論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Applications of the differential game framework with the linear-quadratic form and an infinite time horizon
著者(和文)	黄淵侃
Author(English)	YUANKAN HUANG
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11034号, 授与年月日:2019年3月26日, 学位の種別:課程博士, 審査員:猪原 健弘,後藤 美香,坂野 達郎,大和 毅彦,中丸 麻由子
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11034号, Conferred date:2019/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	要約
Type(English)	Outline

Applications of the differential game framework with the linear-quadratic form and an infinite time horizon

Yuankan Huang (黄　淵侃)

Supervisor: Professor Takehiro Inohara

Department of Value and Decision Science
Graduate School of Decision Science and Technology
Tokyo Institute of Technology

A Thesis for the Degree of Doctor of Science

February 2019

Abstract

In this thesis, we refine structure and present applications of the differential game framework over an infinite time horizon. The differential game can be viewed as a repeated game (or a dynamic game) in which the stage game is played repeatedly in continuous time and players' strategies dynamically manipulate the evolution of the external state over time. In order to derive optimal solutions to the game, we generalize the differential game at each instantaneous moment of the infinite time horizon, and show a long-run optimal solution to the differential game can be characterized by clarifying an optimal control path of each instantaneous moment. This result suggests both the differential game and the long-run optimal solution can be expressed in the recursive form. Making use of the recursive structure, we verify how the Hamilton-Jacobi-Bellman equation (or HJB equation for short) and Pontryagin's maximum principle work in characterizing an optimal control path (i.e., an optimal solution) of the differential game over an infinite time horizon. In particular, the HJB equation provides a sufficient condition for the optimal control of the differential game, whereas Pontryagin's maximum principle only yields a necessary condition. In this thesis, the HJB equation is the main technique that is used to characterize the equilibria in differential game models.

There exist two kinds of strategies in the differential game: the open-loop strategy and the Markovian strategy. In the open-loop strategy, each player determines a control path (i.e., a sequence of actions over the infinite time horizon) before the starts of the game, and then sticks to the control path throughout the entire duration of the game. In this sense, none of the players revises its action according to the external state after the start of the game. Different from the open-loop strategy, when the Markovian strategy is chosen, the information on the external state is always available to the players, and each player dynamically conditions its action on the external state for each instantaneous moment. We use the HJB equation to show the Markovian strategy can form a Markov perfect equilibrium (MPE) in which each player's dynamic actions satisfy the subgame perfectness (i.e., each player dynamically reoptimizes its strategy along the control path of the equilibrium). This result suggests the MPE is more acceptable than the open-loop equilibrium, and thereby we are concerned with precise characterization and examination of analytic MPEs in applications of the infinite-time-horizon differential game.

In theoretical research in economics, the differential game with quadratic utility function and linear transition function of the external state receives large attention, because such a linear-quadratic form (or LQ-form) permits the characterization of analytic MPEs. In the rest of the thesis, we employ the LQ-form differential game framework to formulate the following three subjects. In particular, we restrict attention to the stable MPE in which the equilibrium strategies guide the external state to converge to a steady state.

The best stable Markov perfect equilibrium with the highest long-run utility: In the dynamic renewable resource duopoly, we propose a new technique to show that for each feasible initial state, the long-run profit is increasing in the steady-state resource. This

result implies the MPE with the highest steady state provides the highest long-run profit, and thereby is the best equilibrium of the game. Different from the previous studies so far, our new technique provides a new method to evaluate all stable MPEs in terms of the long-run profits. We also propose an approach to numerically compute the long-run profit of each stable MPE by means of the composite Simpson's rule. This approach enables the numerical verification of the results derived by the new technique. Moreover, we generalize the LQ-form differential game framework, and show the above the new technique is possibly valid to other LQ-form differential game models such as e.g., the transboundary pollution control model and the dynamic public goods provision problem.

Public goods provision under dynamic budgets: We formulate the public goods provision problem subject to dynamic budget constraints. In the game, each individual (i.e, each player) is assumed to earn income over time, and allocates the income between the private consumption and the public goods provision subject to dynamic budgets. Individuals' contribution of the public good results in a positive stock externality of the game. We characterize all stable MPEs including a linear one and infinitely many nonlinear ones, and specify a simple equation that describes the relation between the level of the income and the amount of the public contribution in any stable MPE. Making use of this relation equation, we show the effect of the public-good level on the budget and the allocation of the income can be easily analyzed. In other words, the relation equation enables the explicit examination on whether each individual raises or curtails its level of income and the amount of public-good contribution as the state variable is approaching the steady state during the play of the game. Additionally, we also show the new technique proposed by the previous subject also applies to the present model. Making use of the new technique, we evaluate and rank the stable MPEs in terms of the long-run utilities.

A dynamic policy game of the combination use of emission taxation and a pollution-removal technique: The game comprises multiple polluting oligopolistic firms and a government who takes charge of the pollution management. The government seeks to optimally control the pollution by means of emission taxation together with a pollution-removal technique. During the play of the game, we assume the government first collects the emission tax from the firms, and then spends tax revenue on the expense of the pollution-abatement efforts or equally redistributes the tax revenue to each firm. We show there only exists a unique stable linear MPE in the game. In the MPE, when the pollution level is sufficiently close to the steady state, the tax revenue exceeds the cost of the pollution-removal efforts, so that the government can reimburse the surplus of the tax revenue to the firms.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text.

Yuankan Huang
February 2019

Acknowledgements

First of all, I would like to express my sincere gratitude to my academic advisor Prof. Takehiro Inohara for the continuous support of my Ph.D. study and related research. During the ten years I spent in Tokyo Tech, he has provided patience, motivation and immense knowledge which guided me in all time of study and research.

Besides my academic advisor, I would like to thank the other members of my thesis committee: Prof. Mika Goto, Prof. Mayuko Nakamaru, Prof. Takehiko Yamato, and Prof. Tatsurou Sakano, for their insightful comments which greatly helped me in revising this thesis.

I would like to thank the secretary of my lab, for her care which has helped me not only in dealing with the academic affairs but also understanding the Japanese culture and social norms. Also, I would like to thank my friends in Tokyo Tech, for the stimulating discussions about research and the encouragement along the way.

Contents

Chapter 1

Introduction

1.1 The linear-quadratic differential game framework

In theoretical research in economics, the differential game framework (Dockner et al. (2000, Chapter 3 and Chapter 4)) is widely employed to formulate the long-run externality issues. In the differential game, a one-shot game (or a stage game) is played repeatedly over a continuous time horizon, [1] and the players' actions dynamically manipulate the external state over time. The long-run utility could be influenced by the external state. Each player seeks to make optimal trade-offs between choice of the actions and the external state for each time point, so that it can maximize the long-run utility in the game.

The differential game framework is defined over a continuous time horizon, and all functions (namely the utility function and the transition function of the external state) are assumed to be continuously differentiable. This setup makes it possible to employ calculus to derive and examine the equilibria of the game. In particular, many studies concentrate on the differential game framework in which the utility function and the transition function are respectively defined in the quadratic form and the linear form. Such a linear-quadratic (or LQ) setup permits characterization of the analytic equilibria. For example, in pioneer studies, Fershtman and Nitzan (1991), Dockner and Long (1993), Benchekroun (2003) and Reynolds (1987) employ the LQ-form differential game framework to formulate the problems of the dynamic public goods provision, the transboundary pollution control, the resource-extraction duopoly and investment in production capacity, respectively.

1.2 Markov perfect equilibria of the differential game

There exist two kinds of equilibria in the above LQ-form differential game framework:

(i) The open-loop equilibrium: Each player chooses a strategy (i.e., a list of the actions over the continuous time horizon) at the very beginning of the game, and plays the game along the control path specified by the strategy for the entire duration of the

---

[1] The differential game can be viewed as a repeated game with an externality (or multiple externalities) and a continuous time horizon. In other words, the differential game is equivalent to a continuous-time-version dynamic game (Mailath and Samuelson (2006, Section 5.5~5.7)).

game. Hence, the players base their strategies only on the information available at the beginning, but do not revise their strategies using the new information that becomes available in time.

(ii) The Markov perfect equilibrium: For each time point of the game, the information on the external state is available to the players, and thereby each player conditions its choice of actions on the external state variable. In other words, in the Markov perfect equilibrium, each player is allowed to reoptimize its strategy conditional on the current state for each time point.

It is well known that the Markov perfect equilibrium (or MPE for short) satisfies subgame perfectness and thereby it is more acceptable than the open-loop equilibrium in the research of the differential game. Moreover, a differential game could have multiple or even an infinite number of Markov perfect equilibria (or MPEs for short). Many studies restrict attention to the characterization and investigation of the stable MPEs that guide the external state to converge to a specific steady state. Tsutsui and Mino (1990) proposes a technique to graphically characterize all nonlinear stable MPE of a differential game. Itaya and Shimomura (2001) develops a simpler technique by means of the envelope theorem. Fujiwara (2008) further generalizes the Itaya and Shimomura technique and specifies the best level of the external state.

1.3 The main subjects of the thesis

In this thesis, to increase the understanding of the differential game framework, we first compare the repeated game and the dynamic game in terms of the structure. The two games are common in that the players engage in a one-shot game (or a stage game) repeatedly and seek to maximize the long-run utility by choosing an optimal strategy profile. On the other hand, in the dynamic game, the players' actions also manipulate the external state. Thus, different from the repeated game, the players in the dynamic game has to take into the account the external state when addressing the optimization issue. The dynamic game is played over a discrete time horizon, whereas the differential game is a special version of the dynamic game over a continuous time horizon. In the present thesis, we first refine the differential game framework with an infinite time horizon, and then formulate three models with the use of the differential game framework:

Dockner et al. (2000, Chapter 3) generalizes a single-player differential game model with a finite time horizon. The single-player structure simplifies explanation of the optimization issue of the differential game. Different from Dockner et al. (2000, Chapter 3), we refine the structure of the differential game on the infinite time horizon, and precisely show the infinite time horizon can also be recursively split into an infinite number of the instantaneous moments. Such a recursive structure suggests the optimization issue of the single-player differential game can be solved by maximizing the utility for each instantaneous moment. On the basis of these results, we further clarify the processes how an optimal control path can be characterized by means of the Hamilton-Jacobi-Bellman equation (HJB equation) and Pontryagin's maximum principle on the infinite time horizon. Such processes exhibit that any optimal control path can be characterized by the HJB equation and confirmed by Pontryagin's maximum principle.

Moreover, in the practical formulation of, e.g., the dynamic public goods provision issue and the transboundary pollution control issue, etc., this process makes it possible to confirm whether a given HJB equation can be correctly transformed into the maximization of the long-run utility of the differential game.

Different from Dockner et al. (2000, Chapter 4), we further refine the multiple-player differential game framework on the infinite time horizon. For each player in the game, the individual optimization issue is dependent on the belief that the other players must stick to their optimal strategies. This setup is different from the above single-player version and complicates structure of game. Nevertheless, we show the process examined in the single-player version is still valid to the present multiple-player differential game. This result suggests any Markovian Nash equilibrium can still be characterized by HJB equation stated above. Finally, we focus on the Markov perfect equilibrium (MPE) in which in any subgame the control path derived by the optimal strategy profile forms a Markovian Nash equilibrium. In other words, the MPE is indeed a subgame perfect equilibrium of the multiple-player differential game. We further show any MPE can be characterized by the HJB equation, and the HJB equation can be transformed into the maximization of the long-run utility of the multiple-player differential game on the infinite time horizon.

On the basis of the above results, we employ the multiple-player differential game framework to formulate the following three subjects. In particular, we are concerned with the characterization and examination of the MPEs of the game by means of the HJB equation.

Subject-1: The best stable Markov perfect equilibrium in differential games with linear quadratic form

In this subject, we generalize the LQ-form differential game framework, and propose a new technique to clarify how the equilibrium strategies and the steady state influence the long-run profit for each stable MPE. Applying this technique to the dynamic renewable resource duopoly (Fujiwara (2008)), we show for each feasible initial state, the long-run profit is increasing in the steady-state resource. Hence, the MPE with the highest steady state provides the highest long-run profit, and thereby is the best equilibrium of the game. Moreover, we show the best MPE is independent of the discount rate but possibly changes with the initial state. We also propose an approach to numerically compute the long-run profit of each stable MPE by means of the composite Simpson's rule. This approach enables the numerical verification of the results derived by the new technique.

So far many previous studies, e.g., Docker (1993), Wirl (1996), Itaya and Shimomura (2001), Rubio and Casino (2002), Fujiwara (2008) and Lambertini and Motovani (2014) are all concerned with characterization of the MPE that provides the best steady state. In contrast, our technique provides a new method to evaluate and rank all stable MPEs in terms of the long-run profits. Moreover, we also show the new technique possibly applies to all LQ-form differential games. Hence, in the model formulated by Subject-2 and Subject-3, the long-run utility of each MPE is also computed with the use of the new technique.

Subject-2: Public goods provision under dynamic budgets

In this subject, we incorporate dynamic budget constraints into the public goods provision problem presented by Fershtman and Nitzan (1991), Wirl (1996) and Itaya and Shimomura (2001). In the game, each individual is assumed to make efforts to earn income over time, and allocates the income between the private consumption and the public-good contribution subject to dynamic budget constraints. Accumulation of the individuals' contribution of public goods results in a positive stock externality of the game. In this setup, we drive all stable MPEs including a linear one and an infinite number of nonlinear ones. Due to the dynamic-budget setup, we show it is possible to clarify the relation equation between the level of income and the amount of public-good contribution. The relation equation enables the specification of the interior-solution condition of the stable MPEs, and the analyses on the effect of the public good level on the dynamic budgets as well as the allocation of incomes. In other words, making use of the above relation equation, it can be precisely examined if each individual raises or curtails its level of income and the amount of public-good contribution, as the state variable is approaching the steady state during the play of the game. This result is different from the fixed-income setup stated by Itaya and Shimomura (2001), Ihori and Itaya (2001) and Yanase (2006). Furthermore, we also show the technique proposed by Subject-1 also applies to the present model. That is, it is possible to evaluate all stable MPEs in terms of long-run utilities. As a consequence, we show in the set of the nonlinear stable MPEs, the long-run utility is increasing in the steady state, and thereby the stable MPE with the highest level of the public goods is the best equilibrium (i.e., the most socially efficient equilibrium) of the game.

Subject-3: A dynamic policy game of the combination use of emission taxation and a pollution-removal technique

We consider a policy game in which the government seeks to optimally control pollution by means of emission taxation together with a pollution-removal technique. In the game, multiple firms engage in a polluting oligopoly and the government takes charge of the pollution management. The emission taxation reduces the pollution emission arising from the firms' production but it hurts the industry profit and the consumer surplus. In contrast, utilization of the pollution-removal technique directly diminishes the pollution level but it could be quite costly. We show there only exists a unique stable linear MPE in the game. In the MPE, we find it is convenient to examine the impact of the efficiency of the pollution-removal technique and the impact of the harmfulness of pollution stock on the social welfare at the steady state. Moreover, we also show when the pollution level is sufficiently close to the steady state, the tax revenue exceeds the cost of the pollution-removal efforts, and thereby the government can reimburse the surplus of the tax revenue to the firms.

This model is analogous to Yanase (2007) and Yanase (2009) in terms of the assumption that neither the representative consumer nor the firms are not concerned with the damage resulting from the pollution. Different from these two previous studies, our model incorporates the emission taxation (Wirl (1994) and Rubio and Escriche (2001)) and the use of a pollution abatement technique (Yeung (2007) and Manoussi and Xepapadeas (2017)) into the policy game. Due to the simple structure of game, we characterize a unique analytic stable MPE that describes the optimal combination use of the dynamic emission taxation together with the pollution abatement technique. In particular, the stable MPE also enables precise examination on how the government transfer its cash flow from the tax revenue resulting from the emission taxation to the expense of the pollution abatement technique.

