



PROJET ACADÉMIQUE APPROFONDI

Optimisation du retour à l'état d'équilibre

Étude variationnelle et numérique d'un pendule contrôlé

Auteur :

Paul-Emmanuel Meyrignac

Avec l'accompagnement scientifique de :

Badre Mounda (chercheur en géométrie algébrique)

Domaines :

Analyse numérique — Contrôle optimal — Analyse fonctionnelle

Résumé

Ce projet étudie la stabilisation en temps fini d'un pendule simple par une commande externe, tout en minimisant un coût quadratique sur l'effort de contrôle. Après modélisation par une équation différentielle non linéaire, la dynamique est simulée numériquement puis linéarisée dans le régime des petites oscillations afin de reformuler le problème comme une optimisation quadratique sous contraintes dynamiques. Une approche variationnelle permet ensuite d'établir existence et unicité de la solution optimale.

avril–juin 2025

Table des matières

1	Introduction	2
2	Modélisation du pendule simple	3
2.1	Équation différentielle	3
2.2	Réécriture sous forme d'un système	3
3	Résolution numérique : Euler explicite	4
3.1	Principe de la méthode	4
3.2	Application au pendule	4
3.3	Étude de différentes fonctions de contrôle	5
4	Formulation du problème d'optimisation	7
4.1	Objectif	7
4.2	Discrétisation du problème	7
4.3	Linéarisation du problème dans le cas des petites oscillations	8
4.4	Caractéristiques du problème simplifié	10
4.4.1	Etude de la fonction objectif	10
4.4.2	Étude de l'ensemble des contrôles admissibles	11
4.4.3	Conclusion sur l'existence et l'unicité de la solution du problème . .	12
5	Méthode du gradient projeté pour la commande optimale discrète	13
5.1	Principe de la méthode du gradient	13
5.2	Projection sur l'ensemble admissible	13
5.3	Algorithme du gradient projeté	14
5.4	Résolution numérique du problème d'optimisation	16
6	Formulation variationnelle rigoureuse	17
6.1	Formulation variationnelle	18
6.2	Énoncé et démonstration du théorème de Lax–Milgram	19
6.3	Application de Lax–Milgram	20
6.3.1	Conclusion : unicité de la solution optimale	21
7	Caractérisation optimale et suite naturelle après Lax–Milgram	22
8	Approche Hamiltonienne : cohérence avec la solution variationnelle	24
9	Analyse fonctionnelle avancée : compacité, spectre et stabilité	26
A	Étude complémentaire des fonctions de contrôle	28
B	Gradient et hessienne d'une forme quadratique	30
C	Projeté orthogonal sur le sous-espace affine	32
D	Sur l'égalité $\text{rang}(MM^T) = \text{rang}(M)$	34
E	Bibliographie	34

1 Introduction

Ce projet traite du retour à l'état d'équilibre stable en temps fini d'un pendule simple via une force de contrôle $t \mapsto h(t)$, que l'on souhaite optimiser afin de minimiser l'énergie consommée. Après avoir modélisé le système dynamique par une équation différentielle non linéaire, nous en dérivons une version discrétisée à l'aide du schéma d'Euler explicite. Cette phase initiale permet d'analyser le comportement du pendule sous différentes fonctions de contrôle choisies empiriquement (constante, sinusoïdale, impulsion, rétroaction). Nous nous plaçons par la suite dans le cadre des petites oscillations, ce qui permet de linéariser la dynamique du système.

Une fois le problème d'optimisation discrétisé, il prend la forme d'une minimisation quadratique sous contraintes dynamiques linéaires. Nous en proposons une résolution numérique via la méthode du gradient projeté à pas fixe, en détaillant sa mise en œuvre, sa convergence, et la construction de la projection sur l'ensemble admissible.

Enfin, nous établissons une formulation variationnelle rigoureuse du problème d'optimisation dans un espace de Sobolev adapté, et montrons, via le théorème de Lax–Milgram, l'existence et l'unicité de la solution optimale. L'analyse conduit à une équation d'Euler–Lagrange d'ordre 4 que nous résolvons explicitement. Nous concluons en reliant cette solution à une approche hamiltonienne via le principe de Pontryagin, assurant la cohérence entre méthode variationnelle et contrôle optimal.

2 Modélisation du pendule simple

2.1 Équation différentielle

L'étude est conduite sur l'intervalle temporel $[0,1]$.

Le pendule simple sans frottement soumis à une force de contrôle externe $t \mapsto h(t)$ est modélisé par l'équation :

$$\forall t \in [0, 1] \quad \ddot{\theta}(t) + \omega^2 \sin(\theta(t)) = h(t).$$

avec :

- $\theta(t)$ l'angle du pendule avec la verticale à l'instant t ,
- $\omega := \sqrt{\frac{g}{l}}$
- g l'accélération gravitationnelle,
- l la longueur du fil,
- $h(t)$ la force de contrôle externe, au moins continue.

2.2 Réécriture sous forme d'un système

Transformons l'équation en un système d'ordre 1. Pour cela on utilise d'abord la relation $\frac{d\theta}{dt} = \dot{\theta}$ qui n'est rien de plus que la définition de $\frac{d\theta}{dt}$. On utilise ensuite la relation $\ddot{\theta} = -\omega^2 \sin(\theta) + h$ issue de l'équation différentielle, que l'on réécrit, par définition de $\dot{\theta}$, $\frac{d\dot{\theta}}{dt} = -\omega^2 \sin(\theta) + h$. On obtient le système suivant :

$$\begin{cases} \frac{d\theta}{dt}(t) = \dot{\theta}(t) \\ \frac{d\dot{\theta}}{dt}(t) = -\omega^2 \sin(\theta(t)) + h(t) \end{cases}$$

Cette réécriture sous la forme d'un système d'ordre 1 i.e ne faisant intervenir que des dérivées d'ordre 1 par rapport au temps, permet la résolution numérique de ce problème grâce la méthode d'Euler explicite.

3 Résolution numérique : Euler explicite

3.1 Principe de la méthode

Si l'on considère une équation différentielle ordinaire d'ordre 1 avec une condition initiale :

$$\begin{cases} y'(t) = f(t, y(t)) \text{ si } t \in [0, 1] \\ y(0) = y_0 \end{cases}$$

Connaissant la valeur de y à un instant $t_n \in [0, 1[$, on peut, sous réserve que y soit suffisamment régulière (par exemple de classe \mathcal{C}^1), exprimer la valeur de y à l'instant suivant $t_{n+1} \in]t_n, 1]$ à l'aide de la formule :

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

donc :

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt$$

Tout l'enjeu réside dans l'approximation de l'intégrale $\int_{t_n}^{t_{n+1}} f(t, y(t)) dt$.

La méthode d'Euler explicite consiste à approximer cette intégrale par l'aire algébrique du rectangle de hauteur $f(t_n, y(t_n))$ et de largeur $t_{n+1} - t_n$. Soit :

$$\int_{t_n}^{t_{n+1}} f(t, y(t)) dt \approx f(t_n, y(t_n))(t_{n+1} - t_n)$$

D'où le schéma itératif :

$$\begin{cases} y_{n+1} = y_n + (t_{n+1} - t_n)f(t_n, y_n) \\ y_0 = y(0) \end{cases}$$

Dans notre cas,

$$y := \begin{pmatrix} \theta \\ \dot{\theta} \end{pmatrix} \quad \text{et pour tout } t \in [0, 1], \quad f(t, y(t)) := \begin{pmatrix} \dot{\theta}(t) \\ -\frac{g}{l} \sin(\theta(t)) + h(t) \end{pmatrix}.$$

3.2 Application au pendule

On obtient ce schéma itératif, permettant de calculer les valeurs de $\theta(t)$ et $\dot{\theta}(t)$ de proche en proche :

$$\begin{cases} \theta^{n+1} = \theta^n + (t_{n+1} - t_n) \cdot \dot{\theta}^n \\ \dot{\theta}^{n+1} = \dot{\theta}^n + (t_{n+1} - t_n) \cdot \left(-\frac{g}{l} \sin(\theta^n) + h(t_n) \right) \end{cases}$$

Pour simplifier les calculs, on introduit une subdivision régulière de l'intervalle $[0, 1]$, définie par la suite $(t_n)_{0 \leq n \leq N}$, où $N \in \mathbb{N}^*$ est le nombre de sous-intervalles. Ainsi, pour tout $n \in \{0, \dots, N-1\}$ la différence $t_{n+1} - t_n$ est constante.

On pose alors :

$$\Delta t := t_{n+1} - t_n = \frac{1}{N}.$$

De plus,

$$\forall n \in \{0, \dots, N\} \quad t_n = \frac{n}{N}$$

Le schéma devient :

$$\begin{cases} \theta_{n+1} = \theta_n + \Delta t \cdot \dot{\theta}_n, \\ \dot{\theta}_{n+1} = \dot{\theta}_n + \Delta t \cdot \left(-\frac{g}{l} \sin(\theta_n) + h(t_n) \right). \end{cases}$$

3.3 Étude de différentes fonctions de contrôle

Paramètres de la simulation

Nous choisissons les conditions initiales suivantes pour le pendule :

$$\theta(0) = 0.8 \text{ rad}, \quad \dot{\theta}(0) = 0 \text{ rad.s}^{-1}.$$

L'intervalle temporel de la simulation est fixé à $[0, T]$ avec $T = 10$ secondes. Ce choix est motivé par le fait qu'un intervalle trop court, comme $[0, 1]$, ne permet pas d'observer efficacement le comportement du pendule soumis à différentes fonctions de contrôle h . Nous effectuons une subdivision régulière de cet intervalle à l'aide de $N = 20000$ points, ce qui permet d'assurer une bonne précision numérique.

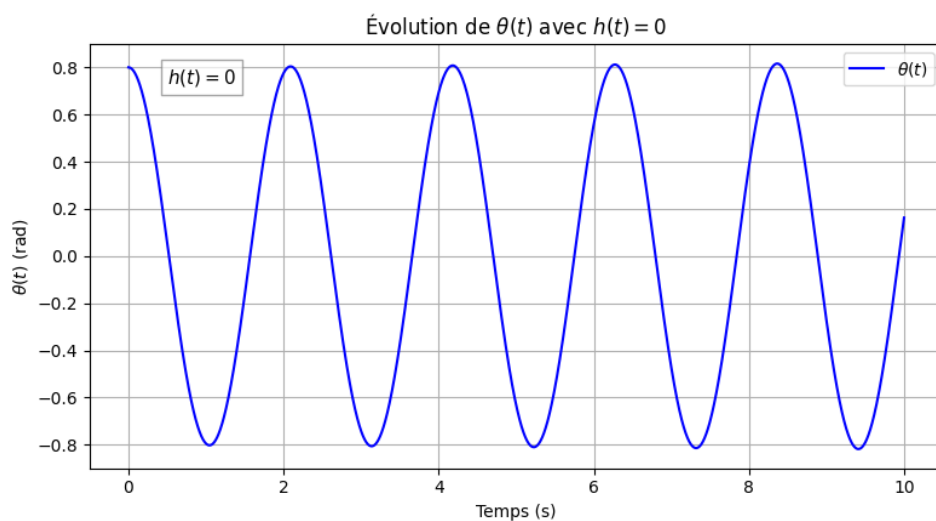
Les paramètres physiques du système sont les suivants :

- Accélération gravitationnelle : $g = 9,81 \text{ m} \cdot \text{s}^{-2}$
- Longueur du pendule : $l = 1,0 \text{ m}$

Nous testons différentes fonctions de contrôle $t \mapsto h(t)$ afin d'observer leur effet sur le comportement du pendule. Chaque graphique présente l'évolution de $\theta(t)$ pour un choix donné de $t \mapsto h(t)$.

