Université de Bordeaux

Master Mathématiques Appliquées et Statistiques
Parcours Ingénierie des Risques Économiques et Financiers

Systèmes dynamiques

Réalisé par : **Barkiré Douramane Moussa**

Enseignant : **Emmanuelle Augeraud**

Année académique 2024–2025

Remerciements

Je tiens à exprimer ma profonde gratitude à **Madame Emmanuelle Augeraud-Véron** pour l'organisation de ce projet, ainsi que pour son accompagnement tout au long de l'UE *Systèmes Dynamiques*. Je remercie également chaleureusement **Madame Daria Ghilli** (*Dipartimento di Scienze Economiche e Aziendali, Università di Pavia*) pour la qualité exceptionnelle des cours supplémentaires qu'elle nous a dispensés lors de son séjour à l'Université de Bordeaux. Ses enseignements, notamment sur le contrôle optimal, la programmation dynamique, les jeux différentiels et les applications des *Mean Field Games* en économie et finance, ont grandement enrichi notre compréhension théorique et appliquée de ces outils analytiques fondamentaux. Les notes de cours fournies, claires et rigoureuses, ont été un soutien précieux pour approfondir les concepts du contrôle optimal (méthodes de Pontryagin et Bellman), ainsi que pour découvrir les structures d'équilibre dans les jeux à nombreux joueurs. Ce projet a ainsi bénéficié d'un cadre académique stimulant et d'une ouverture vers des recherches modernes en économie mathématique, que je souhaite vivement prolonger dans mes travaux futurs.

Université de Bordeaux, Avril 2025

Table des matières

I Modèle à joueur unique (Contrôle optimal) 5

I.1 Formulation du problème 5

I.2 Résolution par la méthode de Pontryagin (PMP) 5

I.2.a Construction de l'Hamiltonien 6

I.2.b Conditions nécessaires du maximum 6

I.2.c Système dynamique complet 7

I.2.d Recherche de l'état stationnaire 7

I.2.e Interprétation économique 8

I.3 Résolution par la méthode de Bellman (HJB) 8

I.3.a Détermination du contrôle optimal 8

I.3.b Substitution dans l'équation HJB 9

I.3.c Hypothèse de solution quadratique 9

I.3.d Résolution et expressions des coefficients 9

I.3.e Dynamique en boucle fermée 9

I.3.f Calcul de l'état stationnaire 10

I.3.g Interprétation économique 10

I.4 Comparaison des deux méthodes (cas joueur unique) 10

I.4.a Équivalence conceptuelle fondamentale 10

I.4.b Démonstration de l'équivalence 11

I.4.c Convergence pratique 11

I.4.d Avantages comparés 11

II Jeu différentiel à plusieurs joueurs 12

II.1 Formulation du jeu 12

II.1.a Notions d'équilibre 12

II.2 Résolution de l'équilibre Open-Loop 13

II.2.a Hamiltonien de l'agent i 13

II.2.b Conditions nécessaires du maximum 13

II.2.c Système dynamique complet 14

II.2.d Recherche de l'état stationnaire 14

II.2.e Interprétation économique 15

II.3 Résolution de l'équilibre Markov-Perfect 15

II.3.a Principe général 15

II.3.b Fonction valeur et équation de Bellman 15

II.3.c Détermination du contrôle optimal 16

II.3.d Dynamique de l'état sous feedback 16

II.3.e Résolution de l'équation de Bellman 16

II.3.f Calcul de l'état stationnaire 17

II.3.g Interprétation économique 17

II.4 Comparaison des équilibres Open-Loop et Markov-Perfect 17
II.4.a Nature des stratégies 18
II.4.b Niveau de pollution stationnaire 18
II.4.c Interprétation économique 20

III Lien avec les Mean Field Games **20**
III.1 Présentation générale des Mean Field Games 20
III.2 Application au modèle de pollution 21
III.3 Comportement en limite $n \to +\infty$ 21
III.4 Lien avec le jeu différentiel classique 22
III.5 Interprétation économique 22

Introduction

L'optimisation dynamique constitue un outil fondamental en économie pour analyser les problèmes de décision intertemporelle. Dans de nombreux domaines économiques — tels que la gestion des ressources naturelles, le contrôle de la pollution, l'accumulation de capital ou encore la définition de politiques publiques — il est crucial de comprendre comment les choix présents influencent l'évolution future d'un système.

Le **contrôle optimal** fournit le cadre mathématique permettant à un agent unique de déterminer la séquence d'actions maximisant son bien-être ou son profit sur un horizon infini. Ce type d'analyse repose sur deux approches principales : d'une part, le **Principe du Maximum de Pontryagin (PMP)**, qui mène à des stratégies dites en *boucle ouverte* (open-loop), fixées au début du processus ; d'autre part, la **programmation dynamique**, qui aboutit à la résolution de l'**équation de Hamilton-Jacobi-Bellman (HJB)** et permet de caractériser des stratégies en *boucle fermée* (feedback), dépendant de l'état du système à chaque instant.

Cependant, dans de nombreuses situations économiques réelles, plusieurs agents agissent simultanément, chacun poursuivant ses propres objectifs tout en influençant l'état du système commun. Dans ce contexte, l'analyse doit passer du cadre du contrôle optimal individuel à celui des **jeux différentiels**, qui combinent dynamiques intertemporelles et interactions stratégiques.