1.4 Outline of the thesis

The present thesis is organized as follows: In Chapter 2, we simply review the structure of the repeated game and the dynamic game, and discuss the relation between the dynamic game and the differential game. In Chapter 3, we redefine the fundamental model of a single-player differential game over an infinite time horizon, and verify how the HJB equation and Pontryagin's maximum principle work in the characterization of an optimal solution to the game. In Chapter 4, we extend the differential game stated in Chapter 3 to the general version with multiple players, and verify the characterization of the MPE with the use of the HJB equation. In Chapter 5~7, we mainly employ the LQ differential game framework together with the Itaya and Shimomura technique to deal with the three subjects stated in Section 1.3.

Chapter 2.

The repeated game and the dynamic game

2.1 Introduction

To enhance the understanding of differential game, this chapter introduces the frameworks of the repeated game and the dynamic game. In both of them, each player engages in a simultaneous stage game (or a simultaneous one-shot game) repeatedly over an infinite but discrete time horizon.[2] Each player seeks to choose an optimal strategy to play the game. Along the action path derived by the optimal strategy profile, the players' actions possibly do not provide the maximum payoff of the stage game for each period, but the sequence of actions describes the players' optimal long-run behaviors of the whole game. In other words, the actions sequence derived by the optimal strategy profile maximizes the long-run utility. The dynamic game is indeed a repeated game with a state variable (or multiple state variables). In the dynamic game, as the players are choosing the action for each period, the players' actions dynamically manipulate the evolution of the external state over time.

In this chapter, using the notation convention of Mailath and Samuelson (2006, Chapter 2), we first generalize the framework of the repeated game. We show the repeated game takes the recursive form. Therefore, the optimization issue of the whole game can be solved by maximizing the net utility consisting of the current and the future payoffs for each period. In order to further explain the image of the subgame perfect equilibrium (or SPE for short), we consider the two-player repeated prisoner dilemma and assume each player chooses a fixed mixed strategy. Herein each player's mixed strategy expresses its behavior pattern (i.e., the probability at which it cooperates and defects). According to Huang and Inohara (2015-a), making use of the folk theorem, the players can have a SPE strategy profile with a sufficiently high discount rate. In particular, it can be interpreted that the two players can have a continual interaction if their strategy profile forms a SPE, whereas their interaction will be closed if any player has incentive to trigger the punishment (i.e., to defect perpetually) during play of the game. On the basis of this result, we model the group separation with the use of two-player repeated prisoners' dilemma and the folk theorem: There are multiple players engaging in the game, and all players differ in the behavior patterns (i.e., the mixed strategies).We show with the common discount rate varying, the group separation emerges in diverse patterns. Thus, given different patience rates on the future play, the players form different groups.

As stated before, the dynamic game is indeed a repeated game with an external state (or multiple external states). In this chapter, we use the above notation convention to present the framework of the dynamic game with an infinite time horizon. In particular,

[2] The terminology "simultaneous" refers to all players move simultaneously in the one-shot game

we restrict attention to a specific dynamic game model named *dynamic extraction game of fish resources* (Mailth and Samuelson (2006, page 178)). In the game, each player benefits from extracting fish resource from a common pool, and the players' amounts of the extraction influence level of fish in the subsequent periods. In this setup, we derive a linear stationary Markov equilibrium via the recursive structure. Moreover, we point out that this dynamic game can be simply transformed to a differential game by replacing the discrete time horizon with a continuous one.

This chapter is arranged as follows: Section 2.2 defines a notation convention of the repeated game. Making use of the notation convention, Section 2.3 gives two simple examples of the repeated two-player prisoners' dilemma game. Section 2.4 introduces the framework of the dynamic game by incorporating a state variable into the repeated game. Section 2.5 states the relation between the dynamic game and the differential game.

2.2 The repeated game with an infinite horizon

In repeated games, players play a one-shot game (or a stage game) for multiple periods. For example, if the stage game is defined to be the prisoners' dilemma, then the players are supposed to play the prisoners' dilemma for each period, and each of the players seeks to choose a long-run behavior to maximize sum of the utilities over the entire time horizon.

Consider the repeated game that consists of an infinite number of periods. Then the stage game is played for each period $t \in \{0,1,\dots\}$. Assume there exist n players. In the stage game, for each player $i \in \{1,2,\dots,n\}$, A^i denotes i's action set and $a^i \in A^i$ expresses an action. The set of action profiles is given by $A \equiv \prod_i A^i$ where $\prod$ stands for the Cartesian product operator. Stage game payoffs are measured by a continuous function $F^i\colon A \to \mathbb{R}$ where $\mathbb{R}$ denotes the set of real numbers. Letting $\rho \in [0,1)$ express the player's common discount rate on the future payoffs and supposing each player i's action set does not change with the period t, player i's average discounted payoff is given by

$$\frac{\sum_{t=0}^{\infty} \rho^t\, F^i\left(a^i(t), a^{-i}(t)\right)}{\sum_{t=0}^{\infty} \rho^t} = (1-\rho)\sum_{t=0}^{\infty} \rho^t\, F^i\left(a^i(t), a^{-i}(t)\right) \qquad (2.2.1)$$

where $a^{-i}(t) = (a^1(t), a^2(t), \dots, a^{i-1}(t), a^{i+1}(t), \dots, a^n(t))$, and $a^i(t) \in A^i$ for each $i \in \{1,2,\dots,n\}$ and each $t \in \{0,1,\dots\}$. On the left-hand side of (2.2.1), the numerator $\sum_{t=0}^{\infty} \rho^t\, F^i(a^i(t), a^{-i}(t))$ reveals the long-run utility is measured by summing up the stage-game payoffs over the infinite time horizon, and the long-run utility is normalized by being divided by the sum of the common discount rate $\sum_{t=0}^{\infty} \rho^t$. In (2.2.1), $(a^i(t), a^{-i}(t)) = (a^1(t), \dots, a^n(t))$ stand for an outcome at the period t. Let $a(t) = (a^1(t), \dots, a^n(t))$ for any $t \in \{0,1,\dots,n\}$. Then the play of the repeated game generates an outcome path $\mathbf{a} = (a(0), a(1), a(2), \dots) \in A^{\infty}$ where $A \equiv \prod_i A^i$.

Let us consider the repeated game in which all players observe the opponents' actions so far at end of each period.[3] At the period $t \geq 0$, the set of histories is given by $\mathcal{H}^t \equiv A^t$

[3] Such a kind of the repeated game is formally named *the repeated game of perfect monitoring*.

where we define the initial history to be the null set, $A^0 \equiv \{\emptyset\}$, and A^t to be the t-fold product of A. Thus, a history $h^t \in \mathcal{H}^t$ is a list of t action profiles and h^t identifies the actions chosen from the period 0 to period $t-1$. For example, an output come path $h^t = (a(0), a(1),, \dots, a(t-1))$ must lie in the set $\mathcal{H}^t$. At the beginning of the period $t+1$, the actions profiles chosen from period 0 to period t yields the history h^{t+1} such that $h^{t+1} \in \mathcal{H}^{t+1} = \mathcal{H}^t \times A$. The set of all possible histories is given by $\mathcal{H} \equiv \bigcup_{t=0}^{\infty} \mathcal{H}^t$. A repeated game strategy (or strategy for short) is defined to be $\sigma^i : \mathcal{H} \to A^i$ and a strategy profile of the whole repeated game is given by $\sigma = (\sigma^1, \sigma^2, \dots, \sigma^n)$.

Using the notation convention defined above, the strategy profile $\sigma = (\sigma^1, \dots, \sigma^n)$ yields an outcome path $\mathbf{a}(\sigma) \equiv (a(\sigma)(0), a(\sigma)(1), a(\sigma)(2), \dots)$ where $a(\sigma)(t)$ stands for an outcome $a(t) = (a^1(t), \dots, a^n(t))$ yielded by σ at each period $t \in \{0,1,2,\dots\}$. Accordingly, σ generates an outcome as follows: In the first period, the action profile is played:

$$a(\sigma)(0) = \left(\sigma^1(\emptyset), \dots, \sigma^n(\emptyset)\right). \tag{2.2.2}$$

In the second period, the history $a(\sigma)(0)$ is observed and the players choose the following action profile

$$a(\sigma)(1) = \left(\sigma^1\big(a(\sigma)(0)\big), \dots, \sigma^n\big(a(\sigma)(0)\big)\right) \tag{2.2.3}$$

Recursively, in the third period, the history becomes $(a(\sigma)(0), a(\sigma)(1))$ and thereby the action profile

$$a(\sigma)(2) = \left(\sigma^1\big(a(\sigma)(0), a(\sigma)(1)\big), \dots, \sigma^n\big(a(\sigma)(0), a(\sigma)(1)\big)\right) \tag{2.2.4}$$

is played, and so on. Particularly, in the repeated game that begins in period t,[4] in the strategy profile σ conditional on the history $h^t \in \mathcal{H}$, player i chooses the action $\sigma^i|_{h^t} = \sigma^i(h^t)$ and thereby the strategy profile starts at period t is given by $\sigma|_{h^t} = \sigma(h^t)$. Incorporating the players' strategy notation into the long-run utility (2.2.1) yields

$$U^i(\sigma) = (1-\rho)\sum_{k=0}^{\infty} \rho^t \, F^i\big(a^i(\sigma)(k), a^{-i}(\sigma)(k)\big). \tag{2.2.5}$$

where k represents the period index.

In the period t, the long-run utility turns out to be

$$U^i(\sigma|_{h^t}) = (1-\rho)\sum_{k=t}^{\infty} \rho^{k-t} \, F^i\big(a^i(\sigma|_{h^k})(k), a^{-i}(\sigma|_{h^k})(k)\big). \tag{2.2.6}$$

(2.2.6) can be rewritten in the following recursive form:

$$U^i(\sigma|_{h^t}) = (1-\rho)F^i\left(a^i(\sigma|_{h^t})(t), a^{-i}(\sigma|_{h^t})(t)\right) + \rho U^i(\sigma|_{h^{t+1}}). \tag{2.2.7}$$

Proposition 2.2.1

In the repeated game, σ is a subgame perfect equilibrium (SPE) if and only if for any t, any history $h^t \in \mathcal{H}$ and any action $\tilde{a}^i \in A^i$, it holds that

$$\begin{aligned}(1-\rho)F^i\left(a^i(\sigma|_{h^t})(t), a^{-i}(\sigma|_{h^t})(t)\right) &+ \rho U^i(\sigma|_{h^{t+1}}) \\ &\geq (1-\rho)F^i\left(\tilde{a}^i, a^{-i}(\sigma|_{h^t})(t)\right) + \rho U^i(\sigma|_{\tilde{h}^{t+1}})\end{aligned} \tag{2.2.8}$$

where $\tilde{h}^{t+1}$ is the history corresponding to the action profile $\left(\tilde{a}^i, a^{-i}(\sigma|_{h^t})(t)\right)$.

[4] The repeated game starting at the period t is a subgame of the whole repeated game.

Proposition 2.1.1 states that during the play of the repeated game, if it is unprofitable for any player to deviate at the start of any subgame, then the strategy profile forms a SPE.

2.3 Simple examples: The folk theorem in repeated prisoners' dilemma games

Define the stage game to be the two-player prisoners' dilemma game with the payoff matrix Table 2.3.1. In the repeated prisoners' dilemma, if the game is played with pure action, then we have $A^i = \{C, D\}$ for each $i \in \{1,2\}$. Next we define a pure strategy profile of the repeated prisoners' dilemma game.

Definition 2.3.1:
The definition of the strategy profile σ_p

(1) If no player deviates, the two players play the stage game with the action profile (C, C) *for each period.*

(2) At any period t, *if player* i *deviates from the action profile* (C, C), *then the game will be played following the action profile* (P, P) *for each period thereafter.*

		Player 2	
		C	*D*
Player 1	*C*	*R,R*	*S,T*
	D	*T, S*	*P,P*

Table. 2.3.1: *C* and *D* stand for the actions *cooperate* and *defect* of the prisoners' dilemma game. *R*, *T*, *S* and *P* are the rational numbers and satisfy the constraints $T > R > P > S$ and $R > (T + S)/2$.

Recall (2.2.8), set $(a^i(\sigma_p|_{h^t})(t), a^{-i}(\sigma_p|_{h^t})(t)) = (C, C)$, $(\tilde{a}^i, a^{-i}(\sigma_p|_{h^t})(t)) = (D, C)$, $\sigma_p|_{h^{t+1}} = (EE, EE, EE, \dots)$ and $\sigma_p|_{\tilde{h}^{t+1}} = (DD, DD, DD, \dots)$. σ_p states the players will play the game as promised if there is no deviation arising. Otherwise, when player i deviates by choosing $\tilde{a}^i$, its opponent j will choose Defect perpetually to punish the deviation. σ_p is called the trigger strategy in the context of repeated prisoner dilemma games. Substituting the corresponding value defined by Table 2.3.1 into (2.2.8) yields

$$(1-\rho)R + \rho R \geq (1-\rho)T + \rho P. \qquad (2.3.2)$$

In addition, let $(a^i(\sigma_p|_{h^t})(t), a^{-i}(\sigma_p|_{h^t})(t)) = (D, D)$, $(\tilde{a}^i, a^{-i}(\sigma_p|_{h^t})(t)) = (C, D)$, $\sigma_p|_{h^{t+1}} = \sigma_p|_{\tilde{h}^{t+1}} = (DD, DD, DD, \dots)$. Then (2.2.8) can be rewritten as

$$(1-\rho)P + \rho P \geq (1-\rho)S + \rho P. \qquad (2.3.3)$$

It is clear that (2.3.2) and (2.3.3) both holds with $\rho \geq (T-R)/(T-P)$. In other words, when the common discount rate satisfies $\rho \geq (T-R)/(T-P)$, σ is a SPE of the game.

In contrast to the pure-strategy folk theorem stated above, we next construct a repeated game strategy with the fixed mixed actions. The set of mixed actions is defined to be $A^i = [0,1] \times [0,1]$, and player i chooses the mixed action $a^i = (\delta^i, 1-\delta^i) \in A^i$ where $\delta^i \in [0,1]$ and $1-\delta^i \in [0,1]$ stand for probabilities assigned to the action C and D

respectively. The repeated game strategy of the mixed action is defined as follows:

Definition 2.3.2:

The definition of the strategy profile σ_m:

(1) Assume each player plays the stage game following a fixed mixed action $a^i = (\delta^i, 1-\delta^i)$ where $\delta^i \in (0,1)$. a^i stands for player i's behavior pattern in which player i chooses C (resp. D) at the probability δ^i (resp. $1-\delta^i$)

(2) If no player deviates, then the two players play the stage game with action profile $(a^1, a^2) = ((\delta^1, 1-\delta^1), (\delta^2, 1-\delta^2))$ for each period.

(3) At any period t, if player i deviates from the action profile (a^1, a^2), then the game will be played following the action profile $(\hat{a}^1, \hat{a}^2) = ((0,1), (0,1)) = (P, P)$ for each period thereafter.