Cas 1 : $h(t) = 0$ (aucun contrôle)

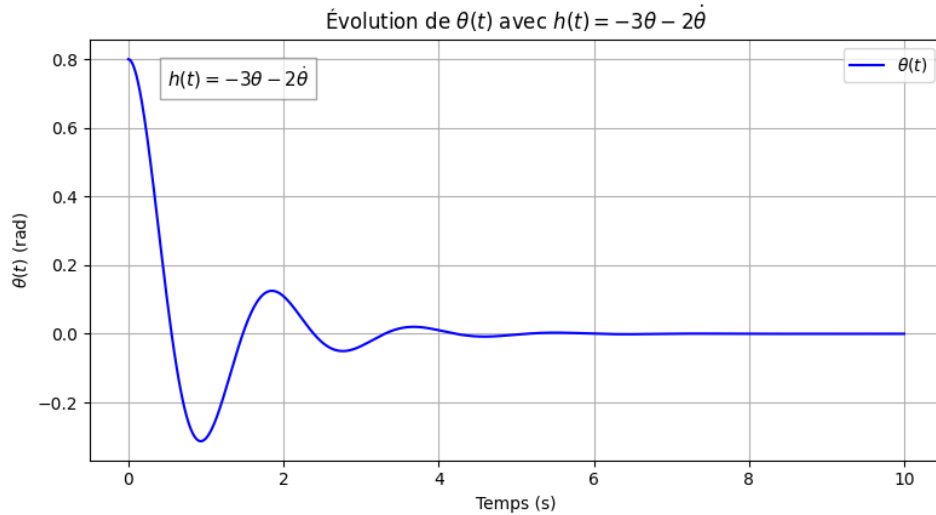
Le pendule oscille librement sans intervention externe.



Évolution naturelle sans contrôle externe $h(t) = 0$

Cas 2 : $h(t) = -3\theta(t) - 2\dot{\theta}(t)$ (rétroaction stabilisante)

Ce contrôle agit comme un correcteur proportionnel-dérivé : il cherche à ramener θ vers zéro.



Contrôle par rétroaction $h(t) = -3\theta(t) - 2\dot{\theta}(t)$

Les autres fonctions de contrôle testées, ainsi que l'ensemble des simulations, sont présentées en Annexe A.

Conclusion

Les simulations précédentes montrent que certaines formes de contrôle, notamment par rétroaction, permettent de stabiliser efficacement le pendule. Toutefois, ces approches reposent sur des choix empiriques de la fonction $t \mapsto h(t)$, sans garantie que l'énergie dépensée soit minimale.

Cela motive une démarche plus rigoureuse, dans laquelle nous cherchons à déterminer, parmi toutes les fonctions admissibles, celle qui ramène le pendule à l'équilibre tout en minimisant l'énergie fournie.

Nous passons donc à une formulation mathématique du problème d'optimisation, en définissant précisément le critère de coût et les contraintes dynamiques associées.

4 Formulation du problème d'optimisation

4.1 Objectif

Nous cherchons maintenant à déterminer une fonction de contrôle $t \mapsto h(t)$ définie sur $[0, 1]$, permettant de ramener le pendule à l'état d'équilibre $(\theta, \dot{\theta}) = (0, 0)$ tout en minimisant l'énergie nécessaire pour y parvenir.

L'énergie consommée est modélisée par une fonctionnelle de coût J :

$$\forall J \in \mathcal{C}^0([0, 1], \mathbb{R}) \quad J(h) := \int_0^1 h(t)^2 dt.$$

Ce critère quadratique est standard en contrôle optimal car il pénalise les efforts excessifs et conduit à un problème convexe bien posé.

Plutôt que de traiter immédiatement le problème dans son cadre fonctionnel continu, nous en étudions d'abord une version discrétisée, adaptée à une résolution numérique. Cette première formulation prépare le passage à une approche plus abstraite, qui sera développée en fin de rapport à travers une formulation variationnelle rigoureuse dans un espace de Sobolev adapté.

4.2 Discrétisation du problème

En reprenant le schéma d'Euler explicite, nous avons déjà obtenu une discrétisation de l'équation différentielle du pendule sous la forme :

$$\begin{cases} \theta_{n+1} = \theta_n + \Delta t \cdot \dot{\theta}_n, \\ \dot{\theta}_{n+1} = \dot{\theta}_n + \Delta t \cdot \left(-\frac{g}{l} \sin(\theta_n) + h_n \right), \end{cases}$$

où $h_n := h(t_n)$.

Soit $N \in \mathbb{N}^*$. On discrétise la fonctionnelle de coût par la N -ième somme de Riemann associée à la subdivision $0 = t_0 < t_1 < \dots < t_N = 1$. La fonctionnelle de coût devient alors :

$$J(h) \approx \sum_{n=0}^{N-1} \Delta t \cdot h_n^2$$

Cette approximation est justifiée par le fait que, pour h continue,

$$\sum_{i=0}^{N-1} \Delta t \cdot h(t_i)^2 \xrightarrow{N \rightarrow +\infty} \int_0^1 h(t)^2 dt$$

Nous cherchons donc à résoudre le problème d'optimisation discret suivant :

$$\left\{ \begin{array}{l} \min_{(h_0, \dots, h_{N-1}) \in \mathbb{R}^N} \sum_{n=0}^{N-1} \Delta t \cdot h_n^2 \\ \text{sous les contraintes :} \\ \forall n \in \{0, \dots, N-1\}, \quad \theta_{n+1} = \theta_n + \Delta t \cdot \dot{\theta}_n, \\ \forall n \in \{0, \dots, N-1\}, \quad \dot{\theta}_{n+1} = \dot{\theta}_n + \Delta t \cdot \left(-\frac{g}{l} \sin(\theta_n) + h_n \right), \\ \theta_0, \dot{\theta}_0 \in \mathbb{R} \quad \theta_N = 0, \quad \dot{\theta}_N = 0 \end{array} \right. \quad (1)$$

Complexité du problème non linéaire

Ce problème d'optimisation est difficile à résoudre en raison de la non-linéarité introduite par le terme $\sin(\theta_n)$ dans les contraintes dynamiques. Cette non-linéarité rend l'ensemble des trajectoires admissible non convexe, ce qui complique à la fois l'analyse théorique (existence et unicité de la solution) et la résolution numérique (présence possible de minima locaux, convergence non garantie des algorithmes classiques).

Afin de contourner ces difficultés, nous nous plaçons par la suite dans le régime des petites oscillations, c'est-à-dire lorsque l'angle reste suffisamment proche de zéro ($|\theta_n| \leq 0,2$ rad). Dans ce cas, l'approximation $\sin(\theta_n) \approx \theta_n$ est valable, et les contraintes dynamiques deviennent affines en $\begin{pmatrix} \theta_n \\ \dot{\theta}_n \end{pmatrix}$. Cette simplification permet de reformuler le problème comme une minimisation quadratique sous contraintes affines, bien mieux adapté à une résolution numérique rigoureuse et efficace.

4.3 Linéarisation du problème dans le cas des petites oscillations

Comme expliqué dans le dernier paragraphe, on se place dans le cas de petites oscillations ($|\theta_n| \leq 0,2$ rad) et on utilise $\sin(\theta_n) \approx \theta_n$.

On considère \mathbb{R}^N muni de son produit scalaire canonique :

$$\forall (x, y) \in \mathbb{R}^N \times \mathbb{R}^N \quad \langle x, y \rangle := \sum_{n=1}^N x_n y_n = x^T y$$

où x^t désigne la transposée de x vu comme élément de $\mathcal{M}_{n,1}(\mathbb{R})$ et le produit utilisé est le produit matriciel usuel. Le problème devient :

$$\begin{cases} \min_{h=(h_0, \dots, h_{N-1}) \in \mathbb{R}^N} \Delta t \|h\|^2 \\ \forall n \in \{0, \dots, N-1\}, \quad \theta_{n+1} = \theta_n + \Delta t \cdot \dot{\theta}_n \\ \forall n \in \{0, \dots, N-1\}, \quad \dot{\theta}_{n+1} = \dot{\theta}_n + \Delta t \cdot \left(-\frac{g}{l}\theta_n + h_n\right) \\ \theta_0, \dot{\theta}_0 \in \mathbb{R}, \quad \theta_N = 0, \quad \dot{\theta}_N = 0 \end{cases}$$

Notons :

$$A := \begin{pmatrix} 1 & \Delta t \\ -\frac{g}{l}\Delta t & 1 \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R}), \quad B := \begin{pmatrix} 0 \\ \Delta t \end{pmatrix} \in \mathbb{R}^2 \quad \text{et} \quad \forall n \in \{0, \dots, N-1\} \quad Y_n := \begin{pmatrix} \theta_n \\ \dot{\theta}_n \end{pmatrix} \in \mathbb{R}^2.$$

Le problème se réécrit :

$$\begin{cases} \min_{h=(h_0, \dots, h_{N-1}) \in \mathbb{R}^N} \Delta t \|h\|^2 \\ \forall n \in \{0, \dots, N-1\}, \quad Y_{n+1} = AY_n + h_n B \\ Y_0 \in \mathbb{R}^2 \\ Y_N = \begin{pmatrix} \theta_N \\ \dot{\theta}_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$$

Notons que A est inversible. En effet,

$$\det(A) = 1 + \Delta t^2 \cdot \frac{g}{l} > 0$$

Soit $n \in \{0, \dots, N-1\}$.

En multipliant la relation $Y_{n+1} = AY_n + h_n B$ à gauche par $A^{-(n+1)}$, on obtient :

$$A^{-(n+1)}Y_{n+1} = A^{-n}Y_n + h_n A^{-(n+1)}B.$$

En sommant cette égalité pour n allant de 0 à $N-1$, on obtient par télescopage :

$$A^0 Y_0 + \sum_{n=0}^{N-1} h_n A^{-(n+1)} B = A^{-N} Y_N \quad (*)$$

Comme $Y_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, on obtient :

$$\sum_{n=0}^{N-1} h_n A^{-(n+1)} B = -A^0 Y_0$$

Il est plus commode de travailler avec les puissances de A qu'avec les puissances de A^{-1} , donc on multiplie par A^N à gauche. Cela donne :

$$\sum_{n=0}^{N-1} h_n A^{N-(n+1)} B = -A^N Y_0.$$

Posons :

$$M := \begin{pmatrix} A^{N-1}B & A^{N-2}B & \dots & B \end{pmatrix} \in \mathbb{R}^{2 \times N}$$

La contrainte devient :

$$\boxed{Mh = -A^N Y_0}$$

Résoudre le problème initial implique la résolution de ce problème de minimisation :

$$\begin{cases} \min_{h \in \mathbb{R}^N} \Delta t \|h\|^2 \\ Mh = -A^N Y_0 \\ Y_0 \in \mathbb{R}^2 \end{cases}$$

Ces problèmes sont en fait équivalents.