Le présent travail, s'appuyant sur l'article de Aart de Zeeuw (2024) intitulé *"A Crash Course in Differential Games and Applications"*, propose d'étudier successivement :

- d'abord, le problème de **contrôle optimal pour un joueur unique**, en appliquant les deux méthodes fondamentales : le **Principe du Maximum de Pontryagin** et l'**équation de Hamilton-Jacobi-Bellman** ;
- ensuite, l'extension à un **jeu différentiel à plusieurs joueurs**, avec la caractérisation des équilibres dynamiques (open-loop et Markov-perfect) ;
- enfin, le lien conceptuel avec les **jeux à champ moyen (Mean Field Games)**, pertinents lorsque le nombre d'agents devient très grand.

Dans chaque cas, une attention particulière sera portée à la structure des trajectoires optimales et aux différences entre stratégies fixées initialement (open-loop) et stratégies adaptatives en fonction de l'état courant (feedback).

I Modèle à joueur unique (Contrôle optimal)

I.1 Formulation du problème

Avant d'analyser des situations stratégiques impliquant plusieurs agents, nous considérons d'abord un modèle simple de **contrôle optimal** avec un seul décideur. Ce dernier peut être interprété comme un planificateur social, une entreprise isolée ou un pays unique cherchant à optimiser ses décisions économiques sur un horizon infini.

Le décideur cherche à maximiser un **bénéfice net actualisé** dans le temps. À chaque instant, il reçoit un bénéfice de sa production $y(t)$, représenté par une fonction concave, tout en subissant un coût lié à l'accumulation d'un **stock de pollution** $s(t)$.

Le problème d'optimisation dynamique est formalisé de la manière suivante : il s'agit de maximiser la fonction objectif

$$\max_{y(\cdot)} \int_0^{+\infty} e^{-rt} \left(\beta y(t) - \frac{1}{2} y(t)^2 - \frac{1}{2} \gamma s(t)^2 \right) dt$$

où :

- $\beta > 0$ représente le bénéfice marginal initial de la production,
- $y(t) \geq 0$ est le niveau de production ou d'émission choisi par l'agent à l'instant t,
- $\gamma > 0$ mesure l'intensité du dommage associé au stock de pollution,
- $r > 0$ est le taux d'actualisation reflétant la préférence pour le présent.

La dynamique du système est donnée par une équation différentielle décrivant l'évolution du stock de pollution :

$$\dot{s}(t) = y(t) - \delta s(t), \quad s(0) = s_0 \geq 0$$

où :

- $\delta > 0$ est le taux d'assimilation naturelle de la pollution par l'environnement,
- s_0 est la quantité initiale de pollution au début de l'horizon.

Ainsi, la pollution s'accroît par l'activité productive $y(t)$ mais décroît naturellement au rythme proportionnel $\delta s(t)$. Ce problème modélise donc le compromis fondamental entre **bénéfices économiques immédiats** et **coûts environnementaux futurs**. Le décideur doit déterminer la trajectoire optimale de production $y(t)$ qui maximise son bien-être total, tout en tenant compte de l'effet cumulatif de ses décisions sur l'environnement.

I.2 Résolution par la méthode de Pontryagin (PMP)

Pour résoudre le problème de contrôle optimal présenté précédemment, nous appliquons le **Principe du Maximum de Pontryagin** (PMP). Cette approche repose sur l'introduction d'une variable auxiliaire appelée **co-état**, notée $\lambda(t)$, qui mesure la valeur économique marginale du stock de pollution accumulé.

Nous procédons en trois étapes principales : la définition du Hamiltonien, l'obtention des conditions du maximum, puis la résolution du système dynamique associé.

I.2.a Construction de l'Hamiltonien

Nous construisons d'abord l'**Hamiltonien**, qui agrège à chaque instant le gain immédiat et la contribution de l'évolution de l'état :

$$H(s, y, \lambda) = \beta y - \frac{1}{2}y^2 - \frac{1}{2}\gamma s^2 + \lambda(y - \delta s)$$

où :

- $\beta y - \frac{1}{2}y^2$ représente le bénéfice net de la production,
- $\frac{1}{2}\gamma s^2$ est le coût environnemental associé à la pollution,
- $\lambda(y - \delta s)$ capture l'effet marginal de l'évolution du stock de pollution.

I.2.b Conditions nécessaires du maximum

Premièrement, nous maximisons l'Hamiltonien par rapport au contrôle $y(t)$. En dérivant H par rapport à y, nous obtenons :

$$\frac{\partial H}{\partial y} = \beta - y + \lambda$$

En posant cette dérivée égale à zéro, la condition de premier ordre donne :

$$y(t) = \beta + \lambda(t)$$

Ainsi, la production optimale dépend directement du bénéfice marginal immédiat et de la valeur marginale de l'état.

Deuxièmement, la dynamique de l'état est donnée par l'équation :

$$\dot{s}(t) = y(t) - \delta s(t)$$

En remplaçant l'expression de $y(t)$, cela devient :

$$\dot{s}(t) = \beta + \lambda(t) - \delta s(t)$$

Troisièmement, la dynamique du co-état est déterminée par :

$$\dot{\lambda}(t) = r\lambda(t) - \frac{\partial H}{\partial s}$$

En calculant la dérivée de H par rapport à s :

$$\frac{\partial H}{\partial s} = -\gamma s - \lambda\delta$$

nous obtenons finalement :

$$\dot{\lambda}(t) = (r+\delta)\lambda(t) + \gamma s(t)$$

Cette équation décrit comment évolue la valeur marginale du stock en fonction du temps.

