- $\varphi(0) = \varphi(0^+)$ (limite à gauche de $\varphi(\cdot)$ en 0).

L'ensemble des fonctions $\varphi(\cdot) : [0, T] \to \mathbb{R}^d$ ayant les propriétés décrites ci-dessus est un espace vectoriel réel, noté $C_M([0, T], \mathbb{R}^d)$. Il peut être structuré en espace préhilbertien grâce au produit scalaire :

$$(\varphi|\psi) := \int_0^T \langle \varphi(t), \psi(t) \rangle \, \mathrm{d}t,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^d$.

Une fonction $x(\cdot) : [0, T] \to \mathbb{R}^d$ est dite C^1 *par morceaux* sur $[0, T]$ si elle est C^1 sur $[0, T]$ ou s'il existe un nombre fini de points $t_0 = 0 < t_1 < \ldots < t_{k-1} < t_k = T$ tels que :

- $x(\cdot)$ est C^1 sur chaque intervalle $]t_{i-1}, t_i[$, $i = 1, \ldots, k$;
- en tout t_i ($i = 1, \ldots, k - 1$), $x(\cdot)$ a une dérivée à droite $\dot{x}_d(t_i)$ et cette dérivée à droite est égale à $\dot{x}(t_i^+)$;

– en tout t_i ($i = 1, \ldots, k-1$), $x(\cdot)$ a une dérivée à gauche $\dot{x}_g(t_i)$ et cette dérivée à gauche est égale à $\dot{x}(t_i^-)$.

Une telle fonction est nécessairement continue sur $[0,T]$, dérivable sur $[0,T]$ privé des points $t_1, \ldots, t_{k-1}$, et $\dot{x}$ est continue par morceaux sur $[0,T]$. On prolonge $\dot{x}(\cdot)$ aux points t_i en posant :

$$\dot{x}(0) := \dot{x}(0^+) \text{ et } \dot{x}(t_i) := \dot{x}(t_i^-) \text{ pour tout } i = 1, \ldots, k$$

de sorte que la nouvelle fonction ainsi définie sur $[0,T]$ se trouve dans $C_M([0,T], \mathbb{R}^d)$. L'ensemble des fonctions $x(\cdot) : [0,T] \to \mathbb{R}^d$ C^1 par morceaux sur $[0,T]$ est un espace vectoriel réel noté $C^1_M([0,T], \mathbb{R}^d)$. Toute fonction $\varphi(\cdot) \in C_M([0,T], \mathbb{R}^d)$ est (bornée et) intégrable sur $[0,T]$ (au sens de RIEMANN) ; on peut donc définir :

$$x(\cdot) : t \in [0,T] \mapsto x(t) := \int_0^t \varphi(s)\,\mathrm{d}s.$$

Alors, $x(\cdot) \in C^1_M([0,T], \mathbb{R}^d)$ avec :

$$\begin{array}{rcl} \dot{x}(t) & = & \varphi(t) \text{ si } t \neq t_i; \\ \dot{x}_+(t_i) & = & \varphi(t_i^+) \text{ pour tout } i = 0, \ldots, k-1; \\ \dot{x}_-(t_i) & = & \varphi(t_i^-) = \varphi(t_i) \text{ pour tout } i = 1, \ldots, k. \end{array}$$

Inversement, si $x(\cdot) \in C^1_M([0,T], \mathbb{R})$, on a toujours la représentation intégrale

$$x(t) = x(0) + \int_0^T \dot{x}(s)\,\mathrm{d}s, \quad t \in [0,T].$$

1.1.2. Première partie — Soient $A \in \mathscr{M}_n(\mathbb{R})$, $b(\cdot) \in C_M([0,T], \mathbb{R}^n)$, et $x_0 \in \mathbb{R}^n$. On définit :

$$t \in [0,T] \mapsto x(t) := \mathrm{e}^{tA} x_0 + \int_0^t \mathrm{e}^{(t-s)A} b(s)\,\mathrm{d}s. \tag{1.1}$$

1. — a) On suppose dans un premier temps $b(\cdot)$ continue sur $[0,T]$. Rappeler brièvement pourquoi $x(\cdot)$ définie en (1.1) est l'unique solution dans $C^1([0,T], \mathbb{R}^n)$ du problème de CAUCHY suivant :

$$\begin{cases} \dot{x}(t) & = & Ax(t) + b(t) \quad \text{pour tout } \; t \in [0,T]; \\ x(0) & = & x_0. \end{cases}$$

b) On suppose à présent $b(\cdot) \in C_M([0,T],\mathbb{R}^n)$. Montrer que $x(\cdot)$ définie en (1.1) est l'unique fonction de $C^1_M([0,T],\mathbb{R}^n)$ vérifiant :

$$\begin{cases} \dot{x}(t) &= Ax(t) + b(t) \quad \text{en tout point } t \text{ où } x(\cdot) \text{ est dérivable ;} \\ x(0) &= x_0. \end{cases}$$

