College of Natural Sciences
Department of Mathematics

Optimal control and Differential game

BY: Waltengus Dagne
Advisor: Semu Mitiku(Ph.D)

Graduate project Report submitted to the department of Mathematics in partial fulfillment of the requirements for the Degree of Master of Science in Mathematics.

Stream: Optimization

June 26, 2016

Department of Mathematics
College of Natural Sciences
Addis Ababa University

The undersigned hereby certify that they have read and recommend to the school of graduate studies for acceptance of a project entitled on **Optimal control and differential games** by Walltengus Dagne in partial fulfillment of the requirements for the degree of master of Science.

Signed by the Examining Committee:

Advisor	Signature	Date
Examiner	Signature	Date
Examiner	Signature	Date

Abstract

A differential game problem is a generalized optimal control problem in cases where there are more than one decision makers, called players. However, differential games are conceptually far more complex than optimal control problems in the sense that it is no longer obvious what constitutes a solution. If the number of players are two, one player chooses the control $u_1(t) \in \Omega_{u_1} \subseteq \Re^{mu_1}$ and tries to maximize his cost functional $J_1(u_1, u_2)$, while the other player chooses the control $u_2(t) \in \Omega_{u_2} \subseteq \Re^{mu_2}$ and tries to maximize his cost functional $J_2(u_1, u_2)$ within a common dynamic system which is described by differential equation of the form $\dot{x}(t) = f(t, x, u_1, u_2)$. Where $u_1(t)$ and $u_2(t)$ are the controls implemented by the first and second player respectively. In this cases, solving a differential game problem is mathematically quite tricky in order to satisfy the absolute need of the two players. Depending on the willingness of the two players they may cooperate or may not cooperate each other. In this project, we discussed only non cooperative differential game in particular by adopting open-loop strategies and feedback(closed-loop) strategies to solve the problem of the game. Solution procedures for two person differential games are also discussed.

Acknowledgements

First, I would like to express my deepest gratitude to **God** for giving me patience; I would also like to extend my gratitude to my advisor **Dr. Semu Mitiku** for his helpful discussion, comments and providing the necessary materials in the preparation of this paper. This paper would not have been possible without his support.

Finally, I would like thanks to **Department of mathematics, Addis Ababa University** for giving the necessary materials throughout the preparation of this paper.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Mathematical Notations

ODEs	ordinary differential equations
PDEs	partial differential equations
ELDE	Euler Lagrange differential equation
OCP	Optimal Control Problem
DP	dynamic programming
A_i	Assumption i for $i \in \{1, 2, 3, 4\}$
HJB	Hamilton-Jacobi-Bellman
PMP	pontriagns maximum or minimum principle
t_0, T	initial time , final time
$u(t)$	control vector variable
$x(t)$	state vector variable
$\Omega \subseteq \mathbb{R}^m$	control constraints
$\psi(x(T), T)$	terminal payoff
$J(t, x)$	optimal cost-to-go function
$V(t, x)$	value function for the optimization problem
$\nabla_x \psi_i(x(T))$	partial derivative of ψ_i with respect to x for $i = 1, 2$
$\nabla_t V_i \quad , \quad \nabla_x V_i$	partial derivative of V_i with respect to t and x respec. for $i = 1, 2$
$f(t, x(t), u(t))$	right-hand side of the state differential equation
$L(x(t), u(t), t)$	integrand of the running cost functional
$\mathbb{L}(x, u, \lambda, t)$	Lagrangian function
$\mathbb{H}(x, u, \lambda, \lambda_0, t)$	Hamiltonian function
$\lambda(t)$	Lagrange multipliers
$\phi^A_{ij}(a_i, b_j)$	payoff function for player A
$\phi^B_{ij}(a_i, b_j)$	payoff function for player B
$R^A(a), R^B(b)$	set of best possible replies of player A and B respectively
$\tilde{\phi}^A(\mu, v)$	expected values of the payoffs player A
$\tilde{\phi}^B(\mu, v)$	expected values of the payoffs player B
$u_1(\cdot) \in U_1, u_2(\cdot) \in U_2$	controls implemented by player 1 and player 2 respectively

Chapter 1

Introduction

1.1 What is game theory?

Mathematically, a game is a process of decision making with more than one agent or decision maker called players. Game theory deals with strategic interactions among multiple decision makers, called players (and in some context agents), with each player's preference ordering among multiple alternatives captured in an objective function for that player, which he either tries to maximize (in which case the objective function is a utility function or benefit function) or minimize (in which case we refer to the objective function as a cost function or a loss function). For a non-trivial game, the objective function of a player depends on the choices (actions) of at least one other player, and generally of all the players, and hence a player cannot simply optimize his own objective function independent of the choices of the other players. This thus brings in a coupling between the actions of the players, and binds them together in decision making even in a non-cooperative environment.

A decision making process with multiple agents is said to be non cooperative if each decision makers optimizes his own objective function for each choice of every other decision maker with out considering their common output. In the case of cooperative games, players may enforce cooperative behavior and agree in how to share the extra output, which is called a coalation formation.

If the players were able to enter into a cooperative agreement so that the selection of actions or decisions is done collectively and with full trust, so that all players would benefit to the extent possible, and no inefficiency would arise, then we would be in the realm of cooperative game theory, with issues of bargaining, coalition formation, excess utility distribution, etc.[2]

If no cooperation is allowed among the players, then we are in the realm of non-cooperative game theory, where first one has to introduce a satisfactory solution concept. To first order, such a solution point should have the property that if all players but one stay, then the player who has the option of moving away from the solution point should not have any incentive to do so because he cannot improve his payoff. Note that we cannot allow two or more players to move collectively from the solution point, because such a collective move requires cooperation, which is not allowed in a non-cooperative game. Such a solution point where none of the players can improve her payoff by a unilateral move is known as a non-cooperative equilibrium or Nash equilibrium.
We say that a non-cooperative game is nonzero-sum if the sum of the players' objective functions cannot be made zero after appropriate positive scaling and/or translation that do not depend on the players' decision variables. We say that a two-player game is zero-sum if the sum of the objective functions of the two players is zero or can be made zero by appropriate positive scaling and/or translation that do not depend on the decision variables of the players. If the two players' objective functions add up to a constant (without scaling or translation), then the game is sometimes called constant sum, but according to our convention such games are also zero sum. A game is a finite game if each player has only a finite number of alternatives, that is the players pick their actions out of finite sets (action sets); otherwise the game is an infinite game; finite games are also known as matrix games. An infinite game is said to be a continuous-kernel game if the action sets of the players are continua, and the players' objective functions are continuous with respect to action variables of all players.

If the decision variables of each players vary with time, then we have the so called dynamic games. Moreover , if time is continiously measured in dynamic games, the problem is called a differential game.
In differential games, each of the players solve optimal control problems in their own control(decision) variables. Hence such problems are set of (or sequence of) optimal control problems, where the order of optimizing problem (herarchical or non heriarchical).

The general statement of differential game is formulated as; find an admissible optimal controls $u_1 : [0,T] \to U_1$ and $u_2 : [0,T] \to U_2$ such that the dynamic system described by the differential equation

$$\begin{aligned} \dot{x} &= f(t, x, u_1, u_2), \qquad t \in [0,T] \\ x(0) &= x_0. \end{aligned}$$

is achieved. where $u_1(\cdot), u_2(\cdot)$ are the controls implemented by the two players that satisfy the pointwise constraints

$$u_1(t) \in U_1, \qquad u_2(t) \in U_2$$

for some given sets $U_1, U_2 \subseteq \Re^m$
For $i = 1, 2$ the goal of the i^{th} player is to maximize his own payoff, namely

$$J_i(u_1, u_2) = \psi_i[x(T)] - \int_0^T L_i(t, x(t), u_1(t), u_2(t))dt$$

where ψ_i is a terminal payoff, while L_i accounts for a running cost.
To mathematically address such type of problem we need to apply the solution procedure for optimal control problems.
Therefore in the priminary of this project, a set of necessary conditions for the optimality of a solution of an optimal control problem is derived using the calculus of variations. This set of necessary conditions is known by the name "Pontryagin's Minimum Principle" [9]. By exploiting Pontryagin's Minimum Principle, several optimal control and differential games problems are solved completely.

Solving an optimal control problem using Pontryagin's Minimum Principle typically proceeds in the following (possibly iterative) steps:

- Formulate the optimal control problem.
- Existence: Determine whether the problem can have an optimal solution.
- Formulate all of the necessary conditions of Pontryagin's Minimum Principle.
- Globally minimize the Hamiltonian function H:

$$u^*(x^*(t), \lambda^*(t), \lambda_0, t) = arg \min_{u \in Omega} H(x^*(t), u^*(t), \lambda^*(t), \lambda_0), t) \quad \text{for all} \quad t \in [t_0, T].$$

- Singularity: Determine whether the problem can have a singular solution.
- Solve the two-point boundary value problem for $x^*(\cdot)$ and $\lambda^*(\cdot)$.
- Eliminate locally optimal solutions which are not globally optimal.

In this project; we discussed a brief introduction to the theory of optimal control problems and the theory of static games and different concepts of solution are included, including Pareto optima, Nash and Stackelberg equilibria. In chapter 4 we introduced the basic framework of differential games for two players within an Open-loop solution strategy where the controls are allowed to depend only on time. Nash and Stackelberg solutions which can be computed by solving a two-point boundary value problem for a system of ODEs are also derived from the Pontryagin maximum principle.In addition, solutions of two person differential games in feedback form are also where the controls are allowed to depend on time and on the current state of the system are included.

Chapter 2

Preliminaries

2.1 Optimal Control Problem

A general unconstrained optimal control problem can be formulated as follows:
Find a picewise continuous admissible optimal control $u : [t_0, T] \to \Re^m$ such that the dynamic system described by the differential equation

$$\dot{x}(t) = f(x(t), u(t), t) \tag{2.1}$$

is transferred from the initial state

$$x(t_0) = x_0 \tag{2.2}$$

to the final state

$$x(T) = x_T \quad \text{or free} \tag{2.3}$$

such that the cost functional

$$J(x, u) = \psi(x(T), T) + \int_{t_0}^{T} L(x(t), u(t), t)dt \tag{2.4}$$

is minimized or maximized depending on the problem.
Here the function f, ψ, and L are assumed to be at least once continuously differentiable with respect to all of their arguments. The variable $x \in \Re^n$ describes the state of the system, and $u \in \Re^m$ represents the input of the decision maker, which is called the control.
Depending upon the type of the optimal control problem, the final time T could be fixed or free. When T is not specified (i.e when it is free) one needs to find the optimal time T , that gives an optimal control of the system.
In this project, we only consider optimal control problems where the initial time t_0 and the initial state $x(t_0) = x_0$ are specified.

NOTE:

1. If T is finite then the problem is called finite horizon optimal control problem.
2. If T is infinity the problem is called infinite horizon optimal control problem.

2.1.1 Lagrange Multiplier and the Hamiltonian Function

IF the optimal control problem is given by;

$$\min_u J(x,u) = \int_{t_0}^{T} L(x(t), u(t), t)dt$$

subjected to;

$$\dot{x}(t) = f(t, x(t), u(t)); \qquad x(t_0) = x_0; \qquad x(T) = x_T$$

where the state and control variables are not constrained. Thus the only constraint of the problem will be the differential equation together with the appropriate boundary or initial conditions

$$\dot{x}(t) = f(t, x(t), u(t)); \qquad x(t_0) = x_0; \qquad x(T) = x_T$$

This implies that;

$$f(t, x(t), u(t)) - \dot{x}(t) = 0$$

Corresponding to this equality constraint we define a multiplier function $\lambda : [t_0; T] \to \Re^n$ and then the Lagrangian of the problem will be ;

$$\mathbb{L}(x, u, \lambda, \dot{x}) = \int_{t_0}^{T} \left[L(x(t), u(t), t) + \lambda(t)\left(f(t, x, u) - \dot{x}(t)\right)\right] dt$$

Theorem 2.1.1. *If $(x^*, u^*, \lambda^*, \dot{x}^*)$ is a minimizer of $\mathbb{L}$, then (x^*, u^*) is also a minimizer of J(x,u).*

Proof. Define $p(x, u, \lambda) = \int_{t_0}^{T} \lambda(t)(f(t, x, u) - \dot{x}(t))dt$
Let λ^* corresponds to the pair (x^*, u^*) such that $(x^*, u^*, \lambda^*, \dot{x}^*)$ is a minimum for $\mathbb{L}$.
Then $\mathbb{L}(x^*, u^*, \lambda^*, \dot{x}^*) = J(x, u) + p(x, u, \lambda)$ and $\mathbb{L}(x^*, u^*, \lambda^*, \dot{x}^*) \leq \mathbb{L}(x, u, \lambda, \dot{x})$ for all (x, u).
If the pair (x^*, u^*) is feasible, the state condition must be satisfied. That is: $\dot{x}(t) = f(t, x, u)$, which implies $p(x^*, u^*, \lambda) = 0$.
$J(x^*, u^*) + p(x^*, u^*, \lambda^*) \leq J(x, u) + p(x, u, \lambda)$ for all (x, u, λ)
$\Rightarrow J(x^*, u^*) \leq J(x, u)$ for all feasible pair (x, u). i.e $J(x^*, u^*)$ is minimum.
Therefore the Theorem holds true. □

Now, in the above problem, let $G(x,u,\lambda,\dot{x},\dot{u},\dot{\lambda},t) = L(t,x,u) + \lambda(t)(f(t,x,u) - \dot{x}(t))$. Thus by the Euler Lagrangian Differential Equation (ELDE) we have:

$$\begin{aligned}
&1. \quad \frac{d}{dt}\left(\frac{\partial G}{\partial \dot{x}}\right) = \frac{\partial G}{\partial x} \Rightarrow -\dot{\lambda} = L_x + \lambda f_x \\
&2. \quad \frac{d}{dt}\left(\frac{\partial G}{\partial \dot{u}}\right) = \frac{\partial G}{\partial u} \Rightarrow 0 = L_u + \lambda f_u \\
&3. \quad \frac{d}{dt}\left(\frac{\partial G}{\partial \dot{\lambda}}\right) = \frac{\partial G}{\partial \lambda} \Rightarrow 0 = f(t,x,u) - \dot{x}
\end{aligned}$$

These three conditions are necessary for (x^*, u^*, λ^*) to be optimal for G.
Now we define the Hamiltonian of the problem;

Definition 2.1.1. *Hamiltonian function* $\mathbb{H} : \Re^n \times \Re^m \times \Re^n \times [t_0, T] \to \Re$,
$\mathbb{H}(x(t), u(t), \lambda(t), t) = L(x(t), u(t), t) + \lambda^T(t) f(x(t), u(t), t)$

Remark 2.1.1. *In terms of the Hamiltonian, the above necessary conditions for optimality can be written as:*

1. *The Adjoint Condition:* $\dot{\lambda}(t) = -\frac{\partial \mathbb{H}}{\partial x}, \qquad \dot{\lambda}(t) = -L_x - \lambda f_x$
2. *The Optimality Condition:* $\frac{\partial \mathbb{H}}{\partial u} = 0, \qquad L_u + \lambda f_u = 0$
3. *The State Condition:* $\dot{x} = f(t,x,u)$ *with the given ending point Conditions,*
   $x(t_0) = x_0$ *and* $x(T) = x_T$.

From the above Remark 2.1.1 and Theorem 2.1.1 , one can conclude that, if a pair (x^*, u^*) is optimal for the optimal control problem, then the following conditions hold at the point (x^*, u^*) for some $C^{(1)}$ function $\lambda^*(t)$.

1. $\dot{\lambda}^*(t) = -L_x(t, x^*, u^*) - \lambda f_x(t, x^*, u^*)$
2. $0 = L_u(t, x^*, u^*) + \lambda^* f_u(t, x^*, u^*)$
3. $\dot{x}^* = f(t, x^*, u^*) \qquad x^*(t_0) = x_0 \quad , x^*(T) = x_T.$

2.2 Pontryagin's Minimum Principle

Optimal control problem with free final state

The optimal control problem with free final state is given by;

$$\min_{u\in U} J(x,u) = \psi(x(T),T) + \int_0^T L(t,x,u)dt$$
$$\text{s.t} \qquad \dot{x} = f(t,x,u), \quad x(0)=x_0, \quad x(T) \quad \text{is free} \tag{2.5}$$

Given an optimal control problem (OCP) with fixed or free ending point, we can formulate the following neccesary conditions for optimality.

Theorem 2.2.1. *If the pair (x^*, u^*) is optimal for the optimal control problem given above, then there exists a function $\lambda(t)$ such that the following conditions are satisfied:*

(i). $\dot{\lambda}(t) = -\frac{\partial \mathbb{H}}{\partial x}(t, x^*(t), u^*(t), \lambda(t)) = -L_x - \lambda f_x \quad \longrightarrow$ *The Adjoint Condition*

(ii). $\frac{\partial \mathbb{H}}{\partial u}(t, x^*(t), u^*(t), \lambda(t)) = L_u + \lambda f_u = 0 \quad \longrightarrow$ *The Optimality Condition*

If the choice of u is constrained the more general form of this condition is given by,

$\mathbb{H}(t, x^*(t), u^*(t), \lambda(t)) = \min_{u\in\Omega} \mathbb{H}(t, x(t), u(t), \lambda(t))$ *where Ω is the constraint set for u*

(iii). $\dot{x}^*(t) = \frac{\partial \mathbb{H}}{\partial \lambda}(t, x^*(t), u^*(t), \lambda(t)) = f(t, x^*(t), u^*(t)) \quad \longrightarrow$ *The State Condition*

(iv). $x(t_0) = x_0, \quad x(T) = x_T \longrightarrow$ *boundary conditions or,*

$\lambda(T) = \psi_x(x^*(T), T) \longrightarrow$ *Transversality condition*

Proof. Assume that there is a piecewise continuous optimal control u^* with corresponding state: $x^* = x(u^*)$.
Let $h(t)$ be a piecewise continious variation function and $\epsilon \in \Re$ be a constant then define

$$u_\epsilon(t) = u^*(t) + \epsilon h(t)$$

Let $x_\epsilon(t)$ be the state that corresponds to $u_\epsilon(t)$

$$\Rightarrow \frac{d}{dt} x_\epsilon(t) = f(t, x_\epsilon(t), u_\epsilon(t))$$

Whenever $u_\epsilon(t)$ is continuous and such that $x_\epsilon(0) = x_0$ then

$$\lim_{\epsilon\to 0} u_\epsilon(t) = u^*(t)$$

for all t and

$$\frac{\partial}{\partial \epsilon} u_\epsilon(t)\Big|_{\epsilon=0} = h(t).$$

Therefore, it follows that

$$\Rightarrow x_\epsilon(t) \to x^*(t) \qquad \text{as} \quad \epsilon \to 0 \quad \text{for each } t.$$

And assume that $\frac{\partial}{\partial\epsilon}x_\epsilon(t)\big|_{\epsilon=0}$ exists for each t.
Then, with respect to the variation function $h(t)$ and the parameter ϵ the objective function becomes:

$$J(x_\epsilon, u_\epsilon) = \psi(x_\epsilon(T), T) + \int_0^T L(t, x_\epsilon(t), u_\epsilon(t)dt$$

Let $\lambda(t)$ be a piecewise differentiable function defined on $[0, T]$ whose actual expression is to be determined. Then by the fundamental theorem of calculus we have;

$$\begin{aligned} \int_0^T \frac{d}{dt}(\lambda(t)x_\epsilon(t))dt &= \lambda(T)x_\epsilon(T) - \lambda(0)x_\epsilon(0) \\ \Rightarrow \int_0^T \frac{d}{dt}(\lambda(t)x_\epsilon(t))dt &- \lambda(T)x_\epsilon(T) + \lambda(0)x_\epsilon(0) = 0 \end{aligned}$$

Then

$$\begin{aligned} J(x_\epsilon, u_\epsilon) &= \psi(x_\epsilon(T), T) + \int_0^T \left[L(t, x_\epsilon(t), u_\epsilon(t)) + \frac{d}{dt}\left(\lambda(t)x_\epsilon(t)\right)\right] dt - \lambda(T)x_\epsilon(T) \\ &\quad + \lambda(0)x_\epsilon(0) \\ &= \psi(x_\epsilon(T), T) + \int_0^T \left[L(t, x_\epsilon(t), u_\epsilon(t)) + \lambda'(t)x_\epsilon + \lambda(t)f(t, x_\epsilon, u_\epsilon)\right] dt - \lambda(T)x_\epsilon(T) \\ &\quad + \lambda(0)x_\epsilon(0) \end{aligned}$$

Since u^* is aminimizer for J , we have

$$0 = \frac{d}{d\epsilon}J(x_\epsilon, u_\epsilon)|_{\epsilon=0} = \lim_{\epsilon\to 0}\frac{J(x_\epsilon, u_\epsilon) - J(x^*, u^*)}{\epsilon}$$

Applying the Lebesgue Dominance integral theorem we can move the limit (or derivative) inside the integral .