I.2.c Système dynamique complet

Le système d'équations différentielles couplées est donc :

$$\begin{cases} \dot{s}(t) = \beta + \lambda(t) - \delta s(t) \\ \dot{\lambda}(t) = (r+\delta)\lambda(t) + \gamma s(t) \end{cases}$$

Ce système caractérise l'évolution conjointe du stock de pollution et de sa valeur marginale associée.

I.2.d Recherche de l'état stationnaire

À long terme, on suppose que le système converge vers un état stationnaire (s^*, λ^*), pour lequel :

$$\dot{s} = 0 \quad \text{et} \quad \dot{\lambda} = 0$$

En posant ces conditions dans le système, nous obtenons :

$$\beta + \lambda^* - \delta s^* = 0$$

$$(r+\delta)\lambda^* + \gamma s^* = 0$$

De la première équation, on tire :

$$\lambda^* = \delta s^* - \beta$$

En substituant dans la seconde :

$$(r+\delta)(\delta s^* - \beta) + \gamma s^* = 0$$

$$(\delta(r+\delta) + \gamma)s^* = \beta(r+\delta)$$

$$s^* = \frac{\beta(r+\delta)}{\gamma + \delta(r+\delta)}$$

Le co-état stationnaire λ^* s'obtient en remplaçant s^* :

$$\lambda^* = \delta s^* - \beta$$

et la production optimale au steady-state est :

$$y^* = \beta + \lambda^*$$

I.2.e Interprétation économique

L'état stationnaire met en évidence un arbitrage entre bénéfice économique et coût environnemental :

- Plus le bénéfice marginal initial β est élevé, plus le stock de pollution optimal s^* est grand.
- Plus l'importance des dommages environnementaux γ est forte, plus le stock de pollution optimal est faible.

La politique optimale reflète donc un compromis entre produire aujourd'hui et préserver l'environnement pour demain.

I.3 Résolution par la méthode de Bellman (HJB)

Nous proposons maintenant de résoudre le même problème d'optimisation dynamique en utilisant la **méthode de programmation dynamique**, reposant sur l'**équation de Hamilton-Jacobi-Bellman** (HJB).

Cette approche repose sur la définition d'une **fonction valeur** $V(s)$, représentant la valeur optimale de l'objectif intertemporel maximal que l'on peut atteindre en partant d'un stock de pollution donné s.

La fonction valeur satisfait l'équation HJB suivante :

$$rV(s) = \max_y \left\{ \beta y - \frac{1}{2}y^2 - \frac{1}{2}\gamma s^2 + V'(s)(y - \delta s) \right\}$$

où :

- $rV(s)$ représente le rendement actualisé de posséder l'état s,
- $\beta y - \frac{1}{2}y^2 - \frac{1}{2}\gamma s^2$ est le gain instantané,
- $V'(s)(y - \delta s)$ correspond à la variation marginale de la valeur liée à l'évolution du stock.

I.3.a Détermination du contrôle optimal

Afin de déterminer la stratégie optimale, nous maximisons le membre de droite de l'équation HJB par rapport au contrôle y.

La condition du premier ordre est obtenue en dérivant par rapport à y et en posant cette dérivée égale à zéro :

$$\frac{\partial}{\partial y}\left(\beta y - \frac{1}{2}y^2 + V'(s)y\right) = 0 \quad \Rightarrow \quad \beta - y + V'(s) = 0 \quad \Rightarrow \quad y(s) = \beta + V'(s)$$

Ainsi, à chaque état s, le contrôle optimal y dépend du bénéfice marginal direct β et de la valeur marginale future $V'(s)$.

I.3.b Substitution dans l'équation HJB

En remplaçant l'expression de $y(s)$ dans l'équation HJB, nous obtenons :

$$rV(s) = \left(\beta(\beta + V'(s)) - \frac{1}{2}(\beta + V'(s))^2 - \frac{1}{2}\gamma s^2\right) + V'(s)\left((\beta + V'(s)) - \delta s\right)$$

Cette équation relie la fonction valeur $V(s)$ à sa dérivée $V'(s)$.

I.3.c Hypothèse de solution quadratique

Compte tenu de la structure quadratique du problème, nous cherchons une solution sous la forme :

$$V(s) = as^2 + bs + c$$

où a, b et c sont des constantes à déterminer.

En dérivant :

$$V'(s) = 2as + b$$

et en remplaçant dans l'équation HJB, nous obtenons un système qui permet d'identifier les coefficients.

I.3.d Résolution et expressions des coefficients

Après identification terme à terme, les coefficients sont donnés par (voir De Zeeuw, 2024) :

$$a = \frac{(r + 2\delta) - \sqrt{(r + 2\delta)^2 + 4\gamma}}{4} \quad \text{avec} \quad a < 0$$

$$b = \frac{2\beta a}{r + \delta - 2a} \quad \text{et} \quad c = \frac{(\beta + b)^2}{2r}$$

Ces expressions assurent que la fonction valeur est concave et que l'optimisation est bien posée.