In σ_m, when the stage game is played following the mixed action $(a^1, a^2) = ((\delta^1, 1-\delta^1), (\delta^2, 1-\delta^2))$, player i gains the payoff:[5]

$$F^i(a^i, a^j) = \delta^i[\delta^j R + (1-\delta^j)S] + (1-\delta^i)[\delta^j T + (1-\delta^j)P] \tag{2.3.4}$$

Rewrite (2.2.8) as

$$\begin{aligned}(1-\rho)F^i\left(a^i(\sigma_m|_{h^t})(t), a^{-i}(\sigma_m|_{h^t})(t)\right) + \rho U^i(\sigma_m|_{h^{t+1}}) \\ \geq (1-\rho)\max_{\tilde{a}^i \in A^i} F^i\left(\tilde{a}^i, a^{-i}(\sigma_m|_{h^t})(t)\right) + \rho U^i(\sigma_m|_{\tilde{h}^{t+1}}).\end{aligned} \tag{2.3.5}$$

Definition 2.3.1 assumes the stage game is played following the action profile $(a^1, a^2) = ((\delta^1, 1-\delta^1), (\delta^2, 1-\delta^2))$ if no deviation arises so far. Therefore, according to Table 2.3.1, it can be easily shown that

$$\underset{\tilde{a}^i \in A^i}{\operatorname{argmax}} F^i\left(\tilde{a}^i, a^{-i}(\sigma|_{h^t})(t)\right) = \{(0,1)\}. \tag{2.3.6}$$

That is, due to assumption $T > R > P > S$, it holds that

$$\begin{aligned}\delta^j T + (1-\delta^j)P \geq \delta^i[\delta^j R + (1-\delta^j)S] + (1-\delta^i)[\delta^j T + (1-\delta^j)P] \\ \Leftrightarrow \delta^j T + (1-\delta^j)P \geq \delta^j R + (1-\delta^j)S, \quad j \neq i\end{aligned} \tag{2.3.7}$$

Therefore, if player i has an incentive to deviate then it will take the mixed action in which the cooperation and the defect are respectively chosen at the probability 0 and 1 (i.e., the probability 0% and 100%).

Analogous to σ_p, it follows that σ_m is a SPE if and only if (2.3.5) and (2.3.3) both hold for each player $i \in \{1,2\}$.

Plugging (2.3.4) into (2.3.5) gives

$$\begin{aligned}\delta^i[\delta^j R + (1-\delta^j)S] + (1-\delta^i)[\delta^j T + (1-\delta^j)P] \\ \geq (1-\rho)[\delta^j T + (1-\delta^j)P] + \rho P\end{aligned} \tag{2.3.8}$$

Define a function $\theta: [0,1] \times [0,1] \to [0,1]$ to be

$$\theta(\delta^i, \delta^j) = \delta^i \frac{(T-R)-(P-S)}{(T-P)} + \frac{\delta^i}{\delta^j}\frac{(P-S)}{(T-P)}. \tag{2.3.9}$$

It can be shown that if $\rho \geq \theta(\delta^i, \delta^j)$ then (2.3.8) and (2.3.3) both hold for player $i \in \{1,2\}$. Because $\theta(\delta^i, \delta^j)$ does not always equal $\theta(\delta^j, \delta^i)$, it can be easily shown that if $\rho \geq \max\{\theta(\delta^i, \delta^j), \theta(\delta^j, \delta^i)\}$ then σ_m is a SPE. In other words,

[5] In the game with mixed actions, the payoff is measured in expected value.

$\max\{\theta(\delta^i, \delta^j), \theta(\delta^j, \delta^i)\}$ is the lower bound of the common discount rate for σ_m to form a SPE.

The above result implies that given a common discount rate $\rho \in (0,1)$, whether σ_m is a SPE is dependent on the players' behavior patterns $(a^1, a^2) = ((\delta^1, 1-\delta^1), (\delta^2, 1-\delta^2))$.

Assumption 2.3.1:
In Table 2.3.1, assume $T > 0$, $R > 0$, $P = 0$, $S < 0$.

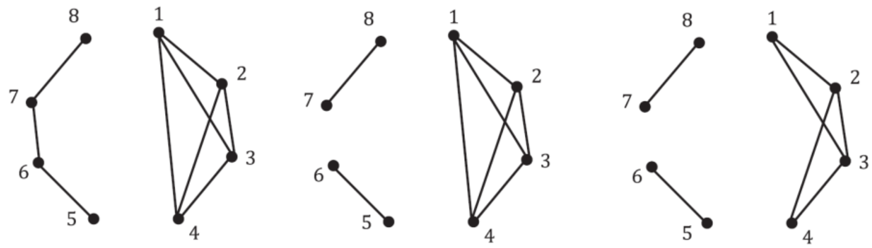
In Assumption 2.3.1, $P = 0$ implies when the punishment is triggered, each player gains nothing (i.e., gains the payoff $P = 0$) from the play of the game, so that the interaction between the players is closed. In contrast, if the stage game is played following the players' behavior patterns (a^1, a^2), then each player gains a positive utility via the continual interaction between them.

Under Assumption 2.3.1, Huang and Inohara (2015-a) proposes a group-separation model by means of the two-player repeated prisoners' dilemma game. In the model, each player is assumed to take a distinct fixed behavior pattern (i.e., a distinct fixed mixed strategy). Any two players (i.e., any pair of the players) engage in the two-player repeated prisoners' dilemma, and thereby the interval of the discount rate of the SPE varies with the pair of the players. This outcome also suggests for each player, the set of the continual interactions and the set of the closed interactions possibly respectively changes with the common discount rate. Define that a set of players form a group if any player has at least a continual interaction with the other players of the set, whereas any two players who are respectively inside and outside the set only can have a closed interaction. Then, according to Huang and Inohara (2015-a), the following numerical example exhibits the players form different groups with the variation of the common discount rate.

Example 2.3.1.
The set of players is given as $N = \{1,2,3,4,5,6,7,8\}$. *For each player* $i \in N$, *the behavior pattern* $a^i = (\delta^i, 1-\delta^i)$ *is abbreviated as* δ^i *for simplicity. The mixed actions of all players are given as* $\delta^1 = 0.99$, $\delta^2 = 0.93$, $\delta^3 = 0.86$, $\delta^4 = 0.8$, $\delta^5 = 0.6$, $\delta^6 = 0.5$, $\delta^7 = 0.4$ *and* $\delta^8 = 0.34$. *The payoff matrix of Table 2.3.1 is given as* $T = 3$, $R = 2$, $P = 0$ *and* $S = -1$. *For each* $i \in N$ *and each* $j \in N - \{i\}$, i *and* j *play the two-player repeated prisoners' dilemma game with the stage game Table 2.3.1. The strategy profile chosen by* i *and* j *is denoted by* $\sigma_m^{i,j}$. *The stage-game payoff is given by* $F^i(\delta^i, \delta^j) = 3\delta^j - \delta^i$, *and the lower bound of SPE between* i *and* j *is* $max\{\theta(\delta^i, \delta^j), \theta(\delta^j, \delta^i)\}$.

In Example 2.3.1, it is clear that given a specific common discount rate $\rho \in (0,1)$, some pairs of player can choose a SPE strategy profile to play the game, while the strategy profile of the other pairs cannot form a SPE. In such a case, following Assumption 2.3.1, it turns out that some pairs of the players gain positive utilities from their continual interaction, meanwhile the interaction of the other pairs is closed and thereby provides no payoffs. To illustrate this outcome further, we give three common discount rates $\rho = 0.440$, $\rho = 0.414$ and $\rho = 0.410$, and exhibit players' strategy profiles in the following Figure.

Fig.2.3.1 shows the players' strategy profile could generate multiple groups of the players: The players form two groups under $\rho = 0.440$, while three groups emerge when $\rho = 0.414$ and $\rho = 0.410$. In particular, as $\rho = 0.414$, the group-separation generates three complete graphs (refer to Chartrand and Zhang (2012, page 19) for the definition of the complete graph).



$\rho = 0.440$ (A) $\rho = 0.414$ (B) $\rho = 0.410$ (C)

Fig. 2.3.1: Panel (A), (B) and (C) plot the players' strategy profiles under the common discount rate $\rho = 0.440$, $\rho = 0.414$ and $\rho = 0.410$ respectively. In each panel, the solid line between two players expresses that their strategy profile forms a SPE. The non-SPE strategy profiles stand for a closed interaction and thereby they are not displayed in each panel.

2.4 The Dynamic game with an infinite horizon

The dynamic game is a repeated game with a state variable (or multiple state variables). The players' action profile affects the transition of the state variable for each period, and the state evolves over time throughout the play of the repeated game. Mailath and Samuelson (2006, page 178) employs the dynamic game framework to formulate an extraction game of fish resources. In the game, the level of the fish resources is defined to be a state variable. Each player gains utility by extracting fish resources and the players' actions dynamically manipulate the evolution of the fish resources.

Before introducing the above simple example explicitly, we first incorporate the state variable into the notation system defined in Section 2.2.

Let S denote the set of states, and each element of S represents a specific state. We let $s \in S$ stand for a state variable of the dynamic game. Recall $A \equiv \prod_i A^i$ (refer to Section 2.2). The stage game payoff function of player i is defined to be $F^i: A \times S \to \mathbb{R}$. The evolution of the state variable is described by a transition function $f: A \times S \to S$. Different from the repeated game, it is assumed that each player can condition its action on the state variable rather than the history of the play. Hence we do not give definition of the history in the dynamic game.

Let us go back to the fish-resource-extraction game stated above. Suppose two players, named player 1 and player 2, extract fish resources from a common pool. In each period t, the pool contains fish resources of level $s(t) \in \mathbb{R}_+$ where $\mathbb{R}_+$ stands for the set of positive real numbers. In period t, player $i \in \{1,2\}$ extracts $a^i(t) \geq 0$ units of the fish from the pool, and its utility is measured by the function $F^i(a^i(t)) = \ln(a^i(t))$. For each period, after the extraction, the remaining stock of fish doubles before the next period. In this context, the transition function can be given by $s(t+1) = f\big(s(t), a^1(t), a^2(t)\big) = 2\big(s(t) - a^1(t) - a^2(t)\big)$. According to the above statements, the maximization of player i's long-run utility is described as

$$\max_{\left(a^i(t)\right)_{t=0}^{\infty}} \left\{(1-\rho)\sum_{t=0}^{\infty} \rho^t \ln(a^i(t))\right\} \tag{2.4.1}$$

subject to the dynamic rule

$$s(t+1) = f\big(s(t), a^1(t), a^2(t)\big) = 2\big(s(t) - a^1(t) - a^2(t)\big). \tag{2.4.2}$$

As in Mailath and Samuelson (2006, page 178), under the symmetry, suppose the solution to (2.4.1) is $a^1(s) = a^2(s) = a(s)$ where $s = s(t)$ for each $t = 1,2, \ldots$, and $a(s)$ suggests that each firm chooses its action dependent on the state variable s in the equilibrium. Given an initial state $s^0 = s(0)$, the equilibrium long-run profit of firm i is given by

$$V(s^0) = (1-\rho)\sum_{t=0}^{\infty} \rho^t \ln\Big(a\big(s(t)\big)\Big) \tag{2.4.3}$$

where $V(s^0)$ is the value function that measures the equilibrium long-run profits.

Using the above value function, the Bellman equation of the dynamic game can be given as

$$a(s) = \underset{\tilde{a} \in A^i}{\operatorname{argmax}} \left\{(1-\rho)\ln(\tilde{a}) + \rho V\Big(2\big(s - \tilde{a} - a(s)\big)\Big)\right\}. \tag{2.4.4}$$

where $\tilde{a}$ stands for an arbitrary amount of extraction and $2\big(s - \tilde{a} - a(s)\big)$ expresses the state posterior to the extraction. The first order condition yields

$$\frac{(1-\rho)}{a(s)} = 2\rho V'\Big(2\big(s - 2a(s)\big)\Big). \tag{2.4.5}$$

Guess $a(s) = ks$ where k is the undetermined coefficient. Subsuming $a(s) = ks$ into the dynamic rule (2.4.2) yields $s(t) = 2(1-2k)s(t-1)$. This further implies $s(t) = 2(1-2k)s(t-1)$, $s(t-1) = 2(1-2k)s(t-2)$,…, $s(3) = 2(1-2k)s(2)$, $s(2) = 2(1-2k)s(1)$, $s(1) = 2(1-2k)s(0)$. Let $s(0) = s$. Then we have $s(t) = (2(1-2k))^t s$. Since $s = s(0) = s^0$, (2.4.3) can be rewritten as

$$V(s) = (1-\rho)\sum_{t=0}^{\infty} \rho^t \ln\Big(k\big(2(1-2k)\big)^t s\Big). \tag{2.4.7}$$

Differentiating both sides of (2.4.7) with respect to s gives $V'(s) = 1/s$. Plugging $V'(s) = 1/s$ and $a(s) = ks$ into (2.4.5), and solve the resulting equation for k yields $k = (1-\rho)/(2-\rho)$. Consequently the equilibrium action is characterized as

$$a(s) = \frac{(1-\rho)}{(2-\rho)} s. \tag{2.4.8}$$

In other words, the dynamic game have an equilibrium strategy profile σ_d in which each firm chooses the action $a(s(t)) = ((1-\rho)/(2-\rho))s(t)$ for each period t. In the context of dynamic game, σ_d is called stationary Markov equilibrium (Mailath and Samuelson (2006, Definition 5.5.2 and page 180)).

2.5 Summary

This chapter introduces the structure of the repeated game and the dynamic game. In both games, the stage game is played for an infinite number of periods. The two games differ in that: in the former one, the players play the stage game repeatedly and the action profile directly results in utilities for each period, whereas in the latter one, the action profile also manipulates the evolution of the state for each period, which influences the players' subsequent utilities and actions.

The extraction game of fish resources stated in Section 2.4 suggests the dynamic game framework can be used to formulate economy with a stock externality (or multiple stock externalities). In particular, Section 2.4 also characterizes an analytic equilibrium due to the quite simple structure of the game. Nevertheless, the characterization of analytic equilibria is usually impossible in dynamic games, because it is rather hard to guess each player's equilibrium action explicitly (refer to Section 2.4 for $a(s) = ks$) and the Bellman equation does not always behaves perfectly in discrete time. To overcome these technical difficulties, game theorists and economists usually employ the dynamic game framework with a continuous time horizon to formulate the stock-externality issue. Such a dynamic game framework is formally named differential game. The subsequent chapters of this dissertation will restrict attention to refinement and application of the differential game framework.

Chapter 3.

The differential game with an infinite time horizon and techniques for the optimal control problem

3.1 Introduction

Dockner et al. (2000, Chapter 3) presents a single-player differential game with a finite time horizon. In order to simplify the interpretation of the general differential game framework involving multiple players stated later, in this chapter we first refine the fundamental model of the single-player differential game over an infinite time horizon. We precisely show the infinite time horizon can also be recursively split into an infinite number of the instantaneous moments. Making use of the recursive structure, we further exhibit that the optimization of the single-player differential game can be solved by characterizing the optimal action for each instantaneous moment. On the basis of the results, we clarify the process how an optimal control path can be characterized by means of the Hamilton-Jacobi-Bellman equation (HJB equation). The HJB equation is a partial differential equation involves a value function and control variables on the right-hand side. The above process implies the solution to the right-hand side of the HJB equation forms an optimal strategy of the whole differential game with the infinite time horizon. In other words, the first order conditions of the HJB equation generates sufficient conditions for the characterization of the optimal strategy (see Section 3.3 for Theorem 3.3.1). Furthermore, the value function of the HJB equation also equals the maximized long-run utility resulting from the optimal control path. In the practical formulation of, e.g., the dynamic public goods provision issue and the transboundary pollution control issue (see Chapter 6 and 7 of the present thesis), etc., the above process also makes it possible to confirm whether a given HJB equation can be correctly transformed into the maximization of the long-run utility of the differential game over an infinite time horizon.

Additional to HJB equation, we also verify how Pontryagin's maximum principle works in the single-player differential game over an infinite time horizon. We use costate variable to construct the current-value Hamiltonian function. By the use of the current-value Hamiltonian function, we clarify the conditions under which the optimal control path can be characterized from Pontryagin's maximum principle. In the practical application of the differential game framework, Pontryagin's maximum principle are generally adopted to confirm the optimal strategy derived by the HJB equation.

3.2 The fundamental model

Assume the differential game is defined over the infinite time horizon $[0, \infty)$. A player takes actions for each time point $t \in [0, \infty)$, thereby influencing the evolution of the state as well as the benefit over time.