Par extension, on dira encore que $x(\cdot)$ est solution du problème de CAUCHY noté en abrégé :

$$\begin{cases} \dot{x}(t) &= Ax(t) + b(t) \quad \text{sur } [0,T] \text{ ;} \\ x(0) &= x_0. \end{cases}$$

Soit $K \in \mathscr{M}_{n,m}(\mathbb{R})$ avec $m \leq n$. A toute fonction $u(\cdot) \in C_M([0,T],\mathbb{R}^m)$ on associe le problème de CAUCHY que voici :

$$(\mathscr{C}_u) \begin{cases} \dot{x}(t) &= Ax(t) + Ku(t) \quad \text{sur } [0,T] \text{ ;} \\ x(0) &= x_0. \end{cases} \tag{1.2}$$

On dit que $u(\cdot)$ est la *commande* (ou le *contrôle*) du système décrit par l'équation différentielle (1.2). On désigne par $x_u(\cdot)$ l'unique solution de $(\mathscr{C}_u)$.

2. — Soit $v(\cdot) \in C_M([0,T],\mathbb{R}^m)$.
Montrer qu'il existe $y_v(\cdot) \in C^1_M([0,T],\mathbb{R}^n)$ tel que :

$$\forall \lambda \in \mathbb{R}, \quad \forall u(\cdot) \in C_M([0,T],\mathbb{R}^m), \qquad x_{u+\lambda v} = x_u + \lambda y_v.$$

Préciser le problème de CAUCHY dont y_v est solution.

Soient α, β, γ des réels positifs ou nuls. On considère la fonction $J : C_M([0,T],\mathbb{R}^m) \to \mathbb{R}$ définie par :

$$J(u) := \int_0^T \left[\alpha\|x_u(t)\|^2 + \beta\|u(t)\|^2\right] \,\mathrm{d}t + \gamma\|x_u(T)\|^2$$

modélisant un coût que l'on cherche à rendre minimal.

3. — Montrer que $\lambda \in \mathbb{R} \mapsto J(u + \lambda v)$ est une fonction polynomiale du second degré en λ dont on explicitera les coefficients. Quel est le signe du coefficient de λ^2 ?

4. — a) Montrer qu'il existe une unique fonction $z_u(\cdot) \in C^1([0,T],\mathbb{R}^n)$ telle que

$$\begin{cases} \dot{z}_u(t) &= -A^T z_u(t) - 2\alpha x_u(t) \quad \text{sur } [0,T] \text{ ;} \\ z_u(T) &= 2\gamma x_u(T). \end{cases}$$

b) Exprimer

$$\langle z_u(T), y_v(T)\rangle + 2\alpha \int_0^T \langle x_u(t), y_v(t)\rangle \,\mathrm{d}t$$

à l'aide d'une intégrale faisant intervenir K, v et z_u.

5. — a) Déduire des questions précédentes que la dérivée en 0 de la fonction $\lambda \mapsto J(u + \lambda v)$ est

$$\int_0^T \left\langle K^T z_u(t) + 2\beta u(t), v(t)\right\rangle \,\mathrm{d}t.$$

b) Montrer que $u_0(\cdot) \in C_M([0, T], \mathbb{R}^m)$ est un minimiseur de J sur $C_M([0, T], \mathbb{R}^m)$ si et seulement si :

$$K^T z_{u_0}(t) + 2\beta u_0(t) = 0$$

pour tout $t \in [0, T]$.

1.1.3. Seconde partie — Soit U une partie convexe fermée (non vide) de $\mathbb{R}^m$; on définit $\mathscr{U}_{ad}$ comme l'ensemble des $u(\cdot) \in C_M([0, T], \mathbb{R}^m)$ tels que $u(t) \in U$ en tout point t de continuité de $u(\cdot)$.

$\mathscr{U}_{ad}$ est *l'ensemble des commandes admissibles* pour notre système gouverné par l'équation différentielle (1.2)[1.1].

6. — a) Montrer l'équivalence suivante, relative à la commande $u(\cdot) \in C_M([0, T], \mathbb{R}^m)$:

$$u(\cdot) \in \mathscr{U}_{ad} \Leftrightarrow u(t) \in U \text{ pour tout } t \in [0, T].$$

b) Montrer que $\mathscr{U}_{ad}$ est une partie convexe fermée (non vide) de $C_M([0, T], \mathbb{R}^m)$ (pour la topologie d'espace préhilbertien définie à partir du produit scalaire $(.|.)$ rappelé au début).