En effet, si un vecteur $h \in \mathbb{R}^N$ vérifie $Mh = -A^N Y_0$, alors on peut reconstruire la trajectoire complète $(Y_n)_{0 \leq n \leq N}$ par récurrence en posant :

$$\begin{cases} Y_0 \in \mathbb{R}^2 \\ \forall n \in \{0, \dots, N-1\}, \quad Y_{n+1} = AY_n + h_n B \end{cases}$$

En reprenant l'équation (*), une telle trajectoire vérifie :

$$\begin{aligned} Y_N &= A^N Y_0 + \sum_{k=0}^{N-1} h_k A^{N-1-k} B \\ \iff Y_N &= A^N Y_0 + Mh \\ \iff Y_N &= A^N Y_0 + (-A^N Y_0) \\ \iff Y_N &= 0 \end{aligned}$$

Ainsi, la trajectoire $(Y_n)_{0 \leq n \leq N}$ définie par récurrence à partir de Y_0 et du vecteur h respecte bien toutes les contraintes du problème initial.

$$\left\{ \begin{array}{l} \min_{h=(h_0, \dots, h_{N-1}) \in \mathbb{R}^N} \Delta t \cdot \|h\|^2 \\ \forall n \in \{0, \dots, N-1\} \ Y_{n+1} = AY_n + h_n B \\ Y_0 \in \mathbb{R}^2 \\ Y_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \min_{h \in \mathbb{R}^N} \Delta t \cdot \|h\|^2 \\ Mh = -A^N Y_0 \\ Y_0 \in \mathbb{R}^2 \end{array} \right.$$

Dans le problème de droite, les contraintes dynamiques et la condition finale ont été condensées en une unique contrainte affine en $h \in \mathbb{R}^n$ de la forme $Mh = -A^N Y_0$. On passe ainsi d'un ensemble de N contraintes couplées et une condition finale à une seule équation affine en $h \in \mathbb{R}^n$, ce qui simplifie considérablement l'analyse du problème. On étudie donc ce nouveau problème.

4.4 Caractéristiques du problème simplifié

4.4.1 Etude de la fonction objectif

Soit $h \in \mathbb{R}^N$.

Notons I_N la matrice identité de $\mathcal{M}_N(\mathbb{R})$.

On peut écrire :

$$\begin{aligned} \Delta t \|h\|^2 &= \Delta t (h^T h) \\ \Delta t \|h\|^2 &= h^T \Delta t I_N h \quad (\text{car } I_N h = h) \\ \Delta t \|h\|^2 &= \frac{1}{2} h^T (2\Delta t I_N) h \end{aligned}$$

Posons :

$$Q := 2\Delta t I_N = (Q_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$$

Alors Q est clairement symétrique et définie positive, car elle est diagonale et ses coefficients diagonaux sont strictement positifs.

On a alors :

$$\Delta t \|h\|^2 = \frac{1}{2} h^T Q h$$

La fonction objectif se réécrit :

$$J : \begin{cases} \mathbb{R}^N \rightarrow \mathbb{R} \\ h \mapsto \frac{1}{2} h^T Q h \end{cases}$$

Par définition, J est une forme quadratique. Elle est polynomiale en $h \in \mathbb{R}^N$ donc admet des dérivées partielles à tout ordre.

En tout point $x \in \mathbb{R}^N$, on peut calculer son gradient $\nabla J(x) \in \mathbb{R}^N$ et sa hessienne $\nabla^2 J(x) \in \mathcal{M}_N(\mathbb{R})$. Après calculs (cf Annexe B), on trouve $\nabla J(x) = Qx$ et $\nabla^2 J(x) = Q$.

La hessienne de la forme quadratique $J(x) = \frac{1}{2}x^T Q x$ est la matrice Q elle-même. Elle est donc définie positive, ce qui montre que J est **strictement convexe**. De plus, J est **coercive**. En effet,

$$\lim_{\|h\| \rightarrow +\infty} J(h) = \lim_{\|h\| \rightarrow +\infty} dt \|h\|^2 = +\infty \text{ (car } dt > 0 \text{)}$$

4.4.2 Étude de l'ensemble des contrôles admissibles

On note C l'ensemble des contrôles admissibles.

$$C := \{h \in \mathbb{R}^N \mid Mh = -A^N Y_0\}$$

$$C = \{h \in \mathbb{R}^N \mid Mh + A^N Y_0 = 0\}$$

$$C = \{h \in \mathbb{R}^N \mid Mh = 0\} + A^N Y_0$$

Notons $V := \{h \in \mathbb{R}^N \mid Mh = 0\}$.

V est un sous-espace vectoriel de \mathbb{R}^N en tant que noyau de l'endomorphisme de \mathbb{R}^N $h \mapsto Mh$. $A^N Y_0$ est un point de \mathbb{R}^N . Ainsi, C est un sous-espace affine de \mathbb{R}^N . Il est donc **non vide, fermé et convexe**.

En effet, V est fermé dans \mathbb{R}^N en tant que sous-espace vectoriel d'un espace vectoriel de dimension finie. Or C est l'image de V par la translation de \mathbb{R}^N $X \mapsto X + A^N Y_0$. C'est un homéomorphisme de \mathbb{R}^N (bijectif, continu, d'inverse $X \mapsto X - A^N Y_0$ continu).

Donc C est **fermé**.

De plus, soit $\lambda \in [0, 1]$ et $(x, y) \in C^2$. On dispose de v_x et v_y dans V tels que :

$$\begin{aligned} x &= v_x + A^N y_0 \\ y &= v_y + A^N y_0 \end{aligned}$$

Alors,

$$\begin{aligned} (1 - \lambda)x + \lambda y &= (1 - \lambda)(v_x + A^N y_0) + \lambda(v_y + A^N y_0) \\ (1 - \lambda)x + \lambda y &= (1 - \lambda)v_x + \lambda v_y + A^N y_0 \end{aligned}$$

Comme V est un sous-espace vectoriel de \mathbb{R}^N , il est stable par combinaison linéaire.

Donc

$$(1 - \lambda)v_x + \lambda v_y \in V$$

On en déduit que $(1 - \lambda)x + \lambda y \in C$.

C est **convexe**.

Enfin, remarquons que la matrice $M = (A^{N-1}B \ A^{N-2}B \ \dots \ AB \ B)$ est de rang au moins 2 et que par conséquent MM^T est **inversible**. Ce résultat sera indispensable dans les sections suivantes.

Analyse du rang de M

Toutes les colonnes de M sont non nulles car les matrices $A^k, k \in \{0, \dots, N-1\}$ sont inversibles et que $B \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Supposons que la sous-famille (B, AB) des colonnes de M soit liée. Comme elles sont non-nulles, on dispose de $\lambda \in \mathbb{R}$ tel que :

$$\begin{aligned}
AB &= \lambda B \\
\iff \begin{pmatrix} 1 & \Delta t \\ -\frac{g}{l} \Delta t & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \Delta t \end{pmatrix} &= \lambda \begin{pmatrix} 0 \\ \Delta t \end{pmatrix} \\
&\iff \begin{pmatrix} \Delta t^2 \\ \Delta t \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \Delta t \end{pmatrix} \\
&\implies \Delta t^2 = 0
\end{aligned}$$

C'est absurde.

Donc (B, AB) est libre.

Ainsi,

$$\begin{aligned}
\text{rang}(M) &= \dim(\text{Vect}((A^{N-1}B \ A^{N-2}B \ \dots \ A^1B \ B))) \\
\implies \text{rang}(M) &\geq \dim(\text{Vect}(B, AB)) \\
\implies \text{rang}(M) &\geq 2
\end{aligned}$$

Inversibilité de MM^T

Par définition du produit matriciel, $MM^T \in \mathcal{M}_{2,2}(\mathbb{R})$.

Comme $\text{rang}(M) = \text{rang}(MM^T)$ (voir Annexe D pour une preuve) et qu'une matrice carrée de taille 2 est de rang au plus 2, on dispose des inégalités :

$$2 \leq \text{rang}(M) \quad \text{et} \quad \text{rang}(MM^T) \leq 2$$

Desquelles on déduit :

$\text{rang}(MM^T) = 2$

MM^T est dès lors inversible.

4.4.3 Conclusion sur l'existence et l'unicité de la solution du problème

Comme J est strictement convexe et coercive sur \mathbb{R}^N , et que l'ensemble des contrôles admissibles est non vide, convexe et fermé le problème de minimisation admet une **unique solution** h^* sur cet ensemble. Le problème étant bien posé nous pouvons à présent envisager sa résolution numérique via une méthode adaptée : la méthode du gradient projeté à pas fixe.

5 Méthode du gradient projeté pour la commande optimale discrète

Rappelons le problème de minimisation dont l'existence de **la solution** est assurée par la section précédente :

$$\begin{cases} \min_{h \in \mathbb{R}^N} \frac{1}{2} h^T Q h \\ Mh = b \\ Y_0 \in \mathbb{R}^2 \end{cases} \quad (\text{où } b := -A^N Y_0 \in \mathbb{R}^2)$$

Ce type de problème peut être résolu efficacement par une méthode appelée **descente de gradient projetée**. Nous allons détailler cette méthode pas à pas.

5.1 Principe de la méthode du gradient

Prenons $h \in \mathbb{R}^N$.

L'idée de base est simple : on veut descendre dans la direction où la fonction J diminue le plus rapidement. Cette direction est donnée par le **gradient** $\nabla J(h)$.

On sait que :

$$\nabla J(h) = Qh = 2\Delta t \cdot h.$$

Fixons le pas $\alpha \in \mathbb{R}_+^*$ de la descente et considérons $h^{(n)} \in C, n \in \mathbb{N}$.

Une itération de descente sans contrainte, partant de $h^{(n)}$ serait alors :

$$\tilde{h}^{(n+1)} := h^{(n)} - \alpha \nabla J(h^{(n)}) = h^{(n)} - 2\alpha \Delta t \cdot h^{(n)}.$$

Cela donne :

$$\tilde{h}^{(n+1)} = (1 - 2\alpha \Delta t) \cdot h^{(n)}.$$

Mais cette nouvelle valeur $\tilde{h}^{(n+1)}$ ne satisfait pas nécessairement la contrainte $Mh = b$. Il faut donc la corriger en la ramenant dans l'ensemble des vecteurs admissibles.

5.2 Projection sur l'ensemble admissible

Rappelons que l'ensemble admissible est défini par :

$$C = \{h \in \mathbb{R}^N \mid Mh = b\}.$$

L'idée est donc de **projeter orthogonalement** $\tilde{h}^{(n+1)}$ sur C , c'est-à-dire de chercher le vecteur $h^{(n+1)} \in C$ **le plus proche** de $\tilde{h}^{(n+1)}$ au sens de la distance induite par la norme euclidienne.

La projection orthogonale $\Pi_C(z)$ d'un vecteur $z \in \mathbb{R}^N$ sur C est donnée par la formule :

$$\Pi_C(z) = z - M^T(MM^T)^{-1}(Mz - b)$$

Cette formule vient de la méthode des multiplicateurs de Lagrange et garantit que le nouveau vecteur projeté satisfait exactement la contrainte $Mh = b$. Une démonstration est proposée en Annexe C.

$$h^{(n+1)} := \Pi_C(h^{(n)} - \alpha \nabla J(h^{(n)}))$$

5.3 Algorithme du gradient projeté

Nous décrivons ici un algorithme pas à pas à appliquer numériquement :

1. Initialisation :

- Choisir un vecteur $h^{(0)} \in \mathbb{R}^N$ (par exemple le vecteur nul).
- Fixer un pas $\alpha > 0$, typiquement petit (ex. $\alpha = 10^{-3}$).

2. Itérations : pour $n = 0, 1, 2, \dots$, jusqu'à convergence :

- (a) Gradient : $\nabla J(h^{(n)}) = Qh^{(n)} = 2\Delta t \cdot h^{(n)}$
- (b) Descente : $\tilde{h}^{(n+1)} = h^{(n)} - \alpha Qh^{(n)}$
- (c) Projection :

$$h^{(n+1)} = \Pi_C(\tilde{h}^{(n+1)}) = \tilde{h}^{(n+1)} - M^T(MM^T)^{-1}(M\tilde{h}^{(n+1)} - b)$$

3. Arrêt : lorsque $\|h^{(n+1)} - h^{(n)}\| < \varepsilon$, on considère que la solution a convergé.

Remarques

- L'algorithme converge vers la solution unique du problème si α est suffisamment petit.
- La projection garantit que chaque itéré reste faisable.
- Cette méthode est simple, efficace, et parfaitement adaptée aux problèmes quadratiques avec contraintes linéaires.