Let $S \subset \mathbb{R}^+$ stand for the space of the states, and let $s(t) \in S$ express the state of the game at the time point t. Analogue to the state transition function (2.4.2), we assume the evolution of the state is described by the ordinary differential equation:

$$\dot{s}(t) = f\big(s(t), u(t)\big) \tag{3.2.1}$$

with the initial state

$$s(0) = s^0 \in S \tag{3.2.2}$$

where $u(t) \in \mathbb{R}$ is the actions chosen by the player at the time point t. Supposing the current state equals s, the set of all actions that are feasible at time point t is given by $U(s(t)) \subseteq \mathbb{R}$. This suggests that during the play of the game, the choice of the actions is subject to constraint

$$u(t) \in U\big(s(t)\big) \tag{3.2.3}$$

for each time point $t \in [0, \infty)$. The function $f(.)$ given in (3.2.1) is defined on the set $\Omega = \{(s, u) |\ s \in S, u \in U(s)\}$. Equation (3.2.1) is the dynamic rule (i.e., the transition function of the state) that describes how the current state $s(t)$ and the player's action at the time point t influence the rate of the state at t.

The goal of each player is to choose the control path $u: [0, \infty) \to \mathbb{R}$ to maximize the long-run benefit. More precisely, each player i's long-run benefit is described by the following definite integral over the infinite time horizon

$$J\big(u(.)\big) = \int_0^{\infty} e^{-\rho t} F(s(t), u(t))\, dt \tag{3.2.4}$$

where $\rho > 0$ denotes the common discount rate. Particularly, (3.2.4) and (2.4.1) sum up stage-game benefits over a continuous time horizon and a discrete time horizon respectively. In (3.2.4), the optimal control path $u(.)$ represents the sequence $u(t)_0^\infty$ where $u(t)$ stands for the optimal action that the players choose at the time point t. In other words, an optimal control path is a sequence of optimal actions. See (3.2.4) again. The lower ρ makes the players more concerned with the outcomes in the future and thereby it is also referred to as impatience rate on the future play. Depending on the context, the objective function $F(.)$ is called utility function (i.e., stage game utility function), profit function or benefit function. In any case, F maps the set Ω given above into the real numbers. Note that $F(.)$ is the rate at which profits or utility flow so that it has the dimension *profit per unit of time*, or *utility per unit of time*, respectively.

A basic optimal control problem consists of maximizing the long-run utility (3.2.4) over all control paths $u(.)$ that satisfy (3.2.3) while the state evolves subject to the dynamic rule (3.2.1) and the initial state (3.2.2).

It is well known that solutions to (3.2.1) and (3.2.2) need not exist or may be nonunqiue, and the integral in (3.2.4) may not be defined unless the right-hand side of

(3.2.1) and the integrand in (3.2.4) behave sufficiently well. In order to solve our control problem, it is needed to give the minimal requirement to restrict the set of control paths $u(.)$ and well defined long-run utility $J\big(u(.)\big)$. See the following definition.

Definition 3.2.1
A control path $u: [0,\infty) \to \mathbb{R}^n$ is feasible for the above optimal control problem if the initial value problem (3.2.1)-(3.2.2) has a unique absolutely continuous solution $s(.)$ such that the constraint $s(t) \in S$ and $u(t) \in U(s(t))$ hold for all t and the integral in (3.2.4) is well defined. The control path $u(.)$ is optimal if it is feasible and if the inequality $J\big(u(.)\big) \geq J\big(\tilde{u}(.)\big)$ holds for all feasible control path $\tilde{u}(.)$.

It is clear that the optimal control problem possibly have none, one, multiple or infinitely many feasible control paths and there also exist none, one, multiple or infinitely many optimal control paths for the player to choose in the game. In the next section, we introduce the method to derive the optimal control path of the game.

3.3 The Hamilton-Jacobi-Bellman equation

The Hamilton-Jacobi-Bellman (HJB) equation is one of the main techniques to characterize the optimal control path $u(.)$ that maximizes the long-run benefit (3.2.4). Dockner et al. (2000) deals with the HJB equation of the differential game over a finite time horizon. In this section, we refine the HJB equation and precisely show how an optimal control path can be characterized by the HJB equation over an infinite time horizon.

The HJB equation is based on the principles of embedding and recursion, which is analogue to the repeated game and the dynamic game stated in Chapter 2. To explain the recursion of the HJB equation, consider a differential game that starts at time 0 and the initial state is given as s^0. Let the notation $P(s^0, 0)$ denote the problem, and accordingly, let $P(s,t)$ denote the problem that the differential game starts at time t and the current state is s. According to the principle of embedding, we need to solve not only the present problem $P(s^0, 0)$ but also the entire family of the problem $\{P(s,t) | s \in S, t \in [0,\infty)\}$. Using the above notation, the maximization of the long-run benefit of the problem $P(s,t)$ is described as

$$\max_{u(y)} \int_t^\infty e^{-\rho(y-t)} F(s(y), u(y))\, dy \tag{3.3.1}$$

subject to

$\dot{s}(y) = f\big(s(y), u(y)\big)$
$s(t) = s$
$u(y) \in U\big(s(y)\big).$

Therefore the problem $P(s^0, 0)$ is embedded in the family $\{P(s,t) | s \in S, t \in [0,\infty)\}$.

Combing the principle of embedding with the principle of recursion yields the very powerful HJB equation. Due to the recursive form of the HJB equation, the entire maximization of the long-run benefit (or equivalently the entire long-run optimization) over the infinite time horizon can be converted to the optimization issues over each instantaneous moment. Solving this instantaneous optimization issues, we can obtain the optimal control path over the entire infinite time horizon. To see how the HJB equations work recursively, we next give explicit explanation.

Let $V: S \to \mathbb{R}$ measure the optimal value of the long-run benefit of the problem $P(s,t)$. In other words, we have

$$V(s) = \max_{u(y)} \int_t^{\infty} e^{-\rho(y-t)} F(s(y), u(y))\, dy. \tag{3.3.2}$$

That is, $V(s)$ equals the long-run benefit in the differential game $P(s,t)$ and the game is played along the optimal control path $u(.)$. At the same time, the value function $V(s)$ satisfies the following ordinary differential equations

$$\rho V\big(s(t)\big) = \max_{u(t)} \{F\big(s(t), u(t)\big) + V'\big(s(t)\big) f\big(s(t), u(t)\big)\}. \tag{3.3.3}$$

(3.3.3) is the HJB equation stated in the beginning of the present section. In (3.3.3), we assume $V(s)$ is continuously differentiable.

In the next step, our goal is to solve $P(s,t)$ where $s \in S$ and $t \in [0, \infty)$. That is, the game starts at the time point t and the current state is given as s. Instead of attempting to characterize the entire control path $u(.)$, we focus on the instantaneous moment $[t, t+\Delta)$ where $\Delta > 0$ is very small. Every feasible control path $u: [t, t+\Delta) \to \mathbb{R}$ takes the instantaneous game from the present state $s(t)$ to the state $s(t+\varDelta)$ that takes place in the subsequent instantaneous moment. If each player behave optimally from $t+\varDelta$ onwards, the whole benefit obtained during the interval $[t+\varDelta, \infty)$ is given by $V(s(t+\varDelta))$. This follows from the way in which the value function $V(.)$ is defined. Therefore, at the time point t, the total benefit that each player earns by choosing the optimal control path $u(.)$ until time $t+\Delta$, and behaving optimally thereafter, is equal to

$$\int_t^{t+\Delta} e^{-\rho(y-t)} F(s(y), u(y))\, dy + e^{-\rho\Delta} V\big(s(t+\varDelta)\big). \tag{3.3.4}$$

If the player chooses the control path $u(.)$ over the instantaneous moment $[t, t+\Delta)$, then the above discounted benefit equals the maximal benefit $V(s(t))$. This can be described as

$$V(s(t)) = \max_{u(.)} \left\{ \int_t^{t+\Delta} e^{-\rho(y-t)} F\big(s(y), u(y)\big)\, dy + e^{-\rho\Delta} V\big(s(t+\varDelta)\big) \right\} \tag{3.3.5}$$

subject to the dynamic law (3.2.1).

Subtracting $V(s)$ from both sides of (3.3.5) and dividing the resulting equation by $\varDelta$ gives

$$0 = \max_{u(.)} \left\{ \frac{1}{\varDelta} \int_t^{t+\Delta} e^{-\rho(y-t)} F\big(s(y), u(y)\big)\, dy + \frac{e^{-\rho\Delta} V\big(s(t+\varDelta)\big) - V(s(t))}{\varDelta} \right\}. \tag{3.3.6}$$

Assuming all the functions appearing in (3.3.6) are sufficiently smooth, it is worth examining what happens if $\varDelta$ approaches zero. The mean value theorem implies

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{t}^{t+\Delta} e^{-\rho(y-t)}F\big(s(y),u(y)\big)\,dy = F\big(s(t),u(t)\big).$$

Taking the same limit of the second term on the right side of (3.3.6) results in

$$\lim_{\Delta\to 0}\frac{e^{-\rho\Delta}V\big(s(t+\Delta)\big)-V\big(s(t)\big)}{\Delta} \tag{3.3.7}$$

Since $e^{-\rho\Delta}V\big(s(t+\Delta)\big)-V\big(s(t)\big)\to 0$ as $\Delta\to 0$, using L'Hôpital's rule, we obtain

$$\lim_{\Delta\to 0}\frac{e^{-\rho\Delta}V\big(s(t+\Delta)\big)-V\big(s(t)\big)}{\Delta}$$

$$=\lim_{\Delta\to 0}\frac{\big[e^{-\rho\Delta}V\big(s(t+\Delta)\big)\big]'}{1}$$

$$=\lim_{\Delta\to 0}(-\rho)e^{-\rho\Delta}V\big(s(t+\Delta)\big)+e^{-\rho\Delta}V'\big(s(t+\Delta)\big)\dot{s}(t+\Delta)$$

$$=-\rho V\big(s(t)\big)+V'\big(s(t)\big)\dot{s}(t) \tag{3.3.8}$$

Putting everything together, we see that as $\Delta\to 0$, (3.3.6) becomes

$$0=\max_{u(.)}\big\{F\big(s(t),u(t)\big)-\rho V\big(s(t)\big)+V'\big(s(t)\big)\dot{s}(t)\big\}. \tag{3.3.9}$$

Note the maximization of the long-run benefit has to be carried out with respect to all feasible control paths $u:[0,\infty)\to\mathbb{R}$ and subject to the dynamic rule $\dot{s}(t)=f(s(t),u(t))$, the current state $s(t)=s$, and the decision variable $u(y)\in U(s(y))$ for any $y\in[t,t+\Delta t)$. Since we let Δ approach zero, the decision variable $u(y)=u$ at the time point t. Based on the above facts, (3.3.9) can be rewritten as

$$0=\max_{u(.)}\{F(s,u)-\rho V(s)+V'(s)f(s,u)\,|\,u\in U(s)\}. \tag{3.3.10}$$

(3.3.10) is indeed identical to the HJB equation (3.3.3).

In the next step, we show under the assumption that the value function is bounded, the control variable that maximizes the right side of the HJB equation (3.3.3) forms the optimal control path for the differential game over an infinite time horizon. Moreover, we also point out that the boundedness of the value function does not results in the loss of generality if the object function $F(.)$ is given in the linear quadratic form.

Assumption 3.3.1

The value function $V(.)$ is bounded on the domain S.

Theorem 3.3.1

Let $V:S\to\mathbb{R}$ be a continuously differentiable function which follows Assumption 3.3.1 and satisfies the HJB equation

$$\rho V(s(t))=\max_{u(t)}\{F(s(t),u(t))+V'(s(t))f(s(t),u(t))\}. \tag{3.3.12}$$

for any $s\in S$.

Suppose the initial state is s^0. Pick up a feasible control path $u(.)$ with corresponding state trajectory $s(.)$. If $u(t)$ maximizes the right-hand side of (3.3.12) for any $t\in[0,\infty)$, then $u(.)$ is an optimal control path. Moreover, $V(s)$ is the optimal value of problem $P(x,t)$.

Proof:

Pick up any feasible control path $\tilde{u}(.)$ with corresponding state trajectory $\tilde{s}(.)$. Herein Our goal is to show that $J\big(u(.)\big)\geq J(\tilde{u}(.))$. According to the HJB equation (3.3.12), it holds that

$$\rho V(\tilde{s}(t)) \leq F(\tilde{s}(t), \tilde{u}(t)) + V'(\tilde{s}(t)) f(\tilde{s}(t), \tilde{u}(t)). \tag{3.3.13}$$

Multiplying two sides of the inequality by $e^{-\rho t}$ and using the dynamic rule $\dot{s}(t) = f(s(t), u(t))$, (3.3.13) is rewritten as follows:

$$e^{-\rho t}\rho V(\tilde{s}(t)) \leq e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t)) + e^{-\rho t} V'(\tilde{s}(t))\dot{\tilde{s}}(t)$$

$$\Leftrightarrow \rho e^{-\rho t} V(\tilde{s}(t)) - e^{-\rho t} V'(\tilde{s}(t))\dot{\tilde{s}}(t) \leq e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t))$$

Since

$$\rho e^{-\rho t} V(\tilde{s}(t)) - e^{-\rho t} V'(\tilde{s}(t))\dot{\tilde{s}}(t) = -\frac{d}{dt}\left[e^{-\rho t} V(\tilde{s}(t))\right],$$

We have

$$-\frac{d}{dt}\left[e^{-\rho t} V(\tilde{s}(t))\right] \leq e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t))$$

In the above inequality, integrating both sides over the infinite time horizon yields

$$-\int_0^\infty d\left[e^{-\rho t} V(\tilde{s}(t))\right] \leq \int_0^\infty e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t)) dt$$

$$\Leftrightarrow -\left[\lim_{t\to\infty} e^{-\rho t}\, V(\tilde{s}(t)) - V(\tilde{s}(0))\right] \leq \int_0^\infty e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t)) dt \tag{3.3.14}$$

Recall the optimal control path $u(.)$. It is clear that

$$\rho V(s(t)) = F(s(t), u(t)) + V'(s(t))\dot{s}(t).$$

Analogue to the above case, multiplying the two sides by $e^{-\rho t}$ gives

$$e^{-\rho t}\rho V(s(t)) = e^{-\rho t} F(s(t), u(t)) + e^{-\rho t} V'(s(t))\dot{s}(t)$$

$$\Leftrightarrow e^{-\rho t}\rho V(s(t)) - e^{-\rho t} V'(s(t))\dot{s}(t) = e^{-\rho t} F(s(t), u(t)).$$

$$\Leftrightarrow -\frac{d}{dt}\left[e^{-\rho t} V(s(t))\right] = e^{-\rho t} F(s(t), u(t)).$$

Integrating both side over the infinite time horizon yields

$$-\left[\lim_{t\to\infty} e^{-\rho t}\, V(s(t)) - V(s(0))\right] = \int_0^\infty e^{-\rho t} F(s(t), u(t)) dt = J(u(.)) \tag{3.3.15}$$

Subtracting (3.3.14) from (3.3.15) results in

$$\begin{aligned} J(u(.)) - J(\tilde{u}(.)) &= \int_0^\infty e^{-\rho t} F(s(t), u(t)) dt - \int_0^\infty e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t)) dt \\ &\geq -\lim_{t\to\infty} e^{-\rho t}\, V(s(t)) + V(s(0)) + \lim_{t\to\infty} e^{-\rho t}\, V(\tilde{s}(t)) - V(\tilde{s}(0)) \end{aligned} \tag{3.3.16}$$

Since $V(.)$ is a specific function and $s(0) = \tilde{s}(0) = s^0$, it is clear that $V(s(0)) = V(\tilde{s}(0))$. According to Assumption 3.3.1, due to the boundedness of $V(.)$, suppose $V(.) \in [\underline{v}, \overline{v}]$. Then, we can see

$$\begin{aligned} J(u(.)) - J(\tilde{u}(.)) &= \int_0^\infty e^{-\rho t} F(s(t), u(t)) dt - \int_0^\infty e^{-\rho t} F(\tilde{s}(t), \tilde{u}(t)) dt \\ &\geq \lim_{t\to\infty} e^{-\rho t}\, V(\tilde{s}(t)) - \lim_{t\to\infty} e^{-\rho t}\, V(s(t)) \end{aligned}$$

$$\geq \underline{v} \lim_{t\to\infty} e^{-\rho t} - \overline{v} \lim_{t\to\infty} e^{-\rho t} = 0. \tag{3.3.17}$$

Consequently $J\big(u(.)\big) - J\big(\tilde{u}(.)\big) \geq 0$. Moreover, $J\big(u(.)\big) = V\big(s(0)\big)$ and hence $V\big(s(0)\big)$ measures the optimal long-run benefit. Following the above process of proof, it is straightforward to show that $V\big(s(t)\big)$ measures optimal long-run benefit of the problem $P(x,t)$ (refer to (3.3.1)). □

Assumption 3.3.1 imposes boundedness on the value function $V(.)$, and accordingly, Theorem 3.3.1 shows given an arbitrary initial state s^0,

$$V(s^0) = \max_{\tilde{u}(y)} \int_t^\infty e^{-\rho(y-t)} F(\tilde{s}(y), \tilde{u}(y))\, dy = \int_t^\infty e^{-\rho(y-t)} F(s(y), u(y))\, dy \tag{3.3.18}$$

where $s^0 = \tilde{s}(t) = s(t)$ and $u(y)_{y=0}^{\infty} \in argmax_{\tilde{u}(y)}\{\int_t^\infty e^{-\rho(y-t)} F(\tilde{s}(y), \tilde{u}(y))\, dy\}$. Throughout this dissertation, the analyses are restricted to the optimal control path $u(.)$. In the linear quadratic (LQ) context, the objective function $F(.,.)$ is assumed to be concave in the argument u, and thereby we restrict ourselves to the interior optimal control path $u(.)$ such that the marginal utility $\partial F(s(t), u(t))/\partial u \geq 0$ for any $t \in [0, \infty)$. Along the interior optimal control path $u(.)$, it is straightforward to show that $V(s)$ is bounded for any feasible initial state $s^0 \in S$ and thereby Assumption 3.3.1 does not result in loss of generality. In particular, along any interior optimal control path $u(.)$, it holds that

$$\lim_{t\to\infty} e^{-\rho t} V(s(t)) = 0. \tag{3.3.19}$$

(3.3.19) is called terminal condition. In the subsequent chapters (refer to Chapter 4~Chapter 7), the characterization of the optimal control path is based on the terminal condition.