$$\begin{aligned} 0 &= \frac{d}{d\epsilon}J(x^*, u^*)|_{\epsilon=0} \\ &= \frac{\partial}{\partial\epsilon}\left(\psi(x_\epsilon(T), T)\right)|_{\epsilon=0} + \int_0^T \frac{\partial}{\partial\epsilon}\left[L(t, x_\epsilon, u_\epsilon) + \lambda'(t)x_\epsilon + \lambda(t)f(t, x_\epsilon, u_\epsilon)\right] dt|_{\epsilon=0} \\ &\quad - \frac{\partial}{\partial\epsilon}\left(\lambda(T)x_\epsilon(T)\right)|_{\epsilon=0} + 0 \\ &= \int_0^T \left[L_x\frac{\partial x_\epsilon}{\partial\epsilon} + L_u\frac{\partial u_\epsilon}{\partial\epsilon} + \lambda'\frac{\partial x_\epsilon}{\partial\epsilon} + \lambda\left(f_x\frac{\partial x_\epsilon}{\partial\epsilon} + f_u\frac{\partial u_\epsilon}{\partial\epsilon}\right)\right] dt|_{\epsilon=0} \\ &\quad \psi_{x_\epsilon}(x_\epsilon(T), T)\frac{\partial x_\epsilon(T)}{\partial\epsilon} - \lambda(T)\frac{\partial x_\epsilon(T)}{\partial\epsilon}|_{\epsilon=0} \\ &= \int_0^T \left[L_x\frac{\partial x_\epsilon}{\partial\epsilon} + L_u\frac{\partial u_\epsilon}{\partial\epsilon} + \lambda'\frac{\partial x_\epsilon}{\partial\epsilon} + \lambda\left(f_x\frac{\partial x_\epsilon}{\partial\epsilon} + f_u\frac{\partial u_\epsilon}{\partial\epsilon}\right)\right] dt|_{\epsilon=0} \\ &\quad \left[\left(\psi_{x_\epsilon}(x_\epsilon(T), T) - \lambda(T)\right)\frac{\partial x_\epsilon(T)}{\partial\epsilon}\right]|_{\epsilon=0} \quad \text{since} \quad \frac{\partial u_\epsilon}{\partial\epsilon} = h \quad \text{and then} \\ &= \int_0^T \left[\left(L_x + \lambda' + \lambda f_x\right)\frac{\partial x_\epsilon}{\partial\epsilon} + \left(L_u + \lambda f_u\right)h\right] dt|_{\epsilon=0} - \left[\left(\psi_{x_\epsilon}(x_\epsilon(T), T) - \lambda(T)\right)\frac{\partial x_\epsilon(T)}{\partial\epsilon}\right]|_{\epsilon=0} \end{aligned}$$

Now we want to choose the adjoint function $\lambda(t)$ so that the coefficients of $\frac{\partial x_\epsilon}{\partial \epsilon}$ vanish
Hence $\lambda(t)$ satisfy

$$\begin{aligned}\lambda'(t) &= -\left[L_x(t,x^*,u^*) + \lambda f_x(t,x^*,u^*)\right] \longrightarrow \textbf{Adjoint condition} \\ \lambda(T) &= \psi_x(x^*(T),T) \longrightarrow \textbf{Transversality condition}\end{aligned}$$

With the above choice of λ we have ,

$$0 = \int_0^T \left[L_u(t,x^*,u^*) + \lambda(t) f_u(t,x^*,u^*)\right] h(t)dt$$

holds for any variational function $h(t)$.
In particular for $h = L_u(t,x^*,u^*) + \lambda(t)f_u(t,x^*,u^*)$ we have

$$0 = \int_0^T \left[L_u(t,x^*,u^*) + \lambda(t) f_u(t,x^*,u^*)\right]^2 dt$$

which implies that

$$L_u(t,x^*,u^*) + \lambda(t) f_u(t,x^*,u^*) = 0 \quad \forall t \in [0,T] \longrightarrow \textbf{Obtimality condition}$$

and

$$\dot{x}^*(t) = f(t,x^*(t),u^*(t)) \qquad \longrightarrow \textbf{The State Condition}$$

Hence, the conditions in the Pontryagin's Mininimum Principle (PMP) are satisfied. □

2.2.1 Sufficient Conditions for Optimality

As described in [9], in general, the conditions stated in the above theorem 2.1.1 (the PMP) are necessary but not sufficient. However, under suitable concavity or convexity condition, it turns out that every pairs(x^*,u^*) satisfying the PMP is optimal. Let us consider the optimal control problem given by:

$$J(x,u) = \min_u \int_{t_0}^T L(x(t),u(t),t)dt \tag{2.6}$$

Where the state variable satisfy the conditions:

$$\begin{aligned}x(t_0) &= x_0 \\ \dot{x}(t) &= f(x(t),u(t),t) \qquad \text{for all} \quad t \in [t_0,T] \\ x(T) &= x_T, \quad \text{or free}\end{aligned}$$

Theorem 2.2.2. *If L and f in the above optimal control problem* (2.6) *are convex functionals with respect to (x,u) for each fixed $t \in [t_0,T]$, then every solution of the optimal control problem* (2.6) *that satisfy the necessary conditions is an optimal solution.*

Proof. Assume that the pair $(x^*; u^*)$ satisfy all the necessary conditions and let (x, u) be any other admissible pair. then

$$J(x,u) - J(x^*,u^*) = \int_{t_0}^{T} [L(x,u,t) - L(x^*,u^*,t)]\, dt \tag{2.7}$$

by convexty of L we have;

$$L(x,u,t) - L(x^*,u^*,t) \;\geq\; \frac{\partial L}{\partial x}(x - x^*) - \frac{\partial L}{\partial u}(u - u^*) \tag{2.8}$$

and from the necessary conditions of optimality we have $\lambda' = -L_x - \lambda f_x \Rightarrow L_x = -(\lambda' + \lambda f_x)$ and $0 = L_u + \lambda f_u \Rightarrow L_u = -\lambda f_u$ with equation (2.8) and (2.7) imply that

$$\begin{aligned}
J(x,u) - J(x^*,u^*) \;&\geq\; \int_{t_0}^{T} \left[\frac{\partial}{\partial x} L(t,x,u)(x - x^*) - \frac{\partial L}{\partial u}(t,x,u)(u - u^*)\right] dt \\
&= \int_{t_0}^{T} \Big[-(\dot{\lambda} + \lambda \frac{\partial f}{\partial x}(t,x^*,u^*))(x - x^*) \\
&\quad -\lambda \frac{\partial f}{\partial u}(t,x^*,u^*)(u - u^*)\Big]\, dt \\
&\quad \text{(by the adjoint and optimalitty conditions of the PMP)} \\
&= -\int_{t_0}^{T} \lambda f_x(t,x^*,u^*)(x - x^*)dt - \int_{t_0}^{T} \dot{\lambda}(x - x^*)dt \\
&\quad -\int_{t_0}^{T} \lambda f_u(t,x^*,u^*)(u - u^*)dt \\
&= -\int_{t_0}^{T} \lambda f_x(x - x^*)dt - [\lambda x]_{t_0}^{T} + \int_{t_0}^{T} \lambda \dot{x} dt + [\lambda x^*]_{t_0}^{T} \\
&\quad -\int_{t_0}^{T} \dot{\lambda} x^* dt - \int_{t_0}^{T} \lambda f_u(u - u^*)dt \\
&\quad -\int_{t_0}^{T} \lambda \left[f^* - f + f_x(x - x^*) + f_u(u - u^*)\right] dt
\end{aligned}$$

$$\begin{aligned}
&= -\int_{t_0}^{T} \lambda \left[f(t,x,u) - f(t,x^*,u^*)\right] dt - \int_{t_0}^{T} \lambda \left[f_x(t,x^*,u^*)(x - x^*)\right. \\
&\quad \left. + f_u(t,x^*,u^*)(u - u^*)\right] dt \quad \text{and by convexity of f} \\
&\geq -\int_{t_0}^{T} \lambda \left[f(t,x^*,u^*) - f(t,x,u)\right] dt - \int_{t_0}^{T} \lambda \left[f(t,x,u) - f(t,x^*,u^*)\right] dt \\
&= 0
\end{aligned}$$

Hence we have

$$J(x,u) - J(x^*,u^*) \geq 0 \Rightarrow J(x^*,u^*) \leq J(x,u)$$

Therefore, (x^*, u^*) is a minimizer of the Optimal Control Problem. $\square$

2.2.2 Principle of Dynamic Programming and Hamilton-Jacobi-Bellman equation

The basic principle of dynamic programming, called the principle of optimality, can be put as follows. An optimal path has the property that whatever the initial conditions and control values over some initial period, the control (or decision variables) over the remaining period must be optimal for the remaining problem, with the state resulting from the early decisions considered as the initial condition. [10]
Let us consider the maximization for of optimal control problem:

$$maximize : J(u, t_0, x_0) = \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))dt. \tag{2.9}$$

with,

$$\begin{aligned} \dot{x} &= f(t, x, u) \\ x(t_0) &= x_0 \end{aligned}$$

Where $x \in \Re^n$ is the state variable, the control function $u(\cdot)$ is assumed to be measurable, taking values in a compact domain $U \to \Re^m$.
Here ψ describes a terminal payoff, while $L(\cdot)$ is a running cost. For a given initial data (t_0, x_0), the payoff J should be maximized over all measurable control functions $u : [t_0, T] \to U$.
Let for each $(t_0, x_0) \in [t_0, T] \times \Re^n$

$$V(t_0, x_0) = \sup_{u:[t_0,T]\to U} J(u : t_0, x_0) \tag{2.10}$$

be the maximum payoff that can be achieved starting from the state x_0 at time t_0. The function V in (2.10) is called the value function for the optimization problem (2.9).
When t = T we have;

$$V(T, x) = \psi(x(T)) \tag{2.11}$$

A basic property of this value function is given in the following theorem

Theorem 2.2.3. ***(Principle of Dynamic Programming).*** *For any initial data $x_0 \in \Re$ and $0 \leq t_0 \leq t_1 \leq T$, one has*

$$V(t_0, x_0) = \sup_{u:[t_0,T]\to U} \left\{ V(t_1, x(t_1; t_0, x_0, u) - \int_{t_0}^{t_1} L\left(t, x(t, t_0, x_0, u(t)), u(t)\right) dt \right\} \tag{2.12}$$

In other words (see figure below),the optimization problem on the time interval $[t_0, T]$ can be split into two separate problems:

- As a first step, we solve the optimization problem on the sub-interval $[t_1, T]$, with running cost L and terminal payoff ψ. In this way, we determine the value function $V(t_1, \cdot)$, at time t_1.
- As a second step, we solve the optimization problem on the sub-interval $[t_0, t_1]$, with running cost L and terminal payoff $V(t_1, \cdot)$, determined by the first step.

At the initial time t_0, according to (2.12) the value function $V(t_0, \cdot)$ obtained in step 2 is the same as the value function corresponding to the global optimization problem over the whole interval $[t_0, T]$.

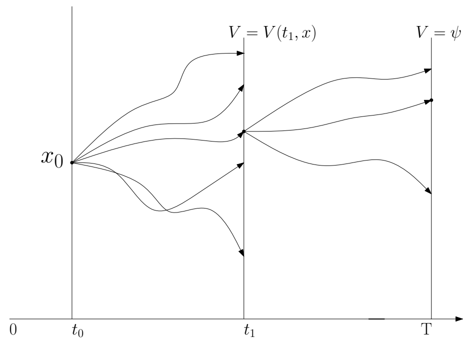

Figure 2.1: The dynamic programming principle.

Proof. Let call J_0 the right hand side of (2.12).

i. To prove that $V(t_0, x_0) \leq J_0$, fix $\epsilon \geq 0$, choose a control $u : [t_0, T] \to U$ such that

$$J(u; t_0, x_0) \geq V(t_0, x_0) - \epsilon,$$

and call $x_1 = x(t_1; t_0, x_0, u)$. Observing that $x(t; t_0, x_0, u) = x(t; t_1, x_1, u)$, we obtain

$$\begin{aligned} V(t_0, x_0) - \epsilon &\leq J(u; t_0, x_0) = J(u; t_1, x_1) - \int_{t_0}^{t_1} L\left(t, x(t, t_0, x_0, u), u(t)\right) dt \\ &\leq V(t_1, x_1) - \int_{t_0}^{t_1} L\left(t, x(t, t_0, x_0, u), u(t)\right) dt \leq J_0 \end{aligned}$$

Since $\epsilon \geq 0$ is arbitrary, the first inequality is proved.

ii. To prove that $V(t_0, x_0) \geq J_0$, fix $\epsilon \geq 0$, then there exists a control $u : [t_0, t_1] \to U$ such that

$$V\left(t_1, x(t_1, t_0, x_0, u_0)\right) - \int_{t_0}^{t_1} L\left(t, x(t, t_0, x_0, u_0), u_0(t)\right) dt \geq J_0 - \epsilon. \tag{2.13}$$

call $x_1 = x(t_1; t_0, x_0, u_0)$. We can now find a control $u_1 : [t_1, T] \to U$ such that

$$J(u_1, t_1, x_1) \geq V(t_1, x_1) - \epsilon \tag{2.14}$$

Consider the new control $u : [t_0, T] \to U$ defined as the concatenation of u_0, u_1 :

$$u(t) = \begin{cases} u_0(t) & if \quad t \in [t_0, t_1] \\ u_1(t) & if \quad t \in [t_1, T] \end{cases}$$

By (2.13) and (2.14) it now follows

$$V(t_0, x_0) \geq J(u, t_0, x_0) = J(u_1, t_1, x_1) - \int_{t_0}^{t_1} L\left(t, x(t, t_0, x_0, u_0), u_0\right) dt \geq J_0 - 2\epsilon.$$

Since $\epsilon \geq 0$ can be arbitrarily small,the second inequality is also proved.

$\square$

Hamilton-Jacobi-Bellman (HJB) equation of dynamic programming

Consider the **Ocp:**

$$maximize : J(x(t), u) = \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))dt.$$

subjected to:

$$\begin{aligned} \dot{x} &= f(t, x, u) \\ x(t_0) &= x_0 \end{aligned}$$

Suppose that a value function or a maximum cost function $V(x(\cdot), \cdot)$ exists for all admissible state $x(t), \quad t \leq T$ i.e

$$V(x(t), t) = \max_{u(s): t \leq s \leq T} \left\{ \psi(x(T)) - \int_{t}^{T} L(s, x(s), u(s))ds \right\}.$$

for all $s \geq t$

$$\frac{dx}{ds} = f(s, x(s), u(s)), \qquad x(t) = x.$$

To apply Belleman's principle for the discrete case, we need to discritize the time interval $[t, T]$ in to $[t, t + \triangle t]$, and $[t + \triangle t, T]$ and we obtain

$$V(x(t), t) = \max_{u} \left\{ - \int_{t}^{t+\triangle t} L(t, x(t), u(t))dt + V(x(t + \triangle t), t + \triangle t) \right\}.$$

where $V(x(t+\triangle t), t+\triangle t)$ is the maximum cost of the process for the time interval $[t+\triangle t, T]$ with initial state $x(t+\triangle t)$.
Assume that $V(x(t+\triangle t), t+\triangle t)$ is twice partially differentiable function. Then the Taylor serious expansion of this function gives us

$$\begin{aligned} V(x(t),t) &= \max_u \left\{ -\int_t^{t+\triangle t} L(t,x(t),u(t))dt + V(x(t),t) + \frac{\partial V(x(t),t)}{\partial t}\triangle t \right. \\ &\quad \left. + \frac{\partial V(x(t),t)}{\partial x}\left(x(t+\triangle t) - x(t)\right) + \text{terms of higher order} \right\}. \end{aligned}$$

Now, for small $\triangle t$ we have

$$\begin{aligned} V(x(t),t) &= \max_u \{ -L(t,x(t),u(t))\triangle t + V(x(t),t) + V_t(x(t),t)\triangle t \\ &\quad + V_x(x(t),t)\cdot f(t,x,u)\triangle t + O(\triangle t) \}. \end{aligned}$$

$$\begin{aligned} \Rightarrow V(x(t),t) - V(x(t),t) &= \max_u \{ [-L(t,x(t),u(t)) + V_t(x(t),t) \\ &\quad + V_x(x(t),t)\cdot f(t,x,u)]\triangle t + O(\triangle t) \} \\ \Rightarrow 0 &= \max_u \{ [-L(t,x(t),u(t)) + V_t(x(t),t) \\ &\quad + V_x(x(t),t)\cdot f(t,x,u)]\triangle t + O(\triangle t) \} \\ &= V_t(x(t),t)\triangle t + \max_u \{ [-L(t,x(t),u(t)) \\ &\quad + V_x(x(t),t)\cdot f(t,x,u)]\triangle t + O(\triangle t) \} \end{aligned}$$

Dividing both sides of the above equations by $\triangle t$ and taking limits as
$\triangle t \to 0$ we get;

$$0 = V_t(x(t),t) + \max_u \{ -L(t,x(t),u(t)) \tag{2.15}$$

$$+ V_x(x(t),t)\cdot f(t,x,u) \} \qquad \text{since} \quad \lim_{\triangle t\to 0} \frac{O(\triangle t)}{\triangle t} = 0 \tag{2.16}$$

Moreover, we have

$$V(x(T),T) = \psi(x(T))$$

We denote the marginal cost vector along the optimal trajoctory x^* by the adjoint vector $\lambda(t)$.
Thus we have;

$$\lambda(t) = V_x(x^*(t),t)$$

We define the Hamiltonian;

$$\mathbb{H}(t,x,u,\lambda) := -L(t,x,u) + V_x(x^*(t),t)\cdot f(t,x,u) \quad \text{where} \quad \lambda(t) = V_x(x^*(t),t) \qquad \text{and}$$

$$\mathbb{H}(t, x^*, u^*, \lambda) = \max_u \mathbb{H}(t, x, u, \lambda)$$

Then

$$\begin{aligned} V_t(x(t), t) &= -\mathbb{H}(t, x^*, u^*, \lambda) = \max_u\{-L(t, x, u) + V_x(x^*(t), t) \cdot f(t, x, u)\} \\ \Rightarrow -V_t(x(t), t) &= \max_u\{-L(t, x, u) + V_x(x^*(t), t) \cdot f(t, x, u)\} \qquad (2.17) \\ \text{where} \qquad \lambda(t) &= V_x(x^*(t), t) \end{aligned}$$

The above equation (2.17) is called the **Hamiltonian Jacobi Beilman's (HJB)** equation.

Theorem 2.2.4. ***(Sufficient conditions for optimality).*** *Let $W : [0, T] \times \Re^n \to \Re$ be a C^1 solution of the terminal value problem*

$$W_t + \mathbb{H}(t, x, W_x) = 0 \qquad W(T, x) = \psi(x) \qquad (2.18)$$

Then W coincides with the value function. In particular, for any given initial data (t_0, x_0), a control $u^*(\cdot)$ that achieves the payoff $J(u^*, t_0, x_0) = W(t_0, x_0)$ is optimal.

Proof. Let $V = V(t_0, x_0)$ be the value function for the optimal control problem (2.9).

1. We first show that $V \leq W$
   Given an initial data (t_0, x_0), consider any control $u : [t_0, T] \to U$, and let $t \to x(t) = x(t; t_0, x_0, u)$ be the corresponding trajectory. We claim that

$$J(u, t_0, x_0) = \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))dt \leq W(t_0, x_0) \qquad (2.19)$$

   Observe that

$$\begin{aligned} &\frac{d}{dt}\left[W(t, x(t)) - \int_{t_0}^{t} L(s, x(s), u(s)))ds\right] \\ &\quad = W_t(t, x(t)) + W_x \cdot f(t, x(t), \omega) - L(t, x(t), u(t)) \\ &\quad \leq W_t(t, x(t)) + \max_{\omega \in U} \{W_x \cdot f(t, x(t), \omega) - L(t, x(t), \omega)\} = 0 \end{aligned}$$

   Integrating over the time interval $[t_0, T]$ we obtains

$$\begin{aligned} W(t_0, x_0) &\geq W(T, x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))dt \\ &= \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))dt = J(u, t_0, x_0). \qquad (2.20) \end{aligned}$$

   Therefore,

$$V(t_0, x_0) = \sup_{u(\cdot)} J(u, t_0, x_0) \leq W(t_0, x_0)$$

   By assumption, the control $u^*(\cdot)$ yields the payoff $J(u^*) = W(t_0, x_0)$. On the other hand, (2.20) shows that any other control $u(\cdot)$ yields a payoff $J(u) \leq W(t_0, x_0)$. Therefore, u^* is optimal.