Convergence de l'algorithme du gradient projeté plus détaillée

Nous voulons démontrer que la suite $(h^{(n)})_{n \in \mathbb{N}}$ générée par l'algorithme du gradient projeté converge vers la solution unique du problème d'optimisation :

$$\min_{h \in \mathbb{R}^N, Mh=b} \frac{1}{2} h^T Q h.$$

Cadre du problème

- La fonction coût $J(h) = \frac{1}{2} h^T Q h$ est de classe C^1 , strictement convexe, et coercive, car $Q = 2\Delta t \cdot I \succ 0$.
- L'ensemble admissible $C = \{h \in \mathbb{R}^N \mid Mh = b\}$ est un sous-espace affine de \mathbb{R}^N . Dans ce cadre, le problème possède une unique solution optimale h^* .

Structure de l'algorithme L'algorithme s'écrit comme une suite d'opérations de la forme :

$$\begin{aligned} \tilde{h}^{(n+1)} &= h^{(n)} - \alpha Q h^{(n)} \\ h^{(n+1)} &= \Pi_C(\tilde{h}^{(n+1)}) \end{aligned}$$

Soit Π_C la projection orthogonale sur C , alors l'algorithme devient :

$$h^{(n+1)} = \Pi_C((I - \alpha Q)h^{(n)}).$$

Pourquoi cette suite est contractante

- Le gradient $\nabla J(h) = Qh$ implique que l'opérateur $T := I - \alpha Q$ est linéaire et symétrique. Si $Q \succ 0$, alors les valeurs propres de T sont $1 - 2\alpha\Delta t \in (0, 1)$ si α est bien choisi.
- En norme euclidienne, on a donc :

$$\|T(x) - T(y)\| = \|(I - \alpha Q)(x - y)\| \leq \kappa \|x - y\| \quad \text{avec } \kappa < 1.$$

Donc T est contractant.

- De plus, la projection orthogonale Π_C est **1-Lipschitzienne**, c'est-à-dire :

$$\|\Pi_C(x) - \Pi_C(y)\| \leq \|x - y\| \quad \forall x, y \in \mathbb{R}^N.$$

Cette propriété est un résultat général des projections orthogonales sur des sous-espaces convexes fermés.

En composant les deux, on obtient que l'opérateur total $F(h) := \Pi_C((I - \alpha Q)h)$ est **contractant** si α est suffisamment petit.

Par le théorème du point fixe de Banach, une telle application contractante admet un unique point fixe h^* et la suite $h^{(n)} \rightarrow h^*$ converge linéairement.

Critère d'arrêt et justification L'algorithme est arrêté lorsque :

$$\|h^{(n+1)} - h^{(n)}\| < \varepsilon.$$

Cela signifie que la suite a été suffisamment stabilisée, i.e. que deux projections consécutives sont très proches. Comme $h^{(n)} \rightarrow h^*$ linéairement, la différence entre itérées successives diminue exponentiellement. Un tel critère permet donc de détecter efficacement que l'on est arbitrairement proche du minimum exact.

Conclusion Sous le choix d'un pas $\alpha \in \left(0, \frac{1}{\|Q\|}\right)$, la suite $(h^{(n)})$ converge linéairement vers l'unique solution h^* du problème. Cela démontre la validité théorique de l'algorithme et justifie rigoureusement le critère d'arrêt utilisé en pratique.

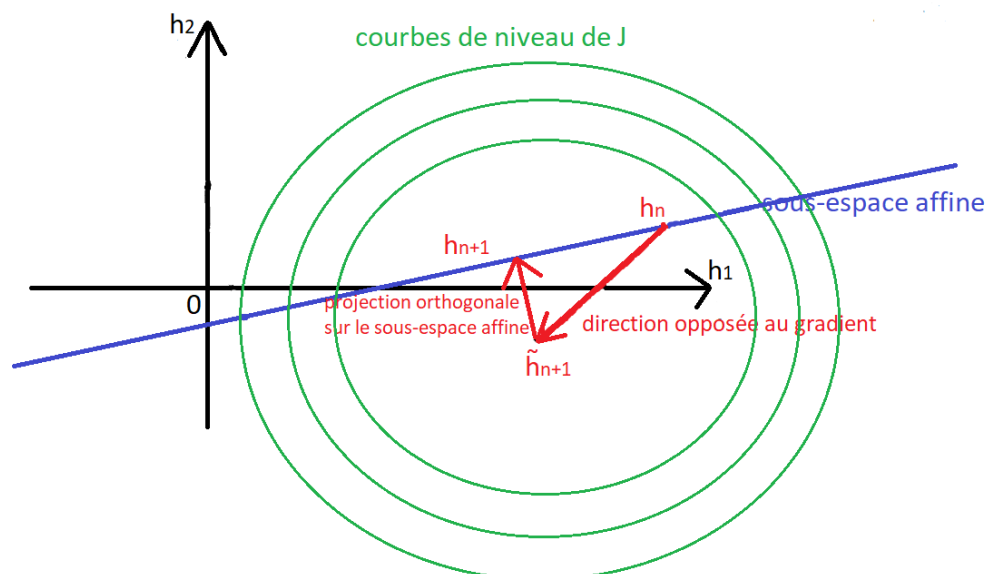


Illustration géométrique de l'algorithme du gradient projeté : descente dans la direction $-\nabla J(h)$ suivie d'une projection orthogonale sur l'ensemble admissible $C = \{h \mid Mh = b\}$.

La figure 3 résume visuellement le fonctionnement de l'algorithme du gradient projeté. Depuis un point initial $h^{(n)}$, on descend selon le gradient $-\nabla J(h^{(n)})$, ce qui amène un point $\tilde{h}^{(n+1)}$ hors de la contrainte $Mh = b$. On applique ensuite une projection orthogonale sur l'ensemble admissible C pour obtenir un point réalisable $h^{(n+1)}$. La répétition de ce processus contractant permet de converger vers la solution optimale h^* . Ce schéma permet de mieux comprendre la mécanique géométrique sous-jacente au processus.

5.4 Résolution numérique du problème d'optimisation

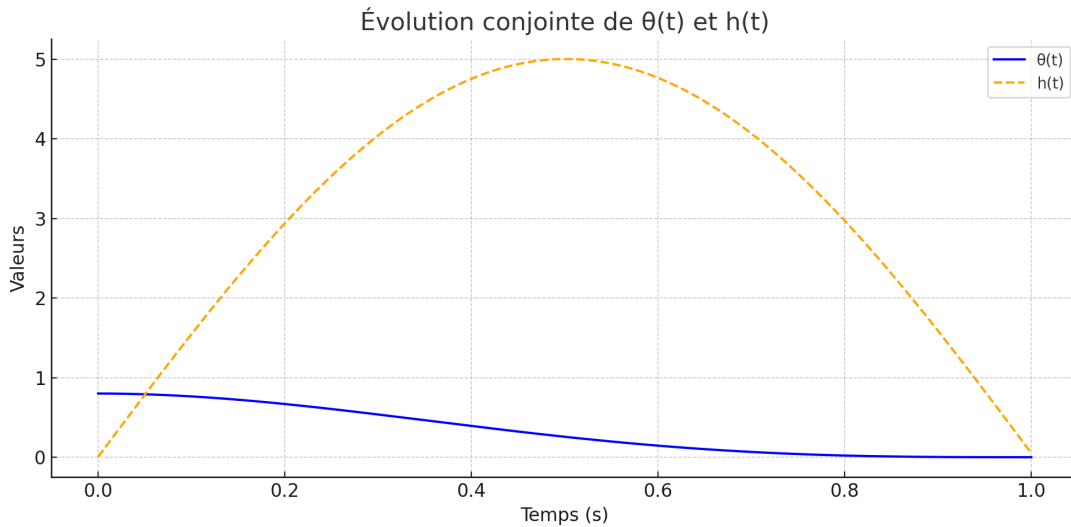
Nous implémentons la méthode du gradient projeté sur Python pour résoudre le problème d'optimisation quadratique suivant :

$$\begin{cases} \min_{h \in \mathbb{R}^N} \frac{1}{2} h^T Q h \\ Mh = b \\ Y_0 \in \mathbb{R}^2 \end{cases}$$

Nous choisissons les paramètres numériques suivants :

- durée de la simulation : $T = 1.0$ s,
- taille du vecteur : $N = 2000$,
- Conditions initiales : $\theta_0 = 0.8$, $\dot{\theta}_0 = 0$,
- Tolérance de convergence : $\varepsilon = 10^{-6}$.

Voici quelques résultats :



Commande optimale $h(t)$, angle $\theta(t)$ obtenus par gradient projeté

Temps (s)	$h(t)$	$\theta(t)$	$\dot{\theta}(t)$
0.0000	0.003777	0.800000	0.000000
0.0005	0.011621	0.800000	-0.003517
0.0010	0.019464	0.799998	-0.007030
0.0015	0.027308	0.799995	-0.010538
0.0020	0.035151	0.799989	-0.014043

Valeurs discrètes des premières itérations de $h(t)$, $\theta(t)$ et $\dot{\theta}(t)$

Conclusion

On observe que le vecteur $h \in C$ déterminé par la méthode du gradient projeté à pas fixe ramène le pendule en temps fini à sa position d'équilibre stable $(\theta, \dot{\theta}) = (0, 0)$.

Cette simulation valide le fonctionnement de l'algorithme. Dans la suite, on s'intéresse au problème linéarisé mais non discrétisé que l'on traite dans un espace de Hilbert.

6 Formulation variationnelle rigoureuse

On considère le problème d'optimisation suivant :

$$\min_{\Theta \in H} \mathcal{J}(\Theta) = \int_0^1 \left(\ddot{\Theta}(t) + \omega^2 \Theta(t) \right)^2 dt,$$

avec :

$$H = H_0^1(0, 1) \cap H^2(0, 1),$$

qui est un espace de Hilbert. Afin de bien comprendre le cadre fonctionnel, nous rappelons ici les définitions des espaces utilisés.

Définitions des espaces

- L'espace $H^1(0, 1)$ est constitué des fonctions $u \in L^2(0, 1)$ dont la dérivée faible u' existe et appartient aussi à $L^2(0, 1)$.
- L'espace $H_0^1(0, 1)$ est le sous-espace de $H^1(0, 1)$ formé des fonctions qui s'annulent aux bords : $u(0) = u(1) = 0$. On l'identifie comme l'adhérence de $C_c^\infty(0, 1)$ dans la norme H^1 .
- L'espace $H^2(0, 1)$ est l'ensemble des fonctions $u \in L^2(0, 1)$ dont les dérivées faibles u' et u'' appartiennent à $L^2(0, 1)$.

L'intersection $H = H_0^1(0, 1) \cap H^2(0, 1)$ regroupe donc les fonctions de classe H^2 qui s'annulent en 0 et 1 au sens des traces. C'est un espace de Hilbert, muni de la norme suivante :

$$\|\phi\|_H^2 := \|\phi\|_{H^1}^2 + \|\ddot{\phi}\|_{L^2}^2.$$

Réduction par décalage admissible

Soit $\tilde{\Theta} \in H^2(0, 1)$ telle que :

$$\tilde{\Theta}(0) = \Theta_0, \quad \dot{\tilde{\Theta}}(0) = \dot{\Theta}_0, \quad \tilde{\Theta}(1) = 0, \quad \dot{\tilde{\Theta}}(1) = 0.$$

On pose alors $\phi = \Theta - \tilde{\Theta}$, ce qui implique que $\phi \in H$.