3.4 Pontryagin's maximum principle

Section 3.3 states a sufficient condition for the optimal control problem. That is, along the control path, if the action always solves the right-hand side of the HJB equation, then the player's long-run benefit will be maximized by the flow of the actions. In addition to the HJB equation, Pontryagin's maximum principle is another technique that is generally employed to examine the optimal control path of the differential game. Because Pontryagin's maximum principle describes a necessary condition for the optimal control problem, any interior optimal control path $u(.)$ that is derived from the HJB equation must satisfy Pontryagin's maximum principle, but this statement does not hold in the reverse direction. Nevertheless, in this section, our goal is to clarify the situation in which Pontryagin's maximum principle can also be a sufficient condition for the optimal control problem over the infinite time horizon.

Let us start the analysis by defining a real valued function H as

$$H(s, u, \lambda) = F(s, u) + \lambda f(s, u). \tag{3.4.1}$$

The domain of $H(.,.,.)$ is given as $\{(s,u,\lambda)|s \in S,\ u \in U(s),\ \lambda \in \mathbb{R}\}$. The function H is called the current-value Hamiltonian function and plays a prominent role in Pontryagin's maximum principle. The variable λ is called the (current-value) costate variable associated with the state variable s. The maximized Hamiltonian function $H^*: S \times \mathbb{R} \to \mathbb{R}$ is given by

$$H^*(s,\lambda) = \max\{H(s,u,\lambda)|u \in U(s)\}. \quad (3.4.2)$$

Turn back to the HJB equation (3.3.3). It follows from Theorem 3.3.1 that the control variables which maximize the right-hand side of (3.3.3) generates an optimal control path of the differential game. Let $\lambda = V_s(s)$. Using the definition of the maximized Hamiltonian function, the HJB equation can be rewritten as

$$\rho V(s(t)) = H^*(s(t), V_s(s(t))) \quad (3.4.3)$$

Since (3.4.3) holds for any $t \in [0,\infty)$, differentiating (3.4.3) with respect to s gives

$$\rho V_s(s) = H_s^*(s, V_s(s)) + V_{ss}(s) H_\lambda^*\big(s, V_s(s)\big). \quad (3.4.4)$$

Turn to (3.4.2). Suppose u solves the right-hand side of (3.4.2). That is, the control path u maximizes the right-hand side of (3.4.2). Then the envelope theorem gives

$$H_\lambda^*\big(s, V_s(s)\big) = H_\lambda\big(s, u, V_s(s)\big). \quad (3.4.5)$$

Plugging (3.4.5) into (3.4.4) results in

$$\rho V_s(s) = H_s^*(s, V_s(s)) + V_{ss}(s) H_\lambda\big(s, u, V_s(s)\big). \quad (3.4.6)$$

On the other hand, differentiating (3.4.1) with respect to the state variable λ gives

$$H_\lambda(s,u,\lambda) = f(s,u). \quad (3.4.7)$$

Plugging (3.4.7) into (3.4.6), we have

$$\rho V_s(s) = H_s^*(s, V_s(s)) + V_{ss}(s) f(s,u). \quad (3.4.8)$$

The equation $\lambda = V_s(s)$ implies $\lambda(t) = V_s(s(t))$. Differentiating the resulting equation with respect to t yields

$$\dot{\lambda}(t) = \frac{d}{dt} V_s\big(s(t)\big) = V_{ss}\big(s(t)\big)\dot{s}(t) \quad (3.4.9)$$
$$= V_{ss}\big(s(t)\big) f\big(s(t), u(t)\big)$$

Plugging $\lambda = V_s(s)$ and (3.4.9) into (3.4.8), and rearranging the resulting equation gives

$$\dot{\lambda} = \rho\lambda - H_s^*\big(s, V_s(s)\big). \quad (3.4.10)$$

(3.4.10) is called the adjoint equation or the costate equation of the optimal control problem. Applying the envelope theorem to (3.4.2) gives

$$H_s^*(s,\lambda) = H_s(s, u, V_s(s)). \quad (3.4.11)$$

Accordingly, the adjoint condition can also be expressed as follows:

$$\dot{\lambda} = \rho\lambda - H_s(s, u, V_s(s)). \quad (3.4.12)$$

The maximum principle states for every optimal control path $u(.)$, there exists a costate trajectory $\lambda(.)$ such that the maximum condition (3.4.2) and adjoint equation (3.4.10) both hold. In the next step, we clarify the condition for the maximum principle to be a sufficient condition under which an optimal control path can be specified.

Assumption 3.4.1

Let $u(.)$ be an optimal control path with corresponding state trajectory $s(.)$. In contrast, $\tilde{u}(.)$ stands for an arbitrary control path with the corresponding state trajectory $\tilde{s}(.)$. It is assumed that $\lim_{t\to\infty} e^{-\rho t}\lambda(t)[\tilde{s}(t) - s(t)] \geq 0$.

In Assumption 3.4.1, $\lim_{t\to\infty} e^{-\rho t}\lambda(t)[\tilde{s}(t) - s(t)] \geq 0$ is called the transversality condition for Pontryagin's maximum principle. Theorem 3.4.1 shows Pontryagin's maximum principle can be the sufficient condition for the optimal control problem if the transversality condition is satisfied.

Theorem 3.4.1

Consider the optimal control problem of Section 3.3 and use the Hamiltonian function H together with the maximized Hamiltonian function H^* defined in Section 3.4. Assume the state space S is a convex set. Let $u(.)$ be a feasible control path with corresponding state trajectory $s(.)$. If there exists an absolutely continuous function $\lambda: [0, \infty] \to \mathbb{R}$ such that Assumption 3.4.1 is followed and the maximum condition

$$H\big(s(t), u(t), \lambda(t)\big) = H^*\big(s(t), \lambda(t)\big), \tag{3.4.13}$$

the adjoint equation

$$\dot{\lambda}(t) = \rho\lambda(t) - H_s^*\big(s(t), \lambda(t)\big) = \rho\lambda - H_s\big(s(t), u(t), \lambda(t)\big) \tag{3.4.14}$$

are satisfied, then $u(.)$ is an optimal path.

Proof:

Pick up an arbitrary feasible control path $\tilde{u}(.)$ with corresponding state trajectory $\tilde{s}(.)$. Our goal is to show $J\big(u(.)\big) \geq J(\tilde{u}(.))$. Subtracting $J(\tilde{u}(.))$ from $J\big(u(.)\big)$ gives

$$J\big(u(.)\big) - J\big(\tilde{u}(.)\big) = \int_0^\infty e^{-\rho t} F\big(s(t), u(t)\big)\, dt - \int_0^\infty e^{-\rho t} F\big(\tilde{s}(t), \tilde{u}(t)\big)\, dt \tag{3.4.16}$$

Recall (3.4.1). It holds that $H(s(t), u(t), \lambda(t)) = F(s(t), u(t)) + \lambda\dot{s}(t)$ and $H(\tilde{s}(t), \tilde{u}(t), \lambda(t)) = F(\tilde{s}(t), \tilde{u}(t)) + \lambda\dot{\tilde{s}}(t)$. Thus (3.4.16) can be extended as

$$J\big(u(.)\big) - J\big(\tilde{u}(.)\big)$$
$$= \int_0^\infty e^{-\rho t} [H\big(s(t), u(t), \lambda(t)\big) - \lambda\dot{s}(t)]\, dt - \int_0^\infty e^{-\rho t} [H\big(\tilde{s}(t), \tilde{u}(t), \lambda(t)\big) - \lambda\dot{\tilde{s}}(t)]\, dt \tag{3.4.17}$$

It follows from (3.4.2) that $H^*(s(t), \lambda(t)) = H(s(t), u(t), \lambda(t))$ (note it has been supposed that u maximizes the right-hand side of (3.4.2)) and $H^*(\tilde{s}(t), \lambda(t)) \geq H(\tilde{s}(t), \tilde{u}(t), \lambda(t))$. Thus (3.4.16) can be further rewirtten as

$$J\big(u(.)\big) - J\big(\tilde{u}(.)\big)$$
$$\geq \int_0^\infty e^{-\rho t} [H^*(s(t), \lambda(t)) - \lambda\dot{s}(t)]\, dt - \int_0^\infty e^{-\rho t} [H^*(\tilde{s}(t), \lambda(t)) - \lambda\dot{\tilde{s}}(t)]\, dt \tag{3.4.18}$$

Since $H^*(.,.)$ is concave, following Taylor approximation, it holds that

$$H_s^*(s(t), \lambda(t))\, [s(t) - \tilde{s}(t)] \leq H^*\big(s(t), \lambda(t)\big) - H^*\big(\tilde{s}(t), \lambda(t)\big) \tag{3.4.19}$$

Due to the adjoint condition (3.4.14), (3.4.19) can be rewirtten as

$$H^*\big(s(t), \lambda(t)\big) - H^*\big(\tilde{s}(t), \lambda(t)\big) \geq \big[\rho\lambda(t) - \dot{\lambda}(t)\big]\, [s(t) - \tilde{s}(t)] \tag{3.4.20}$$

Using (3.4.20) to (3.4.18) yields

$$J\big(u(.)\big) - J\big(\tilde{u}(.)\big) \geq \int_0^\infty e^{-\rho t} [H^*\big(s(t), \lambda(t)\big) - H^*\big(\tilde{s}(t), \lambda(t)\big) - \lambda\dot{s}(t) + \lambda\dot{\tilde{s}}(t)]\, dt$$

$$\geq \int_0^\infty \left\{ e^{-\rho t} \left[\rho\lambda(t) - \dot{\lambda}(t) \right] [s(t) - \tilde{s}(t)] - e^{-\rho t} \lambda(t) \left[\dot{s}(t) - \dot{\tilde{s}}(t) \right] \right\} dt$$

$$= \int_0^\infty \left\{ \left[e^{-\rho t} \rho\lambda(t) - e^{-\rho t} \dot{\lambda}(t) \right] [s(t) - \tilde{s}(t)] - e^{-\rho t} \lambda(t) \left[\dot{s}(t) - \dot{\tilde{s}}(t) \right] \right\} dt$$

$$= \int_0^\infty \left\{ - \left[-\rho e^{-\rho t} \lambda(t) + e^{-\rho t} \dot{\lambda}(t) \right] [s(t) - \tilde{s}(t)] - e^{-\rho t} \lambda(t) \left[\dot{s}(t) - \dot{\tilde{s}}(t) \right] \right\} dt$$

$$= \int_0^\infty \left\{ \left[-\rho e^{-\rho t} \lambda(t) + e^{-\rho t} \dot{\lambda}(t) \right] [\tilde{s}(t) - s(t)] + e^{-\rho t} \lambda(t) \left[\dot{\tilde{s}}(t) - \dot{s}(t) \right] \right\} dt$$

$$= \int_0^\infty \frac{d}{dt} \{ e^{-\rho t} \lambda(t) [\tilde{s}(t) - s(t)] \} dt$$

$$= \int_0^\infty d\{ e^{-\rho t} \lambda(t) [\tilde{s}(t) - s(t)] \}$$

$$= \lim_{t \to \infty} e^{-\rho t} \lambda(t) [\tilde{s}(t) - s(t)] \qquad (3.4.21)$$

Assumption 3.4.1 states $\lim_{t\to\infty} e^{-\rho t}\lambda(t)[\tilde{s}(t) - s(t)] \geq 0$. Therefore $J(u(.)) \geq J(\tilde{u}(.))$. □

3.5 Summary

In this chapter, we refine the framework of a single-player differential game over an infinite time horizon, and present two techniques, the HJB equation and Pontryagin's maximum principle, for derivation of the optimal solution to the differential game. In particular, analogue to the repeated game and the dynamic game examined in Chapter 2, the infinite time horizon can be recursively split into an infinite number of instantaneous moments, so that the optimization of the differential game can be solved by optimizing the play of the game for each instantaneous moment. On the basis of this structure, we show that the HJB equation works recursively in solving the optimal control problem of the game. Given specific the terminal condition or the transversality condition, we also show the control path that solves the HJB equation or satisfies Pontryagin's maximum principle can form an optimal solution to the differential game.

In the next Chapter, we apply the above two techniques to multiple-player differential game with an infinite time horizon, and show how the techniques work in characterizing Markov perfect equilibria of the game.

Chapter 4.

Markovian equilibria in differential games with simultaneous play

4.1 Introduction

Chapter 3 generalizes a single-player differential game framework on an infinite time horizon, and precisely shows the optimization of the game can be solved by means of either the HJB equation or the Pontryagin's maximum principle.

In this Chapter, we refine the framework of the multiple-player differential game on an infinite time horizon. In the game, each player is assumed to optimize its control on the basis of the belief that the other players stick to their optimal strategies during the play of the game. All players' actions dynamically influence the level of the external state of the next instantaneous moment. Assuming the information on external state is always available to each player throughout the play of the game, we are concerned with the Markovian strategy profile in which each player bases its action on the level of the external state for any instantaneous moment. The optimal Markovian strategy profile constitutes a Markovian Nash equilibrium. We show Theorem 3.4.1 also applies to the multiple-player differential game framework on an infinite time horizon. That is, for each player, the solution to the right-hand side of the HJB equation forms a Markovian Nash equilibrium strategy.