7. — Montrer que $u_0(\cdot) \in \mathscr{U}_{ad}$ est un minimiseur de J sur $\mathscr{U}_{ad}$ si et seulement si : pour tout $v(\cdot) \in \mathscr{U}_{ad}$ et tout $t \in [0, T]$,

$$\left\langle K^T z_{u_0}(t) + 2\beta u_0(t), v(t) - u_0(t)\right\rangle \geq 0. \tag{1.3}$$

[1.1] $U = \mathbb{R}^m$ dans la première partie du problème ; $U = [-a, +a]^m$ est un exemple fréquent dans les applications.

8. Application. — Dans l'exemple qui suit, on aura $n = 2$ et $m = 1$. On choisit $U = [-a, +a]$, et $k > 0$ est un coefficient donné. La position d'un modèle sur la droite réelle à l'instant $t \in [0, T]$ est repérée par $p(t)$, et l'évolution de $p(t)$ est régie par l'équation différentielle scalaire du second ordre suivante : $\dot{A}$

$$\ddot{p}(t) = -ku(t). \tag{1.4}$$

Ici $u(\cdot) \in C_M([0, T], \mathbb{R})$, $u(t) \in [-a, +a]$ pour tout $t \in [0, T]$, et on entend par (1.4) une fonction $p : [0, T] \to \mathbb{R}$ dérivable et dont la dérivée $\dot{p}(\cdot)$ est C^1 par morceaux sur $[0, T]$, avec :

$$\dot{\dot{p}}\ (t) = -ku(t) \text{ sur } [0, T].$$

a) Écrire ce problème sous la forme (1.2) avec des matrices A et K que l'on déterminera.

Soient p_0 et $\dot{p}_0$ deux nombres réels. Montrer qu'il existe une unique fonction $p_u(\cdot)$ solution de ce problème telle que $p_u(0) = p_0$ et $\dot{p}_u(0) = \dot{p}_0$.

b) On choisit $J(u) := [p_u(T)]^2 + [\dot{p}_u(T)]^2$ comme fonction coût. Montrer que $z_u : [0, T] \to \mathbb{R}^2$ est une fonction affine de t.

c) Soit $u_0(\cdot) \in \mathscr{U}_{ad}$ un minimiseur de J sur $\mathscr{U}_{ad}$, que l'on suppose tel que $(p_{u_0}(T), \dot{p}_{u_0}(T)) \neq (0, 0)$. Montrer que $u_0(\cdot)$ est constante par morceaux, avec deux morceaux au plus, les constantes étant a ou $-a$.

d) On suppose ici que $p_0 = 1 + \frac{T^2}{2}(1 + \frac{ka}{2})$ et $\dot{p}_0 = -T/2$. On considère alors :

$$t \in [0, T] \mapsto u_0(t) = a \text{ si } t \in [0, \frac{T}{2}], \quad -a \text{ si } t \in]\frac{T}{2}, T].$$

- Calculer $p_{u_0}(T)$ et $\dot{p}_{u_0}(T)$.
- Vérifier que $u_0(\cdot)$ est un minimiseur de J sur $\mathscr{U}_{ad}$.
- On particularise encore en faisant $ka = 1/4$ et $T = 4$. La fonction $u_0(\cdot)$ proposée au-dessus est-elle l'unique minimiseur de J sur $\mathscr{U}_{ad}$?

1.2. Corrigé —

1.2.1. Première partie —

1. — a) Dans la mesure où $b(\cdot)$ est supposée continue sur $[0, T]$, le théorème de CAUCHY-LIPSCHITZ pour les équations différentielles (vectorielles) linéaires permet de dire que le problème de CAUCHY

$$\begin{cases} \dot{x}(t) &= Ax(t) + b(t) \quad \text{pour tout } t \in [0, T] \\ x(0) &= x_0 \end{cases}$$

admet une et une seule solution $x(\cdot) \in C^1([0,T],\mathbb{R}^n)$. Il n'y a plus qu'à dériver la fonction $x(\cdot)$ proposée

$$t \mapsto x(t) = e^{tA}x_0 + e^{tA}\int_0^t e^{-sA}b(s)\,\mathrm{d}s \tag{1.5}$$

pour constater que

$$\begin{aligned} \dot{x}(t) &= Ae^{tA}x_0 + Ae^{tA}\int_0^t e^{-sA}b(s)\,\mathrm{d}s + e^{tA}e^{-tA}b(t) \\ &= Ax(t) + b(t). \end{aligned}$$