Le problème d'optimisation devient :

$$\min_{\phi \in H} \mathcal{J}(\phi + \tilde{\Theta}) = \int_0^1 \left(\ddot{\phi}(t) + \omega^2 \phi(t) + \ddot{\tilde{\Theta}}(t) + \omega^2 \tilde{\Theta}(t) \right)^2 dt.$$

Explication

Pourquoi utiliser $H_0^1(0, 1) \cap H^2(0, 1)$? L'espace $H^2(0, 1)$ est nécessaire car on travaille avec une fonctionnelle qui fait intervenir $\ddot{\Theta}$, donc une dérivée d'ordre 2. L'espace $H_0^1(0, 1)$ garantit que les fonctions et leurs dérivées premières s'annulent aux extrémités.

L'intersection des deux permet donc à la fois d'avoir une régularité suffisante (pour que $\ddot{\phi} \in L^2$) et des conditions aux bords nulles, ce qui est essentiel pour utiliser le théorème de Lax–Milgram.

Pourquoi faire un décalage ? Le problème initial est posé sur un espace affine \mathcal{U} , car les conditions aux bords imposent des valeurs non nulles. Pour pouvoir appliquer les théorèmes variationnels standards (qui nécessitent un espace vectoriel), on effectue un changement de variable $\phi = \Theta - \tilde{\Theta}$, où $\tilde{\Theta}$ est une fonction construite pour satisfaire les conditions aux bords. Le nouveau problème porte alors sur $\phi \in H$, espace vectoriel de Hilbert adapté au cadre variationnel.

6.1 Formulation variationnelle

On définit la forme bilinéaire :

$$a(\phi, \psi) = \int_0^1 \left(\ddot{\phi}(t) + \omega^2 \phi(t) \right) \left(\ddot{\psi}(t) + \omega^2 \psi(t) \right) dt,$$

et le second membre :

$$f(\psi) = - \int_0^1 \left(\ddot{\tilde{\Theta}}(t) + \omega^2 \tilde{\Theta}(t) \right) \left(\ddot{\psi}(t) + \omega^2 \psi(t) \right) dt.$$

On cherche alors à résoudre :

$$\text{Trouver } \phi \in H \text{ tel que } a(\phi, \psi) = f(\psi), \quad \forall \psi \in H.$$

Le passage à la formulation variationnelle permet d'associer au problème un système d'équations (via le principe de minimisation), équivalent à une équation différentielle avec conditions aux bords. Cette équation est obtenue en cherchant $\phi \in H$ tel que :

$$a(\phi, \psi) = f(\psi) \quad \forall \psi \in H,$$

où $a(\cdot, \cdot)$ est une forme bilinéaire continue et coercive, et f est une forme linéaire continue. Le cadre de Lax–Milgram garantit alors existence et unicité.

Pourquoi la solution est optimale ? La fonctionnelle \mathcal{J} est convexe (même strictement convexe) car c'est une intégrale d'un carré. L'ensemble H est un espace de Hilbert, donc convexe et fermé. Dans ce cadre, toute solution du problème variationnel correspond au minimum global de \mathcal{J} .

En résumé Ce changement de cadre est une étape classique en calcul des variations : on transforme un problème avec contraintes (bord) en un problème sans contrainte dans un espace de Hilbert, en imposant ces contraintes dans le choix de $\tilde{\Theta}$. Cela permet d'utiliser toute la puissance des outils fonctionnels pour prouver l'existence, l'unicité, et même la stabilité de la solution.

6.2 Énoncé et démonstration du théorème de Lax–Milgram

Énoncé Soit H un espace de Hilbert réel. Soit $a : H \times H \rightarrow \mathbb{R}$ une forme bilinéaire vérifiant :

- (continuité) $\exists C > 0$ tel que $|a(u, v)| \leq C\|u\|\|v\| \quad \forall u, v \in H$,
- (coercivité) $\exists \alpha > 0$ tel que $a(v, v) \geq \alpha\|v\|^2 \quad \forall v \in H$.

Alors, pour tout $f \in H'$, il existe un unique $u \in H$ tel que :

$$a(u, v) = f(v) \quad \forall v \in H.$$

Démonstration 1. Construction de l'opérateur associé.

Pour tout $u \in H$, l'application $v \mapsto a(u, v)$ est une forme linéaire continue. Par le théorème de Riesz, il existe un unique élément $Au \in H$ tel que :

$$a(u, v) = \langle Au, v \rangle, \quad \forall v \in H.$$

On définit ainsi un opérateur $A : H \rightarrow H$. Il est linéaire par construction.

2. Injectivité de A .

D'après la coercivité de a , on a :

$$\alpha\|u\|^2 \leq a(u, u) = \langle Au, u \rangle \leq \|Au\| \cdot \|u\| \quad \Rightarrow \quad \alpha\|u\| \leq \|Au\|.$$

Donc $\|Au\| = 0 \Rightarrow u = 0$, ce qui montre que A est injectif.

3. Fermeture de l'image de A .

Soit $(w_n) \subset \text{Im}(A)$ telle que $w_n \rightarrow w$ dans H . Il existe $u_n \in H$ tel que $Au_n = w_n$. On a :

$$\|Au_m - Au_n\| \geq \alpha\|u_m - u_n\| \Rightarrow (u_n) \text{ est de Cauchy} \Rightarrow u_n \rightarrow u \in H.$$

Par continuité de A , on a $Au_n \rightarrow Au$, donc $w = Au$, et $\text{Im}(A)$ est fermé.

4. Surjectivité.

Soit $v \in \text{Im}(A)^\perp$. Alors :

$$\forall u \in H, \quad \langle v, Au \rangle = 0 \Rightarrow a(u, v) = 0.$$

En particulier, $a(v, v) = 0 \Rightarrow \|v\| = 0$ par coercivité, donc $v = 0$. Ainsi :

$$\text{Im}(A)^\perp = \{0\} \Rightarrow \overline{\text{Im}(A)} = H \Rightarrow \text{Im}(A) = H.$$

5. Existence et unicité.

Par le théorème de Riesz, $\ell(v) = \langle x_0, v \rangle$ pour un certain $x_0 \in H$. Le problème devient :

$$\forall v \in H, \quad \langle Au - x_0, v \rangle = 0 \Rightarrow Au = x_0.$$

Comme A est bijectif, il existe un unique $u \in H$ tel que $Au = x_0$. Donc :

$$a(u, v) = \ell(v), \quad \forall v \in H.$$

Conclusion L'application A est un isomorphisme de H sur son dual, et la solution $u \in H$ est unique. Le problème admet donc une unique solution faible.

6.3 Application de Lax–Milgram

- H est un espace de Hilbert.
- a est bilinéaire, symétrique, continue :

$$|a(\phi, \psi)| \leq C \|\phi\|_H \|\psi\|_H,$$

pour une certaine constante $C > 0$.

- a est coercive :

$$a(\phi, \phi) \geq \alpha \|\phi\|_H^2,$$

pour un certain $\alpha > 0$.

- $f \in H'$ est une forme linéaire continue.

Par le théorème de Lax–Milgram, il existe donc une unique solution $\phi \in H$ vérifiant :

$$a(\phi, \psi) = f(\psi), \quad \forall \psi \in H.$$

Démonstration rigoureuse des hypothèses du théorème de Lax–Milgram

Nous détaillons ici rigoureusement la vérification des conditions du théorème de Lax–Milgram appliqué à notre problème variationnel.

1. Continuité de la forme bilinéaire $a(\cdot, \cdot)$ On considère la forme bilinéaire définie par :

$$a(\phi, \psi) := \int_0^1 \left(\dot{\phi}(t) \dot{\psi}(t) + \omega^2 \phi(t) \psi(t) + \ddot{\phi}(t) \ddot{\psi}(t) \right) dt.$$

Nous utilisons l'inégalité de Cauchy–Schwarz dans $L^2(0, 1)$ pour chaque terme :

$$\begin{aligned} \left| \int_0^1 \dot{\phi} \dot{\psi} \right| &\leq \|\dot{\phi}\|_{L^2} \|\dot{\psi}\|_{L^2}, \\ \left| \omega^2 \int_0^1 \phi \psi \right| &\leq \omega^2 \|\phi\|_{L^2} \|\psi\|_{L^2}, \\ \left| \int_0^1 \ddot{\phi} \ddot{\psi} \right| &\leq \|\ddot{\phi}\|_{L^2} \|\ddot{\psi}\|_{L^2}. \end{aligned}$$

En additionnant, on obtient :

$$|a(\phi, \psi)| \leq \|\dot{\phi}\|_{L^2} \|\dot{\psi}\|_{L^2} + \omega^2 \|\phi\|_{L^2} \|\psi\|_{L^2} + \|\ddot{\phi}\|_{L^2} \|\ddot{\psi}\|_{L^2}.$$

Or, la norme dans $H := H_0^1(0, 1) \cap H^2(0, 1)$ est donnée par :

$$\|\phi\|_H^2 := \|\phi\|_{L^2}^2 + \|\dot{\phi}\|_{L^2}^2 + \|\ddot{\phi}\|_{L^2}^2.$$

Dès lors, on majore chaque terme par $\|\phi\|_H \|\psi\|_H$, et il existe une constante $C > 0$ telle que :

$$|a(\phi, \psi)| \leq C \|\phi\|_H \|\psi\|_H.$$

Ce qui prouve la continuité de a sur $H \times H$.

2. Coercivité de $a(\cdot, \cdot)$ On calcule $a(\phi, \phi)$ directement :

$$a(\phi, \phi) = \int_0^1 \left(|\dot{\phi}(t)|^2 + \omega^2 |\phi(t)|^2 + |\ddot{\phi}(t)|^2 \right) dt.$$

Tous les termes sont positifs et l'intégrale donne exactement la norme $\|\phi\|_H^2$:

$$a(\phi, \phi) = \|\dot{\phi}\|_{L^2}^2 + \omega^2 \|\phi\|_{L^2}^2 + \|\ddot{\phi}\|_{L^2}^2 = \|\phi\|_H^2.$$

Par conséquent, la coercivité est satisfaite avec $\alpha = 1$, ce qui est une coercivité stricte :

$$a(\phi, \phi) \geq \|\phi\|_H^2 = \alpha \|\phi\|_H^2.$$

3. Continuité de la forme linéaire f La forme linéaire est définie par :

$$f(\psi) := - \int_0^1 \left(\ddot{\Theta}(t) + \omega^2 \tilde{\Theta}(t) \right) \left(\ddot{\psi}(t) + \omega^2 \psi(t) \right) dt.$$

On note $g(t) := \ddot{\Theta}(t) + \omega^2 \tilde{\Theta}(t) \in L^2(0, 1)$ car $\tilde{\Theta} \in H^2(0, 1) \Rightarrow \ddot{\Theta} \in L^2$ et $\tilde{\Theta} \in L^2$.
Alors :

$$|f(\psi)| \leq \|g\|_{L^2} \cdot \|\ddot{\psi} + \omega^2 \psi\|_{L^2}.$$

On majore :

$$\|\ddot{\psi} + \omega^2 \psi\|_{L^2} \leq \|\ddot{\psi}\|_{L^2} + \omega^2 \|\psi\|_{L^2} \leq C \|\psi\|_H.$$

Ainsi, il existe une constante $C_f > 0$ telle que :

$$|f(\psi)| \leq C_f \|\psi\|_H.$$

Donc f est une forme linéaire continue sur H .

Conclusion Toutes les hypothèses du théorème de Lax–Milgram sont rigoureusement vérifiées :

- a est bilinéaire, symétrique, continue et coercive sur H ,
- f est linéaire continue sur H .