Additional to the above relation between the Markovian Nash equilibrium and HJB equation, we also verify the conditions for a Markovian Nash equilibrium to satisfy two properties: time consistency and subgame perfectness in the infinite-time-horizon differential game. Time consistency is satisfied if the players' optimal control paths form a Markovian Nash equilibrium for the subgame starting from any instantaneous moment (see Definition 4.4.1). Time consistency is the minimal requirement for the credibility of a Markovian Nash equilibrium in the sense that if a player has an incentive to deviate in the subgame starting from a certain time point (i.e., a certain instantaneous moment), then the opponents will detect the incentive of the deviation in the first place, and thereby choose other optimal strategies to form a new Markovian Nash equilibrium. Different from time consistency, subgame perfectness requires that the players' optimal control paths form a Markovian Nash equilibrium at the subgame starting from any instantaneous moment and any feasible external state. The Markovian Nash equilibrium that satisfies subgame perfectness is called Markov perfect equilibrium (MPE). It is clear that a MPE is a SPE of the differential game. Finally, we clarify the conditions under which a MPE can be characterized by Pontryagin's maximum principle (i.e., Theorem 3.4.1) in the multiple-player differential game on an infinite time horizon. In particular, in the practical application of the multiple-player differential game framework (see Chapter 5~7), we employ the HJB equation to derive infinitely many MPEs and examine each MPE precisely.

This Chapter is arranged as follows: In Section 4.2, we define Markovian Nash equilibria and present the relation between the optimal control paths and the Markovian Nash equilibria. In Section 4.3, we give the sufficient conditions for a Markovian strategy to form a Markovian Nash equilibrium. This section employs the techniques of the HJB equation and the Pontryagin's maximum principle presented by Chapter 3. In Section 4.4, we discuss two important properties of Markovian Nash equilibria: time consistency and subgame perfectness. We show that any Markovian Nash equilibrium is time consistent. The Markovian Nash equilibrium that satisfies subgame perfectness is called MPE. Moreover, we show a MPE can be characterized via the HJB equation, which is analogous to Theorem 3.4.1 given by Chapter 3.

4.2 Markov equilibria with simultaneous play

Consider a differential game over an infinite time horizon $[0, \infty)$. The state at each time point $t \in [0, \infty)$ is described by $s(t) \in S$ where $S \subseteq \mathbb{R}$ expresses the state space of the game. The initial state is given by s^0. The game involves n players $i = 1,2, \dots, n$. We use superscript to denote each specific player. At each time point $t \in [0, \infty)$, each player $i \in \{1,2, \dots, n\}$ chooses a control variable $u^i(t)$ (i.e., an action) from its set of feasible controls $U^i(s(t), u^{-i}(t)) \subseteq \mathbb{R}$. In general, the set of control variables depends on the current state $s(t)$, and the vector

$$u^{-i}(t) = \left(u^1(t), u^2(t), \dots, u^{i-1}(t), u^{i+1}(t), \dots, u^n(t)\right). \tag{4.2.1}$$

The state variable evolves subject to the following dynamic rule

$$\dot{s}(t) = f\big(s(t), u^1(t), u^2(t), \dots, u^n(t)\big), \qquad s(0) = s^0, \tag{4.2.2}$$

where the function f are defined on the set

$$\Omega = \left\{(s, u^1, u^2, \dots, u^n) \mid s \in S, t \in [0, \infty), u^i \in \left(s, u^{-i}\right), i = 1,2, \dots, n\right\} \tag{4.2.3}$$

and $f{:}\, \Omega \to \mathbb{R}$. Each player $i \in \{1,2, \dots, n\}$ seeks to maximize its long-run utility by choosing an optimal control path. The maximization of the long-run utility is denoted by

$$J^i\left(u^i(.)\right) = \int_0^\infty e^{-\rho t} F^i(s(t), u^1(t), u^2(t), \dots, u^n(t)) dt \tag{4.2.4}$$

where $F^i{:}\, \Omega \to \mathbb{R}$ measures player i's instantaneous utility (i.e., $F^i(.)$ is stage game utility function).

A Nash equilibrium is an n-tuple of strategies $(\phi^1, \phi^2, \dots, \phi^n)$. In the Nash equilibrium, given the opponents' equilibrium strategies, it is unprofitable for any player to change its own strategy. Employing the above notational convention, player i's decision problem can be rewritten as

$$\max \int_0^\infty e^{-\rho t} F^i(s(t), \phi^1(t), \phi^2(t), \dots, \phi^{i-1}(t), u^i(t), \phi^{i+1}(t), \dots, \phi^n(t)) dt \tag{4.2.5}$$

subject to

$u^i(t) \in U^i(s(t), \phi^1(t), \phi^2(t), \dots, \phi^{i-1}(t), \phi^{i+1}(t), \dots, \phi^n(t)),$

$$\dot{s}(t) = f\left(s(t), \phi^1(t), \phi^2(t), \dots, \phi^{i-1}(t), u^i(t), \phi^{i+1}(t), \dots, \phi^n(t)\right), \quad s(0) = s^0.$$

Let $\phi^{-i}(t) = \left(\phi^1(t), \phi^2(t), \dots, \phi^{i-1}(t), \phi^{i+1}(t), \dots, \phi^n(t)\right)$. Then (4.2.5) can be simplified as

$$\max_{u^i(t)\in U^i\left(s(t),\phi^{-i}(t)\right)} \int_0^\infty e^{-\rho t} F^i\left(s(t), u^i(t), \phi^{-i}(t)\right) dt \tag{4.2.6}$$

subject to

$$\dot{s}(t) = f\left(s(t), u^i(t), \phi^{-i}(t)\right), \quad s(0) = s^0.$$

In the differential game described by (4.2.5), a Markovian Nash equilibrium is defined as follows:

Definition 4.2.1

Define a function $\phi^i: S \to \mathbb{R}$ *for each* $i \in \{1,2,\dots,n\}$. *The* n*-tuple* $(\phi^1, \phi^2, \dots, \phi^n)$ *is called a Markovian Nash equilibrium, if an optimal control path* $u^i(.)$ *of the problem (4.2.5) exists and is characterized by the Markovian strategy* $u^i(t) = \phi^i(s(t))$.

Definition 4.2.1 implies seeking a Markovian Nash equilibrium of an *n*-player differential game is equivalent to the specification of each player's optimal control path of the game. Therefore, the techniques developed by Section 3.3 and Section 3.4 are valid in the characterization of a Markovian Nash equilibrium.

In Definition 4.2.1, replacing the assumption that optimal paths are defined by Markovian strategies by the assumption that the optimal paths are given by open-loop strategies, then we obtain the following definition

Definition 4.2.2

Define a function $\phi^i: [0,\infty) \to \mathbb{R}$ *for each* $i \in \{1,2,\dots,n\}$. *The n-tuple* $(\phi^1, \phi^2, \dots, \phi^n)$ *is called an open-loop equilibrium, if an optimal control path* $u^i(.)$ *of the problem (4.2.5) exists and is characterized by the open-loop strategy* $u^i(t) = \phi^i(t)$.

In Definition 4.2.1, the setup $u^i(t) = \phi^i(s(t))$ shows in order to choose the Markovian Nash equilibrium strategies to play the game, each player needs to be informed about the current state for each instantaneous moment. In contrast, the open-loop Nash equilibrium strategies does not require information concerning the states throughout the play of the game. Generally speaking, the difference between the Markovian Nash equilibrium and the open-loop Nash equilibrium can be summarized as: in the Markovian Nash equilibria, each player dynamically conditions its action (i.e., its control variable) on the external state, and the entire play of the game generates an optimal control path (i.e., $u(.)$) with the state trajectory (i.e., $s(.)$); meanwhile when the game is played along open-loop Nash equilibria, each player is supposed to pre-commit an optimal control path in advance, and then chooses an action for each time point following the optimal control path. In particular, the Markovian Nash equilibrium strategy enables the control variables stick to the control path for the entire duration of the game. In this sense, compared to the open-loop Nash equilibria, the Markovian Nash equilibria receives much more attention in the context of economics because each player is allowed to reoptimizie at each time point after observing the state variables.

4.3 Equilibrium condition

This section employs the techniques stated by Section 3.3 and Section 3.4 to generalize the conditions for Markovian Nash equilibria of the n-player game defined by Section 4.2.

Theorem 4.3.1
Let $(\phi^1, \phi^2, \dots, \phi^n)$ stand for a given n-tuple of functions $\phi^i: S \to \mathbb{R}$ and make the following assumptions:
(i) there exists a unique continuously solution $s: [0, \infty) \to S$ of the differentiation equation

$$\dot{s}(t) = f\left(s(t), \phi^1(s(t)), \phi^2(s(t)), \dots, \phi^n(s(t))\right), \quad s(0) = s^0, \tag{4.3.1}$$

(ii) for all $i \in \{1,2, \dots, n\}$, there exists a continuously differentiable function $V^i: S \to \mathbb{R}$ such that the HJB equations

$$\rho V^i(s) = \max_{u^i \in U^i\left(s, \phi^{-i}(s)\right)} \left\{F^i\left(s, u^i, \phi^{-i}(s)\right) + V_s^i(s) f\left(s, u^i, \phi^{-i}(s)\right)\right\} \tag{4.3.2}$$

holds true for all $s \in S$.
(iii) The function $V^i(.)$ is bounded on the domain S.
The $u^i(.)$ is an arbitrary control path in which $u^i(t) \in U^i(s(t), \phi^{-i}(s(t)))$ and $u^i(t)$ maximizes the right-hand side of (4.3.2) for any $t \in [0, \infty)$. If $\phi^i(s(t)) = u^i(t)$ for each $i \in \{1,2, \dots, n\}$ and any $t \in [0, \infty)$ then $(\phi^1, \phi^2, \dots, \phi^n)$ is a Markovian Nash equilibrium.
Proof:
According to Definition 3.3.1, in a Markovian Nash equilibrium, the relation between the optimal control rule and equilibrium strategy is given as $u^i(t) = \phi^i(s(t))$. Therefore, it is clear that the proof of Theorem 3.3.1 applies to Theorem 4.3.1.□

Theorem 4.3.1 states for any player, given the opponents' Markovian strategies, the player's Markovian strategy that satisfies the HJB equation (4.3.2) generates an optimal control path. Therefore, the Markovian strategies profile such that each player's strategy follows the claims of Theorem 4.3.1 constitutes a Markovian Nash equilibrium of the game. Next we examines the relation between the Markovian Nash equilibria and Pontryagin's maximum principle.

Theorem 4.3.2
Let $(\phi^1, \phi^2, \dots, \phi^n)$ stand for a given n-tuple of functions $\phi^i: S \to \mathbb{R}$ and assume under the n-tuple, the state variable evolves subject to the dynamic rule (4.3.1). Define for all $i \in \{1,2, \dots, n\}$ the Hamiltonians $H^i: S \times \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ by

$$H^i\left(s, u^i, \phi^{-i}(s), \lambda^i\right) = F^i\left(s, u^i, \phi^{-i}(s)\right) + \lambda^i f^i\left(s, u^i, \phi^{-i}(s)\right) \tag{4.3.3}$$

and the maximized Hamiltonians $H^{i*}: S \times \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ by

$$H^{i*}\left(s, \phi^{-i}(s), \lambda^i\right) = max\left\{H^i\left(s, u^i, \phi^{-i}(s), \lambda^i\right) \mid u^i \in U^i\left(s, \phi^{-i}(s)\right)\right\}. \qquad (4.3.4)$$

Assume that the state space is convex, and there exist n *continuous functions* $\lambda^i: [0, \infty) \to \mathbb{R}$ *such that*

(i) for any player $i \in \{1,2, \dots, n\}$ *and each time point* $t \in [0, \infty)$*, the maximum condition* $H^i\left(s(t), \phi^i\left(s(t)\right), \phi^{-i}\left(s(t)\right), \lambda^i(t)\right) = H^{i*}\left(s(t), \phi^{-i}\left(s(t)\right), \lambda^i(t)\right)$ *holds;*

(ii) for any player $i \in \{1,2, \dots, n\}$ *and each time point* $t \in [0, \infty)$*, the adjoint equation* $\dot{\lambda}^i(t) = \rho\lambda^i(t) - (\partial/\partial s)H^{i*}\left(s(t), \phi^{-i}\left(s(t)\right), \lambda^i(t)\right)$ *holds;*

(iii) either $\lim_{t\to\infty} e^{-\rho t}\lambda^i(t)\tilde{s}(t) = 0$ *holds for any* $i \in \{1,2, \dots, n\}$ *and any feasible state trajectory* $\tilde{s}(.)$*, or there exists a real number* $a \in \mathbb{R}$ *such that* $s \geq a$ *for all* $s \in S$*,* $\lambda^i(t) \geq 0$ *for all* $i \in \{1,2, \dots n\}$ *and all sufficiently large* t*, and* $\lim_{t\to\infty} e^{-\rho t}\lambda^i(t)[s(t) - a] \leq 0$.

Then $(\phi^1, \phi^2, \dots, \phi^n)$ *is a Markovian Nash equilibrium.*

Proof:

It can be immediately seen that the proof of Theorem 3.4.1 applies to the present theorem.□

If the game is symmetric in the sense that all players have the same utility function, the same value function, the same sets of feasible controls, and the same discount rate, and if the dynamic rule is symmetric with respect to the players' controls, then it is possible to find a symmetric Nash equilibrium $(\phi^1, \phi^2, \dots, \phi^n)$ in which $\phi^i = \phi^j$ for all $i, j \in \{1,2, \dots, n\}$. In other words, in the game with the above setup, one can take advantage of the symmetry to precisely characterize the Markovian Nash equilibria (see Chapter 5~7).

4.4 Time consistency and subgame perfectness

The previous section shows the Nash equilibria can be defined in two distinct ways: open-loop Nash equilibria and Markovian Nash equilibria, and the two kinds of Nash equilibria differ in the accessibility of the information on the externality.

In this section, we continue to restrict ourselves to Markovian equilibria and discuss two important properties of Markovian equilibria: time consistency and subgame perfectness.

Let the notation $\Gamma(s^0, 0)$ denote the differential game studied in Section 3.2 where each player $i \in \{1,2, \dots, n\}$ seeks an optimal control path to maximize its long-run utility as follows:

$$\int_0^\infty e^{-\rho t}F^i(s(t), u^i(t), u^{-i}(t))\, dt \qquad (4.4.1)$$

subject to the constraints $u^i(t) \in U^i\left(s(t), u^{-i}(t)\right)$ and the dynamic rule

$\dot{s}(t) = f\left(s(t),\ u^i(t),\ u^{-i}(t)\right), \quad s(0) = s^0.$

At any arbitrary time point t, let $s \in S$ be the corresponding state. Thus, the subgame that starts from the state $s(t)$ is defined to be

$$\int_t^\infty e^{-\rho(y-t)} F^i(s(y), u^i(y), u^{-i}(y))\, dy \tag{4.4.2}$$

subject to

$\dot{s}(t) = f\big(s(t),\ u^i(t),\ u^{-i}(t)\big),\quad s(t) = s.$

Hence $\Gamma(s, t)$ is a differential game defined on the timer interval $[t, \infty)$ with the initial state $s(t) = s$.

Definition 4.4.1

Let $(\phi^1, \phi^2, \dots, \phi^n)$ be a Markovian Nash equilibrium for the game $\Gamma(s^0, 0)$ and let $s(.)$ denote the unique state trajectory generated by this equilibrium. The equilibrium is called time consistent if, for each $t \in [0, \infty)$, the subgame admits a Markovian Nash equilibrium $(\psi^1, \psi^2, \dots, \psi^n)$ such that $\psi^i\big(s(t)\big) = \phi^i\big(s(t)\big)$ holds for all $i \in \{1,2, \dots, n\}$ and any $t \in [0, \infty)$.