2. We shall prove the opposite inequality , assuming that the map

$$(t,x) \to u^{\#}(t,x) = arg \max_{u\in U} \{W_x(t,x)\cdot f(t,x,u) - L(t,x,u)\}$$

is uniquely defined and continuous. In this case, given the initial condition (t_0, x_0), let $t \leftarrow x^*(t)$ be a solution to the Cauchy problem

$$\dot{x} = f(t,x,u^{\#}(t,x)) \qquad x(t_0) = x_0$$

Notice that, by our assumption, the above ODE has a continuous right hand side. Hence a solution exists. Calling $u^*(t) = u^{\#}(t, x^*(t))$, we have

$$\begin{aligned}
&\frac{d}{dt}\left[W(t,x^*(t)) - \int_{t_0}^{t} L(s,x^*(s),u^*(s)))ds\right] = W_t(t,x^*(t))\\
&\quad +W_x\cdot f\left(t,x^*(t),u^{\#}(t,x^*(t))\right) - L\left(t,x^*(t),u^{\#}(t,x^*(t))\right)\\
&\quad = W_t(t,x(t)) + \max_{\omega\in U}\{W_x\cdot f(t,x^*(t),\omega) - L(t,x^*(t),\omega)\} = 0
\end{aligned}$$

Integrating over the time interval $[t_0, T]$ we obtains

$$\begin{aligned}
W(t_0,x_0) &= W(T,x^*(T)) - \int_{t_0}^{T} L(t,x^*(t),u^*(t))dt\\
&= \psi(x^*(T)) - \int_{t_0}^{T} L(t,x^*(t),u^*(t))dt\\
&= J(u^*,t_0,x_0) \le V(t_0,x_0).
\end{aligned}$$

The above equation implies that

$$W \le V$$

Therefore, from 1 and 2 above we can conclude that,
$W = V$or W coincides with the value function. □

2.3 Optimal control with discounted functional objective

Sometimes the integrand function is multiplied by a discount factor. In particular, this may be useful to ensure the convergence of the integral in the infinite horizon problems. Therefore, the optimal control of infinite horizon with discount rate r: is given by

$$\max \int_{t_0}^{\infty} e^{-rt} L(x,u)dt \tag{2.21}$$

$$\text{subjected to} \qquad x' = f(x,u), \quad x(t_0) = x_0. \tag{2.22}$$

In terms of the Hamiltonian

$$\mathbb{H} = e^{-rt}L(x,u) + \lambda f(x,u), \tag{2.23}$$

we require (x, u, λ) to satisfy

$$\mathbb{H}_u = e^{-rt}L_u + \lambda f_u \tag{2.24}$$

$$\lambda' = -\mathbb{H}_x = -e^{-rt}L_x - \lambda f_x, \qquad \lim_{t\to\infty} \lambda(t) = 0. \tag{2.25}$$

Now from equation (2.23) we have

$$\mathbb{H} = e^{-rt}L(x,u) + \lambda f(x,u) = e^{-rt}(L(x,u) + e^{rt}\lambda f(x,u) \tag{2.26}$$

and define

$$m(t) = e^{rt}\lambda(t) \tag{2.27}$$

where $m(t)$ is the current value multiplier associated with (2.22) and $e^{-rt}m(t) \to 0$ as $t \to \infty$ Let

$$\mathbf{H} = e^{rt}H = L(x,u) + mf(x,u) \tag{2.28}$$

We call $\mathbf{H}$ the current value Hamiltonian. Differentiating (2.27) with respect to time gives

$$\begin{aligned} m' &= re^{rt}\lambda(t) + \lambda' e^{rt} \\ &= rm - e^{rt}\mathbb{H}_x \end{aligned} \tag{2.29}$$

on substituting from (2.27) and (2.25). In view of (2.28), $\mathbb{H} = e^{-rt}\mathbf{H}$, so (2.29) becomes

$$\begin{aligned} m' &= rm - e^{rt}\frac{\partial(e^{-rt}\mathbf{H})}{\partial x} \\ &= rm - e^{rt}e^{-rt}\mathbf{H}_x \\ &= rm - L_x - mf_x \end{aligned} \tag{2.30}$$

In addition, (2.24) can be written as

$$\mathbb{H}_u = \frac{\partial(e^{-rt}\mathbf{H})}{\partial u} = e^{-rt}\frac{\partial \mathbf{H}}{\partial u} = 0$$

which implies

$$\frac{\partial \mathbf{H}}{\partial u} = 0 \tag{2.31}$$

Finally, (2.22) may be recovered in terms of the current value Hamiltonian:

$$x' = \frac{\partial \mathbf{H}}{\partial m} = f \tag{2.32}$$

therefore equation (2.23)-(2.25) can be equivalently stated as

$$\mathbf{H} = L(x,u) + mf(x,u), \tag{2.33}$$

$$\frac{\partial \mathbf{H}}{\partial u} = L_u + mf_u = 0 \tag{2.34}$$

$$m' = rm - \frac{\partial \mathbf{H}}{\partial x} = rm - L_x - mf_x. \tag{2.35}$$

2.3.1 Hamilton-Jacobi-Bellman (HJB) equation with discounted functional objective

To form of the optimality conditions for infinite horizon optimal control problems, let us consider the following **OCP**

$$\begin{aligned} \max \int_{t_0}^{\infty} & e^{-rt} L(x,u)dt \\ \text{subjected to;} \quad \dot{x} &= f(x,u) \\ x(t_0) &= x_0 \end{aligned} \tag{2.36}$$

Let $J(t_0, x_0)$ be the maximum value at time $t = t_0$ and $x(t_0) = x_0$
Hence,

$$\begin{aligned} J(t_0, x_0) &= \max \int_{t_0}^{\infty} e^{-rt} L(x,u)dt \\ &= e^{-rt_0} \max \int_{t_0}^{\infty} e^{-r(t-t_0)} L(x,u)dt \end{aligned}$$

The value of the integral on the right depends on the initial state, but is independent of the initial time, i.e. it only depends on elapsed time. Now let

$$V(t_0, x_0) = \max \int_{t_0}^{\infty} e^{-r(t-t_0)} L(x,u)dt$$

Then

$$\begin{aligned} J(t,x) &= e^{-rt} V(x) \\ \Rightarrow J_t &= -re^{-rt} V(x) \quad \text{and} \\ J_x &= e^{-rt} V_x(x) \end{aligned}$$

By using equation (2.15) we get;

$$\begin{aligned} 0 &= J_t(x(t),t) + \max_u \left\{ e^{-rt} L(x(t),u(t)) + J_x(x(t),t) \cdot f(t,x,u) \right\} \\ \Rightarrow -J_t(x(t),t) &= \max_u \left\{ e^{-rt} L(t,x(t),u(t)) + J_x(x(t),t) \cdot f(x,u) \right\} \\ \Rightarrow \quad -(-re^{-rt} V(x)) &= \max_u \left\{ e^{-rt} L(x(t),u(t)) + e^{-rt} V_x(x) \cdot f(x,u) \right\} \\ \Rightarrow \quad rV(x) &= \max_u \left\{ L(x(t),u(t)) + V_x(x) \cdot f(x,u) \right\} \end{aligned}$$

Therefore ,

$$rV(x) = \max_u \left\{ L(x(t),u(t)) + V_x(x) \cdot f(x,u) \right\}$$

obeyed by the optimal current value function V(x) associated with problem (2.36)

Chapter 3

Static Games

As described in [10] , in a basic form, a game for two players, say "Player A" and "Player B", is given by:

- The two sets of strategies: A and B, available to the players.
- The two payoff functions: $\phi^A : A \times B \mapsto \Re$ and $\phi^B : A \times B \mapsto \Re$.

If the first player chooses a strategy $a \in A$ and the second player chooses $b \in B$, then the payoffs achieved by the two players are $\phi^A(a, b)$ and $\phi^B(a, b)$, respectively. The goal of each player is to maximize his own payoff. We shall always assume that each player has full knowledge of both payoff functions ϕ^A, ϕ^B, but he may not know in advance the strategy adopted by the other player.
If $\phi^A(a, b) + \phi^B(a, b) = 0$ for every pair of strategies (a, b), the game is called a zero sum game.
A zero-sum game is determined by one single payoff function $\phi = \phi^A = -\phi^B$.

Throughout the following, our basic assumption will be:
Assumption
The sets A and B are compact metric spaces. The payoff functions ϕ^A, ϕ^B are continuous functions from $A \times B$ into $\mathbb{R}$.

The simplest class of games consists of **bi-matrix games**, where each player has a finite set of strategies to choose from. Say,

$$A = \{a_1, a_2, \cdots, a_m\}, \qquad B = \{b_1, b_2, \cdots, b_n\}. \tag{3.1}$$

In this case, each payoff function is determined by its $m \times n$ values

$$\phi^A_{ij} = \phi^A(a_i, b_j) \qquad \phi^B_{ij} = \phi^B(a_i, b_j) \tag{3.2}$$

Of course, these numbers can be written as the entries of two $m \times n$ matrices. The game can also be conveniently represented by an $m \times n$ "bi-matrix", where each entry consists of the pair of two numbers: ϕ^A_{ij}, ϕ^B_{ij} as shown in table 3.1.

		Player B				
	Strategies	b_1	b_2	b_3	$\cdots$	b_n
	a_1	$(\phi^A_{11}, \phi^B_{11})$	$(\phi^A_{12}, \phi^B_{12})$	$(\phi^A_{13}, \phi^B_{13})$	$\cdots$	$(\phi^A_{1n}, \phi^B_{1n})$
Player A	a_2	$(\phi^A_{21}, \phi^B_{21})$	$(\phi^A_{22}, \phi^B_{22})$	$(\phi^A_{23}, \phi^B_{23})$	$\cdots$	$(\phi^A_{2n}, \phi^B_{2n})$
	a_3	$(\phi^A_{31}, \phi^B_{31})$	$(\phi^A_{32}, \phi^B_{32})$	$(\phi^A_{33}, \phi^B_{33})$	$\cdots$	$(\phi^A_{3n}, \phi^B_{3n})$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	a_m	$(\phi^A_{m1}, \phi^B_{m1})$	$(\phi^A_{m2}, \phi^B_{m2})$	$(\phi^A_{m3}, \phi^B_{m3})$	$\cdots$	$(\phi^A_{mn}, \phi^B_{mn})$

Table 3.1: The Bi-matrix payoff representation for two players with finite strategies

3.1 Solution concepts

In general, we cannot speak about the optimal solution of such a game. Indeed, an outcome that is optimal for one player can be very bad for the other one. Therefore, we are expected to see various concepts of solutions. These can provide appropriate models in specific situations, depending on the information available to the players and on their willingness to cooperate.

I. Pareto optimality.

Definition 3.1.1. *A pair of strategies (a^*, b^*) is said to be* ***Pareto optimal*** *if there exists no other pair $(a, b) \in A \times B$ such that*

$$\phi^A(a,b) > \phi^A(a^*,b^*) \qquad \textit{and} \qquad \phi^B(a,b) \geq \phi^B(a^*,b^*).$$

or

$$\phi^B(a,b) > \phi^B(a^*,b^*) \qquad \textit{and} \qquad \phi^A(a,b) \geq \phi^A(a^*,b^*).$$

In general, a game can admit several Pareto optima. In order to construct a pair of strategies which is Pareto optimal, choose any number $\lambda \in (0,1)$ and consider the optimization problem

$$\max_{(a,b)\in A\times B} \lambda\phi^A(a,b) + (1-\lambda)\phi^B(a,b). \tag{3.3}$$

since the sets A and B are compact metric spaces and the payoff functions ϕ^A, ϕ^B are continuous functions from $A \times B$ into $\Re$, an optimal solution exist.
Any pair (a^*, b^*) where the maximum is attained yields a Pareto optimum.

Further concepts of solution can be formulated in terms of the best reply maps. For a given choice $b \in B$ of player B, consider the set of best possible replies of player A:

$$R^A(b) = \left\{ a \in A; \quad \phi^A(a,b) = \max_{\omega \in A} \phi^A(\omega, b) \right\}. \tag{3.4}$$

Similarly, for a given choice $a \in A$ of player A, consider the set of best possible replies of player B:

$$R^B(a) = \left\{ b \in B; \quad \phi^B(a,b) = \max_{\omega \in B} \phi^B(a, \omega) \right\}. \tag{3.5}$$

Since the sets A and B are compact metric spaces and the payoff functions $\phi^A, \phi^B$are continuous from $A \times B$ into $\Re$, the above sets are non-empty. However, in general they need not be single-valued. Indeed, our assumptions imply that the maps $a \mapsto R^B(a)$ and $b \mapsto R^A(b)$ are upper semicontinuous, with compact values.

II. Stackelberg equilibrium.

This models a situation with asymmetry of information. We assume that player A (the leader) announces his/her strategy in advance, and then player B (the follower) makes his/her choice accordingly. In this case, the game can be reduced to a pair of nested optimization problems, solved one after the other. In connection with the strategy a adopted by the first player, the second player needs to maximize his/her payoff function $b \mapsto \phi^B(a,b)$. She/He will thus choose a best reply $b^* \in R^B(a)$.
Assuming that this reply is unique, say $b^* = \beta(a)$, the goal of Player A is now to maximize the composite function $a \mapsto \phi^A(a, \beta(a))$.

More generally, we shall adopt the following definition, which does not require uniqueness of the best reply map. In case where player B has several best replies to a value $a \in A$, we take here the optimistic view that She/he will choose the one which is most favorable to Player A.

Definition 3.1.2. *A pair of strategies $(a_S, b_S) \in A \times B$ is called a **Stackelberg equilibrium** if $b_S \in R^B(a_S)$ and*

$$\phi^A(a,b) \le \phi^A(a_S, b_S) \qquad \textit{for every pair } (a,b) \textit{ with } b \in R^B(a).$$

If the sets A and B are compact metric spaces and the payoff functions $\phi^A, \phi^B$are continuous functions from $A \times B$ into $\Re$, then a Stackelberg equilibrium always exists.
Indeed, consider the domain

$$R = \{(a,b); \quad b \in R^B(a)\} \subseteq A \times B.$$

By the compactness of A, B and the continuity of ϕ^B, the set R is closed, hence compact. Therefore, the continuous function ϕ^A attains its global maximum at some point $(a_S, b_S) \in R$.

III. Nash equilibrium.

This models a symmetric situation where the players have no means to cooperate and do not share any information about their strategies.

Definition 3.1.3. *The pair of strategies (a^*, b^*) is a **Nash equilibrium** of the game if, for every $a \in A$ and $b \in B$, one has*

$$\phi^A(a, b^*) \leq \phi^A(a^*, b^*) \qquad \phi^B(a^*, b) \leq \phi^B(a^*, b^*) \tag{3.6}$$

In other words, no player can increase his/her payoff by single-mindedly changing his/her strategy, as long as the other player sticks to one equilibrium strategy. Observe that a pair of strategies (a^*, b^*) is a Nash equilibrium if and only if it is a fixed point of the best reply map:

$$a^* \in R^A(b^*), \qquad b^* \in R^B(a^*).$$

The following examples show that:

(i) In general, a Nash equilibrium may not exist,

(ii) Nash equilibrium need not be unique,

(iii) Different Nash equilibria can yield different payoffs to each player,

(iv) A Nash equilibrium may not be a Pareto optimum.

Example 3.1.1. *Assume that each player draws a coin, choosing to show either head or tail. If the two coins match, player A earns \$1 and player B loses \$1. If the two coins do not match, player B earns \$1 and player A loses \$1. This is a zero-sum game, described by the bi-matrix in the Table below. By direct inspection, one checks that it does not admit any Nash equilibrium solution.*

Player B

Player A

Strategies	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Strategies	H	T
H	1	-1
T	-1	1

Table 3.2: The bi-matrix of the payoffs for the "head and tail" game. Since this is a zero-sum game,it can be represented by a single matrix (right), containing the payoffs for the first player.

Example 3.1.2. *Consider the game whose bi-matrix of payoffs is given in the Table below. The pair of strategies* $(a_1, b_3) = (5, 4)$ *is a Nash equilibrium, as well as a Pareto optimum. On the other hand, the pair of strategies* $(a_2, b_1) = (3, 3)$ *is a Nash equilibrium but not a Pareto optimum. Indeed,* $(a_1, b_3) = (5, 4)$ *is the unique Pareto optimum.*

Player B

Player A

Strategies	b_1	b_2	b_3
a_1	0,0	0,0	5,4
a_2	3,3	0,0	0,0

Table 3.3: A bi-matrix of payoffs, with two Nash equilibrium points and only one Pareto optimum.

Example 3.1.3. ***(prisoners' dilemma).*** *Consider the game with payoffs described by the bimatrix in the Table below. This models a situation where two prisoners are separately interrogated.*
Each one has two options: either (C) *confess and accuse the other prisoner, or* (N) *not confess.*
If anyone of them confesses, the police rewards him by reducing his sentence. None of the prisoners, while interrogated, knows about the behavior of the other. The negative payoffs account for the number of years in jail faced by the two prisoners, depending on their actions. Taking the side of player A, one could argue as follows. If player B confesses, the two options result in either 6 *or* 8 *years in jail, hence confessing is the best choice. On the other hand, if player B does not confess, then the two options result in either* 0 *or* 1 *years in jail. Again, confessing is the best choice. Since player B can argue exactly in the same way, the outcome of the game is that both players confess, and get a* 6 *years sentence is the best option for*

Player B

Player A

Strategies	C	N
C	-6,-6	0,-8
N	-8,0	-1,-1

Table 3.4: The bi-matrix payoffs for the "prisoners' dilemma".

both. In a sense, this is paradoxical because an entirely rational argument results in the worst possible outcome: the total number of years in jail for the two prisoners is maximal.
If they cooperated, they could both have achieved a better outcome, totaling only 2 *years in jail. Observe that the pair of strategies* (C, C) *is the unique Nash equilibrium, but it is not Pareto optimal. On the other hand, all three other pairs* $(C, N), (N, C), (N, N)$ *are Pareto optimal, but not Nash equilibrium.*

Example 3.1.4. *Let* $A = B = [0, 4]$ *and consider the payoff functions*

$$\phi^A(a, b) = 2a + 2b - \frac{a^2}{2}, \qquad \phi^B(a, b) = a + b - \frac{b^2}{2} \tag{3.7}$$

If (a^*, b^*) *is a Nash equilibrium, then*

$$\begin{aligned} a^* &= arg\max_{a\in A}\left\{2a + 2b^* - \frac{a^2}{2}\right\} = 2. \\ b^* &= arg\max_{b\in B}\left\{a^* + b - \frac{b^2}{2}\right\} = 1. \end{aligned}$$

Hence $(2, 1)$ *is the unique Nash equilibrium solution. This is not a Pareto optimum. Indeed,*

$$\phi^A(2, 1) = 4, \qquad \phi^B(2, 1) = \frac{5}{2},$$

while the pair of strategies $(3, 2)$ *yields a strictly better payoff to both players:*

$$\phi^A(3, 2) = \frac{11}{2}, \qquad \phi^B(3, 2) = 3,$$

To find Pareto optimal points, for any $0 < \lambda < 1$ *we consider the optimization problem*

$$\max_{(a,b)\in A\times B}\left\{\lambda\phi^A(a, b) + (1-\lambda)\phi^B(a, b)\right\} = \max_{a,b\in[0,4]}\left\{(\lambda+1)a + (\lambda+1)b - \frac{\lambda a^2 + (1-\lambda)b^2}{2}\right\}.$$

This yields the Pareto optimal point (a^*, b^*)*, with*

$$\begin{aligned} a_\lambda &= arg\max_{a\in[0,4]}\left\{(\lambda+1)a - \frac{\lambda a^2}{2}\right\} = \min\left\{1 + \frac{1}{\lambda}, 4\right\}, \\ b_\lambda &= arg\max_{b\in[0,4]}\left\{(\lambda+1)b - \frac{(1-\lambda)b^2}{2}\right\} = \min\left\{\frac{1+\lambda}{1-\lambda}, 4\right\} \end{aligned}$$

3.2 Existence of Nash equilibria

Theorem 3.2.1. ***(Existence of Nash equilibria).*** *Assume that the sets of strategies A,B are compact, convex subsets of $\Re^n$. Let the payoff functions ϕ^A, ϕ^B be continuous and assume that*

$$\begin{aligned} a &\mapsto \phi^A(a,b) \;\; \textit{is a concave function of } a, \textit{ for each fixed } b \in B, \\ b &\mapsto \phi^B(a,b) \;\; \textit{is a concave function of } b, \textit{ for each fixed } a \in A. \end{aligned}$$

Then the non-cooperative game admits a Nash equilibrium.

Proof. Consider the best reply maps R^A, R^B, defined at (3.4)- (3.5).

1. The compactness of B and the continuity of ϕ^B imply that the function

$$a \mapsto m(a) = \max_{b\in B} \phi^B(a,b)$$

   is continuous. Therefore, the set

$$graph(R^B) = \big\{(a,b); b \in R^B(a)\big\} = \big\{(a,b); \phi^B(a,b) = m(b)\big\}$$

   is closed. Having closed graph, the multifunction $a \mapsto R^B(a) \subseteq B$ is upper semicontinuous

2. We claim that each set $R^B(a) \subseteq B$ is convex. Indeed, let $b_1, b_2 \in R^B(a)$, so that

$$\phi^B(a,b_1) = \phi^B(a,b_2) = m(a)$$

   and let $\theta \in [0,1]$. Using the concavity of the function $b \mapsto \phi^B(a,b)$we obtain

$$m(a) \geq \phi^B(a, \theta b_1 + (1-\theta)b_2) \geq \theta\phi^B(a,b_1) + (1-\theta)\phi^B(a,b_2) = m(a)$$

   Since B is convex, one has $\theta b_1 + (1-\theta)b_2 \in B$. Hence $\theta b_1 + (1-\theta)b_2 \in R^B(a)$, proving our claim.

3. By the previous steps, the multifunction $a \mapsto R^B(a) \subseteq B$ is upper semicontinuous, with compact, convex values. Of course, the same holds for the multifunction $b \mapsto R^A(b) \subseteq A$. We now consider the multifunction on the product space $A \times B$, defined as

$$(a,b) \mapsto R^A(b) \times R^B(a) \subseteq A \subseteq B.$$

   By the previous arguments, this multifunction is upper semicontinuous, with compact convex values. Applying Kakutani's fixed point theorem, we obtain a pair of strategies $(a^*, b^*) \in \big(R^A(b^*), R^B(a^*)\big)$, i.e. a Nash equilibrium solution.