I.3.e Dynamique en boucle fermée

La politique de contrôle optimale exprimée en fonction de l'état devient :

$$y(s) = \beta + V'(s) = \beta + 2as + b$$

La dynamique du stock de pollution sous stratégie feedback est donnée par :

$$\dot{s}(t) = y(s(t)) - \delta s(t) \quad \Rightarrow \quad \dot{s}(t) = (2a - \delta)s(t) + (\beta + b)$$

I.3.f Calcul de l'état stationnaire

L'état stationnaire s^* est atteint lorsque $\dot{s}(t) = 0$, ce qui donne :

$$(2a - \delta)s^* + (\beta + b) = 0 \quad \Rightarrow \quad s^* = \frac{\beta + b}{\delta - 2a}$$

En utilisant les expressions précédentes de a et b, on retrouve que s^* coïncide avec celui obtenu via la méthode de Pontryagin.

I.3.g Interprétation économique

La méthode de Bellman permet de construire une stratégie de contrôle en **feedback**, qui s'ajuste dynamiquement à l'état courant du système.

Ainsi, le décideur module son niveau de production en fonction de la quantité de pollution présente, assurant ainsi une flexibilité maximale face aux évolutions environnementales.

La convergence vers le même état stationnaire que celui obtenu par la méthode de Pontryagin confirme l'équivalence des deux méthodes dans ce contexte.

I.4 Comparaison des deux méthodes (cas joueur unique)

Après avoir appliqué les deux grandes méthodes d'optimisation dynamique – le **Principe du Maximum de Pontryagin (PMP)** et l'**Équation de Hamilton-Jacobi-Bellman (HJB)** – nous analysons leurs implications théoriques et pratiques.

I.4.a Équivalence conceptuelle fondamentale

- **Pontryagin** : Introduit un *co-état* $\lambda(t)$, dont l'évolution est donnée par :

  $$\dot{\lambda}(t) = r\lambda(t) - \frac{\partial \mathcal{H}}{\partial s}$$

  où $\mathcal{H}$ est l'Hamiltonien du problème.
- **Bellman** : Définit une *fonction valeur* $V(s)$, dont la dérivée première correspond à la valeur marginale du stock :

  $$V'(s(t)) = \lambda(t)$$

I.4.b Démonstration de l'équivalence

Le long de la trajectoire optimale, on observe formellement :

$$\underbrace{\dot{\lambda}(t) = (r+\delta)\lambda(t) + \gamma s(t)}_{\text{Équation adjointe (PMP)}} \quad \Leftrightarrow \quad \underbrace{rV'(s) = \frac{d}{ds}\left(\beta y(s) - \frac{1}{2}y(s)^2\right) + V''(s)\dot{s}(t)}_{\text{Différenciation de l'HJB}}$$

En effet, en différenciant l'équation de Hamilton-Jacobi-Bellman par rapport à l'état s, on retrouve une équation sur V' qui correspond dynamiquement à l'évolution du co-état $\lambda(t)$ dans Pontryagin.

Cela établit l'équivalence complète entre la méthode du co-état et la méthode de programmation dynamique.

I.4.c Convergence pratique

- **État stationnaire identique** :

$$s^* = \frac{\beta(r+\delta)}{\gamma + \delta(r+\delta)}$$

  obtenu par les deux méthodes, confirmant la cohérence interne des résultats.
- **Trajectoires équivalentes** : La solution en boucle ouverte $(y(t))$ obtenue par Pontryagin coïncide avec la stratégie feedback $(y(s))$ issue de Bellman lorsqu'on suit la trajectoire $s(t)$.

I.4.d Avantages comparés

	Pontryagin	**Bellman**
Force	Résolution analytique simple pour les problèmes linéaires-quadratiques	Modélisation des stratégies adaptatives en temps réel
Limite	Stratégies rigides (open-loop)	Complexité mathématique accrue
Usage typique	Optimisation centrale, planificateur social	Systèmes décentralisés avec comportements dynamiques

Cette analyse confirme que le choix méthodologique dépend du contexte opérationnel plutôt que d'une supériorité théorique intrinsèque. L'équivalence des résultats valide la cohérence interne des deux approches.

Comme nous le verrons dans la suite de ce projet, la capacité de Bellman à produire des stratégies en fonction de l'état sera essentielle pour définir les **équilibres de Nash en**

feedback (Markov-Perfect Equilibria) dans les jeux différentiels à plusieurs joueurs.

II Jeu différentiel à plusieurs joueurs

II.1 Formulation du jeu

Après avoir étudié le problème d'un décideur unique, nous généralisons maintenant le cadre au cas de plusieurs agents interagissant stratégiquement sur un même environnement.

Nous considérons n agents (par exemple, des pays ou des entreprises), numérotés $i = 1, \ldots, n$. Chaque agent choisit son propre niveau de production ou d'émission $y_i(t)$ afin de maximiser son bien-être intertemporel, tout en prenant en compte l'impact collectif de l'ensemble des décisions sur l'évolution du stock de pollution.

Le problème de contrôle optimal individuel de chaque agent i est formulé comme suit :

$$\max_{y_i(\cdot)} \int_0^{+\infty} e^{-rt} \left(\beta y_i(t) - \frac{1}{2} y_i(t)^2 - \frac{1}{2} \gamma s(t)^2 \right) dt$$

où :

- $y_i(t)$ est le niveau d'émission de l'agent i à l'instant t,
- $s(t)$ est le stock de pollution commun affectant tous les agents,
- $\beta > 0$ est le bénéfice marginal initial de la production,
- $\gamma > 0$ mesure l'importance des dommages liés au stock de pollution,
- $r > 0$ est le taux d'actualisation.