In order to interpret the property of time consistency, let us assume the n players are playing the game $\Gamma(s^0, 0)$ by choosing the Markovian Nash equilibrium $(\phi^1, \phi^2, \dots, \phi^n)$ where ϕ^i is a mapping from S to $\mathbb{R}$. This means at the time point $y \in [0, \infty)$, each player $i \in \{1,2, \dots, n\}$ chooses its action $u^i(y)$ subject to the rule $u^i(y) = \phi^i(s(y))$ where $s(y)$ is the state variable observed at y. Although the players possibly choose different control variables $u(y)$ as the state variable $s(y)$ is evolving, the rule $\phi^i(s)$ remains unchanged throughout the play of the game. With time passing, at the time point t, the players are supposed to play the subgame $\Gamma(s(t), t)$. If the decision rule $(\phi^1, \phi^2, \dots, \phi^n)$ of the entire game $\Gamma(s^0, 0)$ still constitutes a Markovian Nash equilibrium in the subgame $\Gamma(s(t), t)$, then $(\phi^1, \phi^2, \dots, \phi^n)$ is time consistent. Simply speaking a Markovian Nash equilibrium is time consistent if it is a Markovian Nash equilibrium of every subgame along the original equilibrium state trajectory $s(.)$.

Theorem 4.4.1

Every Markovian Nash equilibrium of a differential game is time consistent.

Proof:

In the entire game $\Gamma(s^0, 0)$. Suppose $(\phi^1, \phi^2, \dots, \phi^n)$ is the Markovian Nash equilibrium that generates the optimal control path $(u^1(.), u^2(.), \dots, u^n(.))$ with the state trajectory $s(.)$. If the time consistency is not satisfied, then at the time t, we suppose player i can improve its utility by choosing the strategy $\tilde{\phi}^i$ to play the subsequent subgame $(\Gamma(s(t), t))$, given that its opponents stick to the strategies profile ϕ^{-i}. More precisely, player i's strategy consists of the following two parts:

$$\pi^i(s(y)) = \begin{cases} \phi^i(s(y)) & if \quad y \in [0, t), \\ \tilde{\phi}^i(s(y)) & if \quad y \in [t, \infty). \end{cases} \tag{4.4.3}$$

The strategies profile $(\phi^1, \phi^2, \dots, \phi^{i-1}, \pi^i, \phi^{i+1}, \dots, \phi^n)$ forms a new Markovian Nash equilibrium. We suppose the Markovian Nash equilibrium generates an optimal control $(v^1(.), v^2(.), \dots, v^n(.))$ with the state trajectory $x(.)$. For the player i stated above, it holds that

$$J^i\left(v^i(.)\right) = \int_0^t e^{-\rho t} F^i(s(y), u^i(y), u^{-i}(y))\, dy + \int_t^\infty e^{-\rho t} F^i(x(y), v^i(y), v^{-i}(y))\, dy. \tag{4.4.4}$$

Since it is assumed that $\tilde{\phi}^i$ is preferred to ϕ^i in the subgame $(\Gamma(s(t), t))$, it turns out that

$$\int_0^t e^{-\rho t} F^i\left(s(y), u^i(y), u^{-i}(y)\right) dy + \int_t^\infty e^{-\rho t} F^i\left(x(y), v^i(y), v^{-i}(y)\right) dy$$

$$> \int_0^t e^{-\rho t} F^i\left(s(y), u^i(y), u^{-i}(y)\right) dy + \int_t^\infty e^{-\rho t} F^i\left(s(y), u^i(y), u^{-i}(y)\right) dy,$$

which implies $J^i(v^i(.)) > J^i(u^i(.))$. This inequality is a contradiction to the assumption $(\phi^1, \phi^2, \ldots, \phi^n)$ is a Markovian Nash equilibrium of the game $\Gamma(s^0, 0)$. The contradiction proves the claim. □

Time consistency is the minimal requirement for the credibility of a Markovian Nash equilibrium in a differential game. If player i has incentive to deviate from its promised action at one or multiple time points in its Markovian strategy ϕ^i, then its opponents would not believe its announcement in the first place. In such a case, the opponents will compute their own strategies by taking player i's deviation into the account. Consequently, this will lead to a new Markovian strategies profile instead of the original one. Theorem 4.4.1 ensures this situation does not occur in a Markovian Nash equilibrium.

On the basis of the time consistency addressed above, subgame perfectness of Markovian Nash equilibria is defined as follows:

Definition 4.4.2

Let $(\phi^1, \phi^2, \ldots, \phi^n)$ be a Markovian Nash equilibrium for the game $\Gamma(s^0, 0)$. The equilibrium is called subgame perfect if, for each $t \in [0, \infty)$ and any $s \in S$, supposing the current state $s(t) = s$, the subgame $\Gamma(s, t)$ admits a Markovian Nash equilibrium $(\psi^1, \psi^2, \ldots, \psi^n)$ such that $\psi^i(s(y)) = \phi^i(s(y))$ for each $i \in \{1,2,\ldots,n\}$ and all $y \in [t, \infty)$. A subgame perfect Markovian Nash equilibrium is called a Markov perfect (Nash) equilibrium.

In contrast to time consistency, subgame perfectness not only requires $(\phi^1, \phi^2, \ldots, \phi^n)$ is a Markovian Nash equilibrium of the subgame $\Gamma(s(t), t)$, where $s(.)$ is the state trajectory generated by the $(\phi^1, \phi^2, \ldots, \phi^n)$, but also it is a Markovian Nash equilibrium for all subgame $\Gamma(s, t)$ for any $t \in [0, \infty)$ and any $s \in S$. In the reverse direction, of course, a Markov perfect equilibrium is time consistent.

The next theorem clarifies the precise condition for a strategy profile to be a Markov perfect equilibrium.

Theorem 4.4.2

Let $(\phi^1, \phi^2, \ldots, \phi^n)$ be a given *n*-tuple in which $\phi^i: S \to \mathbb{R}$ and make the following assumptions

(i) *for any time point* $y \in [0, \infty)$ *and any state* $x \in S$*, there exists a unique continuously solution* $s: [0, \infty) \to S$ *of the differentiation equation*

$$\dot{s}(t) = f\left(s(t), \phi^1\big(s(t)\big), \phi^2\big(s(t)\big), \dots, \phi^n\big(s(t)\big)\right), \quad s(y) = x, \tag{4.4.5}$$

(ii) for all $i \in \{1,2,\dots,n\}$*, there exists a continuously differentiable function* $V^i: S \to \mathbb{R}$ *such that the HJB equations*

$$\rho V^i(s) = \max_{u^i \in U^i\left(s, \phi^{-i}(s)\right)} \left\{F^i\left(s, u^i, \phi^{-i}(s)\right) + V_s^i(s) f\big(s, u^i, \phi^{-i}(s)\big) \right\} \tag{4.4.6}$$

holds true for all $s \in S$.

(iii) The function $V^i(.)$ *is bounded on the domain S.*

At an arbitrary time point t*, pick up an arbitrary* $s \in S$*. Suppose the* $u^i(.)$ *denotes the control path such that* $u^i(t) \in U^i(s, \phi^{-i}(s))$ *and* $u^i(y)$ *maximizes the right-hand side of (4.4.6) for any* $y \in [t, \infty)$*. If* $\phi^i(s(y)) = u^i(y)$ *holds for each* $i \in \{1,2,\dots,n\}$ *and any* $y \in [t, 0)$*, then* $(\phi^1, \phi^2, \dots, \phi^n)$ *is a Markov perfect equilibrium.*

Proof:

Different from Theorem 4.3.1, we need to show at any $t \in [0, \infty)$, for any $s \in S$, supposing the present state $s(t) = s$, the control path $u^i(.)$ along which $u^i(t) \in U^i(s, \phi^{-i}(s))$ and $u^i(y)$ maximizes the right-hand side of (3.4.6) for any $y \in [t, \infty)$ is an optimal control path. According to Theorem 4.3.1, it is straightforward to show that the above statement holds true. Since the claim gives the hypothesis $\phi^i(s(y)) = u^i(y)$ holds for each $i \in \{1,2,\dots,n\}$ and any $y \in [t, 0)$, $(\phi^1, \phi^2, \dots, \phi^n)$ is a Markov perfect equilibrium. □

Following the proof of Theorem 3.3.1, it can be shown from (4.4.6) that along the optimal control path characterized by the MPE, the action profile solves the right-hand side of (4.4.6) for any feasible external state, and the HJB equation can be transformed into the maximized long-run utility of the multiple-player different game over the infinite time horizon. On the other hand, if it is possible to characterize an optimal control path via the right-hand side of the HJB equation (4.4.6), then such an optimal control path forms a MPE of the game. In the practical application of the differential game framework (see Chapter 5~7), the optimal control path can be specified via the first order conditions of the right-hand side of (4.4.6) and the symmetry.[6]

4.5 Summary

This Chapter mainly restricts attention to Markovian Nash equilibria of differential games. Definition 4.2.1 clearly presents the relation between the optimal control path and the Markovian Nash equilibria. We show that the techniques of the HJB equation and the Pontryagin's maximum principle given in Chapter 3 are both valid to Markovian Nash equilibria (refer to Theorem 4.3.1 and Theorem 4.3.2). Additionally, we also clarify two properties of Markovian Nash equilibria: time consistency and subgame perfectness. As stated in Theorem 4.4.1, it can be easily shown that any Markovian Nash equilibrium is time consistent, so that in any subgame along the control path, the corresponding part of

[6] The right-hand side of the HJB equation (4.4.6) has to be concave in the control variable u^i.

the entire Markovian strategies also constitutes a Markovian Nash equilibrium as well. In contrast to time consistency, a Markovian Nash equilibrium is subgame perfect if in any possible subgame (rather than only along the control path), the corresponding Markovian strategies forms a Markovian Nash equilibrium. In the context of differential games, such a Markovian Nash equilibrium is called Markov perfect equilibrium for ease, and it is clear that any Markov perfect equilibrium satisfies time consistency.

We also show the Markov perfect equilibrium can be characterized from HJB equation. The subsequent chapters focus attention on the Markov perfect equilibria of three specific differential games with linear quadratic form: the dynamic renewable resource duopoly, the dynamic public goods provision problem, and a policy game concerning the combination use of the emission tax and the use of a pollution abatement technique. In each of them, we are concerned with the characterization and examination of Markov perfect equilibria of the game.

Chapter 5.

The best stable Markov perfect equilibrium in differential games with linear quadratic form

5.1 Introduction

The differential games that admit analytic Markov perfect equilibria (or MPEs for short) receives large attention in the research of the dynamic economics and management science. Particularly, as shown by, e.g., Dockner and Long (1993), Wirl (1996), Itaya and Shimomura (2001), Rubio and Casino (2002), Fujiwara (2008) and Lambertini and Mantovani (2014), in the differential games with the well-defined linear quadratic (LQ) forms,[7] it is possible to graphically characterize an infinite number of stable MPEs, including a unique linear MPE and infinitely many nonlinear ones. When the game is played along the optimal control path generated by a stable MPE, the state variable asymptotically converges to a steady state.

The main goal of this chapter is to propose a new technique to evaluate all stable MPEs in terms of the long-run benefits in the LQ-form differential games stated above. The new technique clarifies connections among the players' equilibrium strategies, the steady states and the long-run benefits. We first generalize the LQ form differential game framework on an infinite time horizon, and roughly express the basic idea of the new technique. Moreover, in order to further present the idea precisely, we employ the dynamic renewable resource duopoly (Fujiwara (2008)),[8] a well-known model of the LQ-form differential games. In the dynamic renewable resource duopoly, with the use of the new technique, we show for each feasible initial state, the long-run profit is increasing in steady-state resource in the stable MPEs. This result suggests each firm (i.e., each player) can earn the highest long-run profit (i.e., long-run benefit) by choosing the stable MPE that guides the resource to reach the highest level at the steady state. In other words, the stable MPE that results in the highest steady state is the best equilibrium of the game. Additionally, we find the best MPE is independent of the discount rate. This finding suggests the best MPE does not change with the extent to which the firms value the future profits. Finally, we also show that for some large initial states, the equilibrium output that results in a constant resource forms the best MPE of the game. To verify the result

[7] The terminology *LQ (linear-quadratic) form* indicates that the differential game comprises a LQ objective function together with a linear dynamic rule.

[8] The dynamic renewable resource duopoly is first formulated by Benchekroun (2003) and Benchekroun (2008). In Fujiwara (2008) and Lambertini and Mantovani (2016), simplifying the linear dynamic rule of the game, the authors precisely characterize and examine each stable MPE. In particular, in these two studies, the stable MPEs are evaluated by the levels of the steady states.

derived by the new technique, we employ the composite Simpson's rule to numerically rank the stable MPEs in terms of the long-run profits.

So far many previous studies, e.g., Docker and Long (1993), Wirl (1996), Itaya and Shimomura (2001), Rubio and Casino (2002), Fujiwara (2008) and Lambertini and Motovani (2014) are all concerned with characterization of the MPE that provides the best steady state. In contrast, our technique provides a new method to evaluate and rank all stable MPEs in terms of the long-run profits. Moreover, we show that the new technique possibly applies to all LQ-form differential games. Hence, in the model formulated by Chapter 6 and Chapter 7, the long-run utility of each MPE is also computed with the use of the new technique presented in this chapter.

The rest of the present chapter is arranged as follows: Section 5.2 generalizes the structure of the LQ-form differential games, and employs the envelope-theorem-based technique (Itaya and Shimomura (2001), Page 161) to form connections among the MPE strategies, the steady state and the long-run benefits for each stable MPE. In order to further present the basic idea of this chapter, Section 5.3 restricts attention to the dynamic renewable resource duopoly (Fujiwara (2008)), and defines the fundamental models of the game. Section 5.4 graphically characterizes all stable MPEs of the game. Using the connections derived by Section 5.2, Section 5.5 and Section 5.6 present a new technique to evaluate each stable MPE in terms of the long-run profit. Using this technique, we specify the best MPE for each feasible initial state. Section 5.7 compares the stable MPEs and the unstable ones. Section 5.8 gives multiple numerical examples for verification. Section 5.9 summarizes. Note that although the basic idea of this chapter is precisely presented in the dynamic renewable resource duopoly, Section 5.2 implies the new approach possibly applies to the dynamic public-goods provision problem presented by Chapter 6 and the policy game (Dockner and Long (1993)).

Chapter 6.

Public goods provision under dynamic budgets

6.1 Introduction

In the public economics, the formulation of dynamic public good provision is one of the main issues that receives much attention. Fershtman and Nitzan (1991) is the pioneer research that addresses the dynamic-public-good-provision issue by means of the differential game approach. Wirl (1996) employs the Tutusi and Mino (1990) technique to further specify infinitely many nonlinear stable Markov perfect equilibria (MPEs) of the game. Itaya and Shimomura (2001) and Ihori and Itaya (2001) introduce a fixed budget constraint to the model and use the envelope theorem to develop a new technique for the graphic characterization of the nonlinear stable MPEs. Using the Itaya and Shimomura (2001) model, Yanase (2006) proposes a tax-subsidy scheme to enhance the steady state for the stable MPEs.

More recently, Benchekroun and Long (2008) defines the stock of public goods as a social capital of trustworthiness. The authors adopt Itaya and Shimomura (2001) technique to examine analytic stable MPEs. Claude et al. (2012) employs the Benchekroun and Long (1998) technique to derive a social-optimum tax-subsidy scheme that regulates the individuals' public goods contribution.

In this chapter, we use the differential game approach to examine the public goods provision problem under dynamic budgets. In the model, we assume each individual makes efforts to earn income over time. Different from the fixed-income setup proposed by Itaya and Shimomura (2001), Ihori and Itaya (2001) and Yanase (2006), each individual in the present model dynamically chooses the level of income, and allocates the income between private consumption and public-good contribution subject to dynamic budgets. During the play of game, the individuals' public-good contribution accumulates over time, which generates a positive stock externality. In this setup, we employ the Itaya and Shimomura (2001) technique to drive infinite many stable nonlinear Markov perfect equilibria (or stable nonlinear MPEs for short). Due to the dynamic budget constraints, we show it is possible to clarify the relation between the level of the income and the amount of the public-good contribution in each MPE. Making use of this relation equation, we show the effect of the level of the public goods on the budget and the allocation of the income can be easily analyzed. In other words, the relation equation enables the explicit examination on whether each individual raises or curtails its level of income and the amount of public-good contribution as the state variable is approaching the steady state during the play of the game. Furthermore, we employ the relation equation to clarify the interior-solution conditions for the stable nonlinear MPEs, and show the level of public-good contribution is increasing in the amount of income. Therefore, the stable MPE with higher incomes and higher public-good contribution guides the level of the public goods to converge to a higher steady state. This outcome suggests that in order to attain a higher level of public goods, each individual has to makes more efforts to increase its income so that it can raise the amount of the public-good contribution.