□

3.3 Randomized strategies

If the convexity or concavity assumptions fail, the previous theorem does not apply. Clearly, the above result cannot be used if one of the players can choose among a finite number of strategies.
There are games which do not admit any Nash equilibrium solution.
To achieve a general existence result, one needs to relax the definition of solution, allowing the players to choose randomly among their sets of strategies.

Definition 3.3.1. *A randomized strategy for player A is a probability distribution μ on his/her set of strategies A. Similarly, a randomized strategy for player B is a probability distribution v on the set B.*

Given two randomized strategies μ, v for players A and B respectively, the corresponding payoff functions are defined as

$$\tilde{\phi}^A(\mu, v) = \int_{A\times B} \phi^A(a, b) d\mu \times dv \qquad \tilde{\phi}^B(\mu, v) = \int_{A\times B} \phi^B(a, b) d\mu \times dv \tag{3.8}$$

Remark 3.3.1. *The above quantities $\tilde{\phi}^A(\mu, v)$ and $\tilde{\phi}^B(\mu, v)$ are the* ***expected values of the payoffs****, if the two players choose random strategies, independent of each other, according to the probability distributions μ, v , respectively.*

In the following, by $P(A), P(B)$ we denote the family of all probability measures on the sets A, B, respectively. Notice that to each $a \in A$ there corresponds a unique probability distribution concentrating all the mass at the single point a. This will be called a pure strategy. **Pure strategies** are a subset of all randomized strategies.

Remark 3.3.2. *If $A = a_1, a_2, \ldots, a_m$ is a finite set, a probability distribution on A is uniquely determined by a vector $x = (x_1, \ldots, x_m) \in \triangle_m$, where*

$$\triangle_m = \left\{ x = (x_1, \ldots, x_m); \quad x_i \in [0, 1], \quad \sum_{i=1}^{m} x_i = 1 \right\} \tag{3.9}$$

where x_i is the probability that player A chooses the strategy a_i. Given the bi-matrix game described at (3.1)-(3.2), the corresponding randomized game can be represented as follows. The two players choose from the sets of strategies

$$\tilde{A} = \triangle_m, \qquad \tilde{B} = \triangle_n \tag{3.10}$$

Given probability vectors $x = (x_1, \ldots, x_m) \in \triangle_m$ and $y = (y_1, \ldots, y_n) \in \triangle_n$, the payoff functions are

$$\tilde{\phi}^A(x, y) = \sum_{ij} \phi^A_{ij} x_i y_j, \qquad \tilde{\phi}^B(x, y) = \sum_{ij} \phi^B_{ij} x_i y_j \tag{3.11}$$

Theorem 3.3.1. ***(Existence of Nash equilibria for randomized strategies).*** *If the sets A and B are compact metric spaces and the payoff functions ϕ^A, ϕ^B are continuous functions from $A \times B$ into $\Re$. Then there exist probability measures $\mu^* \in P(A)$ and $v^* \in P(B)$ such that*

$$\tilde{\phi}^A(\mu, v^*) \leq \tilde{\phi}^A(\mu^*, v^*) \qquad \text{for all} \quad \mu \in P(A). \tag{3.12}$$
$$\tilde{\phi}^B(\mu^*, v) \leq \tilde{\phi}^B(\mu^*, v^*) \qquad \text{for all} \quad v \in P(B). \tag{3.13}$$

Proof. The theorem will first be proved for a bi-matrix game, then in the general case.

1. Consider the bi-matrix game described at (3.1)-(3.2). We check that all assumptions of the above theorem are satisfied. The sets of randomized strategies, defined at (3.9)-(3.10), are compact convex simplexes. The payoff functions $\phi^A, \phi^B : \triangle_m \times \triangle_n \mapsto \Re$, defined at (3.11), are bilinear, hence continuous.
   For each given strategy $y \in \triangle_n$ chosen by the second player, the payoff function for the first player
   $$x \mapsto \phi^A(x, y) = \sum_{ij} \phi^A_{ij} x_i y_j$$
   is linear, hence concave. Similarly, for each $x \in \triangle_m$, the payoff function for the second player
   $$y \mapsto \phi^B(x, y) = \sum_{ij} \phi^B_{ij} x_i y_j$$
   is linear, hence concave.
   We can thus apply the above Theorem and obtain the existence of a Nash equilibrium solution $(x^*, y^*) \in \triangle_m \times \triangle_n$.

2. In the remainder of the proof, using an approximation argument we extend the result to the general case whereA, B are compact metric spaces. Let $\{a_1, a_2, \dots\}$ be a sequence of points dense in A, and let $\{b_1, b_2, \dots\}$ be a sequence of points dense in B. For each $n \geq 1$, consider the game with payoffs ϕ^A, ϕ^B but where the players can choose only among the finite sets of strategies $A_n = \{a_1, \dots, a_n\}$ and $B_n = \{b_1, \dots, b_n\}$. By the previous step, this game has a Nash equilibrium solution, given by a pair of randomized strategies (μ_n, v_n). Here μ_n , and v_n are probability distributions supported on the finite sets A_n and B_n, respectively. Since both A and B are compact, by possibly extracting a subsequence we can achieve the weak convergence
   $$\mu_n \to \mu^*, \qquad v_n \to v^* \qquad \text{as,} \;\; n \to \infty \tag{3.14}$$
   for some probability measures $\mu^* 2 P(A)$ and $v \in P(B)$.

3. We claim that the pair (μ^*, v^*) in (3.14) provides a Nash equilibrium solution, i.e. (3.12)- (3.13) hold. This will be proved by showing that

$$\int_{A\times B} \phi^A(a,b) d\mu^* \times dv^* = \max_{\mu \in P(A)} \int_{A\times B} \phi^A(a,b) d\mu \times dv^* \tag{3.15}$$

together with the analogous property for ϕ^B Let $\epsilon > 0$ be given.Since the sets A and B are compact metric spaces and the payoff functions $\phi^A, \phi^B$are continuous from $A \times B$ into $\Re$, then there exists $\delta > 0$ such that

$$d(a,a') \leq \delta \quad \text{and} \quad d(b,b') \leq \delta \quad \text{imply} \quad |\phi^A(a,b) - \phi^A(a',b')| < \epsilon \tag{3.16}$$

Since the sequences $a_k : k \geq 1$ and $b_k : k \geq 1$ are dense in A and B respectively, we can find an integer $N = N(\delta)$ such that the following holds. The set A is covered by the union of the open balls $B(a_i, \delta), i = 1, \ldots, N$, centered at the points a_i with radius $\delta > 0$. Similarly, the set B is covered by the union of the open balls $B(b_j, \delta)$,$j = 1, \ldots, N$, centered at the points b_j with radius $\delta > 0$. Let $\{\varphi_1, \ldots, \varphi_N\}$ be a continuous partition of unity on A, subordinated to the covering $\{B(a_i, \delta); i = 1, \ldots, N\}$, and let $\{\phi_1, \ldots, \phi\}$ be a continuous partition of unity on B,subordinated to the covering $\{B(a_j, \delta); j = 1, \ldots, N\}$ Any probability measure $\mu \in P(A)$ can now be approximated by a probability measure $\hat{\mu}$ supported on the discrete set $A_N = a_1, \ldots, a_N$. This approximation is uniquely defined by setting

$$\hat{\mu}(\{a_i\}) = \int \varphi_i d\mu \qquad i = 1, \ldots i = 1, \ldots N.$$

Similarly, any probability measure $v \in P(B)$ can now be approximated by a probability measure $\hat{v}$ supported on the discrete set $B_N = \{b_1, \ldots, b_N\}$. This approximation is uniquely defined by setting

$$\hat{v}(\{b_j\}) = \int \psi d\mu \qquad i = 1, \ldots i = 1, \ldots N.$$

For every pair of probability measures (μ, v), by (3.16) the above construction yields

$$\begin{aligned} &\left| \int_{A\times B} \phi^A(a,b) d\mu \times dv - \int_{A\times B} \phi^A(a,b) d\hat{\mu} \times d\hat{v} \right| \\ &\leq \int_{A\times B} \sum_{ij} \varphi_i(a_i)\psi_j(b_j) \left| \phi^A(a,b) - \phi^A(a_i, b_j) \right| d\mu \times dv \\ &\leq \int_{A\times B} \epsilon d\mu \times dv \epsilon = \epsilon \end{aligned} \tag{3.17}$$

4. For all $i, j = 1, \ldots, N$, as $n \to \infty$ the weak convergence (3.14) yields

$$\hat{\mu}_n(\{b_j\}) = \int \varphi_i d\mu_n \to \int \varphi_i d\mu^* = \hat{\mu}^*(\{a_i\}) \tag{3.18}$$

Similarly, $\hat{v}_n(\{b_j\}) \to \hat{v}^*(\{b_j\})$

Observe that, for every $\mu \in P(A)$ and $n \geq N$, one has

$$\overline{\phi}^A(\hat{\mu}, v_n) \leq \overline{\phi}^A(\mu_n, v_n). \tag{3.19}$$

Indeed, $\hat{\mu}$ is a probability measure supported on the finite set $A_N = \{a_1, \ldots, a_N\} \subseteq A_n$, and the pair of randomized strategies (μ_n, v_n) provides a Nash equilibrium to the game restricted to $A_n \times B_n$. Using (3.18), (3.18), and (3.19), for every $\mu \in P(A)$ we obtain

$$\begin{aligned}
&\overline{\phi}^A(\mu, v^*) - \epsilon \leq \overline{\phi}^A(\hat{\mu}, \hat{v}^*) = \lim_{n\to\infty} \overline{\phi}^A(\hat{\mu}, \hat{v}) \\
&\leq \limsup_{n\to\infty} \overline{\phi}^A(\hat{\mu}, \hat{v}) + \epsilon \leq \overline{\phi}^A(\mu_n, v_n) + \epsilon = \overline{\phi}^A(\mu^*, v^*) + \epsilon
\end{aligned}$$

Since $\mu \in P(A)$ and $\epsilon > 0$ were arbitrary, this proves (3.12). The proof of (3.13) is entirely similar. □

3.4 Zero-sum games

Consider again a game for two players, with payoff functions $\phi^A, \phi^B : A \times B \mapsto \Re$. In the special case where $\phi^B = -\phi^A$, we have a zero-sum game, described by a single function

$$\phi : A \times B \mapsto \Re. \tag{3.20}$$

Given any couple (a, b) with $a \in A$ and $b \in B$, we think of $\phi(a, b)$ as the amount of money that B pays to A, if these strategies are chosen. The goal of player A is to maximize this payoff, while player B wishes to minimize it. As before, we assume the domains A, B are compact metric spaces and the function $\phi : A \times B \mapsto \Re$ is continuous. In particular, this implies that the maps

$$b \mapsto \max_{a\in A} \phi(a, b), \qquad a \mapsto \max_{b\in B} \phi(a, b) \tag{3.21}$$

are both continuous.

In a symmetric situation, each of the two players will have to make his/her choice without a priori knowledge of the action taken by his opponent. However, one may also consider cases where one player has this advantage of information.

CASE 1: The second player chooses a strategy $b \in B$, then the first player makes his/her choice, depending on b.

This is clearly a situation where player A has the advantage of knowing his/her opponent's strategy.
The best reply of player A will be some $\alpha(b) \in A$ such that

$$\phi(\alpha(b), b) = \max_{a\in A} \phi(a, b).$$

As a consequence, the minimum payment that the second player can achieve is

$$V^+ = \min_{b\in B} \phi(\alpha(b), b) = \min_{b\in B} \max_{a\in A} \phi(a, b) \tag{3.22}$$

CASE 2: The first player chooses $a$strategy $a \in A$, then the second player makes his choice, depending on a.

In this case, it is player B who has the advantage of knowing his opponent's strategy. The best reply of player B will be some $\beta(a) \in B$ such that

$$\phi(a, \beta(a)) = \min_{b\in B} \phi(a, b)$$

As a consequence, the maximum payment that the first player can secure is

$$V^- = \max_{a\in A} \phi(a, \beta(a)) = \max_{a\in A} \min_{b\in B} \ \phi(a, b). \tag{3.23}$$

Lemma 3.4.1. *In the above setting, one has*

$$V^- = \max_{a\in A} \min_{b\in B} \ \phi(a, b) \le \min_{b\in B} \max_{a\in A} \phi(a, b) = V^+. \tag{3.24}$$

Proof. Consider the (possibly discontinuous) map $a \mapsto \beta(a)$, i.e. the best reply map for player B. Since

$$V^- = \sup_{a\in A} \phi(a, \beta(a).$$

given any $\epsilon > 0$ there exists $a_\epsilon \in A$ such that

$$\phi(a_\epsilon, \beta(a_\epsilon)) > V^- - \epsilon. \tag{3.25}$$

In turn, this implies

$$V^+ = \min_{b\in B} \phi(\alpha(b), b) \ge \min_{b\in B} \phi(a_\epsilon, \beta(a_\epsilon)) > V^- - \epsilon$$

Since $\epsilon > 0$ was arbitrary, this proves the lemma. □

In general, one may have the strict inequality $V^- < V^+$. In the case where equality holds, we say that this common value $V = V^- = V^+$ is the **value of the game**.

Moreover, if there exist strategies $a^* \in A$ and $b^* \in B$ such that

$$\min_{b\in B} \phi(a^*, b) = \phi(a^*, b^*) = \max_{a\in A} \phi(a, b^*). \tag{3.26}$$

then we say that the pair (a^*, b^*) is a saddle point of the game. Calling V the common value of the two quantities in (3.26), the following holds:

- If A adopts the strategy a^*, he is guaranteed to receive no less than V.
- If B adopts the strategy b^*, he is guaranteed to pay no more than V .

For a zero-sum game, the concept of saddle point is thus the same as a Nash equilibrium.

Theorem 3.4.1. ***(value and saddle point).*** *Under the assumptions of the domains A, B are compact metric spaces and the function $\phi : A \times B \mapsto \Re$ is continuous, the zero-sum game* (3.20) *has a value V if and only if a saddle point (a^*, b^*) exists. In the positive case, one has*

$$V = V^- = V^+ = \phi(a^*, b^*). \tag{3.27}$$

Proof. 1. Assume that a saddle point (a^*, b^*) exists. Then

$$V^- = \max_{a\in A} \min_{b\in B} \phi(a, b) \geq \min_{b\in B} \phi(a^*, b) = \max_{a\in A} \phi(a, b^*) \geq \min_{b\in B} \max_{a\in A} \phi(a, b) = V^+.$$

By (3.24) this implies $V = V = V^+$, showing that the game has a value.

2. Next, assume $V = V = V^+$. Let $a \mapsto \beta(a)$ be the best reply map for player B. For each $\epsilon > 0$ choose $a_\epsilon \in A$ such that (3.25) holds. Since the sets A and B are compact, we can choose a subsequence $\epsilon_n \to 0$ such that the corresponding strategies converge, say

$$a_{\epsilon_n} \to a^*, \qquad \beta(\epsilon_n) \to b^*$$

We claim that (a^*, b^*) is a saddle point. Indeed, the continuity of the payoff function ϕ yields

$$\phi(a^*, b^*) = \lim_{n\to\infty} \phi(a_{\epsilon_n}, \beta(a_{\epsilon_n})).$$

From

$$V^- - \epsilon_n < \phi(a_{\epsilon_n}, \beta(a_{\epsilon_n})) \leq \sup_{a\in A} \phi(a, \beta(a)) = V^+$$

letting $\epsilon \to 0$ we conclude

$$V^- \leq \lim_{n\to\infty} \phi(a_{\epsilon_n}, \beta(a_{\epsilon_n})) = \phi(a^*, b^*) \leq V^+.$$

Since we are assuming $V = V^+$, this shows that (a^*, b^*) is a saddle point, concluding the proof.

□

Remark 3.4.1. *A non-zero-sum game may admit several Nash equilibrium solutions, providing different payoffs to each players. However, for a zero-sum game, if a Nash equilibrium exists, then all Nash equilibria yield the same payoff. Indeed, this payoff (i.e., the value of the game) is characterized as*

$$V = \min_{b\in B} \max_{a\in A} \phi(a, b) = \max_{a\in A} \min_{b\in B} \phi(a, b).$$

Corollary 3.4.1. ***(Existence of a saddle point)****. Consider a zero-sum game.*
If the domains A, B are compact metric spaces and the function $\phi : A \times B \mapsto \Re$ are continuous.
Assume that the sets A, B are convex, and moreover

$$\begin{aligned} a &\mapsto \phi(a,b) \text{ is a concave function of } a, \text{ for each fixed } b \in B. \\ b &\mapsto \phi(a,b) \text{ is a convex function of } b, \text{ for each fixed } a \in A. \end{aligned}$$

Then the game admits a Nash equilibrium, i.e. a saddle point.

Since a game always admits a Nash equilibrium in the class of randomized strategies. Specializing this result to the case of zero-sum games one obtains

Corollary 3.4.2. ***(Existence of a saddle point in the class of randomized strategies).***
If the domains A, B are compact metric spaces and the function $\phi : A \times B \mapsto \Re$ are continuous, a zero-sum game always has a value, and a saddle point, within the class of randomized strategies.

Otherwise stated, there exists a pair (μ^*, v^*) of probability measures on A and B respectively,such that

$$\int_{A\times B} \phi(a,b) d\mu \times dv^* \leq \int_{A\times B} \phi(a,b) d\mu^* \times dv^* \leq \int_{A\times B} \phi(a,b) d\mu^* \times dv$$

for every other probability measures $\mu \in P(A)$ and $v \in P(B)$. If the game already has a value with in the class of pure strategies, the two values of course coincide.

We now specialize this result to the case of a matrix game, where $A = \{a_1, \dots, a_m\}$ and $B = b_1, \dots, b_n$. The sets of randomized strategies can now be identified with the simplexes $\triangle_m, \triangle_n$ defined at (3.9). Let $\phi_{ij} = \phi(a_i, b_j)$. According to the above corollary , there exist $x^* \in \triangle_m$ and $y^* \in \triangle_n$ such that

$$\max_{x\in\triangle_m} \left(\sum_{ij} \phi_{ij} x_i y_j^* \right) \leq \sum_{ij} \phi_{ij} x_i^* y_j^* \leq \min_{y\in\triangle_n} \left(\sum_{ij} \phi_{ij} x_i^* y_j \right).$$

To compute the optimal randomized strategies x^*, y^* we observe that any linear function on a compact domain attains its global maximum or minimum at an extreme point of the domain.
Therefore

$$\begin{aligned} \max_{x\in\triangle_m} \left(\sum_{ij} \phi_{ij} x_i y_j \right) &= \max_{i\in\{1,\dots,m\}} \left(\sum_{j} \phi_{ij} y_j \right). \\ \min_{y\in\triangle_n} \left(\sum_{ij} \phi_{ij} x_i y_j \right) &= \min_{j\in\{1,\dots,n\}} \left(\sum_{i} \phi_{ij} x_i \right). \end{aligned}$$

The value $x^* = (x_1^*, \ldots, x_m^*) \in \triangle_m$ is thus the point where the function

$$x \mapsto \phi^{min}(x) = \min_j \left(\sum_i \phi_{ij} x_i \right) \tag{3.28}$$

attains its global maximum. Similarly, the value $y^* = (y_1^*, \ldots, y_n^*) \in \triangle_m$ is thus the point where the function

$$y \mapsto \phi^{max}(y) = \max_i \left(\sum_j \phi_{ij} y_j \right) \tag{3.29}$$

attains its global minimum.

Example 3.4.1. ***(rock-paper-scissors game).*** *This is a zero-sum matrix game. Each player has a set of three choices, which we denote as $\{R, P, S\}$. The corresponding matrix of payoffs for player A is given in table 3.5. The upper and lower values of the game are $V^+ = 1, V = 1$. No saddle point exists within the class of pure strategies. However, the game has a saddle point within the class of randomized strategies, where each player chooses among his three options with equal probabilities $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. In this case, the value of the game is $V = 0$.*

Player B

	Strategies	R	P	S
Player A	R	0	-1	1
	P	1	0	-1
	S	-1	1	0

Table 3.5: The matrix describing the "rock-paper-scissors" game.
Its entries represent the payments from player B to player A.

Example 3.4.2. *Player B (the defender) has two military installations. He can defend one, but not both. Player A (the attacker) can attack one of the two. An installation which is attacked but not defended gets destroyed. The first installation is worth three times more than the second one. Each player must decide which installation to attack (or defend), without knowledge of the other player's strategy.*

This situation can be modeled as a zero-sum game, where the payoff matrix is given in table 3.6.

$$\begin{aligned} \phi^{min}(x) &= \min_{j \in 1,2} (\phi_{1j} x_1 + \phi_{2j} x_2) = \min\{3x_1, x_2\} \qquad (x_1 = 1 - x_2). \\ \phi^{max}(y) &= \max_{i \in 1,2} (\phi_{i1} y_1 + \phi_{i2} y_2) = \min\{y_1, 3y_2\} \qquad (y_1 = 1 - y_2). \end{aligned}$$

A saddle point, in randomized strategies, is provided the pair (x^*, y^*), *where*

$$x^* = \left(\frac{1}{4}, \frac{3}{4}\right), \qquad y^* = \left(\frac{3}{4}, \frac{1}{4}\right).$$

In other words, Player B should favor defending his first (and more valuable) installation with odds $3:1$. *Player A should favor attacking the second (less valuable) installation with odds* $3:1$. *The more valuable installation is destroyed with probability* $1/16$, *while the less valuable one gets destroyed with probability* $9/16$. *The value of the game is* $3/4$.