La dynamique du stock de pollution est influencée par la somme des émissions de tous les agents :

$$\dot{s}(t) = \sum_{i=1}^{n} y_i(t) - \delta s(t) \quad \text{avec} \quad s(0) = s_0 \geq 0$$

où $\delta > 0$ est le taux d'assimilation naturelle de la pollution.

Chaque agent maximise son propre objectif en tenant compte de la dynamique de l'état mais sans chercher à coordonner ses actions avec celles des autres. Le comportement est donc purement **non-coopératif**.

II.1.a Notions d'équilibre

Deux grands types d'équilibres stratégiques sont envisageables dans ce contexte dynamique :

- **Équilibre de Nash en open-loop** : chaque agent choisit au départ un plan d'actions dépendant du temps uniquement, sans adaptation ultérieure basée sur l'évolution du stock.

- **Équilibre de Nash en feedback (Markov-Perfect Equilibrium)** : chaque agent ajuste dynamiquement ses décisions en fonction de l'état courant $s(t)$ du système.

Dans un premier temps, nous allons caractériser l'**équilibre open-loop** en utilisant le Principe du Maximum de Pontryagin, puis nous étudierons l'**équilibre feedback** à l'aide de la méthode de Bellman.

II.2 Résolution de l'équilibre Open-Loop

Dans cette section, nous déterminons l'équilibre de Nash lorsque les agents choisissent leurs trajectoires dès le départ, sans s'ajuster dynamiquement en fonction de l'état : c'est ce qu'on appelle une stratégie en **open-loop**.

Chaque agent maximise son objectif propre, en tenant compte de la dynamique commune du stock de pollution.

Pour résoudre ce problème, nous appliquons le **Principe du Maximum de Pontryagin** à chaque agent i.

II.2.a Hamiltonien de l'agent i

L'Hamiltonien de l'agent i s'écrit comme la somme du bénéfice courant et de la valeur actualisée du changement d'état :

$$H_i(s, y_1, \ldots, y_n, \lambda_i) = \beta y_i - \frac{1}{2}y_i^2 - \frac{1}{2}\gamma s^2 + \lambda_i \left(\sum_{j=1}^{n} y_j - \delta s \right)$$

Ici, le premier terme représente le gain net, le second le dommage environnemental, et le dernier la valeur marginale associée à l'évolution du stock.

Chaque agent considère les émissions des autres comme données.

II.2.b Conditions nécessaires du maximum

Nous obtenons maintenant les conditions d'optimalité.

Maximisation par rapport à y_i : En dérivant H_i par rapport à y_i :

$$\frac{\partial H_i}{\partial y_i} = \beta - y_i + \lambda_i \quad \Rightarrow \quad y_i(t) = \beta + \lambda_i(t)$$

Chaque agent ajuste donc sa production selon son propre co-état $\lambda_i(t)$.

Dynamique de l'état : Le stock de pollution évolue selon :

$$\dot{s}(t) = \sum_{j=1}^{n} y_j(t) - \delta s(t)$$

En remplaçant les expressions optimales des contrôles, cela devient :

$$\dot{s}(t) = \sum_{j=1}^{n} (\beta + \lambda_j(t)) - \delta s(t)$$

Dynamique du co-état : La dynamique du co-état pour chaque agent i suit :

$$\dot{\lambda}_i(t) = r\lambda_i(t) - \frac{\partial H_i}{\partial s}$$

Calculons :

$$\frac{\partial H_i}{\partial s} = -\gamma s - \lambda_i \delta \quad \Rightarrow \quad \dot{\lambda}_i(t) = (r+\delta)\lambda_i(t) + \gamma s(t)$$

Ainsi, la valeur marginale de la pollution augmente avec le stock et décroît avec l'actualisation.

II.2.c Système dynamique complet

Le système dynamique complet pour l'agent i s'écrit :

$$\begin{cases} y_i(t) = \beta + \lambda_i(t) \\ \dot{\lambda}_i(t) = (r+\delta)\lambda_i(t) + \gamma s(t) \end{cases}$$

et pour l'état global :

$$\dot{s}(t) = \sum_{i=1}^{n} (\beta + \lambda_i(t)) - \delta s(t)$$

II.2.d Recherche de l'état stationnaire

À l'état stationnaire ($\dot{s} = 0$ et $\dot{\lambda}_i = 0$ pour tout i) :

$$\beta + \lambda_i^* = \text{constant} \quad \forall i$$

$$(r+\delta)\lambda_i^* + \gamma s^* = 0 \quad \forall i$$

En supposant la symétrie entre les agents ($\lambda_1^* = \lambda_2^* = \cdots = \lambda_n^* = \lambda^*$), la dynamique stationnaire devient :

$$0 = n(\beta + \lambda^*) - \delta s^*$$

En utilisant :

$$\lambda^* = -\frac{\gamma}{r+\delta} s^*$$

nous obtenons :

$$0 = n\left(\beta - \frac{\gamma}{r+\delta}s^*\right) - \delta s^*$$

En développant :

$$n\beta - \frac{n\gamma}{r+\delta}s^* - \delta s^* = 0$$

ce qui donne :

$$\left(\frac{n\gamma}{r+\delta} + \delta\right)s^* = n\beta \quad \Rightarrow \quad s^* = \frac{n\beta(r+\delta)}{n\gamma + \delta(r+\delta)}$$

II.2.e Interprétation économique

À l'équilibre open-loop :

- Plus il y a d'agents n, plus le stock de pollution s^* est élevé.
- Chaque agent internalise seulement l'impact de ses propres décisions.
- La pollution collective est donc supérieure à celle d'une situation coopérative.