On reconnaît en $\left\{t \mapsto e^{tA}x_0, x_0 \in \mathbb{R}^n\right\}$ le sous-espace vectoriel réel (de dimension n) constituant l'ensemble des solutions de l'équation différentielle linéaire «homogénéisée» $\dot{x}(t) = Ax(t)$, et en $t \mapsto \int_0^t e^{(t-s)A}b(s)\,\mathrm{d}s$ une solution de l'équation différentielle complète $\dot{x}(t) = Ax(t) + b(t)$.

b) Dans le cas où $b(\cdot)$ est continue par morceaux sur $[0,T]$, la fonction $x(\cdot)$ proposée en (1.5) est C^1 par morceaux sur $[0,T]$ avec :

$$\begin{cases} \dot{x}(t) = Ax(t) + b(t) & \text{en tout point où } b(\cdot) \text{ est continue ;} \\ \dot{x}_d(t_i) = Ax(t_i) + b({t_i}^+) & \text{et} \\ \dot{x}_g(t_i) = Ax(t_i) + b({t_i}^-) & \text{en tout point } t_i \text{ où } b(\cdot) \text{ est discontinue.} \end{cases}$$

Les points de non dérivabilité de $x(\cdot)$ sont exactement les points de discontinuité de $b(\cdot)$. Donc $x(\cdot)$ est une fonction solution dans $C^1_M([0,T],\mathbb{R}^n)$ du problème de CAUCHY suggéré.

Reste à s'assurer de l'unicité. Si $y(\cdot)$ et $z(\cdot)$ dans $C^1_M([0,T],\mathbb{R}^n)$ vérifient ce qui est demandé, la fonction $y(\cdot) - z(\cdot)$ est continue, dérivable et de dérivée nulle sur une suite d'intervalles $]t_i, t_{i+1}[$ couvrant $[0,T]$ à l'exception d'un nombre fini de points t_i (d'accord ?). Par conséquent, $y(\cdot) - z(\cdot)$ est constante sur $[0,T]$ (donc nulle puisque $y(0) = z(0)$).

2. — On a par définitions :

$$\begin{cases} \dot{x}_{u+\lambda v}(t) = Ax_{u+\lambda v}(t) + K\left[u(t) + \lambda v(t)\right] \text{ sur } [0,T], \\ x_{u+\lambda v}(0) = x_0 \end{cases}$$

$$\begin{cases} \dot{x}_u(t) = Ax_u(t) + Ku(t) \text{ sur } [0,T], \\ x_u(0) = x_0. \end{cases}$$

Par suite $\theta := x_{u+\lambda v} - x_u$ ($\in C^1_M([0,T],\mathbb{R}^n)$) vérifie

$$\begin{cases} \dot{\theta}(t) = A\theta(t) + \lambda Kv(t) \text{ sur } [0,T], \\ \dot{\theta}(0) = 0. \end{cases}$$

Ainsi, de par les résultats d'existence et d'unicité de (*1*.a), $\theta = \lambda y_v$, où $y_v(\cdot) \in C^1_M([0,T], \mathbb{R}^n)$ est la solution du problème de CAUCHY suivant :

$$\begin{cases} \dot{y}_v(t) &= Ay_v(t) + Kv(t) \text{ sur } [0,T]\,; \\ y_v(0) &= 0. \end{cases} \tag{1.6}$$

3. — Le coût $J(u)$ comprend une partie répartie tout au long de l'évolution du système (via l'intégrale de 0 à T) et une partie ne dépendant que de l'état terminal $x_u(T)$ du système.

Le développement de $J(u + \lambda v)$ ne pose pas de difficultés ; après quelques calculs élémentaires, on arrive à :

$$\begin{aligned} &J(u + \lambda v) = J(u) \\ &+\lambda^2 \left\{ \int_0^T \left[\alpha \left\| y_v(t) \right\|^2 + \beta \left\| v(t) \right\|^2 \right] \mathrm{d}t + \gamma \left\| y_v(T) \right\|^2 \right\} \\ &+2\lambda \left\{ \int_0^T \left[\alpha \left\langle x_u(t), y_v(t) \right\rangle + \beta \left\langle u(t), v(t) \right\rangle \right] \mathrm{d}t + \gamma \left\langle x_u(T), y_v(T) \right\rangle \right\}. \end{aligned} \tag{1.7}$$

De par sa forme même, le coefficient de λ^2 est positif.