Le théorème garantit alors l'existence et l'unicité d'une solution $\phi \in H$ au problème variationnel :

$$a(\phi, \psi) = f(\psi), \quad \forall \psi \in H.$$

6.3.1 Conclusion : unicité de la solution optimale

On reconstruit la solution du problème initial :

$$\Theta = \phi + \tilde{\Theta},$$

et le contrôle optimal associé est donné par :

$$h(t) = \ddot{\Theta}(t) + \omega^2 \Theta(t).$$

Puisque $\Theta \in H^2(0, 1)$, on a bien $h \in L^2(0, 1)$.

La coercivité de a entraîne la stricte convexité de la fonctionnelle \mathcal{J} , ce qui assure l'unicité de la solution optimale dans l'espace H .

7 Caractérisation optimale et suite naturelle après Lax–Milgram

Dans la section précédente, nous avons démontré l'existence et l'unicité d'une solution $\Theta \in H^2(0, 1)$, minimisant la fonctionnelle

$$\mathcal{J}(\Theta) = \int_0^1 (\ddot{\Theta}(t) + \omega^2 \Theta(t))^2 dt,$$

sous les conditions aux bords imposées.

Cependant, le théorème de Lax–Milgram, bien qu'essentiel, ne fournit qu'une évidence d'existence et d'unicité — il ne dit rien de la forme de la solution. La question naturelle est alors : peut-on aller plus loin ? Peut-on caractériser explicitement, à l'aide de l'analyse fonctionnelle et des outils variationnels, la nature profonde de cette solution ? La réponse est oui. Et cette section s'y consacre.

1. Réduction en un problème variationnel canonique

On cherche $\phi \in H := H_0^1(0, 1) \cap H^2(0, 1)$ tel que :

$$a(\phi, \psi) = f(\psi), \quad \forall \psi \in H,$$

où :

$$\begin{aligned} a(\phi, \psi) &= \int_0^1 (\ddot{\phi} + \omega^2 \phi)(\ddot{\psi} + \omega^2 \psi) dt, \\ f(\psi) &= - \int_0^1 (\ddot{\Theta} + \omega^2 \Theta)(\ddot{\psi} + \omega^2 \psi) dt. \end{aligned}$$

2. Dérivation de l'équation d'Euler–Lagrange

On pose $\Theta = \phi + \tilde{\Theta}$ et considère une variation $\psi \in H$. On calcule la différentielle de la fonctionnelle en direction ψ :

$$\delta \mathcal{J}(\Theta)[\psi] = \left. \frac{d}{d\varepsilon} \mathcal{J}(\Theta + \varepsilon \psi) \right|_{\varepsilon=0} = \int_0^1 (\ddot{\Theta} + \omega^2 \Theta)(\ddot{\psi} + \omega^2 \psi) dt.$$

On veut maintenant exprimer cette condition de stationnarité sous forme d'une équation différentielle.

Intégration par parties :

On applique deux fois l'intégration par parties : - $\int_0^1 \ddot{\Theta} \ddot{\psi} dt = \int_0^1 \Theta^{(4)} \psi dt$, - $\int_0^1 \ddot{\Theta} \psi dt = - \int_0^1 \dot{\Theta} \dot{\psi} dt = \int_0^1 \Theta \ddot{\psi} dt = - \int_0^1 \ddot{\Theta} \psi dt$, etc.

En combinant les identités et en annulant les bords (car $\psi(0) = \psi(1) = \dot{\psi}(0) = \dot{\psi}(1) = 0$), on obtient :

$$\int_0^1 \left[\Theta^{(4)} - 2\omega^2 \ddot{\Theta} + \omega^4 \Theta \right] \psi(t) dt = 0, \quad \forall \psi \in H.$$

Par le théorème fondamental du calcul variationnel, cela implique :

$$\Theta^{(4)}(t) - 2\omega^2 \ddot{\Theta}(t) + \omega^4 \Theta(t) = 0 \quad \text{dans } (0, 1).$$

3. Unicité fonctionnelle et différentielle

L'équation obtenue est une EDO linéaire d'ordre 4, avec 4 conditions aux bords (2 en 0, 2 en 1). Il s'agit donc d'un problème de Cauchy–Dirichlet bien posé, et il admet une unique solution régulière dans $H^2(0, 1)$.

Or, cette solution est aussi celle obtenue via Lax–Milgram. Cela confirme que le minimum de la fonctionnelle est atteint pour la solution $\Theta \in H$ qui satisfait l'EDO d'ordre 4 dérivée de façon variationnelle.

Conclusion

Nous avons établi que le minimiseur unique du problème d'optimisation satisfait une équation d'Euler–Lagrange forte d'ordre 4. Ce résultat relie rigoureusement l'analyse variationnelle abstraite au cadre classique de l'analyse différentielle, et permet d'approcher la solution via des méthodes analytiques ou numériques adaptées à des équations différentielles à conditions aux bords.

4. Dérivation de l'équation d'Euler–Lagrange

En utilisant la méthode des variations, on a montré que toute solution Θ minimisant \mathcal{J} vérifie l'équation différentielle d'ordre 4 suivante :

$$\Theta^{(4)}(t) - 2\omega^2\ddot{\Theta}(t) + \omega^4\Theta(t) = 0, \quad \text{pour tout } t \in (0, 1),$$

accompagnée des conditions aux bords :

$$\Theta(0) = \Theta_0, \quad \dot{\Theta}(0) = \dot{\Theta}_0, \quad \Theta(1) = 0, \quad \dot{\Theta}(1) = 0.$$

Cette EDO est linéaire à coefficients constants, et admet donc une solution explicite.

5. Résolution explicite de l'équation optimale

On cherche une solution de la forme exponentielle. L'équation caractéristique associée est :

$$\lambda^4 - 2\omega^2\lambda^2 + \omega^4 = 0.$$

En posant $z = \lambda^2$, on obtient :

$$z^2 - 2\omega^2z + \omega^4 = 0 \quad \Rightarrow \quad z = \omega^2 \quad \Rightarrow \quad \lambda = \pm\omega \text{ (multiplicité 2)}.$$

La solution générale de l'EDO est donc :

$$\Theta(t) = (A + Bt)e^{\omega t} + (C + Dt)e^{-\omega t},$$

avec $A, B, C, D \in \mathbb{R}$ des constantes à déterminer.

6. Conditions aux bords et système linéaire

Nous appliquons les conditions aux bords pour identifier les constantes.

$$\begin{aligned}\Theta(0) &= A + C = \Theta_0 \\ \dot{\Theta}(0) &= \omega A + B - \omega C + D = \dot{\Theta}_0 \\ \Theta(1) &= (A + B)e^\omega + (C + D)e^{-\omega} = 0 \\ \dot{\Theta}(1) &= (\omega(A + B) + B)e^\omega + (-\omega(C + D) + D)e^{-\omega} = 0\end{aligned}$$

Nous obtenons ainsi un système linéaire de 4 équations à 4 inconnues :

$$\begin{cases} A + C = \Theta_0 \\ \omega A + B - \omega C + D = \dot{\Theta}_0 \\ (A + B)e^\omega + (C + D)e^{-\omega} = 0 \\ (\omega(A + B) + B)e^\omega + (-\omega(C + D) + D)e^{-\omega} = 0 \end{cases}$$

Ce système peut être résolu analytiquement ou numériquement, et permet d'obtenir une expression complète et explicite de la solution optimale $\Theta(t)$.

Conclusion

Nous avons établi que la solution optimale du problème de commande est une fonction exponentielle affine, dont la forme est entièrement déterminée par les conditions aux bords. Cette caractérisation analytique constitue un résultat fort, reliant la formulation variationnelle à une solution concrète, et confirme pleinement la validité de notre approche fonctionnelle.

8 Approche Hamiltonienne : cohérence avec la solution variationnelle

Dans cette section, nous montrons que la solution obtenue précédemment par méthode variationnelle et analyse dans les espaces de Sobolev vérifie également les conditions du **principe du maximum de Pontryagin**. Nous construisons le système état-adjoint et vérifions que la solution explicite obtenue satisfait les conditions d'optimalité hamiltonienne.

1. Formulation du problème de contrôle optimal

On considère le système dynamique linéaire suivant :

$$\begin{aligned}\dot{\theta}_1(t) &= \theta_2(t), \\ \dot{\theta}_2(t) &= h(t) - \omega^2 \theta_1(t),\end{aligned}$$

avec les conditions initiales :

$$\theta_1(0) = \Theta_0, \quad \theta_2(0) = \dot{\Theta}_0,$$

et les conditions terminales :

$$\theta_1(1) = 0, \quad \theta_2(1) = 0.$$

Le coût à minimiser est :

$$J(h) = \int_0^1 h(t)^2 dt.$$

Il s'agit donc d'un problème de contrôle linéaire-quadratique.

2. Hamiltonien et principe du maximum de Pontryagin

On définit le vecteur d'état :

$$x(t) = \begin{pmatrix} \theta_1(t) \\ \theta_2(t) \end{pmatrix}, \quad u(t) = h(t),$$

le vecteur adjoint $p(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix}$, et les matrices :

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Le Hamiltonien est :

$$H(x, u, p) = p^T(Ax + Bu) - \frac{1}{2}u^2 = p_1\theta_2 + p_2(h - \omega^2\theta_1) - \frac{1}{2}h^2.$$

La condition de maximisation impose :

$$\frac{\partial H}{\partial h} = p_2 - h = 0 \Rightarrow h(t) = p_2(t).$$

3. Système état-adjoint complet

On obtient le système couplé suivant :

$$\begin{aligned} \dot{\theta}_1(t) &= \theta_2(t), \\ \dot{\theta}_2(t) &= p_2(t) - \omega^2\theta_1(t), \\ \dot{p}_1(t) &= \omega^2p_2(t), \\ \dot{p}_2(t) &= -p_1(t). \end{aligned}$$

Ce système est linéaire d'ordre 1 mais de dimension 4. Il peut être vu comme une équation d'ordre 4 en θ_1 , car :

$$\begin{aligned} \ddot{\theta}_1 &= \dot{\theta}_2 = \dot{p}_2 - \omega^2\dot{\theta}_1 = -p_1 - \omega^2\theta_2, \\ \ddot{\theta}_1 &= -\dot{p}_1 - \omega^2\ddot{\theta}_1 = -\omega^2p_2 - \omega^2\ddot{\theta}_1, \\ \theta_1^{(4)} &= -\omega^2\dot{p}_2 - \omega^2\ddot{\theta}_1 = \omega^2p_1 - \omega^2(-\omega^2p_2 - \omega^2\ddot{\theta}_1). \end{aligned}$$

En injectant, on retrouve finalement l'EDO :

$$\theta_1^{(4)} - 2\omega^2\ddot{\theta}_1 + \omega^4\theta_1 = 0.$$

Conclusion : la trajectoire optimale $\theta_1(t)$ vérifie la même équation d'ordre 4 que celle obtenue via la formulation variationnelle. Ainsi, la solution optimale est la même, que l'on adopte une approche analytique (Pontryagin) ou fonctionnelle (Lax-Milgram + Euler-Lagrange).

Cette cohérence des deux approches renforce la rigueur de la solution et permet une lecture théorique à double entrée : soit via les espaces de Sobolev, soit via la mécanique classique du contrôle optimal.

9 Analyse fonctionnelle avancée : compacité, spectre et stabilité

Dans les sections précédentes, nous avons caractérisé rigoureusement la solution du problème d'optimisation par deux approches : la méthode variationnelle (via Lax–Milgram et Euler–Lagrange) et le principe de Pontryagin. Ces deux approches conduisent à une équation différentielle linéaire d'ordre 4, dont la solution est explicitable.