Player B

Player A

Strategies	b_1	b_2
a_1	0	3
a_2	1	0

Table 3.6: The payoff matrix for the zero-sum game for the next Example.

Chapter 4

Differential Game

4.1 The Differential Game Problem

We consider differential game problems, where the initial time t_0 and the initial state $x(t_0) = x_0$ are specified.
Without loss of generality, let us assume that $t_0 = 0$ hence, a general form of differential game problem for two players can be formulated as follows:
Find admissible optimal controls $u_1 : [0,T] \to U_1$ and $u_2 : [0,T] \to U_2$ such that the dynamic system described by the differential equation

$$\begin{aligned} \dot{x} &= f(t, x, u_1, u_2), \qquad t \in [0,T] \\ x(0) &= x_0. \end{aligned} \tag{4.1}$$

is achieved. where $u_1(\cdot), u_2(\cdot)$ are the controls implemented by the two players. We assume that the controls satisfy the pointwise constraints

$$u_1(t) \in U_1, \qquad u_2(t) \in U_2 \tag{4.2}$$

for some given sets $U_1, U_2 \subseteq \Re^m$
For $i = 1, 2$ the goal of the i^{th} player is to maximize his own payoff, namely

$$J_i(u_1, u_2) = \psi_i[x(T)] - \int_0^T L_i(t, x(t), u_1(t), u_2(t))dt \tag{4.3}$$

where ψ_i is a terminal payoff, while L_i accounts for a running cost.

In order to completely describe the game, it is essential to specify the information available to the two players. Indeed, the strategy adopted by a player depends on the information available to him at each time t. Therefore, different information structures result in different game situations.

In the following, we shall assume that each player has perfect knowledge of

- The function f determining the evolution of the system,
- The sets U_1, U_2 of control values available to the two players.
- The two payoff functions J_1, J_2.
- The instantaneous time $t \in [0, T]$ (i.e. both players have a clock).
- The initial state x_0.

However, we shall consider different cases concerning the information that each player has, regarding:

(i) the current state of the system $x(t)$, and

(ii) the control $u(\cdot)$ implemented by the other player

CASE 1 (Open loop strategies):

Apart from the initial data, Player i cannot make any observation of the state of the system, or of the strategy adopted by the other player. In this case, his strategy must be open loop, i.e. it can only depend on time $t \in [0, T]$. The set S_i of strategies available to the i^{th} player will thus consist of all measurable functions $t \to u_i(t)$ from $[0, T]$ into U_i

CASE 2 (Markovian strategies):

Assume that, at each time $t \in [0, T]$, Player i can observe the current state $x(t)$ of the system. However, he has no additional information about the strategy of the other player. In particular, he cannot predict the future actions of the other player.
In this case, each player can implement a **Markovian** strategy (i.e., of **feedback** type): the control $u_i = u_i(t, x)$ can depend both on time t and on the current state x. The set S of strategies available to the i^{th} player will thus consist of all measurable functions $(t, x) \to u_i(t, x)$ from $[0, T] \times \Re^n$ into U_i.

CASE 3 (Hierarchical play):

Player 1 (the leader) announces his strategy in advance. This can be either open loop $u_1 = u_1^\#(t)$, or feedback $u_1 = u_1^\#(t,x)$. At this stage, the game yields an optimal control problem for Player 2 (the follower). Namely

$$maximize: \qquad = \psi_2[x(T)] - \int_0^T L_2(t, x(t), u_1^\#(t,x), u_2(t))dt \tag{4.4}$$

subjected to

$$\dot{x} = f(t, x, u_1^\#(t,x), u_2(t)), \qquad x(0) = x_0, \qquad u_2(t) \in U_2 \tag{4.5}$$

Notice that in this case the knowledge of the initial point x together with the evolution equation (4.5) provides Player 2 with complete information about the state of the system for all $t \in [0,T]$. From the point of view of Player 1, the task is to devise a strategy $u_1 = u_1^\#(t,x)$ such that the reply u_2 of the other player yields a payoff (for Player 1) as large as possible.

4.2 Open loop strategies

In this section we consider solutions to the differential game (4.1) - (4.3) , in the case where the strategies implemented by the players must be functions of time alone.

4.2.1 Open-loop Nash equilibrium solutions

Definition 4.2.1. *A pair of control functions $t \to (u_1^*(t), u_2^*(t))$ is a Nash equilibrium for the game* (4.1) - (4.3) *within the class of open-loop strategies if the following holds.*

(i) The control $u_1^*(\cdot)$ provides a solution to the optimal control problem for Player 1:

$$maximize: J_1(u_1, u_2^*) = \psi_1[x(T)] - \int_0^T L_1(t, x(t), u_1(t), u_2^*(t))dt \tag{4.6}$$

over all controls $u_1(\cdot)$, for the system with dynamics

$$x(0) = x_0 \in \Re^n \qquad \dot{x} = f(t, x, u_1, u_2^*(t)), \qquad u_1(t) \in U_1 \qquad t \in [0,T]. \tag{4.7}$$

(ii) The control $u_2^*(\cdot)$ provides a solution to the optimal control problem for Player 2:

$$maximize: J_1(u_1^*, u_2) = \psi_1[x(T)] - \int_0^T L_1(t, x(t), u_1^*(t), u_2(t))dt \tag{4.8}$$

over all controls $u_2(\cdot)$, for the system with dynamics

$$x(0) = x_0 \in \Re^n \qquad \dot{x} = f(t, x, u_1^*, u_2(t)), \qquad u_2(t) \in U_2 \qquad t \in [0,T]. \tag{4.9}$$

To find Nash equilibrium solutions, we thus need to simultaneously solve two optimal control problems. The optimal solution $u_1^*(\cdot)$ of the first problem enters as a parameter in the second problem, and viceversa. Assuming that all functions $f, \psi_1, \psi_2, L_1, L_2$ are continuously differentiable, necessary conditions for optimality are provided by the Pontryagin Maximum Principle.
Based on the **PMP**, we now describe a procedure for finding a pair of open-loop strategies $t \to (u_1^*(t), u_2^*(t))$ yielding a Nash equilibrium. Toward this goal, we need to assume that a family of point-wise maximization problems can be uniquely solved. More precisely, we assume

Assumption (A1)

for every $(t,x) \in [0,T] \times \Re^n$ and any two vectors $q_1, q_2 \in \Re^n$ there exists a unique pair $(u_1^{\#}, u_2^{\#}) \in U_1 \times U_2$ such that

$$u_1^{\#} = arg \max_{\omega \in U_1} \{q_1 \cdot f(t,x,\omega,u_2^{\#}) - L_1(t,x,\omega,u_2^{\#})\}, \tag{4.10}$$

$$u_2^{\#} = arg \max_{\omega \in U_2} \{q_2 \cdot f(t,x,u_1^{\#},\omega) - L_2(t,x,u_1^{\#},\omega\}, \tag{4.11}$$

The corresponding map will be denoted by

$$(t,x,q_1,q_2) \to (u_1^{\#}(t,x,q_1,q_2), u_2^{\#}(t,x,q_1,q_2)) \tag{4.12}$$

The assumption $(A1)$ can be interpreted as follows. For any give $(t,x,q_1,q_2) \in [0,T] \times \Re^n \times \Re^n \times \Re^n$, consider the one-shot game where the players choose strategies $u_i \in U_i$ in order to maximize their instantaneous payoffs

$$\phi_i(u_1,u_2) = q_i \cdot f(t,x,u_1,u_2) - L_i(t,x,u_1,u_2) \qquad i = 1,2. \tag{4.13}$$

According to $(A1)$, for every t, x, q_1, q_2 this one-shot game has a unique Nash equilibrium solution. Notice that, if the sets U_1, U_2 of control values are compact, from this uniqueness property it follows that the map in (4.12) is continuous. We now describe an important class of problems where this assumption is satisfied.

Lemma 4.2.1. *Assume that the dynamics and the running costs take the decoupled form*

$$f(t,x,u_1,u_2) = f_0(t,x) + M_1(t,x)u_1 + M_2(t,x)u_2, \tag{4.14}$$

$$L_i(t,x,u_1,u_2) = L_{i1}(t,x,u_1) + L_{i2}(t,x,u_2). \tag{4.15}$$

Assume that

(i) The domains U_1, U_2 are closed and convex subsets of $\Re^m$, possibly unbounded.

(ii) M_1, M_2 are $N \times m$ matrices, continuously depending on t, x,

(iii) The functions $u_1 \to L_{11}(t,x,u_1)$ and $u_2 \to L_{22}(t,x,u_2)$ are strictly convex,

(iv) For each $i = 1,2$, either U_i is compact, or L_{ii} has superlinear growth, i.e.

$$\lim_{|\omega|\to+\infty} \frac{L_{ii}(t,x,\omega)}{|\omega|} = +\infty.$$

Then the assumption (A1) holds.

Indeed, for any given (t,x,q_1,q_2), the control values $u_1^\#, u_2^\#$ are determined

$$\begin{aligned} u_1^\# &= arg \max_{\omega \in U_1} \{q_1 \cdot M_1(t,x)\omega - L_{11}(t,x,\omega\}, \\ u_2^\# &= arg \max_{\omega \in U_2} \{q_2 \cdot M_2(t,x)\omega - L_{22}(t,x,\omega)\} \end{aligned} \tag{4.16}$$

The assumptions $(i)(iv)$ guarantee that the above maximizers exist and are unique.

Finding a Nash equilibrium using the PMP. Assume that (A1) holds, and let $x^*(\cdot), u_1^*(\cdot), u_2^*(\cdot)$ be respectively the trajectory and the open-loop controls of the two players, in a Nash equilibrium. By definition, the controls u_1^* and u_2^* provide solutions to the corresponding optimal control problems for the two players. Applying the Pontryagin Maximum Principle we obtains the following set of necessary conditions.

$$\begin{cases} \dot{x} = f(t,x,u_1^*,u_2^*), \\ \dot{q}_1 = -q_1 \dfrac{\partial f}{\partial x}(t,x,u_1^*,u_2^*) + \dfrac{\partial L_1}{\partial x}(t,x,u_1^*,u_2^*) \\ \dot{q}_2 = -q_2 \dfrac{\partial f}{\partial x}(t,x,u_1^*,u_2^*) + \dfrac{\partial L_2}{\partial x}(t,x,u_1^*,u_2^*) \end{cases} \tag{4.17}$$

with initial and terminal conditions

$$\begin{cases} x(0) = x_0 \\ q_1(T) = \nabla_x \psi_1(x(T)) \\ q_2(T) = \nabla_x \psi_2(x(T)) \end{cases} \tag{4.18}$$

Notice that in (4.17) the variables $u_1^\#, u_2^\#$ are functions of (t,x,q_1,q_2), defined at (4.10)-(4.12). One can use the above system in order to compute a Nash equilibrium solution to the differential game. Notice that (4.17) consists of three **ODEs** in $\Re^n$. This needs to be solved with the mixed boundary data (4.18). Here the value of variable x (the state of the system) is explicitly given at the initial time $t = 0$. On the other hand, since $x(T)$ is not a priori known, the values for q_1, q_2(the adjoint variables) are only determined by two implicit equations at the terminal time $t = T$. Together with the strong non linearity of the maps $u_1^{\#}, u_1^{\#}$ in (4.12), this makes the problem (4.12)-(4.13) hard to solve, in general. If a solution $t \to (x(t), q_1(t), q_2(t))$ to the two-point boundary value problem (4.17)-(4.18) is found, the trajectory x^* and the controls u_1^*, u_2^* are determined by

$$x^*(t) = x(t), \qquad u_1^*(t) = u_1^{\#}(t, x(t), q_1(t), q_2(t)) \quad \text{and} \quad u_2^*(t) = u_2^{\#}(t, x(t), q_1(t), q_2(t))$$

The above Pontryagin maximum principle is only a necessary condition, not sufficient for optimality. In other words, any pair $t \to (u_1^*(t), u_2^*(t))$ of open-loop strategies which is a Nash equilibrium must provide a solution to (4.17)-(4.18). On the other hand, being a solution of (4.17)-(4.18) does not guarantee that the pair $(u_1^*(t), u_2^*(t))$ is a Nash equilibrium.
Based on the following Theorem on sufficient Conditions for Optimality of the obtimal control problem we have a remark for sufficient Conditions for open loop Nash equilibrium solution.

Sufficient Conditions for Optimality

In general, the conditions stated by PMP are necessary but not sufficient for a control $u^*(\cdot)$ to be optimal. However, under a suitable concavity condition, it turns out that every control $u^*(\cdot)$ satisfying the PMP is optimal. Consider the Hamiltonian function

$$H(t, x, u, p) = p \cdot f(t, x, u) - L(t, x, u) \tag{4.19}$$

and the reduced Hamiltonian

$$H(t, x, p) = \max_{\omega \in U} \{p(t) \cdot f(t, x, \omega) - L(t, x, \omega)\} \tag{4.20}$$

Theorem 4.2.1. ***(PMP + concavity ⇒ optimality).***

Let $t \to u^*(t)$ be an optimal control function, and let $t \to x^*(t)$ be the corresponding optimal trajectory for the maximization problem

$$\textit{maximize:} \quad J(u : x_0) \;=\; \psi(x(T)) - \int_0^T L(t, x(t), u(t))dt.$$

subjected to :

$$\begin{aligned} \dot{x} &= f(t, x, u), \\ x(0) &= x_0 \quad u(t) \in U \quad \textit{for all} \quad t \in [0, T] \end{aligned}$$

Consider a measurable function $t \to u^*(t) \in U$ *and two absolutely continuous functions* $x^*(\cdot), p(\cdot)$ *satisfying the boundary value problem*

$$\begin{cases} \dot{x} = f(t, x, u^*(t)), \\ \dot{p} = -p(t)\frac{\partial f}{\partial x}(t, x, u^*(t)) + \frac{\partial L}{\partial x}(t, x, u^*(t)). \end{cases} \qquad \begin{cases} x(0) = x_0, \\ p(T) = \nabla\psi(x(T)). \end{cases} \tag{4.21}$$

together with the maximality condition;

$$\begin{aligned} p(t) \cdot f(t, x^*(t), u^*(t)) - L(t, x^*(t), u^*(t)) &= \max_{u \in U}\{p(t) \cdot f(t, x^*(t), u(t)) \\ &\quad -L(t, x^*(t), u(t))\} \end{aligned} \tag{4.22}$$

. Assume that the set U *is convex and that the functions*

$$x \to H(t, x, p(t)) \qquad x \to \psi(x)$$

are concave. Then $u^*(\cdot)$ *is an optimal control, and* $x^*(\cdot)$ *is the corresponding optimal trajectory.*

Proof. Let $u : [0, T] \to U$ be any measurable control function. Then

$$\begin{aligned} J(u) - J(u^*) &= \psi(x(T)) - \psi(x^*(T)) - \int_0^T [L(t, x(t), u(t)) - L(t, x^*(t), u^*(t))]\, dt \\ &= \psi(x(T)) - \psi(x^*(T)) + \int_0^T \{[H(t, x(t), p(t), u(t)) - p(t)\dot{x}(t)] \\ &\quad - [H(t, x^*(t), p(t), u^*(t)) - p(t)\dot{x}^*(t)]\}\, dt \end{aligned} \tag{4.23}$$

Since u^* satisfies the maximality condition (4.22), for $a.e. t \in [0, T]$ one has

$$H(t, x^*(t), p(t), u^*(t)) = H(t, x(t), p(t)), \qquad H(t, x(t), p(t), u(t)) \le H(t, x(t), p(t)).$$

Using these inequalities in (4.23) we obtain

$$\begin{aligned} J(u) - J(u^*) &\le \psi(x(T)) - \psi(x^*(T)) + \int_0^T \{[H(t, x(t), p(t)) - p(t)\dot{x}(t)] \\ &\quad - [H(t, x^*(t), p(t)) - p(t)\dot{x}^*(t)]\}\, dt \end{aligned} \tag{4.24}$$

If the map $x \to H(t, x, p(t))$ is differentiable, the concavity assumption implies

$$\begin{aligned} H(t, x(t), p(t)) &\le H(t, x^*(t), p(t)) + \frac{\partial H}{\partial x}(t, x^*(t), p(t))\, [x(t) - x^*(t)] \\ &= H(t, x^*(t), p(t)) - \dot{p}(t)|x(t) - x^*(t)| \end{aligned}$$

The same conclusion can be reached also if H is not differentiable, using arguments from convex analysis. Inserting the above inequality in (4.24) we finally obtain

$$\begin{aligned} J(u) - J(u^*) &\le \psi(x(T)) - \psi(x^*(T)) - \int_0^T \{\dot{p}(t)|x(t) - x^*(t)| + p(t)|x(t) - x^*(t)|\}\, dt \\ &= \psi(x(T)) - \psi(x^*(T)) - \{p(T)|x(T) - x^*(T)| - p(0)|x(0) - x^*(0)|\} \\ &\le 0. \end{aligned}$$

Indeed, the initial and terminal conditions in (4.21), and the concavity of ψ imply

$$x(0) = x^*(0) = x_0, \qquad \psi(x(T)) \le \psi(x^*(T)) + \nabla\psi(x^*(T))[x(T)x^*(T)].$$

□

Remark 4.2.1. *Depending on the definition 4.2.1 we can define the Hamiltonian function;*

$$\begin{aligned} H_1(t, x, q_1, u_1, u_2^*) &= q_1 \cdot f(t, x, u_1, u_2^*) - L(t, x, u_1, u_2^*) \\ H_2(t, x, q_2, u_1^*, u_2) &= q_2 \cdot f(t, x, u_1^*, u_2) - L(t, x, u_1^*, u_2) \end{aligned}$$

and the reduced Hamiltonian;

$$\begin{aligned} \mathbb{H}_1(t, x, q_1, u_2^*) &= \max_{\omega \in u_1} \{q_1 \cdot f(t, x, \omega, u_2^*) - L(t, x, \omega, u_2^*)\} \\ \mathbb{H}_2(t, x, q_2, u_1^*) &= \max_{\omega \in u_2} \{q_2 \cdot f(t, x, u_1^*, \omega) - L(t, x, u_1^*, \omega)\} \end{aligned}$$

If U_1 and U_2 are convex , $\psi_1(x(T))$ and $\psi_2(x(T))$ are concave with respect to x , $\mathbb{H}_1(t, x, q_1, u_2^*)$ and $\mathbb{H}_2(t, x, q_2, u_1^*)$ are concave with respect to (x, u_1) and (x, u_2) respectively then the pair (u_1^*, u_2^*) is an open loop Nash equilibrium solution for the given differential game (4.1) - (4.3) *within the class of open-loop strategies.*

Example 4.2.1. *Given the differential game*

$$\begin{aligned} \max_{u_i(\cdot)} J_i\left(u_1(\cdot), u_2(\cdot)\right) &= \max_{u_i(\cdot)} \int_{\ln(2)}^{1} x(t)u_i(t) - \frac{u_i^2(t)}{2} dt \\ & \quad \dot{x} = u_1(t) + u_2(t) \\ & \quad x(\ln(2)) = 1 \end{aligned} \tag{4.25}$$

the Hamiltonian function of player i is

$$\mathbb{H}_i(t, x, u_1, q_i, u_2) = xu_i - \frac{u_i^2}{2} + q_i(u_1 + u_2) \qquad \textit{for} \quad i = \{1, 2\} \tag{4.26}$$

and then from

- *optimality condition we have*

$$\begin{aligned} \frac{\partial \mathbb{H}_i}{\partial u_i} = 0 \;\Rightarrow\; & \frac{\partial\left(xu_i - \frac{u_i^2}{2} + q_i(u_1 + u_2)\right)}{\partial u_i} = 0 \\ & x - u_i + q_i u_i = 0 \\ & u_i^{\#}(t, x, q_1, q_2) = x + q_i \end{aligned} \tag{4.27}$$

- *adjoint condition*

$$\dot{q}_i = -\frac{\partial \mathbb{H}_i}{\partial x} \quad \Rightarrow \quad \dot{q}_i = -\frac{\partial\left(xu_i - \frac{u_i^2}{2} + q_i(u_1+u_2)\right)}{\partial x}$$

$$\dot{q}_i = -u_i$$

$$\dot{q}_1 = -u_1 \qquad \text{and} \quad \dot{q}_2 = -u_2$$

$$\Rightarrow \quad \dot{q}_1 = -x - q_1 \qquad \text{and} \quad \dot{q}_2 = -x - q_2 \tag{4.28}$$

- *state law*

$$\begin{aligned} \dot{x} &= u_1(t) + u_2(t) \\ &= x + q_1 + x + q_2 \\ &\Rightarrow \ \dot{x} = 2x + q_1 + q_2 \end{aligned} \tag{4.29}$$