4. — a) Comme $-2\alpha x_u(\cdot)$ est continue, il suffit d'appliquer comme en (*1*.a) le théorème de CAUCHY-LIPSCHITZ pour les équations différentielles linéaires ; prendre garde toutefois au fait que la condition imposée dans le problème de CAUCHY associé est une condition *finale*, $z_u(T) = 2\gamma x_u(T)$. C'est la continuité de $-2\alpha x_u(\cdot)$ dans le second membre de l'équation différentielle qui est importante pour garantir que z_u est bien de classe C^1 sur $[0,T]$. La forme explicite de $z_u(t)$, adaptation de celle proposée en (1.5), est :

$$\forall t \in [0,T], z_u(t) = 2\gamma e^{(T-t)A^T} x_u(T) + 2\alpha \int_t^T e^{(s-t)A^T} x_u(s)\,\mathrm{d}s.$$

b) La fonction $t \mapsto \sigma(t) := \langle z_u(t), y_v(t) \rangle$ est dérivable sauf en un nombre fini de points de $[0,T]$, avec :

$$\dot{\sigma}(t) = \langle \dot{z}_u(t), y_v(t) \rangle + \langle z_u(t), \dot{y}_v(t) \rangle\,,$$

ce qui, au vu des relations différentielles vérifiées par z_u et y_v, conduit à :

$$\begin{aligned} \dot{\sigma}(t) &= \left\langle -A^T z_u(t), y_v(t) \right\rangle - 2\alpha \left\langle x_u(t), y_v(t) \right\rangle \\ &\qquad + \left\langle z_u(t), Ay_v(t) \right\rangle + \left\langle z_u(t), Kv(t) \right\rangle \\ &= -2\alpha \left\langle x_u(t), y_v(t) \right\rangle + \left\langle z_u(t), Kv(t) \right\rangle. \end{aligned}$$

Comme $\sigma(0) = 0$ (car $y_v(0) = 0$),

$$\sigma(T) - \sigma(0) = \sigma(T) = \int_0^T \dot{\sigma}(t)\,\mathrm{d}t,$$

soit

$$\langle z_u(T), y_v(T)\rangle = \int_0^T \left[-2\alpha \langle x_u(t), y_v(t)\rangle + \langle z_u(t), Kv(t)\rangle\right] \mathrm{d}t.$$

En définitive,

$$\langle z_u(T), y_v(T)\rangle + 2\alpha \int_0^T \langle x_u(t), y_v(t)\rangle\, dt = \int_0^T \langle z_u(t), Kv(t)\rangle\, dt. \quad (1.8)$$

5. — a) D'après l'expression de $J(u + \lambda v)$ obtenue en (1.7), si $c(\lambda)$ désigne $J(u + \lambda v)$, on a :

$$\frac{c(\lambda) - c(0)}{\lambda} = 2\left\{\int_0^T \left[\alpha \langle x_u(t), y_v(t)\rangle + \beta \langle u(t), v(t)\rangle\, \mathrm{d}t \right.\right.$$
$$\left.\left. + \gamma \langle x_u(t), y_v(t)\rangle\right]\right\} + o(1).$$

D'après le résultat de la question précédente et le fait que $z_u(T) = 2\gamma x_u(T)$, on simplifie le quotient différentiel ci-dessus en

$$\begin{aligned}\frac{c(\lambda) - c(0)}{\lambda} &= \int_0^T \langle z_u(t), Kv(t)\rangle\, \mathrm{d}t + 2\beta \int_0^T \langle u(t), v(t)\rangle\, \mathrm{d}t + o(1) \\ &= \int_0^T \left\langle K^T z_u(t) + 2\beta u(t), v(t)\right\rangle \mathrm{d}t + o(1).\end{aligned}$$

Au final, c est dérivable en 0 avec

$$\dot{c}\,(0) = \int_0^T \left\langle K^T z_u(t) + 2\beta u(t), v(t)\right\rangle \mathrm{d}t. \quad (1.9)$$

b) D'après l'expression de J le long de toute droite passant par u et dirigée par v (fonction trinôme $a\lambda^2 + b\lambda + c$ avec $a \geq 0$), J est convexe sur $C_M([0, T], \mathbb{R}^m)$. Se rappelant la structure d'espace préhilbertien réel de $C_M([0, T], \mathbb{R}^n)$, ce que dit le résultat de la question précédente est que J a une sorte de «gradient en u», noté $\nabla J(u_0)$, qui est

$$\nabla J(u_0) = K^T z_u(\cdot) + 2\beta u(\cdot) \qquad (\in C_M([0, T], \mathbb{R}^m))\,.$$

Ainsi, $u_0(\cdot)$ est un minimiseur de J sur $C_M([0, T], \mathbb{R}^m)$ si et seulement si $\nabla J(u_0) = 0$.