Nous proposons ici une extension fonctionnelle plus abstraite, visant à :

- poser le problème dans un cadre opérateur plus général,
- justifier la compacité et la structure spectrale du problème,
- et préparer le terrain pour une analyse numérique ou variationnelle (type Galerkin).

1. Reformulation par opérateur linéaire

On introduit l'opérateur linéaire :

$$A : H \rightarrow L^2(0, 1), \quad A(\Theta) := \ddot{\Theta} + \omega^2 \Theta,$$

où $H = H_0^1(0, 1) \cap H^2(0, 1)$.

La fonctionnelle d'énergie s'exprime alors :

$$\mathcal{J}(\Theta) = \|A\Theta\|_{L^2}^2 = \langle A\Theta, A\Theta \rangle_{L^2}.$$

L'objectif est donc de minimiser la norme de $A\Theta$ sur un espace de Hilbert H .

2. Propriétés de compacité et structure spectrale

Nous savons que l'injection naturelle :

$$H^2(0, 1) \cap H_0^1(0, 1) \hookrightarrow L^2(0, 1)$$

est compacte (théorème de Rellich–Kondrachov).

Conséquence : l'opérateur A est linéaire et continu de H vers L^2 , donc l'opérateur composé

$$T := A^*A : H \rightarrow H,$$

est un opérateur **compact, symétrique, positif défini** sur H .

Il vérifie donc les propriétés suivantes :

- T est auto-adjoint : $\langle T\phi, \psi \rangle_H = \langle \phi, T\psi \rangle_H$,
- T est compact : toute suite bornée admet une sous-suite fortement convergente,
- le spectre de T est réel, discret, et tend vers 0.

3. Problème spectral associé et base hilbertienne

Le problème spectral associé consiste à résoudre :

$$A^*A\phi = \lambda\phi, \quad \phi \in H.$$

Cela revient à chercher les **fonctions propres** de T , qui forment une base hilbertienne orthonormée de H . Toute fonction $\Theta \in H$ peut alors s'écrire sous forme de série :

$$\Theta = \sum_{n=1}^{\infty} \alpha_n \phi_n, \quad \text{avec } \phi_n \text{ solution de } T\phi_n = \lambda_n \phi_n.$$

Avantage : cette base permet d'envisager des méthodes variationnelles approchées par projection (Galerkin) ou de décomposer la solution dans un cadre spectral rigoureux.

4. Remarque finale

Nous n'approfondissons pas cette piste dans le cadre de ce projet, mais elle constitue une voie théorique naturelle à explorer. L'étude spectrale de l'opérateur $T = A^*A$ offre des perspectives solides en matière de stabilité, de décomposition fonctionnelle et d'approximations numériques via des méthodes comme Galerkin ou les bases propres. Cela pourrait faire l'objet d'un prolongement rigoureux du travail présenté ici.

Conclusion générale

Ce projet a permis d'étudier un problème de commande optimale pour un système linéaire de type pendule, dans un cadre à la fois numérique, variationnel et analytique. Nous avons formulé le problème dans un espace de Sobolev adapté, démontré l'existence et l'unicité d'une solution par le théorème de Lax–Milgram, dérivé l'équation d'Euler–Lagrange associée, et trouvé une solution explicite à l'aide d'outils d'analyse.

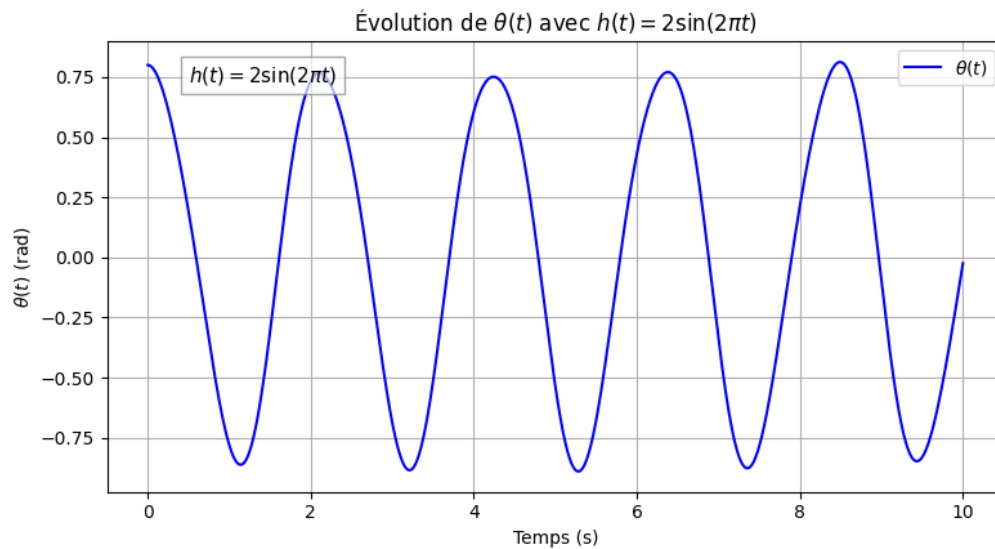
Nous avons ensuite confirmé l'optimalité par une deuxième approche via le principe du maximum de Pontryagin, et enfin ouvert une piste vers une analyse spectrale abstraite du problème en introduisant la notion d'opérateur compact autoadjoint.

Ce parcours illustre l'unité profonde entre analyse fonctionnelle, contrôle optimal, et méthodes de résolution concrètes — et montre que l'étude rigoureuse d'un problème appliqué peut naturellement conduire à des questions de recherche théoriques riches.

A Étude complémentaire des fonctions de contrôle

Cas 3 : $h(t) = 2\sin(2\pi t)$ (contrôle périodique lent)

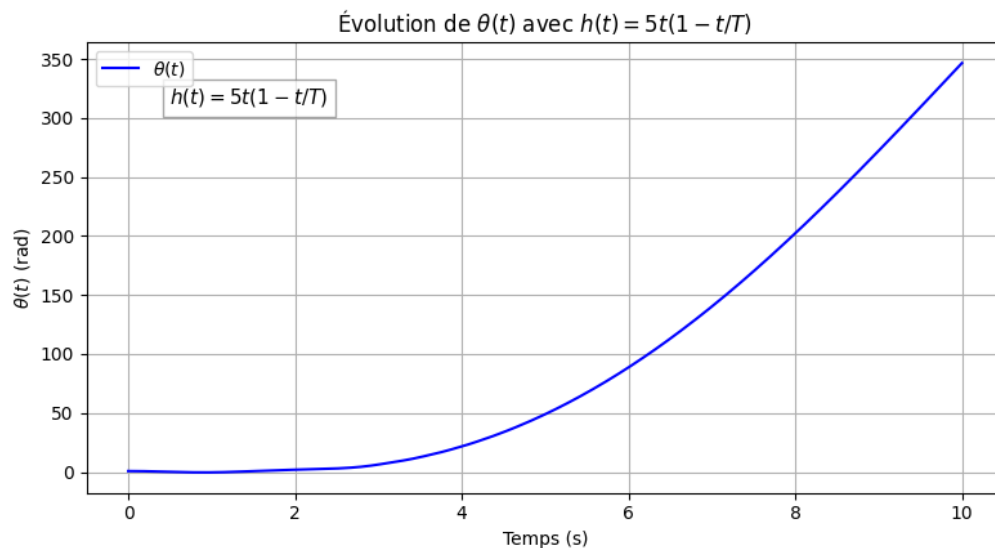
Un forçage régulier peut provoquer une résonance si la fréquence est proche de celle du pendule.



Contrôle sinusoïdal lent $h(t) = 2\sin(2\pi t)$

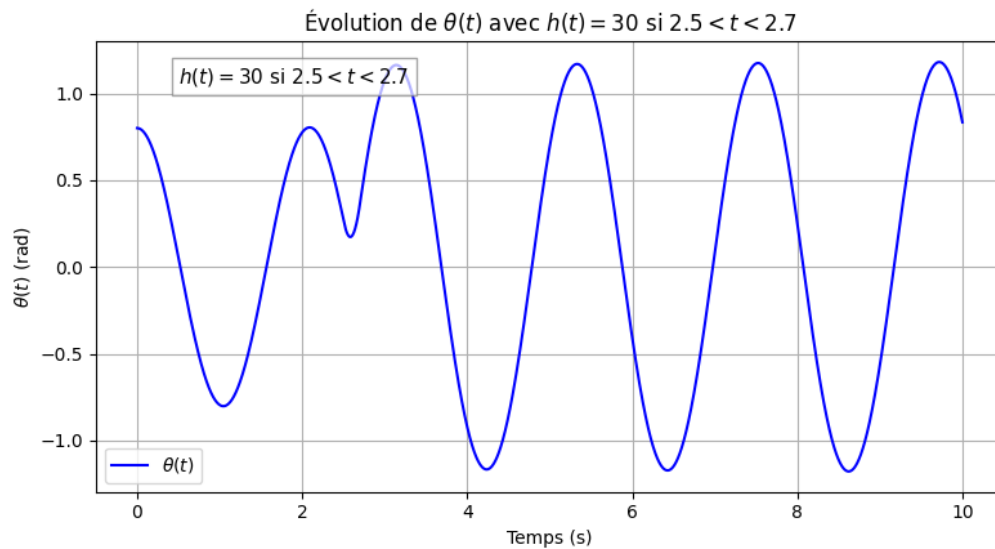
Cas 4 : $h(t) = 5t(1 - t/T)$ (contrôle en cloche)

Cette impulsion transitoire simule une poussée progressive, puis décroissante, centrée autour de $T/2$.



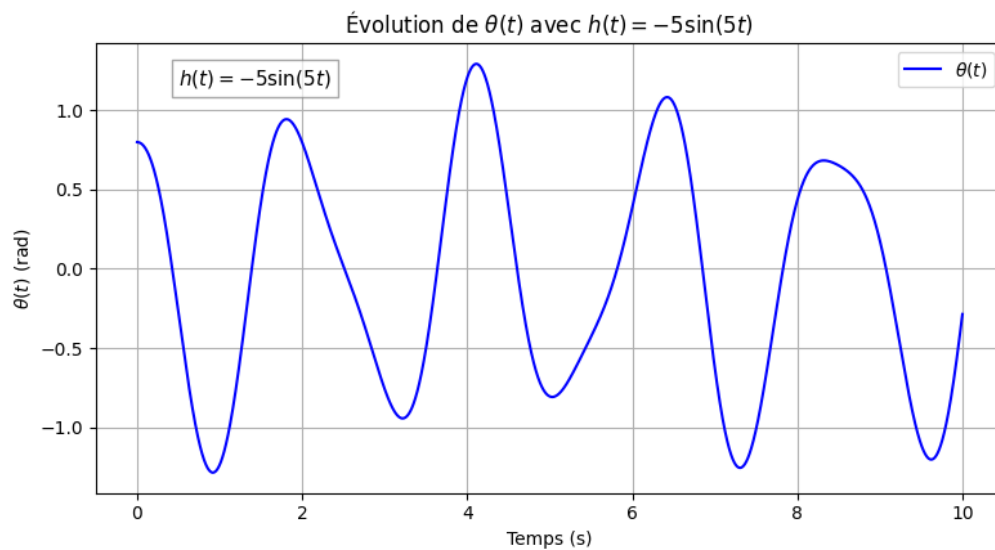
Contrôle en cloche $h(t) = 5t(1 - t/T)$

Cas 5 : $h(t) = 30$ si $2.5 < t < 2.7$ (**choc ponctuel**) Un choc bref et intense produit une perturbation nette du mouvement.