Now from equation (4.28)-(4.29) *we can determine the state* $x(\cdot)$ *and the adjoint variables* $q_1(\cdot), q_2(\cdot)$ *by solving the following boundary value problem*

$$\begin{aligned} \dot{x} &= 2x + q_1 + q_2 \\ \dot{q}_1 &= -x - q_1 \\ \dot{q}_2 &= -x - q_2 \end{aligned}$$

with initial and terminal conditions

$$\begin{cases} x(\ln(2)) = 1 \\ q_1(1) = 0 \\ q_2(1) = 0 \end{cases} \tag{4.30}$$

therefore the above systems of ordinary differential equations can be written as follows

$$\underbrace{\begin{pmatrix} \dot{x} \\ \dot{q}_1 \\ \dot{q}_2 \end{pmatrix}}_{\dot{Y}} = \underbrace{\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ q_1 \\ q_2 \end{pmatrix}}_{Y} \tag{4.31}$$

To solve,

$$\dot{Y} = AY$$

first find the eigenvalues and the eigenvectors of a matrix A

$$\begin{aligned}\det(A-\lambda I) &= 0\\ \begin{vmatrix} 2-\lambda & 1 & 1\\ -1 & -1-\lambda & 0\\ -1 & 0 & -1-\lambda \end{vmatrix} &= 0\\ \lambda^3-\lambda &= 0\\ \lambda(\lambda-1)(\lambda+1) &= 0\end{aligned}$$

Therefore

$$\lambda_1 = 0, \quad \lambda_2 = 1, \quad \lambda_3 = -1, \quad \textit{are eigenvalues of matrix} \quad A$$

for $\lambda_1 = 0, v_1 = (-1,1,1)^t$
for $\lambda_2 = 1, v_2 = (-2,1,1)^t$ *and*
for $\lambda_3 = -1, v_3 = (0,-1,1)^t$ *are linearly independent eigenvectors to the corresponding eigenvalue of a matrix A*
This implies that $Y(t) = c_1v_1 + c2e^tv_2 + c_3e^{-t}v_3$ *is the general solution of* $\dot{Y} = AY$.
Therefore

$$\begin{aligned} Y(t) &= c_1v_1 + c2e^tv_2 + c_3e^{-t}v_3\\ \begin{pmatrix} x(t)\\ q_1(t)\\ q_2(t)\end{pmatrix} &= c_1\begin{pmatrix}-1\\1\\1\end{pmatrix} + c_2e^t\begin{pmatrix}-2\\1\\1\end{pmatrix} + c_3e^{-t}\begin{pmatrix}0\\-1\\1\end{pmatrix}\\ &= \begin{pmatrix} -c_1-2c_2e^t\\ c_1+c_2e^t-c_3e^{-t}\\ c_1+c_2e^t+c_3e^{-t}\end{pmatrix}\end{aligned}$$

$$\begin{cases} x(t) = -c_1 - 2c_2e^t\\ q_1(t) = c_1 + c_2e^t - c_3e^{-t}\\ q_2(t) = c_1 + c_2e^t + c_3e^{-t}\end{cases}$$

with initial and terminal conditions stated in equation (4.30) *we can get the value of* c_1, c_2, c_3

$$c_1 = \frac{-e}{e-4}, \qquad c_2 = \frac{1}{e-4}, \qquad c_3 = 0$$

this implies

$$\begin{cases} x^*(t) = \dfrac{2e^t-e}{4-e}\\ q_1^*(t) = \dfrac{e-e^t}{4-e}\\ q_2^*(t) = \dfrac{e-e^t}{4-e}\end{cases} \tag{4.32}$$

and from equation (4.27)

$$\begin{aligned} u_i^{\#}(t,x,q_1,q_2) &= x + q_i \\ &= \frac{2e^t - e}{4-e} + \frac{e - e^t}{4-e} \\ u_i^{\#}(t) &= \frac{e^t}{4-e} \\ \therefore u_1^{\#}(t) = u_2^{\#}(t) &= \frac{e^t}{4-e} \end{aligned} \tag{4.33}$$

Since in case of open loop strategies the controls depends only as a function of time t we are expected to show that the Hamiltonian function $\mathbb{H}_i = \mathbb{H}_i(t, x^*, u_i, u_j^*, q_i^*, q_j^*)$ the Hamiltonian function $\mathbb{H}_i(t; x; u_i)$ is concave in (x, u_i) .
Therefore its Hessian is constant and equal to

$$\begin{pmatrix} \mathbb{H}_{xx} & \mathbb{H}_{xu_i} \\ \mathbb{H}_{u_ix} & \mathbb{H}_{u_iu_i} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{which is negative definite}$$

this implies that $\mathbb{H}$ is strictly concave .

$$\therefore u_1^{\#}(t) = u_2^{\#}(t) = \frac{e^t}{4-e} \qquad \text{is the optimal solution of the problem}$$

Therefore, the optimal value of the problem will be :

$$\begin{aligned} J_1^* = J_2^* = J^* &= \int_{\ln(2)}^{1} \left(\frac{(2e^t - e)}{(4-e)^2} e^t - \frac{e^{2t}}{2(4-e)^2} \right) dt \\ &= \frac{1}{2(4-e)^2} \left(-\frac{e^2}{2} + 4e - 6 \right) \end{aligned}$$

4.2.2 Open-loop Stackelberg equilibrium solution

We now assume that the strategies of the players are not chosen simultaneously, but in two stages.
First, Player 1 (the leader) chooses his strategy $t \to u_1(t)$, and communicates it to Player 2. In a second stage, Player 2 (the follower) chooses his control function $u_2(\cdot)$. maximizing his own payoff, relative to the strategy $u_1(\cdot)$ already chosen by the first player.
Given any admissible control $u_1^{\#} : [0,T] \to U_1$ for the first player, we denote by $R_2(u_1^{\#})$ the set of best replies for the second player. More precisely, $R_2(u_1^{\#})$ is the set of all admissible control functions $u_2 : [0,T] \to U_2$ for Player 2, which achieve the maximum payoff in connection with $u_1^{\#}$. Namely, they solve the optimal control problem

$$maximize : \psi_2[x(T)] - \int_0^T L_2(t, x(t), u_1^{\#}(t), u_2(t))dt \tag{4.34}$$

Over all control functions $u_2(\cdot)$, subject to

$$\dot{x} = f(t, x(t), u_1^{\#}(t), u_2(t)), \qquad x(0) = x_0, \qquad u_2(t) \in U_2. \tag{4.35}$$

Definition 4.2.2. *We say that a pair of control functions $t \to (u_1^*(t), u_2^*(t))$ is a Stackelberg equilibrium for the game* (4.1)-(4.3) *within the class of open-loop strategies if the following holds.*

(i) $u_2^*(t) \in R_2(u_1^*)$

(ii) Given any admissible control $u_1(\cdot)$ for Player 1 and every best reply $u_2(\cdot) \in R_2(u_1)$ for Player 2, one has

$$\psi_1(x(T, u_1, u_2)) - \int_0^T L_1(t, x(t, u_1, u_2), u_1(t), u_2(t))dt$$
$$\leq \psi_1(x(T, u_1^*, u_2^*)) - \int_0^T L_1(t, x(t, u_1^*, u_2^*), u_1^*(t), u_2^*(t))dt \tag{4.36}$$

To find a Stackelberg solution, Player 1 has to calculate the best reply of Player 2 to each of his controls $u_1(\cdot)$, and choose the control function $u_1^*(t)$ in order to maximize his own payoff J_1. We are here taking the optimistic view that, if Player 2 has several best replies to a strategy $u_1^*(\cdot)$, he will choose the one which is most favorable to Player 1. Necessary conditions in order that a pair of open-loop strategies (u_1^*, u_2^*) be a Stackelberg equilibrium can be derived by variational analysis. Let $t \to x^*(t)$ be the trajectory of the system determined by the controls u_1^*, u_2^*. Since $u_2^*(\cdot)$ is an optimal reply for Player 2, the Pontryagin maximum principle yields the existence of an adjoint vector $q_2^*(\cdot)$ such that

$$\begin{cases} \dot{x^*}(t) = f(t, x^*(t), u_1^*(t), u_2^*(t), \\ \dot{q_2^*}(t) = -q_2 \dfrac{\partial f}{\partial x}(t, x^*(t), u_1^*(t), u_2^*(t)) + \dfrac{\partial L_2}{\partial x}(t, x^*(t), u_1^*(t), u_2^*(t)) \end{cases} \tag{4.37}$$

with boundary conditions

$$\begin{cases} x^*(0) = x_0 \\ q_2^*(T) = \nabla_x \psi_2(x^*(T)) \end{cases} \tag{4.38}$$

Moreover, the following optimality conditions hold:

$$u_2^*(t) \in arg \max_{\omega \in U_2} \{q_2^*(t) \cdot f(t, x^*(t), u_1^*(t), \omega) - L_2(t, x^*(t), u_1^*(t), \omega)\}, \text{ for a.e } t \in [0, T] \tag{4.39}$$

We now consider the problem of the first player. To derive a set of necessary conditions for optimality, our main assumption is:

Assumption (A2)

for each $(t, x, u_1, q_2) \in [0, T] \times \Re^n \times U_1 \times \Re^n$ there exists a unique optimal choice $u_2^b \in U_2$ for Player 2, namely

$$u_2^b(t, x, u_1, q_2) = arg \max_{\omega \in U_2} \{q_2 \cdot f(t, x, u_1, \omega) - L_2(t, x, u_1, \omega)\}, \tag{4.40}$$

The optimization problem for Player 1 can now be formulated as an optimal control problem in an extended state space, where the state variables are $(x, q_2) \in \Re^n \times \Re^n$.

$$maximize : \psi_1(x(T)) - \int_0^T L_1(t, x(t), u_1(t), u_2^b(t, x(t), u_1(t), q_2(t)))dt \tag{4.41}$$

for the system on $\Re^{2n}$ with dynamics

$$\begin{cases} \dot{x}(t) = f(t, x(t), u_1, u_2^b(t, x, u_1, q_2)) \\ \dot{q_2}(t) = -q_2 \dfrac{\partial f}{\partial x}(t, x, u_1, u_2^b(t, x, u_1, q_2)) + \dfrac{\partial L_2}{\partial x}(t, x, u_1, u_2^b(t, x, u_1, q_2)) \end{cases} \tag{4.42}$$

and with boundary conditions

$$x(0) = x_0, \qquad q_2(T) = \nabla_x \psi_2(x(T)). \tag{4.43}$$

This is a standard problem in optimal control. Notice, however, that the state variables (x, q_2) are not both assigned at time $t = 0$. Instead, we have the constraint $x = x_0$ valid at $t = 0$ and another constraint $q_2 = \nabla_x \psi_2(x)$ valid at $t = T$. In order to apply the **PMP**, we need to assume that all functions in (4.41)-(4.43) are continuously differentiable w.r.t. the new state variables x, q_2. More precisely

Assumption(A3)

For every fixed $t \in [0, T]$ and $u_1 \in U_1$, the maps

$$\begin{aligned} (x, q_2) &\to \tilde{L}_1(t, x, u_1, q_2) = L_1(t, x, u_1, u_2^b(t, x, u_1, q_2)) \\ (x, q_2) &\to F(t, x, u_1, q_2) = f(t, x, u_1, u_2^b(t, x, u_1, q_2)) \\ (x, q_2) &\to G(t, x, u_1, q_2)) = -q_2 \frac{\partial f}{\partial x}(t, x, u_1), u_2^b(t, x, u_1, q_2)) + \\ &\qquad \frac{\partial L_2}{\partial x}(t, x, u_1, u_2^b(t, x, u_1, q_2)) \qquad (4.44) \\ x &\to \nabla_x \psi_2(x). \qquad (4.45) \end{aligned}$$

are continuously differentiable.

An application of the **PMP** to the above optimal control problem with initial and terminal state constraints yields

Consider the more general optimization problem

$$maximize: \quad J = \varphi(x(0)) + \psi(x(T)) - \int_0^T L(t, x(t), u(t))dt \tag{4.46}$$

assuming that both the initial and the terminal point can vary. Here φ is an initial payoff,ψ is a terminal payoff, while $L(\cdot)$ accounts for a running cost. The payoff J has to be maximized among all measurable control functions $u : [0, T] \to U$ and all choices of an initial and terminal data, satisfying the constraints

$$x(0) \in S_0, \qquad x(T) \in S_T. \tag{4.47}$$

In the following, we assume that the functions f, L, φ, ψ are continuously differentiable, while $S_0, S_T \subset \Re^n$ are two C^1 embedded manifolds. A set of necessary conditions for optimality is provided by the following more general version of the Pontryagin Maximum Principle.

Theorem 4.2.2. ***(PMP, constrained initial and terminal points).*** *Let $t \to u^*(t)$ be a bounded, optimal control function and let $t \to x^*(t)$ be the corresponding optimal trajectory for the problem* (4.46)*, with dynamics*

$$x = f(t, x, u),$$

and initial and terminal constraints (4.47)*.*
Then the following holds.

(i) There exists an absolutely continuous adjoint vector $t \to p(t) = (p_0, p_1, ..., p_n)(t)$ which never vanishes on $[0, T]$, with $p_0 \geq 0$ constant, satisfying

$$\dot{p}_i((t) = -\sum_{j=1}^{n} p_j(t)\frac{\partial f}{\partial x_i}(t, x*(t), u*(t)) + p_0\frac{\partial L}{\partial x_i}(t, x^*(t), u^*(t)), \quad i = 1, 2 \ldots, n \tag{4.48}$$

(ii) The initial and terminal values of p satisfy

$$\begin{cases} (p_1, ..., p_n)(0) = p_0\varphi_x(x^*(0)) + n_0 \\ (p_1, ..., p_n)(T) = p_0\psi_x(x^*(T)) + n_T \end{cases} \tag{4.49}$$

for some vector n_0 orthogonal to manifold S_0 at the initial point $x^*(0)$ and some vector n_T orthogonal to manifold S_T at the terminal point $x^*(T)$).

(iii) The maximality condition

$$\begin{aligned} &\sum_{i=1}^{n} p_i(t) \cdot f(t, x^*(t), u^*(t)) - p_0L(t, x^*(t), u^*(t)) \\ &= \max_{u \in U} \{p_i(t) \cdot f(t, x^*(t), u(t)) - p_0L(t, x^*(t), u(t))\} \end{aligned} \tag{4.50}$$

holds for a.e.$t \in [0, T]$

Remark 4.2.2. *In a standard situation, the sets S_0, S_T are described in terms of a finite number of smooth scalar functions. For example*

$$S_0 = \{x \in \Re^n; \quad \alpha_1(x) = 0, \quad \dots \quad , \alpha_k(x) = 0\}$$

Assuming that the gradients of the functions α_i are linearly independent at the point $x^*(0) \in S_0$, any normal vector to S_0 can be represented as

$$n_0 = \sum_{i=1}^{k} \lambda_i \alpha_{i,x}(x^*(0))$$

for some numbers $\lambda_1, \dots, \lambda_k \in \Re$. The first condition in (4.49) *can thus be written as*

$$(p_1, \dots, p_n)(0) = p_0 \varphi_x(x^*(0)) + \sum_{i=1}^{k} \lambda_i \alpha_{i,x}(x^*(0))$$

for some scalar multipliers $\lambda_1, \dots, \lambda_k$.

Remark 4.2.3. *In the special case where the initial point is fixed and the terminal point is free,*
we have $S_0 = x_0$ while $S_T = \Re^n$. Therefore, $n_0 \in \Re^n$ can be arbitrary while $n_T = 0$. The boundary conditions for p become

$$(p_1, \dots, p_n)(0) \quad \textit{arbitrary}, \qquad (p_1, \dots, p_n)(T) = p_0 \psi_x(x^*(T))$$

The condition that $(p_0, p_1, \dots, p_n)(T)$ should be non-zero implies $p_0 > 0$. Since p is determined up to a positive constant, it is thus not restrictive to assume $p_0 = 1$.

Relying on the PMP, the computation of the optimal control can be achieved in two steps:

STEP 1: solve the pointwise maximixation problem

$$\begin{aligned} p(t) \cdot f(t, x^*(t), u^*(t)) - L(t, x^*(t), u^*(t)) \quad &= \quad \max_{u \in U} \{p(t) \cdot f(t, x^*(t), u(t)) \\ &\qquad -L(t, x^*(t), u(t))\} \end{aligned}$$

And obtaining the optimal control $u^\#$ as a function of t, x, p :

$$u^\#(t, x, p) = \arg\max_{u \in U} \{p \cdot f(t, x, u) - L(t, x, u)\} \tag{4.51}$$

STEP 2: solve the two-point boundary value problem on the interval $[0, T]$:

$$\begin{cases} \dot{x} = f(t, x, u^\#(t, x, p)) \\ \dot{p} = -p(t) \frac{\partial f}{\partial x}(t, x, u^\#(t, x, p)) + \frac{\partial L}{\partial x}(t, x, u^\#(t, x, p)) \end{cases} \begin{cases} x(0) = x_0 \\ p(T) = \psi_x(x(T)) \end{cases} \tag{4.52}$$

Theorem 4.2.3. ***(Necessary conditions for an open-loop Stackelberg equilibrium).*** *Let the assumptions* $(A2)-(A3)$ *hold. Let* $t \to (u_1^*(t), u_1^*(t))$ *be open-loop strategies yielding a Stackelberg equilibrium for the differential game* (4.1)-(4.3). *Let* $x^*(\cdot), q_2^*(\cdot)$ *be the corresponding trajectory and adjoint vector for Player 2, satisfying* (4.37)-(4.39). *Then there exists a constant* $\lambda_0 \geq 0$ *and two absolutely continuous adjoint vectors* $\lambda_1(\cdot), \lambda_1(\cdot)$*(not all equal to zero), satisfying the equations*

$$\begin{cases} \dot{\lambda}_1 = \lambda_0 \dfrac{\partial \tilde{L_1}}{\partial x} - \lambda_1 \dfrac{\partial F}{\partial x} - \lambda_2 \dfrac{\partial G}{\partial x} \\ \dot{\lambda}_2 = \lambda_0 \dfrac{\partial \tilde{L_1}}{\partial q_2} - \lambda_1 \dfrac{\partial F}{\partial q_2} - \lambda_2 \dfrac{\partial G}{\partial q_2} \end{cases} \tag{4.53}$$

for a.e. $t \in [0, T]$ *together with the boundary conditions*

$$\lambda_2(0) = 0 \qquad \lambda_1(T) = \lambda_0 \nabla_x \psi_1(x^*(T)) - \lambda_2(T) D^2 \psi_2(x^*(T)). \tag{4.54}$$

Moreover, for a.e. $t \in [0, T]$ one has

$$\begin{aligned} u_1^*(t) \quad &= \quad arg \max_{\omega \in u_1} \{ -\lambda_0 \tilde{L}_1(t, x^*(t), q_2^*(t), \omega) + \lambda_1(t) F(t, x^*(t), q_2^*(t), \omega) + \\ &\quad \lambda_2(t) G(t, x^*(t), q_2^*(t), \omega) \end{aligned} \tag{4.55}$$

In the ODEs (4.53), it is understood that the right hand sides are computed at $(t, x^*(t), q_2^*(t), u_1^*(t))$. In (4.54), by $D^2\psi_2(x)$ we denote the Hessian matrix of second derivatives of $\psi_2(x)$, at the point x.

Observe that the initial data is constrained to the set

$$S_0 = \{(x, q_2) \in \Re^{n+n}; \qquad x = x_0\}.$$

Since there is no cost associated with the initial condition, the initial values of the adjoint vector $\lambda = (\lambda_1, \lambda_2) \in \Re^{n+n}$ can be any vector perpendicular to S_0. Hence

$$\lambda_1(0) \in \Re^n \qquad \lambda_2(0) = 0$$

On the other hand, the terminal data is constrained to the set

$$S_T = \{(x, q_2) \in \Re^{n+n} : q_2 - \nabla_x \psi_2(x) = 0\}.$$

A vector $(v_1, v_2) \in \Re^{2n}$ is tangent to the manifold S_T at the point (x, q_2) provided that

$$v_2 = D_x^2 \psi_2(x) v_1.$$

Hence a vector $(n_1, n_2) \in \Re^{2n}$ is normal to S_T provided that

$$n_1 = -D_x^2 \psi_2(x) n_2.$$

$$\lambda_1(T) = \lambda_0 \nabla_x \psi_1(x^*(T)) = \lambda_0 \nabla_x \psi_1(x^*(T)) - \lambda_2(T) D^2 \psi_2(x(T)).$$

for some constant $\lambda_0 \geq 0$.