Cette situation illustre la présence d'une **externalité environnementale non internalisée** dans un jeu différentiel ouvert.

II.3 Résolution de l'équilibre Markov-Perfect

Après avoir étudié l'équilibre open-loop, nous abordons maintenant l'équilibre où chaque agent ajuste dynamiquement ses décisions en fonction de l'état courant du système. Ce type d'équilibre est appelé **Markov-Perfect Nash Equilibrium (MPNE)** et repose sur la notion de stratégie feedback.

Chaque agent choisit une stratégie optimale qui dépend directement du stock courant de pollution $s(t)$.

II.3.a Principe général

Contrairement à l'open-loop, ici chaque agent i définit sa politique optimale sous forme d'une fonction :

$$y_i(t) = Y_i(s(t))$$

Le contrôle dépend donc explicitement de l'état courant, permettant une adaptation continue au fil du temps.

Cette approche repose sur la programmation dynamique, avec l'utilisation de l'équation de Hamilton-Jacobi-Bellman (HJB) adaptée au jeu différentiel.

II.3.b Fonction valeur et équation de Bellman

Chaque agent associe à chaque état s une **fonction valeur** $V_i(s)$, représentant son bien-être total attendu à partir de cet état.

L'équation de Bellman de l'agent i est donnée par :

$$rV_i(s) = \max_{y_i} \left\{ \beta y_i - \frac{1}{2}y_i^2 - \frac{1}{2}\gamma s^2 + V_i'(s) \left(y_i + \sum_{j \neq i} Y_j(s) - \delta s \right) \right\}$$

L'agent i choisit y_i pour maximiser son gain immédiat et la variation marginale de sa valeur future, tout en anticipant le comportement des autres agents.

II.3.c Détermination du contrôle optimal

Pour déterminer la politique optimale $Y_i(s)$, nous dérivons le membre de droite de l'équation de Bellman par rapport à y_i.

La condition du premier ordre donne :

$$\beta - y_i + V_i'(s) = 0 \quad \Rightarrow \quad y_i(s) = \beta + V_i'(s)$$

Ainsi, chaque agent choisit son niveau d'émission en fonction de son bénéfice marginal immédiat et de la valeur marginale qu'il associe au stock de pollution.

II.3.d Dynamique de l'état sous feedback

En supposant que les n agents utilisent des stratégies symétriques :

$$V_1(s) = V_2(s) = \cdots = V_n(s) = V(s)$$

on en déduit que tous adoptent la même stratégie :

$$y(s) = \beta + V'(s)$$

La dynamique du stock de pollution devient alors :

$$\dot{s}(t) = n(\beta + V'(s)) - \delta s(t)$$

Le stock évolue sous l'effet de la somme des décisions optimales individuelles.

II.3.e Résolution de l'équation de Bellman

En remplaçant la politique optimale dans l'équation de Bellman, nous obtenons :

$$rV(s) = \beta(\beta + V'(s)) - \frac{1}{2}(\beta + V'(s))^2 - \frac{1}{2}\gamma s^2 + V'(s)(n(\beta + V'(s)) - \delta s)$$

Cette équation différentielle relie la fonction valeur $V(s)$ à sa dérivée $V'(s)$.

Pour résoudre cette équation, nous faisons l'hypothèse naturelle que $V(s)$ est une fonction quadratique :

$$V(s) = as^2 + bs + c \quad \text{d'où} \quad V'(s) = 2as + b$$

En substituant dans l'équation précédente et en identifiant les termes, il est possible de déterminer les coefficients a, b, et c.

II.3.f Calcul de l'état stationnaire

L'état stationnaire est atteint lorsque $\dot{s}(t) = 0$, soit :

$$n(\beta + V'(s^*)) = \delta s^*$$

En résolvant cette égalité, nous obtenons :

$$s^* = \frac{n(\beta + b)}{\delta - 2na}$$

ce qui donne l'expression du stock de pollution stationnaire en fonction des paramètres économiques et du nombre d'agents.

II.3.g Interprétation économique

L'équilibre Markov-Perfect présente plusieurs caractéristiques importantes :

- Chaque agent adapte continuellement sa décision en fonction du stock actuel de pollution, offrant plus de flexibilité qu'en open-loop.
- Toutefois, chaque agent prend seulement en compte son propre impact marginal sur le stock, sans internaliser pleinement l'effet de ses décisions combinées avec celles des autres.
- Par conséquent, le niveau stationnaire de pollution s^* est généralement plus élevé qu'en open-loop.

Cette situation illustre une **aggravation de l'externalité environnementale** sous feedback : la décentralisation dynamique rend l'optimisation individuelle plus myope vis-à-vis des effets globaux. Malgré une capacité d'adaptation plus fine, le résultat collectif est moins efficace que dans le cadre d'une stratégie open-loop coordonnée.