Impulsion localisée : $h(t) = 30$ si $2.5 < t < 2.7$

Cas 6 : $h(t) = -5 \sin(5t)$ (**forçage rapide**) Une excitation rapide qui provoque des oscillations complexes, voire chaotiques.



Forçage rapide $h(t) = -5 \sin(5t)$

B Gradient et hessienne d'une forme quadratique

On considère la forme quadratique $J : \begin{cases} \mathbb{R}^N \rightarrow \mathbb{R} \\ h \mapsto \frac{1}{2}h^T Q h \end{cases}$ et on démontre que

$$\nabla J(x) = Qx \text{ et } \nabla^2 J(x) = Q.$$

Soit $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{R}^N$.

Explicitons $J(x)$,

$$J(x) = \frac{1}{2}x^T Q x = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n x_i Q_{i,j} x_j$$

Soit $k \in \llbracket 1, N \rrbracket$.

On calcule la dérivée partielle par rapport à x_k .

Par linéarité de la dérivation :

$$\frac{\partial J}{\partial x_k} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n Q_{i,j} \frac{\partial (x_i x_j)}{\partial x_k}$$

Par dérivation d'un produit et en utilisant le symbole de Kronecker δ_{ik} qui vaut 1 si $i=k$ et 0 sinon,

$$\frac{\partial J}{\partial x_k}(x) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n Q_{i,j} (\delta_{ik} x_j + \delta_{jk} x_i)$$

On sépare en deux sommes doubles :

$$\frac{\partial J}{\partial x_k}(x) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n Q_{i,j} \delta_{ik} x_j + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n Q_{i,j} \delta_{jk} x_i$$

En simplifiant :

$$\frac{\partial J}{\partial x_k}(x) = \frac{1}{2} \sum_{j=1}^n Q_{k,j} x_j + \frac{1}{2} \sum_{i=1}^n Q_{i,k} x_i$$

Comme Q est symétrique, pour tout $(k, j) \in \{1, \dots, N\}^2$ $Q_{k,j} = Q_{j,k}$. On a donc deux fois la même somme et on obtient :

$$\frac{\partial J}{\partial x_k}(x) = \sum_{j=1}^n Q_{k,j} x_j$$

On en déduit le gradient de J au point x :

$$\nabla J(x) := \begin{pmatrix} \frac{\partial J}{\partial x_1}(x) \\ \frac{\partial J}{\partial x_2}(x) \\ \vdots \\ \frac{\partial J}{\partial x_N}(x) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^N Q_{1,j}x_j \\ \sum_{j=1}^N Q_{2,j}x_j \\ \vdots \\ \sum_{j=1}^N Q_{N,j}x_j \end{pmatrix}$$

$$\boxed{\nabla J(x) = Qx}$$

Soit $(k, \ell) \in \llbracket 1, N \rrbracket^2$.

On calcule la dérivée seconde :

$$\frac{\partial^2 J}{\partial x_k \partial x_\ell}(x) = \frac{\partial}{\partial x_k} \left(\sum_{j=1}^n Q_{\ell,j}x_j \right) = \left(\sum_{j=1}^n Q_{\ell,j} \frac{\partial x_j}{\partial x_k} \right) = Q_{\ell,k} = Q_{k,\ell}$$

Ainsi, la hessienne est :

$$\nabla^2 J(x) := \left(\frac{\partial^2 J}{\partial x_k \partial x_\ell} \right)_{1 \leq k, \ell \leq N} = (Q_{k,\ell})_{1 \leq k, \ell \leq N} = Q$$

$$\boxed{\nabla^2 J(x) = Q}$$

C Projeté orthogonal sur le sous-espace affine

On considère un vecteur $z := (z_1, \dots, z_n)^T \in \mathbb{R}^N$ et le sous-espace affine $C = \{h \in \mathbb{R}^N \mid Mh = 0\} + A^N Y_0$. On démontre que le projeté orthogonal de z sur C est donné par $h^* := z - M^T(MM^T)^{-1}(Mz - b)$.

Démonstration :

Projeter z orthogonalement sur C , revient à résoudre le problème :

$$\begin{cases} \min_{h \in \mathbb{R}^N} \|h - z\| \\ Mh = b \end{cases}$$

Ce problème est équivalent à :

$$\begin{cases} \min_{h \in \mathbb{R}^N} \|h - z\|^2 \\ Mh = b \end{cases}$$

En effet, la fonction $t \mapsto t^2$ est strictement croissante sur $[0, +\infty)$, donc le minimisant est inchangé. Cette reformulation est avantageuse car la fonction $h \mapsto \|h - z\|^2$ est différentiable, convexe et coercive sur \mathbb{R}^N , ce qui facilite l'analyse.

L'ensemble admissible C est **non vide, fermé et convexe** et la fonction objectif est **strictement convexe**. Par conséquent, le problème admet une **solution unique**.

Pour résoudre ce problème, on utilise la méthode des multiplicateurs de Lagrange. $\mathbb{R}^N \times \mathbb{R}^2$ est identifié à \mathbb{R}^{N+2} via la concaténation de vecteur qui est un isomorphisme. Le lagrangien associé \mathcal{L} est donné par :

$$\forall (h, \lambda) \in \mathbb{R}^N \times \mathbb{R}^2 \quad \mathcal{L}(h, \lambda) := \|h - z\|^2 + \lambda^T(Mh - b),$$

Soit $(h_1, \dots, h_n) \in \mathbb{R}^N$, $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ et $(k, l) \in \{1, \dots, N\} \times \{1, 2\}$.

Calculons les dérivées partielles de \mathcal{L} par rapport à ses différentes variables.

En explicitant les termes définissant \mathcal{L} :

$$\frac{d\mathcal{L}}{dh_k}(h, \lambda) = \frac{d}{dh_k} \left[\sum_{i=0}^N (h_i - z_i)^2 + \lambda_1 \left(\sum_{i=0}^N m_{1,i} h_i - b_i \right) + \lambda_2 \left(\sum_{i=0}^N m_{2,i} h_i - b_i \right) \right]$$

$$\boxed{\frac{d\mathcal{L}}{dh_k}(h, \lambda) = 2(h_k - z_k) + \lambda_1 m_{1,k} + \lambda_2 m_{2,k}}$$

$$\frac{d\mathcal{L}}{d\lambda_l}(h, \lambda) = \frac{d}{d\lambda_l} \left[\sum_{i=0}^N (h_i - z_i)^2 + \lambda_1 \left(\sum_{i=0}^N m_{1,i} h_i - b_i \right) + \lambda_2 \left(\sum_{i=0}^N m_{2,i} h_i - b_i \right) \right]$$

$$\boxed{\frac{d\mathcal{L}}{d\lambda_l}(h, \lambda) = \sum_{i=1}^N (m_{l,i} h_i - b_i)}$$

D'où,

$$\nabla \mathcal{L}(h, \lambda) = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial h_1}(h, \lambda) \\ \frac{\partial \mathcal{L}}{\partial h_2}(h, \lambda) \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial h_N}(h, \lambda) \\ \frac{\partial \mathcal{L}}{\partial \lambda_1}(h, \lambda) \\ \frac{\partial \mathcal{L}}{\partial \lambda_2}(h, \lambda) \end{bmatrix} = \begin{bmatrix} 2(h_1 - z_1) + \lambda_1 m_{1,1} + \lambda_2 m_{2,1} \\ 2(h_2 - z_2) + \lambda_1 m_{1,2} + \lambda_2 m_{2,2} \\ \vdots \\ 2(h_N - z_N) + \lambda_1 m_{1,N} + \lambda_2 m_{2,N} \\ \sum_{i=1}^N (m_{1,i} h_i - b_i) \\ \sum_{i=1}^N (m_{2,i} h_i - b_i) \end{bmatrix}$$

Ce qui se réécrit :

$$\boxed{\nabla \mathcal{L}(h, \lambda) = \begin{bmatrix} 2(h - z) + M^T \lambda \\ Mh - b \end{bmatrix}}$$

Les conditions de régularité sont satisfaites, car les contraintes sont affines et M est de rang 2, donc ses lignes sont linéairement indépendantes. La théorie des multiplicateurs de Lagrange s'applique, et il suffit alors de résoudre le système des conditions stationnaires de premier ordre :

$$\nabla \mathcal{L}(h_1, \dots, h_n, \lambda_1, \lambda_2) = 0 \quad \Longleftrightarrow \quad \begin{cases} 2(h - z) + M^T \lambda = 0 \\ Mh = b \end{cases}$$

Ce système linéaire admet une solution unique grâce à la coercivité de la fonctionnelle et à la régularité des contraintes. La résolution de ce système fournit donc directement la projection orthogonale recherchée, sans qu'il soit nécessaire d'examiner des conditions du second ordre.

$$\begin{aligned} \begin{cases} 2(h - z) + M^T \lambda = 0 \\ Mh = b \end{cases} &\Longleftrightarrow \begin{cases} h = z - \frac{1}{2} M^T \lambda \\ Mh = b \end{cases} \\ \Longleftrightarrow \begin{cases} h = z - \frac{1}{2} M^T \lambda \\ M \left(z - \frac{1}{2} M^T \lambda \right) = b \end{cases} &\Longleftrightarrow \begin{cases} h = z - \frac{1}{2} M^T \lambda \\ MM^T \lambda = 2(Mz - b) \end{cases} \end{aligned}$$

La matrice MM^T étant inversible,

$$\Longleftrightarrow \begin{cases} h = z - \frac{1}{2} M^T \lambda, \\ \lambda = 2(MM^T)^{-1}(Mz - b) \end{cases} \implies \boxed{h^* = z - M^T(MM^T)^{-1}(Mz - b)}$$

□

D Sur l'égalité $\text{rang}(MM^T) = \text{rang}(M)$

Pour démontrer l'inversibilité de MM^T dans la sous-section 4.4.2 nous avons utilisé le résultat d'algèbre linéaire non trivial $\text{rang}(M) = \text{rang}(MM^T)$ que l'on redémontre ci-dessous.

Démonstration :

On commence par montrer l'égalité $\ker(MM^T) = \ker(M^T)$.

On a déjà trivialement l'inclusion $\ker(M^T) \subseteq \ker(MM^T)$.

Réciproquement, soit $x \in \ker(MM^T)$. Alors :

$$MM^T x = 0 \iff x^T MM^T x = 0$$

Or :

$$x^T MM^T x = (M^T x)^T M^T x = \|M^T x\|^2$$

donc

$$\begin{aligned} \|M^T x\|^2 &= 0 \\ \implies M^T x &= 0 \quad (\text{séparation de la norme}) \\ \implies x &\in \ker(M^T) \end{aligned}$$

Ce qui prouve $\ker(M^T) \supseteq \ker(MM^T)$.

Donc

$$\boxed{\ker(MM^T) = \ker(M^T)}$$

On obtient par égalité des dimensions :

$$\dim(\ker(MM^T)) = \dim(\ker(M^T))$$

D'après le théorème du rang,

Pour toute matrice $A \in \mathcal{M}_{p,q}(\mathbb{R})$, on a : $\text{rang}(A) = q - \dim(\ker A)$.

D'où les égalités successives :

$$\text{rang}(MM^T) = m - \dim(\ker(MM^T)) = m - \dim(\ker(M^T)) = \text{rang}(M^T)$$

De $\text{rang}(M^T) = \text{rang}(M)$, on déduit finalement

$$\boxed{\text{rang}(MM^T) = \text{rang}(M)}$$

Ce qui conclut la preuve. □

E Bibliographie

- Emmanuel Trélat, *Contrôle optimal : théorie et applications*, Vuibert, 2008.
- Collectif, *Mathématiques L3 – Mathématiques Appliquées*, Pearson, 2009.