Example 4.2.2. ***(economic growth).*** *Let $x(t)$ describe the total wealth of capitalists in a country, at time t. Assume that this quantity evolves according to*

$$\dot{x} = ax - u_1 x - u_2, \qquad x(0) = x_0, \qquad t \in [0, T]. \tag{4.56}$$

Here $a > 0$ is a constant growth rate, $u_2(t)$ is the instantaneous amount of consumption, and u_1 is the capital tax rate imposed by the government. The payoffs for the government and for the capitalists are given by

$$J_1 = bx(T) + \int_0^T \phi_1\left(u_1(t)x(t)\right) dt \tag{4.57}$$

$$J_2 = x(T) + \int_0^T \phi_2\left(u_2(t)\right) dt \tag{4.58}$$

Here ϕ_1, ϕ_2 are utility functions. To fix the ideas, assume $\phi_i(s) = k_i \ln s$.
We seek a Stackelberg equilibrium for this differential game, where the government is the leader, announcing in advance the tax rate $u_1()$ as a function of time, and the capitalists are the followers. For this example, the functions considered in (A3)-(A4) take the form

$$\begin{aligned} u_2^b(x, u_1, q_2) &= arg \max_{\omega \geq 0} \{-q_2\omega + k_2 \ln \omega\} = \frac{k_2}{q_2} \\ \tilde{L}_1(x, q_2, u_1) &= \phi_1\left(u_1(t)x(t)\right) = k_1 \ln(u_1 x), \\ F(x, q_2, u_1) &= ax - u_1 x - \frac{k_2}{q_2}, \\ G(x, q_2, u_1) &= -q_2(x - u_1). \end{aligned}$$

The government, playing the role of the leader, now has to solve the following optimization problem.

$$\text{maximize:} \qquad bx(T) + \int_0^T k_1 \ln\left(u_1(t)x(t)\right) dt \tag{4.59}$$

for a system with two state variables (x, q_2), with dynamics

$$\begin{cases} \dot{x} = ax - u_1 x - \dfrac{k_2}{q_2}, \\ \dot{q}_2 = -q_2(x - u_1), \end{cases} \tag{4.60}$$

and boundary conditions

$$x(0) = x_0, \qquad q_2(T) = 1. \tag{4.61}$$

By the above Theorem , an optimal control can be found as follows.
STEP 1: *For any constants $\lambda_0 \geq 0, \lambda_1, \lambda_2$ compute the optimal feedback control*

$$u_1^{\#}(x, q_2, \lambda_0, \lambda_1, \lambda_2) = arg \max_{\omega \geq 0} \{\lambda_1(-\omega x) + \lambda_2 q_2 \omega + \lambda_0 k_1 \ln(\omega x)\} = \frac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}$$

STEP 2: *Solve the boundary value problem for the system of ODEs*

$$\begin{cases} \dot{x} = (a - u_1^{\#})x - \dfrac{k_2}{q_2} = \left(a - \dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}\right) x - \dfrac{k_2}{q_2}, \\ \dot{q}_2 = -q_2(x - u_1^{\#}) = \left(\dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2} - a\right) q_2, \\ \dot{\lambda}_1 = -\lambda_0 \dfrac{k_1}{x} - \lambda_1(a - u_1^{\#}) = -\lambda_0 \dfrac{k_1}{x} + \lambda_1 \left(\dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2} - a\right), \\ \dot{\lambda}_2 = -\lambda_1 \dfrac{k_2}{q_2} - \lambda_2(a - u_1^{\#}) = -\lambda_1 \dfrac{k_2}{q_2} + \lambda_2 \left(a - \dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}\right). \end{cases}$$

with initial and terminal conditions

$$x(0) = x_0, \qquad q_2(T) = 1 \qquad \lambda_1(T) = \lambda_0 b \qquad \lambda_2(0) = 0.$$

From step two we can get the value of $x(t), q_2(t), \lambda_1(t), \lambda_0$ *and* $\lambda_2(t)$ *and then the two optimal controls are determined as;*

$$u_1^*(t) = u_1^{\#}(x, q_2, \lambda_0, \lambda_1, \lambda_2) = \frac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}, \qquad u_2^*(t) = u_2^b(x, u_1^*, q_2) = \frac{k_2}{q_2}$$

4.3 Markovian strategies

We consider here the case where both players can observe the current state of the system. Their strategies will thus be functions $u_i = u_i(t, x)$ of time t and of the state x. Observe that, in the open-loop case, the optimal controls $u_i = u_i^*(t)$ strongly depend on the initial data x_0 and time only. On the other hand, in the Markovian case, it is natural to look for optimal feedback strategies $u_i = u_i^*(t, x)$ that are optimal for the problems (4.1), (4.3), simultaneously for any choice of initial data

$$x(\tau) = y \tag{4.62}$$

with $\tau \in [0, T], y \in \Re^n$.

In the following, we say that a control $(t, x) \to u(t, x) \in U$ is an **optimal feedback** for the optimization problem

$$\max_u \{\psi(x(T)) - \int_\tau^T L(t, x, u))dt\} \tag{4.63}$$

with dynamics

$$\dot{x}(t) = f(t, x, u)) \qquad u(t) \in U, \tag{4.64}$$

if, for every initial data $(\tau, y) \in [0, T] \times \Re^n$, every Caratheodory solution of the Cauchy problem.

$$\dot{x}(t) = f(t, x, u) \qquad x(\tau) = y$$

is optimal, i.e. it achieves the maximum payoff in (4.63).

Definition 4.3.1. ***(feedback Nash equilibrium).*** *A pair of control functions $(t,x) \to (u_1^*(t,x), u_2^*(t,x)$ is a Nash equilibrium for the game (4.1),(4.2), (4.3) within the class of feedback strategies if the following holds.*

(i) The control $(t,x) \to u_1^*(t,x)$ provides an optimal feedback in connection with the optimal control problem for Player 1:

$$\max_{u_1}\{\psi_1(x(T)) - \int_0^T L_1(t,x(t),u_1,u_2^*(t,x(t)))dt\} \tag{4.65}$$

for the system with dynamics

$$\dot{x}(t) = f(t,x(t),u_1,u_2^*(t,x(t))) \qquad u_1(t) \in U_1. \tag{4.66}$$

(ii) The control $(t,x) \to u_2^*(t,x)$ provides an optimal feedback in connection with the optimal control problem for Player 2:

$$\max_{u_2}\{\psi_2(x(T)) - \int_0^T L_2(t,x(t),u_1^*(t,x(t)),u_2)dt\} \tag{4.67}$$

for the system with dynamics

$$\dot{x}(t) = f(t,x(t),u_1^*(t,x(t)),u_2)) \qquad u_2 \in U_2. \tag{4.68}$$

4.3.1 Finding feedback Nash equilibria by solving a system of PDEs.

Assume that the pair of feedback controls (u_1^*, u_2^*) provides a Nash equilibrium. Given an initial data $(\tau, y) \in [0,T] \times \Re^n$, call $t \to x^*(t;\tau,y)$ the solution of

$$\dot{x} = f(t,x(t),u_1^*(t,x),u_2^*(t,x)) \qquad x(\tau) = y$$

Then the corresponding value functions V_1, V_2 will be;

$$V_i(\tau,y) = \psi_i(x^*(T) - \int_\tau^T L_i(t,x^*(t),u_1^*(t,x(t)),u_2^*(t,x(t)))dt, \qquad \text{for} \quad i = 1,2.$$

where $x^*(t) = x^*(t,\tau,y)$. Notice that $V_i(\tau,y)$ is the total payoff achieved by Player i if the game starts at y , at time τ.
Let the assumption (A1) hold. On a region where V_1, V_2 are C^1 and by the dynamic programming principle they satisfy the system of Hamilton-Jacobi PDEs

$$\begin{cases} \nabla_t V_1 + \nabla_x V_1 \cdot f(t,x,u_1^\#,u_2^\#) = L_1(t,x,u_1^\#,u_2^\#) \\ \nabla_t V_2 + \nabla_x V_2 \cdot f(t,x,u_1^\#,u_2^\#) = L_2(t,x,u_1^\#,u_2^\#) \end{cases} \tag{4.69}$$

This system is closed by the equations

$$u_i^\# = u_i^\#(t, x, \nabla_x V_1, \nabla_x V_2) \qquad i = 1, 2, \tag{4.70}$$

introduced at (4.12), and complemented by the terminal conditions

$$V_1(T, x) = \psi_1(x), \qquad V_2(T, x) = \psi_2(x) \tag{4.71}$$

Because of the non linearity of the functions $(t, x, q^1, q^2) \to u_i^\#(t, x, q^1, q^2)$, the system (4.69) is a strongly non-linear system of two scalar **PDEs,** and difficult to solve.

The well-posedness of the Cauchy problem can be studied by looking at a linearized equation. Let $V = (V_1, V_2)$ be a smooth solution of (4.69), and let

$$V^\epsilon(t, x) = V(t, x) + \epsilon Z(t, x) + o(\epsilon) \tag{4.72}$$

describe a small perturbation. Here the Landau symbol $o(\epsilon)$ denotes a higher order infinitesimal, as $\epsilon \to 0$. Assuming that V^ϵ is also a solution, we can insert (4.72) in the equation (4.69) and compute a linearized equation satisfied by the first order perturbation $Z = (Z_1, Z_2)$. Writing $f = (f_1, \ldots, f_n)$, $\quad q_1 = (q_{11}, \ldots, q_{1n})$, $\quad q_2 = (q_{21}, \ldots q_{2n})$, we find

$$Z_{i,t} + \sum_{\alpha=1}^{n} f_\alpha Z_{i,x_\alpha} + \sum_{\alpha=1}^{n} \sum_{j=1}^{2} (\nabla_x V_i \cdot \frac{\partial f}{\partial u_j} - \frac{\partial L_i}{\partial u_j})(\frac{\partial u_j^\#}{\partial q_{1\alpha}} Z_{1,x_\alpha} + \frac{\partial u_j^\#}{\partial q_{2\alpha}} Z_{2,x_\alpha}) = 0 \tag{4.73}$$

Observe that, if the maxima in (4.5)-(4.6) are attained at interior points of the domains U_i, then the necessary conditions for a maximum yield

$$\nabla_x V_1 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_i}{\partial u_j} = 0, \qquad \nabla_x V_2 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_2}{\partial u_2} = 0 \tag{4.74}$$

Therefore, these terms drop off from the right hand sides of (4.73). In matrix notation, this homogeneous linear system can be written as

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} + \sum_{\alpha=1}^{n} A^\alpha \begin{pmatrix} Z_{1,x_\alpha} \\ Z_{2,x_\alpha} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4.75}$$

where the 2×2 matrices A^α are given by

$$A^\alpha = \begin{pmatrix} f_\alpha + \left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) \frac{\partial u_2^\#}{\partial q_{1\alpha}} & \left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) \frac{\partial u_1^\#}{\partial q_{2\alpha}} \\ \left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \frac{\partial u_1^\#}{\partial q_{1\alpha}} & f_\alpha + \left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \frac{\partial u_1^\#}{\partial q_{2\alpha}} \end{pmatrix} \tag{4.76}$$

Fix a point $(\bar{t}, \bar{x})$, and freeze the coefficients of the above matrices at the corresponding point $\left(\bar{t}, \bar{x}), V(\bar{t}, \bar{x}), D_x V(\bar{t}, \bar{x})\right)$. In this way we obtain a linear system of two first order linear homogeneous PDEs with constant coefficients.

A necessary condition in order that the system (4.75) be hyperbolic (and hence that the linear Cauchy be well posed), is that for all $\zeta \in \Re^n$ the matrix

$$A(\zeta) = \sum_{\alpha} A^{\alpha}\zeta_{\alpha} \tag{4.77}$$

has real eigenvalues.
To understand whether this condition can be satisfied, consider first the simpler situation where the dynamics and the payoff functions can be decoupled, i.e.

$$f = f^{(1)}(t, x, u_1) + f^{(2)}(t, x, u_2), \qquad L_i = L_i^{(1)}(t, x, u_1) + L_i^{(2)}(t, x, u_2).$$

In this case the function $u_1^{\#}$ does not depend on q_2 , and similarly the function $u_2^{\#}$ does not depend on q_1 . The 2×2 matrix $A(\zeta)$ thus takes the simpler form

$$A^{\zeta} = \sum_{\alpha=1}^{n} \begin{pmatrix} f_{\alpha}\zeta_{\alpha} & \left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) \frac{\partial u_2^{\#}}{\partial q_{2\alpha}} \zeta_{\alpha} \\ \left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \frac{\partial u_1^{\#}}{\partial q_{1\alpha}} \zeta_{\alpha} & f_{\alpha}\zeta_{\alpha} \end{pmatrix}$$

Consider the two vector

$$V = (v_1, \dots, v_n), \qquad V_{\alpha} = \left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) \frac{\partial u_2^{\#}}{\partial q_{2\alpha}}, \tag{4.78}$$

$$W = (w_1, \dots, w_n), \qquad W_{\alpha} = \left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \frac{\partial u_1^{\#}}{\partial q_{1\alpha}}. \tag{4.79}$$

Observe that the matrix $A(\zeta)$ in (4.77) has real eigenvalues if and only if the two inner products satisfy

$$(V \cdot \zeta)(W \cdot \zeta) \geq 0. \tag{4.80}$$

The condition (4.80) is satisfied for all $\zeta \in \Re^n$ if and only if the two vectors v, w are linearly dependent and have the same orientation. That is, if and only if there exist scalar coefficients $a, b \geq 0$, not both zero, such that $au = bw$.
In any dimension $n \geq 2$, this condition generically fails. Indeed, if v, w are linearly independent, we can find a vector of the form $\zeta = v - \theta w$ which is perpendicular to $v + w$, so that (4.80) fails. Hence the system (4.69) is NOT hyperbolic, and the linearized Cauchy problem is ill-posed, both forward and backward in time.
Going back to the general case (4.76), recall that a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has real eigenvalues if and only if $(a - d)^2 + 4bc \geq 0$. Introduce the vector

$$Z = (z_1, \dots, z_n), \qquad Z_{\alpha} = \left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) \frac{\partial u_2^{\#}}{\partial q_{1\alpha}} - \left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \frac{\partial u_1^{\#}}{\partial q_{2\alpha}}. \tag{4.81}$$

This holds if and only if

$$(Z\cdot\zeta)^2 + 4(V\cdot\zeta)(W\cdot\zeta) \geq 0. \qquad \text{for all} \quad \zeta\in\Re^n. \tag{4.82}$$

In space dimension $n\geq 3$, the condition (4.82) generically fails. Indeed, assume that the vectors v, w, z are linearly independent. Then we can find a nonzero vector

$$\zeta \in \{Z, V+W\}^{\perp} \cap span\{V, W, Z\}$$

With this choice, the quantity in (4.82) is strictly negative.
In space dimension $n = 2$, however, one may find situations where

$$\min_{\zeta\in\Re^2,\quad |\zeta|=1} \left\{(Z\cdot\zeta)^2 + 4(V\cdot\zeta)(W\cdot\zeta)\right\} > 0$$

.

For example, if the vectors in (4.78) , (4.79), (4.81) happen to be

$$V = (1,0), \qquad W = (1,1), \qquad Z = (0,2),$$

then the system (4.75)-(4.76), in two space dimensions, would be locally hyperbolic. Indeed,for any $\zeta = (\zeta_1, \zeta_2)$, one has

$$(Z\cdot\zeta)^2 + 4(V\cdot\zeta)(W\cdot\zeta) = 4\zeta_1^2 + 4\zeta_1(\zeta_1+\zeta_2) = 3(\zeta_1+\zeta_2)^2 + (\zeta_1-\zeta_2)^2 \geq 0.$$

Remark 4.3.1. *In the special case of a zero-sum game, we have $\psi_2 = -\psi_1, L_2 = -L_1$, and $V_2 = -V_1$. The matrices A^α in* (4.76) *should be computed only at points (t, x, q_1, q_2) where $q_2 = \nabla_x V_2 = -\nabla_x V_1 = -q_1$ By* (4.74)*, this yields*

$$\begin{aligned}
\nabla_x V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1} &= -\left(\nabla_x V_1 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_1}{\partial u_1}\right) = 0\\
\nabla_x V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2} &= -\left(\nabla_x V_2 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_2}{\partial u_2}\right) = 0
\end{aligned}$$

Therefore, in the case of zero sum games we simply have

$$A^\alpha = \begin{pmatrix} f_\alpha & 0 \\ 0 & f_\alpha \end{pmatrix}$$

and the system is clearly hyperbolic.

Apart from zero-sum games, to find relevant cases where the backward Cauchy problem (4.69)-(4.71) is well posed, one has to restrict the attention to games in one space dimension. An existence theorem of Nash equilibria in feedback form, valid for one-dimensional non cooperative games. This result is obtained differentiating the equations (4.69) w.r.t. the space variable x. This yields a nonlinear system of conservation laws for the variables $q_1 = \nabla_x V_1$ and $q_2 = \nabla_x V_2$. If this system is hyperbolic, well known PDE theory yields the existence and uniqueness of an entropy weak solution to the Cauchy problem. In turn, this yields a Nash equilibrium solution to the non-cooperative game, in feedback form.

Example 4.3.1. *Suppose two firms produce an identical product. The cost of producing it is governed by the total cost function*

$$C(u_i) = cu_i + \frac{u_i^2}{2}, \qquad i = 1, 2 \tag{4.83}$$

where $u_i(t)$ refers to the i^{th} firm's production level at time t. Each firm supplies all it produces at time t to the market. At each point in time the firms face a common price $p(t)$ at which they can sell their product. However, the amount they jointly supply at time $t, u_1(t) + u_2(t)$, determines the rate at which price changes at time t, $\frac{dp(t)}{dt} = \dot{p}(t)$. The relationship between the total amount supplied and the change in price at time t is described by the differential equation,

$$\dot{p}(t) = s\left[a - u_1(t) - u_2(t) - p(t)\right], \qquad p(0) = p_0. \tag{4.84}$$

Thus, $p(t)$ is the state variable. The parameter 5 *refers to the speed at which the price adjusts to the price corresponding to the total quantity supplied on the demand function. The full meaning of this will become apparent shortly. At this point it should be observed that if $s = 0$, then $\dot{p}(t) = 0$. Each firm chooses its level of output $u_i(t)$ so as to maximize*

$$J_i = \int_0^\infty e^{-rt}\left(p(t)u_i(t) - C_i(u_i(t))\right)dt, \qquad i = 1, 2 \tag{4.85}$$

subject to (4.83) *and* (4.84)*.and $s > 0$*

Solution

To find the feedback strategies $u_i(t, p(t))$ we form for each player the value function

$$rV^i(p) = \max_{u_i}\left\{(p-c)u_i - \frac{u_i^2}{2} + sV_p^i(p)\left[a - p - u_i - u_j\right]\right\}, \qquad i = 1, 2, i \neq j \tag{4.86}$$

where V_p^i refers to the derivative of V^i with respect to p. The maximization with respect to u_i yields

$$\begin{aligned} 0 &= \frac{\partial}{\partial u_i}\left\{(p-c)u_i - \frac{u_i^2}{2} + sV_p^i(p)\left[a - p - u_i - u_j\right]\right\} \\ &= p - c - u_i - sV_p^i(p) \\ \Rightarrow u_i(p) &= p - c - sV_p^i, \qquad i = 1, 2. \end{aligned} \tag{4.87}$$

Substitution from (4.87) into (4.86) yields

$$\begin{aligned} rV^i &= (p-c)(p-c-sV_p^i) - \frac{(p-c-sV_p^i)^2}{2} \\ &+sV_p^i\left[a-p-(2p-2c-sV_p^i-sV_p^j)\right], \qquad i=1,2, \qquad i\neq j. \end{aligned} \tag{4.88}$$

Solving the system of differential equations represented by (4.88) means finding value functions $V^i(p)$ that satisfy them.
∴ The following value functions are proposed as solutions,

$$V^i(p) = g_i - E_i p + K_i\frac{p^2}{2}, \qquad i=1,2. \tag{4.89}$$

which implies

$$\begin{aligned} \frac{\partial}{\partial p}V^i = V_p^i &= \frac{\partial}{\partial p}\left(g_i - E_i p + K_i\frac{p^2}{2}\right) \\ &= K_i p - E_i. \end{aligned} \tag{4.90}$$

In order for the value functions proposed in (4.89) to be solutions to (4.88), the coefficients E_i, K_i and the constant g must have the "right" values. To find them, substitute from (4.90) into (4.88)

$$\begin{aligned} rV^i &= (p-c)(p-c-sV_p^i) - \frac{(p-c-sV_p^i)^2}{2} \\ &+sV_p^i\left[a-p-(2p-2c-sV_p^i-sV_p^j)\right] \\ r\left(g_i - E_i p + K_i\frac{p^2}{2}\right) &= (p-c)(p-c-s(K_ip-E_i)) - \frac{(p-c-s(K_ip-E_i))^2}{2} \\ &+s(K_ip-E_i)\left[a-p-(2p-2c-sV_p^i(K_ip-E_i)-s(K_jp-E_j))\right] \\ \frac{1}{2}rK_ip^2 - rE_ip + rg_i &= \left(\frac{1}{2} - 3sK_i + s^2K_iK_j + \frac{1}{2}s^2K_i^2\right)p^2 \\ &+\left[3sE_i - s^2K_iE_i - 2s^2K_iE_j - c + sK_i(a+2c)\right]p \\ &+\frac{1}{2}c^2 + \left(\frac{1}{2}sE_i + sE_j - a - 2c\right)E_i, \qquad i=1,2. \end{aligned} \tag{4.91}$$

Now the coefficients of p^2 , p as well as the constant terms on both side of (4.91) must be equal,
Equating the coefficients of p^2 yields

$$\begin{aligned} \frac{1}{2} - 3sK_i + s^2K_iK_j + \frac{1}{2}s^2K_i^2 &= \frac{1}{2}rK_i \\ 1 - 6sK_i + 2s^2K_iK_j + s^2K_i^2 &= rK_i \\ s^2K_i^2 + 2s^2K_iK_j - 6sK_i - rK_i + 1 &= 0 \\ s^2K_i^2 + (2s^2K_j - 6s - r)K_i + 1 &= 0, \qquad i=1,2, \qquad i\neq j. \end{aligned} \tag{4.92}$$