II.4 Comparaison des équilibres Open-Loop et Markov-Perfect

Après avoir caractérisé les équilibres open-loop et Markov-Perfect, il est important de comparer leurs propriétés économiques et dynamiques. Bien que les deux types d'équilibres reposent sur des stratégies optimales dans un cadre non-coopératif, ils diffèrent par la nature de l'information utilisée et par le niveau final de pollution atteint.

II.4.a Nature des stratégies

Dans l'**équilibre Open-Loop**, chaque agent détermine au début de la période un plan d'actions optimal, en fonction du temps, sans s'adapter à l'évolution du stock de pollution en cours de route.

- Les décisions sont **pré-commitées** et fixées au départ.
- Les agents ne tiennent pas compte de l'état courant du système dans leurs choix futurs.

Dans l'**équilibre Markov-Perfect**, chaque agent ajuste dynamiquement ses décisions en fonction du stock de pollution au moment où l'action est prise.

- Les stratégies sont **adaptatives** et **réactives** à l'état actuel.
- Les agents re-optimisent en temps réel, tenant compte de l'information disponible.

II.4.b Niveau de pollution stationnaire

Un point essentiel de la comparaison porte sur le stock de pollution stationnaire s^* atteint dans chaque cas.

Les formules précédentes montrent que :

- Le stock stationnaire est **plus élevé en Markov-Perfect** qu'en Open-Loop.
- Formellement : $s^*_{\text{MPNE}} > s^*_{\text{Open-Loop}}$.

Cette différence résulte du fait que :

- En Open-Loop, les agents prennent des engagements sur l'ensemble de leur trajectoire, ce qui limite leur propension à polluer.
- En Markov-Perfect, les agents ajustent leurs décisions à court terme, ce qui réduit leur incitation à internaliser les coûts globaux futurs.

Ainsi, la stratégie feedback mène à une pollution plus importante, malgré la flexibilité qu'elle procure.

Illustration numérique

Pour illustrer concrètement la différence de niveaux de pollution entre l'équilibre Open-Loop et l'équilibre Markov-Perfect, considérons les paramètres suivants :

$$\beta = 1, \quad \gamma = 0.5, \quad \delta = 0.1, \quad r = 0.05$$

Calcul dans l'équilibre Open-Loop :

L'état stationnaire s^*_{OL} est donné par :

$$s^*_{\text{OL}} = \frac{n\beta(r+\delta)}{n\gamma + \delta(r+\delta)}$$

Pour $n = 2$ joueurs :

$$s^*_{\mathrm{OL}} = \frac{2 \times 1 \times (0.05 + 0.1)}{2 \times 0.5 + 0.1(0.05 + 0.1)} = \frac{0.3}{1 + 0.015} = \frac{0.3}{1.015} \approx 0.2956$$

Calcul dans l'équilibre Markov-Perfect :

Dans le cas Markov-Perfect, l'état stationnaire s^*_{MP} est obtenu à partir de la résolution de l'équation de Hamilton-Jacobi-Bellman.

On suppose que la fonction valeur $V(s)$ est quadratique :

$$V(s) = as^2 + bs + c$$

Le coefficient a est solution de l'équation quadratique :

$$4a^2 + 2(r + 2\delta)a + \gamma = 0$$

En remplaçant les paramètres donnés :

$$r + 2\delta = 0.25, \quad (r + 2\delta)^2 + 4\gamma = 2.0625$$

$$a = \frac{0.25 - \sqrt{2.0625}}{4} \approx \frac{0.25 - 1.436}{4} \approx -0.2965$$

Puis, en utilisant la formule de b :

$$b = \frac{2\beta a}{r + \delta - 2a} \quad \Rightarrow \quad b = \frac{2 \times 1 \times (-0.2965)}{0.05 + 0.1 - 2 \times (-0.2965)} = \frac{-0.593}{0.743} \approx -0.7981$$

Ainsi, les coefficients sont :

$$a \approx -0.2965, \quad b \approx -0.7981$$

Le stock stationnaire Markov-Perfect est alors donné par :

$$s^*_{\mathrm{MP}} = \frac{n(\beta + b)}{\delta - 2na} = \frac{2(1 - 0.7981)}{0.1 - 2 \times 2 \times (-0.2965)} = \frac{2(0.2019)}{0.1 + 1.186} = \frac{0.4038}{1.286} \approx 0.314$$

Nous obtenons :

$$s^*_{\mathrm{OL}} \approx 0.296, \quad s^*_{\mathrm{MP}} \approx 0.314$$

confirmant que :

$$s^*_{\mathrm{MP}} > s^*_{\mathrm{OL}}$$

Autrement dit, la pollution est plus importante en Markov-Perfect qu'en Open-Loop, conformément à l'analyse théorique.

II.4.c Interprétation économique

La comparaison entre les deux équilibres illustre un principe classique des externalités environnementales dynamiques :

- Dans des contextes non-coopératifs, l'absence de coordination entraîne une surexploitation du stock commun (ici, la pollution).
- L'engagement initial (open-loop) permet paradoxalement d'obtenir un meilleur résultat collectif, car les agents anticipent mieux l'effet total de leurs décisions.
- L'adaptation myope sous feedback aggrave l'effet externe en favorisant les gains immédiats au détriment du coût futur.