In addition to the clarification and the use of the above relation equation, we also adopt the tangency-point technique (Fujiwara (2008)) to characterize the stable MPE that results in the highest steady state for any given initial state. In particular, we show given a sufficiently large initial state, the stable MPE in which the public-good level remains constant throughout the play of the game generates the highest steady state.

Chapter 5 proposes a long-run-utility-based technique for evaluation of nonlinear stable MPEs, and generalizes the technique to all LQ-form differential game framework (refer to Section 5.2 for the definition of LQ form). In this chapter, we show the technique (more specifically Proposition 5.5.1) also applies to the present model. We find given a sufficiently low initial state, the stable nonlinear MPE that results in the higher level of public goods provides a higher long-run utility. This result implies the stable MPE characterized by the above tangency-point technique is the best equilibrium in the sense that it provides the highest long-run utility and thereby is the most socially efficient stable MPE of the game.

The present chapter is arranged as follows: Section 6.2 defines the fundamental models. Section 6.3 employs Itaya and Shimomura (2001) technique to specify the ordinary differential equations (ODEs) for the graphic characterization of the stable MPEs. Section 6.4 derives all stable MPEs by means of the ODEs, and examines the effect of the pubic-good level on each individual's budget and allocation of the income. Section 6.5 clarifies the stable MPE that results in the highest steady state. Section 6.6 characterize the best stable MPE by means of the new evaluation technique proposed by Chapter 5. Section 6.7 compares the MPE with the highest steady state and the Nash equilibrium of the static setting in terms of the contribution of the public good and the level the public good. Section 6.8 concludes.

Chapter 7.

A dynamic policy game of the combination use of emission taxation and a pollution-removal technique

7.1 Introduction

In the context of differential games, the policy game approach is employed to formulate the governments' dynamic policy on the optimal trade-offs between firms' production and the management of the stock externality. Yanase (2007) and Yanase (2009) develop a policy game in which the government chooses a dynamic tax rate to induce the firms to lessen the pollutant emission with an end-of-the-pipe technique. Fujiwara et al. (2011) employs the policy-game concept to formulate an optimal tariff policy between two countries engaging in an international duopoly with a dynamic global productive asset.

In this chapter, we consider a differential game in which the government seeks to optimally control pollution by means of a policy instrument including emission taxation and a pollution-removal technique. In the game, there exist multiple firms engaging in a Cournot competition and each firm's production generates pollutants emission to the environment. As in Yanase (2007) and Yanase (2009), we assume the information about the stock externality is only available to the government, and thereby each firm does not take the stock externality into consideration when it chooses an optimal output. This assumption can be interpreted as: the firms are not required to autonomously participate in the pollution-control activities, and instead, the government directly takes charge of the whole pollution management. This assumption simplifies the framework of the differential game, which permits characterization of the analytic Markov perfect equilibrium.

In the above setup, the government is the player of the differential game (i.e., the policy game), and the emission tax rate and the pollution-removal efforts (i.e., the use of the pollution-removal technique) are the control variables. The emission taxation reduces the pollution emission but it hurts the industry profit and the consumer surplus. In contrast, the pollution-removal effort directly alleviates the pollution level, but it could be quite costly. For each instantaneous moment of the differential game, the government first gains a tax revenue via the emission taxation, and then equally redistributes the tax revenue to each firm or use the tax revenue to cover the cost of the pollution-abatement efforts.

We characterize a linear Markov perfect equilibrium (or linear MPE for short) that results in a steady state, and show it is the unique stable MPE of the game. Along the control path of the MPE, the government dynamically varies the emission tax rate and the level of the pollution-removal efforts to maximize the long-run social welfare, and the government's MPE behavior guides the pollution stock to asymptotically converge to a steady state. When the steady state is reached, we show it is very convenient to examine the impacts of the efficiency of the pollution-removal technique and the impacts of harmfulness of the pollution stock on the government's MPE behavior and the social welfare. Moreover, we also clarify the condition under which the government's tax revenue exceeds the cost of the pollution-abatement efforts at the steady state. According to this result, we show when the pollution stock is high, the total amount of the tax revenue is insufficient to cover the cost of the pollution-abatement efforts, and thereby the government has to incur the additional expense of the pollution management. In contrast, as the pollution stock is sufficiently close to the steady state, the tax revenue exceeds the cost of the pollution-abatement efforts, and thereby the government can reimburse the surplus of the tax revenue to the firms.

In recent years, the formulation of the pollution-removal efforts receives much attention in the literature of dynamic pollution control. Ouardighi et al. (2014), Ouardighi et al. (2016), Ouardighi et al. (2018-a) and Ouardighi et al. (2018-b) assume the players invest in a technology to restore the capacity of the environmental pollution absorption. These models comprise two stock externalities: the pollution stock and the stock of players' investments (i.e., the stock of the pollution-removal efforts). In contrast, Yeung (2007), Yeung (2014) and Manoussi and Xepapadeas (2017) assume the players' pollution-removal efforts instantly abates the pollution level or the damage caused by the pollution stock. This setup can be interpreted, e.g., as: the players remove the greenhouse gasses from the atmosphere with the use of geoengineering methods. Our model is analogous to Yeung (2007) and Manoussi and Xepapadeas (2017) in terms of the setup of the pollution-removal efforts.

The formulation of dynamic emission taxation is a main subject in the literature of the differential game as well. In the pioneer studies, Wirl (1994) and Rubio and Escriche (2001) presents a dynamic pigouvian tax in the differential game played by the benevolent government and the cartel of manufactures. The recent developments of this subject can be confirmed by e.g., Wirl (2014-a) and Wirl (2014-b).

The present chapter is organized as follows: Section 7.2 defines the fundamental models. Section 7.3 characterizes the linear stable MPE and shows it is the unique stable MPE of the game. Section 7.4 examines how the efficiency of the pollution-removal technique and the harmfulness of the pollution stock influence the government's behavior and the social welfare, and shows the stable linear MPE possibly generates surplus tax revenue that can be reimbursed to each firm. Section 7.5 concludes.

Chapter 8.

Conclusions and future work

8.1 Summary of contributions

In this research, we examine the structure of the linear quadratic (LQ) differential game with an infinite time horizon, and adopt the LQ differential game framework to formulate three specific models.

In order to enhance the understanding of structure of the differential game, we starts the research by introducing and reviewing the repeated game and the dynamic game, and state the relations among the three games (i.e., the differential game, the repeated game and the dynamic game). The three games are identical in that the stage game is play repeatedly and the whole game can be described in the recursive form. However, the differential game are played in continuous time whereas the others are considered over a discrete time horizon.

We refine structure of differential game framework over infinite time horizon and examine the Hamilton-Jacobi-Bellman (HJB) equation and Pontryagin's maximum principle, by deriving the recursive form of the differential game. On the basis of the recursive form, we verify how the Hamilton-Jacobi-Bellman equation (or HJB equation for short) and Pontryagin's maximum principle work in characterizing an optimal control path (i.e., an optimal solution) of the differential game. In particular, the HJB equation provides sufficiency conditions for the optimal control problem of the differential game. Therefore, we employ the HJB equation to derive the analytic Markov perfect equilibria (MPEs) of the specific differential-game models formulated subsequently.

In the third part of our research, we generalize the LQ differential game framework over an infinite horizon, and propose a new technique to characterize the best stable MPE that provides the highest long-run utility in the differential game. We show the MPE with the best steady state provides the highest long-run profit. We apply this new technique to the resource-extraction dynamic duopoly and state that it is also possibly valid to the other kinds of the LQ differential game such as transboundary pollution control problem and the dynamic public goods provision problem. Different from the previous studies, the new technique clarifies connections among the players' equilibrium strategies, the steady states and the long-run benefits, and provides a new method to evaluate and rank all stable MPEs in terms of the long-run profits.

Making use of the LQ differential game framework, we formulate a public goods provision problem subject to dynamic budget constraints (i.e., dynamic income), and show in any stable MPE, the level of public goods provision is increasing in the amount of income. This result suggests in order to enhance the level of public goods at the steady state, the individuals need to choose a stable MPE with higher income. Due to dynamic budget constraints, we find it is possible to analyze on the effect of the public good level on the dynamic budgets as well as the allocation of incomes. Moreover, we also employ the new technique proposed by Chapter 5 to evaluate the stable MPEs in terms of the long-run utilities, and exhibit given a relatively small initial state, the long-run utility is increasing in the steady-state public-good level.

In addition to the public goods provision problem, we also formulate a policy game by means of the LQ differential game framework. The game comprises multiple polluting oligopolistic firms and a government who takes charge of the pollution management. The government seeks to optimally control the pollution by means of emission taxation together with a pollution-removal technique. We characterize a unique stable linear MPE of the game. We conclude that when the pollution level is sufficiently close to the steady state, the government can possibly reimburse the surplus of the tax revenue to the firms.

8.2 Future work

In the future work, we will continue the extensive research on the public goods problem subject to dynamic budget constraints and the policy game of the dynamic pollution control problem. In the former subject, because level of the public provision is increasing in the amount of the income, it could be interesting to propose a policy to raise the level of the incomes. An increase in the individuals' income can stimulate both the private consumption and the public goods provision, and thereby efficiency of the economy can be improved. In the latter subject, we plan to consider a policy game on the pollution management by means of emission taxation and pollution-abatement-technique research and development (R&D). In the game, the government dynamically makes trade-offs between the emission taxation and the investments in the R&D. It is worth clarifying whether an optimal policy (i.e., an analytic MPE) can be derived in the game, and how much tax revenue can be reimbursed to the firms after the investments.

Reference.

Benchekroun, H., N. van Long (1998) Efficiency inducing taxation for polluting oligopolists. Journal of public economics. 70, 325-342.

Benchekroun, H. (2003) Unilateral production restrictions in a dynamic duopoly. Journal of Economic Theory. 111, 214-239.

Benchekroun, H. (2008) Comparative dynamics in a productive asset oligopoly. Journal of Economic Theory. 138, 237-261.

Benchekroun, H. and N.V. Long (2008) The build-up of cooperative behavior among non-cooperative selfish agents. Journal of Economic Behavior & Organization, 67,239–252.

Chartrand, G. and P. Zhang (2012) A first course in graph theory. Dover publication, Inc.

Claude, D., C. Figuieres and M. Tidball (2012) Regulation of investments in infrastructure: the interplay between strategic behaviors and initial endowments. Journal of Public Economic Theory. 14 (1), pp.35-66.

Dockner, E., J., N. V. Long (1993) International Pollution Control. Cooperative versus Noncooperative Strategies. Journal of environmental economics and management 25,1, 13-29.

Dockner, E., J., S. Jorgensen, N. V. Long, and G. Sorger: Differential Games in Economics and Management Science. Cambridge: Cambridge University Press (2000).

Fershtman, C. and S. Nitzan (1991), Dynamic voluntary provision of public goods. European Economic Review 35, 1057-1067.

Fujiwara, K. (2008) Duopoly can be more anti-competitive than monopoly. Economics letters. 101,3,217-219.

Fujiwara, K., T. Shinozaki and A. Yanase (2011) Dynamic Interactions in Trade Policy in a Differential Game Model of Tariff Protection. Review of Development Economics, 14 (4), 689-698.

Hackner, J. (2000) A Note on Price and Quantity Competition in Differentiated Oligopolies. Journal of Economic Theory 93, 233-239.

Ihori, T. and J. Itaya (2001) A dynamic model of fiscal reconstruction. European Jorunal of Political Economy. 17,4, 779-797.

Huang, Y., Takehiro Inohara

(2015-a) Group-separations based on the repeated prisoners' dilemma games, Applied Mathematics and Computation, Volume 256, pp. 267-275.

(2015-b) Steady-state stock and group size: An approach of dynamic voluntary provisions of public goods. Applied Mathematics and Computation, Volume 270, pp. 505-510.

Itaya, J., K.Shimomura (2001) A dynamic conjectural variations model in the private provision of public goods: a differential game approach. Journal of public economics 81,1, 153-172.

Lambertini, L. and A. Mantovani (2009) Process and product innovation by a multiproduct monopolist: A dynamic approach. International Journal of Industrial Organization 27, 508–518.
Lambertini, L. and A. Mantovani (2014) Feedback equilibria in a dynamic renewable resource oligopoly: Pre-emption, voracity and exhaustion. Journal of Economic Dynamics and Control. 47, 115-122.
Lambertini, L. and A. Mantovani (2016) On the (in)stability of nonlinear feedback solutions in a dynamic duopoly with renewable resource exploitation. Economics letters. 143, 9-12.
Mailath, G.J. and L.Samuelson (2006) Repeated Games and Reputations. Oxford University Press.
Manoussi, V. and A. Xepapadeas (2017) Cooperation and Competition in Climate Change Policies: Mitigation and Climate Engineering when Countries are Asymmetric. Environmental and Resource Economics, 66, 605-627.
Ouardighi, F.E., H. Benchekroun and D. Grass (2014) Controlling pollution and environmental absorption capacity. Annals of Operations Research, 220:110-133.
Ouardighi, F.E., K. Kogan, G. Gnecco and M. Sanguineti
(2018-a) Transboundary pollution control and environmental absorption efficiency management. Ann Oper Res, 1-29.
(2018-b) Commitment-Based Equilibrium Environmental Strategies Under Time-Dependent Absorption Efficiency. Group Decis Negot, 27:235-249.
Reynolds, S. (1987). Capacity investment, pre-emption and commitment in an infinite horizon model. International Economic Review 28(1), 69–88.
Rowat, C. (2007) Non-linear strategies in a linear quadratic differential game. Journal of Economic Dynamics and Control. 31,3179-3202.
Rubio, S.J, B. Casino (2002) A note on cooperative versus non-cooperative strategies in international pollution control. Resource and energy economics,3, 24, 251-261.
Singh, N. and X. Vives (1984) Price and quantity competition in a differentiated duopoly. Rand Journal of Economics. 15,4, 546-554.
Tamai, T. (2010) Public goods provision, redistributive taxation, and wealth accumulation. Journal of Public Economics. 94,1067-1072.
Tsutsui, K. and K. Mino (1990) Nonlinear strategies in dynamic duopolistic competition with sticky prices. Journal of Economic Theory, 52,1, 136-161.
Uler, N. (2009) Public goods provision and redistributive taxation. Journal of Public Economics. 93, 440-453.
Wirl, F. (1996) Dynamic voluntary provision of public goods: Extension to nonlinear strategies, European Journal of Political Economy 12, 555–560.
Wirl, F.
(2014-a) Taxes versus permits as incentive for the intertemporal supply of a clean technology by a monopoly. Resource and Energy Economics. 36, 248–269.
(2014-b) Optimal introduction of carbon taxes. Energy System. 5:3–17.
Yanase, A. (2006) Dynamic voluntary provision of public goods and optimal steady-state subsidies. Journal of Public Economic Theory 8, (1), pp.171-179.
Yanase, A. (2007) Dynamic games of Environmental Policy in a Global Economy: Taxes versus Quotas. Review of International Economics, 15(3), 592-611.
Yanase, A. (2009) Global environment and dynamic games of environmental policy in an

international duopoly. Journal of Economics, 97,2,121-140.
Yeung, D.W.K. (2007) Dynamically consistent cooperative solution in a differential game of transboundary industrial pollution. Journal of Optimization Theory and Applications 134, 143–160.
Yeung, D.W.K. (2014) Dynamically consistent collaborative environmental management with production technique choices. Annals of Operations Research, 220,181-204.