This means that

$$\begin{aligned}
s^2K_1^2 + (2s^2K_2 - 6s - r)K_1 + 1 &= 0, \quad \text{and} \\
s^2K_2^2 + (2s^2K_1 - 6s - r)K_2 + 1 &= 0 \\
\Rightarrow s^2(K_1^2 - K_2^2) + 2s^2K_2K_1 - 2s^2K_1K_2 - (6s + r)(K_1 - K_2) &= 0 \\
\Rightarrow s^2(K_1 + K_2)(K_1 - K_2) - (6s + r)(K_1 - K_2) &= 0 \\
\Rightarrow (K_1 - K_2)\left(s^2(K_1 + K_2) - (6s + r)\right) &= 0 \qquad (4.93) \\
\Rightarrow \quad \text{either} \quad K_1 = K_2 \quad \text{or} \quad s^2(K_1 + K_2) = 6s + r & \qquad (4.94)
\end{aligned}$$

Substituting from (4.90) into (4.87) and then into the state equation (4.84) gives

$$\begin{aligned}
u_i(p) &= p - c - sV_p^i \qquad \text{since} \quad V_p^i = K_ip - E_i \\
&= p - c - s(K_ip - E_i) \qquad (4.95) \\
\Rightarrow \dot{p}(t) &= s\left[a - u_1(t) - u_2(t) - p(t)\right] \\
&= s\left[a - (p - c - s(K_1p - E_1)) - (p - c - s(K_2p - E_2)) - p(t)\right] \\
\dot{p}(t) - sp\left[s(K_1 + K_2) - 3\right] &= s\left[a + 2c - s(E_1 + E_2)\right] \qquad (4.96)
\end{aligned}$$

A particular solution to this first order differential equation is obtained by setting $\dot{p}(t) = 0$ and solving for $p(t) = \tilde{p}$, a constant. The solution to the homogeneous part is

$$p(t) = ce^{Dt} \qquad (4.97)$$

where $D = s\left[s(K_1 + K_2) - 3\right]$, and c is the constant of integration.
The general solution to (4.96) is

$$p(t) = \tilde{p} + (p_0 - \tilde{p})e^{Dt} \qquad (4.98)$$

where $c = p_0 - p$ is obtained by setting $t = 0$ in (4.98). Now in order for $p(t)$ to converge to p as $t \to \infty, D < 0$ is required

$$\begin{aligned}
\Rightarrow D = s\left[s(K_1 + K_2) - 3\right] &< 0 \\
\Rightarrow s(K_1 + K_2) - 3 &< 0 \\
\Rightarrow s(K_1 + K_2) &< 3 \quad \text{is required}
\end{aligned}$$

If from (4.94) $s^2(K_1 + K_2) = 6s + r$ by substitution for $s^2(K_1 + K_2)$ this means that

$$\begin{aligned}
s^2(K_1 + K_2 &= 6s + r \\
s(s(K_1 + K_2)) &= 6s + r \\
\Rightarrow 6s + r &= s(s(K_1 + K_2)) < 3s \\
\Rightarrow 6s + r &< 3s \\
\Rightarrow r + 3s < 0 \quad & \text{is required}
\end{aligned}$$

Since the values of r and s are nonnegative this leads to $K_1 = K_2$
Let $K_1 = K_2 = K$ and Equating the coefficients of p from equation (4.91)yields

$$\begin{aligned}
3sE_i - s^2K_iE_i - 2s^2K_iE_j - c + sK_i(a+2c) &= -rE_i \\
3sE_i + rE_i - s^2K_iE_i - 2s^2K_iE_j - c + sK_i(a+2c) &= 0 \\
(3s + r - s^2K_i)E_i - 2s^2K_iE_j - c + sK_i(a+2c) &= 0 \\
\frac{2s^2K_iE_j + c - sK_i(a+2c)}{3s + r - s^2K_i} &= E_i \qquad (4.99)
\end{aligned}$$

since $K_1 = K_2 = K$,

$$\begin{aligned}
(3s + r - s^2K)E_1 - 2s^2KE_2 - c + sK(a+2c) &= 0 \\
(3s + r - s^2K)E_2 - 2s^2KE_1 - c + sK(a+2c) &= 0 \\
\Rightarrow \quad (3s + r - s^2K)(E_1 - E_2) + 2s^2K(E_1 - E_2) &= 0 \\
\Rightarrow \quad (3s + r - s^2K + 2s^2K)(E_1 - E_2) &= 0 \\
\Rightarrow \quad (3s + r + s^2K)(E_1 - E_2) &= 0 \quad \text{for} \quad K \neq \frac{-3s - r}{s^2} \\
\Rightarrow \quad E_1 &= E_2
\end{aligned}$$

And Equating the constant term from equation (4.91)yields

$$\begin{aligned}
\frac{1}{2}c^2 + \left(\frac{1}{2}sE_i + sE_j - a - 2c\right)E_i &= rg_i \\
c^2 + (sE_i + 2sE_j - 2a - 4c)E_i &= 2rg_i
\end{aligned}$$

and since $E_1 = E_2 = E$ we have.

$$\begin{aligned}
c^2 + (sE + 2sE - 2a - 4c)E &= 2rg_1 \quad \text{and} \\
c^2 + (sE + 2sE - 2a - 4c)E &= 2rg_2 \\
&\Rightarrow g_1 = g_2
\end{aligned}$$

Since the values of $K_1 = K_2, E_1 = E_2$, and $g_1 = g_2$ this implies that the control $u_1^*(p)$ and $u_2^*(p)$ are asymmetric, $u_1^*(p) = u_2^*(p)$.
Having established that $K_l = K_2 = K$ we can now solve for the roots of equation (4.92).

$$\begin{aligned}
s^2K^2 + (2s^2K - 6s - r)K + 1 &= 0 \\
s^2K^2 + 2s^2K^2 - (6s + r)K + 1 &= 0 \\
3s^2K^2 - (6s + r)K + 1 &= 0
\end{aligned}$$

$$\overline{K}, \underline{K} = \frac{\left\{r + 6s \underline{+} \left[(r+6s)^2 - 12s^2\right]^{\frac{1}{2}}\right\}}{6s^2} \qquad (4.100)$$

To distinguish between the two roots $\overline{K}, \underline{K}$ given in (4.100) we return to equation (4.98), with $K_1 = K_2$ where now $D = s(2sK - 3)$. The requirement that $D < 0$ means that $K < \frac{3}{2s}$ is required. Now the larger root K takes on its smallest value when $r = 0$ in (4.100). But for $r = 0, \overline{K} = \frac{(3+\sqrt{6})}{3s} > \frac{3}{2s}$. Thus, the larger root of (4.100) prevents convergence of $p(t)$. On the other hand, the smaller root K achieves its highest value at $r = 0$, as $\frac{\partial K}{\partial r} < 0$. At $r = 0$, $\underline{K} = \frac{(3-\sqrt{6})}{3s} < \frac{3}{2s}$. Thus, only the smaller root allows for the convergence of $p(t)$. With $K = \underline{K}$ and $E_1 = E_2 = E$ it follows that from equation (4.99)

$$\begin{aligned} \frac{2s^2KE + c - sK(a+2c)}{3s + r - s^2K} &= E \\ \Rightarrow E(3s + r - s^2K - 2s^2K) = c - sK(a+2c) \\ \Rightarrow E(3s + r - 3s^2K) = c - sK(a+2c) \\ \Rightarrow E = \frac{c - sK(a+2c)}{3s + r - 3s^2K} \end{aligned} \tag{4.101}$$

and from equation (4.95) we have,

$$\begin{aligned} u_i^*(p) &= p - c - s(K_i p - E_i) \\ &= p - c - s(Kp - E) \\ &= (1 - sK)p + sE - c) \end{aligned}$$

therefore

$$u^*(p) = (1 - sK)p + (sE - c). \tag{4.102}$$

Expression (4.102)gives the Nash equilibrium feedback strategies. Now, from (4.96), the particular solution with $K_1 = K_2 = K$, and $E_1 = E_2 = E$ is when $\dot{p}(t) = 0$ and this implies that

$$\begin{aligned} \dot{p}(t) - sp\,[s(K_1 + K_2) - 3] &= s\,[a + 2c - s(E_1 + E_2)] \\ 0 - s\tilde{p}\,[s(K + K) - 3] &= s\,[a + 2c - s(E + E)] \\ -s\tilde{p}\,[2sK - 3] &= s\,[a + 2c - 2sE] \\ \Rightarrow \tilde{p} = \frac{a + 2(c - sE)}{2(1 - sK) + 1}. \end{aligned} \tag{4.103}$$

This is the stationary feedback Nash equilibrium price.

4.3.2 Linear-quadratic differential games

A large portion of the literature on Nash feedback solutions for differential games is concerned with n-dimensional games having linear dynamics and quadratic payoff functions. It is assumed that the state of the system evolves according to

$$\dot{x} = A(t)x + B_1(t)u_1 + B_2(t)u_2, \tag{4.104}$$

while the payoff functions are given by quadratic polynomials w.r.t. the variables x, u_1, u_2. To simplify the computations, we consider here a homogeneous case, with

$$J_i = \psi_i(x(T)) - \int_0^T L_i(t, x(t), u_1(t), u_2(t))dt \tag{4.105}$$

$$\psi_i(x) = \frac{1}{2}x^t\overline{M_i}x \tag{4.106}$$

$$L_i(t, x, u_1, u_2) = \frac{|u_i|^2}{2} + \frac{1}{2}x^tP_i(t)x + \sum_{j=1,2} x^tQ_{ij}(t)u_j \tag{4.107}$$

where the superscript t denotes transposition. Here $x \in \Re^n, u_1 \in \Re^{m_1}, u_2 \in \Re^{m_2}$ are column vectors, A is an $n \times n$ matrix, $M_i, P_i$are $n \times n$ symmetric matrices,Q_{ij} and Bj are $n \times m$ matrices.

In this model, it is important that the controls u_1, u_2 range over the entire spaces $\Re^{m_1}, \Re^{m_2}$, without being restricted to a compact subset. Notice that the assumption (A1) certainly holds: the functions $u_i^{\#}$ in (4.10)-(4.11)are explicitly computed as

$$\begin{aligned} u_i^{\#}(t, x, q_i) &= arg \max_{\omega \in \Re^{m_1}} \left\{ q_iB_i(t)\omega - \frac{|\omega|^2}{2} - x^tQ_{ii}(t)\omega \right\} \\ &= (q_iB_i(t) - x^tQ_{ii}(t))^t \qquad i = 1, 2. \end{aligned} \tag{4.108}$$

Even if the backward Cauchy problem (4.69) - (4.71) is ill-posed, in this linear-quadratic case one can always construct a (local) solution within the class of homogeneous second order polynomials w.r.t. the variables $x = (x_1, \ldots, x_n)$, namely

$$V_i(t, x) = \frac{1}{2}x^tM_i(t)x \tag{4.109}$$

Indeed, denoting by un upper dot a differentiation w.r.t. t, let us compute

$$\nabla_xV_i(t, x) = x^tM_i(t) \qquad \nabla_tV_i(t, x) = \frac{1}{2}x^t\dot{M}_i(t)x \tag{4.110}$$

Therefore;

$$u_i^{\#}(t, x, \nabla_xV_i(t, x)) = \left(x^tM_i(t)B_i(t) - x^tQ_{ii}(t)\right)^t. \tag{4.111}$$

By (4.110) and (4.111), the functions V_i in (4.109) solve the system

$$\nabla_tV_i(t, x) = L_i - \nabla_xV_i \cdot f \qquad i = 1, 2.$$

if and only if the following relations are satisfied

$$
\begin{aligned}
\frac{1}{2}x^t\dot{M}_i x &= \Big[\frac{1}{2}\left(x^t M_i B_i - x^t Q_{ii}\right)\left(x^t M_i B_i - x^t Q_{ii}\right) + \frac{1}{2}x^t P_i x \\
&\quad + \sum_{j=1,2} x^t Q_{ij}\left(x^t M_i B_i - x^t Q_{ii}\right)^t\Big] \\
&\quad - x^t M_i\left(Ax + \sum_{j=1,2} B_j\left(x^t M_j B_j - x^t Q_{jj}\right)^t\right) \qquad (4.112)
\end{aligned}
$$

Notice that both sides of (4.112) are homogeneous quadratic polynomials w.r.t. the variable $x = (x_1, \ldots, x_n)$. The equality holds for every $x \in \Re^n$ if and only if the following identity between $n \times n$ symmetric matrices is satisfied:

$$
\begin{aligned}
\frac{1}{2}\dot{M}_i &= \frac{1}{2}\left(M_i B_i - Q_{ii}\right)\left(M_i B_i - Q_{ii}\right)^t + \frac{1}{2}P_i \\
&\quad + \frac{1}{2}\sum_{j=1,2}\left[Q_{ij}(M_i B_i - Q_{ii})^t + \left(M_i B_i - Q_{ii}\right)Q_{ii}^t\right] \\
&\quad - \frac{1}{2}\left(M_i A + A^t M_i\right) - \frac{1}{2}\sum_{j=1,2}\left[M_i B_j (M_j B_j - Q_{jj})^t\right. \\
&\quad \left. + \left(M_j B_j - Q_{jj}\right)B_j^t M_i\right]. \qquad (4.113)
\end{aligned}
$$

The equations (4.113) represent a system of ODEs for the coefficients of the symmetric matrices $M_1(t), M_2(t)$. These ODEs need to be solved backward, with terminal conditions

$$
M_1(T) = \overline{M}_1 \qquad M_2(T) = \overline{M}_2 \qquad (4.114)
$$

This backward Cauchy problem has a unique local solution, defined for t close to T. In general, however, a global solution may not exist because the right hand side has quadratic growth. Hence the solution may blow up in finite time.

If the backward Cauchy problem (4.113)-(4.114) has as solution on the entire interval $[0, T]$, then the formulas (4.110)-(4.111) yield the the optimal feedback controls $u_i^*(t, x) = u_i^\#(t, x, \nabla_x V_i(t, x))$.

Remark 4.3.2. *The above approach can be applied to a more general class of non-homogeneous linear-quadratic games, with dynamics*

$$
\dot{x} = A(t)x + B_1(t)u_1 + B_2(t)u_2 + c(t)
$$

and payo functions (4.105) *, where*

$$
\psi_i(x) = \frac{1}{2}x^t\overline{M_i}x + \overline{a_i}x + \overline{e}
$$

$$L_i(t,x,u_1,u_2) = \frac{1}{2}u_i^t R_i(t) u_i + \frac{1}{2}x^t R_i(t) x + \sum_{j=1,2} x^t Q_{ij}(t) u_j + \sum_{j=1,2} S_{ij}(t) u_j + b_i(t) x$$

Here one needs to assume that R_1, R_2 are strictly positive symmetric matrices, for all $t \in [0,T]$. In this case, the value functions are sought within the class of (non-homogeneous) quadratic polynomials:

$$V_i(t,x) = x^t M_i(t) x + a_i(t) x + e(t) \qquad \textit{for} \quad i = 1,2$$

4.4 Two Person Zero-Sum Differential Games

Consider the state equation

$$\dot{x} = f(t,x,u_1,u_2), \qquad x(0) = x_0 \qquad t \in [0,T] \tag{4.115}$$

where we may assume all variables to be scalar for the time being. Extension to the vector case simply requires appropriate reinterpretations of each of the variables and the equations. In this equation, we let u_1 and u_2 denote the controls applied by players 1 and 2, respectively. We assume that

$$u_1(t) \in U_1, u_2(t) \in U_2, t \in [0,T],$$

where U_1 and U_2 are convex sets in $\Re$. Consider further the objective function

$$J(u_1,u_2) = \psi[x(T)] - \int_0^T L(t,x(t),u_1(t),u_2(t))dt \tag{4.116}$$

which player 1 wants to maximize and player 2 wants to minimize. Since the gain of player 1 represents a loss to player 2, such games are appropriately termed zero-sum games. Clearly, we are looking for admissible control trajectories u_1 and u_2 such that

$$J(u_1^*,u_2) \geq J(u_1^*,u_2^*) \geq J(u_1,u_2^*). \tag{4.117}$$

The solution (u_1^*, u_2^*) is known as the minimax solution. Here u_1^* and u_2^* stand for $u_1^*(t)$, $t \in [0,T]$, and $u_2^*(t)$, $t \in [0,T]$, respectively. The necessary conditions for u_1^* and u_2^* to satisfy (4.117) are given by an extension of the maximum principle. To obtain these conditions, we form the Hamiltonian

$$\mathbb{H} = -L + \lambda f \tag{4.118}$$

with the adjoint variable λ satisfying the equation

$$\dot{\lambda} = -\mathbb{H}_x, \qquad \lambda(T) = \psi_x[x(T)]. \tag{4.119}$$

The necessary condition for u_1^* and u_2^* to be a minimax solution is that for $t \in [0, T]$,

$$\mathbb{H}(x^*(t), u_1^*(t), u_2^*(t), \lambda^*(t), t) = \min_{u_2 \in U_2} \max_{u_1 \in U_1} \mathbb{H}(x^*(t), u_1(t), u_2(t), \lambda^*(t), t) \qquad (4.120)$$

which can also be stated, with suppression of (t), as

$$\begin{aligned} \mathbb{H}(x^*(t), u_1^*(t), u_2(t), \lambda^*(t), t) &\geq \mathbb{H}(x^*(t), u_1^*(t), u_2^*(t), \lambda^*(t), t) \\ &\geq \mathbb{H}(x^*(t), u_1(t), u_2^*(t), \lambda^*(t), t) \end{aligned} \qquad (4.121)$$

for $u_1 \in U_1$ and $u_2 \in U_2$. Note that (u_1^*, u_2^*) is a saddle point of the Hamiltonian function $\mathbb{H}$. Note that if u_1 and u_2 are unconstrained, *i.e.*, when, $U_1 = U_2 = \Re$ condition (4.120) reduces to the first-order necessary conditions

$$\mathbb{H}_{u_1} = 0 \qquad \text{and} \qquad \mathbb{H}_{u_2} = 0 \qquad (4.122)$$

and the second-order conditions are

$$\mathbb{H}_{u_1 u_1} \leq 0 \qquad \text{and} \qquad \mathbb{H}_{u_2 u_2} \geq 0 \qquad (4.123)$$

Chapter 5

Conclusion

In this Project paper,I have introduced the notion of pure and mixed strategies in two-player for static games with different optimal solution concepts, including Pareto obtimality, Stackelber gequilibrium , and nash equilibrium solution of the static games.
In general differential games are closely related with the optimal control problems, where state variable evolves over time according to a differential equation. In an optimal control problem there is single control $u(t)$ with a single decision maker in order to optimize the objectives. But differential game theory generalizes this to two controls $u_1(t), u_2(t)$ with two decision makers, one for each player to achieve there own desired objectives. Each player attempts to control the state of the system so as to achieve its goal;the system respond to the inputs of all players.
and also the open loop Stackelberg solution of the differential games are the concepts, where each player controls several variables some of which s/he uses as leader and others as follower. We have adopted the open-loop information structure where the players' controls are functions of the initial state and time, and for this class of games we have obtained a complete set of equations satisfied by the control laws. These equations are differential equations where some of the differential equations have specified initial conditions while others have their terminal conditions specified.
Therefore, differential games are a group of problems related to the modeling and analysis of conflict in the context of a dynamical system.
In this project, we have discussed how two person non-cooperative differential games can be formulated and the solution mechanisms for some class of such problems are also given.

References

[1] A. Bressan, Noncooperative Differential Games. A Tutorial, Department of Mathematics, Penn State University, 2010

[2] A. Palestini, Lectures on Optimal Control Theory and Differential Games,University of Bologna, Italy, 2007.

[3] A. Palestini Some, basic notes on Optimal Control Theory and Differential Games, Department of Economics, University of Bologna, Italy

[4] L. Grosset Open Loop Nash Equilibrium in Linear-State Differential Games Applied Mathematical Sciences,vol. 8, 2014, no. 145

[5] L.Brickman ,mathematical introduction to linear programming and game theory ,springer, 1989

[6] E. Bryson, Applied optimal control Stanford university ,1975

[7] C. Pierre,Introduction to differential games,Universite de Brest,springer, 2010

[8] Hans P. Geering ,Optimal Control with Engineering Applications,2007

[9] G. Knowles , An Introduction to Applied Optimal Control , Department of Mathematics Carnegie-Mellon University, Pennsylvania ,1981

[10] G. Leitmann, Cooperative and noncooperative many player differential games. Berlin: Springer Verlag. (1974).

[11] M.Kamien and N. Schwartz, Daynamic Optimization,The Calculus of Variations and Optimal Control in Economics and Management, 1991,Second edition

[12] Suresh P. Sethi, Optimal control theory Applications to Management Science and Economics ,Spring, 1981,second Edition.

[13] Suresh p.sethi gerald.thomson. optimal control theory,Applications to Management Science and Economics ,Second Edition

[14] T. Basar , Lecture Notes on Non-Cooperative Game Theory , University of Bologna, Italy, 2010