En résumé :

$$\text{Open-Loop : meilleur pour l'environnement} \quad < \quad \text{Markov-Perfect : plus flexible mais plus polluant}$$

Cette analyse montre que :

- La flexibilité des stratégies feedback ne garantit pas une meilleure performance collective.
- L'absence de coopération conduit systématiquement à un niveau de pollution plus élevé en dynamique feedback.

Ce constat justifie économiquement le recours à des politiques publiques (régulation, taxation, quotas) visant à internaliser les externalités environnementales et à restaurer une allocation plus efficace des ressources sur un horizon dynamique.

III Lien avec les Mean Field Games

Après avoir étudié les jeux différentiels avec un nombre fini d'agents, il est naturel de s'intéresser à la situation où le nombre d'agents devient très grand. Dans cette limite, l'interaction stratégique individuelle perd de son importance, et l'analyse passe au cadre des **jeux à champ moyen** ou **Mean Field Games (MFG)**.

Les Mean Field Games offrent un outil puissant pour étudier les interactions dans de grandes populations, en simplifiant la complexité des jeux stratégiques tout en conservant une dynamique d'optimisation intertemporelle.

III.1 Présentation générale des Mean Field Games

Un **Mean Field Game** modélise la situation où :

- Le nombre d'agents tend vers l'infini,
- Chaque agent individuel est négligeable par rapport à la population totale,

- Les décisions de chacun dépendent uniquement de la distribution agrégée de l'état ou des actions de la population.

Dans un MFG, chaque agent optimise son comportement en considérant la dynamique moyenne du système comme donnée, sans influencer directement cette dynamique.

III.2 Application au modèle de pollution

Dans notre modèle de pollution, lorsque le nombre d'agents n devient très grand :

- Chaque agent a un impact négligeable sur le stock de pollution global,
- Chaque agent optimise sa stratégie en considérant la pollution moyenne $s(t)$ comme donnée.

Le problème d'optimisation individuel devient :

$$\max_{y(\cdot)} \int_0^{+\infty} e^{-rt} \left(\beta y(t) - \frac{1}{2} y(t)^2 - \frac{1}{2} \gamma s(t)^2 \right) dt$$

avec une dynamique agrégée du stock :

$$\dot{s}(t) = \text{moyenne des émissions} - \delta s(t)$$

Chaque agent prend $s(t)$ comme exogène, ce qui simplifie considérablement le problème stratégique.

III.3 Comportement en limite $n \to +\infty$

À mesure que n augmente :

- L'effet de chaque agent sur le stock devient insignifiant,
- La dynamique de $s(t)$ est déterminée uniquement par le comportement agrégé.

Chaque agent résout donc un problème classique de contrôle optimal, où la dynamique de l'état est vue comme extérieure à ses propres décisions.

Le contrôle optimal est déterminé par la solution d'une équation de Hamilton-Jacobi-Bellman dépendante du champ moyen.

III.4 Lien avec le jeu différentiel classique

Comparé au jeu différentiel classique :

- Dans le jeu différentiel, chaque agent internalise son impact stratégique sur le stock.
- Dans le MFG, chaque agent considère le stock comme indépendant de ses propres choix.

Le passage du jeu différentiel au MFG correspond donc à une transition de l'interaction stratégique directe vers une interaction par la dynamique agrégée.

Formellement, à mesure que $n \to \infty$, l'équilibre Markov-Perfect du jeu différentiel converge vers l'équilibre du Mean Field Game.

III.5 Interprétation économique

Les Mean Field Games permettent de comprendre les dynamiques collectives dans de grandes populations :

- Même si chaque agent agit de manière optimale individuellement, le résultat collectif peut être sous-optimal.
- L'absence d'internalisation des externalités agrégées conduit à une pollution excessive.

Dans le contexte de la pollution :

- Les MFG montrent pourquoi l'intervention publique (régulation, taxation écologique) est nécessaire pour aligner les intérêts privés et l'optimum social.
- Ils offrent un cadre analytique robuste pour concevoir des politiques environnementales efficaces à grande échelle.

Conclusion

Dans ce travail, nous avons étudié différentes approches pour modéliser les décisions économiques intertemporelles dans un contexte de pollution environnementale.

Nous avons tout d'abord analysé un problème de **contrôle optimal** avec un unique décideur. En appliquant successivement le **Principe du Maximum de Pontryagin** et la **Programmation Dynamique via l'équation de Hamilton-Jacobi-Bellman (HJB)**, nous avons confirmé que, bien que reposant sur des techniques différentes, les deux méthodes aboutissent à la même politique optimale et au même état stationnaire.

Nous avons ensuite généralisé l'analyse au cadre des **jeux différentiels à plusieurs joueurs**. En étudiant l'**équilibre open-loop** puis l'**équilibre Markov-Perfect**, nous avons mis en évidence que l'adaptabilité des stratégies feedback n'empêche pas une aggravation des externalités environnementales, conduisant à un niveau de pollution plus élevé qu'en open-loop.

Enfin, nous avons établi un lien avec les **Mean Field Games**, qui permettent de modéliser les comportements stratégiques dans des populations très nombreuses. Cette approche montre que même en présence d'agents rationnels, une pollution excessive peut apparaître en l'absence de coordination ou d'intervention publique.

Ainsi, ce projet a permis d'illustrer la puissance des outils dynamiques pour comprendre les décisions économiques en environnement incertain, et de mieux appréhender les enjeux liés aux politiques de régulation environnementale.